Study of the noise aging mechanisms in single-photon avalanche photodiode for time-of-flight imaging

Mathieu Sicre

N°d'ordre NNT : 2023ISAL0104

THESE de DOCTORAT DE L'INSA LYON, membre de l'Université de Lyon

Ecole Doctorale N° ED 160
EEA (Électronique, Électrotechnique et Automatique)

Spécialité/ discipline de doctorat :
Électronique, micro et nanoélectronique, optique et laser

Soutenue publiquement le 06/12/2023, par :
Mathieu SICRE

Study of the noise aging mechanisms in single-photon avalanche photodiode for time-of-flight imaging

Etude des mécanismes de vieillissement du bruit de photodiode à avalanche à photon unique pour l'imagerie par temps de vol

Devant le jury composé de :

PANCHERI Lucio	Professeur	Université de Trente	Rapporteur
RAFHAY Quentin	Maître de Conférences	IMEP-LAHC	Rapporteur
MANEUX Cristell	Professeure des Universités	Université de Bordeaux	Présidente
GOIFFON Vincent	Professeur des Universités	ISAE-SUPAERO	Examinateur
CALMON Francis	Professeur des Universités	INSA LYON	Directeur de thèse
ROY David	Ingénieur	STMicroelectronics	Invité Co-encadrant
FEDERSPIEL Xavier	Docteur	STMicroelectronics	Invité
COIGNUS Jean	Docteur	CEA-Leti	Invité

Département FEDORA – INSA Lyon - Ecoles Doctorales

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	**CHIMIE DE LYON** **https://www.edchimie-lyon.fr** Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage secretariat@edchimie-lyon.fr	**M. Stéphane DANIELE** C2P2-CPE LYON-UMR 5265 Bâtiment F308, BP 2077 43 Boulevard du 11 novembre 1918 69616 Villeurbanne **directeur@edchimie-lyon.fr**
E.E.A.	**ÉLECTRONIQUE, ÉLECTROTECHNIQUE, AUTOMATIQUE** **https://edeea.universite-lyon.fr** Sec. : Stéphanie CAUVIN Bâtiment Direction INSA Lyon Tél : 04.72.43.71.70 secretariat.edeea@insa-lyon.fr	**M. Philippe DELACHARTRE** INSA LYON Laboratoire CREATIS Bâtiment Blaise Pascal, 7 avenue Jean Capelle 69621 Villeurbanne CEDEX Tél : 04.72.43.88.63 **philippe.delachartre@insa-lyon.fr**
E2M2	**ÉVOLUTION, ÉCOSYSTÈME, MICROBIOLOGIE, MODÉLISATION** **http://e2m2.universite-lyon.fr** Sec. : Bénédicte LANZA Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 secretariat.e2m2@univ-lyon1.fr	**Mme Sandrine CHARLES** Université Claude Bernard Lyon 1 UFR Biosciences Bâtiment Mendel 43, boulevard du 11 Novembre 1918 69622 Villeurbanne CEDEX **sandrine.charles@univ-lyon1.fr**
EDISS	**INTERDISCIPLINAIRE SCIENCES-SANTÉ** **http://ediss.universite-lyon.fr** Sec. : Bénédicte LANZA Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 secretariat.ediss@univ-lyon1.fr	**Mme Sylvie RICARD-BLUM** Institut de Chimie et Biochimie Moléculaires et Supramoléculaires (ICBMS) - UMR 5246 CNRS - Université Lyon 1 Bâtiment Raulin - 2ème étage Nord 43 Boulevard du 11 novembre 1918 69622 Villeurbanne Cedex Tél : +33(0)4 72 44 82 32 **sylvie.ricard-blum@univ-lyon1.fr**
INFOMATHS	**INFORMATIQUE ET MATHÉMATIQUES** **http://edinfomaths.universite-lyon.fr** Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage Tél : 04.72.43.80.46 infomaths@univ-lyon1.fr	**M. Hamamache KHEDDOUCI** Université Claude Bernard Lyon 1 Bât. Nautibus 43, Boulevard du 11 novembre 1918 69 622 Villeurbanne Cedex France Tél : 04.72.44.83.69 **hamamache.kheddouci@univ-lyon1.fr**
Matériaux	**MATÉRIAUX DE LYON** **http://ed34.universite-lyon.fr** Sec. : Yann DE ORDENANA Tél : 04.72.18.62.44 yann.de-ordenana@ec-lyon.fr	**M. Stéphane BENAYOUN** Ecole Centrale de Lyon Laboratoire LTDS 36 avenue Guy de Collongue 69134 Ecully CEDEX Tél : 04.72.18.64.37 **stephane.benayoun@ec-lyon.fr**
MEGA	**MÉCANIQUE, ÉNERGÉTIQUE, GÉNIE CIVIL, ACOUSTIQUE** **http://edmega.universite-lyon.fr** Sec. : Stéphanie CAUVIN Tél : 04.72.43.71.70 Bâtiment Direction INSA Lyon mega@insa-lyon.fr	**M. Jocelyn BONJOUR** INSA Lyon Laboratoire CETHIL Bâtiment Sadi-Carnot 9, rue de la Physique 69621 Villeurbanne CEDEX **jocelyn.bonjour@insa-lyon.fr**
ScSo	**ScSo*** **https://edsciencessociales.universite-lyon.fr** Sec. : Mélina FAVETON INSA : J.Y. TOUSSAINT Tél : 04.78.69.77.79 melina.faveton@univ-lyon2.fr	**M. Bruno MILLY** Université Lumière Lyon 2 86 Rue Pasteur 69365 Lyon CEDEX 07 **bruno.milly@univ-lyon2.fr**

*ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

“T’auras beau raconter, s’pas, on t’croira pas. Pas par méchanceté ou par amour de s’ficher d’toi, mais pa’ce qu’on n’pourra pas. Quand tu diras plus tard, si t’es encore vivant pour placer ton mot : “On a fait des travaux d’nuit, on a été sonnés, pis on a manqué s’enliser”, on répondra : “Ah !” ; p’têt’ qu’on dira : “Vous n’avez pas dû rigoler lourd pendant l’affaire.” C’est tout. Personne ne saura. I’ n’y aura qu’toi.”

Le Feu: journal d'une escouade, L’aube, Henri Barbusse

“Then he who spoke sorrowfully, like a bell, said. It'll be no good telling about it, eh? They wouldn't believe you; not out of malice or through liking to pull your leg, but because they couldn't. When you say to 'em later, if you live to say it, “We were on a night job and we got shelled and we were very nearly drowned in mud,” they'll say, “Ah!” And p'raps they'll say. “You didn't have a very spicy time on the job.” And that's all. No one can know it. Only us.”

Under Fire: The Story of a Squad, The Dawn, Henri Barbusse

Acknowledgements

Ce projet de thèse s'inscrit dans le cadre d'une collaboration entre STMicroelectronics Crolles, le CEA-Leti et l'INSA Lyon-Institut des Nanotechnologies de Lyon (INL). Avant d'entrer dans les détails, je tiens à exprimer ma gratitude envers toutes les personnes qui ont contribué à cette recherche.

Tout d'abord, je souhaite adresser mes remerciements aux membres du jury qui ont accepté ce rôle, notamment Lucio Pancheri et Quentin Rafhay en tant que rapporteurs, Cristell Maneux pour avoir présidé le jury, Vincent Goiffon pour avoir examiné, ainsi que Xavier Federspiel et Jean Coignus pour avoir accepté l'invitation.

Je tiens à exprimer ma reconnaissance envers mon directeur de thèse, Francis Calmon, ainsi que mes encadrants David Roy et Caroline Coutier. Un remerciement particulier à Xavier Federspiel et à Bastien Mandy pour leurs échanges fréquents, à Alexandre Vernhet et Jérôme Goy pour leur expertise instrumentale, et à Jessie Muamba et Aurore D'Oria pour leur assistance dans la réalisation des nombreuses expériences menées au cours de ma thèse. Je souhaite témoigner ma reconnaissance à Megan Agnew pour les caractérisations singulières des SPADs. J'exprime ma profonde gratitude envers François Paolini pour avoir automatisé les tests et assuré l'exportation des données.

Je suis reconnaissant envers Christel Buj pour avoir facilité l'accès à l'ensemble des decks Sentaurus, et envers Gabriel Mugny pour avoir pris le temps de partager la simulation de type Full-Band Monte-Carlo. Je souhaite également exprimer ma reconnaissance envers Ali Jouni pour m'avoir donné l'opportunité de participer aux expériences et simulations impliquant des SPADs irradiés par des protons, ainsi qu'à Tidjani Garba pour m'avoir permis de contribuer aux simulations du mode off-state gate-oxide breakdown dans les MOSFETs. Je remercie Stuart McLeod d'avoir pris le temps de m'expliquer les algorithmes sous-jacents aux mesures à temps de vol. Je souhaite également exprimer ma gratitude envers ceux qui ont précédé mon travail et ont contribué à son avancement.

J'exprime ma gratitude envers Barbara Lisei pour son soutien constant lors des conférences, et je remercie Christèle Izoard pour m'avoir offert la possibilité de participer à diverses formations en communication.

Un grand merci à toute l'équipe Silicon Reliability (SiRel) de STMicroelectronics Crolles et à l'Imaging Division de STMicroelectronics Edinburgh pour les échanges fructueux. Enfin, je souhaite exprimer ma gratitude envers J. pour son soutien tout au long des épreuves.

Abstract

Single-Photon Avalanche Diode (SPAD) are used for Time-of-Flight (ToF) sensors to determine distance from a target by measuring the travel time of an emitted pulsed signal. These photodetectors work by triggering an avalanche of charge carriers upon photon absorption, resulting in a substantial amplification which can be detected. However, they are subject to spurious triggering by parasitic generated charge carriers, quantified as Dark Count Rate (DCR), which can compromise the accuracy of the measured distance. Therefore, it is crucial to identify and eliminate the potential source of DCR.

To tackle this issue, a simulation methodology has been implemented to assess the DCR. This is achieved by simulating the avalanche breakdown probability, integrated with the carrier generation rate from defects. The breakdown probability can be simulated either in a deterministically, based on electric-field streamlines, or stochastically, by means of drift-diffusion simulation of the random carrier path. This methodology allows for the identification of the potential sources of pre-stress DCR by comparing simulation results to experimental data over a wide range of voltage and temperature.

To ensure the accuracy of distance range measurements over time, it is necessary to predict the DCR level under various operating conditions. The aforementioned simulation methodology is used to identify the potential sources of post-stress DCR by comparing simulation results to stress experiments that evaluate the principal stress factors, namely temperature, voltage and irradiance. Furthermore, a Monte-Carlo study has been conducted to examine the device-to-device variation along stress duration.

For an accurate Hot-Carrier Degradation (HCD) kinetics model, it is essential to consider not only the carrier energy distribution function but also the distribution of $Si-H$ bond dissociation energy distribution at the Si/SiO_2 interface. The number of available hot carriers is estimated from the carrier current density according to the carrier energy distribution simulated by means of a full-band Monte-Carlo method. The impact-ionization dissociation probability is employed to model the defect creation process, which exhibits sub-linear time dependence due to the gradual exhaustion of defect precursors.

Accurate distance ranging requires distinguishing the signal from ambient noise and the DCR floor, and ensuring the target's accumulated photon signal dominates over other random noise sources. An analytical formula allows to estimate the maximum distance ranging using the maximum signal strength, ambient noise level, and confidence levels. The impact of DCR can be estimated by considering the target's reflectance and the ambient light conditions.

In a nutshell, this work makes use of a in-depth characterization and simulation methodology to predict DCR in SPAD devices along stress duration, thereby allowing the assessment of its impact on distance range measurements.

Keywords: Single-Photon Avalanche Diode (SPAD), Dark Count Rate (DCR), reliability, Hot-Carrier Degradation (HCD), Technology Computer Aided Design (TCAD), Monte-Carlo methods, electro-optical characterization.

Contents

List of figures

List of tables

List of acronyms

AI carrier Acceleration Integral

ALD Atomic Layer Deposition

AP After-Pulsing

APD Avalanche Photo-Diode

AQC Active Quenching Circuit

ARC Anti-Reflective Coating

BDTI Backside Deep Trench Isolation

BTBT Band-to-Band Tunneling

BTE Boltzmann transport equation

BTI Bias Temperature Instability

CDF Cumulative Distribution Function

CESL Contact Etch Stop Layer

CIS CMOS Image Sensor

CMOS Complementary Metal-Oxide-Semiconductor

CPP Charges Per Pulse

CT Computed Tomography

DCR Dark Count Rate

DD Drift Diffusion

DoE Design of Experiment

DTI Deep Trench Isolation

dToF direct Time of Flight

EES Electron-Electron scattering

ESR Electron Spin Resonance

ET Energy Transport

FBMC Full-Band Monte-Carlo

FF Fill Factor

FGA Forming Gas Annealing

FoV Field of View

FWHM Full-Width at Half Maximum

HB How-bias voltage

HCD Hot-Carrier Degradation

HPD High-Pressure Deuterium annealing

HTOL High Temperature Operating Lifetime

HTS High Temperature Storage

HWHM Half Width at Half Maximum

iToF indirect Time of Flight

LB Low-bias voltage

LCAO Linear Combination of Atomic Orbitals

LCR Light Count Rate

LPCVD Low Pressure Chemical Vapor Deposition

LT Laser Triangulation

MOS Metal-Oxide-Semiconductor

MOSFET Metal-Oxide-Semiconductor Field-Effect Transistor

MVE Multiple Vibrational Excitation

NBTI Negative Bias Temperature Instability

NRM Non-Radiative Multiphonon-assisted

OA Optical Absorption

OEM Original Equipment Manufacturer

OV Oxygen Vacancy

PBTI Positive Bias Temperature Instability

PDE Photo-Detection Efficiency

PDP Photo-Detection Probability

PECVD Plasma-Enhanced Chemical Vapor Deposition

PF Poole-Frenkel effect

PMT Photo-Multiplier Tube

PoR Process of Reference

PQC Passive Quenching Circuit

PVD Physical Vapor Deposition

QE Quantum Efficiency

RDF Random Dopant Fluctuations

R-G Recombination-Generation

RoI Region of Interest

RTP Rapid Thermal Processing

SfM Structure from Motion

SHE Spherical Harmonics Expansion

SL Structured Light

SOI Silicon-On-Insulator

SP Single Process

SPAD Single-Photon Avalanche Diode

SRH Shockley-Read-Hall

TAT Trap-Assisted Tunneling

TCAD Technology Computer-Aided Design

TED Transient Enhanced Diffusion

TEOS Tetra Ethyl Ortho Silicate

ToF Time-of-Flight

VCSEL Vertical-Cavity Surface-Emitting Laser

WKB Wentzel-Kramers-Brillouin

List of symbols

Symbol	Measure Unit	Designation
$c^{n,p}$	$[s^{-1}]$	Capture rate for electrons and holes
DCR	$[cps/ţmš]$	Dark Count Rate
$D_d^{n,p}$	$[cm^2 s^{-1}]$	Diffusion coefficient for holes and electrons
E	$[eV]$	Carrier energy
E_B	$[eV]$	Barrier energy
$E_{c,v}$	$[eV]$	Valance and conduction band energy
$E_{cn,vp}$	$[eV]$	Valance and conduction band energy minimum
E_F	$[eV]$	Fermi energy level
$E_F^{n,p}$	$[eV]$	Quasi-Fermi energy potential for electrons and holes
E_g	$[eV]$	Band gap energy
E_i	$[eV]$	Intrinsic Fermi level
E_p	$[eV]$	Optical-phonon energy
E_t	$[eV]$	Energy level of trap
$E_t h$	$[eV]$	Threshold energy for impact-ionization
$E_t h^{n,p}$	$[eV]$	Threshold energy for impact-ionization for electrons and holes
E_0	$[eV]$	Thermal equilibrium energy at zero electric field
$e^{n,p}$	$[s^{-1}]$	Emission rate for electrons and holes
F	$[Vcm^{-2}]$	Electric field
$F_{th}^{n,p}$	$[eV]$	Threshold electric field for impact-ionization for electrons and holes
$f_{c,v}$	$[.]$	Occupancy probability function of electrons and holes
f_{ph}	$[.]$	Occupancy probability function of phonon
f_t	$[.]$	Occupancy probability function of trap
$f(E)$	$[cm^{-3}eV^{-1}]$	Carrier energy distribution function
G_d	$[cm^{-3}]$	Dark generation rate
$G_p h$	$[cm^{-3}]$	Photo-generated generation rate
h	$[m^2 kgs^{-1}]$	Plank's constant
$\hbar$	$[m^2 kgs^{-1}]$	$\frac{h}{2\pi}$

I_d	$[A]$	Dark current
$I(E)$	$[cm^{-3}]$	Carrier flux depending on carrier energy
I_ph	$[A]$	Photo-generated current
k	$[m^{-1}]$	Wave vector
k_B	$[JK^{-1}]$	Boltzmann constant
l_t	$[m^{-1}]$	Tunneling lentgh
m_0	$[kg]$	Rest mass of the electron
$m^*_{c,v}$	$[kg]$	Effective mass of the density of states in the conduction band and in the valence band
$m^*_{n,p}$	$[kg]$	Effective mass for electrons and holes
m^*	$[kg]$	Effective tunneling mass
$n(E)$	$[cm^{-3}eV^{-1}]$	Carrier energy distribution function
n, p	$[cm^{-3}]$	Electron and hole density
$N_{c,v}$	$[cm^{-3}]$	Effective density of states in the conduction band and in the valence band
$N_{D,A}$	$[cm^{-3}]$	Donor and acceptor density
N_{hb}	$[cm^{-3}]$	Initial concentration of $Si-H$ bond precursors
N_t	$[cm^{-3}]$	Trap density
n_t	$[cm^{-3}]$	Number of filled R-G centers
N_{total}	$[cm^{-3}]$	Initial concentration of $Si-H$ bond precursors plus the number of interfacial states
$P(E)$	[.]	Desorption probability depending on carrier energy
P^0_{II}	$[eV^{-\alpha_0}s^{-1}]$	Attempt rate for impact-ionization
P_t	[%]	Avalanche breakdown probability
$P_{e,h}$	[%]	Avalanche breakdown probability for electrons and holes
P_α	[%]	Joint avalanche breakdown probability for both electron and holes
p_t	$[cm^{-3}]$	Number of emptied R-G centers
q	$[C]$	Elementary charge
R	$[cm^{-3}s^{-1}]$	Net steady-state recombination-generation rate
$r_{n,p}$	$[cm^{-3}s^{-1}]$	Net recombination rates of electrons and holes
$S(E)$	$[cm^2]$	Reaction cross-section
t	$[s]$	Time
T	$[K]$	Lattice temperature

T_c	$[K]$	Characterization temperature
T_s	$[K]$	Stress temperature
V_c	$[V]$	Characterization voltage applied above V_{BD}
V_{BD}	$[V]$	Breakdown voltage corresponding to V_{BD}^{LB}
$V_{BD}^{LB,HB}$	$[V]$	Breakdown voltage of the low- and high-voltage bias regions
$v_d^{n,p}$	$[mcm^{-1}]$	Drift velocity for holes and electrons
V_g	$[V]$	Gate voltage
V_s	$[V]$	Stress voltage applied above V_{BD}
$v_{th}^{n,p}$	$[mcm^{-1}]$	Thermal velocity for holes and electrons
V_{th}	$[V]$	Threshold voltage
$\alpha_{n,p}$	$[cm^{-1}]$	Impact-ionization coefficient for electrons and holes
α_0	$[.]$	Energy exponent of the impact-ionization dynamics
γ	$[s^{-1}]$	Backward (passivation) reaction rate
$\Gamma_{TAT,PF}$	$[.]$	Electric field enhancement factor for TAT and PF
ΔDCR	$[cps/ţmš]$	Dark Count Rate drift
ΔG	$[J]$	Change in Gibbs free energy
ΔG^*	$[J]$	Gibbs free energy of activation
ΔH	$[J]$	Change in enthalpy
ΔS	$[JK^{-1}]$	Change in entropy
ΔH^*	$[J]$	Enthalpy of activation
ΔS^*	$[JK^{-1}]$	Entropy of activation
Δt_{ff}	$[s^{-1}]$	Carrier free-flight time
ϵ_0	$[Fm^{-1}]$	Vacuum permittivity
ϵ_r	$[Fm^{-1}]$	Relative permittivity
$Theta_{r,\|\|}^{\alpha}$	$[s^{-1}]$	Electrooptical frequency
κ	$[s^{-1}]$	Forward (depassivation) reaction rate
κ_0	$[s^{-1}]$	Attempt frequency for HCD mechanism
κ_0	$[.]$	Hot-carrier acceleration factor
Λ_c	$[m^{-1}]$	Carrier mean free path between two successive collisions
$\Lambda_{n,p}$	$[m^{-1}]$	Optical-phonon mean free path
Λ_0	$[m^{-1}]$	Carrier mean free path
$\mu_{n,p}$	$[Fm^{-1}]$	Mobility for electrons and holes

$\nu_{SP,MP}^{(e/h)}$	[.]	Attempt frequency for the SP and MVE mechanisms triggered by electrons or holes
ρ_0	[.]	Energy exponent of the impact-ionization dynamics
ρ	$[cm^{-3}]$	Charge density
σ_{abs}	$[cm^2]$	Phonon absorption capture cross section
σ_{emi}	$[cm^2]$	Phonon emission capture cross section
$\sigma_{n,p}$	$[cm^2]$	Capture cross section for electrons and holes
σ_{II}	$[cm^2]$	Keldysh-like impact-ionization reaction cross section
$\sigma_{SP,MP}^{(e/h)}$	$[cm^2]$	Keldysh-like reaction cross section for the SP and MVE mechanisms triggered by electrons or holes
$\tau_E^{n,p}$	$[s^{-1}]$	Effective energy relaxation time
$\tau_m^{n,p}$	$[s^{-1}]$	Effective momentum relaxation time
$\tau_{n,p}^{-1}$	$[s^{-1}]$	Minority carrier lifetime for electrons and holes injected in a p-type and n-type
ϕ	$[V]$	Electrostatic potential
$\Phi_{e,c}^{n,p}(x,t)$	$[cm^{-3}s^{-1}]$	Capture and emission flux for electrons and holes
ω_{op}	$[s^{-1}]$	Optical-phonon angular frequency
ω_e	$[s^{-1}]$	Phonon reciprocal life-time

Chapter 1

General context

Objectives

This chapter provides a historical background to introduce the photoelectric effect and explain why the Single-Photon Avalanche Diode (SPAD) has become a popular technology for detecting photons using solid-state semiconductor-based technologies. The architecture of the SPAD under study, along with its quenching circuit, is described. The key metrics for evaluating SPAD performance and their corresponding characterization techniques are also discussed. The chapter then turns to 3D imaging technologies, describing the latest developments and their widespread applications in various fields. The market shares of these technologies are also discussed, with a particular focus on the mobile market and the different customers and suppliers involved. Finally, the objectives of the thesis are defined.

Contents

1 Historical background

1.1 From electromagnetism to photoelectric effect

In 1873, Maxwell published about electromagnetic disturbances that should propagate through vacuum at the speed of light and should exhibit the wave-like characteristics of light propagation [1]. Maxwell states that oscillation of charges loosens metallic electrons leading of electron emission. This phenomena depends upon the frequency (rate) and intensity (amplitude) of light as illustrated in **Fig. 1.1**. The intensity affects the velocity of ejected electrons and the oscillation affects the time taken for electrons to be ejected.

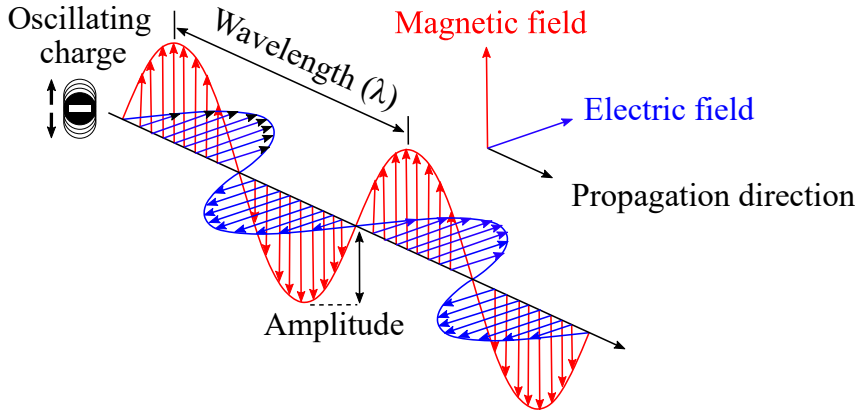


Figure 1.1: Illustration of an electromagnetic wave, characterized by its amplitude and wavelength, encountering a medium which causes an electron to start oscillating.

In 1886, Hertz used a spark-gap transmitter, which is depicted in **Fig. 1.2**, to produce a spark discharge between two pieces of brass to produce radiation of electromagnetic waves [2] according to Maxwell's theory. These waves traveled through the air and created a spark in the middle of a spark-gap receiver made with a piece of copper wire bent into a circle with small brass spheres at the two ends. The waves spreads in all directions straight lined, exactly following the geometric-optical laws of the motion of light, including reflection. Hertz was able to detect the radiation up to fifty feet away before receiver spark starts decaying. He discovered that the length of the sparks produced by his spark gap circuits varied with frequency (oscillation of charges) which could be tuned by inductive and capacitive components. By enclosing the experimental setup in a dark box, the receiver spark length became smaller. The receiver spark length increases in rarefied air. He discovered that a sheet of glass effectively shielded the spark and a slab of quartz did not. He used a quartz prism to break up the light from the transmitter spark into its component colors, and discovered that the wavelength which made the receiver spark more powerful was beyond the visible, in the ultraviolet. The cathode of the spark plug gaps is the actual formation-place of the photoelectric effect.

In 1888, Hallwachs set a circular plate of zinc which was mounted on an insulating stand and was attached by a wire to a gold leaf electroscope charged negatively [3] as described in **Fig. 1.3**. The electroscope lost its charge very slowly in absence of light. However, if the zinc plate was exposed to ultraviolet light from an carbon arc lamp, charge leaked away quickly and

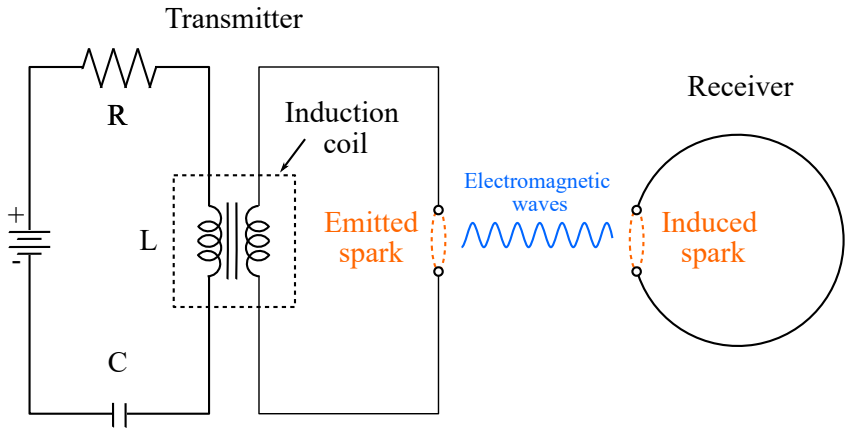


Figure 1.2: Representation of a spark-gap transmitter including a induction coil to modulate the emitted spark from the transmitter to the receiver.

the leaves will open. If the plate was negatively charged, there was fast charge leakage and the leaves will open further. If the plate was positively charged, there was no fast charge leakage and the leaves will close. Plates from different metals show differently strong effect. To sum up, the effect is created at the plate‘s surface by means of ultraviolet light and not by red light or thermal radiation.

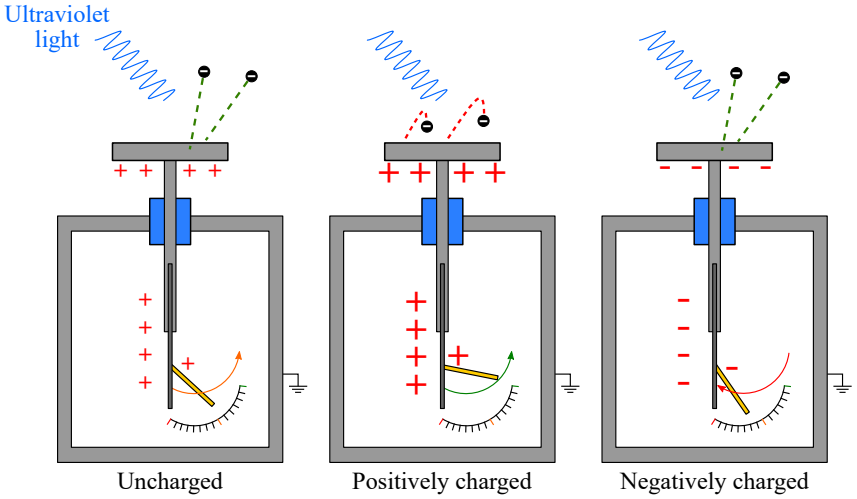


Figure 1.3: Illustration of an electroscope which aims to detect the presence and magnitude of an electromagnetic field. The position of the gold leaf changes in relation to the metal rod due to the interaction between the electromagnetic field and the charges in the metal plate.

In 1889, Thomson identified these sparks as excited electrons that left the metal surface after gaining enough kinetic energy [4] and not an ionization effect because of the air's increased conductivity, since the phenomenon took place in a vacuum. His method was to enclose the metallic surface to be exposed to radiation in a vacuum tube. Electrons were to be ejected from the cathode by the radiation, rather than by the strong electric field used previously.

From 1900 Lenard's experiment [5] followed by Millikan's experiment [6] investigated photoelectric effect using carbon arc light of different frequencies (component color separation) and

intensities shining an evacuated glass tube with several electrodes developed by Elster and Geitel in 1893 [7]. As illustrated in **Fig. 1.4**, to measure the energy of the ejected electrons, collector plate was charged negatively to repel the electrons coming towards it and the emitter plate was charged positively to attract them. Thus, only electrons ejected with enough kinetic energy to get up this potential would contribute to the current. Based on Maxwell's postulates, higher intensity of light results in faster ejected electrons requiring higher voltage to stop them and higher frequency of light means shorter emission time leading to higher current. This experiment showed that light intensity has no effect on stopping voltage and the stopping voltage increases with frequency. The frequency had no effect on current and the current increases with intensity. Light below a certain frequency cannot eject electrons no matter how higher the intensity thus the energy irradiated per second. These conclusions are inconsistent with predictions made by Maxwell's wave model of light. Lenard added to his experiment other electrodes to drive the magnetic field within the vacuum tube. He showed that the current is gifted with inertia determined in a vertically directional magnetic field. He found that the charge-to-mass ratio of the ejected particles is the same as that for electrons.

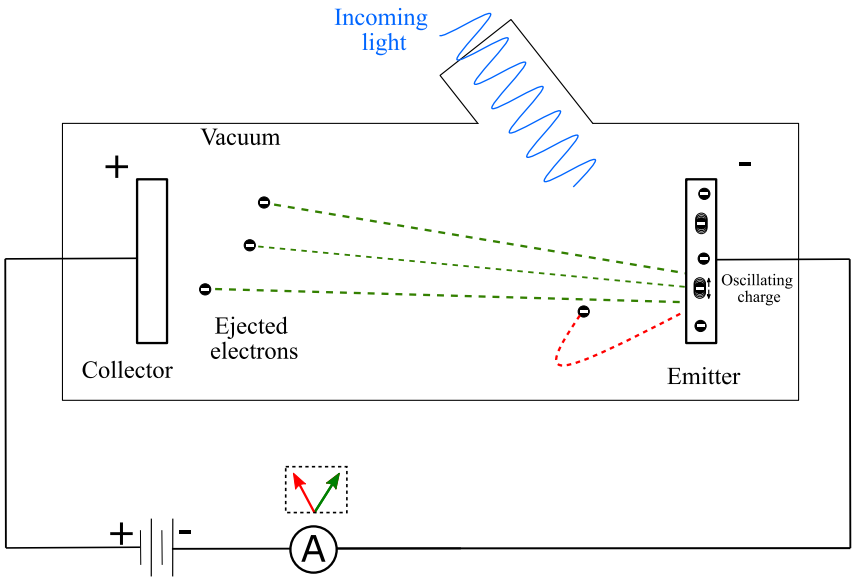


Figure 1.4: Schematic of Lenard's experiment which includes a negatively-charged emitter and a positively-charged collector in an evacuated glass tube aiming to repel and attract electrons ejected by the radiation. A battery and an ammeter are used to modulate the voltage to stop the electron and to measure the current.

Thanks to Einstein who formulated the photoelectric effect in 1905 [8], inventors could turn thing over in their heads to create devices allowing detection of invisible light. The quantum model of light considers light as small packets of energy (hf) called photons as illustrated in **Fig. 1.5**. When the incoming light is incident on a metal surface, the energy transfer occurs between one photon and one electron only. A photon would transfer all its energy to an electron or none at all. A photon cannot transfer its energy partially and the photon's energy determines the kinetic energy of ejected electrons E_k. A photon's energy must exceed the metal's work function (ϕ) which is the minimum energy required to dislodge a single electron as well as the

threshold frequency of light required to meet the work function, as depicted in **Fig. 1.5**. The resulting kinetic energy and velocity of ejected electrons are dependent on energy of photon and work function of metal according to the law of conservation of energy ($E_k = hf - \phi$). Based on Einstein's explanations, photons with higher frequencies have greater energy, which is then transferred into kinetic energy when they interact with electrons. This means that more work, represented by the product of the electron charge and the voltage (qV), is required to stop the electrons when they are exposed to higher frequency photons. Intensity of light has nothing to do with energy of one photon but it is related to the number of photons. This does not affect the kinetic energy of ejected electrons which is only related to the frequency of light thus has no effect on stopping voltage. Higher intensity leads to more electrons ejected therefore the current passing through the ammeter will increase. The frequency affects the energy of one photon and not the number of electrons ejected from the metal. Einstein was able to provide an explanation for the Lenard's experiment based on the photoelectric effect which showed that the number of electrons emitted increased with the intensity of the light, but the kinetic energy of the electrons did not depend on the intensity of the light, but rather on its frequency.

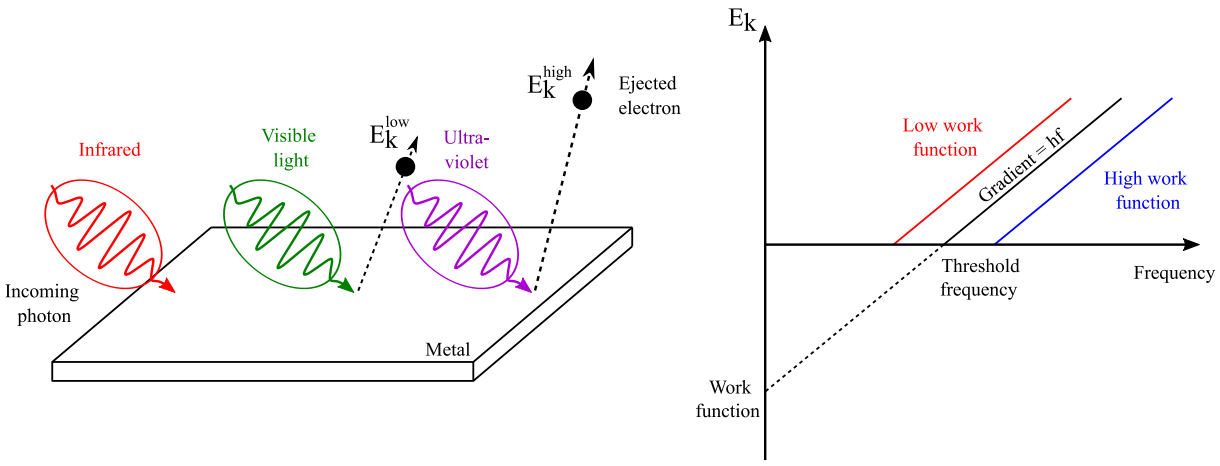


Figure 1.5: Illustration of photoelectric effect with kinetic energy energy of ejected electron as a function of photon frequency. The emission of electrons increases with the intensity of the light, which defined the number of photon packets, but the kinetic energy of the electrons is determined by the frequency of the light.

Conclusions

Maxwell's theory suggests that the oscillation of charges can lead to the emission of electrons from a metal surface, with the intensity of the oscillation affecting the velocity of the ejected electrons and the duration of the oscillation affecting the time taken for electrons to be ejected. Hertz used a spark-gap transmitter to demonstrate this effect, with the cathode of the spark plug gaps being the actual formation-place of the photoelectric effect. Hallwachs used a gold leaf electroscope to show that the effect is created at the plate's surface by means of ultraviolet light and not by red light or thermal radiation. Thomson demonstrated that electrons were ejected from the cathode by radiation rather than by the strong electric field used previously. Lenard and Millikan found that the frequency of light did not affect the current, but the current increased with intensity. Additionally, light below a certain frequency could not eject electrons regardless of the intensity. Lenard discovered that the charge-to-mass ratio of the ejected particles was the same as that for electrons. Einstein later formulated the photoelectric effect, providing an explanation for Lenard's experiment which showed that the number of emitted electrons increased with the intensity of the light, but the kinetic energy of the electrons depended on the frequency of the light rather than its intensity.

1.2 From Photo-Multiplier Tube to semiconductor photodiode

New developments in the detection of photons, triggered by particle and nuclear physics led to photodetector development. Photodetectors are divided in two main categories as illustrated in **Fig. 1.6**: photoemissive based on an incident photon that emits an electron and photoconductive rest upon incidents photons that change the conductivity.

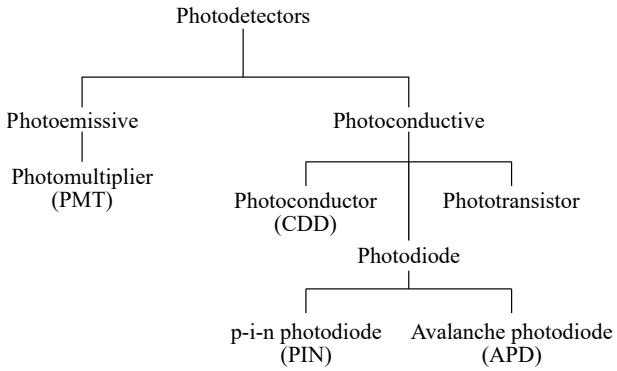


Figure 1.6: Photodetector categories classified based on their operating principles either photoemeissive or photoconductive.

1.2.1 Photo-Multiplier Tube

In 1930, the first Photo-Multiplier Tube (PMT) was invented by Kubetsky [9] and Zworykin developed a PMT with electrostatic focusing in 1939 [10], which is depicted in **Fig. 1.7**. Upon

photon absorption, primary electrons are ejected from the photocathode based on the photoelectric effect. Electron trajectory is influenced by the electric field which is dominated by the electrode configuration. Primary ejected electron are accelerated by the focusing electrode to the first emitter knocking out from its surface secondary lattice electrons based on impact ionization (electrons are excited into higher energy levels during collisions) [11]. The latter are accelerated to the next emitter knocking out secondary electrons. Consecutive secondary electrons are emitted from the dynodes and each subsequent emitter is supplied with higher voltage in comparison with preceding emitter. The degree of multiplication is a function of the interdynodal potentials which accelerate the electron beam. The higher the number of emitter, the higher the gain. The gain in PMT is defined as the ratio of the number of output electrons to the number of input photons. The gain of a PMT can range from a few hundred to millions, depending on the number of dynodes and the voltage applied. The electron flux from the last emitter are collected by collector or anode. The main limitation was the strong magnetic field imposed in high energy experiments which change the carrier trajectories within the PMT.

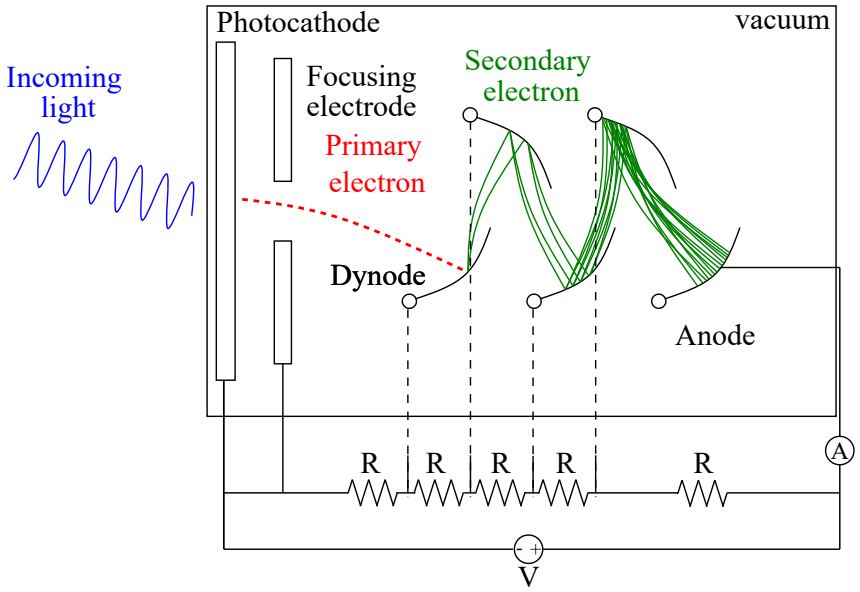


Figure 1.7: Representation of a Photo-Multiplier Tube (PMT) which is composed of a photocathode, which utilizes the photoelectric effect to convert incident photons into primary electrons, a focusing electrode, which accelerates primary ejected electrons towards a series of dynodes, where they undergo impact ionization to produce a cascade of secondary electrons and an anode to collect secondary electrons. The gain of the PMT is determined by the number of dynodes in the tube and the voltage applied to the tube.

1.2.2 p-n junction operating modes

Semiconductor devices are manufactured in mass standard production, consume few room and are less sensitive to magnet fields. Semiconductor materials can be doped by implanting acceptor and donor species. When the material is doped with donors (n-type), more electrons are added in the substrate lattice. Electrons are the majority carriers and holes are the minority carriers. When the material is doped with acceptors (p-type), less electron thus more holes (absence

of electrons) are added in the substrate lattice. Holes are the majority carriers and electrons are the minority carriers. Charge transport in semiconductor flow through diffusion and drift mechanisms as outlined in **Fig. 1.8**. Drift current is generated by an electric field that can be build-in when p-type and n-type compounds are stacked or build-up with an externally applied voltage. Diffusion current is driven by the carrier concentration gradient. When a p- and n-layers are stuck together the majority carriers diffuse into the other region and recombine depleting the transition region. Ionized donors and acceptors that remain build-in an electric field that counterbalance the diffusion through the junction giving no net current flow. Other aspect such as generation and recombination centres from defects that drive creation and annihilation of charge carriers must be considered, which may result in leakage current as explained in the next chapter.

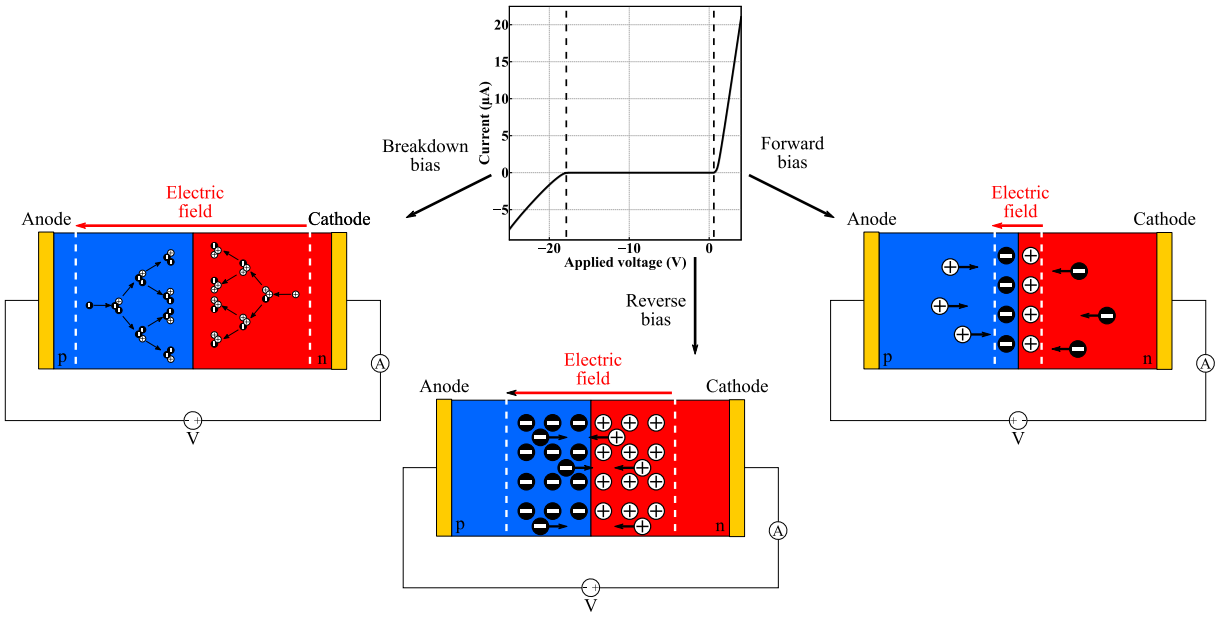


Figure 1.8: Illustration of the different working regimes of a p-n juction. The studied SPAD architecture is characterized by a current-voltage curve that is normalized by a reference voltage. The figure depicts from right to left the diffusion of majority carriers in forward bias (a), the attraction of minority carriers by opposite electrodes in reverse bias (b), and the occurrence of minority carrier avalanche in breakdown bias (c).

The typical operating modes of a p-n junction will now be described. By applying an external forward bias - positive polarity on p-side and negative one on n-side - some majority carriers travel across the depletion region from both sides owing to shorten of the potential barrier of the diode as drawn in **Fig. 1.8**. Without opposite carriers to recombine they diffuse a mean free path before recombining or being collected by the electrodes. The positive terminal attracts n-type majority carriers away from the junction (behaves as cathode) and p-type ones for negative terminal (acts as anode) respectively. Therefore, the diode is a unidirectional device, the current only flow in one direction.

By applying an external reverse bias - positive polarity on n-side and negative one on p-side - minority carriers are generated in the depletion region from both sides as depicted in **Fig. 1.8**.

As the minority carriers are very less with respect to majority carriers the magnitude of the current will be very low compare to the forward bias. Minority carriers experience repulsive force from the external electric field. The positive terminal attracts p-type minority carriers away from the junction (behaves as anode) and n-type ones for negative terminal (acts as cathode) respectively.

When a sufficient reverse bias is applied, the electric field magnitude is high enough to allow charge carriers to gain enough kinetic energy to eject electrons from the valence shell of the atoms resulting in a net increase in current flow. First of all, this carrier must overpass the dead space which the distance a carrier requires to achieve threshold energy to impact ionize an electron-hole pair. This pair and the primary electron together will repeat the same process ionizing other electron-hole pair. This is called avalanche breakdown, and it will be explained in more details in **Chapter 2**.

1.2.3 PIN photodiode

Silicon PIN photodiode was invented by Watanabe and Nishizawa in 1950. This diode has an intrinsic undoped layer sandwiched in between two highly doped p- and n-type semiconductor as represented in **Fig. 1.9**. By applying a positive voltage on the p-side and a negative voltage on the n-side, the diode starts to deplete the intrinsic region which is to say to remove majority carriers diffusing from both side. Electrons generated in the intrinsic region move towards n-side whereas holes generated in the intrinsic region move towards p-side. When a photon is absorbed in the semiconductor bulk, it generates an electron-hole pair. The carrier are then collected at the electrodes as in the absence of light. The wide intrinsic area between the n and p layers allow for more light to be absorbed. The detection of photoabsorption events is constrained by the intrinsic noise, also referred to as dark current, that arises due to the generation of carriers from defect centers and the diffusion of majority carriers. In the case of a PIN photodiode, the gain is limited because it lacks an internal gain mechanism. This means that the current generated by the absorption of photons is directly proportional to the incident light intensity, resulting in a weak output signal that requires external amplification to be useful. However, external amplification can introduce noise and decrease the signal-to-noise ratio, which can limit the effectiveness of the PIN photodiode.

1.2.4 Avalanche photodiode

In 1952, Nishizawa introduced the Avalanche Photo-Diode (APD), which operates based on the principles of carrier multiplication observed in reverse biased p-n junctions without avalanche triggering (as illustrated in **Fig. 1.8**) using a photoabsorption layer as in PIN photodiodes (as depicted in **Fig. 1.9**). The APD combines these two mechanisms to achieve high gain and light sensitivity. In contrast with PIN photodiode, an APD has an internal multiplication gain mechanism that amplifies the signal generated by the absorption of photons, resulting in a larger output signal with higher gain. **Chapter 2** will provide an in-depth explanation of how the carrier multiplication factor, which is a measure of the gain in these photodiodes, is greatly influenced by various factors such as the electric field strength, temperature, and doping profile.

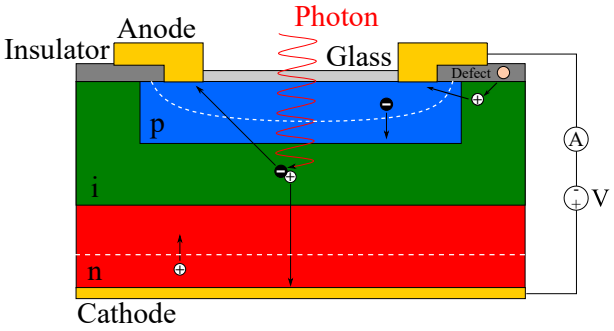


Figure 1.9: Representation of a PIN photodiode which is composed of a p-type layer, an intrinsic layer, and an n-type layer. When light is incident on the PIN photodiode, it generates electron-hole pairs in the intrinsic layer, which are then separated by the electric field created by the p-n junction.

A low-noise preamplifier with a high gain-bandwidth product, is required to detect the brief and weak current pulses generated by a photon [12]. The fluctuation of the multiplied photo-current must be higher than the threshold of the preamplifier.

Single-Photon Avalanche Diode (SPAD) exploits the avalanche current of a reverse biased p-n junction (as shown in **Fig. 1.8**) to detect an incident radiation. The key difference between SPADs and APDs is that SPADs are designed to operate with a reverse bias voltage well above the avalanche breakdown voltage, while APDs operate at a bias voltage lower than the avalanche breakdown voltage. A SPAD operating in this high-gain region is sometimes referred to as being in Geiger mode while the APD is say to be in the linear mode. The resolution of a SPAD is mainly limited by the Dark Count Rate (DCR). DCR refers to the rate at which a SPAD generates a signal in the absence of any incident light. It is important to note that while both Dark Count Rate (DCR) and dark current may originate from the same sources, such as the generation of carriers from defect centers and the diffusion of majority carriers, they differ in that DCR triggers carrier avalanches, whereas dark current is a residual current without avalanche triggering.

The early development of silicon SPADs involved the creation of monolithic single-element devices using custom technological processes. Two main architectures were studied at that time: the planar architecture developed by Cova in the 1980s [13] and the "reach-through" type architecture developed by the McIntyre team [14]. Matrices of SPADs have been developed using dedicated processes, including the work of Kindt in the late 1990s. However, these matrices suffered from optical crosstalk, where photons emitted during the avalanche process could trigger avalanches in neighboring devices. This resulted in false event detections and reduced spatial resolution, which constrained the smallest distinguishable detail in a scene that the imaging system could resolve.. To overcome this issue, Kindt introduced insulated trenches covered with metal between pixels, as described in [15]. Rochas demonstrated in the 2000s [16] that co-integrating SPADs with their processing circuits in standard Complementary Metal-Oxide-Semiconductor (CMOS) technology could lead to low-cost production of SPADs integrated into matrices, taking advantage of technological advances and miniaturization. The pixel pitch, which refers to the distance between adjacent pixels, is closely related to the size of the CMOS

technological node used. As the CMOS node size decreases, the pixel pitch can also be reduced, allowing for photodetectors to be packed more densely and resulting in higher spatial resolution. In addition, this approach also simplified the coupling of quench circuits with SPADs and enables a significant reduction in parasitic capacitances compared to hybrid integrations. This helped to minimize the dead time, which refers to the period during which a detector is recovering from the previous detection event and cannot detect another photon, as elaborated below.

The integration of SPADs into technological nodes began with the 800 nm node, and subsequent research works focused on adapting and optimizing SPADs for increasingly fine technological nodes. Notable contributions came from various teams, including Charbon's team [17][18][19], the Polytechnic School of Milan [20][21], the University of Edinburgh [22][23][24], and the Bruno-Kessler Foundation together with the University of Trento [25][26]. However, as technological nodes decreased below 250 nm, the depletion zones of SPADs became increasingly thin, resulting in noise contributions dominated by tunneling effects. To address these challenges, new architectures, such as Richardson's one [22], were developed. The thinnest technological nodes achieved to date are a 40 nm CMOS technology SPAD from STMicroelectronics by the Pellegrini team [27], a 28 nm CMOS SPAD from STMicroelectronics by Albuquerque [28], and a 22nm CMOS SPAD from Sony by Shimada [29]. These advancements in low CMOS technological node sizes have enabled pixel pitch to decrease to 10.17µm [30], 6.39µm [31], 6µm [29] and 2.5µm [32], thereby significantly enhancing the spatial resolution. Research is currently centered on the integration of SPAD technology into CIS [33][34][35] to take advantage of the latest developments in CMOS imagers. The SPAD electronics, such as the quench circuit, are integrated in a planar manner next to or alongside the SPAD structures resulting in a low Fill Factor (FF), which is the ratio of the active area of the detector to the total area of the detector. The most effective approach for SPADs in the current era of pushing CMOS technology limits is 3D integration microelectronics technology. This method involves using two silicon wafers, with one dedicated to the SPAD matrix structure (top tier process) and the other for the operating electronics (bottom tier process) [36] increasing the FF. Since 2013, research has focused on integrating SPAD into Silicon-On-Insulator (SOI) technology [37] which is based on the use of two electrically isolated silicon layers separated by a buried oxide layer, which helps to reduce parasitic capacitance and crosstalk between components.

Conclusions

Photoemissive devices are based on an incident photon that emits an electron, while photoconductive devices rely on incident photons that change the conductivity. The degree of multiplication in a Photo-Multiplier Tube (PMT) is determined by the interdynodal potentials that accelerate the electron beam, with the main limitation being the strong magnetic field. In contrast, p-n junctions in mass standard production consume less space and are less sensitive to magnetic fields, with operating modes including the diffusion of majority carriers in forward bias, the attraction of minority carriers by opposite electrodes in reverse bias, and the occurrence of minority carrier avalanche in breakdown bias. PIN photodiodes, which are reverse biased with a depleted wide intrinsic area to absorb photons, result in a weak output signal that requires external amplification, which can introduce noise. On the other hand, an Avalanche Photo-Diode (APD) is a reverse biased p-n junction without avalanche triggering, with the gain limited by the carrier multiplication factor. A low-noise preamplifier with a high gain-bandwidth product is required, with the fluctuation of the multiplied photo-current needing to be higher than the threshold of the preamplifier. A Single-Photon Avalanche Diode (SPAD) is a reverse biased p-n junction with avalanche triggering, and their development is related to Complementary Metal-Oxide-Semiconductor (CMOS), which allows for a reduction in pixel pitch increasing spatial resolution. Current research in this area is focused on 3D integration to reduce the fill factor, CMOS Image Sensor (CIS) to take advantage of the latest developments in CMOS imagers, and Silicon-On-Insulator (SOI) to reduce parasitic capacitance and crosstalk between components.

2 SPAD architecture under study

After introducing the historical background and the integration challenge of SPADs into technological nodes, the architecture used in this work will now be discussed. A typical representation of the studied SPAD architecture [38][39] is presented in **Fig. 1.10**. The p-epi layer acts as an absorbing layer that generates carriers when it absorbs photons. The generated electrons move towards the n+ layer, while the generated holes move towards the p++-layer. It is important to keep the photo-generation process confined to this region in order to minimize excess noise that can arise from the generation of carriers from defect centers and the diffusion of majority carriers outside of this region. The application of a sufficiently high voltage to the SPAD results in the collection of drifting electrons by the p+ and n++ layers. These electrons acquire enough kinetic energy to impact-ionize some of the valence shell electrons of the atoms, generating electron-hole pairs and initiating a self-sustaining avalanche of carriers. To minimize absorption from the medium, these layers must be kept thin. The avalanche breakdown voltage of the SPAD is mainly determined by the thickness of this multiplication layer. Guard rings consisting of p- and n-well layers are implemented to increase the required breakdown voltage in the periphery in comparison to the central region, thereby effectively confining the avalanche

at the center. The reduction of doping at the edges leads to an increase in the threshold electric field for impact-ionization. n++ and p++ contact layers guarantee the establishment of ohmic contacts with electrodes to provide a low-resistance path for current.

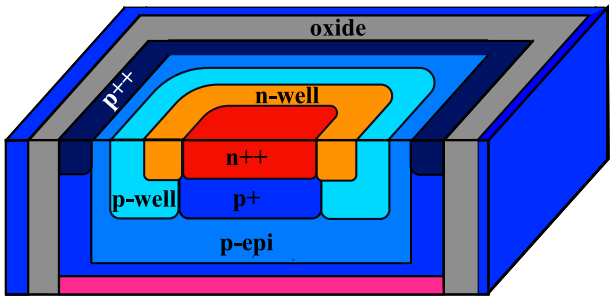


Figure 1.10: Representation of the studied Single-Photon Avalanche Diode (SPAD).

3 Active and passive quenching circuits

When an avalanche is triggered in a SPAD, it generates a large current that results in a potential drop within the device that can be measured as an output signal pulse. However, to avoid physical breakdown, the photodetector signal current must be limited and rapidly reduced. For this purpose, both active and passive quenching circuit techniques have been used [40][41] to suppress the avalanche by lowering the reverse bias of the diode below the breakdown voltage. In a Passive Quenching Circuit (PQC), illustrated in **Fig. 1.11 (a)**, a resistor or a transistor is placed in series with the SPAD. The avalanche current discharges the junction and stray capacitances so that the voltage and current exponentially fall toward the asymptotic steady-state values. The quenching time can be adjusted by changing the values of the resistor or capacitor used in the circuit. However, the adjustment range is limited compared to Active Quenching Circuit (AQC), which can provide more precise control over the quenching time. In a AQC, depicted in **Fig. 1.11 (b)**, a threshold comparator is used to detect the current pulses that pass through the resistor R_S at each avalanche event. The PMOS transistor allows quenching and the application of the quenching potential V_{QUENCH} to the anode contact. The reverse voltage across the diode is brought to a value close to or below the breakdown voltage. Then, in a second phase, the PMOS transistor is deactivated, and the NMOS drives the recharge of the diode by bringing the potential of the anode to ground, resulting in active recharge. AQC allow for fast recharge after the avalanche and offer the possibility of controlling the dead time, but this can reduces the pixel FF depending on the circuit integration.

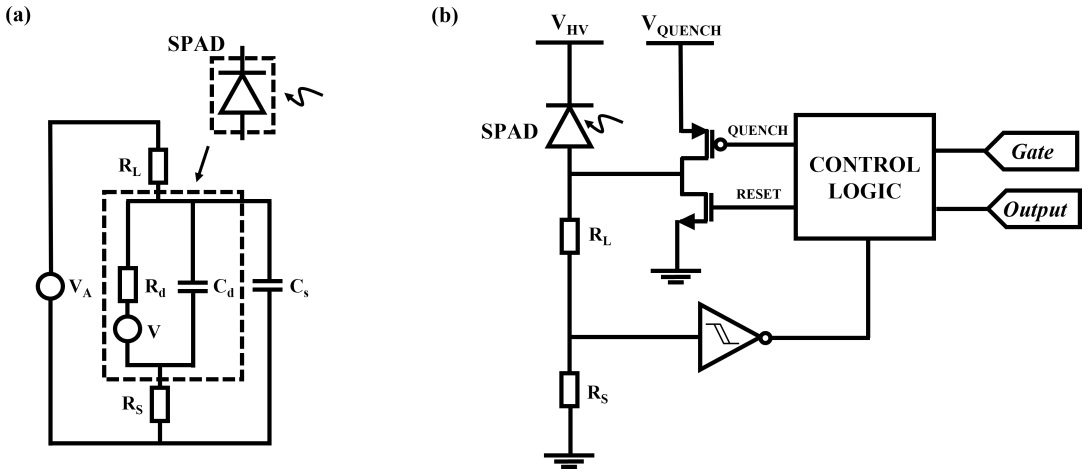


Figure 1.11: Representation of (a) Passive Quenching Circuit (PQC) [40] and (b) Active Quenching Circuit (AQC) [41]. PQC relies on the resistor or capacitor values to determine the quenching time, while AQC uses a threshold comparator to detect the pulse and an inverter to quench and reset quenching time to control dead time.

4 Electrical circuit: From passive-quenching to pulse shaping

As previously stated, it is necessary to design a quenching circuit to ensure that the avalanche is suppressed and the system returns to steady state. The features of the passive quenching circuit employed represented in **Fig. 1.12** are outlined in **Figs. 1.13-1.14**, with the description starting from the lower left and progressing towards the upper right components.

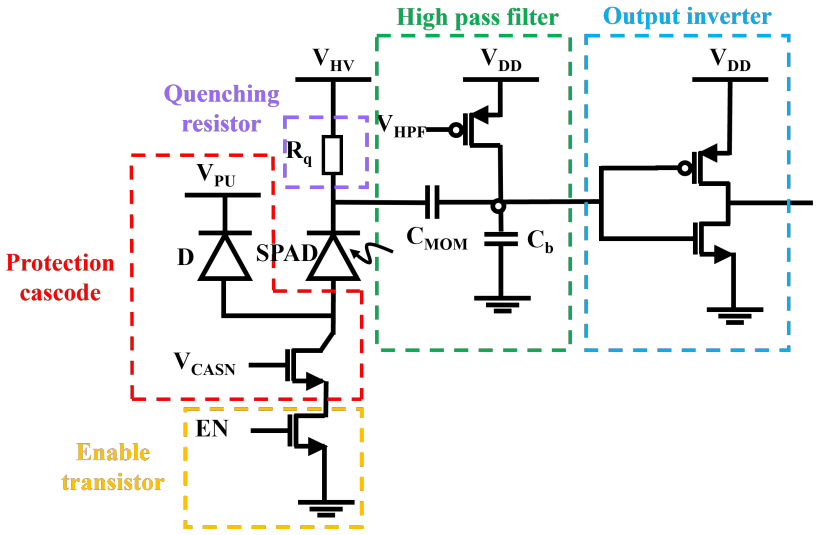


Figure 1.12: Schematic of the passive quenching circuit used which is composed of an enable transistor, which switch on/off SPAD, a protection cascade, which protects the enable transistor (EN), a quenching resistor (R_q), which guarantee return to steady state after avalanche triggering, a high pass filter, which shapes the signal, and an output inverter, which detects the signal.

As shown in **Fig. 1.13**, a combination of a low-voltage n-type enable MOSFET, a high-voltage cascode n-type MOSFET, and a clamp diode, all connected to the anode of the SPAD,

allows to disable the defective pixels exhibiting high DCR and the array sectors that are not addressed or that are non-illuminated. When the enable transistor (EN) is turn off, the anode voltage is allowed to rise more than the excess bias, putting the SPAD safely out of the Geiger mode. However the anode voltage is higher than the transistor voltage tolerance, as shown in **Fig. 1.13 (a)**. A high-voltage transistor V_{CASN} is added to protect the enable transistor and to allow anode to reach high-voltage when enable transistor is off, as illustrated in **Fig. 1.13 (b)**. For the pixel to be completely off, the potential difference between the cathode and the anode of the SPAD must be below V_{BD}. To consider process device-to-device variations, the high-voltage transistor must tolerate voltage higher than the applied voltage plus 3σ of V_{BD} variation. Consequently when the enable transistor is off, anode voltage can reach more than extended drain voltage tolerance. A clamp diode (D) is then used to protect the high-voltage transistor, as depicted in **Fig. 1.13 (c)**. When applied voltage is above V_{PU}, the potential drop at the anode of the SPAD will be clamped to the forward conduction of the pull-up diode to guard components against transient voltages and high current density.

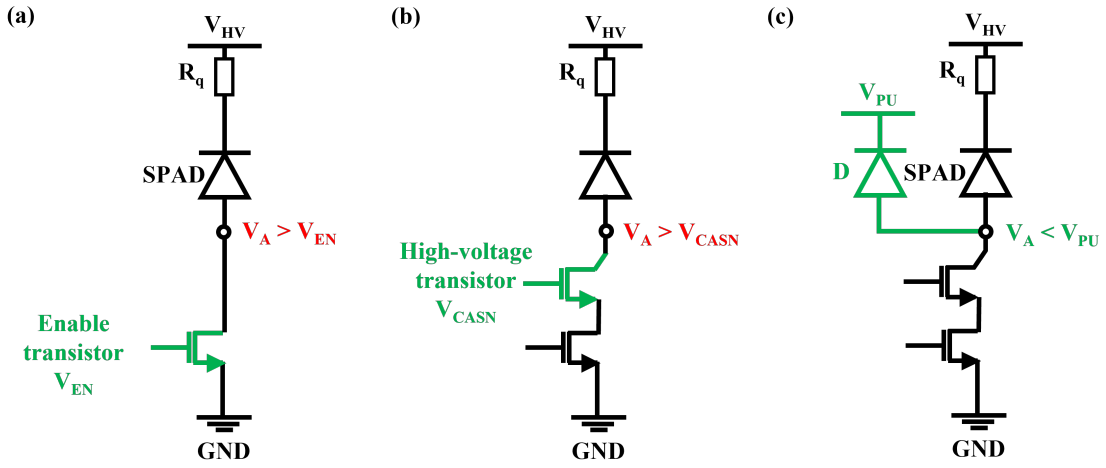


Figure 1.13: Schematic of the protection cascade which is used to protect the enable transistor (EN) and to ensure that anode swing is below the clamp diode (D) to protect the extended drain transistor.

A polysilicon resistor (R_q) is buried in Deep Trench Isolation (DTI) to ensure quenching of avalanches. The cathode node swings swings between V_{HV} and V_{BD} as shown in **Fig. 1.14 (a)**. Sensing directly at the cathode is not possible due to the tolerance of transistors composing the inverter. Adding coupling capacitance C_{MOM} stops high-voltage DC and only couples the high-frequency AC part of the signal, as illustrated in **Fig. 1.14 (b)**. A ballast capacitance C_b can be added in parallel to add more attenuation acting as capacitive voltage divider between the cathode of the SPAD and the input of the inverter, which is used to generate the pulsed output signal of the pixel. In effect, the cathode swing can be adjusted by the ratio of C_{MOM} to C_b max value limit. Floating node at the input of the inverter is pulled up to V_{DD} by a pulse shaper p-type MOS to ensure crossing the inverter threshold, as depicted in **Fig. 1.14 (c)**. When the pulse shaper transistor is switch on, this leads to a large output pulse at the inverter input allowing to filter the signal. Whereas switch off, the output signal is smaller and the signal

filtering is deactivated. This option is preferred for the visualization of the afterpulsing/lag phenomenon, as discussed later on. In effect, the pulse width is depending on V_{HPF} settings, but also on the inverter setup (threshold), and the dynamics of the sensing node (total RC, the ratio C_{MOM}/C_b, the excess bias, etc.).

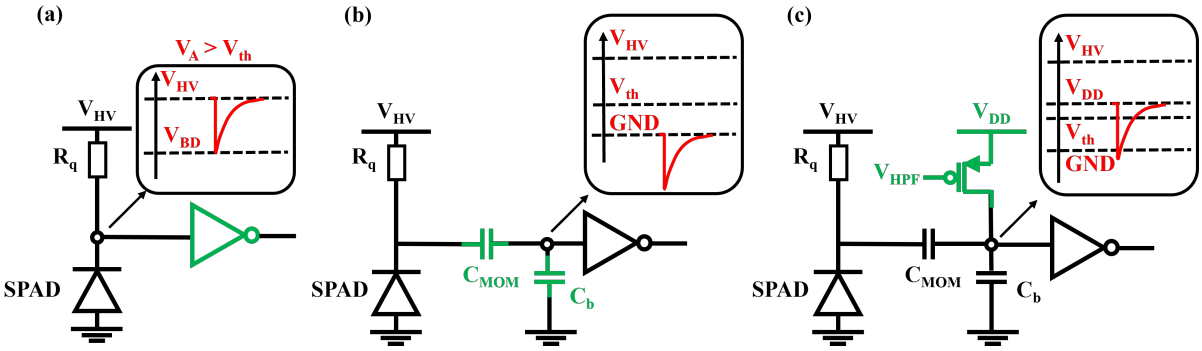


Figure 1.14: Schematic of the high-pass filter which is used to shape the signal and to ensure that the cathode swing is within the the tolerance of transistors composing the output inverter.

The connection of an Agilent 6000 Series Oscilloscope to the inverter output yields typical signals that are displayed in various time windows, as shown in **Fig. 1.15**. Typical signal and afterpulsing pulses are also reported in **Fig. 1.15 (b)**. The signal can be impacted by reflection inside the line as shown on the scope trace by these characteristic oscillations. In effect, each oscillation slope represents the round trip inside the line due to the resonance of parasitic impedance. This impedance arises from the whole area of the probe card which is much greater than the parasitic capacitance from coaxial cable. The line is said to be not adapted owing to the input and output impedances which are not designed to counterbalance the intrinsic one. After several round trips, the reflected signal will lose its amplitude tending toward steady state. The observation of a characteristic bump immediately following the signal recovery indicates the presence of a parasitic line located in close proximity to the probing point. It seems possible to overcome this issue by setting the trigger level at 700mV. A Keysight 53131A Universal Frequency Counter/Timer is plug-in instead of an oscilloscope in the following to average the number of counts on longer time scale. As mentioned earlier, the voltage level of the pulse shaper transistor V_{HPF} can be adjusted, as shown in **Fig. 1.15 (c)**, to increase the width of the output pulse and enable the observation of afterpulsing or lag phenomena.

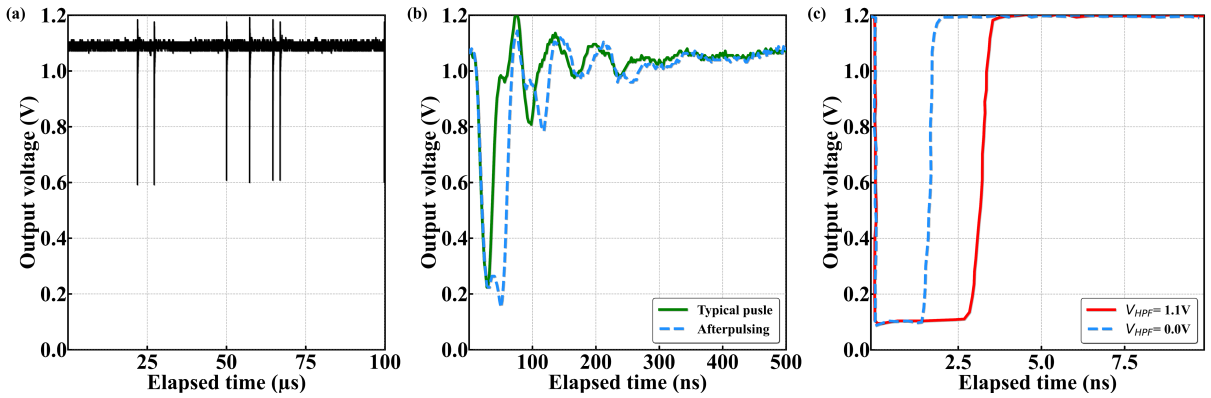


Figure 1.15: Signal pulse shape measured by an oscilloscope in different time windows. (b) Typical signal and afterpulsing pulses are displayed. (c) The voltage level of the pulse shaper transistor V_{HPF} is varied to demonstrate the changes in signal width.

Conclusions

The SPAD architecture under study consists of highly doped n++ and p+ multiplication regions that confine avalanches at low biasing voltage. Lightly doped n and p guard rings prevent lateral breakdown at high biasing voltage, while a p-epi layer acts as an absorbing layer that generates carriers when it absorbs photons. To prevent physical breakdown in a SPAD due to large current, active and passive quenching circuit techniques have been developed. Passive quenching involves using a resistor or transistor in series with the SPAD to discharge the junction and stray capacitances. Active quenching, on the other hand, uses a threshold comparator to provide more precise control over the quenching time and fast recharge after the avalanche. However, active quenching can reduce the pixel fill factor depending on circuit integration. The passive quenching circuit used in this study comprises a low-voltage n-type enable MOSFET, a high-voltage cascode n-type MOSFET, and a clamp diode, all connected to the anode of the SPAD. This circuit disables pixels and guards components against transient voltages and high current density. A high-pass filter shapes the signal and ensures that the cathode swing is within the tolerance of transistors composing the output inverter. The pulse width depends on V_{HPF} settings, enabling the observation of afterpulsing or lag phenomena.

5 SPAD performances and characterization techniques

Previously, the working principle of a Single-Photon Avalanche Diode (SPAD) was explained along with its quenching circuitry. The current focus is on describing the main characteristics of SPADs, as illustrated in **Fig. 1.16**.

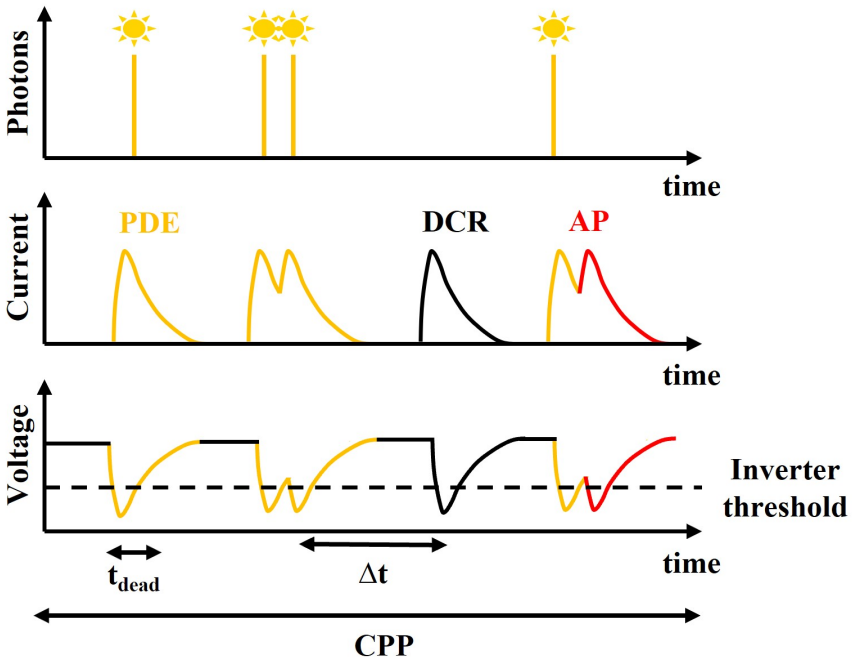


Figure 1.16: Illustration of the characteristics of a SPAD in relation to photon flux, current, and voltage pulse. The detection of photons is dependent on the Photo-Detection Efficiency (PDE), while events that occur in the absence of light are referred to as Dark Count Rate (DCR). When avalanches are triggered immediately after another avalanche, it is known as After-Pulsing (AP). The dead time, represented by t_{dead}, refers to the duration during which the detector is arming or recharging. The Charges Per Pulse (CPP), which indicates the consumption per avalanche, can be calculated by dividing the current measured during the integration time by the number of detected avalanches during that time.

5.1 Avalanche breakdown voltage

The avalanche breakdown voltage V_{BD} represents the applied voltage to trigger the first avalanches at a given temperature. The distinction between the signs of the temperature coefficients for avalanche and Zener breakdown is frequently used to distinguish between the occurrences of the two mechanisms [42]. When the temperature dependency is positive is a consequence of avalanche multiplication. The hot carriers passing through the depletion region under the influence of a high applied electric field lose part of their energy to phonons after travelling each electron-phonon mean-free path. The mean free path decreases with increasing temperature of the crystal lattice. Therefore, the carriers must move through a higher potential until they acquire sufficient ionization energy for electron hole pair. When the temperature dependency is negative, this due to tunneling. The energy gap decreases with temperature thus the probability

of tunnelling will increase. A given breakdown current can be reached at smaller reverse voltage at higher temperature.

Two methodologies can be employed to extract V_{BD} from the count rate and current as a function of the applied voltage, as depicted in **Fig. 1.17 (a)-(b)**. The first one relies on counting the occurrence of events by increasing the applied voltage toward V_{BD} with a rough voltage step. Then refine the voltage step and reduce the applied voltage toward V_{BD}. When single events are framed, V_{BD} is declared to be the reached voltage. An example is given in **Fig. 1.17 (c)** using a dichotomy method to extract V_{BD}. The number of events can be either triggered in the dark or under light reducing the V_{BD} value by higher occurrence number as shown in **Fig. 1.17 (a)**. The second methodology is based on current measurement represented in **Fig. 1.17 (b)**. By differentiating a reference current value without avalanche triggering and an increasing current value, a gain can be extracted as shown in **Fig. 1.17 (d)**. When an avalanche occurs, the gain rockets up and V_{BD} can be extracted accordingly. Again the voltage step matters in V_{BD} estimation. The breakdown voltage V_{BD} extracted from the count rate measurement will be used in the thesis.

The avalanche triggering occurs in the current before it is detected by the counter, resulting in an approximately one-volt difference between the event detection at the inverter output and the current probing at the SPAD cathode. This is explained in detail in section **Section 4** and illustrated in figure **Fig. 1.14**. In order for the inverter to reach its threshold, a minimum voltage swing is required, which is achieved by a significant voltage variation caused by the avalanche. The inverter input used for generating the output signal pulse is not directly connected to the cathode of the SPAD, but instead, capacitances C_{MOM} and C_b act as a capacitive voltage divider. To ensure crossing the inverter threshold, the floating node at the input of the inverter is pulled up to V_{DD} by a pulse shaper p-type MOS. However, when V_{HPF}=0V, the MOS is conducting, and the node is constantly pulled to 1.1V, making it harder for the node to drop sufficiently to trigger the inverter threshold, leading to an offset.

5.2 Dark Count Rate

The Dark Count Rate (DCR) is defined as the rate of counts without any incident light setting the minimum signal above which event counts are really caused by photons. It depends on the avalanche breakdown probability P_t and the Recombination-Generation (R-G). The dark count rate can be reduced by cooling the diode, but this increases the rate of after-pulses caused by trapped carriers. It must be not confused with dark current which is the instantaneous residual current not triggering avalanches in the SPAD without incident light. As illustrated in **Fig. 1.18 (a)**, the origin of DCR traces back through the various R-G mechanisms:

- Shockley-Read-Hall (SRH): an electron in the valence band can absorb a crystal lattice vibration called phonon to reach a trap state before emitting to the conduction band by interaction with another phonon.
- Poole-Frenkel effect (PF): Coulomb potential energy between bands and traps can be reduced by an applied electric field increasing the probability to be thermally excited out

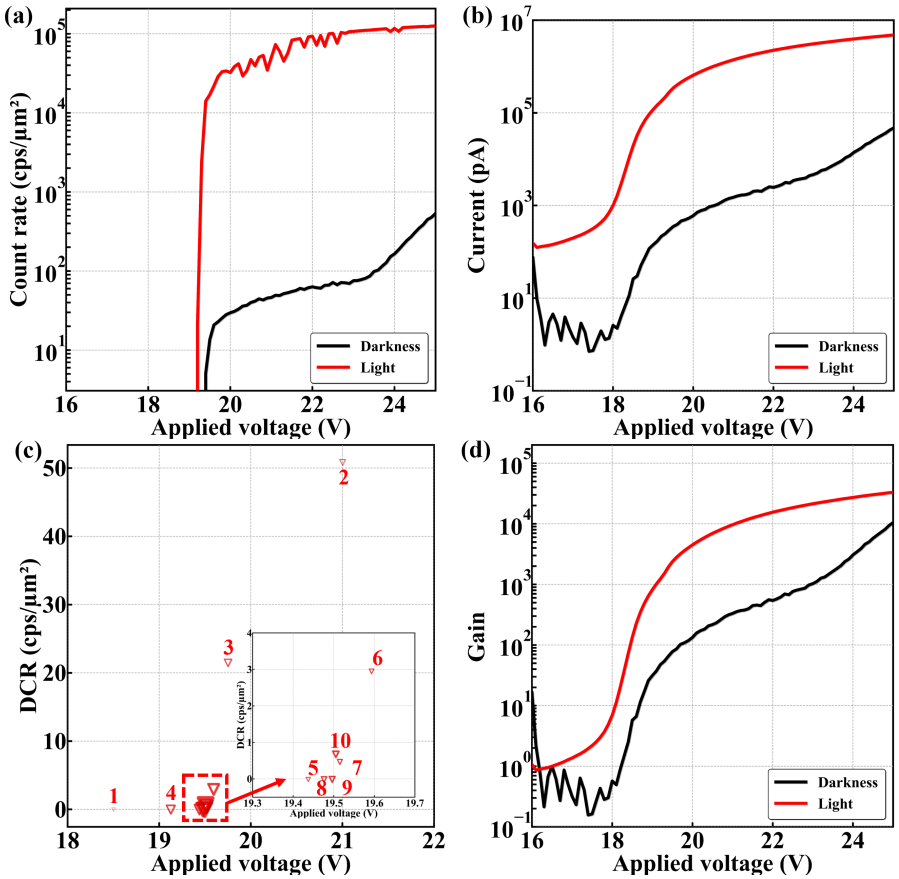


Figure 1.17: The count rate plotted against the applied voltage (a), which is normalized with respect to the avalanche breakdown voltage V_{BD}, enables the derivation of Figure (c), which illustrates the process of obtaining V_{BD} from DCR through dichotomy, with each number representing a step in the methodology. The current plotted against the applied voltage (b) facilitates the derivation of Figure (d), which is generated by taking the derivative of a reference current value obtained without triggering an avalanche, and an increasing current value obtained with avalanche, thus enabling the extraction of V_{BD}.

from one trap sit to another.

- Trap-Assisted Tunneling (TAT): firstly the electron is emitted from the valence band to the trap state by absorbing a phonon. The electron can be partially lifted further and tunnel into the conduction band in an elastic or inelastic manner.

- Band-to-Band Tunneling (BTBT): an electron in the valence band of the semiconductor tunnels across the band gap to the conduction band without the assistance of traps. The band gap acts as the potential barrier that the particle tunnels across. The most probable tunneling path is through the smallest barrier.

The thermal and tunneling contributions of DCR could be observed by a temperature sweep experiment at constant excess bias. By plotting the DCR against one over the temperature multiplied by the boltzmann constant allows to extract the Arrhenius activation energy as shown in **Fig. 1.19**. These effective thermal activation energies refer to the amount of energy

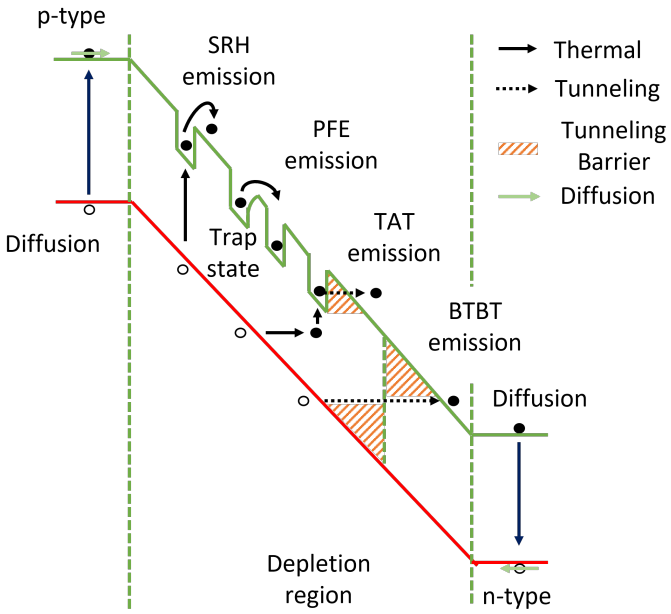


Figure 1.18: Representation of the different Recombination-Generation (R-G) resulting in DCR when coupled with avalanche breakdown probability P_t.

required to carry a carrier based on the different mechanisms. A high energy relates mainly to diffusion and SRH process. A low energy enlightens tunneling. All tunnelling phenomena are noticeable for SPAD operating with electric field profiles under MV/cm range depending on the architecture.

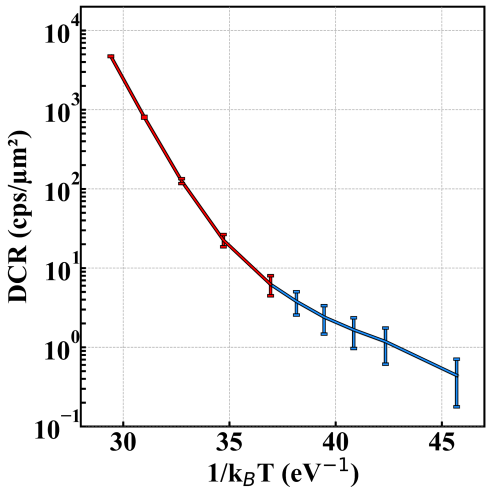


Figure 1.19: DCR as a function of $1/k_BT$ measured at 3V excess bias allows to extract the Arrhenius activation energy, which represents the energy needed to transport a carrier. The thermal activation energy extracted is equivalent to the band gap energy for high temperature and half of the band gap energy for low temperature. These thermal dependencies will be explained in **Chapter 3**.

As mentioned earlier, the Keysight 53131A Universal Frequency Counter/Timer is used to measure the DCR by plugging it in after the inverter output. However, the mean value of the

DCR is dependent on the integration time, the number of times the measure is repeated (points), and the number of devices, as shown in **Fig. 1.20**. The statistical dispersion between devices will be further elaborated in **Chapter 3** through measurement and simulation results.

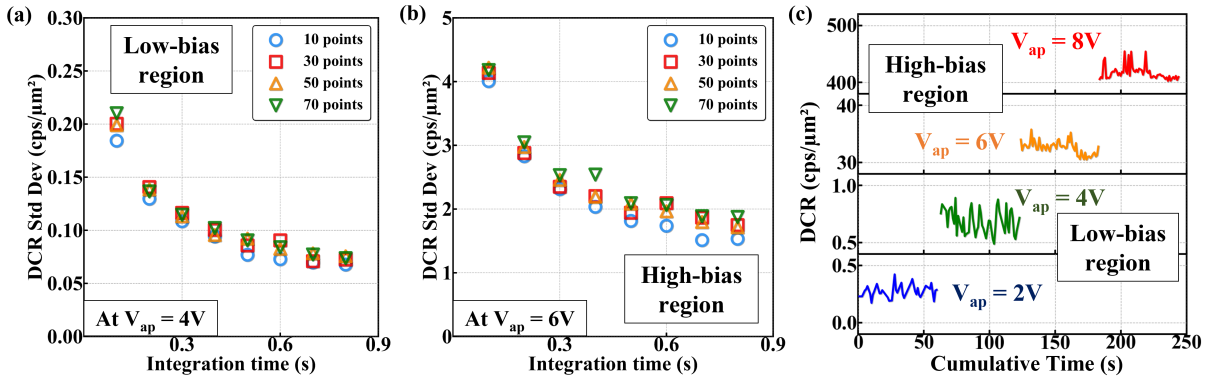


Figure 1.20: Standard deviation of DCR as a function of the integration time and the number of measurements repeated (points) in (a) Low-bias voltage (LB) and (b) How-bias voltage (HB) regions averaged on 64 SPADs. (c) The DCR is shown as a function of the cumulative time to highlight statistical dispersion between measurement.

5.3 Photo-Detection Efficiency

Quantum Efficiency (QE) is the carrier flux generated which contributed to the photocurrent over the incident photon flux. Sometimes, Photodetection Efficiency (PDE) is preferable when the signal measured is not directly correlated to incident photons for example for SPAD. This correlated the number of signal pulses detected with respect to the incident photons. PDE can be expressed as:

$$PDE = OA \cdot P_t \cdot FF \tag{1.1}$$

The Optical Absorption (OA) is the percentage of light transmitted to the media. Optical stacks are used coupling texturization pattern and microlens to optimize the light transmission to the photosensitive volume. The Fill Factor (FF) accounts for the photosensitive area over the total device area. Sometimes Photo-Detection Probability (PDP) is used instead PDE which will accounted for the device photosensitive area. However the microlens are larger that the device photosensitive area. Consequently a higher number of photons accounting for reflection, transmission, and absorption within the device are impinging the device photosensitive area than this area itself. This is why efficiency is preferable other probability because the number of photons impinging the device are higher than the ones considering only the device photosensitive area. Photon outside this area also contributes to the photo-generated signal. In other words, PDE accounts for the fraction of incident photons that are actually detected by the device, while PDP only considers the probability of detecting a single photon.

As previously shown, the Light Count Rate (LCR) is the number of detected events correlated to the light impinging the device. The PDE can be obtained by subtracting the DCR from the

LCR, multiplying the result by the integration time t_{int}, and dividing by the number of incident photons N_{ph}, which yields:

$$PDE = \frac{(LCR - DCR) \cdot t_{int}}{N_{ph}} \tag{1.2}$$

It is assumed that each photon generates only one avalanche without afterpulsing and crosstalk phenomena. The pulse shaping transistor is switch on to reduce the afterpulsing by means of pulse width widening as explained in **Section 4**.

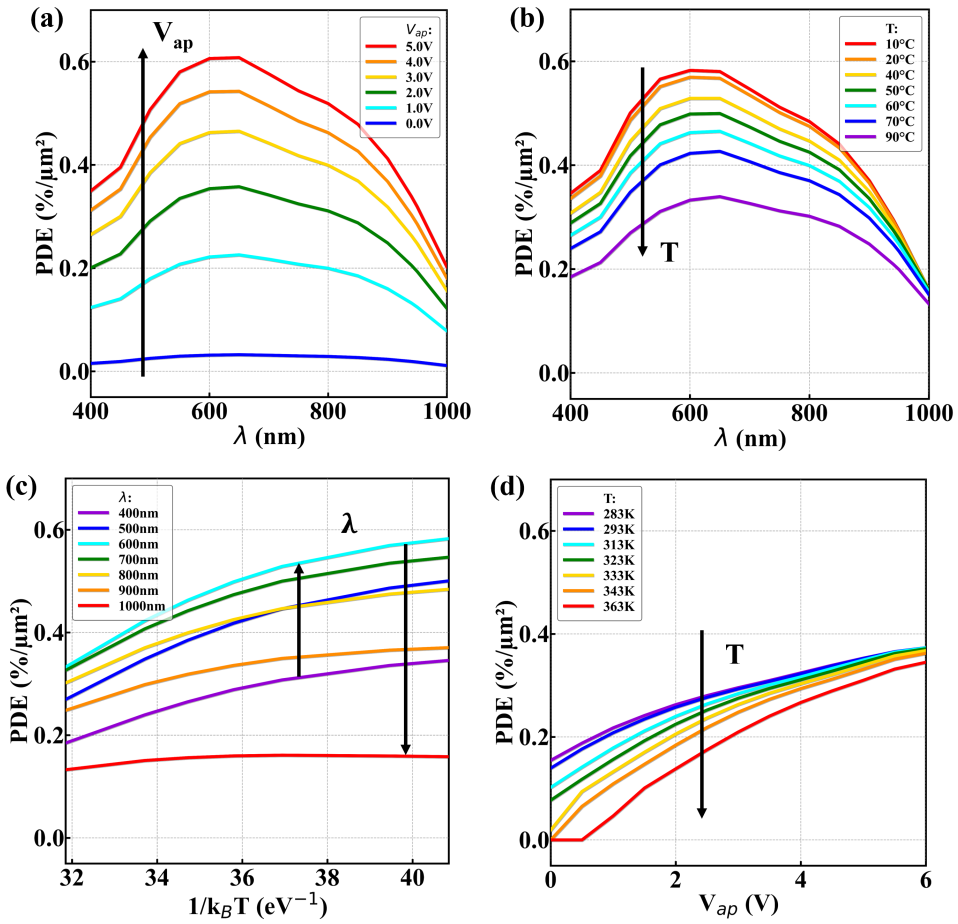


Figure 1.21: PDE as a function of (a) the wavelength λ measured at 60°C for different applied voltage V_{ap}, (b) the wavelength λ measured at V_{ap}=3V for different temperatures, (c) $1/k_BT$ measured at V_{ap}=3V for different λ, and (d) V_{ap} measured at λ=940nm for different temperatures. PDE increases with V_{ap} and $1/k_BT$. PDE increases up to a certain λ and then decreases which relates to photon absorption depth. An increase in temperature causes a decrease in the carrier-phonon mean-free path, resulting in a reduction in avalanche breakdown probability P_t and, consequently, a decrease in PDE. On the other hand, increasing the voltage enhances the electric field, leading to an increase in impact-ionization events, which increases P_t and, in turn, the PDE.

The PDE against wavelength dynamics shown in **Fig. 1.21 (a)** relates to the absorption coefficient α. It determines how far into a material light of a particular wavelength can penetrate before it is absorbed. The absorption coefficient depends on the material and also on the

wavelength of light which is being absorbed. If the lowest energy gap is indirect, such as in silicon, a photon can excite an electron from the valence band to the conduction band with the assistance of a phonon. In in the vicinity of an indirect band gap semiconductor α is proportional to:

$$\alpha \propto \frac{(hv - E_g + E_p)^2}{exp(\frac{E_p}{k_B T}) - 1} + \frac{(hv - E_g - E_p)^2}{1 - exp(\frac{-E_p}{k_B T})} \tag{1.3}$$

The first term relates to phonon absorption and the second one to phonon emission. The positive thermal dependence of the absorption coefficient is reported in **Fig. 1.22**. Light with energy below the band gap does not have sufficient energy to excite an electron into the conduction band from the valence band. Even for photons with an energy above the band gap, the absorption coefficient is not constant, but still depends strongly on wavelength. To sum up, at lower wavelengths, the material absorbs more photons, leading to an increase in PDE. However, at higher wavelengths, the material becomes less efficient at absorbing photons, resulting in a decrease in PDE. At a given applied temperature, PDE increases up to 650nm and then decreases as depicted in **Fig. 1.21 (b)-(c)**. At a given applied voltage, PDE decreases as temperature increases as shown in **Fig. 1.21 (c)**. As the temperature increases, the carrier-phonon mean-free path decreases, resulting in a negative shift of P_t as the temperature increases. This shift fully compensates for the positive shift of OA caused by the decrease in band gap energy as temperature increases. Increasing the applied voltage increases PDE as illustrated in **Fig. 1.21 (d)**. This leads to an increase in the electric field strength, which enhances the impact-ionization mechanisms and, consequently, P_t.

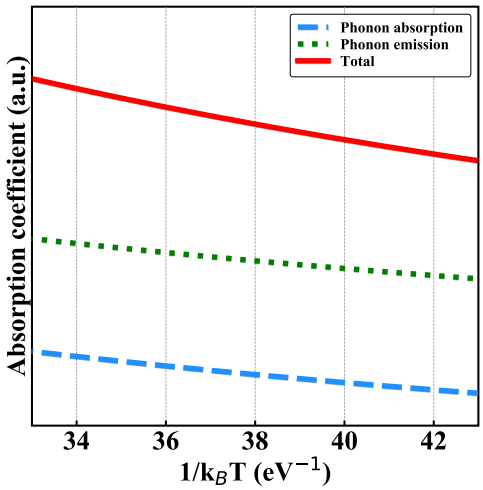


Figure 1.22: Optical Absorption (OA) as a function of $1/k_BT$ for both phonon absorption and emission, as well as the sum of both. With an increase in temperature, the band gap energy decreases, which leads to an increase in OA.

Another characteristic that is related to photon flux in a similar way is the maximum count rate, which plays a role in determining the maximum distance range. One of drawbacks of SPADs with respect to analog detectors is the presence of a deadtime after each photon detection which limits the maximum count rate. The maximum count rate is equal to the inverse of the dead

time. This refers to the dynamic range which is the range of photon flux levels that can be detected without saturation or distortion of the output signal.

5.4 Electrical and optical crosstalk

Crosstalk is an avalanche current that spreads carriers into the interconnections or neighbouring devices. Electrical crosstalk occur in-between parallel interconnections by means of inductive or capacitive coupling as portrayed in **Fig. 1.23**. This propagating electromagnetic wave transfers the signal from one line to another leading to erroneous or false signal. On the other hand, numerous SPADs sharing the same power supply line triggering simultaneously drop the local operational potential inducing a snowball effect over all characteristics. The coupling electrical equations of parallel interconnections are as follow:

$$\frac{d^2v}{dx^2} = -j\omega L\frac{di}{dx} = \omega^2 LCv \tag{1.4}$$

$$\frac{d^2i}{dx^2} = -j\omega L\frac{dv}{dx} = \omega^2 LCi \tag{1.5}$$

The electrical crosstalk relies on the structural parameters such the width, the length and the spacing of the interconnections. As demonstrated by the equations above, it depends on the squared of the frequency which requires MHz range to have significant consequences.

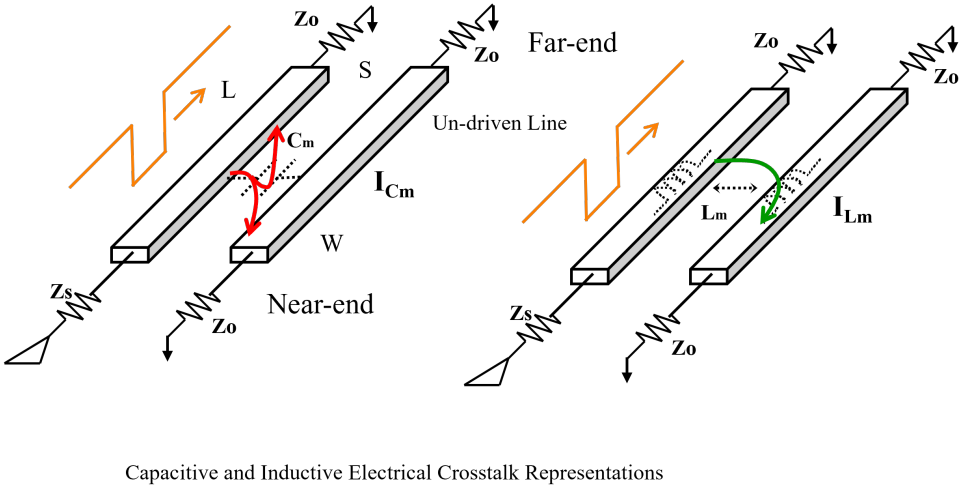


Figure 1.23: Illustration of electrical crosstalk due to capacitive or inductive coupling with the transmission lines, which has the potential to trigger an avalanche effect in nearby devices.

Carrier acceleration and recombination generate photons outside the device leading to other avalanches in other neighbouring devices [43] as outlined in **Fig. 1.24**. Optical crosstalk originates from these photons emitted during an avalanche that can travel hundreds of microns according to their wavelength before generating electron-hole pairs.

The enable transistor EN described in **Fig. 1.12** allows to switch-on/off some nearby SPADs in a pixel array. The optical crosstalk measurement process involves examining the time between two adjacent SPAD pulses, and if this time is less than 5 ns, it is considered a potential crosstalk

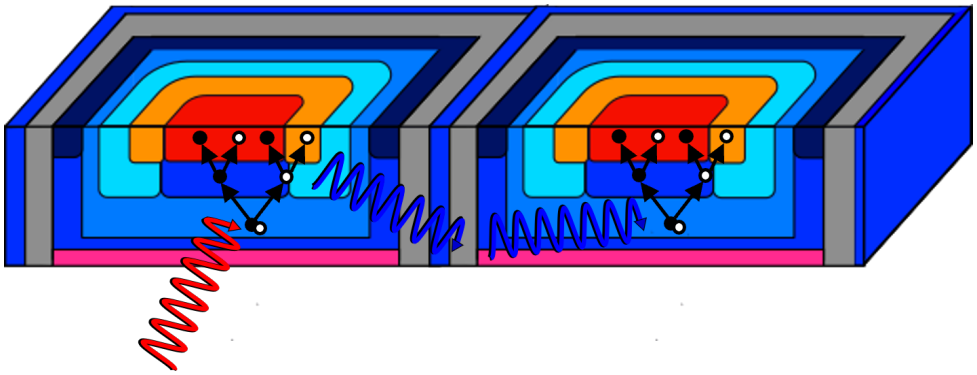


Figure 1.24: Illustration of optical crosstalk due to photons generated during avalanche triggering trigger avalanches in neighboring devices.

event. This time interval corresponds to a photon emission in the SPAD, followed by an avalanche and reabsorption in the neighboring SPAD, and finally triggering an avalanche in the neighboring SPAD. The same time is measured between two distant SPADs, and if there are correlated events, it is interpreted as photons arriving at approximately the same time on the SPADs. Subtracting the second measurement from the first measurement produces the distribution presented in **Fig. 1.23**. The methodology used to analyze adjacent SPADs can be extended to non-adjacent SPADs.

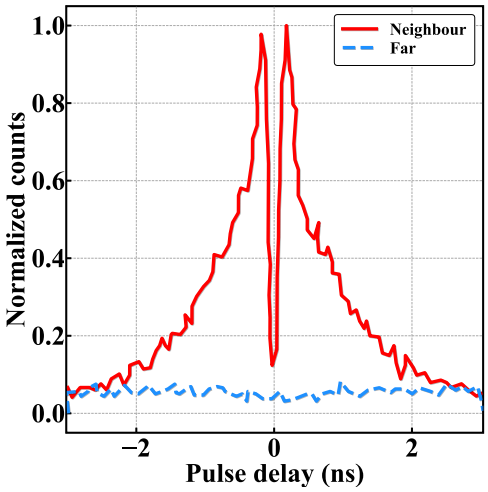


Figure 1.25: Optical crosstalk mesured on two adjacent and non-adjacent SPADs and averaged over tens of thousands of times.

5.5 Jitter, afterpulsing and lag

The time interval between the photon absorption and the beginning of the avalanche presents some statistical fluctuations from localization of photon absorption and avalanche triggering. The jitter is defined as the Full-Width at Half Maximum (FWHM) from the photon absorption time distribution. The stochastic multiplication process gives the peak width. If the laser intensity is high, the SPAD triggers all the time and does not rest, the jitter distribution corresponds

to a Dirac. The evacuation time of carriers photogenerated in the neutral regions and diffusing to the depletion region gives the shape to the jitter tail. These results are discussed in details in **Chapter 2**.

Jitter is measured by stimulating the SPAD by a picosecond laser pulse each microsecond for instance. The time between laser pulse emission and all the successive inverter pulse are measured. Only one SPAD is activated to avoid optical crosstalk described previously. The pulse shaper transistor can be activated to observe the impact of the dead time over the afterpulsing. The presence of afterpulsing can be seen by the appearance of a second distribution after the first photon-correlated distribution and after the dead time. A percentage of afterpulsing can be extracted regarding the number of pulses measured. The height of the the two distribution are different as shown in **Fig. 1.26**.

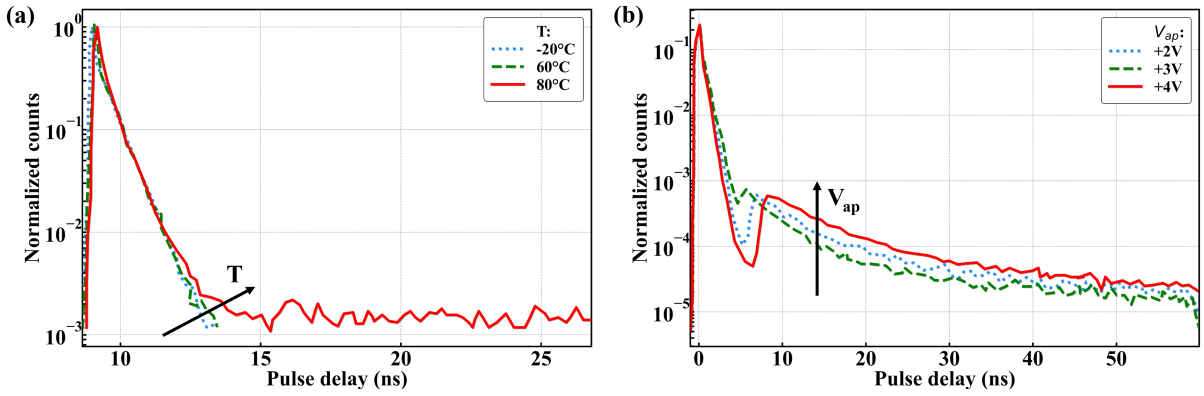


Figure 1.26: Probability density function of the jitter measured for (a) different temperatures and (b) various applied voltage V_{ap}. The process of trapping carriers can be enhanced by increasing the excess bias, while thermal energy plays a crucial role in releasing trapped carriers.

Secondary pulses trace back their origins in two potential mechanisms. The phenomena are called either lag or afterpulsing but both relate to carrier releasing after some nanoseconds when the SPAD is recovering igniting an after pulse into the electrical signal. These effects limit the SPAD dead time. An electrostatic pocket captures some carriers in neutral region or a deep-level trap captures and holds carrier during the avalanche process. **Fig. 1.26 (a)** illustrates that the thermal energy aids in the release of trapped carriers, while **Fig. 1.26 (b)** shows that increasing the excess bias results in more carriers being trapped.

Other techniques exist to measure the afterpulsing. The PDE is measured in two conditions by switching on and off the pulse shaper transistor. The afterpulsing can be calculated as the difference between the PDE in switch off and the PDE in switch on, divided by the PDE in switch on again without any optical crosstalk. The supplementary avalanches measured in switch off mode relate to afterpulsing and assuming that the measure realized in switch on mode is free from afterpulsing thanks to the dead time increasing.

Another method is based on the inter-pulse time measurement. The SPAD is irradiated by not to intense led light to reach high count rate in the order of 10 to 100 kcps to not mask afterpusling pulses by overlapping photon-correlated pulses. Again a single SPAD is switch on and the time between two pusles is measured. A Poisson distribution centered in one over

the count rate exempted of afterpulsing is found thanks to the poissonien nature of led light. A second distribution arises due to afterpulsing secondary pulses. A poissonien fit is used as shown in **Fig. 1.27 (a)** and removed from the overall distribution to left only the afterpulsing distribution. However this technique requires higher measurement time but can be accelerated by increasing the light quantity in right balance to avoid hiding the afterpulsing in the Poisson distribution resulting from light. The same trend can be appreciated in the right-hand side of **Fig. 1.27 (b)** where the distribution of pulse widths is represented.

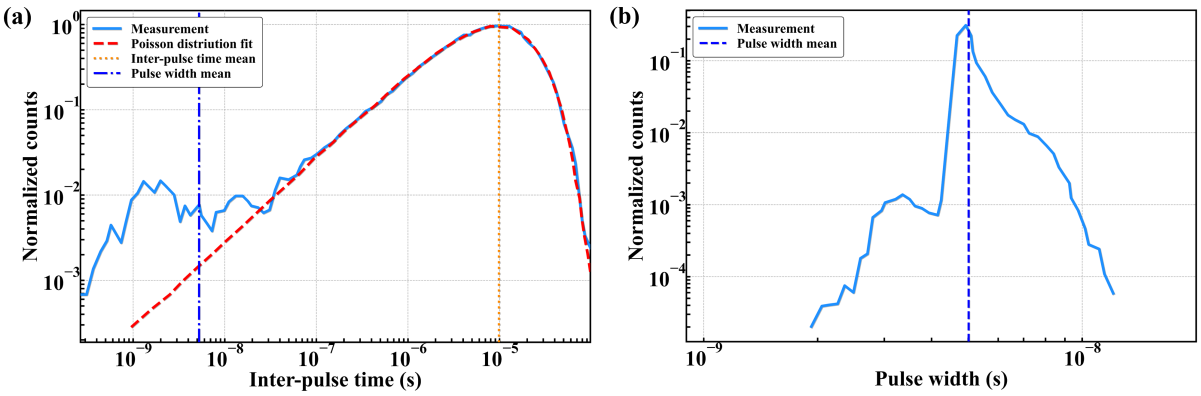


Figure 1.27: (a) Normalized counts of inter-pulse times measured and fitted with a Poisson distribution. The stochastic nature of light is responsible for what is fitted, while the remaining portion is due to afterpulsing. (b) Normalized counts of pulse widths which exhibits a similar trend.

5.6 Performance summary

The main advantages and drawbacks of SPADs are outlines in **Table 1.1**. This table summarizes the impact of various pixel parameters on the system, as well as the trade-offs involved. Pixel pitch affects the spatial resolution and silicon area through the space allocated for digital counting, while V_{BD} and V_{AP} impact power consumption and PDE and jitter. The distribution of DCR affects noise intensity and PDE, and the PDE impacts sensitivity, DCR, jitter, and afterpulsing. Jitter affects temporal noise and PDE, while dead time impacts dynamic range, and quench and afterpulsing. Optical crosstalk affects resolution and PDE, and afterpulsing affects temporal noise and dead time.

Pixel parameter	System impact	Trade-off
Pixel pitch	Resolution	Silicon area (Digital counting)
V_{BD} and V_{AP}	Power consumption	PDE and jitter
DCR distribution	Noise intensity	PDE
PDE at 940nm	Sensitivity	DCR, jitter and afterpulsing
Jitter (FWHM)	Temporal noise	PDE
Dead time	Dynamic range	Quench and afterpulsing
Optical crosstalk	Resolution	PDE
Afterpulsing	Temporal noise	Dead time

Table 1.1: SPAD performances against system criteria.

Conclusions

The features of SPADs were explained, including the avalanche breakdown voltage V_{BD}, which is related to the first avalanche triggered, the Dark Count Rate (DCR), which is caused by parasitic generated carriers in the absence of light, the Photo-Detection Efficiency (PDE), which indicates the number of photons absorbed considering the actual photo-sensitive area, the electrical and optical crosstalk, which are related to the coupling of transmission lines that generates false signals and the photon generated during the avalanche process that triggers avalanches in nearby devices, and jitter, afterpulsing, and lag, which are related to the photon absorption time distribution and secondary pulses generated right after them. These secondary pulses are caused by excess bias resulting in more carriers being trapped or thermal energy, which aids in the release of trapped carriers. The discussion has covered the trade-offs between various metrics and system features, including pixel pitch, V_{BD}, V_{AP}, DCR, PDE, jitter, dead time, optical crosstalk, and afterpulsing, which affect the performance of SPADs in areas such as resolution, power consumption, noise intensity, sensitivity, dynamic range, and temporal noise.

6 3D imaging technologies

6.1 Principles and applications

The objective is to offer readers a thorough comprehension of the market dynamics that revolve around the 3D-imaging market, including technical market segmentation and application range. The main use technologies are presented in **Fig. 1.28** where active solutions involve an additional light source for sensing, while passive solutions utilize ambient light. There are various 3D sensing techniques available that differ in their accuracy, distance range, power consumption, and cost, providing a wide range of options to choose from. In addition to capturing color and luminance, the depth is also recorded.

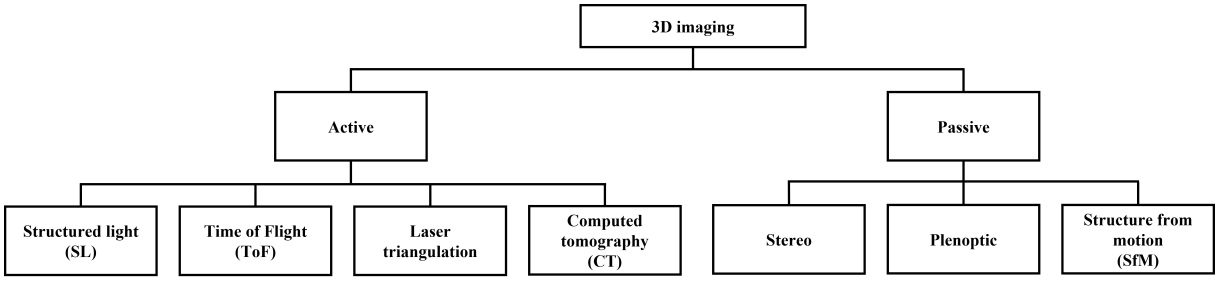


Figure 1.28: Optical 3D sensing technologies with active and passive solutions. Active solutions refer to the use of an additional light source for sensing, while passive solutions rely on ambient light. The choice of technology depends on the desired level of accuracy, distance range, power consumption, and cost.

The passive technologies are:

- Structured Light (SL): projected light pattern and camera to capture reflections and calculate object shape based on pattern deformation.
- indirect Time of Flight (iToF) (shown in **Fig. 1.29**): a laser source transmits an intensity modulated optical wave, which is reflected by a target surface and the difference between the transmitted signal phase and the received signal phase encodes distance to the target.
- direct Time of Flight (dToF) (shown in **Fig. 1.29**): a laser source transmits an optical pulse, which is reflected by a target source and the difference between the transmit time and receive time encodes the distance.
- Laser Triangulation (LT): a laser line projector strongly illuminates a strip of the object which are under motion to be imaged, and the reflected light is imaged on a 2D sensor to detect the profile shape of the illuminated section and by scanning the entire object with the system, accurate object dimensions are delivered.
- Computed Tomography (CT): a series of X-ray images from different angles around the object structure are processed by computation to create cross-sectional images which are be combined to create 3D images .

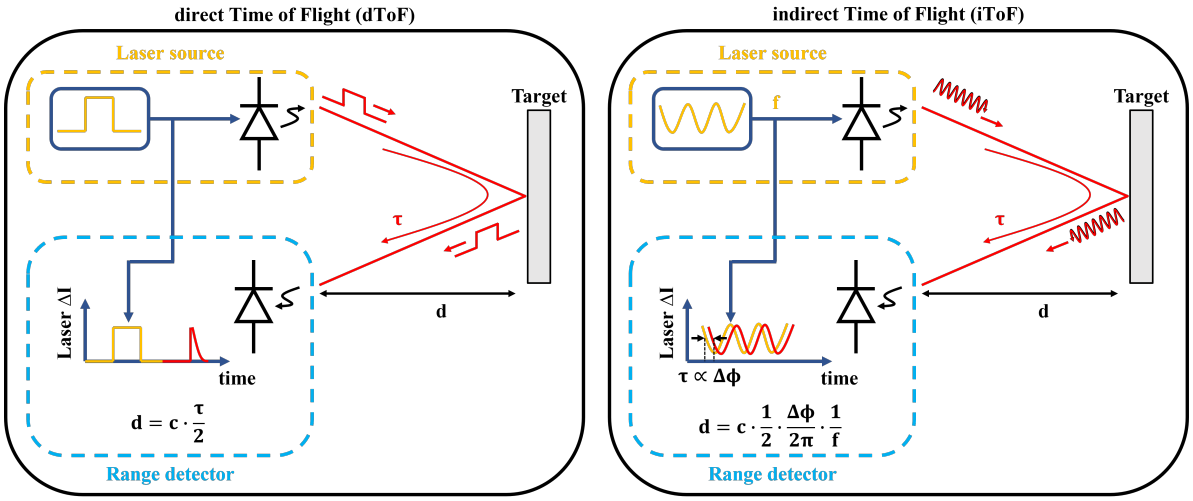


Figure 1.29: Illustration of the Time-of-Flight (ToF) techniques used to determine the distance from the sensor to the target. This can be achieved through direct measurement of photon travel time or indirect measurement of the phase difference between the transmitted and received signal.

The active technologies are:

- Stereo: two cameras positioned at different locations, with their signals processed by a computer to determine the depth and 3D structure of an object or scene.
- Plenoptic: a microlens array is placed between the primary lens and the image sensor which allows the sensor is not only sensitive to light, color, and intensity but also to angle, which enables depth sensing by triangulation
- Structure from Motion (SfM): unordered multiple overlapping frames from moving cameras to reconstruct the 3D scene.

> The data presented in this section are from "3D Imaging and Sensing 2023: Market and Technology Report," YINTR23329, www.yolegroup.com, © Yole Intelligence 2023 [44].

The table **Table 1.2** shows different markets and applications of various 3D sensing technologies, including SL, ToF, stereo, CT, and LT. The markets and applications include mobile devices, automotive, medical, industrial, and defense and aerospace. The table indicates which 3D sensing technologies are suitable for each market and application, with some technologies being more appropriate than others.

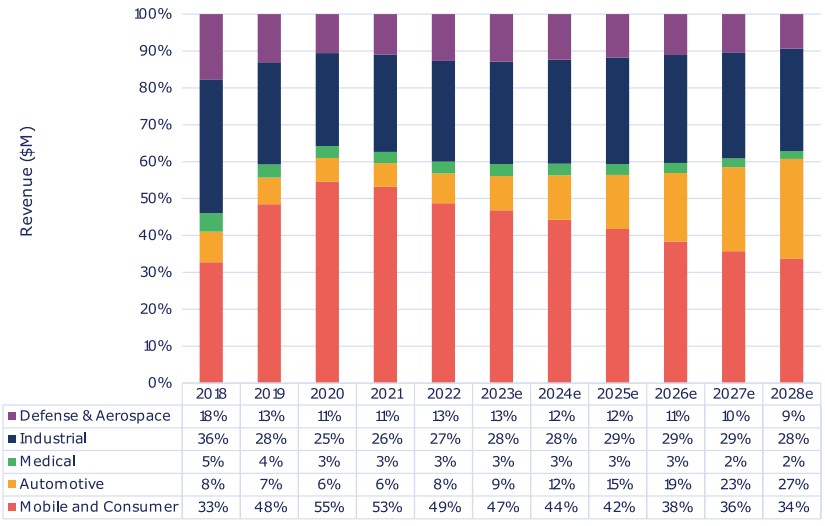
6.2 Market shares and suppliers

Fig. 1.30 shows that the mobile and consumer 3D sensing market grew rapidly from 2018 to 2020 but has since decreased due to disengagement of smartphone Android models from 3D sensing, led by Huawei's loss of market share. This market is expected to steadily decrease, while the automotive 3D sensing market will benefit. The combined mobile and consumer and

Market	Application	SL	ToF	Stereo	CT	LT
Mobile and Consumer	Mobile front	■	■			
	Mobile rear		■			
	Laptop/Tablet Front	■				
	Laptop/Tablet Rear		■			
	AR/VR/MR headset		■	■		
	Wearable	■	■			
	Consumer robot	■	■	■		
	Smart door lock	■	■			
	Consumer drone		■	■		
Automotive	ADAS imaging			■		
	ADAS LiDAR		■			
	In-cabin monitoring	■	■			
	Exterior access	■	■			
Medical	CT scan				■	
	Cone Beam CT				■	
	Optical Coherence Tomography				■	
Industrial	Payment/Access control	■	■			
	Construction and Infrastructure		■			
	Manufacturing and Logistics		■	■	■	■
	Security				■	
Defense and Aerospace	Geographical survey		■			
	Military payload		■			

Table 1.2: Table outlining 3D sensing markets and applications by technology, with significant use and considered market forecast in green and possible emerging use case in orange.

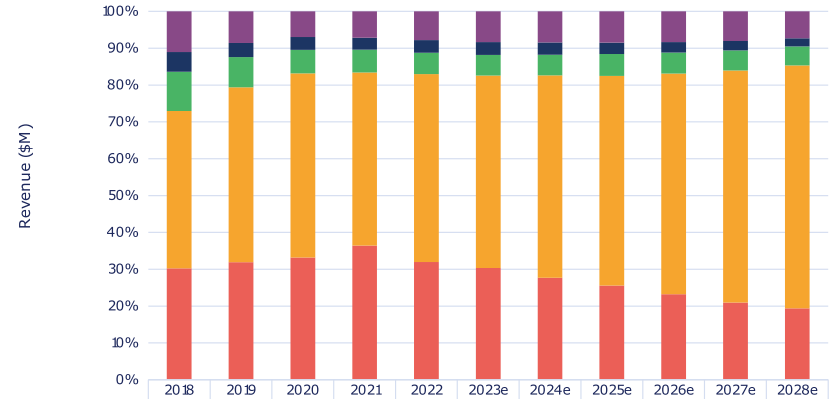
automotive markets will remain relatively stable between 56 and 61%. The medical market will account for 2 to 3% of global 3D sensing revenue, while the industrial market will experience solid growth from 26 to 29% of the revenue. Defense and aerospace applications will remain in the 9 to 13% range.



	2018	2019	2020	2021	2022	2023e	2024e	2025e	2026e	2027e	2028e
Defense & Aerospace	18%	13%	11%	11%	13%	13%	12%	12%	11%	10%	9%
Industrial	36%	28%	25%	26%	27%	28%	28%	29%	29%	29%	28%
Medical	5%	4%	3%	3%	3%	3%	3%	3%	3%	2%	2%
Automotive	8%	7%	6%	6%	8%	9%	12%	15%	19%	23%	27%
Mobile and Consumer	33%	48%	55%	53%	49%	47%	44%	42%	38%	36%	34%

Figure 1.30: 2018-2028 3D Sensing revenue splits by end market.

Fig. 1.31 shows that the revenue from SL technology is decreasing from 36% to 19%, while the revenue from ToF technology is increasing from 47% to 66%, especially in higher-end markets due to the higher-priced modules. The revenue from stereo, CT, and LT will remain stable in a range from 14% to 17%, with established and growing markets such as manufacturing monitoring, medical healthcare devices using CT, and drone products that use stereo.



	2018	2019	2020	2021	2022	2023e	2024e	2025e	2026e	2027e	2028e
Laser triangulation	11%	9%	7%	7%	8%	8%	8%	8%	8%	8%	7%
CT	5%	4%	3%	3%	3%	3%	3%	3%	3%	3%	2%
Stereo	11%	8%	6%	6%	6%	6%	6%	6%	6%	5%	5%
ToF	43%	47%	50%	47%	51%	52%	55%	57%	60%	63%	66%
SL	30%	32%	33%	36%	32%	30%	28%	26%	23%	21%	19%

Figure 1.31: 2018-2028 3D Sensing revenue market shares split by technology.

The trends in the mobile market for Original Equipment Manufacturer (OEM) at sensor-level are depicted in **Fig. 1.32**. STMicroelectronics has been Apple's global supplier of NIR global shutter sensors since 2017, shipping over 923 million units. 3D sensing has become an entrenched technological solution for Apple products with the Face ID feature. Android OEMs will eventually adopt 3D sensing into their models for the high-end mobile segment and to gain traditional Apple customers. They are increasing their leadership in the 3D sensing market with their development of iToF and dToF multizone products, collaboration with Metalenz, and progress on Quantum Dot technology. Sony has emerged as a challenger, providing iToF to Android suppliers and dToF array sensors to Apple, and has established an application center to facilitate the use of iToF for downstream manufacturers. pmd and ams-OSRAM are the remaining challengers, with pmd being a historical leader in iToF products and ams-OSRAM recently introducing its dToF multizone modules. Some image sensor players have disappeared from Mobile 3D sensing, such as OmniVision and Samsung.

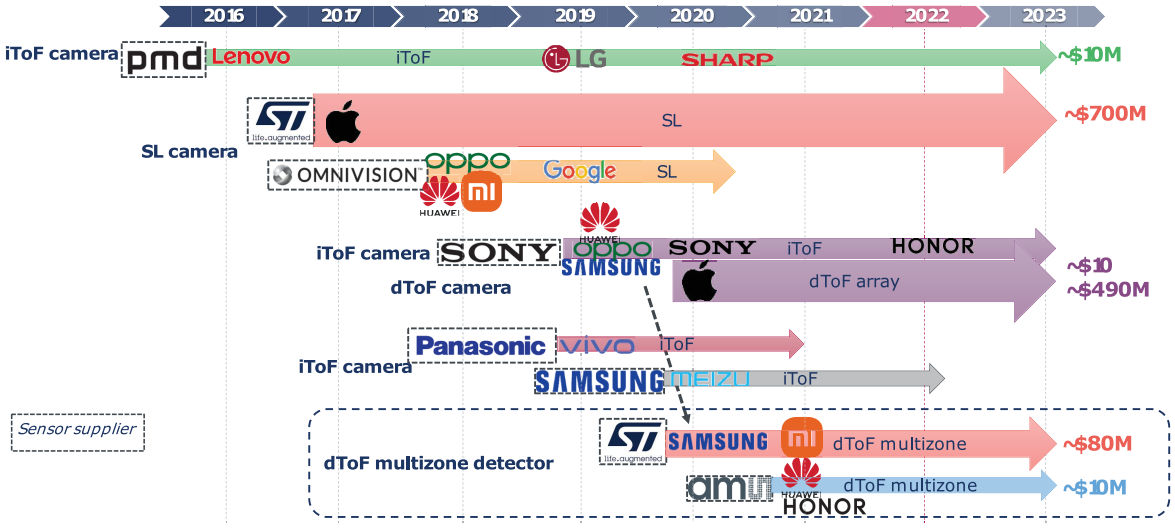


Figure 1.32: Mobile market trends for Original Equipment Manufacturer (OEM) at sensor level.

Conclusions

3D sensing including SL, ToF, stereo, CT, and LT, is expanding to other markets such as consumer, automotive, medical, industrial, defense and aerospace. Time-of-Flight (ToF) technologies are expected to represent 66% of total 3D sensing revenues by 2028, with multizone dToF devices posing a threat to 3D cameras due to their lower price and compactness. STMicroelectronics has a strong foothold in the market, but Sony and other players are emerging as challengers, offering innovative solutions such as iToF and dToF multizone products for both Apple and Android customers.

7 Conclusions

Maxwell's theory suggests that the oscillation of charges can lead to the emission of electrons from a metal surface, with the intensity of the oscillation affecting the velocity of the ejected electrons and the duration of the oscillation affecting the time taken for electrons to be ejected. Hertz used a spark-gap transmitter to demonstrate this effect, with the cathode of the spark plug gaps being the actual formation-place of the photoelectric effect. Hallwachs used a gold leaf electroscope to show that the effect is created at the plate's surface by means of ultraviolet light and not by red light or thermal radiation. Thomson demonstrated that electrons were ejected from the cathode by radiation rather than by the strong electric field used previously. Lenard and Millikan found that the frequency of light did not affect the current, but the current increased with intensity. Additionally, light below a certain frequency could not eject electrons regardless of the intensity. Lenard discovered that the charge-to-mass ratio of the ejected particles was the same as that for electrons. Einstein later formulated the photoelectric effect, providing an explanation for Lenard's experiment which showed that the number of emitted electrons increased with the intensity of the light, but the kinetic energy of the electrons depended on the frequency of the light rather than its intensity.

1

Photoemissive devices are based on an incident photon that emits an electron, while photoconductive devices rely on incident photons that change the conductivity. The degree of multiplication in a Photo-Multiplier Tube (PMT) is determined by the interdynodal potentials that accelerate the electron beam, with the main limitation being the strong magnetic field. In contrast, p-n junctions in mass standard production consume less space and are less sensitive to magnetic fields, with operating modes including the diffusion of majority carriers in forward bias, the attraction of minority carriers by opposite electrodes in reverse bias, and the occurrence of minority carrier avalanche in breakdown bias. PIN photodiodes, which are reverse biased with a depleted wide intrinsic area to absorb photons, result in a weak output signal that requires external amplification, which can introduce noise. On the other hand, an Avalanche Photo-Diode (APD) is a reverse biased p-n junction without avalanche triggering, with the gain limited by the carrier multiplication factor. A low-noise preamplifier with a high gain-bandwidth product is required, with the fluctuation of the multiplied photo-current needing to be higher than the threshold of the preamplifier. A Single-Photon Avalanche Diode (SPAD) is a reverse biased p-n junction with avalanche triggering, and their development is related to Complementary Metal-Oxide-Semiconductor (CMOS), which allows for a reduction in pixel pitch increasing spatial resolution. Current research in this area is focused on 3D integration to reduce the fill factor, CMOS Image Sensor (CIS) to take advantage of the latest developments in CMOS imagers, and Silicon-On-Insulator (SOI) to reduce parasitic capacitance and crosstalk between components.

2

The SPAD architecture under study consists of highly doped n++ and p+ multiplication regions that confine avalanches at low biasing voltage. Lightly doped n and p guard rings prevent lateral breakdown at high biasing voltage, while a p-epi layer acts as an absorbing layer that generates carriers when it absorbs photons. To prevent physical breakdown in a SPAD due to large current, active and passive quenching circuit techniques have been developed. Passive quenching involves using a resistor or transistor in series with the SPAD to discharge the junction and stray capacitances. Active quenching, on the other hand, uses a threshold comparator to provide more precise control over the quenching time and fast recharge after the avalanche. However, active quenching can reduce the pixel fill factor depending on circuit integration. The passive quenching circuit used in this study comprises a low-voltage n-type enable MOSFET, a high-voltage cascode n-type MOSFET, and a clamp diode, all connected to the anode of the SPAD. This circuit disables pixels and guards components against transient voltages and high current density. A high-pass filter shapes the signal and ensures that the cathode swing is within the tolerance of transistors composing the output inverter. The pulse width depends on V_{HPF} settings, enabling the observation of afterpulsing or lag phenomena.

3

4

The features of SPADs are explained, including the avalanche breakdown voltage V_{BD}, which is related to the first avalanche triggered, the Dark Count Rate (DCR), which is caused by parasitic generated carriers in the absence of light, the Photo-Detection Efficiency (PDE), which indicates the number of photons absorbed considering the actual photo-sensitive area, the electrical and optical crosstalk, which are related to the coupling of transmission lines that generates false signals and the photon generated during the avalanche process that triggers avalanches in nearby devices, and jitter, afterpulsing, and lag, which are related to the photon absorption time distribution and secondary pulses generated right after them. These secondary pulses are caused by excess bias resulting in more carriers being trapped or thermal energy, which aids in the release of trapped carriers. The discussion has covered the trade-offs between various metrics and system features, including pixel pitch, V_{BD}, V_{AP}, DCR, PDE, jitter, dead time, optical crosstalk, and afterpulsing, which affect the performance of SPADs in areas such as resolution, power consumption, noise intensity, sensitivity, dynamic range, and temporal noise.

5

3D sensing including SL, ToF, stereo, CT, and LT, is expanding to other markets such as consumer, automotive, medical, industrial, defense and aerospace. Time-of-Flight (ToF) technologies are expected to represent 66% of total 3D sensing revenues by 2028, with multizone dToF devices posing a threat to 3D cameras due to their lower price and compactness. STMicroelectronics has a strong foothold in the market, but Sony and other players are emerging as challengers, offering innovative solutions such as iToF and dToF multizone products for both Apple and Android customers.

8 Thesis objectives and manuscript overview

1

This chapter aims to provide a comprehensive understanding of the probability of avalanche breakdown P_t and the Recombination-Generation (R-G) rate, which are responsible for causing Dark Count Rate (DCR) in SPADs. The chapter begins by introducing computation techniques that can be used to predict P_t in SPADs, including a summary of the simulation workflow employed in Technology Computer-Aided Design (TCAD) and an explanation of the thermal and electric field dependencies of ionization threshold energy and impact-ionization coefficients. The chapter then compares the McIntyre model to an alternative drift-diffusion Monte-Carlo method to determine the avalanche breakdown. Additionally, the chapter investigates the R-G processes from thermal to electric-field enhancement effects, examining the Shockley-Read-Hall (SRH) R-G rate and the Poole-Frenkel effect (PF) effect, among others. The chapter also discusses the thermal and voltage dependence of the R-G mechanisms to draw experimental conclusions regarding their growing rate. Overall, this chapter provides a detailed analysis of the factors contributing to DCR in SPADs, offering insights into the design and optimization of these devices.

2

The third chapter discusses the computation of DCR contribution at each mesh point within the device which is facilitated by the previously introduced avalanche breakdown probability and generation rate of carriers. This generation of carriers may arise from thermal Shockley-Read-Hall (SRH) generation from defects, which may be further enhanced by Trap-Assisted Tunneling (TAT), and Band-to-Band Tunneling (BTBT) according to the electric field strength. This chapter combines insights from measurement and simulation over temperature and voltage to elucidate the origins of average DCR within the low- and high-bias voltage regimes. Additionally, the use of Design of Experiment (DoE) enables the identification of the most significant factors and their interactions that impact DCR. Furthermore, the chapter investigates the statistical dispersion that results in intrinsic and extrinsic DCR populations using a Monte-Carlo approach that randomizes the position and number of defects. Overall, this chapter provides a detailed analysis of the factors contributing to DCR in SPADs, offering insights into the design and optimization of these devices.

3

The fourth chapter introduces insights into the underlying degradation mechanisms of SPADs. The chapter begins by providing an overview of metastable defects, which are known to play a crucial role in temperature, light, and voltage degradation mechanisms. A comprehensive set of experimental results is then presented to identify the specific degradation mechanism in the studied SPAD. Of particular interest is the investigation of the avalanche's contribution to the degradation process, achieved by varying the metal shield extension on the bottom interface to reduce the light aperture. By reducing the number of photons transmitted to the silicon, this technique enables a more precise investigation of the impact of the avalanche on the device's performance. The investigation of degradation mechanisms begins by focusing on the average behavior of the devices under stress and later extends the analysis to encompass device-to-device behavior. Additionally, various samples from different processes are utilized to validate the localization of newly formed defects during stress conditions, adding an extra layer of consistency and reliability to the study's findings. Overall, this chapter provides a comprehensive understanding of the degradation mechanisms in the SPAD under studied, offering insights into the design and optimization of these devices for improved reliability.

4

The fifth chapter discusses the state-of-the-art of Hot-Carrier Degradation (HCD) in MOSFETs and its contributions to the degradation process in SPADs. The chapter aims to explain the time dependence of ΔDCR under device stress to predict device reliability based on the existing literature of HCD in MOSFETs. The use of a Full-Band Monte-Carlo (FBMC) to compute the carrier energy distribution is employed to determine if hot carriers have the threshold energy to dissociate Si-H bonds. The relationship between the degradation of ΔDCR in SPADs and the electric current without avalanches is studied to confirm the role of hot-carriers quantity in the degradation. Furthermore, the currents generated in darkness and under irradiance, along with the impact-ionization multiplication of carriers, are modeled to be incorporated into the degradation rate. The defect creation kinetics is then modeled using the carrier energy distribution against the $Si - H$ bond dissociation distribution to determine the probability of impact-ionization leading to bond dissociation. Overall, this chapter provides a comprehensive understanding of the time dependence of ΔDCR under device stress, offering insights into the design and optimization of these devices for improved reliability.

5

The sixth and last chapter aims to establish a comprehensive understanding of the relationship between parasitic noises, including DCR and ambient light, and the photon signal reflected back from the target. The investigation will consider various factors such as target reflectance and the distance between the sensor and the target. Additionally, the section will explore the attenuation of light intensity and the statistical fluctuations of light. To achieve this, the mechanisms that can disable signal detection at different distances from the target, such as optical crosstalk or pile-up, will be thoroughly investigated. Schematics and simulations will be utilized to describe these mechanisms, while disregarding the external circuitry or photonic algorithm used to process the signal.

Chapter 2

Introduction to SPAD simulation fundamentals

Objectives

The main objective of this chapter is to provide a detailed description of the probability of avalanche breakdown P_t and the Recombination-Generation (R-G) rate, which are responsible for causing Dark Count Rate (DCR) in SPADs. The first step is to introduce some computation techniques that can be used to forecast the probability of avalanche breakdown P_t in SPADs. To begin with, a summary of the simulation workflow employed in Technology Computer-Aided Design (TCAD) is provided. Then the thermal and electric field dependencies of ionization threshold energy and impact-ionization coefficients are explained. Starting from this formalism, P_t is derived using the McIntyre model along electric field streamlines. The simulation results are confronted against experimental data using the definition of Photo-Detection Efficiency (PDE) and breakdown voltage V_{BD}. The aforementioned methodology is then compared to an alternative drift-diffusion Monte-Carlo method. This approach takes into account the stochastic nature of carrier random walk and carrier multiplication as factors in determining the avalanche breakdown. Additionally, the chapter investigates the R-G processes from thermal to electric-field enhancement effects. The Shockley-Read-Hall (SRH) R-G rate is examined, starting from the minority carrier lifetime limited by dopants and progressing to that limited by traps. The Poole-Frenkel effect (PF) effect which induces lowering of the trap potential barrier and the Trap-Assisted Tunneling (TAT) are included in the definition of the minority carrier lifetime to account for the trap description in presence of electric field. Indirect Band-to-Band Tunneling (BTBT) in silicon is also discussed. Finally, the thermal and voltage dependence of the aforementioned R-G meachnisms are studied to draw experimental conclusions regarding their growing rate.

Contents

1 TCAD simulation workflow

Let's first introduce the Technology Computer-Aided Design (TCAD) simulation workflow before delving into a detailed explanation of the simulation methodology for avalanche breakdown probability P_t and Recombination-Generation (R-G) mechanisms. The device structure is built by simulating fabrication processes through the Sentaurus process tool [45] considering the device layout imported from Calibre DESIGNrev [46]. Design of Experiment (DoE) enables to efficiently explore the manufacturing parameters and identify the optimal device design and process conditions as shown in **Fig. 2.1**. The Sentaurus device module [47] is based on analytic models to simulate the physical behavior of semiconductor devices. The Poisson equation and current continuity equations, coupled with the drift-diffusion transport model, are solved using this software to compute the electrical outputs. These equations will be presented in **Section 3.1**. The simulation methodology is divided into three steps. Firstly, the system is solved in steady states considering avalanche generation up to a contact break criteria as represented in **Fig. 2.2**. In steady-state conditions, for each property of the systems, its partial derivative with respect to time is zero, i.e. nothing is changing with time. Afterwards, the avalanche breakdown voltage is extracted using Sentaurus visual module [48]. Starting from the electrostatic profile at this voltage value, the system is solved in transient states increasing the voltage step at each time step while switching-off the avalanche conditions. A transient response is the time-varying response of a system to a change from equilibrium. This allows to study the physics ruling the device beyond the avalanche breakdown voltage without avalanche triggering. As this work involves the use of Sentaurus device tool [47], the available options and parameters for defining a model will be specified.

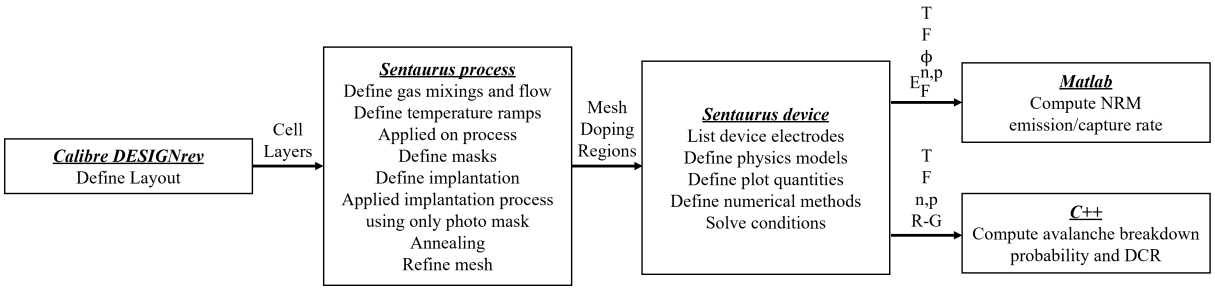


Figure 2.1: TCAD simulation procedure including layout definition from Calibre DESIGNrev [46], process manufacturing from Sentaurus process [45], device physics solving from Sentaurus device [47] and computation using in-house solvers in matlab and C++ languages. T is the lattice temperature, F is the electric field, Ψ is the electrostatic potential, $E_F^{n,p}$ is the quasi-Fermi potential for electrons and holes, n and p are the electron and hole concentrations and R-G is the recombination-generation rate. These inputs are used to compute either the avalanche breakdown probability (**Section 2**), the Non-Radiative Multiphonon-assisted (NRM) generation rate of carriers (**Appendix D**) and the Dark Count Rate (DCR) (**Chapter 3**).

The device structure to be employed in this work is now defined following the establishment of the TCAD environment. The simulation are in two-dimensional space transposed in three-dimensional space with the *Cylindrical* option which enables to simulate a three-dimensional

cylindrical device. In this case, the device is assumed to be rotational symmetric around the vertical axis. By default, a device simulated in two dimensions is assumed to have a thickness in the third dimension of $1\mu m$. The *AreaFactor* option allow to specify the extension of the device in the remaining dimension. The mesh geometry is extracted from Sentaurus simulations as do physical fields at each vertices. These data are parsed with matlab and C++ languages to build-up an internal device representation and compute the avalanche breakdown probability, the NRM generation rate and the DCR as explained in **Section 2**, **Appendix D** and **Chapter 3**.

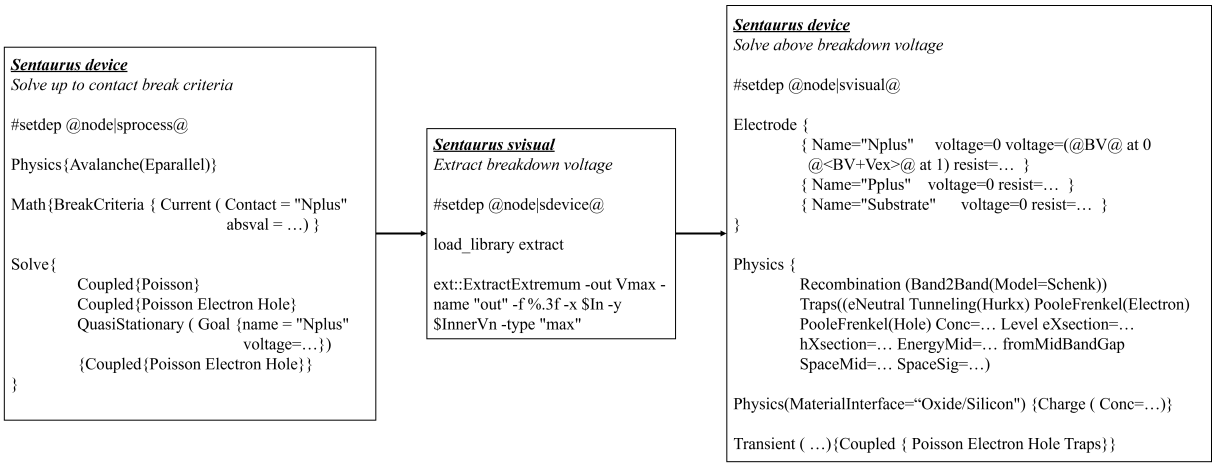


Figure 2.2: Sentaurus simulation scripts to simulate avalanche breakdown, extract avalanche breakdown voltage and solve device physics above breakdown condition without avalanche triggering. Some model examples are given which are introduced hereafter.

2 Computation techniques for avalanche breakdown probability

In this section, the concept of carrier energy threshold for impact-ionization is first introduced to gain insights on the role of the electric field in the carrier multiplication process. The impact-ionization coefficients are then defined and coupled with two different methodologies, one based on electric field streamlines and the other drift-diffusion Monte-Carlo. These approaches are utilized to compute the avalanche breakdown probability P_t either deterministically or stochastically, unveiling carrier contribution to jitter timing.

2.1 Ionization threshold energy against electric field

The threshold energy for impact ionization is the minimum energy required for a carrier to generate an electron-hole pair in a semiconductor material through impact ionization. When a carrier collides with an atom in the semiconductor material, it can transfer some of its energy to the atom, causing one or more bound electrons to be ejected from the valence shell of the atom. Therefore, a new electron-hole pair is generated, which can also gain kinetic energy to ionize more bound electrons. This results in self-sustained chains of impact ionization events. However, not all collisions result in impact ionization. For impact ionization to occur, the electron must gain sufficient kinetic energy under the influence of an electric field in the depletion region to overcome the energy barrier associated with creating an electron-hole pair. The laws of energy conservation and momentum conservation must be fulfilled in the process of impact ionization, which constrains the threshold energy E_{th} to be greater than the bandgap energy E_g. In the parabolic band approximation, represented in **Fig. 2.3 (a)**, the threshold energy is given by [49][50] for electrons:

$$E_{th}^{n} = E_g \left(1 + \frac{m_n^*}{m_n^* + m_p^*}\right) \tag{2.1}$$

and for holes:

$$E_{th}^{p} = E_g \left(1 + \frac{m_p^*}{m_n^* + m_p^*}\right) \tag{2.2}$$

where the effective mass in silicon for electrons is m_n^*=0.19/m_0 and for heavy holes m_p^*=0.49/m_0 with m_0= 9.11 $\cdot 10^{-31}$kg the free electron rest mass [51]. The effective mass is inversely proportional to the curvature of the band $m^* = \frac{\hbar^2}{d^2E/dk^2}$ with k the wave vector and E the energy of a carrier in a crystal lattice. This gives E_{th}^{n} = 1.44eV for electrons and E_{th}^{p} = 1.93eV for holes. The parabolic approximation fails to accurately describe the energy-momentum relationship in some materials because the energy bands near the bandgap have multiple valleys with different curvatures as shown in **Fig. 2.3 (b)**. To consider this anisotropic behavior, the effective mass of the carriers varies depending on the direction of momentum. This can lead to significant variations in the threshold energy depending on the crystallographic orientation of the material. In addition, the calculation of the threshold energy requires the consideration of all four bands of the valence band, which includes the two degenerate heavy hole, light hole, and split-off band.

In silicon, the minimum values reported are 1.1 eV for electrons and 1.8 eV for holes [52][53].

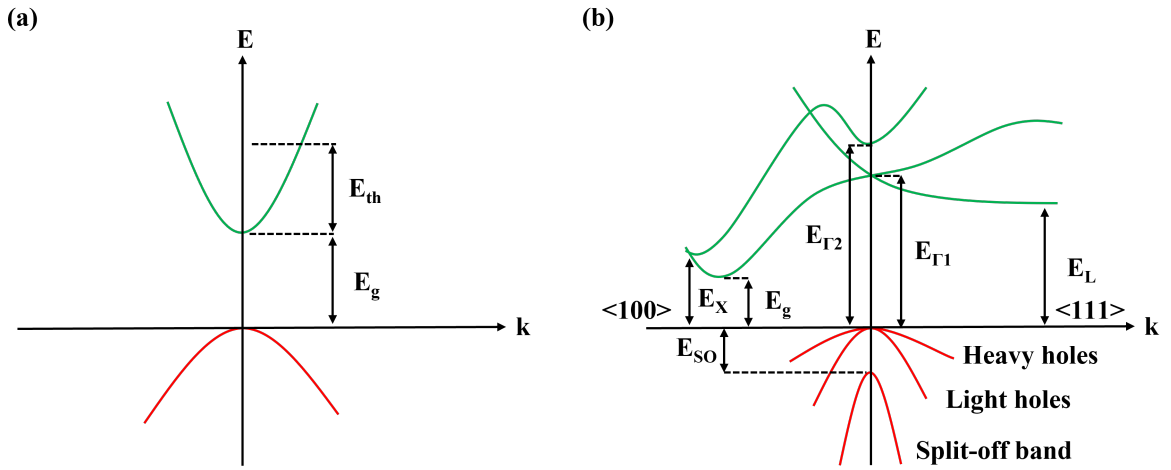


Figure 2.3: (a) Schematic of the parabolic band approximation. At any energy electron and hole are characterized by their isotropic effective mass m_n^* and m_p^*. The narrower the parabola, the smaller the effective mass, the lower the threshold energy. (b) Schematic of the silicon band structure for the conduction and valence bands with the different minima and maxima depending on the crystallographic orientation.

As the carrier gains energy, the probability of impact ionization increases, since the carrier is more likely to have enough energy to experience impact ionization. This rate of impact-ionization probability is modeled by the Keldysh formula [54]:

$$P_{II} = P_{II}^0 (E - E_{th})^{\alpha_0} \tag{2.3}$$

with E and E_{th} the carrier energy and the threshold impact-ionization energy, P_{II}^0 is the attempt rate and α_0 the energy exponent of the impact-ionization dynamics. *ab initio* studies [55][56][57] have demonstrated that the probability of impact-ionization is anisotropic with respect to the direction of momentum, in relation to the threshold energy.

Table 2.1: Parameters of the Keldysh model for silicon extracted from *ab initio* studies [58][59].

Type	P_{II}^0 ($eV^{-\alpha_0} \cdot s^{-1}$)	α_0	E_{th} (eV)
Electron	10^{11}	4.6	1.1
Hole	$1.14 \cdot 10^{12}$	3.4	1.49

The influence of the carrier energy and the band structure has been explained in the process of impact ionization. The upcoming discussion will focus on how the energy acquired by carriers is affected by the strength of the electric field F. The rate at which the energy of electrons changes over time is influenced by both the electric field and scattering mechanisms, which describe the conservation of momentum and energy. This relationship is expressed as [50][60]:

$$\frac{d(m^*_{n,p}v_d^{n,p})}{dt} = q \cdot F - \frac{m^*_{n,p}v_d^{n,p}}{\tau_m^{n,p}} \tag{2.4}$$

$$\frac{dE}{dt} = q \cdot v_d^{n,p} \cdot F - \frac{E - E_0}{\tau_E^{n,p}} \tag{2.5}$$

where $E_0 = \frac{3}{2}k_BT$ is the thermal equilibrium energy at zero electric field and $v_d^{n,p}$ is the drift velocity for holes and electrons. $\tau_m^{n,p}$ and $\tau_E^{n,p}$ are the effective momentum and energy relaxation time which represent the average time it takes for an electron to lose momentum and energy due to collisions with other particles in the material. In other words, the time constants determine the time needed on average to randomize the momentum of an electron and how quickly carriers may dissipate or absorb energy from their surrounding. The energy relaxation time is considerably larger than the momentum relaxation time, meaning that carriers generally lose only a fraction of their kinetic energy in collisions, despite the randomization of their flight direction [61]. It is assumed that the carrier kinetic energy lost by the incoming particle during the impact is constant and equal to the band gap E_g [62][63]. q accounts for negative and positive charge of electrons and holes. In steady state, one derives the carrier energy in an electric field after a characteristic relaxation time has passed as:

$$v_d^{n,p} = \frac{q \cdot \tau_m^{n,p}}{m^*_{n,p}}F = \mu_{n,p} \cdot F \tag{2.6}$$

$$E = E_0 + \frac{(q \cdot F)^2\tau_m^{n,p}\tau_E^{n,p}}{m^*_{n,p}} \tag{2.7}$$

where $\mu_{n,p}$ is the mobility for electrons and holes. The additional energy that is provided by an electric field is directly proportional to the the square of the electric field (**eq. 2.7**). Due to the higher amount of kinetic energy brought to the carrier by the electric field, a higher impact-ionization probability is expected at higher electric fields. The electric field allowing a carrier to reach its ionization energy threshold is then:

$$F_{th}^{n,p} = (\frac{E_{th}^{n,p} - E_0}{q\mu_{n,p}\tau_E^{n,p}})^{1/2} \tag{2.8}$$

Typical values at room temperature are μ_n=0.145m^2/V·s and μ_p=0.045m^2/V·s [64], $\tau_E^{n,p}$=0.1ps [65], E_0=39meV, and considering E_{th} defined by **eq. 2.1** with E_g=1.12eV [66] gives F_{th}^n=0.086MV/cm for electrons and F_{th}^p=0.198MV/cm for holes. These values match pretty well with the threshold electric value measured by van Overstraeten – de Man [67].

2.2 Impact-ionization coefficients against electric field and temperature

The previous discussion has focused on the energy and electric field strength required to trigger an impact-ionization. Now, the focus will shift towards analyzing the rate of impact-ionization. Once the electric field surpasses a certain threshold, it initiates carrier impact ionization of carriers in the valence shell of the atom, causing potentially a self-sustaining multiplication

process as represented in **Fig. 2.4**. The impact-ionization coefficients $\alpha_{n,p}$ represent the number of electron-hole pairs generated per unit distance in semiconductor materials due to impact-ionization. The distance that a carrier must travel without collisions in order to reach the threshold energy E_{th} for impact ionization is given by:

$$\Lambda_c = \frac{E_{th}}{qF} \tag{2.9}$$

The probability of not colliding during a small step of length $\Delta\lambda_c$ can be written as:

$$P(\Delta\lambda_c) = exp(-\frac{\Delta\lambda_c}{\Lambda_0}) \tag{2.10}$$

where Λ_0 is the mean free path of the carrier, which is the average distance a carrier travels between collisions with other particles in the material. This equation assumes that the probability of collision during a small step is proportional to the length of the step. Considering the Taylor series expansion of the exponential function with the limit defined as $\Delta\lambda_c$ approaches zero, one obtains:

$$exp(-\frac{\Delta\lambda_c}{\Lambda_0}) = 1 - \frac{\Delta\lambda_c}{\Lambda_0} + \frac{(\Delta\lambda_c)^2}{2!\Lambda_0^2} - \frac{(\Delta\lambda_c)^3}{3!\Lambda_0^3} + \cdots \tag{2.11}$$

and truncating this series after the first term, under the assumption that $\Delta\lambda_c$ is much smaller than the mean free path Λ_0, one gets:

$$exp(-\frac{\Delta\lambda_c}{\Lambda_0}) \approx 1 - \frac{\Delta\lambda_c}{\Lambda_0} \tag{2.12}$$

To calculate the probability of not colliding during a distance Λ_c, the probability of not colliding during each small step of length $\Delta\lambda_c$ must be considered that adds up to the total distance Λ_c. This is represented by the product:

$$P(\Lambda_c) = \prod_{i=1}^{N} P(\Delta\lambda_c^i) \tag{2.13}$$

where N is the number of steps taken by the carrier to travel the distance Λ_c. Since the probability of not colliding during each small step is given by an exponential function, the product of these probabilities is also an exponential function, this leads to the approximation:

$$P(\Lambda_c) = \prod_{i=1}^{N} P(\Delta\lambda_c^i) \approx \prod_{i=1}^{N} exp(-\frac{\Delta\lambda_c^i}{\Lambda_0}) \approx exp(-\frac{\Lambda_c}{\Lambda_0}) \tag{2.14}$$

and including **eq. 5.23** into this expression results in:

$$P(\Lambda_c) \approx exp(-\frac{E_{th}}{qF\Lambda_0}) \tag{2.15}$$

This equation shows that the probability of impact ionization increases exponentially with increasing electric field strength and decreasing mean free path. Following this approach, the

impact-ionization coefficients can be expressed by the empirical Chynoweth law [68]:

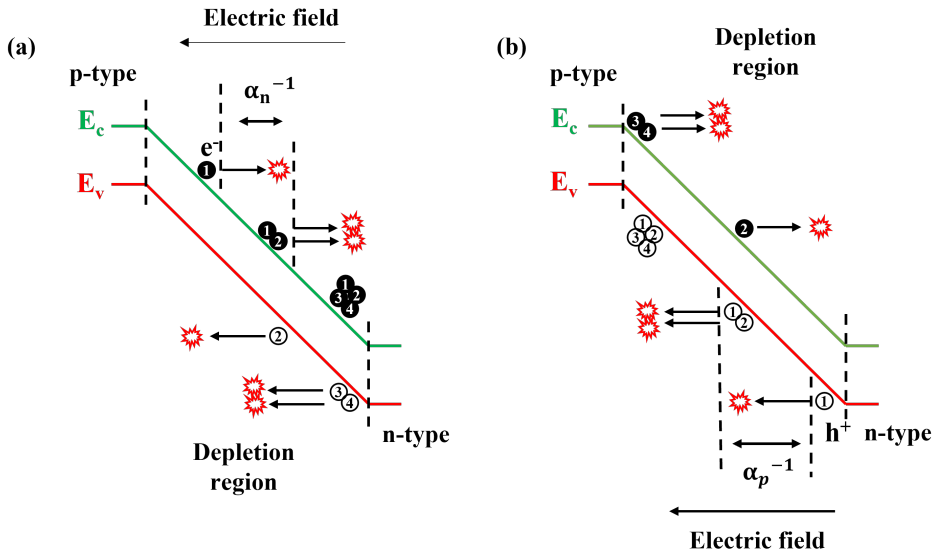


Figure 2.4: Schematic of impact-ionization process represented for (a) electrons and (b) holes. The impact-ionization coefficients represent the number of electron-hole pairs generated per unit distance in semiconductor materials due to impact-ionization. They are higher for electrons than for holes because electrons have a smaller effective mass, which leads to a lower threshold energy for impact ionization.

$$\begin{cases} \alpha_n = \gamma a_n exp(\frac{-\gamma b_n}{E}) \\ \alpha_p = \gamma a_p exp(\frac{-\gamma b_p}{E}) \end{cases} \tag{2.16}$$

The mean free path for the electrons and holes with respect to optical phonons is expressed as:

$$\Lambda_{n,p}(T) = \Lambda^0_{n,p} tanh(\frac{E_p}{2k_B T}) \tag{2.17}$$

The interaction of carriers with the optical phonons is obtained from the parameter γ using $\Lambda_{n,p}$ which is expressed as:

$$\gamma = \frac{\Lambda_{n,p}(T_0)}{\Lambda_{n,p}(T)} = \frac{tanh(\frac{\hbar\omega_{op}}{2k_B T_0})}{tanh(\frac{\hbar\omega_{op}}{2k_B T})} \tag{2.18}$$

T_0=300K and T are the reference and the absolute temperature. The optical-phonon energy E_p is equal to 63 meV in silicon. The carrier-phonon mean free path $\Lambda_{n,p}$ for electrons is 76 Å, and for holes, it is 55 Å, as reported in [69]. The optical-phonon angular frequency ω_{op} is equal to 63meV [67]. The coefficients $a_{n,p}$ and $b_{n,p}$ were measured as a function of electric field and temperature, as illustrated in **Fig. 2.5** and **Fig. 2.6**. Two sets of coefficient parameters are defined in **Table 2.2** for different ranges of electric field: one for low electric field (0.175 to 0.4

MV/cm) and another for high electric field (0.4 to 0.6 MV/cm). The thermal dependence of the bandgap in the threshold energy for impact-ionization is also considered through the coefficient $b_{n,p}$ in **eq. 2.16** which is defined as:

$$b_{n,p} = \frac{\beta_{n,p} E_g}{q \Lambda_{n,p}} \tag{2.19}$$

By using the impact-ionization coefficients $\alpha_{n,p}$, it is possible to determine the number of electron-hole pairs generated per unit distance traveled based on the local electric field and lattice temperature. This value will be employed subsequently to derive the avalanche breakdown probability P_t.

Conclusions

The impact-ionization of electron-hole pairs in semiconductor material requires a certain threshold energy to occur. It depends on the crystallographic orientation and varies with the bandgap energy. The rate of impact ionization is influenced by this threshold energy, which is mainly determined by the kinetic energy brings by the electric field to carriers. A threshold electric field can be defined accordingly to the material properties. The impact-ionization coefficients for electrons and holes consider these factors to calculate the number of electron-hole pairs generated per unit distance. The carrier-phonon mean-free path decreases with increasing the lattice temperature which results in an decrease of the impact-ionization rate.

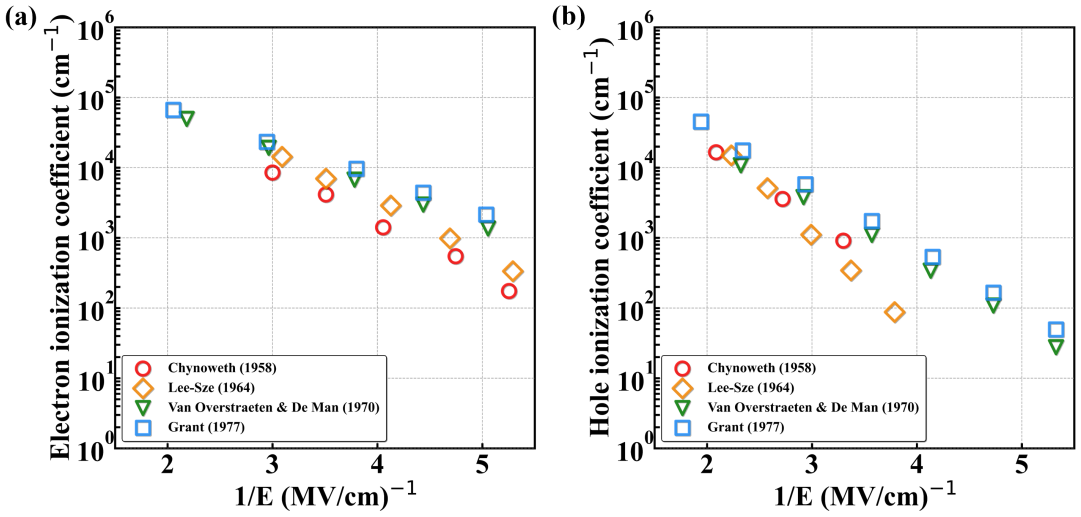


Figure 2.5: Impact-ionization coefficients as a function of the reverse of the electric field for (a) electrons and (b) holes extracted from measurements reported in [68][70][67][71]. At equivalent electric field, electrons impact-ionize more electron-hole pairs than holes. This is due to their smaller effective mass in silicon leading to lower threshold energy for impact-ionization (**eqs. 2.1-2.2**) and to higher acquired energy when drifting in the electric field (**eq. 2.7**), resulting in a impact-ionization rate.

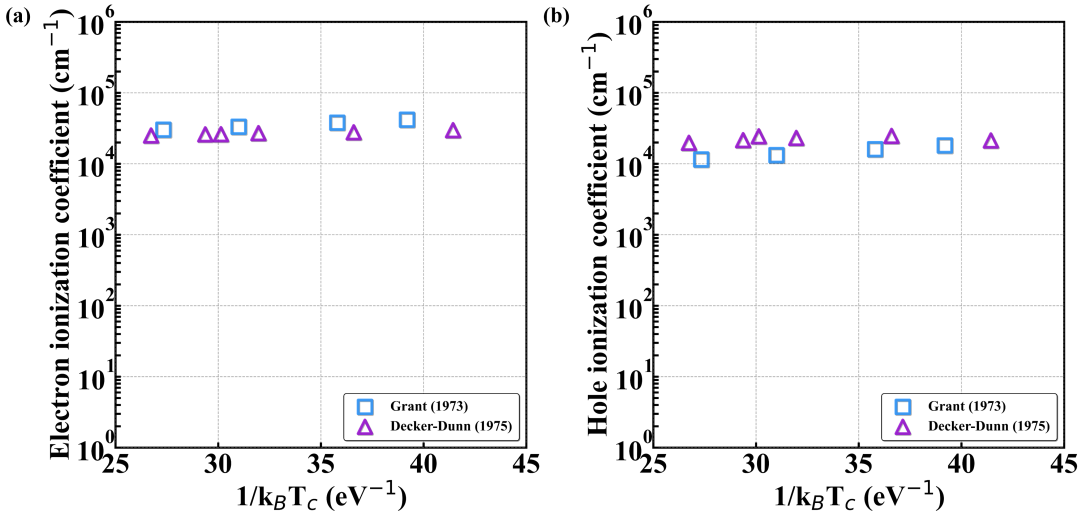


Figure 2.6: Impact-ionization coefficients as a function of $1/k_BT$ for (a) electrons and (b) holes extracted at an electric field of 0.4MV/cm from measurements reported in [71][72]. For the purpose of comparison with electric field dependence, the same scale as **Fig. 2.4** is used. The carrier-phonon mean free path decreases with increasing lattice temperature which results in a lower impact-ionization rate (**eq. 2.19**) under the same electric field.

Table 2.2: Parameters of the van Overstraeten – de Man model **eqs. 2.16-2.19** for silicon [67]. The model is activated by setting *Recombination(Avalanche(vanOverstraeten BandgapDependence))* in physics section and all the parameters are accessible in *vanOverstraetendeMan* parameter set.

Symbol 1	Parameter	Electrons	Holes	Valid range of electric field	Unit
$a_{n,p}$	a(low)	$7.03{\cdot}10^5$	$1.582{\cdot}10^6$	0.175 to 0.4 MV/cm	cm^{-1}
	a(high)	$7.03{\cdot}10^5$	$6.71{\cdot}10^5$	0.4 to 0.6 MV/cm	cm^{-1}
$b_{n,p}$	b(low)	$1.231{\cdot}10^6$	$2.036{\cdot}10^6$	0.175 to 0.4 MV/cm	V/cm
	b(high)	$1.231{\cdot}10^6$	$1.693{\cdot}10^6$	0.4 to 0.6 MV/cm	
$\hbar\omega_{op}$	hbarOmega	0.063	0.063		eV
λ	lambda	$62{\cdot}10^{-8}$	$45{\cdot}10^{-8}$		cm
β	beta(low)	0.678925	0.815009	0.175 to 0.4 MV/cm	cm
	beta(low)	0.678925	0.677706	0.4 to 0.6 MV/cm	1

2.3 Electric field-based derivation of avalanche breakdown probability

This section reports the simulation methodology presented in R. Helleboid et al., "Comprehensive Modeling and Characterization of Photon Detection Efficiency and Jitter Tail in Advanced SPAD Devices," IEEE Journal of the Electron Devices Society, vol. 10, pp. 584–592, Apr. 2022 [73] and its application to DCR modelling presented in M. Sicre et al., "Dark Count Rate in Single-Photon Avalanche Diodes: Characterization and Modeling study", in ESSCIRC 2021 - IEEE 47th European Solid State Circuits Conference (ESSCIRC), pp. 143–146, Sep. 2021 [74].

2.3.1 Non-linear computation

The models to forecast the avalanche breakdown probability P_t focus on the possibility of a carrier successfully initiating an avalanche event by means of self-sustaining carrier multiplication through impact-ionization. In a simple 1-D depletion region of SPAD shown **Fig. 2.7**, $P_n(x)$ is the probability that an electron starting at x in the depletion layer triggers an avalanche and $P_p(x)$ the same probability for an hole starting at x. The triggering probability of electrons and holes can be represented as a function of ionization coefficients and positions. The probability $P_\alpha(x)$ that either a hole or an electron trigger an avalanche is given by [75]:

$$P_\alpha(x) = 1 - (1 - P_n(x))(1 - P_p(x)) = P_n(x) + P_p(x) - P_n(x)P_p(x)$$

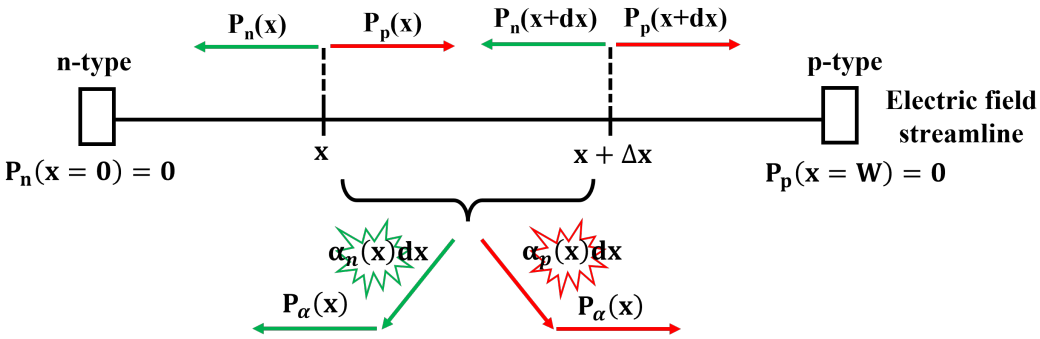


Figure 2.7: Schematic of McIntyre modeling of the avalanche breakdown probability for electrons (green) and holes (red) moving along electric field streamline. Either carriers travel to x then causes an avalanche. Or carriers cause an impact ionization between x+dx and x. The resulting hole-electron pair then causes an avalanche.

The probability that an electron starting at x + dx triggers an avalanche is equal to the probability that the electron reaches the position x and triggers an avalanche in x plus the probability that it triggers an avalanche between x and x + dx less the probability of the intersection of the two previous events. It writes considering **eq. 2.3.1**:

$$\begin{aligned} P_n(x+dx) &= P_n(x) + \alpha_n(x)dxP_\alpha(x) - P_n(x)\alpha_n(x)dxP_\alpha(x) \\ &= P_n(x) + \alpha_n(x)dx\left(P_n(x) + P_p(x) - P_n(x)P_p(x)\right) \\ &\quad - P_n(x)\alpha_n(x)dx\left(P_n(x) + P_p(x) - P_n(x)P_p(x)\right) \\ &= P_n(x) + \alpha_n(x)dx(P_n(x) + P_p(x) - P_n(x)P_p(x))(1 - P_n(x)) \end{aligned}$$

One can rearrange the terms to obtain :

$$\frac{P_n(x+dx) - P_n(x)}{dx} = \alpha_n(x)(P_n(x) + P_p(x) - P_n(x)P_p(x))(1 - P_n(x))$$

Which leads to the first ordinary differential equation :

$$\frac{dP_n}{dx} = \alpha_n(x)(1 - P_n)(P_n + P_p - P_nP_p)$$

The same reasoning applies to the probability that an hole starting at $x - dx$ triggers an avalanche which leads to:

$$\begin{aligned} P_p(x-dx) &= P_p(x) + \alpha_p(x)dxP_\alpha(x) - P_p(x)\alpha_p(x)dxP_\alpha(x) \\ &= P_p(x) + \alpha_p(x)dx\left(P_n(x) + P_p(x) - P_n(x)P_p(x)\right) \\ &\quad - P_p(x)\alpha_p(x)dx\left(P_n(x) + P_p(x) - P_n(x)P_p(x)\right) \\ &= P_p(x) + dx\alpha_p(x)(P_n(x) + P_p(x) - P_n(x)P_p(x))(1 - P_p(x)) \end{aligned}$$

Again rearranging the equality :

$$\frac{dP_p(x) - P_p(x-dx)}{dx} = -\alpha_p(x)(P_n(x) + P_p(x) - P_n(x)P_p(x))(1 - P_p(x))$$

Which leads to the second ordinary differential equation :

$$\frac{P_p}{dx} = -(1 - P_p)\alpha_p(x)(P_n + P_p - P_nP_p)$$

Therefore one can draw up the McIntyre system for $0 \leq x \leq W$:

$$\begin{cases} \dfrac{dP_n}{dx} = \alpha_n(x)(1 - P_n)(P_n + P_p - P_nP_p) & (2.20) \\ \dfrac{dP_p}{dx} = -(1 - P_p)\alpha_p(x)(P_n + P_p - P_nP_p) & (2.21) \end{cases}$$

Adding the couple of boundary value conditions :

$$\begin{cases} P_n(x = 0) = 0 & (2.22) \\ P_p(x = W) = 0 & (2.23) \end{cases}$$

A full 1D coupled and non-linear boundary value problem is obtained. The impact rate coefficients are the only experimental degree of freedom of the McIntyre model. As previously

explained, the probability to trigger an avalanche is constrained at zero due to electric field values that are below the threshold one for impact-ionization. To predict the occurrence of multiplication, the electric field dynamics against voltage in the n++/p+ and n-/p-well regions are reported in **Fig. 2.8 (a)**. Once the electric field threshold for impact-ionization, reported in **Table 2.2**, is reached, carrier multiplications are expected to occur within n++/p+ and n-/p-well regions.

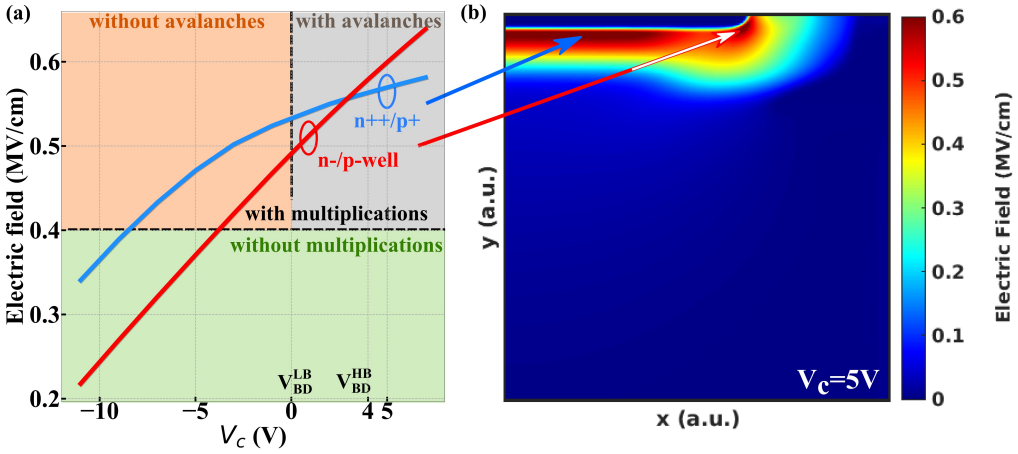


Figure 2.8: (a) Electric field as a function of the characterization voltage V_c normalized by the central avalanche breakdown voltage V_{BD}^{LB} in n++/p+ (blue) and n-/p-well (red) region simulated at the characterization temperature T_c=333K. Depending on the electric field magnitude, there is or not multiplication either in n++/p+ or n-/p-well region (green and orange box) without self-sustaining avalanches. The McIntyre simulation results presented in **Fig. 2.10** indicate the occurrence of avalanches in the n++/p+ region (gray box) above V_{BD}^{LB}. Avalanches in the n-/p-well are triggered at peripheral avalanche breakdown voltage V_{BD}^{HB}. (b) Electric field map simulated at V_c=5V and T_c=333K underlining the electric field activity in both region. For the sake of clarity, V_{BD} will refer to V_{BD}^{LB} in the following.

A damped Newton's method implemented in C++ language is utilized to solve this system of equations along electric field streamlines. Streamlines which tangent at any point in time gives the flow direction. Their values are computed using the Euler method implemented in C++ language, which calculates the position derivative over time with respect to the electric field vector $\overrightarrow{F}$, given by:

$$\frac{dX(s)}{ds} = \overrightarrow{F} \tag{2.24}$$

The streamline is $X(s)$ for $s \in [s_0, s_f]$ with s_0 and s_f the starting and the ending points respectively. The Euler method then reads:

$$\frac{X(s+ds) - X(s)}{ds} = \overrightarrow{F}(X(s)) \tag{2.25}$$

which can be rewritten as:

$$X(s+ds) = X(s) + \overrightarrow{F}(X(s))ds = X(s) + dX \tag{2.26}$$

By discretizing the problem, where X^k represents the approximation of $X(k*ds)$, one obtains:

$$\frac{X^{k+1} - X^k}{ds} = \overrightarrow{F}(X^k) \tag{2.27}$$

which can be reformulated as:

$$X^{k+1} = X^k + \overrightarrow{F}(X^k)ds = X^k + dX^k \tag{2.28}$$

In SPAD simulations, the electric field range can span from 0 to $10^6 V \cdot cm^{-1}$, leading to dX^k being either too small or too large. This can result in streamlines with significant differences in step size. To tackle this problem, the size of dX^k is dynamically adjusted at each iteration by fitting ds, thereby ensuring that dX^k remains within the range of 1nm to 10nm. Prior to proceeding, it is important to note that the definition of streamlines enables the investigation of various physical fields by substituting the electric field vector with others. For example, the gradient of quasi-Fermi Potential $E_F^{n,p}$ and electrostatic potential ϕ, or the current density vector $\overrightarrow{j_{n,p}}$, can be resolved. However, these alternative definitions can result in discrepancies in the simulation, as demonstrated in **Fig. 2.9** under equivalent impact-ionization parameters.

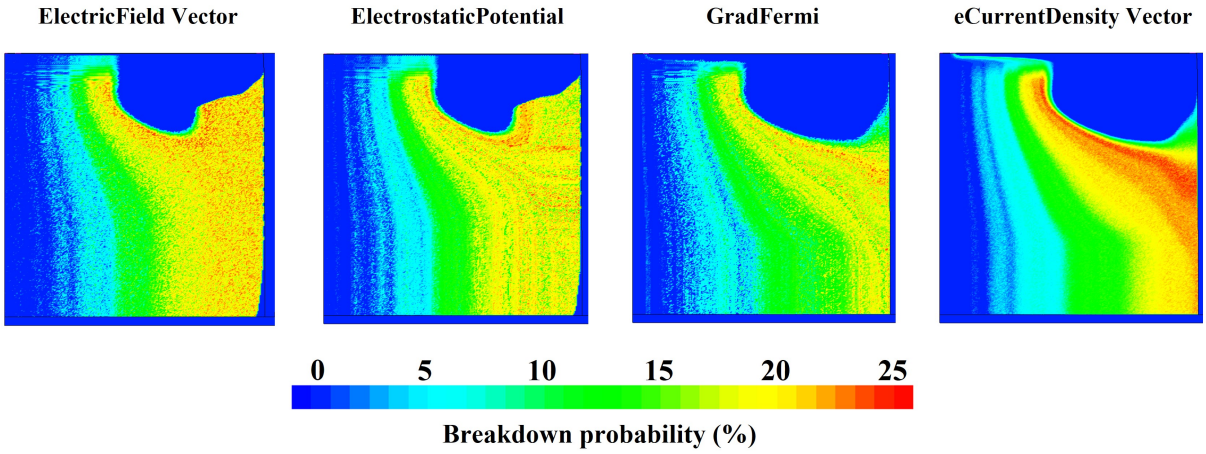


Figure 2.9: Overall avalanche breakdown probability map simulated at T_c=333K and V_c=4V for different streamline definitions from left to right: electric field vector $\overrightarrow{F}$, gradient of electrostatic potential ϕ and of quasi-Fermi Potential $E_F^{n,p}$, and current density vector $\overrightarrow{j_{n,p}}$. The names reported above each figure correspond to the ones define in [47]. The architecture depicted in the simulation deviates from those illustrated in Figures 2.10 and 2.11.

The simulated map of the avalanche breakdown probability P_t are shown in **Fig. 2.10** for voltages and **Fig. 2.11** for temperatures. Different carrier avalanche-leading region are activated depending on the temperature, applied bias and type of carriers. P_t is then extracted for holes as a function of voltage and temperature in **Fig. 2.12** to highlight the thermal and voltage dynamic in central n++/p+ and peripheral n-/p-well regions. The probability for holes has been selected for representation due to its closer proximity to the interface, which is more susceptible to the presence of defects. Two dynamics are observed in the voltage domain related to the strength of the electric field within the n++/p+ and n-/p-well regions as shown in **Fig. 2.8 (a)**).

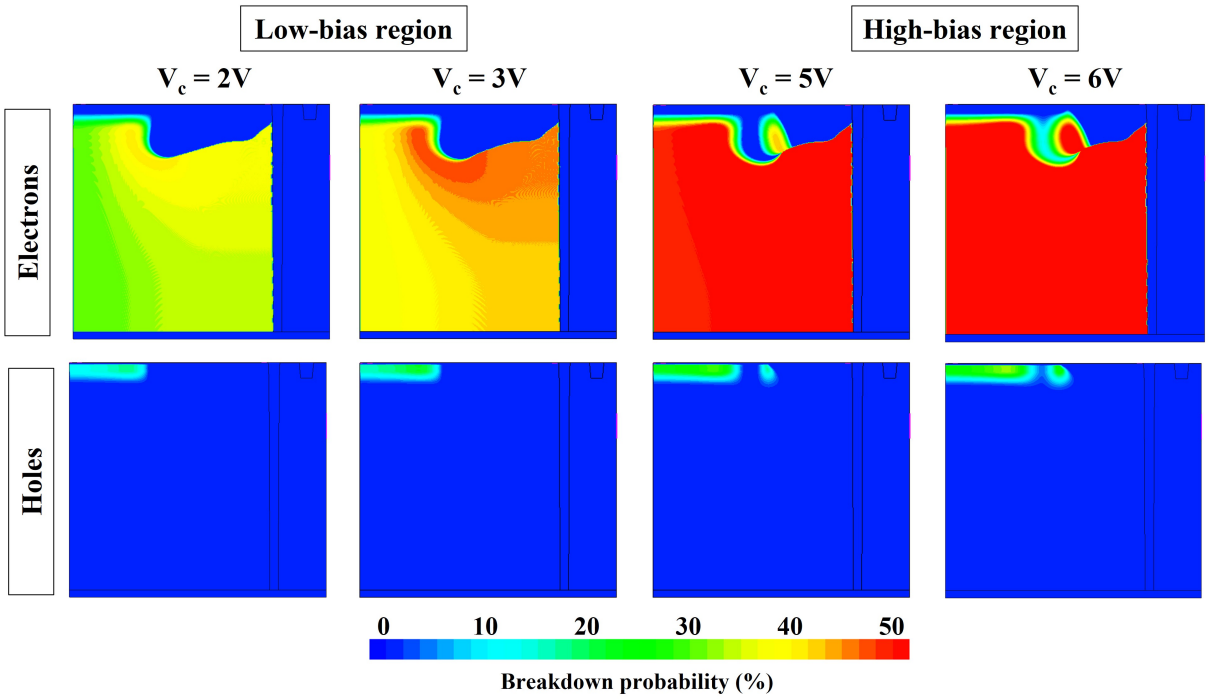


Figure 2.10: Avalanche breakdown probability map simulated at T_c=333K and V_c=2V, 3V, 5V and 6V for electrons (top) and holes (bottom). The avalanche breakdown probability extension and magnitude rise as the applied voltage is increased.

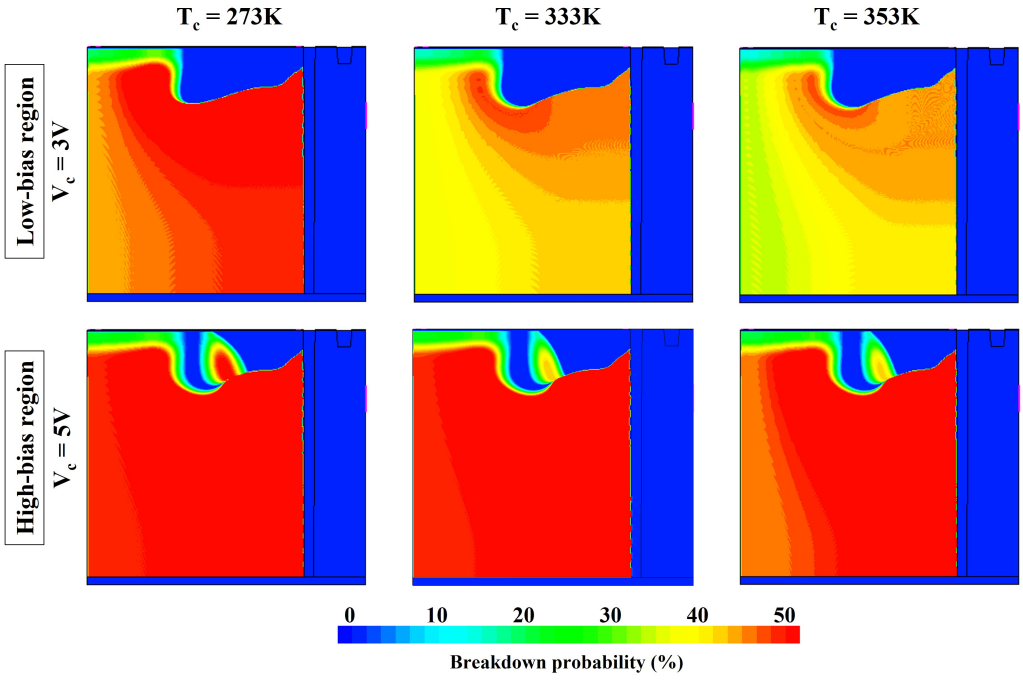


Figure 2.11: Overall avalanche breakdown probability map simulated at T_c=273K, 333K, 353K and V_c=3V (top) and 5V (bottom). The avalanche breakdown probability extension and magnitude fall as the temperature is increased.

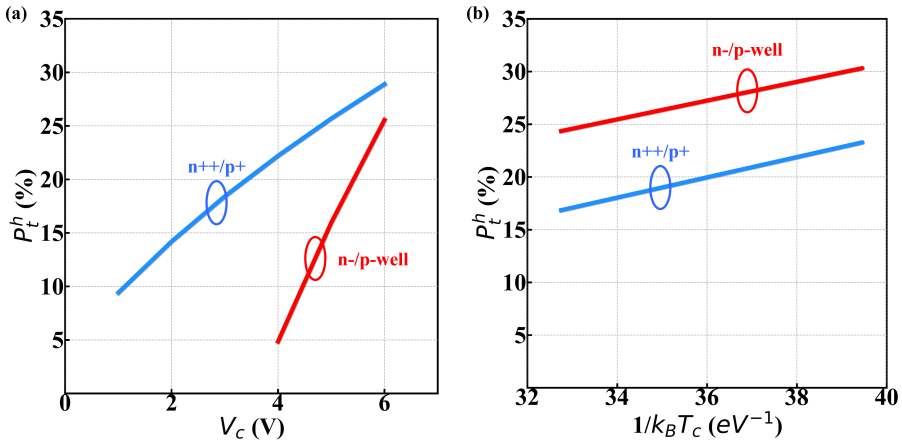


Figure 2.12: (a) Hole breakdown probability as a function of V_c in n++/p+ (blue) and n-/p-well (red) regions simulated at T_c=333K. The electric field trend depicted in **Fig. 2.8** gives rise to two voltage dynamics that can be observed in each region. (b) Hole breakdown probability as a function of $1/k_BT_c$ simulated in n++/p+ region at V_c=3V and in n-/p-well region at V_c=6V. The temperature dependence of impact-ionization coefficients is responsible for the thermal dynamics.

2.3.2 Photo-Detection Efficiency analysis

The simulation's accuracy can be assessed in relation to experimental data by using the **eq. 1.2** equation, which takes into account P_t to calculate PDE. The optical absorption is computed from FDTD Lumerical solver [76]. PDE is captured by the model as a function of temperature, as shown in **Fig. 2.13 (a)**, and as a function of voltage, as shown in **Fig. 2.13 (b)**. As demonstrated in **Fig. 2.6**, the positive thermal activation exhibited by the impact-ionization coefficients is responsible for the positive thermal dependence of PDE in relation to P_t, rather than the negative thermal dependence shown by optical absorption (**eq. 1.3**), as depicted in **Fig. 1.22**.

Another advantage of using this avalanche breakdown probability approach is the ability to determine V_{BD} from the PDE versus voltage curve. The criterion for device breakdown is set at a PDE threshold of 10^{-5} and V_{BD} is identified as the minimum applied voltage at which the threshold is reached. To achieve this, various architectures are compared through the measurement and simulation of PDE and V_{BD}, as illustrated in **Fig. 2.14 (a)-(b)**.

An alternative method, as described in **Fig. 2.2**, for extracting V_{BD} directly in Sentaurus workbench is available. The extracted value corresponds to the maximum voltage ($Vmax$) obtained in the simulation, based on a break criteria that depends on the current value ($BreakCriteria\ Current(Contact = "Nplus" absval = ...)$. The $absval$ is determined by the desired gain, which is defined as the ratio between current with and without avalanches. The unit of the gain depends on the device dimension, being $A\mu m^{-2}$ for 1D, $A\mu m^{-1}$ for 2D, and A for 3D devices. For this particular simulation, the gain is set to 10^5. It is worth noting that a significant correlation is observed between the simulation results obtained using the McIntyre model with a PDE approach and those obtained from the Sentaurus environment using current extraction, as illustrated by **Fig. 2.14 (c)**.

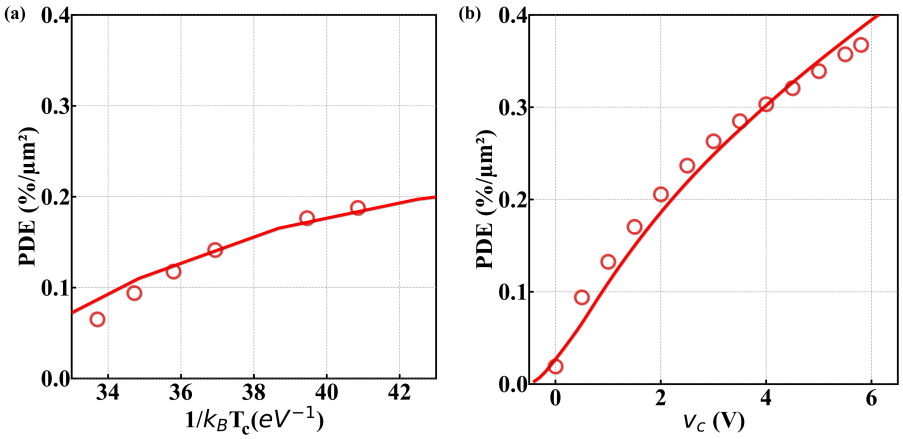


Figure 2.13: PDE measured (symbols) and simulated (bold lines) as a function of (a) $1/k_BT_c$ and (b) V_c for the reference diode. The thermal and voltage dependencies are in accordance with the trends of impact-ionization coefficients as reported in **Fig. 2.5** and **Fig. 2.6**.

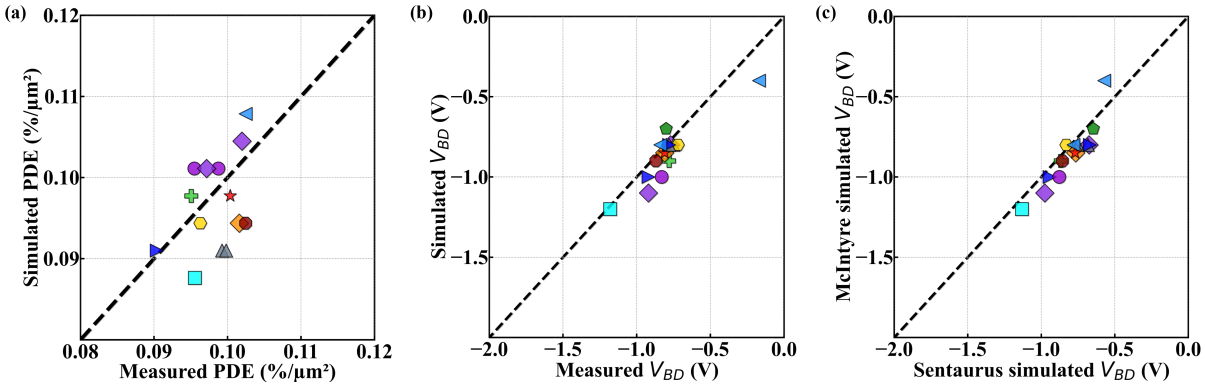


Figure 2.14: Correlation plots between experiments (x-axis) and simulations (y-axis) for (a) PDE and (b) V_{BD} ($= V_{BD}^{LB}$) with respect to a reference voltage at T_c=333K. The simulation results obtained from both the McIntyre model and sdevice tool [47] using the gradient of quasi-Fermi Potential $E_F^{n,p}$ are compared in (c). Each symbol corresponds to a different device architecture. The dashed line is the y=x line highlighting a satisfactory correlation between experiments and simulations.

The solutions of the model are obtained along electric field streamlines according to impact-ionization coefficients which depend predominantly on the electric field strength. Factors influencing the electric field distribution within the device must be carefully considered as they switch-on/off some regions to avalanche triggering and increase the probability for others. The doping profile and device dimensions, for example, influence the electric field distribution within the device, affecting the onset of avalanche breakdown. For the purpose of visibility, the overall probability maps for avalanche breakdown probability P_t corresponding to the variations described below are reported in **Appendix A**. This encompasses several factors, such as the use of all-around oxides to modulate the fixed charge density (**Fig. A.4**), which in turn affects P_t in the vicinity of the interfaces. Epitaxial growth procedures can modify the distribution of dopants in the bulk (**Fig. A.2**), resulting in changes to P_t throughout the volume. Varying

ion implantation doses and energies (**Fig. A.3**) can also reduce or increase P_t contribution in the central or peripheral parts. Additionally, architectural constraints like lateral extension or overlap of regions (**Fig. A.1**) can either diminish or extend P_t contribution in the central or peripheral parts. All these factors will be studied on the light of DCR in the next chapter.

2.3.3 Analytical computation

Another approach for calculating P_t based on the same principle involves differentiating the **eq. 2.20** and **eq. 2.21** [77], which results in:

$$\frac{dP_\alpha}{dx} = (\alpha_n(x) - \alpha_p(x))P_\alpha(1 - P_\alpha) \tag{2.29}$$

which can be integrated to give:

$$\frac{P_\alpha(x)}{1 - P_\alpha(x)} = \frac{P_\alpha(0)}{1 - P_\alpha(0)} exp(\int_0^x (\alpha_n - \alpha_p)dx\prime) \tag{2.30}$$

from boundary conditions reported in **eq. 2.22** and **eq. 2.23**:

$$\frac{P_n(W)}{1 - P_n(W)} = \frac{P_p(0)}{1 - P_p(0)} f(W) \tag{2.31}$$

where $f(x) = exp(\int_0^x (\alpha_n - \alpha_p)dx\prime)$. By solving **eq. 2.30**, $P_\alpha(x)$ is geven by:

$$P_\alpha(x) = \frac{P_p(0) \cdot f(x)}{P_p(0) \cdot f(x) + 1 - P_p(0)} \tag{2.32}$$

The **eq. 2.20** is rewritten in the from:

$$-\frac{1}{1 - P_p}\frac{dP_p}{dx} = \frac{d}{dx} log(1 - P_p) = \alpha_p P_p \tag{2.33}$$

which leads to the integral form:

$$1 - P_p(x) = (1 - P_p(0))exp(\int_0^x \alpha_p(x\prime)P_\alpha(x\prime)dx\prime) \tag{2.34}$$

Since $P_p(W) = 0$:

$$-log(1 - P_p(0)) = \int_0^W \alpha_p(x\prime)P_\alpha(x\prime)dx\prime \tag{2.35}$$

The sdevice tool [47] incorporates this analytical model that utilizes **eq. 2.35** along with **eq. 2.32**. The multiplication region is defined as the region where the electric field is higher than a threshold value which is set at 0.1MV/cm by default. Outside the multiplication region: $P_n(x) = P_p(x) = P_\alpha(x) = 0$. This can be modified by the option *BreakdownProbability(MinElectricField=...)* in the physics section. The option *AvalPostProcessing* must be specified in the math section to compute the avalanche breakdown probability without carriers generated by impact-ionization.

The avalanche breakdown probability map resulting from this methodology are reported in **Fig. 2.15**. The simulation approach, which relies on the impact-ionization coefficient for

thermal and electric field dynamics as demonstrated earlier, is expected to yield similar thermal and voltage trends for P_t. For instance, the avalanche breakdown of n++/p+ and n-/p-well regions are accurately reproduced. However this approach brings some discrepancies in the avalanche breakdown probability results. For instance, without cutting values below an electric field, the value predicted of the integrated P_t over the entire volume is less than 70% the one predicted with the model solving the equations along electric field streamlines, which is consistent with experimental results.

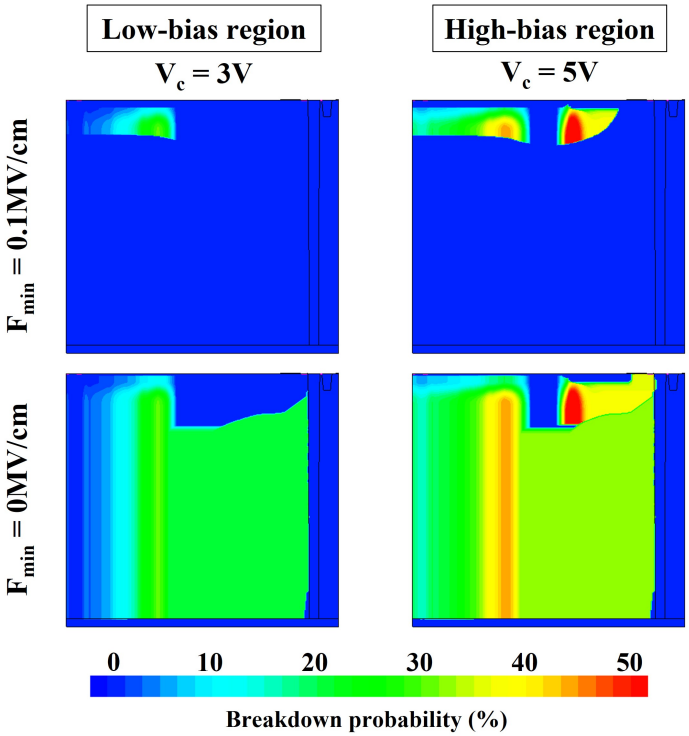


Figure 2.15: Overall avalanche breakdown probability map simulated by sdevice tool [47] at T_c=333K and V_c=3V and 5V for an minimum required electric field of 0.1MV/cm (top) and without (bottom). The avalanche breakdown of n++/p+ and n-/p-well regions are accurately reproduced. Nonetheless, the magnitude of P_t is lower than that predicted by the model that solves the equations along electric field streamlines as illustrated by **Fig. 2.11**.

Conclusions

The non-linear McIntyre model is utilized to simulate the avalanche breakdown probability P_t, by taking into account impact-ionization coefficients and electric field streamlines. The model allows for the examination of different physical fields to define streamlines, resulting in variations in the P_t map. The avalanche breakdown probability maps are dependent on temperature, voltage, and carrier type. Two voltage dynamics are related to the electric field in both central n++/p+ and peripheral n-/p-well regions, and one positive thermal trend is related to the temperature dependence of impact-ionization coefficients. The electric field distribution, and therefore P_t, is altered by the dopant profile within the volume and at the interface, as well as device dimensions. By using the Photo-Detection Efficiency (PDE) equation to compare simulated P_t with experimental results, the model is predictive for the whole range of temperature and voltage. The breakdown voltage V_{BD} is extracted from both the PDE using the model and the current using Sentaurus workbench, and the results show a good level of agreement. Another approach involves an analytical model implemented in the sdevice module to compute P_t, which reveals the activity of both regions but with a lower magnitude when integrated in the entire device.

2.4 Drift-diffusion Monte-Carlo method for avalanche breakdown probability

This section reports the simulation methodology presented in R. Helleboid et al., "A fokker–planck-based monte carlo method for electronic transport and avalanche simulation in single-photon avalanche diodes," Journal of Physics D: Applied Physics, vol. 55, no. 50, p. 505 102, Oct. 2022 [78] and its application to DCR modelling presented in M. Sicre et al., M. Sicre et al., "Statistical measurements and Monte-Carlo simulations of DCR in SPADs," ESSCIRC 2022- IEEE 48th European Solid State Circuits Conference (ESSCIRC), pp. 193-196, Sep. 2022 [79].

2.4.1 Carrier random walk modeling

The approach based on electric field streamlines is a simplified modeling technique that focuses on the macroscopic behavior of carriers. The electric field-based approach is computationally efficient and suitable for analyzing device-level behavior and large-scale simulations. A crucial aspect of device operation is the statistical fluctuations and random processes that affect the behavior of carriers, which cannot be achieved using this deterministic method. Monte-Carlo methods are more detailed and computationally intensive approach that incorporates statistical fluctuations and microscopic interactions of carriers. It uses a particle-based simulation technique where individual carriers are tracked and their trajectories are calculated based on probabilistic rules. The carriers that are governed by the drift-diffusion equation are subject to [80]:

$$\frac{\partial f}{\partial t}(r,t) = -\nabla \cdot (v_d(r,t) f(r,t)) + \nabla \cdot (D(r,t) \nabla f(r,t)) \tag{2.36}$$

In a system of N_e electrons, $f(r,t)$ is the carrier density probability that shows the probability distribution function of electron density at a specific position r and time t. To determine the spatial and temporal distribution of the local charge density $n(r,t)$, $f(r,t)$ can be multiplied by the number of electrons N_e. The drift velocity and diffusion coefficient, $v_d(r,t)$ and $D(r,t)$, respectively, depend on space and time at position r and time t. The solutions of **eq. 2.36** are temporal probability density function of the stochastic process X_t, which is defined according to the Euler-Maruyama Monte Carlo scheme as follows [81]:

$$dX_t = v_d(X_t,t)dt + \sigma(X_t,t)dB_t \tag{2.37}$$

X_t is a random variable that represents the position of a particle at time t. At position X_t and time t, $v_d(X_t,t)$ is the drift velocity of the particle while the diffusion matrix is $\sigma(X_t,t)$. The time-continuous random process known as Brownian motion, denoted as B_t, is the result of collisions with impurities, phonons, and other carriers. It is characterized by a sequence of infinitesimal random intervals that follow a centered Gaussian distribution with a variance of dt. The Box-Muller transform [82] is used to generate random numbers following standard normal distribution from a source uniformly distributed numbers. Each step is independent and identically distributed of the previous ones, resulting in a continuous random walk. Independence means that the occurrence of one event does not affect the probability of the occurrence of another event. Identically distributed means that the probability distribution of each random variable in the sequence is the same. The diffusion factor and diffusion coefficient are related through the following equation:

$$\sigma(X_t,t) = \sqrt{2D(X_t,t)} \tag{2.38}$$

The process can be viewed as a random walk that is associated with the diffusion of carriers and influenced by the drift velocity caused by the electric field. The transport is considered isotropic, allowing the velocity and diffusion coefficients to be simplified into scalars that rely on the electric field strength and doping concentration, via carrier mobility. Only electrons were considered previously, whereas in the following, there is a distinction between electrons and holes. The drift velocity for electrons and holes can be calculated as:

$$v_d^{n,p}(X_t,t) = \pm\mu_{n,p}(T,F)\cdot F \tag{2.39}$$

where $\pm$ stands for electrons and holes. For the diffusion coefficient the Einstein relation is assumed to hold for non-degenerate semiconductor:

$$D_{n,p}(X_t,t) = \frac{\mu_{n,p}(T,F)k_BT}{q} \tag{2.40}$$

The mobility for electrons and holes is computed by using the temperature and doping dependent model of Arora [83]:

$$\mu_0(T, N_{A,D}) = \mu_{\min}(T) + \frac{\mu_d(T)}{1 + \left(\frac{N_A + N_D}{N_0(T)}\right)^{A^*(T)}} \tag{2.41}$$

where

$$\mu_{min}(T) = A_{min} \cdot (\frac{T}{300K})^{\alpha_m} \tag{2.42}$$

$$\mu_d(T) = A_d \cdot (\frac{T}{300K})^{\alpha_d} \tag{2.43}$$

and

$$N_0(T) = A_N \cdot (\frac{T}{300K})^{\alpha_N} \tag{2.44}$$

$$A^*(T) = A_a \cdot (\frac{T}{300K})^{\alpha_a} \tag{2.45}$$

This model of Aurora is coupled with a high-field saturation regime model from Canali's work [84]:

$$\mu_{n,p}(T, F) = \frac{\mu_0(T, N_{A,D})}{(1 + (\mu_0(T, N_{A,D}) \cdot \frac{F}{v_s})^{\beta(T)})^{\frac{1}{\beta(T)}}} \tag{2.46}$$

where

$$\beta = \beta_0(\frac{T}{300K})^{\beta_{exp}} \tag{2.47}$$

Both model constants are reported in **Table 2.3 -2.4**.

Table 2.3: Parameters of the Arora model **eqs. 2.41-2.45** for silicon. The model is activated by setting *Mobility(DopingDependence(Arora))* in physics section and all the parameters are accessible in *DopingDependence* parameter set.

Symbol	Parameter	Electrons	Holes	Unit
A_{min}	Ar_mumin	88	54.3	$cm^2/V{\cdot}s$
α_m	Ar_alm	-0.57	-0.57	1
A_d	Ar_mud	1252	407	$cm^2/V{\cdot}s$
α_d	Ar_ald	-2.33	-2.33	1
A_N	Ar_N0	$1.25{\cdot}10^{17}$	$2.35{\cdot}10^{17}$	cm^{-3}
α_N	Ar_alN	2.4	2.4	1
A_a	Ar_a	0.88	0.88	1
α_a	Ar_ala	-0.146	-0.146	1

At each time step, the numerical scheme implemented in Matlab language updates the state of the system starting from a given position x_0 with a constant time step dt as follows:

Table 2.4: Parameters of the Canali model **eqs. 2.46-2.47** for silicon. The model is activated by setting *Mobility(HighFieldSaturation(Canali))* in physics section and all the parameters are accessible in *HighFieldDependence* parameter set.

Symbol	Parameter	Electrons	Holes	Unit
β_0	*beta0*	1.109	1.213	1
β_{exp}	*betaexp*	0.66	0.17	1

$$X_{n+1} = X_n + v_d(X_n, n \cdot dt)dt + \sigma(X_n, n \cdot dt)W_n \tag{2.48}$$

where $(W^n)_{n\geq 1}$ are n-th random variables in the sequence of the Brownian motion. This model provides a stochastic description of individual carriers submitted to drift and diffusion. An illustration of the random walk obtained is shown on **Fig. 2.16 (a)** for two different starting points and for different simulations. Even when particles are initially placed at the same starting point, they can exhibit distinct diffusion times, paths, and points of entry into junctions, as illustrated in **Fig. 2.16 (b)**. Furthermore, particles can overcome potential barriers in the device due to their thermal energy. This can lead to new avalanche active zones in the device, which can have a significant impact on the device's performance, as demonstrated by the subsequent simulation of avalanche breakdown probability and jitter. But first let's shift from the carrier motion to the carrier multiplication, which may lead to self-sustaining avalanches.

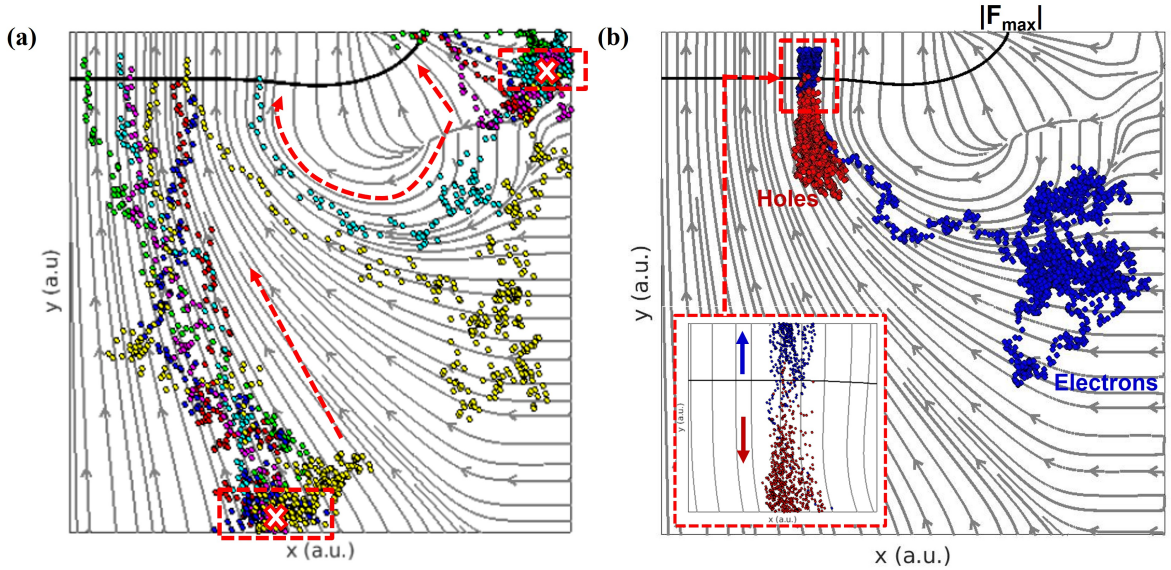


Figure 2.16: (a) Random walk for particles starting from one of the two crosses inside the boxes. Each color represents a potential carrier path. Despite having the same starting point, particles can show variations in their diffusion times, routes, and entry points into junctions. (b) Drift-diffusion Monte-Carlo simulation results for a single electron generated at the device interface. Electrons are represented in blue and holes in red. The inset represents the generation of electron-hole pairs in the multiplication region. Each point corresponds to a different time step. The grey lines represent the electrostatic potential streamlines.

2.4.2 Carrier multiplication toward avalanche breakdown formalism

The position of carriers computed from Monte-Carlo simulations can be used to understand the carrier multiplication process that leads to avalanche breakdown. Carriers reaching the main junction of the device have a low or high probability of triggering the avalanche breakdown due to the presence of a higher electric field in the junction region as shown in **Fig. 2.16 (b)**. Carrier multiplication process is modeled through electric field-dependent rates of impact ionization as previously described in **Section 2.2**. In a one-dimensional scenario, an initial hole current density at the position of $x = 0$, denoted as J_{p_0}, is considered to initiate the avalanche. As a result of impact ionization, the density $J_p(x)$ gradually increases throughout the depletion region, generating holes. Simultaneously, the generated electrons move in the opposite direction, leading to an electron current density J_n. In steady-state, the total current density $J_{total} = J_n + J_p$ must remain constant along the device. At the end of the depletion region at $x = W$, the hole current reaches its maximum. This maximum value can be characterized by defining the multiplication gain factor $M_{n,p}$ for electrons and holes, which is the ratio of the current flowing out to the current flowing in, as follows:

$$M_{n,p} = \frac{J_p(W)}{J_{p_0}} = \frac{J_{total}}{J_{p_0}} \tag{2.49}$$

The change of the J_p passing through the depletion region is evaluated as :

$$dJ_p = J_p\alpha_p dx + J_n\alpha_n dx \tag{2.50}$$

which can be reformulated using the property $J_{total} = J_n + J_p$ in steady state:

$$\frac{dJ_p}{dx} - (\alpha_p - \alpha_n)J_p = \alpha_n J \tag{2.51}$$

Applying the stated boundary conditions, this differential equation can be solved at the position $x = W$, where the hole current density becomes $J_p(W) = J_{total}$ to yield [85]:

$$1 - \frac{1}{M_p} = \int_0^W \alpha_p exp(\int_0^x (\alpha_p - \alpha_n)dx\prime)dx \tag{2.52}$$

and for electrons:

$$1 - \frac{1}{M_n} = \int_0^W \alpha_n exp(\int_0^x (\alpha_n - \alpha_p)dx\prime)dx \tag{2.53}$$

The avalanche breakdown sets in if the multiplication factor $M_{n,p}$ approaches infinity.

Now that the avalanche breakdown condition has been established, it is relevant to determine the probability of experiencing an impact-ionization event and, subsequently, carrier multiplication. The probability that an electron will not create an impact-ionization event between two positions x and x' can be derived from the Poisson probability mass function which is given by :

$$P_k(x|x') = \frac{\lambda^k e^{-\lambda}}{k!} \tag{2.54}$$

where k is the number of events and λ is the average rate of event occurrences. If a small interval of space between positions x and x' is considered, the probability of no ionization events occurring in that interval is given by the Poisson probability mass function with $k = 0$:

$$P_0(x|x') = \frac{\lambda^0 e^{-\lambda}}{0!} = e^{-\lambda} \tag{2.55}$$

Assuming that the ionization events are independent of each other and follow a Poisson distribution, which is to say that the number of ionization events is small relative to the total number of electrons passing through that interval, then the average rate of occurrence of ionization events in the interval between positions x and x' is given by [86][87]:

$$\lambda = \int_x^{x'} \alpha(z)dz \tag{2.56}$$

where $\alpha(z)$ is the ionization coefficient at position z. Substituting this expression for λ (**eq. 2.56**) into the equation for P_0 (**eq. 2.55**):

$$P_0(x|x') = exp(-\int_x^{x'} \alpha(z)dz) \tag{2.57}$$

The larger the value of the integral, the greater the probability of ionization, and the smaller the value of the exponential term, resulting in a lower probability that initiating ionization events occur. The impact-ionization probability is then:

$$P_{II}(x|x') = 1 - P_0(x|x') \tag{2.58}$$

Integrating this equation at each time step and considering the logarithm form ($p_{II} = ln(P_{II})$) results in:

$$p_{II}^{n+1} = p_{II}^n + v_d(X_n, n \cdot dt) \cdot dt \cdot \alpha_{n,p}(X_n) \tag{2.59}$$

In order to model the likelihood of impact ionization, a random number variable R is generated with a uniform distribution defined between 0 and 1. The position x where the probability is greater than or equal to the exponential of R is determined by calculating p_{II} along the carrier's trajectory. If such a position exists, an impact ionization event occurs at that location. Once an impact ionization event has occurred, p_{II} is reset to zero for the particle and a new random number R is generated. Both electrons and holes undergo the same process and may eventually experience impact ionization. The simulation is run 1500 times at each structure mesh point with one initial particle each time. Once the number of active particles reaches the threshold of one hundred particles for each simulation, as shown in **Fig. 2.17**, it is assumed that the self-sustaining avalanche condition has been achieved. The avalanche breakdown probability P_t is defined as the ratio of the number of simulations resulting in an avalanche to the total number of simulations performed. If the P_t value is 30%, it can be inferred that at least fifty thousand particles are produced, given that each successful simulation typically generates a minimum of

one hundred particles. This value represents a substantial multiplication gain factor.

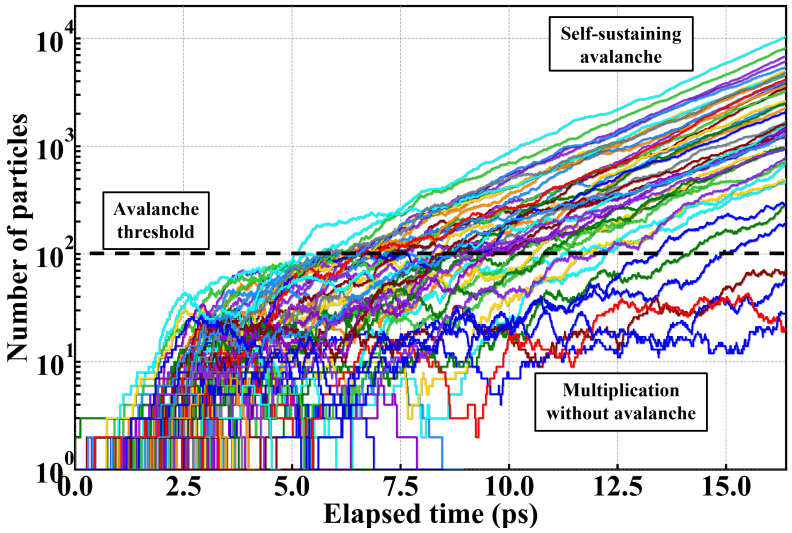


Figure 2.17: Number of carriers as a function of the time elapsed in the simulation. Each color represents a starting particle and all of its descendants, allowing to track the behavior of individual particles over time. Some particles exhibit exponential growth (top) while others multiply and are collected at the electrodes (bottom), which can be observed by the distinct patterns.

The **Fig. 2.18** presents a comparison of the breakdown probabilities obtained from the McIntyre model and the Monte-Carlo method. The results show a significant difference in the near contact region denoted as (B), where the Monte-Carlo method predicts a breakdown probability of up to 5% when integrating over the near contact region, while the McIntyre model predicts zero probability due to an electrostatic potential barrier. An electron, coming from the left ascends the potential. When the potential decreases, it faces a positive electric field that pushes it back towards the left. The barrier cuts the direction of the electric field lines, which prevents the carrier trajectories from following the electric field streamlines in the McIntyre model, leading to the inability to compute P_t. However, the Monte-Carlo method simulates the random movement, which enables carrier diffusion to overcome the electrostatic barrier and provides new possible paths for carriers. These paths include (A) diffusion and collection to the contacts without triggering an avalanche, leading to a reduction in P_t, or (B) reaching an avalanche active region resulting in an increase in P_t. Therefore, the Monte-Carlo approach offers additional path choices for carriers generated within the device, which is in contrast to the carrier trajectory constrained by electric field streamlines in the McIntyre model.

To investigate the relationship between the localization of carrier generation and avalanche triggering, the breakdown probability P_t is divided into two areas at the edge of the peripheral region. This allows to track the carrier's path to the avalanche triggering, either in the central part depicted in **Fig. 2.19 (a)**, or in the peripheral part of the device, as illustrated in **Fig. 2.19 (b)**. The Monte-Carlo method revealed that some carriers generated from the side interface have a probability of triggering peripheral avalanches, which may contributes to the high-bias dynamic of the DCR as elaborated in the subsequent chapter. This finding highlights

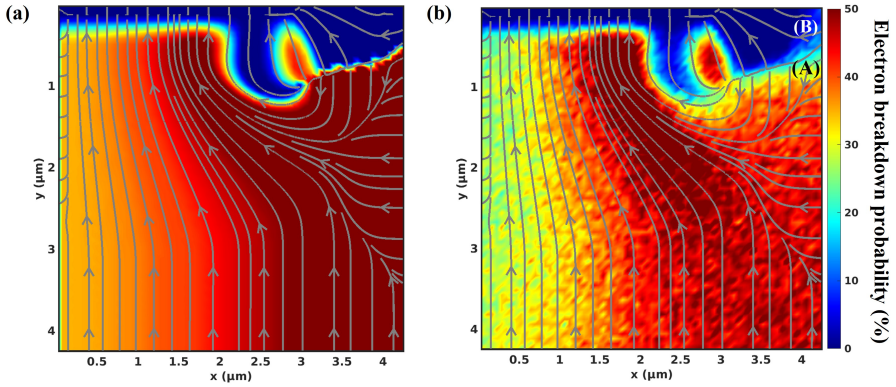


Figure 2.18: Electron breakdown probability with (a) McIntyre model compared to the (b) drift-diffusion Monte Carlo method simulated at V_c=5V and T_c=333K. (A) shows a P_t decrease due to carrier diffusion to contacts and (B) a P_t increase owing to carrier diffusion overcoming electrostatic potential barrier. The superimposed grey lines represent the electrostatic potential streamlines.

the importance of considering both the site of carrier generation and its random diffusion path to comprehend the device's performance. Following the same contribution separation approach, it is possible to compute the jitter by examining the time it takes for carriers to evacuate at each mesh point throughout the device.

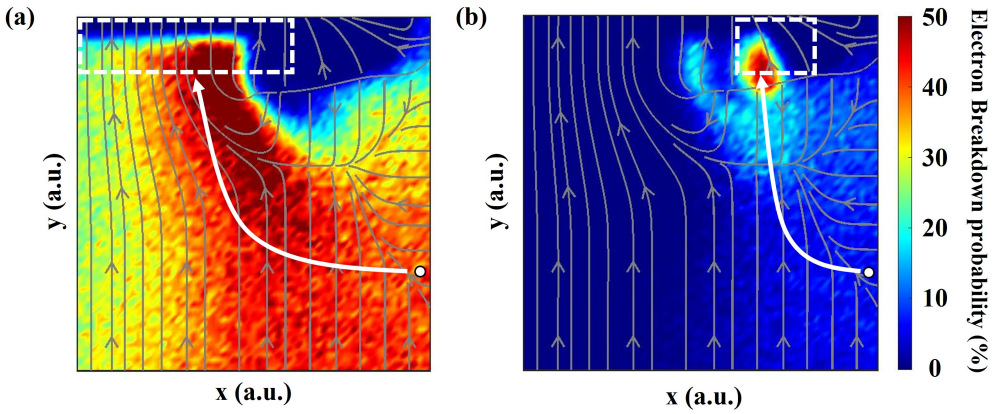


Figure 2.19: Electron breakdown probability simulated with the drift-diffusion Monte-Carlo at V_c=5V and at T_c=333K. The breakdown probability contributions are arbitrary split at the edge of the peripheral region. Each empty circles represented a starting position considering carrier whose ending in (indicated with white arrows): (a) the central region and (b) the peripheral region (represented by dashed white rectangles). The superimposed grey lines represent the electrostatic potential streamlines.

2.4.3 Jitter analysis

The drift-diffusion Monte-Carlo method can be also used to analyze the influence of different photogeneration starting points within the device. Multiple electrons are launched at each mesh

point and their trajectories are simulated. Upon detection of the first impact ionization event, the corresponding time is recorded. By computing the density function of these recorded times, it is possible to depict jitter and compare it with experimental data in **Fig. 2.20 (a)**. The timing distribution exhibits three distinct characteristics, as depicted in **Fig. 2.20 (b)-(c)**. The avalanche region generates a prominent peak near t=0, reflecting the stochastic nature of avalanche buildup [88]. The depletion region yields relatively short average timings associated with the drift transport time, specifically $t = d/v_d$, where d represents the distance to the avalanche region, and v_d corresponds to the drift velocity. In contrast, the undepleted region contributes to a broad, extended tail characterized by an exponential decay. This behavior arises from carriers diffusing from the undepleted region to the depleted side, exhibiting similar properties to carriers generated by photon absorption within the undepleted volume.

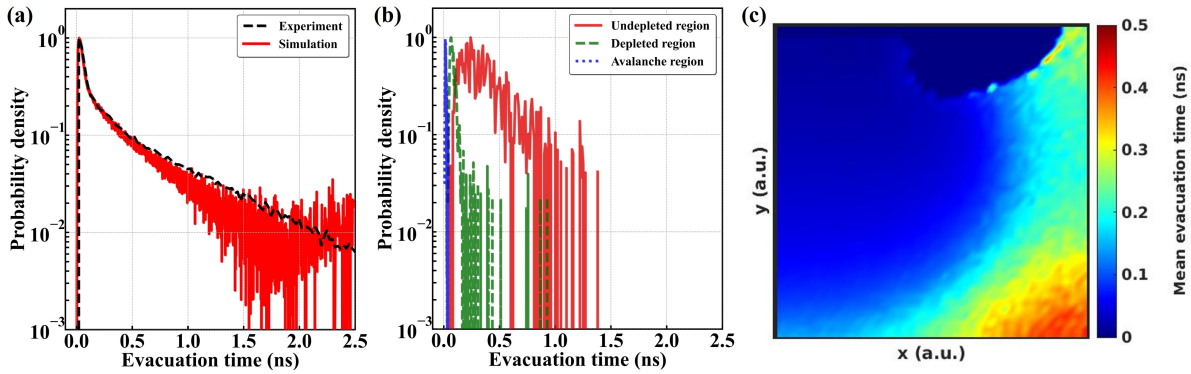


Figure 2.20: (a) Probability density function of the jitter measured and simulated at V_c=4V and T_c=333K. (b) Probability density function of the jitter simulated at V_c=4V and T_c=333K for a starting point from undepleted region (red), depleted region (green) and avalanche region (blue). (c) Evacuation time simulated at V_c=4V and T_c=333K. The timing distribution can be categorized into three distinct characteristics, namely the avalanche region, depletion region, and undepleted region.

Conclusions

The Monte-Carlo method considers both the drift caused by the electric field and the diffusion brought about by Brownian motion, allowing for the capture of individual carrier paths. The multiplication gain factor, which is dependent on the impact-ionization probability, must tend towards infinity for the avalanche breakdown condition to be reached. When examining the studied SPAD architecture, diffusing carriers can activate avalanche-leading regions by overcoming electrostatic potential barriers, which are not captured using the deterministic McIntyre model computation along electric field streamlines. By examining each mesh point within the device, it is possible to differentiate between central and peripheral avalanche contributions and to analyze the statistical behavior of each carrier path. For instance, the jitter shape can be explained from the time it take to evacuate carriers form depleted to undepleted regions.

3 Overview of carrier recombination and generation processes

This section aims to explain the concept of carrier emission and capture, which is used to calculate the Recombination-Generation (R-G) rate. Integrating the Recombination-Generation (R-G) rate with the avalanche breakdown probability across the entire device area yields the Dark Count Rate (DCR), which is further discussed in the following chapter.

3.1 Thermal Shockley–Read–Hall recombination statistics framework

3.1.1 Carrier capture-emission formalism

The discussion begins with the charge density presents within the device, which is essential to understand its impact on the electrostatic potential and carrier flux. The Poisson equation relates the spatial variation of the electrostatic potential to the charge distribution within the semiconductor as follows:

$$\nabla \cdot [\varepsilon_r(x)\nabla\phi(x,t)] = -\frac{\rho(x,t)}{\varepsilon_0} \tag{2.60}$$

where $\varepsilon_r(x)$ and ε_0 are the relative and the vacuum permittivities, $\phi(x,t)$ the electrostatic potential, and $\rho(x,t)$ the charge density, which is given by:

$$\rho(x,t) = q(p(x,t) - n(x,t) + N_D(x) - N_A(x)) - \rho_t(x,t) \tag{2.61}$$

where $n(x,t)$ and $p(x,t)$ are the electron and hole densities, N_D and N_A are the acceptor and donor densities, and $\rho_t(x,t)$ is the charge density contributed by traps. The current continuity equations for electrons and holes write:

$$\frac{\partial n(x,t)}{\partial t} - \frac{1}{q}\frac{\partial J_n}{\partial x} = \Phi_e^n(x,t) - \Phi_c^n(x,t) \tag{2.62}$$

$$\frac{\partial p(x,t)}{\partial t} + \frac{1}{q}\frac{\partial J_p}{\partial x} = \Phi_e^p(x,t) - \Phi_c^p(x,t) \tag{2.63}$$

The rate equation for trap occupation has to be also introduced, which can be expressed as:

$$\frac{\partial p_t(x,t)}{\partial t} = \Phi_c^n(x,t) - \Phi_e^n(x,t) - \Phi_c^p(x,t) + \Phi_e^p(x,t) \tag{2.64}$$

It relates the variation of the trap charge density $\rho_t(x,t)$ in time with the electron (holes, respectively) capture and emission fluxes $\Phi_e^n(x,t)$ and $\Phi_c^n(x,t)$ ($\Phi_e^p(x,t)$ and $\Phi_c^p(x,t)$, respectively) at position x and time t. The charge density $\rho_t(x,t)$ can be calculated by:

$$p_t(x,t) = q\sum_i N_t^i(x,E_t^i) f_t^i(x,t) \tag{2.65}$$

where the sum runs over the traps indexed by the subscript i, and $f_t^i(x,t)$ expresses the occupation function of the trap i, in other words the probability that a trap placed at energy E_t^i

and position x is occupied by an electron at time t. The capture/emission fluxes for electrons are expressed:

$$\begin{cases} \Phi_c^n(x,t) = \sum_i N_t^i(x,E_t^i)(1-f_t^i(x,t))c_n^i(x,t) & (2.66) \\ \Phi_e^n(x,t) = \sum_i N_t^i(x,E_t^i)f_t^i(x,t)e_n^i(x,t) & (2.67) \end{cases}$$

and for holes:

$$\begin{cases} \Phi_c^p(x,t) = \sum_i N_t^i(x,E_t^i)f_t^i(x,t)c_p^i(x,t) & (2.68) \\ \Phi_e^p(x,t) = \sum_i N_t^i(x,E_t^i)(1-f_t^i(x,t))e_p^i(x,t) & (2.69) \end{cases}$$

$c_{n,p}^i(x,t)$ are the capture rates per trap i expressed in s^{-1} of an electron from the conduction band to the trap or a hole from the valence band to the trap. $e_{n,p}^i(x,t)$ are the emission rates per trap i expressed in s^{-1} of an electron from the trap to the conduction band or a hole from the trap to the valence band. The analysis will be simplified by considering only one possible trap at energy E_t at a given position in the following.

The number of available trap states for capture and emission determines the fluxes for these processes as indicated by **eqs. 2.66-2.69**. These traps have the ability to function as Recombination-Generation (R-G) centers, which are depicted in **Fig. 2.21 (a)**. These sites may be crystal lattice dislocations, impurity atoms located interstitially or substitutionally in the crystal lattice [89][90], or surface defects [91][92]. The net recombination rates of electrons and holes from traps can be determined by evaluating the time rate of change of the electron and hole densities as follows:

$$r_n = -\frac{\partial n}{\partial t}|_{R-G} = c_n p_t n - e_n n_t = c_n N_t(1-f_t) - e_n N_t f_t \quad (2.70)$$

$$r_p = -\frac{\partial p}{\partial t}|_{R-G} = c_p n_t p - e_p p_t = c_p N_t f_t - e_p N_t(1-f_t) \quad (2.71)$$

where n_t and p_t represent the number of R-G centers filled with electrons and emptied. It is important to note that the number of emptied traps p_t is distinct from the trap charge density ρ_t. The negative sign in front of the time derivative indicates that the electron and hole concentrations are decreasing due to capture processes. The first term represents the capture rate of electrons (holes) by the R-G centers. The second term represents the emission rate of electrons (holes) from the R-G centers. The equations can also be expressed in terms of the trap occupancy f_t, which is the fraction of R-G centers N_t that are occupied by electrons (holes).

Assuming the detailed balance under thermal equilibrium, which is to say the self-balance of processes $r_n = 0$ and $r_p = 0$ in **eqs. 2.70-2.71** (as shown in **Fig. 2.21 (b)**), the following relationship is obtained:

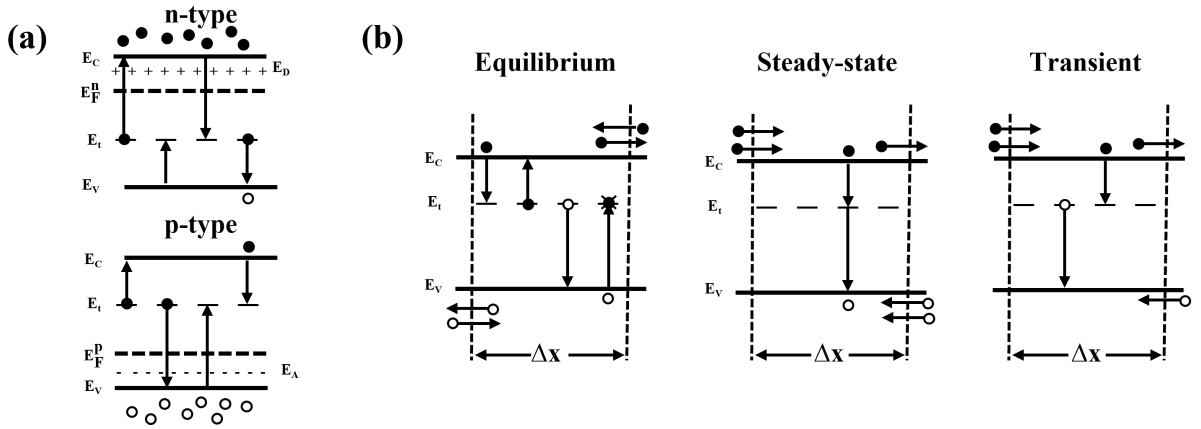


Figure 2.21: (a) The processes of carrier generation and recombination through traps in a n- and p-type semiconductor with electron capture, electron emission, hole capture, hole emission. $E_F^{n,p}$ relates to the non-equilibrium carrier concentration whereas E_F to the equilibrium one. (b) Under equilibrium each process counterbalance its opposite (stationary). Under steady-state carrier density is held constant by balancing effect of distinctly different process (stationary). In transient regime, carrier density is not conserved (not stationary) and no process do not counterbalance.

$$e_n^0 = \frac{c_n^0 p_t^0 n_0}{n_t^0} = c_n^0 n_1 = c_n n_0 \frac{(1-f_t)}{f_t} \tag{2.72}$$

$$e_p^0 = \frac{c_p^0 n_t^0 p_0}{p_t^0} = c_p^0 p_1 = c_p p_0 \frac{f_t}{(1-f_t)} \tag{2.73}$$

with the density of electrons in the conduction band and of holes in the valence band when the Fermi level E_F falls at E_t given by:

$$n_1 = \frac{p_t^0 n_0}{n_t^0} = (\frac{N_t}{n_t^0} - 1) n_0 \tag{2.74}$$

$$p_1 = \frac{n_t^0 p_0}{p_t^0} = (\frac{N_t}{p_t^0} - 1) p_0 \tag{2.75}$$

The subscript/superscript 0 added to point out that it is a relation valid at equilibrium. The occupancy probability by electrons is equal to:

$$f_t = \frac{n_t^0}{N_t} = [1 - g_t exp(\frac{(E_F - E_t)}{k_B T})]^{-1} \tag{2.76}$$

and by holes to:

$$1 - f_t = \frac{p_t^0}{N_t} = [1 + g_t exp(\frac{(E_t - E_F)}{k_B T})]^{-1} \tag{2.77}$$

where E_t is the trap energy with $g_t = 1$ the degeneracy representing the number of possible configurations a defect can assume on a single site. Under non-equilibrium, the emission rates of electrons and holes can be approximated as follows:

$$e_n \approx e_n^0 \approx c_n^0 n_1 \approx c_n n_1 \tag{2.78}$$

$$e_p \approx e_p^0 \approx c_p^0 p_1 \approx c_p p_1 \tag{2.79}$$

By incorporating **eq. 2.78** and **eq. 2.79** into **eq. 2.70** and **eq. 2.71**, the recombination rates of electrons and holes can be expressed as:

$$r_n = c_n(p_t n - n_1 n_t) \tag{2.80}$$

$$r_p = c_p(n_t p - p_1 p_t) \tag{2.81}$$

The first term represents the rate at which electrons (holes) are captured by R-G centers, while the represents the rate at which trapped electrons (holes) recombine with free holes (electrons).

A reminder of the difference between non-degenerate and degenerate semiconductors is necessary before proceeding, specifically with regards to the definition of carrier concentration. Non-degenerate semiconductors, characterized by a moderate level of doping below the intrinsic dopant-free concentration, exhibit distinct properties due to the well-separated dopant atoms and their negligible interactions. In such semiconductors, the dopant atoms occupy discrete energy levels that are typically located either below the conduction band edge or on top of the valence band edge. In contrast, degenerate semiconductors are heavily doped and display significant interaction between dopant atoms. This interaction leads to the formation of donor or acceptor bands, which differ from the discrete energy levels observed in non-degenerate semiconductors. Unlike non-degenerate semiconductors, where the Fermi level is typically located within the energy gaps, the Fermi level is often found near the conduction or valence band edges for degenerate semiconductors, typically within a range of about $3k_BT$ from the band edges. Due to the presence of high doping levels and the position of the Fermi level in the conduction or valence bands, degenerate semiconductors exhibit metallic characteristics at low temperatures, while behaving more like conventional semiconductors at higher temperatures. For a non-degenerate semiconductor:

$$n_0 = n_i exp(\frac{(E_F - E_i)}{k_B T}) \tag{2.82}$$

$$p_0 = n_i exp(\frac{(E_i - E_F)}{k_B T}) \tag{2.83}$$

And for a degenerate semiconductor:

$$n_0 = N_c \Im_{1/2}(\eta_C = \frac{(E_F - E_i)}{k_B T}) \tag{2.84}$$

$$p_0 = N_v \Im_{1/2}(\eta_V = \frac{(E_i - E_F)}{k_B T}) \tag{2.85}$$

$\eta_{C,V}$ measures the location of E_F with respect to the band edges. $\Im_{1/2}(\eta) = \frac{2}{\sqrt{\pi}} F_{1/2}(\eta)$ with $F_{1/2}(\eta)$ the Fermi-Dirac function. $\eta < -3$ for a nondegenerate semiconductor ($\Im_{1/2}(\eta) \approx exp(\eta)$) and $\eta \gg 0$ for a degenerate one ($\Im_{1/2}(\eta) \to \frac{4\eta_F^{3/2}}{3\sqrt{\pi}}$). E_i is the intrinsic Fermi level, and E_v and E_c are valence and conduction band energy. The thermal dependence of the intrinsic density n_i can be expressed as:

$$n_i = \sqrt{N_c N_v} exp(\frac{-E_g}{2k_B T}) \tag{2.86}$$

where the effective densities of states in the conduction band N_c and in the valence band N_v are:

$$N_{c,v} = 2(\frac{2\pi m^*_{c,v} k_B T}{h^2})^{\frac{3}{2}} \tag{2.87}$$

with $m^*_{c,v}$ the effective mass of the density of states in the conduction band and in the valence band. m^*_c=1.08/m_0 and m^*_v=0.56/m_0 in silicon with m_0 the free electron rest mass [51]. The thermal dependence of the band gap follows can be described as [93]:

$$E_g = E_g^0 - \frac{\alpha_g \cdot T^2}{\beta_g + T} \tag{2.88}$$

where E_g^0 is the bandgap at 0K and α_g and β_g some experimental parameters depending on the semiconductor material. The thermal dependence of the Fermi level for an intrinsic semiconductor E_i is:

$$E_i = \frac{E_g}{2} + \frac{3k_B T}{4} ln(\frac{m^*_p}{m^*_n}) \tag{2.89}$$

Assuming that the occupancy of the trap is given by the Boltzmann distribution, the density of electrons in the conduction band and the density of holes in the valence band when the Fermi level E_F is located at E_t are given by:

$$n_1 = n_i(\pm g_t) exp(\frac{(E_t - E_i)}{k_B T}) \tag{2.90}$$

$$p_1 = n_i(\pm g_t) exp(\frac{(E_i - E_t)}{k_B T}) \tag{2.91}$$

Considering **eqs. 2.70-2.71**, the condition of steady-state $r_n = r_p$ can be used to obtain the following expression:

$$\frac{\partial n_t}{\partial t} = -\frac{\partial n}{\partial t}|_{R-G} + \frac{\partial p}{\partial t}|_{R-G} = r_n - r_p = 0 \tag{2.92}$$

Considering **eqs. 2.80-2.81**, the R-G centers filled with electrons is given by:

$$n_t = \frac{c_n N_t n + c_p N_t p_1}{c_n(n+n_1) + c_p(p+p_1)} = \frac{f_t}{N_t} \tag{2.93}$$

The net steady-state R-G rate for a nondegenerate semiconductor can be expressed as:

$$R = \frac{np - n_i^2}{\frac{1}{c_p N_t}(n+n_1) + \frac{1}{c_n N_t}(p+p_1)} \tag{2.94}$$

with n_1 and p_1 given by **eqs. 2.90-2.91**. And for a degenerate semiconductor:

$$R = \frac{np - n_1 p_1}{\frac{1}{c_p N_t}(n+n_1) + \frac{1}{c_n N_t}(p+p_1)} \tag{2.95}$$

with:

$$n_1 = ((\frac{1}{1 - g_t exp(\frac{(E_F - E_t)}{k_B T}} - 1)^{-1}) N_c \Im_{1/2}((E_F - E_c)/k_B T) = n_0 \frac{(1-f_t)}{f_t} \tag{2.96}$$

$$p_1 = ((\frac{1}{1 + g_t exp(\frac{(E_F - E_t)}{k_B T}} - 1)^{-1}) N_v \Im_{1/2}((E_v - E_F)/k_B T) = p_0 \frac{f_t}{(1-f_t)} \tag{2.97}$$

The equations above show that the net R-G rate is affected by the presence of traps, which can act as recombination centers and reduce the efficiency of the semiconductor. The dopant concentration present in the region of interest within the device indicates that the semiconductor material can be treated as nondegenerate, and this assumption will be made in the following.

Another important metric that should be considered for the following analysis is the average time it takes for a free carrier to recombine with another carrier, which is defined as:

$$\tau_{n,p}^{-1} = (c_{n,p} N_t)^{-1} \tag{2.98}$$

where $\tau_{n,p}$ is the minority carrier lifetime for electrons and holes injected in a p-type and n-type specimen respectively. The electric field affects the energy levels of the carriers and traps, which alters the probability of a carrier being captured or emitted by a trap, making $\tau_{n,p}$ an electric-field dependent factor, as explained in further detail below.

Conclusions

By analyzing the charge density within the device, one can gain insight into the carrier flux and electrostatic potential. By taking into account carrier capture/emission and the statistics of occupation function for bands and traps, the thermal Shockley-Read-Hall (SRH) model explains the Recombination-Generation (R-G) of charge carriers (electrons and holes) in a semiconductor material due to filled and emptied defects. This model is applicable in both non-degenerate and degenerate semiconductors, allowing for a comprehensive understanding of the steady-state process. The minority carrier lifetime for electrons and holes injected in a p-type and n-type can be derived to comprehend device's efficiency.

3.1.2 Doping-limiting minority carrier lifetime

Substitutional dopant impurities play a significant role in determining the minority carrier lifetime in a semiconductor. Higher doping concentrations generally lead to shorter lifetimes. This is because an increase in doping concentration results in an increased density of R-G centers, such as traps due to implantation damages or impurity states, which provide more opportunities for minority carriers to recombine. The recombination centers due to the process of doping implantation can be approximated by:

$$N_t = N_0(1 + \frac{N_D + N_A}{N_0}) \tag{2.99}$$

On the other hand, dopant impurities can also influence the dominant recombination mechanisms at stake. For example, in heavily doped semiconductors, Auger recombination becomes more prominent. Auger recombination involves the transfer of excess energy from one carrier to another, resulting in the recombination of two minority carriers.

This has been evidenced by [94][95]. The measurements obtained from these studies can be accurately described using the empirical relation of Scharfetter, which has been enhanced by taking into account temperature and electric-field effects. This relationship is illustrated in is **Fig. 2.22**, and is given by:

$$\begin{aligned} \tau_{n,p} &= \left(\frac{T}{300K}\right)^{-\frac{3}{2}} \cdot [\tau_{min} + \frac{\tau_{max}-\tau_{min}}{1+(\frac{N_{A,0}+N_{D,0}}{N_{ref}})^{\gamma}}] \cdot (1+\Gamma_{TAT}(F))^{-1} \\ &= \tau_0 \cdot \left(\frac{T}{300K}\right)^{-\frac{3}{2}} \cdot [1+(\frac{N_A+N_D}{N_0})]^{-1} \cdot (1+\Gamma_{TAT}(F))^{-1} \end{aligned} \tag{2.100}$$

where $\Gamma_{TAT}(F)$ is an electric field enhancement parameter which models the trap assisted tunneling [96] and will be described in detail below. The Scharfetter model parameters specific to silicon are reported in **Table 2.5**. The definition of dopant-limiting minority carrier lifetime show good agreement at high doping impurity density. However it fails to describe the plateau at low concentration which can be due to the surface defects.

A typical generation rate map simulated from the dopant-limiting lifetime SRH model is

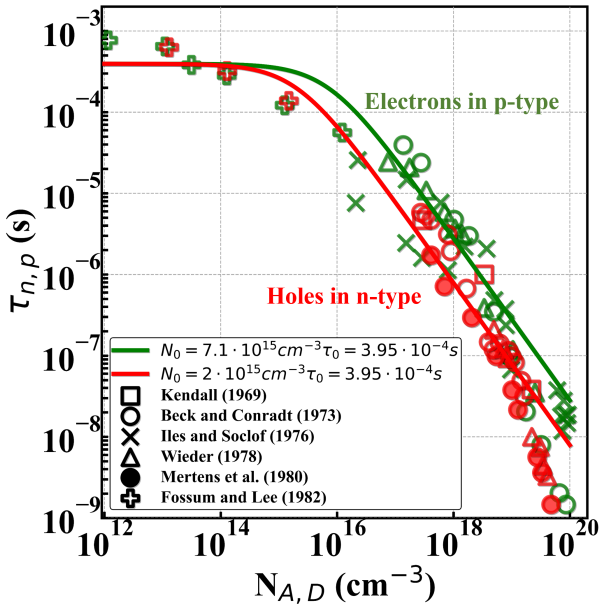


Figure 2.22: Minority carrier recombination lifetime $\tau_{n,p}$ as a function of dopant impurities $N_{A,D}$ (either acceptors or donors) for holes and electrons in n-type and p-type semiconductors respectively obtained from experimental data [94][95]. The results have been fitted with **eq. 2.100**, which exhibits good agreement at high dopant concentration.

Table 2.5: Parameters of the Scharfetter model **eq. 2.100** for silicon. The model is activated by setting *Recombination(SRH(DopingDep TempDependence Tunneling(Hurkx)))* in physics section and all the parameters are accessible in *Scharfetter* parameter set.

Symbol	Parameter	Electrons	Holes	Unit
τ_{min}	taumin	0	0	s
τ_{max}	taumax	10^{-5}	$3 \cdot 10^{-6}$	s
τ_0	tau0	10^{-5}	$3 \cdot 10^{-6}$	s
N_{ref}	Nref	10^{16}	$3 \cdot 10^{16}$	cm^{-3}
γ	gamma	1	1	1

shown in **Fig. 2.23**. Carriers are generated within the depletion region, while recombination predominantly occurs outside of it.

3.1.3 Defect-limiting minority carrier lifetime

To overcome the dopant representation limitation, the defect description is now introduced. In time t an electron will travel a distance $v_{th}^n \cdot t$ and will pass trough a volume $v_{th}^n \cdot t \cdot A$ as represented in **Fig. 2.24**. The thermal velocity $\mathrm{v}_{\mathrm{th}}^{n,p}$ for electrons and holes is equal to:

$$v_{th}^{n,p} = \sqrt{\frac{3k_BT}{m_{n,p}^*}} \tag{2.101}$$

In this volume, there will be p_t empty R-G centers per cm^3 or a total number $v_{th}^n \cdot t \cdot A \cdot n \cdot p_t$ empty R-G centers. The probability of an electron being captured in this volume can be

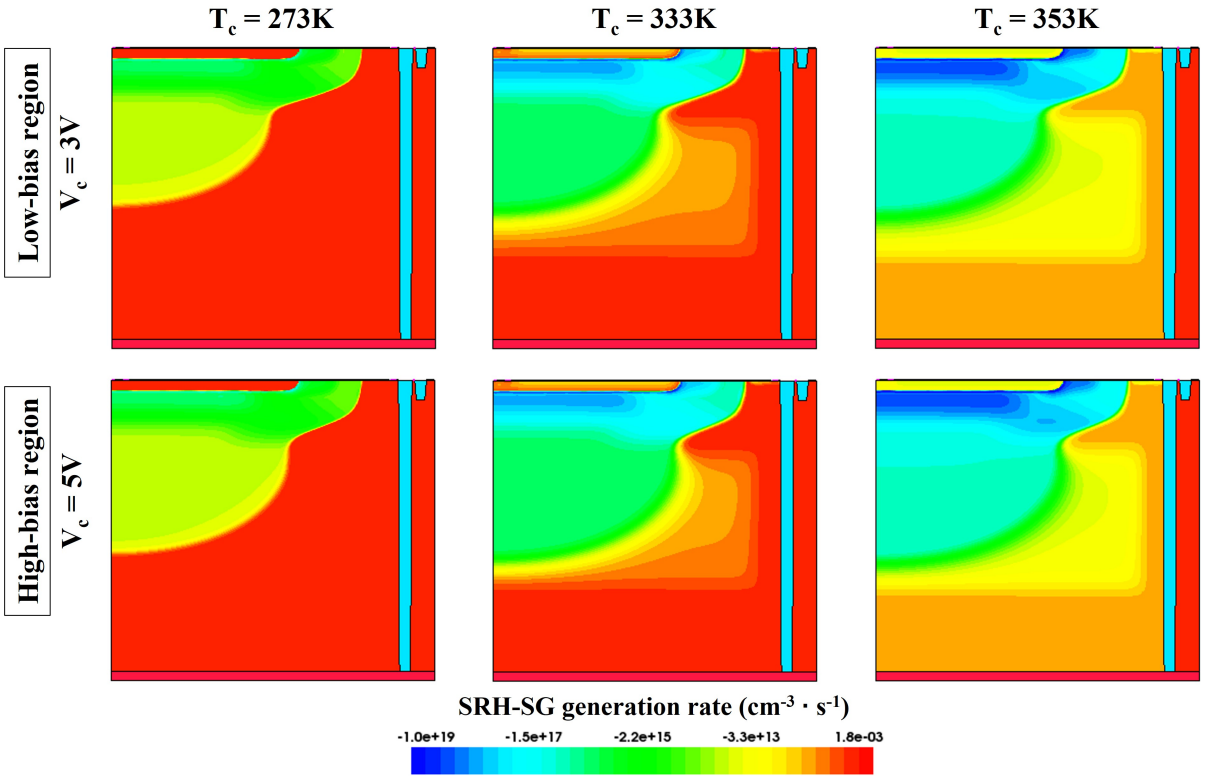


Figure 2.23: Dopant-limiting lifetime SRH generation rate map simulated at T_c=273K, 333K and 353K, and V_c=3V and 5V. As temperature increases, the rate of carrier generation in the depletion region decreases.

determined by conceptually moving the centers to a single-plane in the middle of the volume. The total area blocked by the empty R-G centers is $v_{th}^n \cdot t \cdot A \cdot p_t \cdot \sigma_n$, with σ_n the capture cross section, and the fraction giving rise to capture is $\frac{v_{th}^n \cdot t \cdot A \cdot p_t \cdot \sigma_n}{A} = v_{th}^n \cdot t \cdot p_t \cdot \sigma_{p,n}$ which is the probability of electron to being captured in the volume and the probability rate is given $\frac{v_{th}^n \cdot t \cdot p_t \cdot \sigma_n}{t} = v_{th}^n \cdot p_t \cdot \sigma_n$.

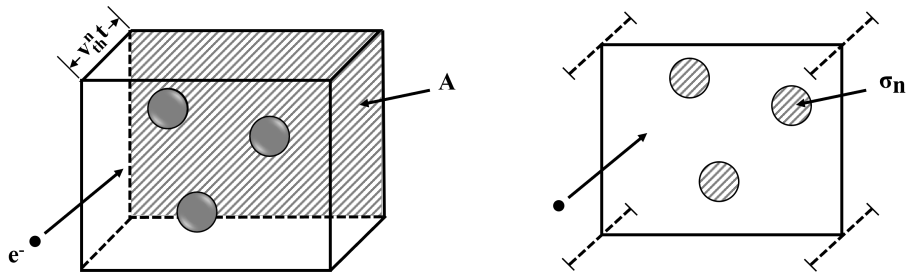


Figure 2.24: The volume is assumed to contain a certain number of empty R-G centers per cm^3 (spheres on the left-hand side). The probability of electron capture is determined by calculating the area blocked by the empty R-G centers (circles on the right-hand side), denoted as the capture cross section σ_n, and the fraction of this area that gives rise to electron capture.

The rate change of electron density captured by R-G centers can rewritten in terms of the capture cross section $\sigma_{n,p}$ in contrast to the dopant-limiting lifetime Scharfetter model, which is

expressed as:

$$\frac{\partial n}{\partial t}|_{R-G} = -c_n p_t n = -\sigma_n v_{th} p_t n \tag{2.102}$$

Similarly for holes:

$$\frac{\partial p}{\partial t}|_{R-G} = -c_p n_t p = -\sigma_p v_{th} n_t p \tag{2.103}$$

where the capture rate is given by:

$$c_{n,p} = \sigma_{n,p} v_{th}^{n,p} \tag{2.104}$$

Substituting this expression into **eq. 2.98** results in a trap-limiting lifetime, which is expressed as:

$$\tau_{n,p}^{-1} = N_t \sigma_{n,p} v_{\mathrm{th}} \tag{2.105}$$

In summary, the capture cross section $\sigma_{n,p}$ describes the probability interaction between electrons or holes and trap states. It represents the effective area given in units of square centimeters for the trap to capture an approaching carrier as shown in **Fig. 2.24**. A larger capture cross section indicates a higher likelihood of capturing carriers, leading to a faster recombination rate or reduced minority carrier lifetime as illustrated by **eq. 2.105**. This definition of minority carrier lifetime allows a flexible investigation of possible generation sites within the device especially at the interface which is not captured by the empirical Sharfetter model. The sdevice tool [47] supports several defect configurations reported in **Table 2.6**.

A typical generation rate map simulated from the defect-limiting lifetime SRH model is shown in **Fig. 2.25**. The defect distribution has been randomized in space by setting *((eNeutral Tunneling(Hurkx) Conc=... Level Randomize = 1 EnergyMid=... fromMidBandGap SpatialShape=Gaussian SpaceMid=... SpaceSig=....))* in physics section. The randomization differs every time the command file is executed. The specified number is used as the seed for the random number generator. As mentioned earlier, the trap definition enables investigation throughout the device, revealing the impact of interface defects on the carrier generation rate.

Conclusions

The minority carrier lifetime can be defined in terms of the dopant concentration within the volume of the material. This definition allows for a focus on the impact of dopants primarily within the depleted volume. Alternatively, the capture cross section can be used to define the minority carrier lifetime, which represents the probability interaction of a carrier being captured by a trap state. This definition enables the investigation of the role of interfaces in carrier generation beyond the depletion region in the studied SPAD.

Defect parameter	Setting
Types	*Donor/hNeutral*: uncharged when unoccupied and becomes positively charged capturing a hole *Acceptor/eNeutral*: uncharged when unoccupied and becomes negatively charged capturing an electron *Fixed Charge*: fully occupied by either electron or hole
Cross section	*eXection* - *hXsection* (cm^{-2})
Electric-field enhancement	*Tunneling(Hurkx)* - *PooleFrenkel(Electron)* - *PooleFrenkel(Hole)*
Concentration	*Conc* (cm^{-3} - cm^{-2} or $eV^{-1}cm^{-3}$ - $eV^{-1}cm^{-2}$ if different from Level)
Energy distribution	*Level* (randomize) - Uniform - Exponential - Gaussian (by default 13 discrete energy levels)
Energy distribution center	*fromCondBand* - *fromValBand* - *fromMidBandGap*
Energy mean and sigma	*EnergyMid* = ± eV and *EnergySig* = ± eV (if different from Level)
Spatial distribution	*SpatialShape* = *Gaussian* - *Uniform*
Spatial mean and sigma	*SpaceMid* = (<x>,<y>,<z>) and *SpaceSig* = (σ_x,σ_y,σ_z)

Table 2.6: Trap parameters available in the sdevice tool [47]. The model is activated by setting *Traps((eNeutral Tunneling(Hurkx) PooleFrenkel(Electron) PooleFrenkel(Hole) Conc=... Level eXsection=... hXsection=... EnergyMid=... fromMidBandGap SpaceMid=... SpaceSig=...)* in physics section and all the parameters are accessible in *Traps*, *HurkxTrapAssistedTunneling* and *PooleFrenkel* parameter sets.

3.1.4 Thermal activation energy concept

The thermal dependence of the SRH R-G rate may be influenced by both the trap energy level, as indicated by **eq. 2.94** including **eqs. 2.90-2.91**, and the minority carrier lifetime, as illustrated by **eq. 2.100**. To comprehend the thermal dependence of the SRH R-G rate, it is necessary to focus on the trap energy level and discuss the concept of thermal activation energy that arises from this definition.

To emphasize the thermal dependence of the SRH generation rate, the focus will be on the depletion region. The principle of carrier depletion has been previously introduced in the context of p-n junction in **Section 1.2**. It is formally defined to be a semiconductor volume where $n \cdot p \ll n_i^2$ in reverse bias, which results in the generation of minority carriers. By incorporating the aforementioned definition into the R-G rate equation **eq. 2.94**, along with the expressions for n_1 and p_1 as stated in equations **eq. 2.90** and **eq. 2.91** respectively, and setting g_t to 1, the generation rate can be expressed as follows, taking into account the trap definition for the capture rate **eq. 2.104**:

$$R = \frac{n_i}{\frac{1}{N_t \sigma_p v_{th}^p} exp(\frac{(E_t - E_i)}{k_B T}) + \frac{1}{N_t \sigma_n v_{th}^n} exp(\frac{(E_i - E_t)}{k_B T})} \tag{2.106}$$

which can be rewritten as:

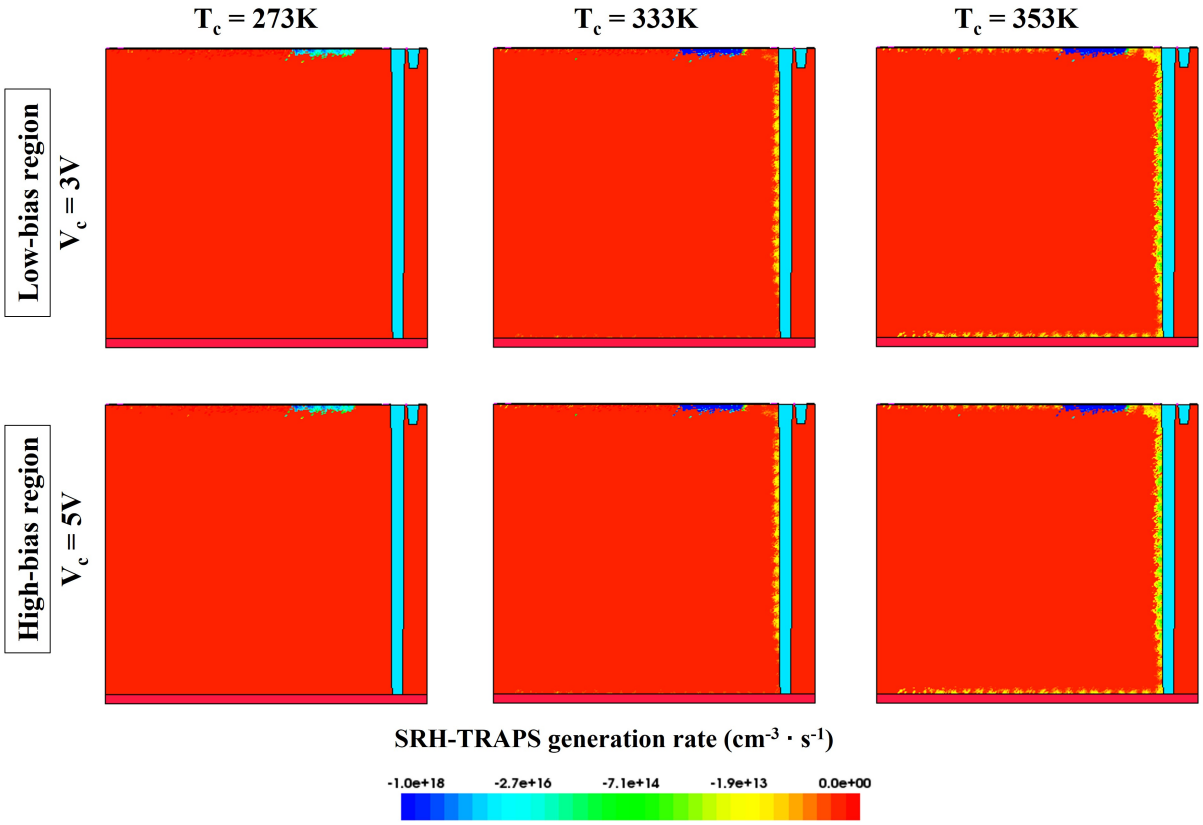


Figure 2.25: Defect-limiting lifetime SRH generation rate map simulated at T_c=273K, 333K and 353K, and V_c=3V and 5V. As temperature increases, the rate of carrier generation decreases.

$$R = \frac{n_i N_t \sigma_n \sigma_p v_{th}^n v_{th}^p}{\sigma_n v_{th}^n exp(\frac{(E_t - E_i)}{k_B T}) + \sigma_p v_{th}^p exp(\frac{(E_i - E_t)}{k_B T})} \tag{2.107}$$

The previous expression shows that the generation rate of an SRH defect located in the central region of the depletion region is maximum when the energy state is located in the middle of the bandgap $E_t = E_i$. By numerical calculus utilizing **eq. 2.101**, one can see that v_{th}^n is approximately 60% higher than v_{th}^p along temperature range considering heavy holes in silicon [51]. $v_{th} \approx v_{th}^n \approx v_{th}^p$ is assumed for the purpose of the demonstration. In this case, the generation rate is then given by:

$$R = \frac{N_t \sigma_n \sigma_p v_{th} n_i}{\sigma_n + \sigma_p} \tag{2.108}$$

Including n_i and v_{th} given in **eq. 2.86** and **eq. 2.101**, the generation rate can be expressed as:

$$R(T) = \frac{N_t \sigma_n \sigma_p}{\sigma_n + \sigma_p} \times \sqrt{\frac{3kT}{m}} \times \sqrt{N_C(T) N_V(T)} \times \exp\left(-\frac{E_g(T)}{2k_B T}\right) \tag{2.109}$$

Parentheses have been used to emphasize the thermal dependence by including the tem-

perature. By including $N_{c,v}$ and E_g from **eq. 2.87** and **eq. 2.88**, the generation rate can be expressed using the following equation, where C_1 is a composite constant:

$$R(T) = C_1 \times T^2 \times \exp\left(-\frac{E_g(T)}{2k_BT}\right) \tag{2.110}$$

which can be rewritten taking the logarithm as:

$$\ln(R(T)) = \ln(C_1) + 2\ln(T) - \frac{E_g(T)}{2k_BT} \tag{2.111}$$

In this case, the generation rate can be formulated as:

$$R(T) = C_2 \times \exp\left(-\frac{E_a(T)}{k_BT}\right) \tag{2.112}$$

where C_2 is an other composite constant and $E_a(T)$ is the activation energy of the generation-recombination rate, which is proportional to the derivative of the natural logarithm of R with respect to $\frac{1}{T}$ [97]:

$$E_a(T) = \frac{d\ln(R(T))}{d\left(\frac{1}{k_BT}\right)} = 2k_BT + \frac{E_g(T)}{2} - \frac{T}{2}\frac{dE_g(T)}{dT}, \quad E_t = E_i \tag{2.113}$$

$E_a(T)$ is the minimum amount of energy required for the generation of a carrier in the material. The first term represents the thermal energy available per charge carrier in the system, the second one is the height of the highest energy barrier that the thermally generated electron needs to overcome, and the third one the variation of the bandgap height with temperature. Their thermal dependencies are depicted in **Fig. 2.26 (a)**, which indicates that the bandgap energy level dominates over the other contribution.

Let now consider the case of a defect with an energy level that is away from the center of the bandgap $|E_t - E_i| > 3k_BT$. In this case, the generation rate can be formulated as follows:

$$R(T) = \frac{N_t\sigma_p v_{th} n_i}{\exp\left(\frac{E_t - E_i}{k_BT}\right)} \tag{2.114}$$

$$R(T) = C_1' \times T^2 \times \exp\left(-\frac{E_g(T)}{2k_BT} - \frac{(E_t - E_i(T))}{k_BT}\right) \tag{2.115}$$

And for $|E_i - E_t| > 3k_BT$:

$$R(T) = \frac{N_t\sigma_n v_{th} n_i}{\exp\left(\frac{E_i - E_t}{k_BT}\right)} \tag{2.116}$$

$$R(T) = C_2' \times T^2 \times \exp\left(-\frac{E_g(T)}{2k_BT} - \frac{(E_i(T) - E_t)}{k_BT}\right) \tag{2.117}$$

By following the same approach used for the activation energy as before, one can derive the subsequent expression:

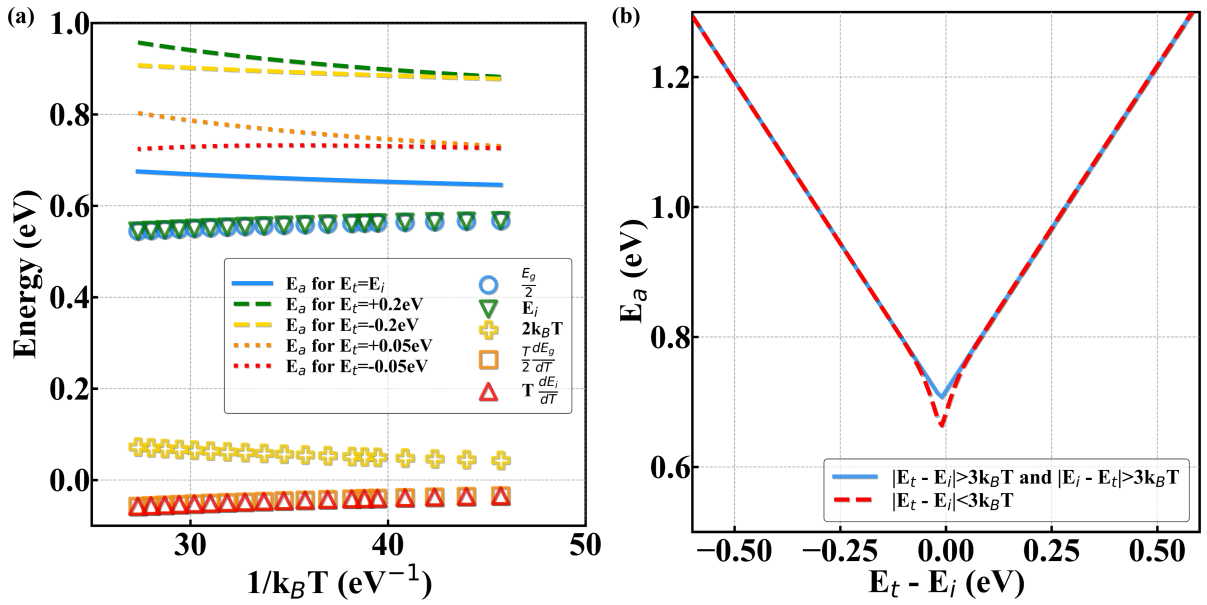


Figure 2.26: (a) Energy of the carrier generation rate as a function of $1/k_BT_c$ for $E_t = E_i$ (**eq. 2.106**), for $|E_t - E_i| > 3k_BT$ or $|E_i - E_t| > 3k_BT$ (**eq. 2.113**) and $|E_t - E_i| < 3k_BT$ (**eq. 2.118**). The different components of these equation are shown independently in symbols. (b) Thermal activation energy as a function of $E_t - E_i$ for $|E_t - E_i| > 3k_BT$ or $|E_i - E_t| > 3k_BT$ (solid lines) and $|E_t - E_i| < 3k_BT$ (dashed lines). The thermal dependence of the carrier generation rate is primarily influenced by the bandgap and intrinsic energy levels. If $|E_t - E_i| < 3k_BT$, the minimum amount of energy required for the generation of a carrier in the material is reduced.

$$E_a = 2k_BT + \frac{E_g(T)}{2} - \frac{T}{2}\frac{dE_g(T)}{dT} + |E_t - E_i| - T\frac{dE_i(T)}{dT}, \quad \begin{matrix} |E_t - E_i| > 3k_BT \\ |E_i - E_t| > 3k_BT \end{matrix} \tag{2.118}$$

Referring to the previous equation given for **eq. 2.113** with $E_t = E_i$, two additional factors are included. These factors account for the difference in energy between the trap and intrinsic energy levels and the change in intrinsic energy level height with temperature. The thermal dependence of these factors is shown in **Fig. 2.26 (a)**, which demonstrates that the bandgap and intrinsic energy levels have a greater impact than the other factors. Since the activation energy of the generation rate is governed by the temperature variation of the limiting process, **eq. 2.106** remains equivalent regardless of the relative position of E_t.

In the case where $E_t \gg E_i$, the generation of a carrier is limited by the electron's transition from the valence band to the energy level of the defect, as it represents the highest energy barrier. Conversely, in the case where $E_t \ll E_i$, it is the emission of the electron from the defect to the conduction band that limits the generation rate. Furthermore, since the generation rate decreases exponentially with $|E_t - E_i|$, defects with $|E_t - E_i| > 3k_BT$ are likely to have much lower generation rates compared to defects with $E_t \approx E_i$. However, significant capture cross-sections $\sigma_{n,p}$ can compensate for this effect.

For defects with energy levels close to the middle of the bandgap $|E_t - E_i| < 3k_BT$, **eq. 2.107** cannot be directly simplified as it depends on both σ_n and σ_p values. For instance, $\sigma_n = 10^{-17}cm^2$ and $\sigma_p = 10^{-16}cm^2$ for vacancy-arsenic and vacancy-boron respectively [89][90]. When the capture cross section for both holes and electrons is assumed to be the same ($\sigma \approx \sigma_n \approx \sigma_p$), the following expression can be obtained:

$$R(T) \approx \frac{\sigma v_{\text{th}} n_i}{2\cosh\left(\frac{E_t - E_i}{kT}\right)} \approx \frac{\sigma v_{\text{th}} n_i}{2} sech(\left(\frac{E_t - E_i}{kT}\right) \tag{2.119}$$

$$\begin{aligned} E_a(T) \approx & 2k_BT + \frac{E_g(T)}{2} - \frac{T}{2}\frac{dE_g(T)}{dT} \\ & + [|E_t - E_i| - T\frac{dE_i(T)}{dT}] \tanh\left(\frac{|E_t - E_i|}{k_BT}\right), |E_t - E_i| < 3k_BT \end{aligned} \tag{2.120}$$

In the context of **eq. 2.113** and **eq. 2.118**, the difference in energy between the trap and intrinsic energy levels and the change in intrinsic energy level height with temperature are multiplied by the hyperbolic tangent of the ratio between the difference in energy and the thermal energy of the system. **Fig. 2.26 (b)** shows the cases where $|E_t - E_i| > 3k_BT$ or $|E_i - E_t| > 3k_BT$, as well as the case where $|E_t - E_i| < 3k_BT$. When $|E_t - E_i| < 3k_BT$, the hyperbolic tangent factor reduces the minimum amount of energy required for the generation of a carrier in the material.

In the case where $\sigma_n \neq \sigma_p$, e.g. $\sigma_n \ll \sigma_p$, instead of being limited by the process with the higher energy barrier (hole emission to the valence band in this case), the generation rate is limited by the process with the lower capture cross section (electron emission to the conduction band in this case).

Conclusions

The concept of thermal activation energy has been introduced to explain the relationship between trap energy and carrier generation from traps. The thermal dependence of the carrier generation rate is primarily influenced by the bandgap and intrinsic energy levels, as well as on the energy level of the trap relative to the intrinsic energy level. When the energy state of the trap is located in the middle of the bandgap or close to it, the minimum energy required for carrier generation is achieved, resulting in the maximum carrier generation rate. The thermal activation energy is not only limited by the trap energy level but also by the capture cross section.

3.2 Electric-field enhancement of minority carrier lifetime

As previously mentioned, the electric field can affect the minority carrier lifetime, which in turn depends on the capture cross section $\sigma_{n,p}$. This can modify the thermal dependence of the carrier generation rate, further decreasing the minimum energy required for carrier generation in the material. This effect is discussed in more detail in this section.

3.2.1 Poole-Frenkel effect

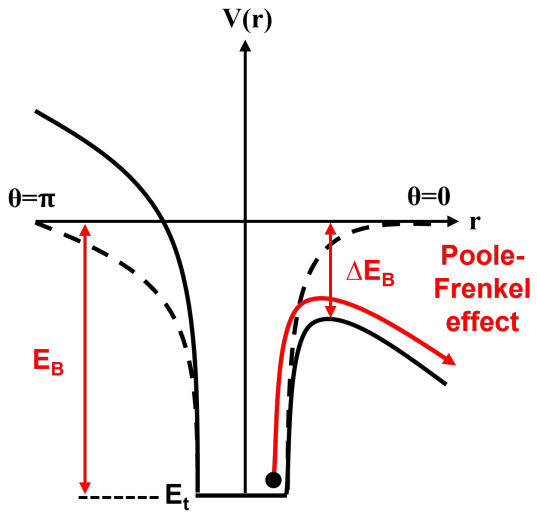


Figure 2.27: Illustration of the Poole-Frenkel effect inducing a lowering of the potential barrier ΔE_B. The dashed line denotes the potential well of the trap without Coulomb interaction and the solid line with Coulomb interaction.

The Poole-Frenkel effect induces a lowering of the potential barrier ΔE_B due to the presence of the electric field which causes a distortion of the energy levels as illustrated on **Fig. 2.27**. The emission is then similar to that of the SRH emission. The emission rate including this effect is written as [98][99]:

$$e_n^{PF} = \sigma_n v_{th} N_c exp(-\frac{E_B - \Delta E_B}{k_B T}) = e_n^{F=0} exp(\frac{\Delta E_B}{k_B T}) \tag{2.121}$$

where $e_n^{F=0}$ is the emission probability in the absence of an electric field. Therefore $\frac{e_n^{PF}}{e_n^{F=0}} = exp(\frac{\Delta E_B}{k_B T})$ is the electric field enhancement of the emission rate. Assuming a Coulombic potential:

$$V(r) = \frac{-q^2}{4\pi\varepsilon_r\varepsilon_0 r} - qF_r cos(\theta) \tag{2.122}$$

The first term on the right-hand side of the equation represents the Coulomb potential energy, which is the electrostatic potential energy between two point charges. The second term on the right-hand side of the equation represents the potential energy due to an external electric field, F_r, acting on the point charge q at an angle θ with respect to the radial direction r. The cosine of the angle $cos(\theta)$ is included in the equation to account for the component of the electric field in the radial direction r. The position at which the potential reaches its maximum value r_{max} is determined by the point at which its derivative equals zero, which can be expressed as:

$$r_{max} = (\frac{q}{4\pi\varepsilon_r\varepsilon_0 F cos(\theta)})^{1/2}, \quad 0 < \theta < \pi/2 \tag{2.123}$$

Therefore, the reduction in potential barrier can be expressed as:

$$\Delta E_B(\theta) = V(r_{max}) = -q(\frac{qFcos(\theta)}{\pi\varepsilon_r\varepsilon_0})^{1/2} \tag{2.124}$$

The electric field enhancement of the emission rate is integrated over the angle θ and the azimuthal angle ϕ as follows:

$$\frac{e_n^{PF}}{e_n^{F=0}} = \frac{1}{4}\int_0^{\pi} exp(\frac{\Delta E_B(\theta)}{k_BT})sin\theta d\theta \int_0^{2\pi} d\phi \tag{2.125}$$

$$\frac{e_n^{PF}}{e_n^{F=0}} = \frac{1}{4}[\int_0^{\pi/2} exp(\frac{\Delta E_B(\theta)}{k_BT}sin\theta d\theta \int_0^{2\pi} d\phi + \int_{\pi/2}^{\pi} sin\theta d\theta \int_0^{2\pi} d\phi] \tag{2.126}$$

$$\frac{e_n^{PF}}{e_n^{F=0}} = \frac{1}{2} + \frac{1 + (\frac{\Delta E_B(\theta)}{k_BT} - 1)exp(\frac{\Delta E_B(\theta)}{k_BT})}{(\frac{\Delta E_B(\theta)}{k_BT})^2} \tag{2.127}$$

The electric field enhancement factor Γ_{PF} for the emission of electrons and holes can be defined as:

$$\frac{\mathrm{e}_n^{PF}}{\mathrm{e}_n^{F=0}} = \frac{n_t}{n} = 1 + \Gamma_{PF} \tag{2.128}$$

$$\frac{\mathrm{e}_p^{PF}}{\mathrm{e}_p^{F=0}} = \frac{p_t}{p} = 1 + \Gamma_{PF} \tag{2.129}$$

By incorporating **eq. 2.104** into the aforementioned equations, it is possible to represent the capture cross section $\sigma_{n,p}$ as:

$$\sigma_{n,p}^{PF} = \sigma_{n,p}^{F=0}(1 + \Gamma_{PF}) \tag{2.130}$$

Including **eq. 2.124** in **eq. 2.127**, Γ_{PF} can be expressed as:

$$\Gamma_{PF} = \frac{1}{\alpha^2}[1 + (\alpha - 1)exp(\alpha)] - \frac{1}{2} \tag{2.131}$$

and:

$$\alpha = \frac{1}{k_BT}\sqrt{\frac{q^3F}{\pi\varepsilon_r\varepsilon_0}} \tag{2.132}$$

The only adjustable parameter, as listed in **Table 2.7**, is the relative permittivity ε_r.

In the depletion region, the SRH rate of generation from defects decreases exponentially with the height of the most important energy barrier to be crossed by the generated electron as explained previously. The decrease in the potential barrier by the Poole-Frenkel effect (PF) leads to further generation of carriers from defects by a factor Γ_{PF} as shown in **Fig. 2.28**.

Table 2.7: Parameters of Poole-Frenkel model **eq. 2.131** for silicon. The model is activated by setting *Traps((PooleFrenkel(Electron) PooleFrenkel(Hole))* in physics section and all the parameters are accessible in *PooleFrenkel* parameter set.

Symbol	Parameter	Electrons	Holes	Unit
ε_r	epsPF	11.7	11.7	1

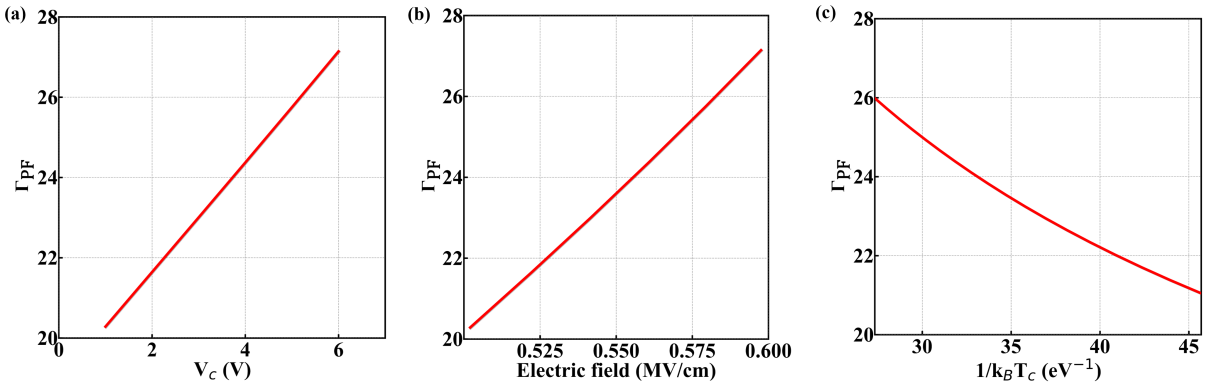


Figure 2.28: Electric field enhancement factor Γ_{PF} for the Poole-Frenkel effect (PF) as a function (a) V_c, (b) the electric field in the peripheral region near the top interface shown in **Fig. 2.8** and (c) $1/k_BT_c$. The strength of the electric field has a significant impact on Γ_{PF}.

3.2.2 Trap-assisted tunneling

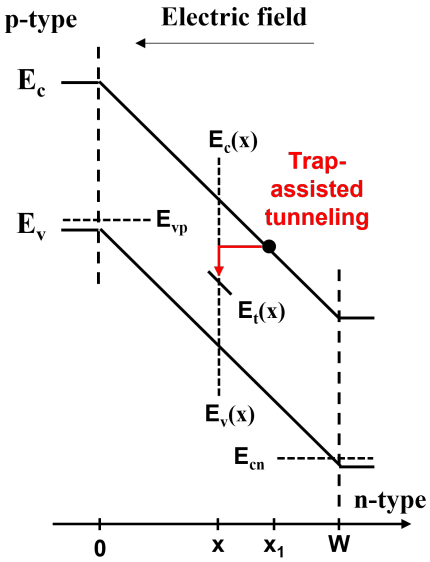


Figure 2.29: Energy-band diagram of the depletion region in a reverse-biased junction. An electron tunnels across the trap potential barrier from x_1 to x.

The presence of a potential barrier between trap states and bands results in Trap-Assisted Tunneling (TAT), as depicted in **Fig. 2.29**. A detailed explanation of the tunneling process is provided in **Appendix B**. Assuming a linear potential $V(x) = qF_x$, with q the elementary charge and F_x the electric field strength in the x-direction, the emission rate enhancement by

an electron tunneling across a triangular trap barrier can be expressed as [96]:

$$\frac{e_n^{TAT}}{e_n^{F=0}} = 1 + \frac{1}{k_B T}\int_0^{\Delta E_n} exp(\frac{E}{k_B T})\frac{A_i^2(2m^*\gamma^{-2}\hbar^{-2}E)}{A_i^2(0)}dE \tag{2.133}$$

where:

$$\gamma = (2qFm^*h^{-2})^{1/3} \tag{2.134}$$

with m^* is the tunneling effective mass. A_i is the Airy function defined in **Appendix B**. The second term of the right-hand side of the above expression represents the excitation probability of a carrier at the trap level to an excited level E multiplied by the probability that an electron at x_1 will tunnel to a trap at x. In forward bias, electron tunneling from trap at location x in **Fig. 2.29** can only occur at an energy level between the local conduction-band minimum $E_c(x)$ and the conduction-band minimum at the neutral n side E_{cn} because there are no available states below this level. In reverse bias, where the trap level is above E_{cn}, the integration interval is the whole trap depth. The same criteria also applies to holes. Considering the relative position of the trap level, the integration intervals for electrons and holes can be written as follows:

$$\Delta E_n(x) = \begin{cases} E_c(x) - E_{cn}, & E_t(x) \leq E_{cn} \\ E_c(x) - E_t(x), & E_t(x) > E_{cn} \end{cases} \tag{2.135}$$

and:

$$\Delta E_p(x) = \begin{cases} E_{vp} - E_v(x), & E_t(x) > E_{vp} \\ E_t(x) - E_v(x), & E_t(x) \leq E_{vn} \end{cases} \tag{2.136}$$

where E_{cn} and E_{vp} are the the conduction- and valence-band minimums at the neutral n- and p-type side. The inclusion of the asymptotic behavior $A_i(y) \sim exp(-2/3y^{3/2})$ in **eq. 2.133** allows for the capture cross section enhancement to be expressed in a similar manner as **eq. 2.130** for TAT:

$$\Gamma_{TAT} = \frac{\Delta E_{n,p}}{k_B T}\int_0^1 exp[\frac{\Delta E_{n,p}}{k_B T}u - K_{n,p}u^{3/2}]du \tag{2.137}$$

with $u = E/\Delta E_{n,p}$ and $K_{n,p} = \frac{4}{3}\frac{\sqrt{2m^*\Delta E_{n,p}^3}}{q\hbar|F|}$. An approximation of this expression can be given by:

$$\Gamma_{TAT} \approx \begin{cases} \sqrt{\pi}\tilde{E}\cdot exp[\frac{1}{3}\tilde{E}^2](2 - erfc[\frac{1}{2}(\frac{\tilde{E}_n}{\tilde{E}} - \tilde{E})]), & \tilde{E} \leq \sqrt{\tilde{E}_n} \\ \sqrt{\pi\tilde{E}}\tilde{E}_n{}^{\frac{1}{4}}exp[-\tilde{E}_n + \tilde{E}\sqrt{\tilde{E}} + \frac{1}{3}\frac{\sqrt{\tilde{E}_n^3}}{\tilde{E}}]erfc[\tilde{E}_n{}^{\frac{1}{4}}\sqrt{\tilde{E}} - \frac{\tilde{E}_n{}^{\frac{3}{4}}}{\sqrt{\tilde{E}}}], & \tilde{E} > \sqrt{\tilde{E}_n} \end{cases} \tag{2.138}$$

where $\tilde{E}$ and $\tilde{E}_n$ are respectively defined as:

$$\tilde{E} = \frac{E}{E_0}, \quad \text{where } E_0 = \frac{\sqrt{8m^* k_B^3 T^3}}{q\hbar} \tag{2.139}$$

$$\tilde{E_n} = \frac{\Delta E_n}{k_B T} = \begin{cases} 0, & k_B T ln(\frac{n}{n_i}) > 0.5E_g \\ \frac{0.5E_g}{k_B T} - ln(\frac{n}{n_i}), & E_t \leq k_B T ln(\frac{n}{n_i}) \leq 0.5E_g \\ \frac{0.5E_g}{k_B T} - \frac{E_t}{k_B T}, & E_t > k_B T ln(\frac{n}{n_i}) \end{cases} \tag{2.140}$$

The only adjustable parameter, as listed in **Table 2.8**, is the tunneling mass m^*.

In the depletion region, where the electric field is high and responsive to the applied voltage, it is necessary to consider carrier tunneling across the potential barrier between the trap state and band. This leads to an enhancement of carrier emission Γ_{TAT} as shown in **Fig. 2.30**.

Table 2.8: Parameters of trap-assisted tunneling model **eq. 2.138** for silicon. The model is activated by setting *Traps((Tunneling(Hurkx))* in physics section and all the parameters are accessible in *HurkxTrapAssistedTunneling* parameter set.

Symbol	Parameter	Electrons	Holes	Unit
m^*	mt	0.5	0.5	1

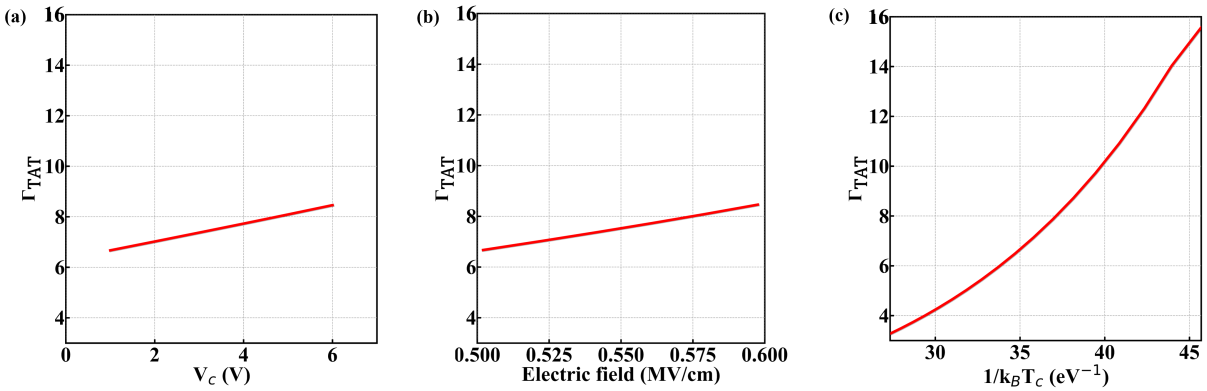


Figure 2.30: Electric field enhancement factor Γ_{TAT} for the trap-assisted tunneling as a function (a) V_c, (b) the electric field in the peripheral region near the top interface shown in **Fig. 2.8** and (c) $1/k_B T_c$. The thermal dependence is more pronounced than the electric field dependence at the electric field magnitudes being considered.

Conclusions

In **Fig. 2.31**, the various contributions to the Shockley-Read-Hall (SRH) Recombination-Generation (R-G) rate are illustrated. The consideration of Trap-Assisted Tunneling (TAT) and Poole-Frenkel effect (PF) results in both electric field and thermal dependence, with the latter being more significant. This is due to the fact that the generation from defects increases exponentially as the energy barrier decreases. The thermal dependence is also influenced by the tunneling effect, which necessitates a lower carrier energy than the barrier energy to tunnel through the barrier that exists between the trap states and bands.

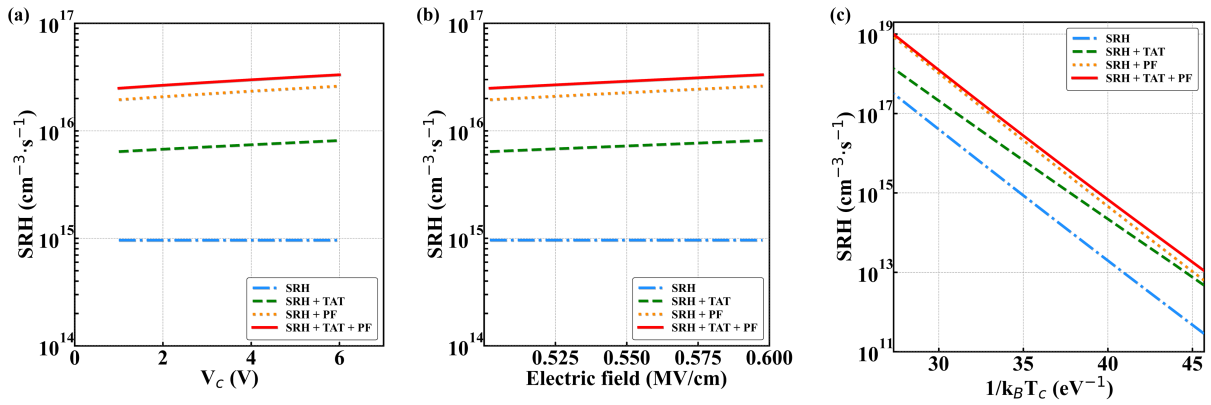


Figure 2.31: Generation rate in a depleted part considering SRH (blue), plus TAT (green), plus PF (orange) and all together (red) as a function (a) V_c, (b) the electric field in the peripheral region near the top interface shown in **Fig. 2.8** and (c) $1/k_BT_c$. The electric field dependence is more significant than the thermal dependence because the generation from defects increases exponentially as the energy barrier decreases due to PF. Additionally, the TAT requires lower carrier energy than the barrier energy to tunnel through the barrier between the trap states and bands.

3.3 Indirect band-to-band tunneling

The process of Band-to-Band Tunneling (BTBT) involves the tunneling of electrons from the valence band to the conduction band of a semiconductor material. However, in the case of silicon, this process is indirect, due the band structure as previously illustrated in **Fig. 2.3**, and requires the consideration of electron-phonon collisions. These collisions serve as a source of momentum for the tunneling electrons, allowing them to tunnel from the valence to the conduction band despite the fact that the minimum energy required for tunneling does not occur at the same momentum in both bands, which is illustrated by **Fig. 2.32**. While the main assumptions of the indirect BTBT modeling are summarized here, the complete derivation can be found in [100]. Phonon-assisted band-to-band tunneling is modeled on the basis of the Kubo formalism for the tunneling conductivity [101][102]. For that purpose the constant field approximation is used, where the electric field is assumed to be constant over the complete tunneling length $l_t = (E_g \pm \hbar\omega)/e|F|$, which is shown in **Fig. 2.32**. With the phonon occupation number

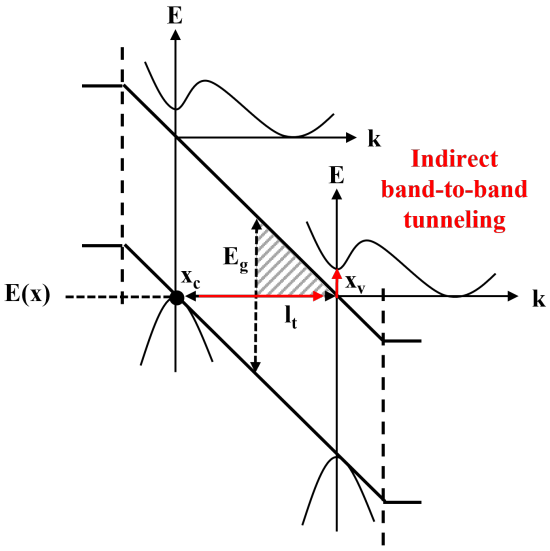


Figure 2.32: Energy-band diagram of the depletion region in a reverse-biased junction considering the band valleys. An electron tunnels through the forbidden band gap and then absorbs phonon to reach the nearest valley.

$f_B = [exp(\hbar\omega/k_BT) - 1]^{-1}$ defined at a single-frequency, this results in an equivalent R-G rate of the form [100][103]:

$$R = \frac{2.47 \cdot 10^{21}}{cm^3/s} \cdot F^2 \cdot \sum_{\alpha=x,y,z} \frac{\sqrt{m_\perp^\alpha m_\parallel^\alpha}}{m} [f_B H(x_\alpha^\mp) + (f_B + 1) H(x_\alpha^\pm)] \cdot (f_v - f_c) \tag{2.141}$$

The first term constant of **eq. 2.141** derives from Kane's two-band model [104] and linear electron-phonon coupling assuming intravalley acoustic scattering [100][105][106][107]. The model takes into account the fact that the effective mass of electrons in the conduction band and the effective mass of holes in the valence band are not constant, but are anisotropic function of the direction of motion. $m_\perp^\alpha$ and $m_\parallel^\alpha$ are the reduced effective tunneling mass transverse or parallel part of the field direction considering the anisotropy of the band valleys. The function H(x), which represents the probability of finding the carriers in the conduction or valence band, is equal to:

$$H(x) = \frac{A_i(x)}{x^2} + \frac{A_i'(x)}{x} + A_{i_1}(x) \tag{2.142}$$

with:

$$x_\alpha^\pm = (\frac{3}{2}\frac{F_c^{\alpha\pm}}{F})^{2/3} \tag{2.143}$$

To account for energy conservation in the electron-phonon system, tunneling under forward bias ($f_v < f_c$ and $np > n_i^2$ resulting in recombination) and reverse bias ($f_v > f_c$ and $np < n_i^2$ resulting in generation) can be described using its lower and upper sign. $A_i(x)$ is the Airy

function, $A_i'(x)$ its derivative and $A_{i_1}(x)$ its integral. An asymptotic approximation of H(x) is given by [108]:

$$H(x) \approx \frac{x^{-9/4}}{2\sqrt{\pi}} \cdot exp(-\frac{2}{3}x^{3/2}), \quad x \to \infty \tag{2.144}$$

The probability of occupation in the conduction and valence bands can be described by Fermi-Dirac distribution functions, which are given by $f_{c,v} = [exp((E(x) - E_{f_{n,p}})/k_BT) + 1]^{-1}$. Here, $E(x)$ represents the energy level of the tunneling transition, which is determined by:

$$E(x) = E_c \pm \hbar\omega - e\phi(x_c^{\pm}) = E_v - e\phi(x_v^{\pm}) \tag{2.145}$$

where $\hbar\omega$ is the phonon energy and ϕ the electrostatic potential at the conduction band $x_c^{\pm}$ and at the valence band $x_v^{\pm}$. In a non-equilibrium state ($np \gg n_i^2$), the concentration of electrons and holes can be expressed through the quasi-Fermi levels $E_{f_{n,p}}$:

$$n = n_i exp(\frac{E_{f_n} - E_i}{k_BT}) \tag{2.146}$$

$$p = n_i exp(\frac{E_i - E_{f_p}}{k_BT}) \tag{2.147}$$

The difference of the Fermi-Dirac distribution functions, which represents the probability of finding an electron in the conduction band and a hole in the valence band, is given by:

$$f_v - f_c = \frac{n_i^2 - np}{(n + n_i) \cdot (p + n_i)} \tag{2.148}$$

The validity of **eq. 2.141** rests on the energy of the interband coupling, which is given by $e \cdot F \cdot a$. For this equation to hold true, this energy must be significantly smaller than the characteristic energy of the intraband coupling, which is defined as $\hbar\Theta_{r,\|}^{\alpha}$. The electrooptical frequency $\Theta_{r,\|}$ is expressed as:

$$\Theta_{r,\|}^{\alpha} = (\frac{e^2F^2}{2\mu_{\|}^{\alpha}\hbar})^{1/3} \tag{2.149}$$

where:

$$\frac{1}{\mu_{\|}^{\alpha}} = \frac{1}{m_{\|}^{\alpha}} + \frac{1}{m_v} \tag{2.150}$$

with m_v the effective mass in the valence band. This aforementioned assertion leads to:

$$\frac{e \cdot F \cdot a}{\hbar\Theta_{r,\|}^{\alpha}} \ll 1 \tag{2.151}$$

The upper limit of the theory is defined by considering the critical field strength at the tunneling length $a = l_t$. The previous equation is then rewritten accordingly as:

$$(E_g \pm \hbar\omega) = (\frac{3}{4})^{2/3}\hbar\Theta_{r,\|}^{\alpha}(F_c^{\alpha\pm}) \tag{2.152}$$

The critical field strength can be expressed as:

$$F_c^{\alpha\pm} = 2^{2/3}\frac{(E_g \pm \hbar\omega)}{\hbar\Theta^{\alpha}_{r,\|}} = \frac{4}{3}\frac{\sqrt{2\mu^{\alpha}_{\|}}}{e\hbar}(E_{g,eff} \pm \hbar\omega)^{3/2} \tag{2.153}$$

When the field direction is assumed to be <111>, the recombination rate can be formulated as:

$$R = A \cdot F^{7/2} \cdot [\frac{(F_c^{\mp})^{-3/2} exp(-\frac{F_c^{\mp}}{-F})}{exp(\frac{\hbar\omega}{k_B T} - 1)} + \frac{(F_c^{\pm})^{-3/2} exp(-\frac{F_c^{\alpha\pm}}{-F})}{1 - exp(-\frac{\hbar\omega}{k_B T})}] \cdot \frac{np - n_i^2}{(n + n_i)(p + n_i)} \tag{2.154}$$

with:

$$F_c^{\pm} = B(E_g \pm \hbar\omega)^{3/2} \tag{2.155}$$

The parameters are listed in **Table 2.9**.

Table 2.9: Parameters of Schenk band-to-band tunneling model **eqs. 2.154-2.155** for silicon. The model is activated by setting *Recombination(Band2Band(Model=Schenk))* in physics section and all the parameters are accessible in *Scharfetter* parameter set.

Symbol	Parameter	Value	Unit
A	A	$8.977{\cdot}10^{20}$	$cm^{-1}s^{-1}V^{-2}$
B	B	$2.14667{\cdot}10^{7}$	$V{\cdot}cm^{-1}eV^{-3/2}$
$\hbar\omega$	hbarOmega	18.6	meV

In the depletion region, where the electric field is high and responsive to the applied voltage, it is necessary to consider carrier tunneling across potential barrier within the forbidden band gap. As depicted in **Fig. 2.33**, the generation rate of indirect BTBT increases exponentially with the electric field. **Fig. 2.34** displays a generation rate map that is typically simulated from the indirect BTBT process, highlighting its significant activity in both the central n++/p+ and peripheral n-/p-well regions.

Conclusions

The tunneling of electrons from the valence band to the conduction band must be considered in high electric field. Electron-phonon collisions play a role in promoting tunneling in indirect bandgap semiconductors like silicon. The generation rate of indirect Band-to-Band Tunneling (BTBT) is shown to increase exponentially with the electric field. The activity occurs in both the central n++/p+ and peripheral n-/p-well regions of the studied SPAD.

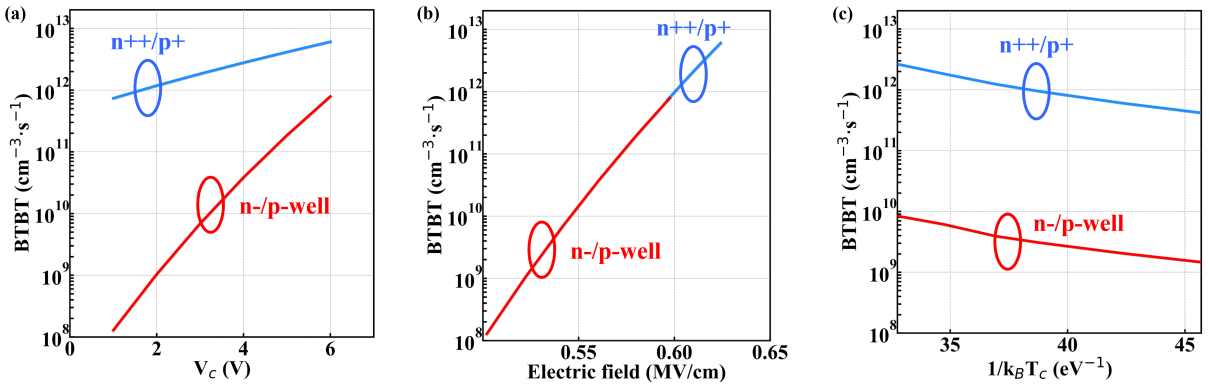


Figure 2.33: Indirect Band-to-Band Tunneling (BTBT) as a function of (a) V_c, (b) the electric field in the central n++/p+ and peripheral n-/p-well regions near the top interface shown in **Fig. 2.8** and (c) $1/k_B T_c$. The generation rate of BTBT increases exponentially with the electric field.

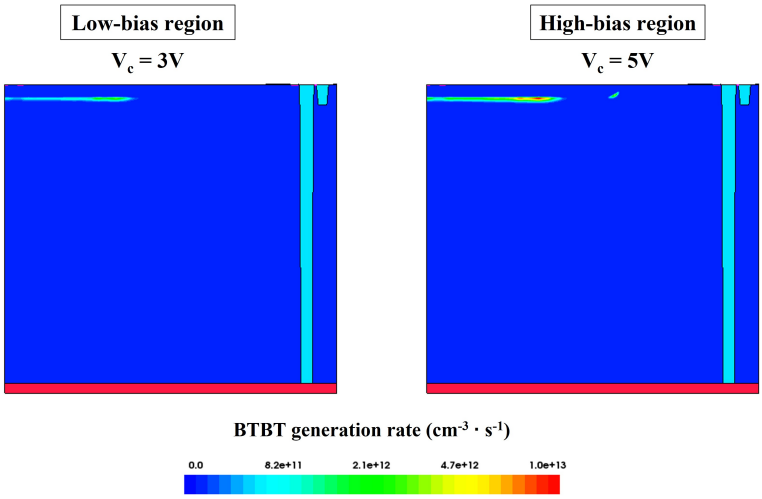


Figure 2.34: Indirect Band-to-Band Tunneling (BTBT) generation rate map simulated at T_c=333K and V_c=3V and 5V.As the applied voltage is increased, the generation rate of carriers also increases, resulting in the activation of both n++/p+ and n-/p-well regions.

3.4 Insights into the generation process through experimental analysis

It would be of experimental interest to determine the temperature or electric field increase required to double the previously mentioned generation processes. It is possible to obtain the result of multiplying the generation rate by 2 raised to the power of n by utilizing the exponential of the natural logarithm. This can be expressed as:

$$R \cdot 2^n = exp(ln(R)) \cdot exp(n \cdot ln(2)) \tag{2.156}$$

The aim is to fit this relationship between the growth rate of generation and the multiplication factor n as shown in **Fig. 2.35** and **Fig. 2.36** to extract the increase required to double the rate. The best fit corresponds to an exponential fit with a constant within the exponential equal to $b = ln(2) \sim 0.69$ according to **eq. 2.156**. The results are reported in **Table 2.10**. The

SRH generation rate is estimated to double each 10°C whatsoever the electric-field enhancement and the BTBT generation rate each 40°C and each $8 \cdot 10^3$V/cm. The saturation observed above 0.5MV/cm and the relatively slow increase make it challenging to draw any significant conclusions regarding the electric field for SRH-TAT and SRH-PF.

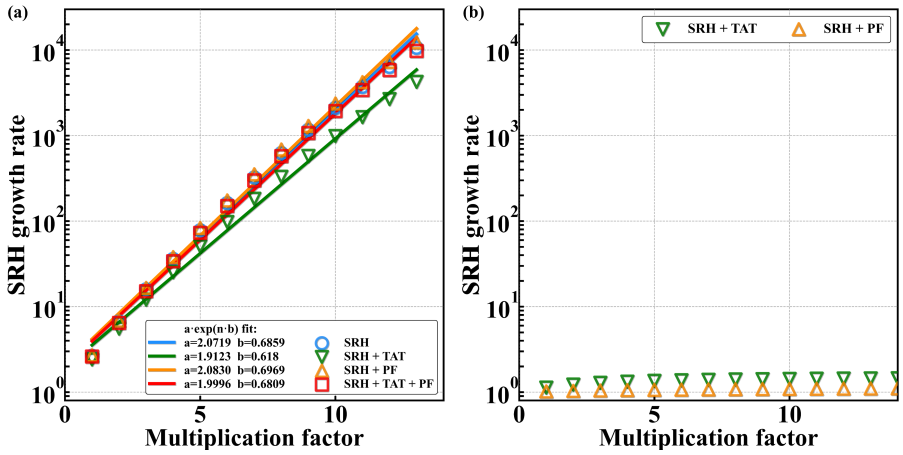


Figure 2.35: Growth rate of the SRH generation mechanism as a function of the multiplication factor. At each step, (a) an increase of 10°C in temperature and (b) an increase of 0.1MV/cm in electric field are applied to double the generation rate, which is not effective for the electric field.

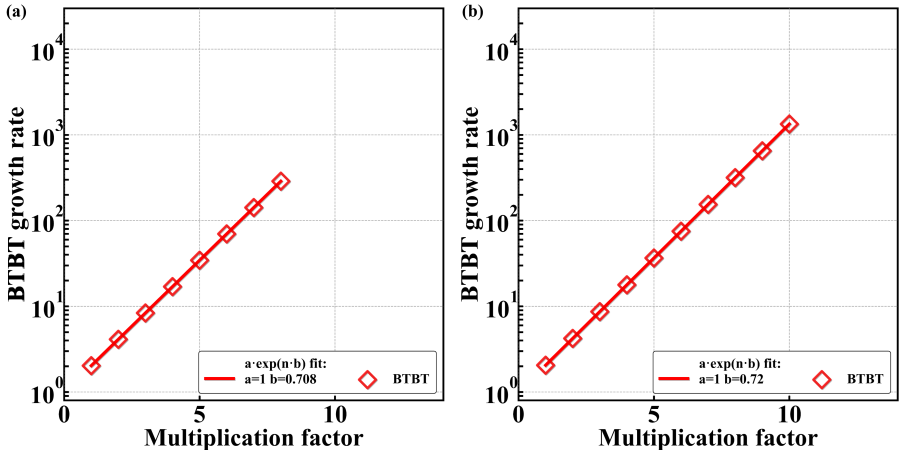


Figure 2.36: Growth rate of the BTBT generation mechanism as a function of the multiplication factor. At each step, (a) an increase of 40°C in temperature and (b) an increase of $8 \cdot 10^3$V/cm in electric field are applied to double the generation rate.

Table 2.10: Corresponding temperature and electric field increase to double the generation rates in the range defined from room temperature to 150°C.

RG	Temperature	Electric field
SRH	10°C	
BTBT	40°C	$8 \cdot 10^3$V/cm

Conclusions

Determining the temperature or electric field increase necessary to double the generation processes would be of experimental interest. The exponential of the natural logarithm of the generation that doubles each temperature or electric field value can be fitted to achieve this. The estimated doubling rate of the SRH generation is 10°C, regardless of the electric-field enhancement, while the BTBT generation rate is expected to double every 40°C and every $8 \cdot 10^3$V/cm.

4 Conclusions

1

The impact-ionization of electron-hole pairs in semiconductor material requires a certain threshold energy to occur. It depends on the crystallographic orientation and varies with the bandgap energy. The rate of impact ionization is influenced by this threshold energy, which is mainly determined by the kinetic energy brings by the electric field to carriers. A threshold electric field can be defined accordingly to the material properties. The impact-ionization coefficients for electrons and holes consider these factors to calculate the number of electron-hole pairs generated per unit distance. The carrier-phonon mean-free path decreases with increasing the lattice temperature which results in an decrease of the impact-ionization rate.

2

The non-linear McIntyre model is utilized to simulate the avalanche breakdown probability P_t, by taking into account impact-ionization coefficients and electric field streamlines. The model allows for the examination of different physical fields to define streamlines, resulting in variations in the P_t map. The avalanche breakdown probability maps are dependent on temperature, voltage, and carrier type. Two voltage dynamics are related to the electric field in both central n++/p+ and peripheral n-/p-well regions, and one positive thermal trend is related to the temperature dependence of impact-ionization coefficients. The electric field distribution, and therefore P_t, is altered by the dopant profile within the volume and at the interface, as well as device dimensions. By using the Photo-Detection Efficiency (PDE) equation to compare simulated P_t with experimental results, the model is predictive for the whole range of temperature and voltage. The breakdown voltage V_{BD} is extracted from both the PDE using the model and the current using Sentaurus workbench, and the results show a good level of agreement. Another approach involves an analytical model implemented in the sdevice module to compute P_t, which reveals the activity of both regions but with a lower magnitude when integrated in the entire device.

3 The Monte-Carlo method considers both the drift caused by the electric field and the diffusion brought about by Brownian motion, allowing for the capture of individual carrier paths. The multiplication gain factor, which is dependent on the impact-ionization probability, must tend towards infinity for the avalanche breakdown condition to be reached. When examining the studied SPAD architecture, diffusing carriers can activate avalanche-leading regions by overcoming electrostatic potential barriers, which are not captured using the deterministic McIntyre model computation along electric field streamlines. By examining each mesh point within the device, it is possible to differentiate between central and peripheral avalanche contributions and to analyze the statistical behavior of each carrier path. For instance, the jitter shape can be explained from the time it take to evacuate carriers form depleted to undepleted regions.

4 By analyzing the charge density within the device, one can gain insight into the carrier flux and electrostatic potential. By taking into account carrier capture/emission and the statistics of occupation function for bands and traps, the thermal Shockley-Read-Hall (SRH) model explains the Recombination-Generation (R-G) of charge carriers (electrons and holes) in a semiconductor material due to filled and emptied defects. This model is applicable in both non-degenerate and degenerate semiconductors, allowing for a comprehensive understanding of the steady-state process. The minority carrier lifetime for electrons and holes injected in a p-type and n-type can be derived to comprehend device's efficiency.

5 The minority carrier lifetime can be defined in terms of the dopant concentration within the volume of the material. This definition allows for a focus on the impact of dopants primarily within the depleted volume. Alternatively, the capture cross section can be used to define the minority carrier lifetime, which represents the probability interaction of a carrier being captured by a trap state. This definition enables the investigation of the role of interfaces in carrier generation beyond the depletion region in the studied SPAD.

6 The concept of thermal activation energy has been introduced to explain the relationship between trap energy and carrier generation from traps. The thermal dependence of the carrier generation rate is primarily influenced by the bandgap and intrinsic energy levels, as well as on the energy level of the trap relative to the intrinsic energy level. When the energy state of the trap is located in the middle of the bandgap or close to it, the minimum energy required for carrier generation is achieved, resulting in the maximum carrier generation rate. The thermal activation energy is not only limited by the trap energy level but also by the capture cross section.

7

In **Fig. 2.31**, the various contributions to the Shockley-Read-Hall (SRH) Recombination-Generation (R-G) rate are illustrated. The consideration of Trap-Assisted Tunneling (TAT) and Poole-Frenkel effect (PF) results in both electric field and thermal dependence, with the latter being more significant. This is due to the fact that the generation from defects increases exponentially as the energy barrier decreases. The thermal dependence is also influenced by the tunneling effect, which necessitates a lower carrier energy than the barrier energy to tunnel through the barrier that exists between the trap states and bands.

8

The tunneling of electrons from the valence band to the conduction band must be considered in high electric field. Electron-phonon collisions play a role in promoting tunneling in indirect bandgap semiconductors like silicon. The generation rate of indirect Band-to-Band Tunneling (BTBT) is shown to increase exponentially with the electric field. The activity occurs in both the central n++/p+ and peripheral n-/p-well regions of the studied SPAD.

9

Determining the temperature or electric field increase necessary to double the generation processes would be of experimental interest. The exponential of the natural logarithm of the generation that doubles each temperature or electric field value can be fitted to achieve this. The estimated doubling rate of the SRH generation is 10°C, regardless of the electric-field enhancement, while the BTBT generation rate is expected to double every 40°C and every $8 \cdot 10^3$V/cm.

Chapter 3

Investigating the origins of pre-stress Dark Count Rate in SPADs

Objectives

The computation of DCR contribution at each mesh point within the device is made possible by the previously introduced avalanche breakdown probability and generation rate of carriers. This generation of carriers may arise from thermal Shockley-Read-Hall (SRH) generation from defects, which may be further enhanced by Trap-Assisted Tunneling (TAT), and Band-to-Band Tunneling (BTBT) according to the electric field strength. By combining insights from measurement and simulation over temperature and voltage, the origins of average DCR within the low- and high-bias voltage regimes are elucidated in this chapter. Furthermore, the use of Design of Experiment (DoE) enables the identification of the most significant factors and their interactions that impact DCR. Additionally, the statistical dispersion that results in intrinsic and extrinsic DCR populations will be investigated using a Monte-Carlo approach that randomizes the position and number of defects.

Contents

1 Comparing existing modeling approaches for Dark Count Rate

This section provides a brief overview of the approaches available in the existing literature for modeling Dark Count Rate (DCR), highlighting the distinctive features of our methodology in comparison to these approaches. PDE and DCR modeling are both based on the avalanche breakdown probability but differ from optical and defect properties, as explained in **Chapter 2**. In other words, the process of exciting carriers can be initiated by any generation processes rather than by the absorption of photons which leads to the following definition:

$$DCR = \int_0^x \int_0^y \int_0^z P_t(x', y', z') \cdot R_{net}(x', y', z') dx' dy' dz' \tag{3.1}$$

Previous studies provide valuable insights into the intricate dynamics of DCR in SPADs based on temperature and voltage investigations. DCR has been derived through analytical models in previous literature and established connections with physical parameters such as dark current or afterpulsing. The results shown in **Fig. 3.1** from these studies do not take into consideration the complete influence of the manufacturing processes, as observed in [109], or the absence of modeling of Band-to-Band Tunneling (BTBT), as seen in [110]. The doping profile utilized in the second modeling was derived from Suprem process simulation [111], and the anode implantation dose was fine-tuned to correspond with the measured breakdown voltage V_{BD}. However, this approach focused only on the depletion region and made the assumption of a single trap energy level. This means that carriers diffusing from the neutral region were neglected. As a consequence, there is a discrepancy with the observed thermal dependence, as depicted in **Fig. 3.1 (b)-(c)**.

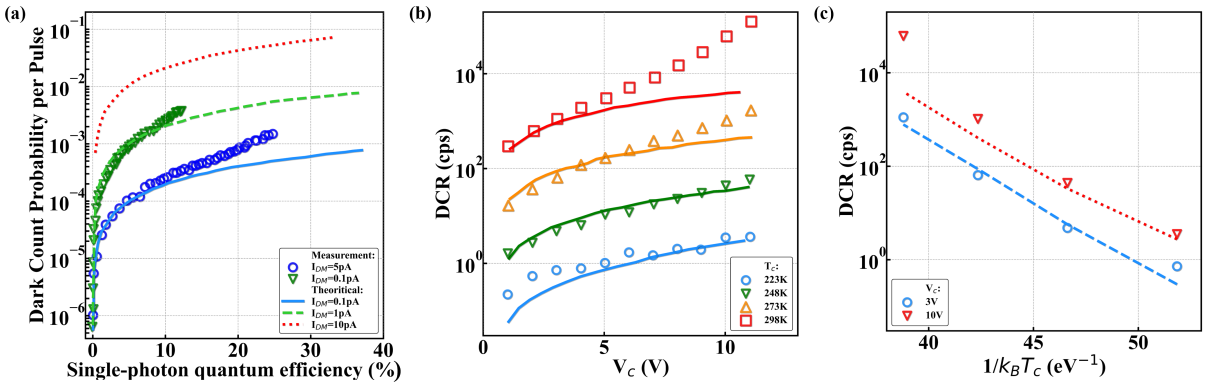


Figure 3.1: (a) Dark Count Probability per pulse as a function of single-photon quantum efficiency for different primary dark current I_{DM} simulated and measured at 240K (circles) and 130K (down triangles) from data reported in [109]. The primary dark current has a substantial impact on the likelihood of dark count occurrences. DCR (b) as a function of V_c for different T_c and (c) as a function of $1/k_BT_c$ for V_c= 3V and 10V from data reported in [110]. Simulation are lines and measurements are symbols. SRH-TAT modeling explain the DCR behavior at low temperature but a deviation at higher temperature and higher bias is observed.

The aforementioned approaches are restricted to analytical models. In contrast, numerical

simulation-based methods employ numerical models to extract physical parameters from outputs. The process is simulated based on literature data and the simulation tool is calibrated by adjusting implantation conditions with literature breakdown voltages [112]. The avalanche triggering probability and electric field are directly extracted from SILVACO simulation tools. From these physical fields, a Verilog-A modeling of DCR is proposed to model DCR as follows. An avalanche pulse is triggered if a random number, drawn from an uniform distribution, is less than the avalanche breakdown probability. And the time duration between two successive carrier generation events follows a Poisson distribution. The focus is on the avalanche depleted region, and the SRH model considers the carrier minority lifetime instead of trap energy. This implies that the thermal dependence of the SRH model is solely determined by the band gap energy E_g, as expressed through the intrinsic concentration n_i in **eq. 2.86**, since $n_i^2 \gg n\dot{p}$ in the depletion region. The trap population and the doping concentration are used as fitting parameters in the SRH and BTBT model, which results in **Fig. 3.2**. To account for the faster growth of DCR with high biasing voltage, a larger tunneling fitting parameter was selected compared to the default value [96]. The electric field map in [112] shows a central and peripheral electrically active region where the electric field is higher than the threshold for impact-ionization phenomena [67]. However, when integrating over the central avalanche region, it is possible to hide the lateral breakdown that occurs at high excess voltage biasing. In such cases, the contribution from the peripheral region becomes significant and cannot be ignored.

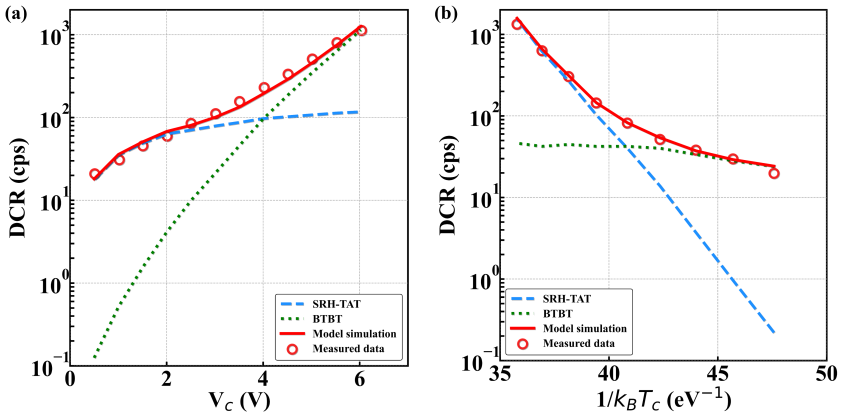


Figure 3.2: DCR as a function of (a) V_c at room temperature and (b) as a function of $1/k_BT_c$ for V_c=3.5V measured and simulated for different generation mechanisms from data reported in [112]. At low temperature the BTBT mechanism is the dominant generation process and with the increase of temperature the SRH-TAT becomes the dominant component. At high excess bias voltage, BTBT becomes the main DCR generation source.

In [113], different guard ring structure are compared assuming a pure Boron process. Both the simulation and experimental data are presented in **Fig. 3.3 (a)**. Simulation results showed that devices with lower doping concentrations in their guard rings had significantly lower DCR compared to those without guard rings or those with guard rings that had higher doping concentrations. The model was improved by incorporating the process-related simulation tool Sprocess [45] instead of relying on simulations based on ideal geometric structures. The trap energy level was assumed to be higher than the middle bandgap by 0.15eV taking into account the considered

doping concentration. In other words, the fitted values of trap energy will vary depending on the doping concentration.

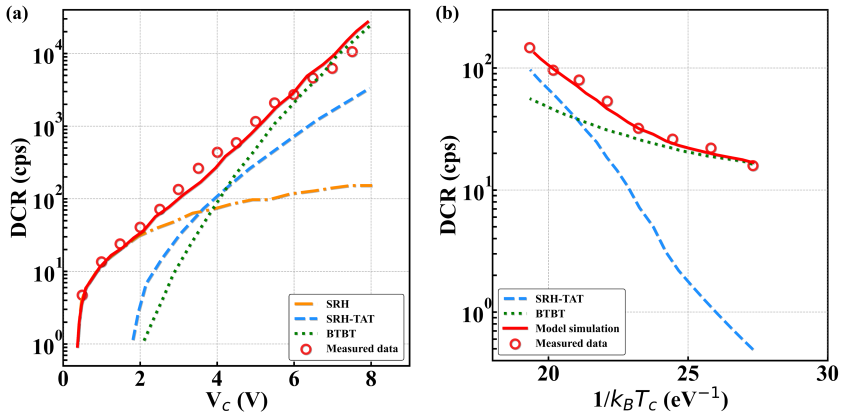


Figure 3.3: (a) DCR as a function of V_c at room temperature measured and simulated for different generation mechanisms from data reported in [113]. At low excess bias voltage, the SRH drives the DCR dynamics while at higher voltage BTBT and SRH-TAT becomes dominant. (b) DCR as a function of $1/k_BT_c$ at V_c=0.3V from data reported in [114]. At low temperature, the DCR is dominated by BTBT while at higher temperature SRH-TAT takes over.

More recently, a simulation methodology [114] was proposed based on two fitting parameters, namely the trap energy and the BTBT fitting parameter. The fitting values were obtained by two measurements at low and ambient temperature, at 125K and 325K respectively, which improve the temperature-dependent DCR performance. The doping and trap information are obtained from measurement. Long recharge time of the quenching circuit conceal the afterpulsing phenomenon. By taking into account the contributions from both the upper and bottom neutral regions, it becomes possible to analyze the respective contributions from different regions. In order to analyze each RoI, a cross-section is created within it and then extended to the entire surface by multiplying the obtained data with the surface area of the corresponding zone. The integration of each contribution, while taking into account the varying generation rate, yields the results shown in **Fig. 3.3 (b)**. The BTBT model had a default value within the observed range [96], and the trap energy level was set to 0.19eV corresponding to the neutral divacancy defect family [90]. Thermal generation was also neglected in the neutral regions. DCR generated from guard ring structures dominates the total DCR of this architecture due to the previous underestimation of the doping concentration.

Conclusions

The main challenge in DCR modeling is to determine the contribution of various generation mechanisms by obtaining data on process quality and physical parameters such as trap energies and BTBT parameters that are associated with the specific process under investigation. For accurate predictions, it is essential that the measured and simulated breakdown voltages match. DCR thermal generation is typically more sensitive to changes in temperature compared to DCR tunneling or field-enhanced generation, which exhibit a greater sensitivity to variations in electric field. The DCR can be inferred from thermal Shockley-Read-Hall (SRH) generation of carriers from defects, potentially enhanced by Trap-Assisted Tunneling (TAT). In addition, Band-to-Band Tunneling (BTBT) and diffusion mechanisms can also be at the origin of DCR.

2 Using mean-based modeling to investigate Dark Count Rate

> The main results of this section have been published in "Dark Count Rate in Single-Photon Avalanche Diodes: Characterization and Modeling study", in ESSCIRC 2021 - IEEE 47th European Solid State Circuits Conference (ESSCIRC), 2021, pp. 143–146 [74].

An accurate physical description of the traps in term of parameters and localizations within the device is critical to accurately predict each individual DCR contribution. This is addressed in this section relying on a dedicated set of measurements and a simulation methodology coupling spatially-dependent breakdown probability to the aforementioned carrier generation and diffusion mechanisms described in **Chapter 2**. This study offers significant value by creating a model using TCAD simulations and measurements on a broad range of architectures and processes, spanning a wide temperature and voltage range. The mean-based model provides valuable insights into the carrier generation mechanisms and trap populations, including their location.

2.1 Analysis of temperature and voltage dynamics

In order to determine the average behavior of DCR, it is necessary to examine the impact of temperature and electric field on the probability of avalanche breakdown and generation rate. The former is heavily influenced by electric field (**Chapter 2 Section 2.3**), while the latter is affected more by either temperature using SRH-TAT (**Chapter 2 Section 3.1**) or electric field using BTBT (**Chapter 2 Section 3.3**). In order to isolate each individual DCR contribution, 64 SPADs are used to measure DCR at various V_c values, which are voltages applied above the avalanche breakdown voltage V_{BD} while the diode is in reverse bias condition, with a 500mV voltage-step. Median DCR values at T_c=253K, 333K, and 353K are chosen to test the model. For experimental comparison, integration over the whole device is carried out resulting in the average DCR in Counts Per Second (cps). Furthermore, to enable comparison between different architectures, both the experimental and simulated DCR are normalized per unit photo-sensitive

area (cps/µm²).

To capture P_t outside the multiplication region, the simulation is realized along the electric field lines according to the McIntyre model presented in **Chapter 2 Section 2.3**. The impact diffusing minority carriers on DCR is resolved using the gradient of the local minority carrier density in order to calculate the minority carrier diffusion current. The BTBT is calibrated using the Schenck model [100] at high bias and low temperature to guarantee the predominance of tunneling contribution over thermal generation. As a first approach, the SRH-TAT model is enriched with the empirical Scharfetter relation defined in **eq. 2.100** (where minority carrier lifetime $\tau_{n,p}$ is expressed as a function of the impurity doping concentration) combined with the Hurkx model for TAT defined in **eq. 2.137** [96] (where $\tau_{n,p}$ decreases with rising temperature). This approach focuses on the dependence of the minority carrier lifetime $\tau_{n,p}$ on the impurity doping density without explicitly defining trap parameters (capture cross section and trap density) except for defect energy level. Typical DCR maps simulated from dopant-limiting lifetime SRH model and BTBT are shown in **Fig. 3.4** and in **Fig. 3.5** respectively. One must pay attention to the area being considered when performing integration because it may conceal certain contributions as shown in **Fig. 3.6**. The presence of two DCR spots can be attributed to the voltage dependence of the avalanche breakdown probability, as illustrated in **Fig. 2.10**.

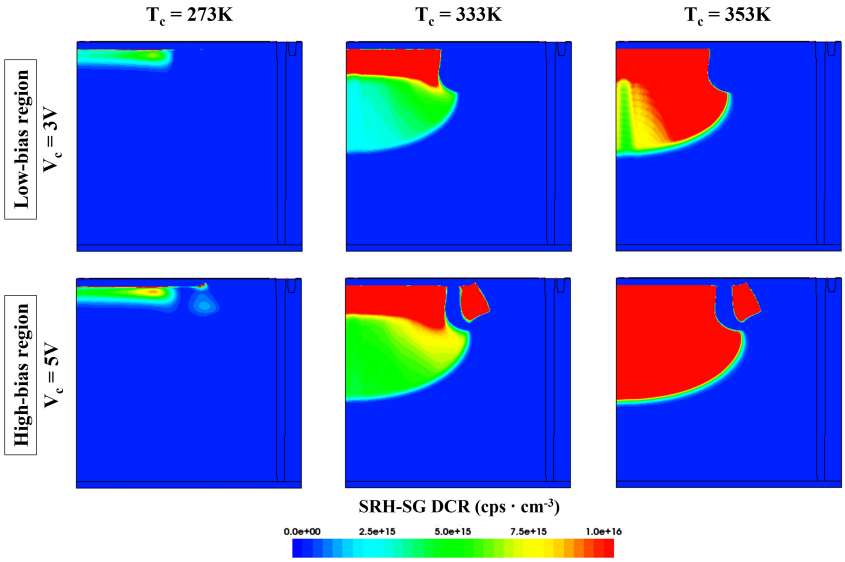


Figure 3.4: Dopant-limiting lifetime SRH DCR map simulated at T_c=273k, 333K and 353K, and V_c=3V and 5V. As the temperature increases, the DCR also increases, with the n++/p+ region being the first to be activated, followed by the n-/p-well regions.

For dopant-limiting lifetime SRH-TAT model, even by tuning the related parameters such as trap energy level, the model is unable to reproduce the DCR tendencies observed against voltage, as illustrated in **Fig. 3.7 (a)**. Even after summing up central and peripheral contributions and fine-tuning the model parameters, the voltage dynamics have not been successfully captured. Despite the fact that this dopant-limiting lifetime SRH-TAT model based on Scharfetter relation has the same dependence shown by the experimental low-bias region, a clear disconnect is still observable in the high-bias region. Attempts to change the BTBT parameters have been

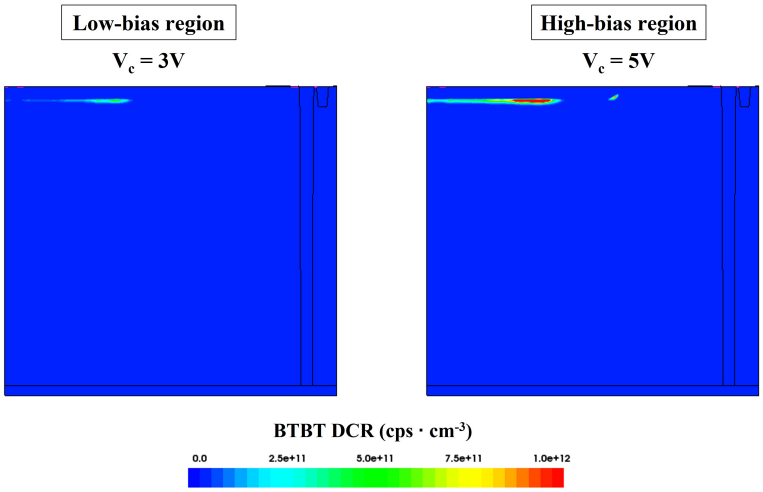


Figure 3.5: Indirect Band-to-Band Tunneling (BTBT) DCR map simulated at T_c=333K and V_c=3V and 5V. As the voltage increases, the DCR also increases, with the n++/p+ region being the first to be activated, followed by the n-/p-well regions.

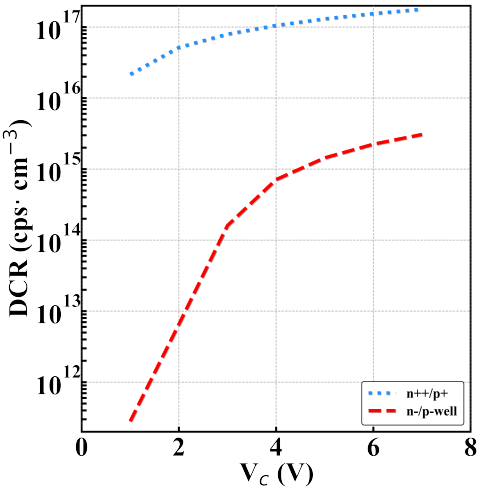


Figure 3.6: DCR as a function of V_c simulated at T_c=333K for both n++/p+ (dashed line) and n-/p-well (dotted line) regions.

unsuccessful in capturing the voltage dependence in this latter region. Each simulated thermal DCR contribution is compared with the experimental DCR thermal activation energy, which represents the combined thermal activation energy of all the different contributions. None of the the aforementioned models have the correct temperature dependence as testified in **Fig. 3.8 (a)**.

Due to the limitations of the previous dopant-limiting lifetime model, the simulation now takes into account the trap charge state, including the localization of traps and the explicit definition of traps through the minority carrier lifetime parameterized by equation **eq. 2.105**. In contrast to the dopant-limiting Scharfetter lifetime model, the $\tau_{n,p}$ is modeled as a product of the trap concentration N_t (rather than the impurity doping density) multiplied by the capture cross section $\sigma_{n,p}$ and the thermal velocity v_{th} for silicon. This configuration allows us to a flexible investigation of possible generation mechanisms inside the SPAD volume or at interfaces. The

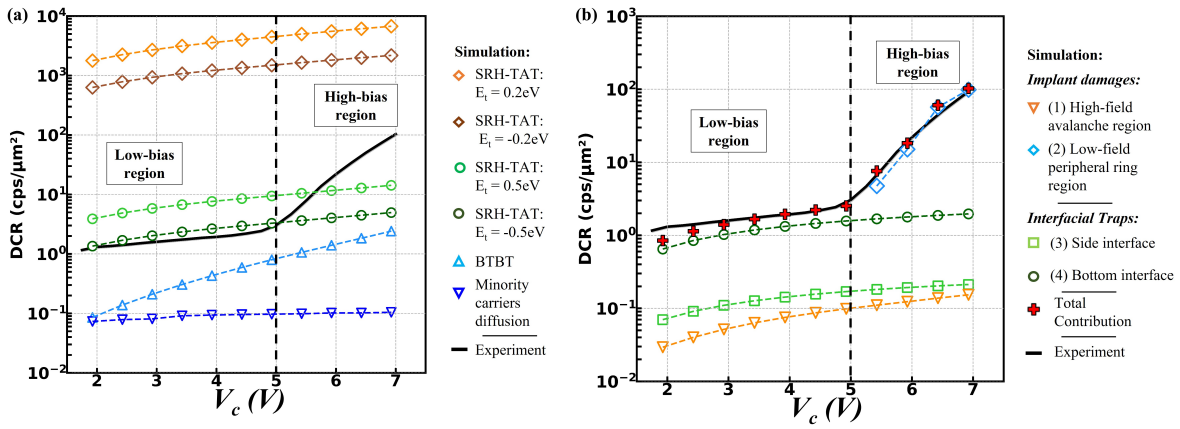


Figure 3.7: DCR as a function of V_c measured (bold line) and simulated (dashed line) at 333K with SRH-TAT within (a) the Scharfetter approximation using various trap energy E_t and (b) the trap description approach for different DCR contribution evaluated for each region. Additionnaly, BTBT and minority carrier diffusion are also superimposed. Without defects, none of these models are able to reproduce the experimental behaviour. N_t is selected as the sole degree of freedom (fitting parameter) and serves as a representative parameter for process-induced defects. Interfacial traps concentration: $10^9 cm^{-2}$ at top, side and bottom interfaces. Bulk traps concentration: $10^8 cm^{-3}$ inside high-field avalanche and $10^{11} cm^{-3}$ inside low-field peripheral ring regions.

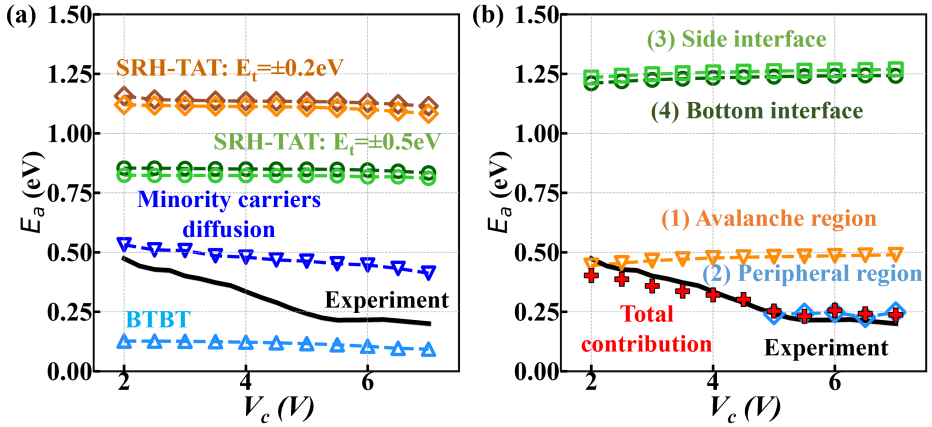


Figure 3.8: DCR thermal activation energy E_a as a function of V_c extracted at 253K, 333K and 353K. The thermal activation energy is equivalent to the slope of the line following the Arhenius law, ln(DCR) vs $1/k_B T_c$. Solid lines are measurements and dashed lines are simulation data. The colours follow the same patterns as **Fig. 3.7 (a)** and **Fig. 3.7 (b)**.

defects are positioned in the bulk and at top, side and bottom interfaces. The emission and capture rates are enhanced with Hurkx tunneling (**eq. 2.137**) and Poole-Frenkel (**eqs. 2.131-2.132**) models using the cross section expressed through **eq. 2.130**. The generation processes at traps discrete localization are coupled consistently with the breakdown probability at this specific position to compute the DCR. A Typical DCR map simulated from defect-limiting lifetime SRH model is shown in **Fig. 3.9**. According to the defect localization, the presence of two DCR spots at the interface can be attributed to the voltage dependence of the avalanche

breakdown probability, as illustrated in **Fig. 2.10**. The inset of the top interface in the lower right-hand corner of the figure enables the observation of defect activation.

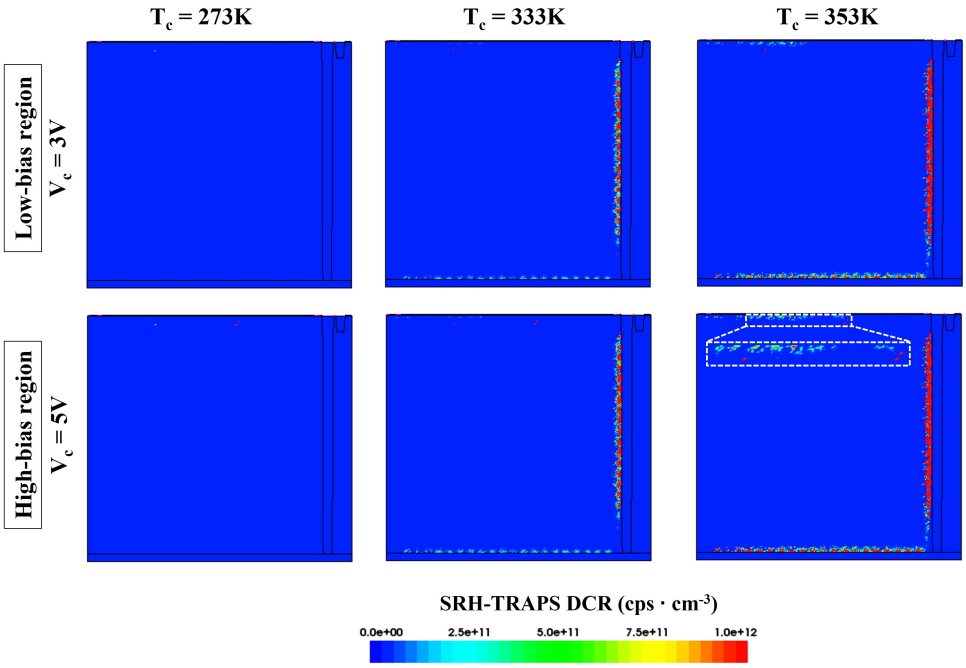


Figure 3.9: Defect-limiting lifetime SRH DCR map simulated at T_c=273K, 333K and 353K, and V_c=3V and 5V. DCR rises as the temperature is increased activating both n++/p+ and n-/p-well regions, as enlighten by the inset at high temperature.

The DCR behavior against voltage is simulated considering defects with single energy level at mid-gap. $\sigma_{n,p}$ is fixed at $10^{-15}cm^2$ and N_t is chosen as the sole degree of freedom (fitting parameter) and representative parameter of process-induced defects (see legends in **Fig. 3.7 (b)**). When the trap state is set to neutral, using either acceptor or donor types leads to equivalent predictions. This implies that at the considered density N_t, the presence of traps inside the device does not significantly alter the device's electrostatic profile. Defects set inside the peripheral ring area and/or at the top interface within this area create the correct effect required to model the high-bias region as depicted in **Fig. 3.7 (b)**. With all contributions included, the whole DCR trend is properly replicated close to the measured data with the right inflexion point when plotted against voltage. Furthermore, as depicted in **Fig. 3.8 (b)**, the thermal dynamics are accurately replicated.

Conclusions

The average DCR behavior is simulated by taking into account temperature and electric field effects in both the low- and high-bias regions. At low temperatures, the low-bias region is influenced by BTBT, while at higher temperatures, thermal SRH generation from the neutral region takes over. In the high-bias region, SRH-TAT from the peripheral n-/p-well region is the primary driving force. However, a simulation that relied on minority carrier lifetime modeled through dopant distribution failed to capture the high-bias regime and required the inclusion of a trap profile within the device.

2.2 Impact of geometrical parameters

Consistent modeling of DCR required an accurate localization of the defects within the structure. In order to deduce their localizations, a wide variety of architectures are designed to allow investigation of a diverse set of DCR contributions. To determine the localization of traps, significant changes are made to the dopant distribution and electric field profile by introducing geometrical variations to the four primary regions that make up the junction, namely the n++, p+, n-well, and p-well regions shown in **Fig. 1.10**. **Fig. 3.10** shows representative DCR modulations against applied voltage for three different architectures. The electric-field distribution brought by different doping profiles unveil specific parts of the device and modulate the avalanche breakdown probability P_t (**Fig. A.1**). As illustrated in **Fig. 3.10**, from experimental and simulation perspectives, the low-bias region of the voltage domain is driven by the central n++/p+ region while the high-bias region of the voltage domain is strongly affected by the peripheral n-/p-well region of the device. The correlation plots of the low-bias region (at V_c=3V) and the high-bias region (at V_c=5V) show in **Fig. 3.11** enlighten the appropriate relationship achieved. The simulations are performed without modifying the doping profile across different architectures. This implies that the DCR constrained by geometrical variations is limited by the electric field distribution, which affects the avalanche breakdown probability.

2.3 Influence of manufacturing processes

2.3.1 Ion implantation

The manufacturing modifications, such as epitaxial growth and ion implantation, alter the avalanche breakdown probability P_t (**Figs. A.2-A.3**) as well as the point defect population modeled by N_t depending of the process modulation strength. The silicon epitaxial growth has been processed with two different methods. The former by keeping constant the gas flow inside the main chamber (Epi. 1) and the latter by changing the flux at specific times during the manufacturing (Epi. 2). The simulations are accounted for in the process simulation implementation without changing the defect parameters. The four process splits shown in **Fig. 3.12** serve as representative examples for each epitaxial growth, emphasizing the influence of ion implantation parameters on the control of the low-bias region. The adjustment of the voltage regimes captures the matching of the breakdown voltage.

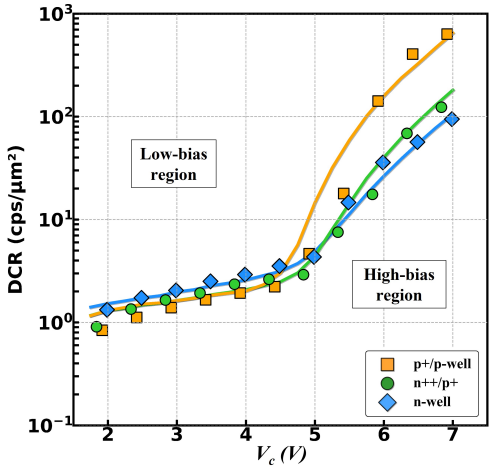


Figure 3.10: DCR as a function of V_c measured (solid line) and simulated (symbol) at T_c=333K for three different architectures varying avalanche region and peripheral ring architectures. Each symbol corresponds to a different device architecture. The central n++/p+ region of the device drives the low-bias region of the voltage domain, while the high-bias region is significantly influenced by the peripheral n-/p-well region.

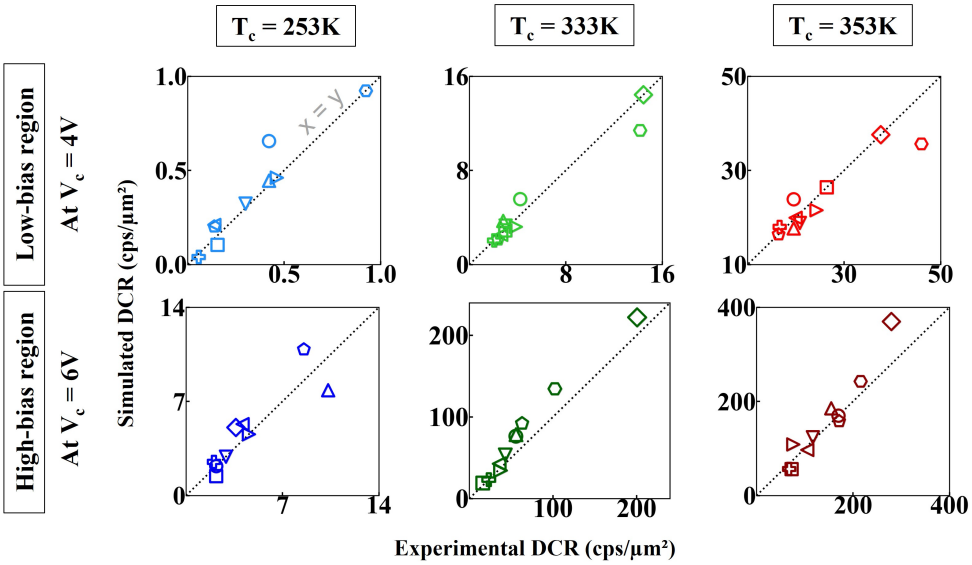


Figure 3.11: DCR correlation plots between experiments and simulations for low-bias and high-bias regions at 253K, 333K and 353K. Each symbol corresponds to a different device architecture. The dash line is the y=x line displaying the perfect relationship achieved between experiments and simulations.

2.3.2 Interface manufacturing

As previously mentioned, silicon-oxide interfaces along bottom and side interface could also have a non-negligible contribution to the DCR from the alteration of the avalanche breakdown probability P_t (**Fig. A.4**). Our simulation results show that these interfaces are the main driver of the low-bias region. This is further confirmed in **Fig. 3.13** where four devices are shown

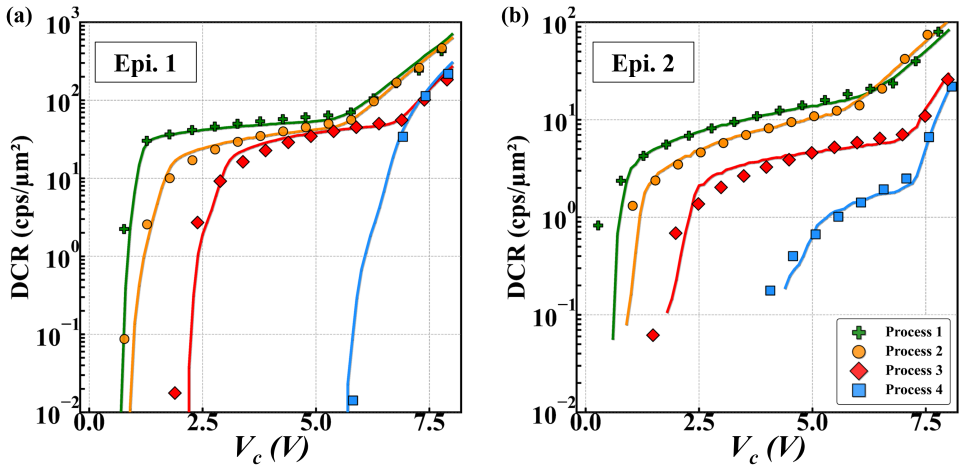


Figure 3.12: DCR as a function of V_c measured (solid line) and simulated (symbol) at 333K for four different architectures varying avalanche region and peripheral ring architectures. Each symbol corresponds to a different junction process. (a) represents Epi. 1. and (b) the Epi. 2. The four process splits showcase epitaxial growth techniques and the influence of ion implantation on controlling the low-bias region, with voltage regime adjustments enabling the matching of the breakdown voltage.

which exhibit different dopant concentration near the side interface or different texturization patterns for the bottom interface keeping the same interfacial defect density constant. For Epi. 1 (**Fig. 3.13 (a)**), the DCR at low-bias is driven by carrier generation at the bottom interface contrary to Epi. 2 (**Fig. 3.13 (b)**) where the carrier generation at the side interface dominates the DCR. This is explained by an overall 30% increase of P_t near the side interface.

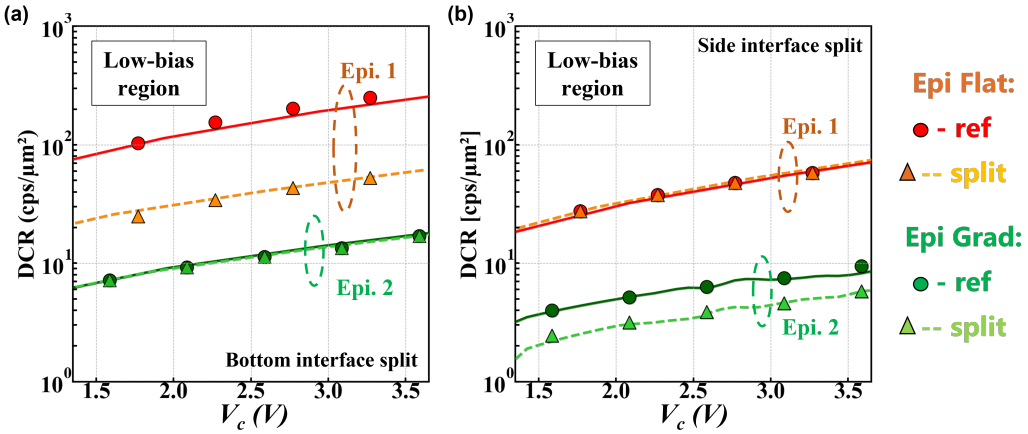


Figure 3.13: DCR as a function of V_c for different interface processes: (a) bottom and (b) side. The up triangle and solid lines are the references, and the down triangle and dashed line are the variations. The legend colors refer to Epi. 1. (red and orange) and Epi. 2. (dark green and green). Lines are measurements and symbols are simulations. In Epi. 1, the low-bias region is influenced by carrier generation at the bottom interface, while in Epi. 2, carrier generation at the side interface dominates.

2.3.3 Process effect partitioning

Design of Experiment (DoE) enables the identification of the most significant factors and their interactions that affect an outcome by carefully selecting and controlling the process factors under investigation, while minimizing the number of experiments required. The goal is to establish the optimal settings for each factor to achieve the desired outcome or identify the factors that require control to enhance the quality and efficiency of a process. In order to determine the optimal compromise between all figures of merit, multiple process splits have been created and analyzed. The focus has been on DCR (**Table 3.1**), which is evaluated by comparing the Process of Reference (PoR) to the various process splits within the same wafer lot. A wafer lot contain wafers that are all subjected to the same manufacturing process steps and conditions, albeit with some degree of variation. The following measurements are taken to provide additional evidence in support of the simulation results obtained previously. **Fig. 1.10** allows for the appreciation of the region where processes vary.

The low-bias region of DCR is significantly affected by changes to the process of interfaces, such as annealing, deposition, or material alteration as shown in **Table 3.1**. Additionally, variations in the ion implantation conditions can alter the local avalanche breakdown probability distribution, leading to modifications in DCR. The high-bias region of DCR is also sensitive to the ion implantation criteria, particularly the contact deposition and the interfacial material used, which may cause embrittlement of diffusing species.

Table 3.1: DCR improvement and degradation in the LB region ($V_c = 3V$) and in the HB region ($V_c = 3V$) for different manufacturing processes.

Process	Improvement		Degradation	
	LB	HB	LB	HB
Deposition of ARC using ALD			x6	
Annealing of the bottom oxide layer	/4			
Filling of DTI using ALD			x10	
Etching of DTI	/2			
Substituting DTI with BDTI	/5	/3		
Deposition of contacts using ALD	/5	/3		
Lower n++ and p-well implant energy	/2			
Lower p-well implant dose			x2	
Higher p-ring implant dose				x2
Higher n-ring implant energy			x2	

* The process details cannot be disclosed due to confidentiality reasons.

Conclusions

Design of Experiment (DoE) allows to pinpoint the most critical factors and their interactions that affect DCR. Architecture constrains and ion implantation conditions shed light on the avalanche breakdown probability contribution to DCR rather than the modulation of defect quantity. Interface manufacturing underlines the role of interfaces in the origin of the low-bias rgion.

3 A Statistical approach to modeling Dark Count Rate

The main results of this section have been published in M. Sicre et al., "Statistical measurements and Monte-Carlo simulations of DCR in SPADs," ESSCIRC 2022- IEEE 48th European Solid State Circuits Conference (ESSCIRC), 2022, pp. 193-196 [79].

3.1 Monte-Carlo simulation of Dark Count Rate

The defect distribution defined constant for all architectures allows to reconstruct the median DCR trends. However, this average value does not capture the intrinsic population (from 1st to 3rd quartiles) and the extrinsic population (outliers) shown in **Fig. 3.17**. The study of local process-induced statistical fluctuations will be performed using the drift-diffusion Monte-Carlo approach outlined in **Chapter 2 Section 2.4**, along with the generation rate of BTBT coupled with Non-Radiative Multiphonon-assisted (NRM) from defects discussed in **Appendix C** rather than SRH framework defined in **Chapter 2 Section 3.1**. The defects will be randomized in terms of their positions and sizes, ranging from single defects to large clusters of up to one hundred defects, all within the previously identified Region of Interest (RoI) in the SPAD device. In each simulation, the defects are uniformly distributed in space and their number follows a Poisson distribution. To confirm the location of defects and ensure agreement between SRH and NRM R-G models, the subsequent analysis will concentrate on the median of DCR. The DoE is composed of four different SPAD architectures, in which geometrical variations are performed without altering the process which are reported in **Table 3.2**.

Table 3.2: SPAD architecture variations

Architecture name	Variation
SPAD 1	p-well overlap p+ contracted
SPAD 2	p-well overlap p+ extended
SPAD 3	n-well contracted
SPAD 4	n-well extended

The experimental results of temperature and field dependent measurements have been used as inputs for simulations. This procedure enables the extraction of thermal activation energy E_a to identify the combination of carrier Recombination-Generation (R-G) mechanisms within RoI (e.g. within the depleted part or at the interfaces). The NRM model is set with mid-gap traps with nonradiative capture cross section computed at $10^{-15}cm^2$ to compare with previous modeling results. The classical definition of capture cross–section has been adopted: $\sigma_n = c_n^{-1}/(v_{th} \cdot n)$ and $\sigma_p = c_p^{-1}/(v_{th} \cdot p)$ for electrons and holes respectively, where $v_{th} = 10^5 m \cdot s^{-1}$ is the carrier saturation velocity and n and p are the total mobile charge densities. The capture/emission rates c_n^{-1} and c_p^{-1} are then computed in **eqs. C.28-C.29** within the NRM framework. From the experimental Arrhenius' plot of median DCR, two main activation energies can be extracted from **Fig. 3.14** in both voltage regimes. This temperature-voltage dependence will be explained based on simulation results. One may observed that there is a shift from an

average thermal activation energy along the voltage range in the previous section to a thermal activation energy along the temperature range in both low- and high-voltage regimes in the current section, enabling better differentiation of the various thermal contributions.

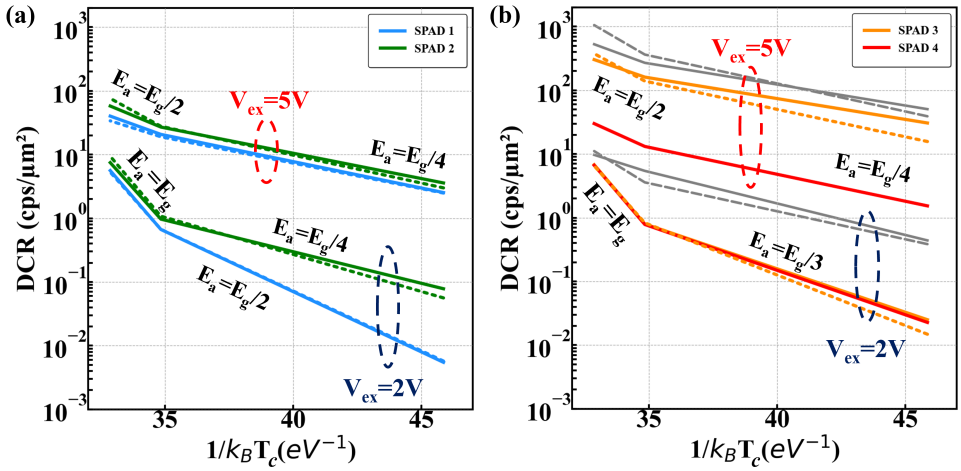


Figure 3.14: DCR as a function of $1/k_BT_c$ measured (bold lines) and simulated (dashed lines) at V_c =2V and 5V for architecture variations of: (a) wells (SPAD 1 and 2) and (b) rings (SPAD 3 and 4). The thermal activation energy E_a is equivalent to the slope of the line following the Arrhenius law ln(DCR) vs $1/k_BT_c$. The activation energies extracted from simulated P_t of **Fig. 2.12 (b)** (-0.08eV at V_c = 2V and -0.03eV at V_c =5V) are removed to the presented DCR activation energy. These results highlight either diffusion ($E_a = E_g$) or NRM and tunneling generation ($E_g/4 < E_a < E_g/2$) mechanisms (with E_g the silicon energy gap). Typical outlier behavior is shown in grey dashed lines.

According to **Fig. 3.14**, the low-bias region is driven by carrier diffusion from interfaces (E_a=E_g) at high temperature, and by tunneling ($E_g/4$<E_a<$E_g/2$) at low temperature. The tunneling part is a combination between NRM and BTBT generation of carriers. For instance, the SPAD 2 which has the highest BTBT contribution at low temperature (**Fig. 3.16 (a)**) has an activation energy E_a of $E_g/4$. Conversely, the SPAD 1 has an activation energy of $E_g/2$ closer to the NRM generation from mid-gap traps. The high-bias region is driven by field-enhanced trap generation in the periphery resulting in a lower activation energy ($E_g/4$<E_a<$E_g/2$). The findings from the simulations presented in both the previous and current sections are consistent and yield the same observations.

3.2 Predicting statistical dispersion

With the generation mechanism well-calibrated on experimental results, the local process-induced statistical fluctuation can be appreciated on **Fig. 3.16** and **Fig. 3.17**. The device architectural constraints allow to verify our assumptions based on different electric field distributions. As a reminder, the defects within the previously identified Region of Interest (RoI) in the SPAD device will be randomized in terms of their positions and sizes, ranging from single defects to large clusters of up to one hundred defects. **Fig. 3.15** illustrates the various components of the simulation methodology proposed to model the process-induced DCR statistical disper-

sion. Convolution between a uniform distribution of random defect positions (**Fig. 3.15 (a)**) associated with different DCR levels (**Fig. 3.15 (b)**) and a number of defects Poisson-distributed in quantity (**Fig. 3.15 (c)**) results in the simulated cumulative plot (**Fig. 3.17**).

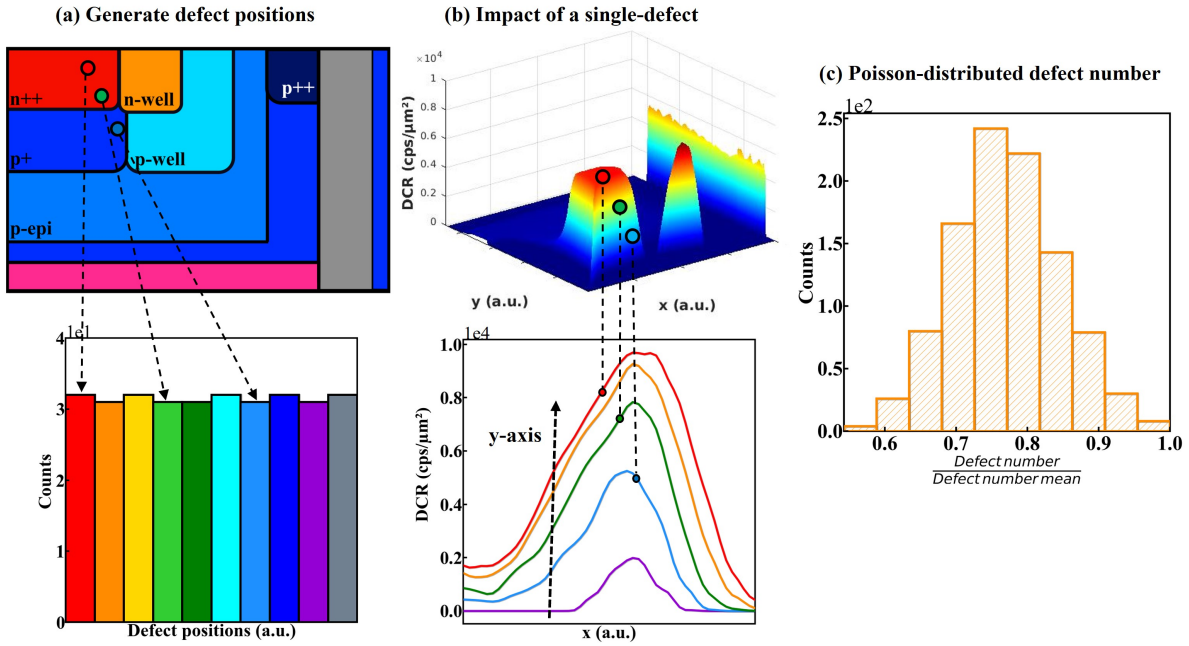


Figure 3.15: Suggested model of process-induced DCR statistical dispersion convoluting (a) the uniformly distributed defect positions (each bin represents 30 potential positions) within the device, (b) the impact of defect position on DCR (each color refer to the defect position reported on (a)) and (c) the defect number Poisson-distributed from device-to-device. The corresponding defect impact as a function of the position in the device with the position histogram are shown in the lower part of the figure. The upper part represents the cross section of the device where defects are drawn and a map of DCR within the device.

On one hand, for the intrinsic population inside the low-bias region, defect clusters randomly set along the side interface (4) together within the near-contact region (5) and within the multiplication region (1) catch this statistical dispersion as illustrated in **Fig. 3.17 (a)**. For the high-bias region, only defects randomly set at the edge of the wells and the rings, where the defects concentration is supposed to be high considering the multiple implant steps, reproduce this trend as enlighten in **Fig. 3.17 (b)**. On the other hand, the extrinsic population is replicated with defect randomly set within the wells (1) and the rings (2). The occurrence of implant damages during the subsequent ion implantation process step is suspected, as described by Bourgoin [90].

Setting the extrinsic condition as equal to three times the median DCR value result in **Fig. 3.18** for several architectures. When the p-well is brought closer to the n-well it increases the extrinsic population as well as the size of the junction. This means that crossing all the implantations at the same suspected localization spot rises the probability to find out with outliers. This provides further evidence to support the hypothesis of the occurrence of implant damages.

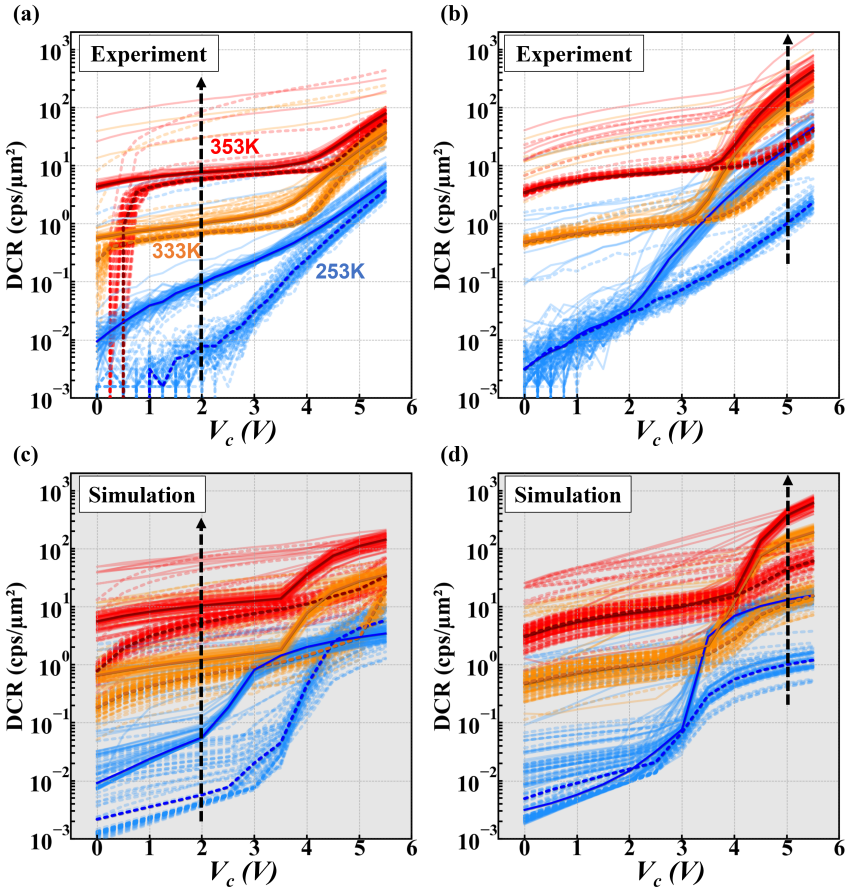


Figure 3.16: Measured and simulated DCR as a function of V_c at T_c=253K (blue), 333K (orange) and 353K (red): (a)(c) wells (SPAD 1 and 2) and (b)(d) rings (SPAD 3 and 4) from reference to variation (arrow direction). The bold lines are the medians. (a) and (b) are experiments and (c) and (d) are simulations. One can observe both intrinsic and extrinsic populations.

Conclusions

Monte-Carlo approach allows to capture the DCR intrinsic and extrinsic population. The statistical dispersion of the intrinsic population in the low-bias region is captured by defect clusters randomly positioned along the side interface, within the near-contact region, and within the multiplication region. In the high-bias region, only defects randomly positioned at the edge of the wells and rings exhibit this trend, where the defect concentration is expected to be high due to multiple implant steps. Conversely, the extrinsic population is reproduced with defects randomly positioned within the wells and rings.

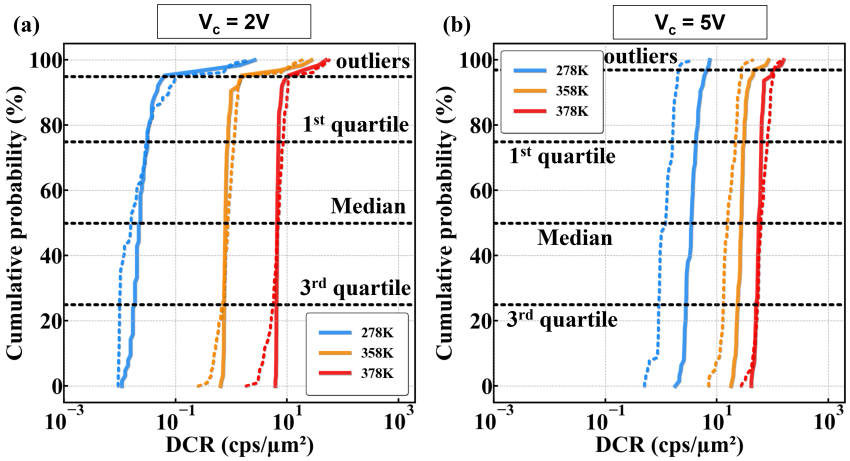


Figure 3.17: Cumulative probability as a function of DCR measured and simulated for SPAD 4 at (a) V_c=2V and (b) V_c=5V. Simulation are in dashed line and experiment in bold line. The measured and simulated data both exhibit a similar pattern, with two distinct groups: the intrinsic population, which includes data points from the 1st to the 3rd quartiles, and the extrinsic population, which contains the outliers.

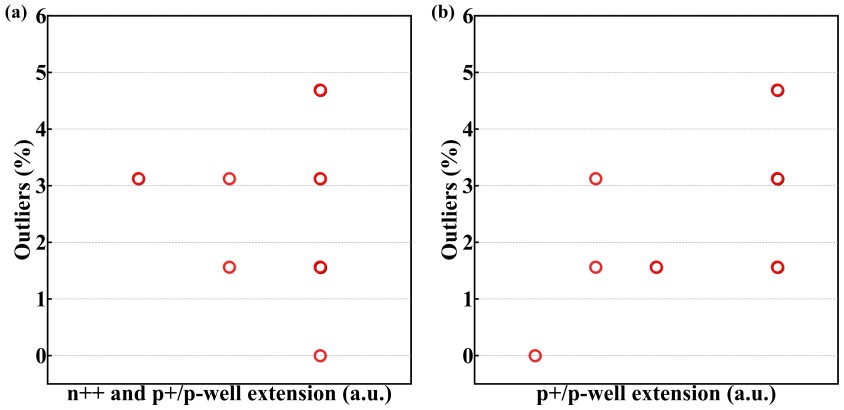


Figure 3.18: Outlier percentages for different architectures as a function of (a) the junction size and (b) the p+/p-well extension. Bringing the p-well closer to the n-well results in an increase in both the extrinsic population and the size of the junction.

4 Conclusions

1 The main challenge in DCR modeling is to determine the contribution of various generation mechanisms by obtaining data on process quality and physical parameters such as trap energies and BTBT parameters that are associated with the specific process under investigation. For accurate predictions, it is essential that the measured and simulated breakdown voltages match. DCR thermal generation is typically more sensitive to changes in temperature compared to DCR tunneling or field-enhanced generation, which exhibit a greater sensitivity to variations in electric field. The DCR can be inferred from thermal Shockley-Read-Hall (SRH) generation of carriers from defects, potentially enhanced by Trap-Assisted Tunneling (TAT). In addition, Band-to-Band Tunneling (BTBT) and diffusion mechanisms can also be at the origin of DCR.

2 The average DCR behavior is simulated by taking into account temperature and electric field effects in both the low- and high-bias regions. At low temperatures, the low-bias region is influenced by BTBT, while at higher temperatures, thermal SRH generation from the neutral region takes over. In the high-bias region, SRH-TAT from the peripheral n-/p-well region is the primary driving force. However, a simulation that relied on minority carrier lifetime modeled through dopant distribution failed to capture the high-bias regime and required the inclusion of a trap profile within the device.

3 Design of Experiment (DoE) allows to pinpoint the most critical factors and their interactions that affect DCR. Architecture constrains and ion implantation conditions shed light on the avalanche breakdown probability contribution to DCR rather than the modulation of defect quantity. Interface manufacturing underlines the role of interfaces in the origin of the low-bias rgion.

4 Monte-Carlo approach allows to capture the DCR intrinsic and extrinsic population. The statistical dispersion of the intrinsic population in the low-bias region is captured by defect clusters randomly positioned along the side interface, within the near-contact region, and within the multiplication region. In the high-bias region, only defects randomly positioned at the edge of the wells and rings exhibit this trend, where the defect concentration is expected to be high due to multiple implant steps. Conversely, the extrinsic population is reproduced with defects randomly positioned within the wells and rings.

Chapter 4

Investigating the origins of post-stress Dark Count Rate drift in SPADs

Objectives

In the previous chapter, DCR was examined using fresh device samples, but it was observed that exposure to voltage, irradiance, and temperature can cause stress and deviations in device performance, including variations in DCR. The main objective of this chapter is to gain valuable insights into the underlying degradation mechanisms. To achieve this, the chapter starts by providing an overview of metastable defects, which are known to play a crucial role in temperature, light, and voltage degradation mechanisms. A comprehensive set of experimental results is then presented to identify the specific degradation mechanism in the studied SPAD. Of particular interest is the investigation of the avalanche's contribution to the degradation process, achieved by varying the metal shield extension on the bottom interface to reduce the light aperture. By reducing the number of photons transmitted to the silicon, this technique enables a more precise investigation of the impact of the avalanche on the device's performance. To validate the localization of newly formed defects during stress conditions, various samples from different processes are utilized, adding an extra layer of consistency and reliability to the study's findings. Similar to **Chapter 3**, the investigation of degradation mechanisms begins by focusing on the average behavior of the devices under stress and later extends the analysis to encompass device-to-device behavior. Overall, this chapter provides a comprehensive understanding of the degradation mechanisms in the SPAD under studied.

Contents

1 Influence of metastable states on device degradation rate

Materials such as pure metals, semiconductors, and dielectrics generally do not exist in nature and must be fabricated into unnatural metastable states. Metastable states are states in which materials exist, but they are only apparently stable and will eventually change or degrade over time. The rate of degradation can be accelerated by subjecting the materials to elevated stress. Metastable states are described by the thermodynamic principles. In details, metastable states will tend to degrade spontaneously in order to reach a lower Gibbs free energy. The Gibbs free energy denotes the quantity of energy that a system can utilize to carry out work under constant temperature and pressure conditions. The change in Gibbs free energy, ΔG, is defined as [115]:

$$\Delta G = \Delta H - T \cdot \Delta S \tag{4.1}$$

where ΔH and ΔS are the change in enthalpy and in entropy respectively. ΔH is the heat energy absorbed or released during the chemical reaction and ΔS is a measure of the degree of disorder in the system. The second law of thermodynamics states that the entropy S (disorder) of isolated system will tend to increase with time. A negative value of Gibbs free energy indicates that the reaction is spontaneous and can proceed without the input of energy, while a positive value indicates that the reaction is non-spontaneous and requires energy input to proceed. Similarly, the Gibbs free energy of activation, ΔG^*, can be written as:

$$\Delta G^* = \Delta H^* - T \cdot \Delta S^* \tag{4.2}$$

where ΔH^* and ΔS^* are the enthalpy and entropy of activation respectively. The key difference between the change in Gibbs free energy ΔG and the Gibbs free energy of activation ΔG^* is that the former deals with the thermodynamic favorability of a reaction, indicating whether a reaction is spontaneous or not, while the latter is related to the kinetic aspects of a reaction, representing the energy barrier that must be overcome for a reaction to occur. The difference in free energy, shown in **Fig. 4.1 (a)**, ΔG, between the initial state (G_1) and the degraded state (G_3) is the driving force for the degradation, while the rate of the degradation reaction is limited by the activation energy, denoted by ΔG^* between the state resulting from the activation energy barrier (G_2). In other words, although the free energy difference ΔG is responsible for driving the degradation process, it is the activation energy ΔG^* that determines the rate at which the degradation occurs. The impact of stress on the free energy of activation ΔG^* and the free energy difference ΔG can be represented by a Maclaurin Series [115]:

$$\begin{aligned} \Delta G^* &\cong (\Delta G^*)_{\xi=0} + [\tfrac{\partial(\Delta G^*)}{\partial \xi}]_{\xi=0}\xi \\ &\cong \Delta G_0^* + a\xi \end{aligned} \tag{4.3}$$

and

$$\begin{aligned} \Delta G &\cong (\Delta G)_{\xi=0} + [\tfrac{\partial(\Delta G^*)}{\partial \xi}]_{\xi=0}\xi \\ &\cong \Delta G_0 + b\xi \end{aligned} \tag{4.4}$$

As shown in **Fig. 4.1 (b)**, external stress ξ can accelerate the degradation rate through two mechanisms: (1) by decreasing the free energy of activation difference ΔG^*, leading to a faster degradation rate, and (2) by increasing the free energy difference ΔG between the initial metastable state and the degraded state, resulting in an even more unstable state.

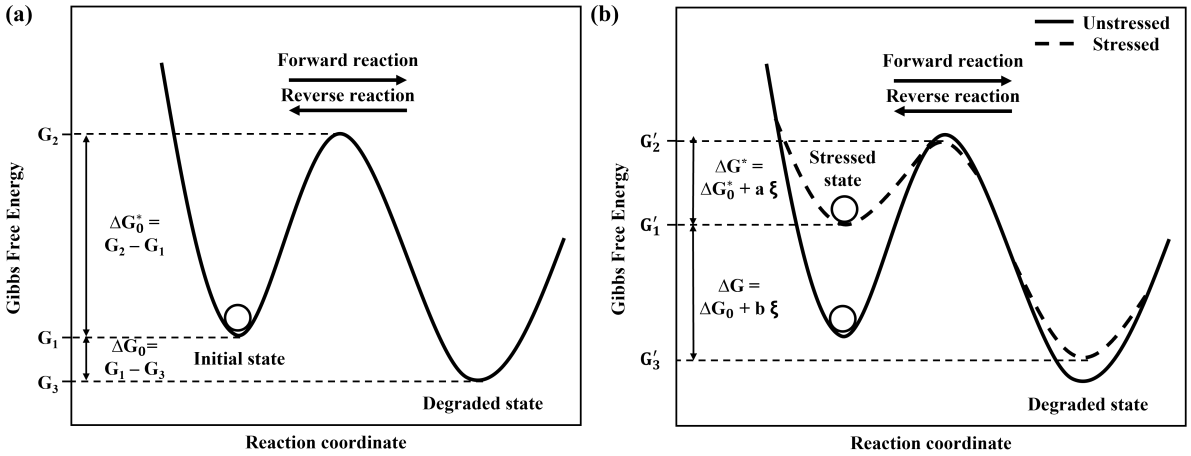


Figure 4.1: (a) The metastability of the initial state is due to the lower free energy of the degraded state, resulting in a difference in free Gibbs energy ΔG_0. The stability of the initial metastable state is affected by the activation energy ΔG_0^* required to transition from the initial metastable state to the degraded state. (b) The application of an external stress ξ on device impacts the free energy in two ways: (1) increasing the Gibbs free energy difference ΔG between the initial and degraded states, thereby making the initial metastable state even more unstable, and (2) it can reduce the activation energy ΔG^*, resulting in an accelerated degradation rate.

In analogy with chemical reaction rates, one can think of the degradation rate in terms of a degradation rate constant defined for the forward reaction:

$$\kappa_f = \kappa_0 \cdot exp(-\frac{\Delta G^*}{k_B T}) \tag{4.5}$$

and for the reverse reaction:

$$\kappa_r = \kappa_0 \cdot exp(-\frac{(\Delta G^* + \Delta G)}{k_B T}) \tag{4.6}$$

The net degradation reaction rate is given by:

$$\kappa_{net} = \kappa_f - \kappa_r = k_0 \cdot exp(-\frac{\Delta G^*}{k_B T})[1 - exp(-\frac{\Delta G}{k_B T})] \tag{4.7}$$

The free energy difference, ΔG, is what drives the net degradation rate. The net degradation rate is zero if the free energy difference, ΔG, is zero, regardless of the activation energy, ΔG^*. However, if the degraded state has a lower or higher free energy than the initial state, the activation energy, ΔG^*, becomes critical in retarding or accelerating the rate of the degradation

reaction.

When $\Delta G \gg k_B T$, which is typically the case in SPAD devices operating at high temperature, the reverse reaction rate becomes negligible and the net degradation reaction rate simplifies to:

$$\kappa_{net} = k_0 \cdot exp(-\frac{\Delta G^*}{k_B T}) \tag{4.8}$$

Similarly to the activation energy described for the carrier generation processes in **Chapter 2 Section 3.1.4**, the activation energy ΔG^* can be extracted as the derivative of the natural logarithm of κ_{net} with respect to $\frac{1}{T}$:

$$\Delta G^* = -\frac{d(ln(\kappa_{net}))}{d(\frac{1}{k_B T})} \tag{4.9}$$

The extraction of the activation energy ΔG^* can be achieved by measuring temperature stress, which is particularly important for identifying the potential degradation mechanism in devices like SPADs.

Conclusions

Metastable materials tend to degrade spontaneously in order to reach a state of lower Gibbs free energy, which takes into account the energy and disorder of the system. This degradation process is driven by a change in Gibbs free energy ΔG, and the rate of degradation is determined by the free energy of activation ΔG^*. Similar to a chemical reaction, it is possible to define a degradation rate constant κ to determine the activation energy ΔG^* for the degradation process. The theoretical framework for the occurrence of degradation enables the comprehension of SPAD aging over the duration of stress.

2 Overview of metastable defects in semiconductors

Previously, the degradation was theorized from a thermodynamic perspective. Now, the potential metastable defects leading to such degradation in SPAD devices are described. In chip manufacturing, silicon oxide SiO_2 and silicon nitride Si_3N_4 are utilized as primary insulators to safeguard the underlying silicon Si surface from oxidation and contamination. This protective layer ensures the reliable operation of the chip by preventing unwanted chemical reactions and electrical interference. Insulators have a large bandgap, which is the energy difference between the valence band (where electrons are tightly bound to the nucleus) and the conduction band (where electrons are free to move). Unlike conductors, insulators cannot conduct electrical current through the movement of free electrons because all their electrons are tightly bound to the nucleus. However, insulators may still have a low residual leakage current due to either inherent or metastable defects in the material, which can be caused by exposure to stress conditions such as changes in temperature, exposure to light, or application of voltage. These defects can create energy states within the bandgap that allow the flow of charge carriers, resulting in a small amount of current flow. As a result, the performance of the SPAD devices can be degraded,

leading to an increase in DCR, which is directly related to the number of charges generated by these defects. By understanding how defects affect the energy bands and electronic properties of the material, one can identify specific degradation mechanisms and explore approaches to limit or counter these negative effects.

2.1 Dangling bond: the case of borderless nitride

The hydrogenated amorphous silicon nitride film ($a-SiN_x:H$) called borderless nitride is obtained through Plasma-Enhanced Chemical Vapor Deposition (PECVD). It is deposited during the manufacturing process after the SiProt layer. SiProt is a protective layer used to prevent the photodiode from being silicided during the contact creation step, as silicide is reflective and would prevent light rays from reaching the SPAD. SiProt is composed of two depositions. First, a Tetra Ethyl Ortho Silicate (TEOS) oxide layer is deposited by Low Pressure Chemical Vapor Deposition (LPCVD) to obtain a good quality $Si-SiO_2$ interface, then a stoichiometric silicon nitride (Si_3N_4) is deposited also by LPCVD.

The evolution of the bandgap width of SiN materials is a function of the N/Si ratio [116]. The gap opening occurs gradually up to N/Si ratios of about 1.1 and rapidly beyond as shown in **Fig. 4.2**. The value of N/Si=1.1 corresponds to the percolation threshold of Si-Si bonds in the $a-SiN_x:H$ network. When N/Si<1.1, the band edges are formed by $Si-Si$ states. With increasing nitrogen concentration, the $Si-Si$ bonds, localized at the top of the valence band, are replaced by $Si-N$ bonds. Thus, when the N/Si ratio >1.3, the valence band edge is formed by $Si-N$ states while the conduction band edge is composed of anti-bonding $Si-Si$ states. For compositions such that 1.1<N/Si<1.3, the properties of the bandgap are intermediate.

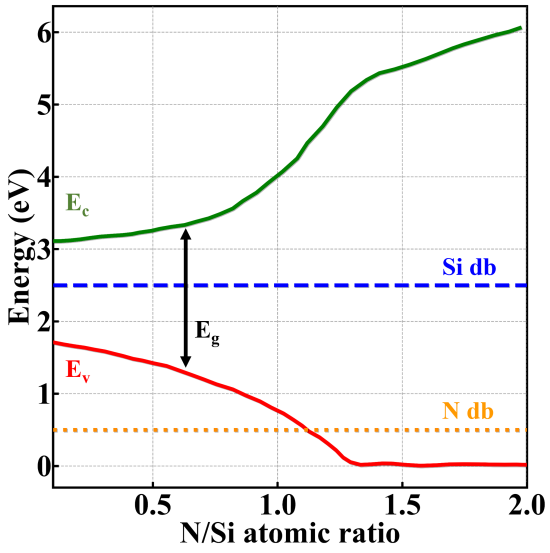


Figure 4.2: Bandgap width of SiN materials as a function of the N/Si ratio and position of Si and N dangling bonds, denoted Si db and N db respectively, from data reported in [116].

In the presence of hydrogen, $Si-H$ and $N-H$ bonds tend to replace $Si-Si$ and $Si-N$ bonds, influencing in some cases the shape of the bands. Indeed, for silicon-rich compositions with N/Si<1.1, $Si-Si$ bonds are replaced by $Si-H$ bonds. Since the energy level of the

latter is deeper, this lowers the valence band level by 0.8eV and increases the gap energy [116]. Therefore, the energy gap of 5.3eV relative to the stoichiometric composition of silicon nitride (N/Si=1.3) is reached for N/Si values greater than 1.5. These electronic properties depend on the metastable defects present in the material.

The different types of defects in silicon nitride are either related to the incorporation of impurities during fabrication or to its structure, in which case they are referred to as dangling bonds. These dangling bonds can be either silicon or nitrogen as shown in **Fig. 4.3** and represent the main defects in silicon nitride [116][117].

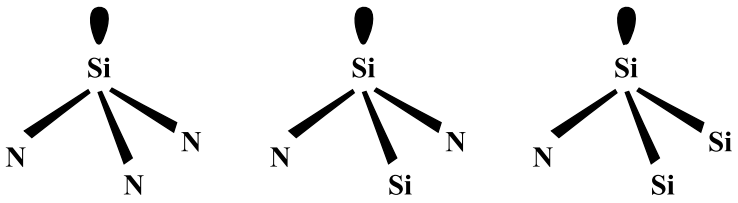


Figure 4.3: Atomic configuration of dangling bonds for SiN materials.

When three nitrogen atoms bond to one silicon atom, it is generally referred to as a K center [118]. The K center has an unpaired spin, making it paramagnetic. It is an amphoteric trap that can become positively charged by bonding with a neighboring silicon atom and losing its unpaired electron, or negatively charged by trapping an electron [119].

By analogy with K centers, nitrogen dangling bonds are referred to as N centers when the nitrogen atom is bonded to two silicon atoms [120]. These centers are also amphoteric and are characterized by a negative effective correlation energy.

Silicon dangling bonds have two energy levels in the middle of the gap [116], while nitrogen dangling bonds have one located near the valence band [121]. The energy levels of these two types of dangling bonds are independent of the composition of silicon nitride [116]. Due to the evolution of the energy gap with the Si/N ratio, nitrogen dangling bonds become electrically active for silicon-rich compositions, while silicon dangling bonds are active regardless of stoichiometry.

Under its amorphous phase, silicon nitride contains hydrogen which passivates dangling bonds and allows for the release of stresses imposed on the microstructure. It is noted as $a - SiN_x : H$ where x is the ratio between the concentration of nitrogen and the concentration of silicon. The bonds formed with hydrogen have energy states outside of the forbidden energy band and are therefore electrically inactive. Thanks to this property, hydrogen is used as a passivating element and its effectiveness would be greater on silicon dangling bonds than on nitrogen ones [122]. It is now accepted that the defect density of silicon nitride films is related to the mobility of hydrogen [123]. However, the presence of this element can lead to a degradation of the insulating properties of the material. Indeed, by desorbing from the bonds it passivates, new dangling bonds can be created, which serve as R-G centers for carriers.

2.2 Lattice mismatch: the case of Si/SiO_2 system

Between adjacent devices, both electrical interference and optical crosstalk can be reduced by using SiO_2 during the implementation of the Deep Trench Isolation (DTI) in SPADs. This

layer serves as an insulator, preventing the passage of electrical current. Additionally, due to its high refractive index, SiO_2 acts as an optical barrier, reflecting light and further reducing crosstalk. One of the issues with the Si/SiO_2 system is the lattice mismatch between Si and SiO_2. Silicon is a tetrahedral atom with a crystalline structure, while silicon oxide (SiO_2 or silica) is amorphous and has a basic building block of an Si atom surrounded by four oxygens (SiO_4) and a $Si-O$ interatomic distance smaller than the $Si-Si$ distance as shown in **Fig. 4.4**. This lattice mismatch results in an interfacial layer of a few atomic layers [124], from 4 to 5Å [125], allowing for a gradual transition from 0 to 2 oxygen atoms per silicon atom. This transition generates defects at the interface and in the oxide, which act as R-G centers for carriers.

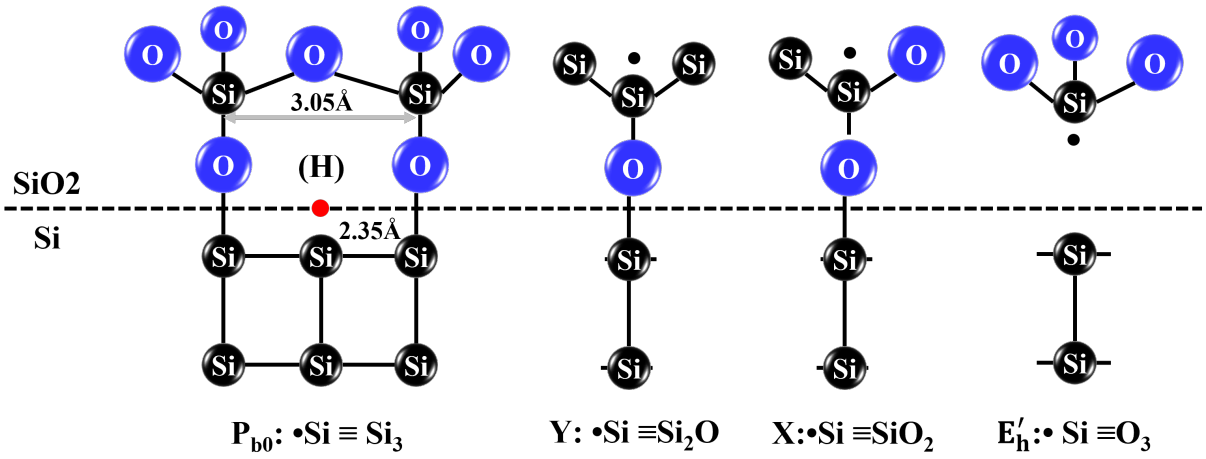


Figure 4.4: Schematic representation of the centers: P_{b0}, Y, X and E'_h, which are described in the text below.

2.2.1 Interface traps

To start, let's focus on the defects at the interface. Due to lattice mismatch, some Si atoms in the substrate bound to three Si atoms cannot form a bond with an oxide atom and therefore have a dangling bond. They were first identified by Electron Spin Resonance (ESR) [126] and named P_b centers (P for paramagnetic and b for resonance site indexing). Two types of P_b defects have since been distinguished: P_{b0} and P_{b1}. The former corresponds to the chemical formula $\bullet Si \equiv Si_3$ [126], while it has more recently been confirmed by calculations that P_{b1} centers are of the type $Si_2 = Si \bullet - Si \equiv Si_2O$ [127]. These dangling bonds of interface defects reduce the interface quality and, consequently, the performance of the device, such as mobility. To reduce the amount of interface defects, a passivation step of P_b centers with hydrogen is required: this involves a post-oxidation annealing under a hydrogen atmosphere [128]. Heating allows for the reorganization of the interface and diffusion of hydrogen towards the interface, with a reaction on the P_b centers to form $Si_3 - SiH$ groups.

2.2.2 Border traps

Defects known as border traps are located within a distance of 3nm on both sides of the interface. As one moves closer to the interfacial layer, they are due to oxygen sub-stoichiometry during the gradual transition from $Si - Si_4$ to $Si - O_4$. Their structure is therefore of the type $\bullet Si \equiv SinO_3 - n$ (n=0, 1, 2) with the $\bullet$ symbolizing the dangling bond resulting from the fact that the central Si forms a bond with only three atoms instead of four, thus leaving the fourth valence electron unpaired [129]. The $-$ symbol in the contracted formula represents a covalent bond, so $\equiv$ means that the central Si is connected to three other atoms, namely nSi atoms and $(3-n)O$ atoms. The paired electron of X and Y centers is delocalized in the atomic orbitals of Si, so it is less strongly bound than in the case of a localized electron as for the E'_h center. X and Y centers can therefore more easily lose their electron than the E'_h center. The study of the electron affinity of X and Y centers has shown that they more easily trap holes than electrons, so they are donors. They can thus play the role of precursor to oxide defects that trap and release carriers known as switching oxide traps [130].

2.2.3 Oxide traps

As one moves away from the interfacial layer, defects appear related to oxygen vacancies or interstitial atoms [131] such as Oxygen Vacancy (OV) $O_3 \equiv Si - Si \equiv O_3$, peroxide bridge $O_3 \equiv Si - O - O - Si \equiv O_3$, non-bonding oxygen $O_3 \equiv Si - O\bullet$, trivalent silicon (E') $O_3 \equiv Si\bullet$ and divalent silicon $O_2 = Si\bullet\bullet$. The most common defects are oxygen vacancies and E' centers as shown in **Fig. 4.5**. The defects that involve two silicon atoms with an unpaired electron localized on one of the two Si atoms, which is bound to three atoms and a hole on the other Si atom spatially separated from the first, are called E' centers [130]. Two primary E' centers, E'_δ and E'_γ, can be distinguished depending on the environment of the defect. E'_δ appears in a small, constrained structure where the electron and hole are equally shared between the two Si atoms. On the other hand, E'_γ is associated with a larger and more deformable structure where the paired electron is entirely localized on the other Si atom, and the center is positively charged. E'_δ centers account for 90% of E' centers, with 80% of them being shallow traps that cannot trap electrons [132]. However, 20% of them are electron traps with energy close to mid-gap. E'_γ centers represent 5% of E' centers [132] and can easily trap and release a hole, making them switching oxide traps [130]. E'_γ centers are associated with a dipole complex [133][134] and become metastable due to Coulombic repulsion between the two electrons located in a restricted volume.

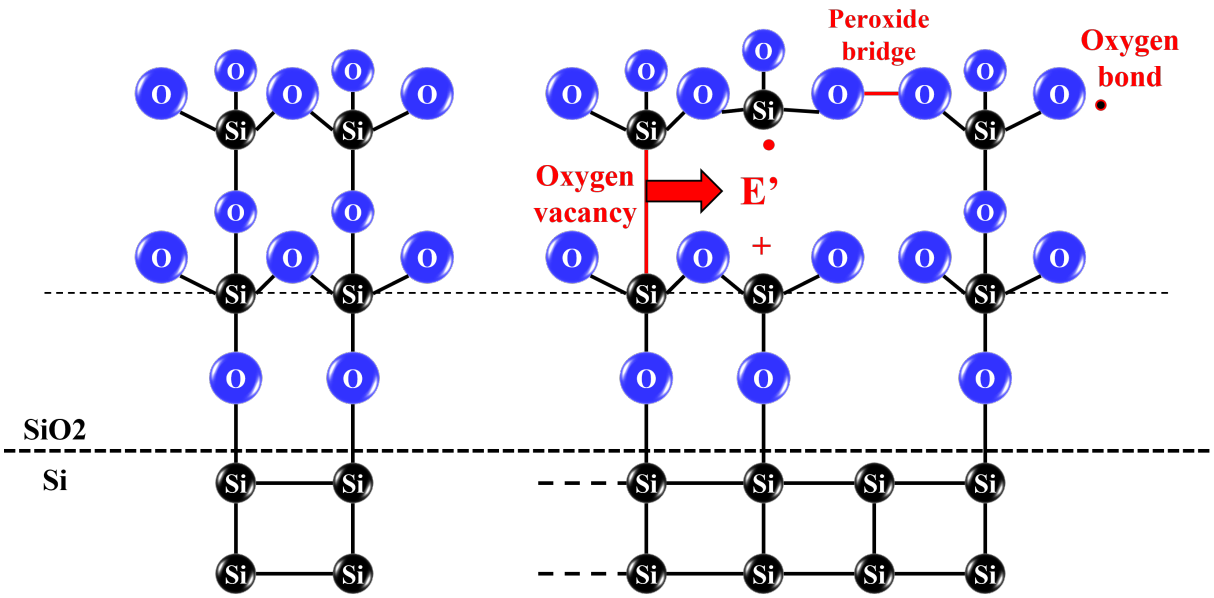


Figure 4.5: Schematic representation of the ideal Si/SiO_2 lattice and the lattice with a mismatch, including the oxygen vacancy and E' center.

Conclusions

Borderless nitride, which is a hydrogenated amorphous silicon nitride film ($a-SiN_x:H$), can adopt several crystallographic configurations depending on the N/Si ratio. In the presence of hydrogen, $Si-H$ and $N-H$ bonds tend to replace $Si-Si$ and $Si-N$ bonds. The main defects in silicon nitride are dangling bonds, which can be either silicon or nitrogen. When three nitrogen atoms bond to one silicon atom, it is generally referred to as a K center, and nitrogen dangling bonds are referred to as N centers when the nitrogen atom is bonded to two silicon atoms. Both K and N centers are amphoteric traps, either acceptor or donor.

The Si/SiO_2 system experiences a lattice mismatch between Si and SiO_2, resulting in defects at the interface and in the oxide. These defects are commonly referred to as P_{b0} and P_{b1}, and they require passivation with hydrogen. Border traps are found within a 3nm range on both sides of the interface, which is caused by oxygen sub-stoichiometry during the gradual transition from $Si-Si_4$ to $Si-O_4$. This leads to the formation of X and Y centers, as well as E'_h centers that are located close to the interfacial layer. Additionally, there are Oxygen Vacancy (OV) and E' centers that are located further from the interfacial layer. These defects act as R-G centers for carriers.

3 Overview of degradation mechanisms of interest in semiconductors

This section does not focus on the degradation of SPADs specifically, but rather aims to provide an overview of the degradation mechanisms that are typically observed in generic CMOS architectures. These mechanisms could also be relevant to SPADs and may be discussed in details later on.

3.1 Bias Temperature Instability

Bias Temperature Instability (BTI) is due to the electric field applied to the gate oxide layer of the MOS device at elevated temperature [135]. Negative Bias Temperature Instability (NBTI) occurs in PMOS when the gate voltage V_g is less than 0, while Positive Bias Temperature Instability (PBTI) occurs in NMOS when V_g is greater than 0. In PMOS, the electric field is directed from the gate to the substrate, and when the gate voltage V_g is negative, it causes the electrons in the oxide to tunnel into the substrate. On the other hand, in NMOS devices, the electric field is directed from the substrate to the gate, and when the gate voltage V_g is positive, it causes the holes in the oxide to tunnel into the substrate. BTI has a permanent component illustrated in **Fig. 4.6** due to the generation of defects at the SiO_2 interface which can occur through several mechanisms, including hole trapping in $Si-H$, $Si-H$ bond breakage, and hydrogen transfer into $Si-O$, which is assisted by holes. These mechanisms can result in the formation of amphoteric interface traps, which can have a positive, neutral, or negative charge. The presence of positive charges close to the interface can be mitigated by fixing them in place.

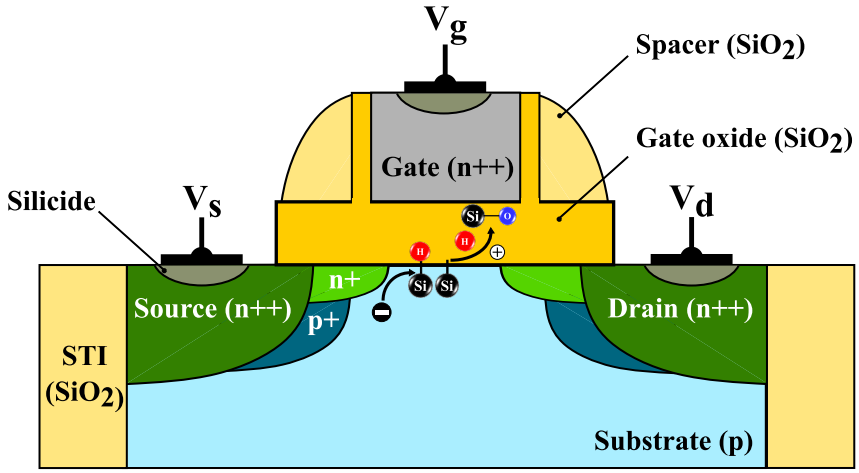


Figure 4.6: Schematic diagram illustrating a Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) and the permanent phenomenon of Bias Temperature Instability (BTI). Defects can be generated at the SiO2 interface through various mechanisms, including hole trapping in Si-H, Si-H bond breakage, and hydrogen transfer into Si-O, which is facilitated by holes.

BTI has also a recoverable component illustrated in **Fig. 4.7** due to defect generation which can also occur in the bulk of SiO_2 through several mechanisms, such as hole capture/emission from oxygen vacancies leading to the formation of E′ centers, as well as hole capture/emission

from K centers and hydrogen retro diffusion, which can passivate the interface (see **Section 2.2**). The defect signature in this case is that the defects are recoverable, meaning that releasing the stress can result in relaxation, even at low temperatures.

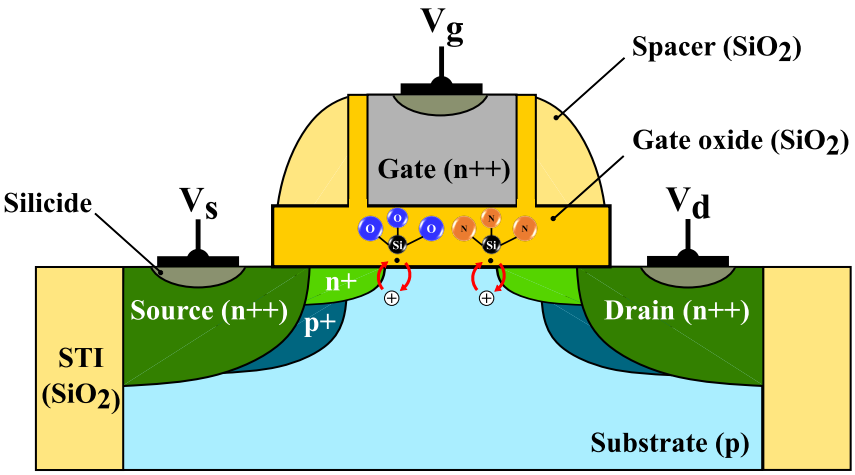


Figure 4.7: Schematic diagram illustrating a Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) and the recoverable phenomenon of Bias Temperature Instability (BTI). Defect generation can also happen in the bulk of SiO_2 through various mechanisms, including hole capture/emission from oxygen vacancies that result in the formation of E′ centers, and hole capture/emission from K centers and hydrogen retro diffusion, which can passivate the interface.

3.2 Hot carrier degradation

When a voltage is applied between the source and the drain (V_{ds}) of a Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET), the electric field accelerates the charge carriers, which can reach high energies depending on the applied bias. Hot-Carrier Degradation (HCD) in MOSFETs is a phenomenon that occurs when these high-energy carriers cause damage to the MOSFET's gate oxide layer and result in a reduction in the device's performance and reliability over time. In HCD of MOSFETs, illustrated in **Fig. 4.8**, there are three types of traps [136][137], as described in **Section 2.2**, that can cause damage to the gate oxide layer: interface traps, border traps, and oxide traps. Interface traps are located at the interface between the gate oxide and the semiconductor material. Border traps are located at the border between the gate oxide and the source/drain regions of the MOSFET (3nm on either side including E′ centers). Oxide traps are located within the gate oxide layer itself. The field dependency of HCD is driven by two fundamental components: carrier density through V_{ds} and carrier energy through V_g. In **Chapter 5 Section 1**, a detailed explanation of the physical mechanisms behind HCD will be provided.

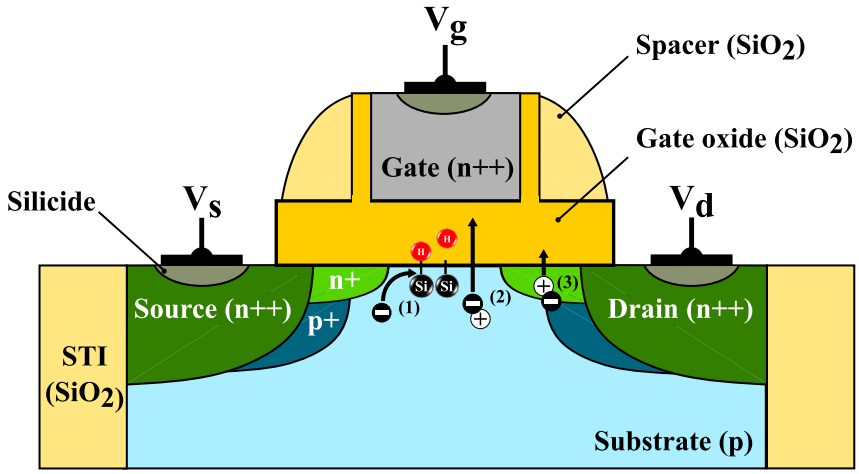


Figure 4.8: Schematic diagram illustrating a Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) and the phenomenon of Hot-Carrier Degradation (HCD). The HCD leads to three types of traps: (1) interface states, which are defects generated at the Si/SiO_2 interface when $Si-H$ bonds break under carrier collision, (2) oxide traps, which occur when high-energy carriers penetrate deeper into the oxide and become trapped, both mechanisms occur within a p-doped layer, and (3) border traps, where the charge is trapped in the oxide/spacer, the latter process takes place within a n-doped layer.

Conclusions

From MOSFET perspective, the key parameters are V_{ds}, which represents carrier energy, and V_g, which represents carrier density and oxide field. The permanent component of BTI involves hole trapping in $Si-H$, $Si-H$ bond breakage, and hydrogen transfer into $Si-O$, which is facilitated by holes. On the other hand, the recoverable component of BTI includes hole capture/emission from oxygen vacancies, leading to the formation of E′ centers, as well as hole capture/emission from K centers and hydrogen retro diffusion, which can passivate the interface.

The physical mechanism underlying HCD can be broken down into three main steps. Firstly, electron defect generation occurs at the Si/SiO_2 interface through $Si-H$ bond breakage. Secondly, electron trapping takes place in the oxide, leading to bulk defect generation. Finally, charge trapping occurs in the oxide/spacer. The field dependency of HCD is driven by two fundamental components: carrier density and carrier energy, which will be further explained in **Chapter 5 Section 1**.

4 Existing Litterature on degradation mechanisms in SPADs

The study reported in [138] involves an analysis of degradation by measuring the DCR over stress time and the current-voltage (I-V) characteristics at various stages of aging of InP-based Geiger-mode avalanche photodiode arrays, as illustrated in **Fig. 4.9**. Initially, the DCR remains relatively stable for the first 20 hours, while the dark current experiences a significant four-order-of-magnitude increase. Subsequently, after 20 hours of aging, the DCR begins to escalate

rapidly, with the dark current showing only moderate continued growth. This pattern suggests that the rising dark current may primarily be a surface leakage along the device sidewalls. In such a scenario, electrons could travel along the device's perimeter without passing through the high-field avalanche region, thus avoiding triggering an avalanche event.

During operation, ions or charges might accumulate along the perimeter of the device. With time, these accumulations could reach a critical quantity, establishing a pathway for charge flow along the device's perimeter and causing a significant surge in leakage current. Initially, surface conduction poses no interference with Geiger-mode operation, as the carriers do not penetrate into the device's avalanche region to trigger a breakdown event. However, as the charge accumulates on the surface, it reaches a point where it bends the electric field within the device. This bending electric field starts drawing carriers into the avalanche region, causing a sudden and substantial increase in the DCR, even with minimal further increases in dark current.

To delve deeper into this degradation mechanism, a high-temperature bake was conducted on a deteriorated device to observe its impact on the DCR. The device reverts to its initial pre-aged DCR and sustains it for a brief duration before undergoing rapid degradation once more. This experiment provides additional evidence in favor of the theory that ions accumulate on the device's sidewall and diffuse somewhat during high-temperature bakes. Although their movement during subsequent rounds is limited, as indicated by the more rapid onset of degradation, the charges disperse far enough to have a negligible effect on the fields within the device. As a result, the device returns to its initial DCR before the onset of degradation. The insights derived from these results will contribute to the subsequent degradation analysis of the studied SPAD.

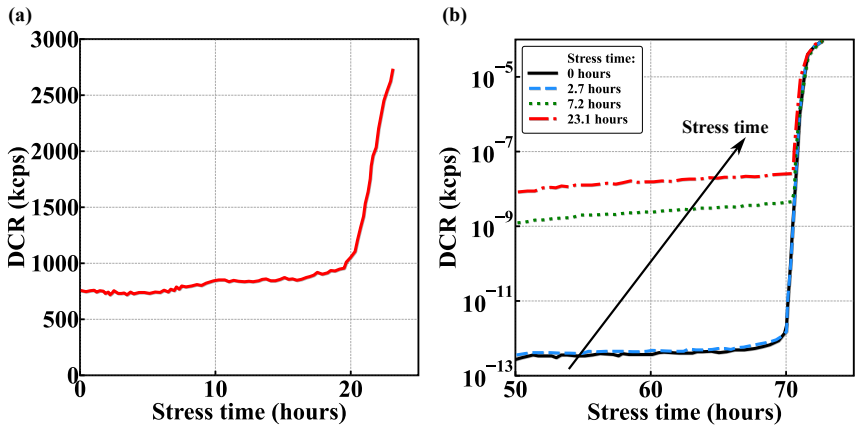


Figure 4.9: (a) DCR as a function of stress time and (b) dark current as a function of voltage at different stress times. The dark current graphs show that the leakage significantly increases by several orders of magnitude before any rise in Dark Count Rate (DCR) is detected.

5 Partitioning the impact of stress

The main results of this section have been published in M. Sicre, X. Federspiel, D. Roy, B. Mamby, C. Coutier and F. Calmon, "Identification of stress factors and degradation mechanisms inducing DCR drift in SPADs," 2022 IEEE International Integrated Reliability Workshop (IIRW), 2022, pp. 1-5. [139].

The DCR must be guaranteed over time as a function of the operating conditions to ensure distance range integrity as it will be explained in **Chapter 6**. To address this concern, this section assesses the principal stress factors, namely voltage, irradiance, and temperature, to identify potential degradations responsible for DCR drift (ΔDCR) in SPADs as shown in **Fig. 4.10**. Each stress in **Table 4.1** is performed on 8 devices stressed for 24 hours with each row representing a stress with the conditions applied together. The measured DCR drift (ΔDCR) is higher at high-bias than at low-bias as illustrated in **Fig. 4.10**. The low-bias region presented similar degradation for the same stress conditions, albeit lower in magnitude. As an initial approach, the device is measured after each stress period at a temperature of $T_c = 333K$ and an excess voltage bias of $V_c = 5V$ for all subsequent stress results.

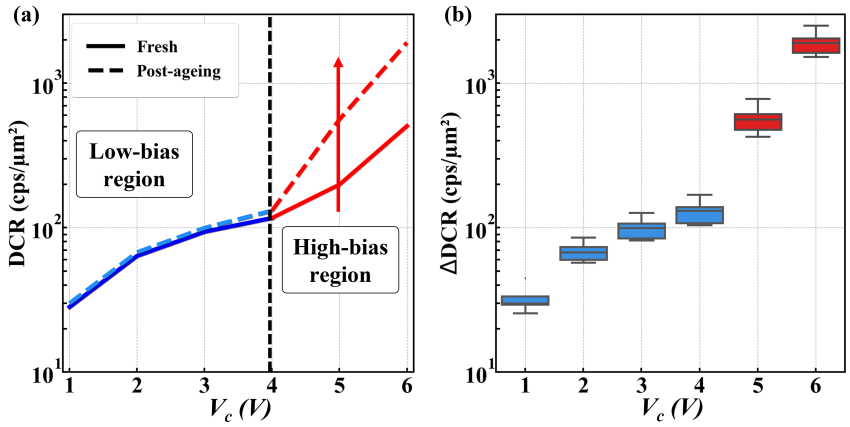


Figure 4.10: (a) Median DCR of 8 SPADs as a function of V_c for initial (bold line) and post-ageing (dashed line) measurements at T_c=333K for a 24-hours stress at T_s=353K, V_s=5V and an irradiance of $0.33W/m^2$. (b) Box plots of the stress resultant ΔDCR versus V_c. The top of the box represents the 75th percentile, the bottom the 25th percentile, and the line in the middle represents the median. The whiskers represent the highest and lowest value.

To capture the effect of each stress on ΔDCR, High Temperature Operating Lifetime (HTOL) experiments (see JEDEC JESD22-A108 standard test method) at various voltage and irradiance stress conditions are performed at a fixed temperature. Afterwards, the temperature stress is studied alone through High Temperature Storage (HTS) experiments (see JEDEC JESD22-A103 standard test method). The absolute ΔDCR from the subtraction of the post-ageing $DCR(t)$ and initial $DCR(t_0)$ is used to compare the different stress conditions:

$$\Delta DCR(t) = DCR(t) - DCR(t_0) \tag{4.10}$$

Table 4.1: Stress conditions to identify degradation localization

Index	T_s (K)	V_s (V)	Irradiance (W/m^2)
(1)	353	-3/4/5	0.0
	423	5	
(2)	353/373/423	-3	0.33
	353	-3	0.01
(3)	353	4	0.1/0.33/0.43
(4)	353	5	0.08/0.33/0.42
	373		0.22
(5)	353/373/423	None	0.0

Through this subtraction, the population of defects that are activated during the stress phase could be specifically studied by isolating them from the initial defect population represented by $DCR(t_0)$.

5.1 Analysis of voltage and temperature degradation dynamics

5.1.1 Voltage effect

Fig. 4.11 (a) displays the relationship between ΔDCR and V_s, which exhibits two exponential patterns: a smoother one at the low bias (LB) and a sharper one at the high bias (HB). These voltage dynamics are consistent with the pre-stress DCR observed in Chapter 3, where carrier multiplication occurs in the n++ and p+ central region for LB, and in the n- and p-well peripheral region for HB. This exponential voltage dependence indicates possible Hot-Carrier Degradation (HCD) through electron-hole pair impact ionization. As such, the study conducted by [140] assumed that the injection of Drain Avalanche Hot-carriers (DACH) was due to impact ionisation occurring at the drain of the MOS device. A detailed explanation of this mechanism will be provided in **Chapter 5 Section 1**. At high irradiance level ($1W/m^2$), where the current derives mainly from photo-generation, and far below V_{BD} (where few carrier multiplications occur), the ΔDCR is overwhelmingly influenced by the high volume of photo-generated carriers, and is therefore subject to constant degradation as indicated by the triangles in **Fig. 4.11 (a)**.

In **Fig. 4.11 (b)**, with (V_s=4V and 5V) or without (V_s=-3V) avalanche triggering, ΔDCR versus irradiance follows a linear relationship. From these results, the same ageing mechanisms can be assumed to relate to the number of available hot carriers, regardless of avalanche triggering. Furthermore, the distributions of the observed ΔDCR between the 8 devices used for each stress scheme show relatively tight distributions which suggest that the diode degradation relates more to intrinsic design factors rather than arbitrary defectivity. These degradation dependencies on the number of carriers with and without avalanches will be explained in the following paragraphs and in **Chapter 5 Section 2.5** based on measurement and modeling of electrical current.

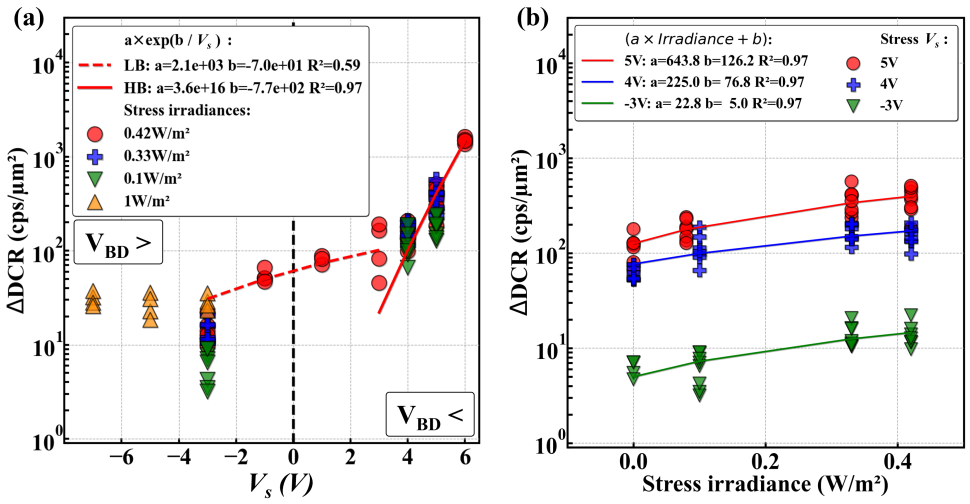


Figure 4.11: (a) ΔDCR measured at T_c=333K and V_c=5V as a function of the stress V_s below and above V_{BD} for different stress irradiances. (b) ΔDCR as a function of the stress irradiances for different stress voltages V_s. The degradation's exponential voltage and linear irradiance dependencies are linked to carrier multiplication and quantity, which are associated with Hot-Carrier Degradation (HCD).

5.1.2 Thermal effect

ΔDCR increases smoothly with temperature stress in darkness without any applied stress voltage as shown in **Fig. 4.12**. The magnitude of ΔDCR at T_s=353K does not explain ΔDCR values obtained by voltage stress below V_{BD} at the same temperature (**Fig. 4.11**), which exhibit higher magnitudes at shorter stress duration. This degradation is probably due to charge accumulation during stress at the borderless nitride or Contact Etch Stop Layer (CESL) [141], which is a hydrogenated amorphous silicon nitride film ($a-SiN_x:H$) formed during manufacturing at the upper interface as explained in **Section 2.1**. At such temperatures, ion migration [142] does not occur.

5.2 Analyzing avalanches as potential stress factor

To evaluate the potential degradation role of avalanches during stress, the ΔDCR is studied as a function of the number of avalanches triggered in darkness and under irradiance. **Fig. 4.13** shows that there is no correlation between the number of avalanches at stress in darkness and the ΔDCR at a given stress voltage. This remains true regardless of whether the avalanche localization during stress develops in the center (V_s=3V) or at the periphery (V_s=6V) of the device.

Some trials with reduced light aperture, by means of metal shield extension on the bottom interface, to lower the number of photons transmitted to the silicon are reported in **Fig. 4.14**. Initial DCR increases as the metal shield is extended as depicted in **Fig. 4.14 (a)**. The ΔDCR difference between post-ageing and initial samples can therefore be expected to decrease as the aperture is reduced. Indeed, the initial DCR will be higher for smaller aperture size, which leads

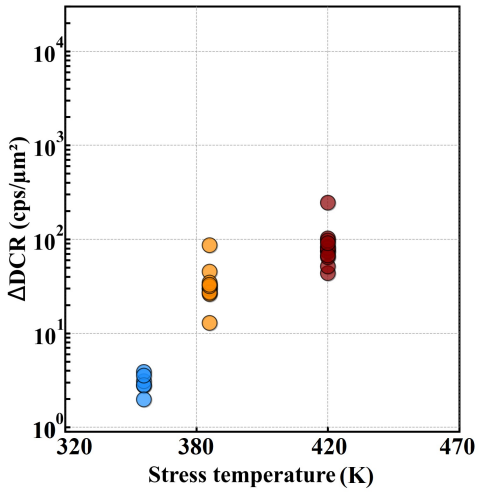


Figure 4.12: ΔDCR measured at T_c=333K and V_c=5V as a function of the stress temperatures T_s measured after 80h of stress without any applied stress voltage. Each color represents a different stress temperature. Instead of ion migration [142], charge accumulation occurs during stress at the borderless nitride or Contact Etch Stop Layer (CESL) [141] at such temperatures.

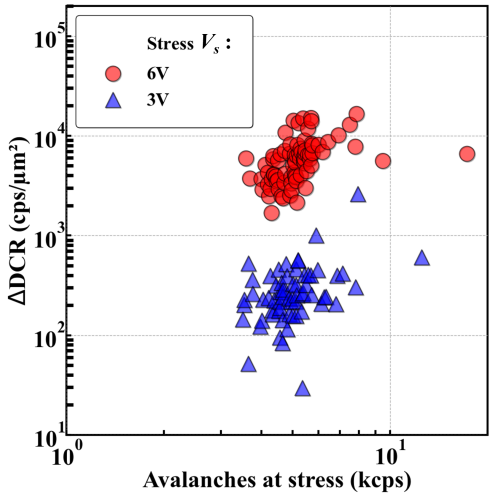


Figure 4.13: ΔDCR measured at T_c=333K and V_c=5V as a function of the number of avalanches measured before stress normalized per unit photo-sensitive area for 24-hour stress voltages V_s of 3V and 6V at 333K in darkness for 64 SPADs. There is no correlation between the number of avalanches during dark stress and the ΔDCR at a given stress voltage.

to smaller ΔDCR from subtraction projection and lower number of impinging photons as well as number of avalanches, which initially seem to represent less stressing condition. This is not the case, however: the lower the aperture size, which is to say the number of avalanches at stress, the higher the ΔDCR as illustrated by **Fig. 4.14 (b)**). This means that the degradation is much higher for smaller aperture sizes because they have higher initial DCR. Therefore, there is no correlation with avalanches triggered by irradiance either. This can be instead linked to the extended metal shield that embrittles hydrogen, preventing its diffusion to the upper interface during subsequent fabrication processes, which in turn increases the amount of dangling bonds and the degradation speed of the Si/SiO_2 interface as described in **Section 2**.

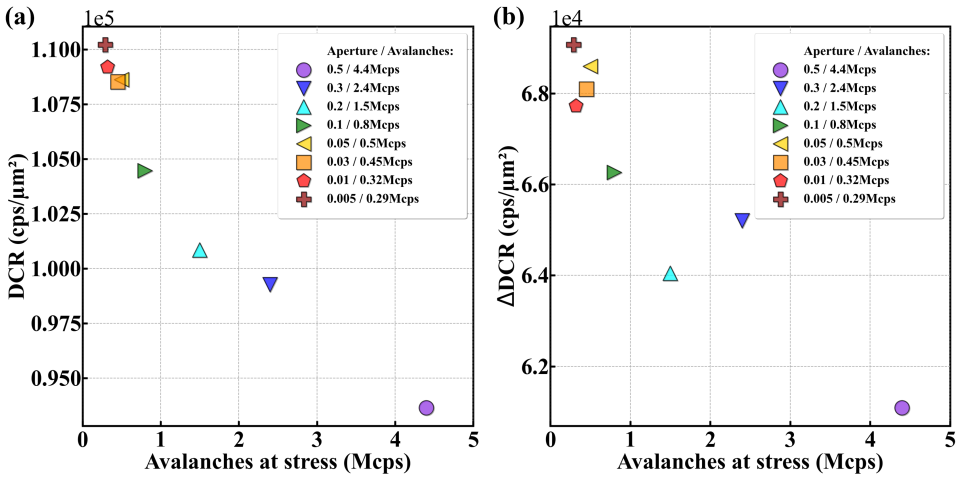


Figure 4.14: (a) Median of the initial DCR measured at T_c=333K and V_c=5V as a function of the number of avalanches measured measured at T_s=353K and V_s=3V before stress for different light aperture reductions with respect to reference architecture, each consisting of 64 SPADs illuminated at the same irradiance of $0.02W/m^2$. (b) ΔDCR measured at T_c=333K and V_c=5V as a function of the number of avalanches before stress for different light aperture reductions for the avalanche conditions reported in x-axis and for stress voltages V_s of 3V at T_s=353K. The lower the aperture size, which is to say the number of avalanches at stress, the higher the ΔDCR.

5.3 Interpreting the temporal dynamics of stress

To study the dynamics of the degradation mechanisms over time, the time accelerating factors are extracted from HTOL experiments at a given voltage (**Fig. 4.15 (a)**) and a given irradiance (**Fig. 4.15 (b)**) at the fixed stress temperature of 353K. Assuming ΔDCR is equal to a constant multiplied by a power of the stress duration ($\Delta DCR = A \cdot t_{stress}^{n}$), as it will be explained in **Chapter 5 Section 1**, a typical stress time power n of 0.6 is extracted for all the stress condition coupling voltages and irradiance levels. This stress-resilient time power provided insight into the occurrence of the same degradation mechanisms. The constant in front of the stress time is therefore dependent on the irradiance and the voltage, which is related to the increase of the number of hot carriers as explained in **Section 5.1.1** and in the next section.

5.4 Device ageing induced by electrical current

From the previous stress results, one of the main degradation assumptions relates to the number of hot-carriers through the increase of current without any influence of avalanches triggered during stress (see **Section 5.2**). To confirm this trend, the current is measured on 16 SPADs below V_{BD}, in order to prevent the excessive avalanche current, as a function of both V_c and irradiance. Both dark- and photo-generated current grows exponentially with the inverse of the V_c (**Fig. 4.16 (a)**), while the photo-generated current increases linearly with the irradiance (**Fig. 4.16 (b)**). The photo-generated current follows a similar linear relationship with respect to the irradiance as it does for ΔDCR (**Fig. 4.11 (b)**). This suggests that the number of hot carriers is the determining factor, rather than the kinetic energy brought to the carriers by the

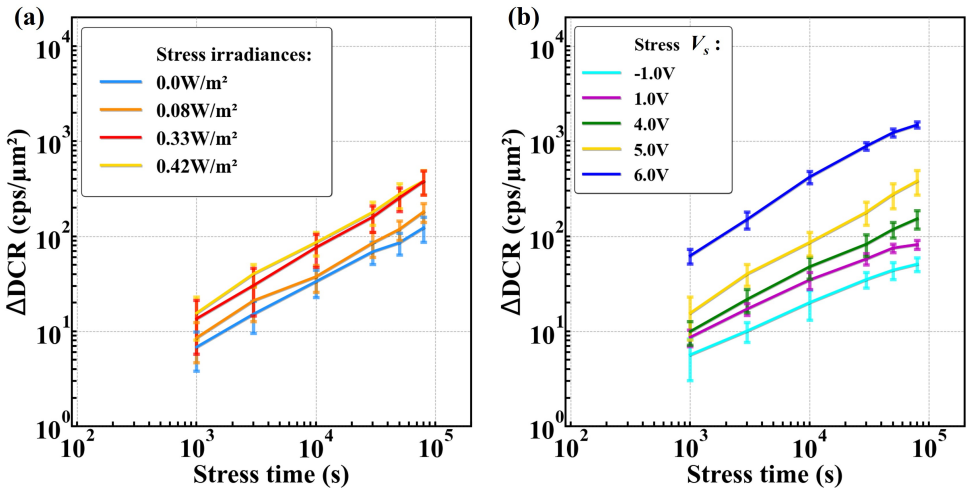


Figure 4.15: (a) ΔDCR measured at T_c=333K and V_c=5V as a function of the stress time for a stress V_s=5V and for different stress irradiances with error bars represented by the standard deviation between each device for the same stress condition. (b) ΔDCR as a function of the stress time for an irradiance level of 0.42W/m² and for different stress V_s with error bars. Error bars represent the standard deviation between each device. Extracting the same typical degradation time power across all stress conditions indicates that the degradation mechanisms are consistent.

electric field.

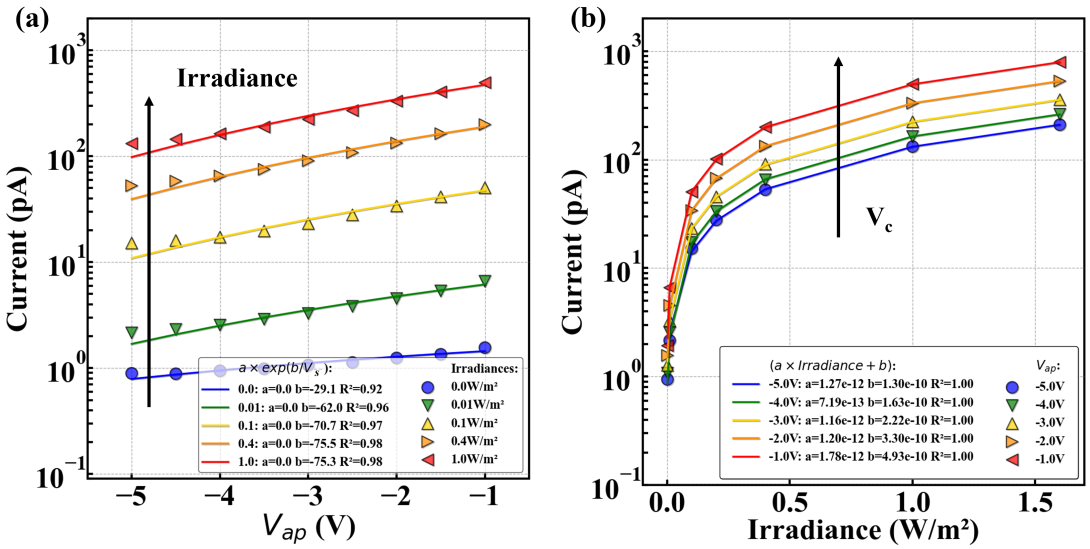


Figure 4.16: (a) Current versus V_c below V_{BD} for different irradiance levels measured (symbols) and fitted (lines) by an exponential relationship. (b) Current versus irradiance for different V_c measured (symbols) and fitted (lines) by a linear relationship. Both current and DCR exhibit a linear correlation with irradiance, which suggests that the degradation factor is the number of hot-carriers.

Now that the degradation factor has been identified as the number of hot-carriers, the potential localization of the degradation spot can be determined through simulation of carrier flow.

The carrier trajectories are simulated thanks to the drift-diffusion Monte-Carlo (described in **Chapter 2 Section 2.4**) and a current density streamline is represented in **Fig. 4.17**. The electrons move along the upper interface toward the cathode, while the holes travel in opposite direction (not shown here). From a degradation perspective, the number of hot carriers increases thanks to multiplication, which deposit their energy on the $Si-H$ bonds of the upper interface leading to their dissociation [143] as it will be explained in **Chapter 5 Section 1**. Two spots of degradation can be observed in both LB and HB regions. Moreover, the high carrier density causes a decrease in the electrostatic potential, which limits the kinetic energy acquired during avalanche events and subsequently restricts the degradation.

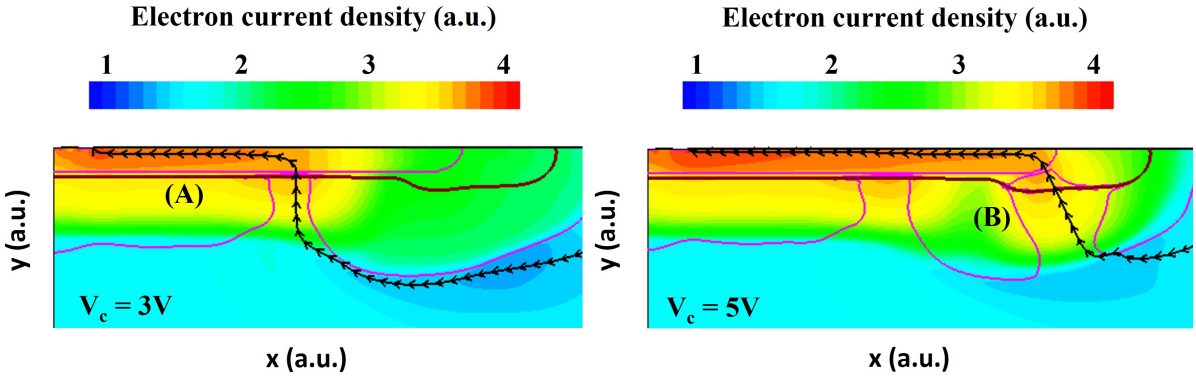


Figure 4.17: Maps of the electron current density to underline multiplication in (A) central part at V_c=3V and (B) peripheral part at V_c=5V. Typical electron density streamlines are shown (black lines with arrows). According to the sdevice tool [47], the purple lines in the plot indicate the boundary of the depletion region. This boundary is defined as $n \cdot \frac{eDensity}{Doping} - p \cdot \frac{hDensity}{Doping}$, where n and p are calculated as $n = max(\frac{Doping}{abs(Doping)+1}, 0)$ and $p = max(\frac{-Doping}{abs(Doping)+1}, 0)$, respectively. Both LB and HB regions exhibit two distinct spots of degradation. An electrostatic potential drop occurs due to the high carrier density, which restricts the kinetic energy acquired during avalanche events and consequently limits the degradation.

5.5 Process effect partitioning

The current simulation results suggest that the degradation spot is likely located at the top interface. To confirm this localization, a comparison with process variation will be conducted. The localization of activated defects during stress at the top interface is confirmed by means of process variations as shown in **Table 4.2**. **Fig. 1.10** allows for the appreciation of the region where processes vary. The key degradation factors are the etching of upper interface, the contacts and the implant dose.

As discussed in both Section **Section 2** and **Section 3**, the roughness quality of the interface can significantly impact the degradation mechanisms. A rough interface can lead to higher metastable state configurations that can break under hot-carrier interaction, resulting in the formation of a new R-G defect center. The contacts also play crucial role in the transport of charge carriers across the interface. Similarly, the implant dose near the interface can alter the electric field distribution by increasing the number of dopants, thereby increasing the built-

in electric field. This can result in a higher number of available energetic hot-carriers that can dissociate $Si-H$ bonds, leading to further degradation. Other processes on other interfaces have not shown significant impact expect for side effects due to species diffusion that can passivate the upper interface or annealing, which decreased the ΔDCR by a factor two. Annealing promotes the diffusion of atoms which can lead to the passivation of dangling bonds. It can also reduce the built-in electric field at the interface by modifying the doping concentration and distribution.

Table 4.2: ΔDCR improvement and degradation in the LB region ($V_c = 3V$) and in the HB region ($V_c = 3V$) for different manufacturing processes.

Process	Improvement		Degradation	
	LB	HB	LB	HB
Variation of W-shield aperture size		/2		
Annealing of the bottom oxide layer		/2		
Filling of DTI using ALD				x2
Substituting DTI with BDTI		/4		
Top etching after n++ and deposition of contacts using PVD				x9
Top etching after n++ and SiProt				x5
Deposition of contacts using ALD				x2
Lower n++ and p+ implant dose		x1/2		
Higher n++ and p+ implant energy				x3/2
Higher p-well implant energy				x4
Higher p+ and p-well implant dose			x3/2	
Higher n-well implant dose			x3/2	
RTP of n++ and n-well anneal or anti-TED of p-well		x1/2		

* The process details cannot be disclosed due to confidentiality reasons.

Conclusions

The objective of the SPAD reliability study was to identify potential degradation mechanisms and their locations. Based on the set of stress conditions, three distinct aging processes were suggested, one related to voltage, another to dark- and photo-generated current, and the last to temperature. It appears that the first two degradation mechanisms are connected through the quantity of hot-carriers present. Based on the lack of correlation observed between device-to-device avalanches at stress and different light aperture sizes, the triggered avalanches were not considered as stress factors. One of the identified degradation mechanisms was the accumulation at the hydrogenated amorphous silicon nitride film ($a - SiN_x : H$) during temperature stress in darkness without any applied voltage. On the other hand, ΔDCR showed a linear dependence on irradiance and an exponential dependence on the inverse of V_s, while fresh current measurements exhibited a similar dependence on voltage and irradiance. Hot-Carrier Degradation (HCD) was proposed as the degradation mechanism related to the growth of carrier multiplications, based on the number of carriers that increase linearly with irradiance and exponentially with the inverse of V_s. The aging mechanism was found to be consistent across stresses at various irradiances and voltages at the same temperature, as evidenced by the same time power. The manufacturing process was analyzed through a Design of Experiment (DoE), which revealed the importance of the upper interface in the degradation process.

6 A Statistical approach to modeling Dark Count Rate drift

6.1 Theoretical background: device-to-device variability in MOSFETs

The source of device-to-device variability in the threshold voltage V_{th} of MOSFETs, which is the minimum voltage required to create a conducting channel between the source and drain terminals, will be introduced. This approach will then be applied to the device-to-device variability of ΔDCR in SPADs. The study of the variability of threshold voltage drift (ΔV_{th}) during stress in CMOS has been conducted with a focus on two degradation mechanisms due to the application of a bias voltage and elevated temperature: Bias Temperature Instability (BTI) and Hot-Carrier Degradation (HCD) (described in **Section 3**). BTI and HCD in CMOS technologies [144][145][146][147] showed that the variability of threshold voltage drift ΔV_{th} along stress can be inferred to uncorrelated individual charged defects. This can be inferred to Random Dopant Fluctuations (RDF) along channel [148][149][150]. The simulation methodology in [144] is further explained to shed light on this phenomenon. Initially, a mesh of random V_{th} values is uniformly distributed to represent variations in local potential. The mesh is solved using SPICE to extract the V_{th}^0 of each device, resulting in a normal distribution. Next, a fixed number of charged defects are inserted, each represented by an additional V_{th} shift of one random mesh point. The resulting change in V_{th}, denoted as ΔV_{th}, is calculated. For a single additional charged defect, the model shows that the ΔV_{th} distribution follows a Weibull distribution. The distribution is characterized by two parameters: the shape parameter, which determines the shape of the distribution curve, and the scale parameter, which determines the location of the curve along the x-axis. The model also predicts that the shape parameter η scales inversely with both the width (W) and length (L) of the device, indicating that a single charge has a relatively larger impact on smaller devices. Since the locations of subsequent charges are uncorrelated, the overall ΔV_{th} distribution can be expressed as a convolution of individual exponential distributions, and the Cumulative Distribution Function (CDF) is described by the equation [144]:

$$F_n(\Delta V_{th}, \eta) = 1 - \frac{\Gamma(n, \frac{\Delta V_{th}}{\eta})}{(n-1)!} \tag{4.11}$$

where the right-hand side fraction is the regularized gamma function. In reality, a population of stressed devices will consist of devices with a different number of visible oxide defects, which will be Poisson distributed [151]. The total ΔV_{th} distribution can be obtained by summing the distributions F_n weighted by the Poisson probability:

$$P_n(n) = \frac{exp(-N_d)N_d^n}{n!} \tag{4.12}$$

where N_d is the mean number of defects of the Poisson distribution. The total CDF of ΔV_{th} is then given by:

$$F_{N_d}(\Delta V_{th}, \eta) = \sum_{n=1}^{\infty} \frac{exp(-N_d)N_d^n}{n!} F_n(\Delta V_{th}, \eta) \tag{4.13}$$

The mean of the derived distribution is:

$$\overline{\Delta V_{th}} = N_d \eta \tag{4.14}$$

and the variance of the distribution is given by:

$$\sigma^2_{\Delta V_{th}} = 2 N_d \eta \tag{4.15}$$

The electrostatic effect of a defect can be extracted from the linear relationship between the variance $\sigma^2_{\Delta V_{th}}$ and the mean $\overline{V_{th}}$, as demonstrated in **Fig. 4.18** [151][147]. This approach is based on the assumptions that ΔV_{th} follows an exponential distribution [144][145][146][147] and the number of defects is Poisson-distributed [151]. The device-to-device variability of ΔDCR in SPADs are compared to this approach in the following.

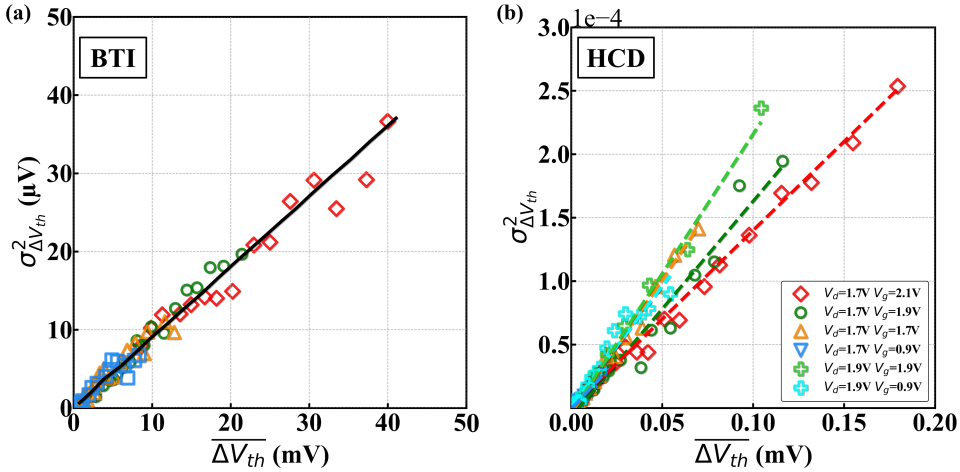


Figure 4.18: $\sigma^2_{\Delta V_{th}}$ as a function of $\overline{\Delta V_{th}}$ for (a) BTI from data reported in [151] and (b) HCD from data reported in [147] where different symbols representing different stress voltages. A linear relationship between $\sigma^2_{\Delta V_{th}}$ and $\overline{V_{th}}$ is fitted which allows the extraction of the electrostatic effect of a defect.

6.2 Monte-Carlo simulation of Dark Count Rate drift

> The main results of this section have been published in M. Sicre, X. Federspiel, B. Mamdy, D. Roy and F. Calmon, "Characterization and modeling of DCR and DCR drift variability in SPADs," 2023 IEEE International Reliability Physics Symposium (IRPS), 2023, pp. 1-5 [152].

6.2.1 Device-to-device measurement insights of Dark Count Rate drift

The objective of the upcoming simulation is to replicate the relationship between ΔDCR device-to-device variance $\sigma^2_{\Delta DCR}$ and its mean $\overline{\Delta DCR}$ in SPADs, trying to use the approach employed for ΔV_{th} device-to-device variability in MOSFETs. Measurements on a large device sample with a whole set of stress and process conditions coupled with a Monte-Carlo simulation methodology sheds light on the potential origin of the relationship between ΔDCR device-to-device variance

$\sigma^2_{\Delta DCR}$ and its mean $\overline{\Delta DCR}$ in SPADs. The Henry plot is used to display the CDF, and it helps to determine whether the variable is Gaussian by observing whether the points (x, y) align on the line with equation $y = (x - \overline{x})/\sigma$. To address ΔDCR variability, a 32x32 SPAD array is stressed at T_s=353K, V_s=3V and 0.02W/m² at 940nm for 67 hours. The ΔDCR measured from device-to-device shown in **Fig. 4.19 (a)** deviates from exponentially distributed hypothesis for CMOS [144][145][146][147]. Attempts to fit the ΔDCR cumulative distributions with exponential distributions have been unsuccessful. Furthermore, a quadratic relationship is found when plotting the device-to-device variance against the ΔDCR as illustrated by **Fig. 4.19 (b)**.

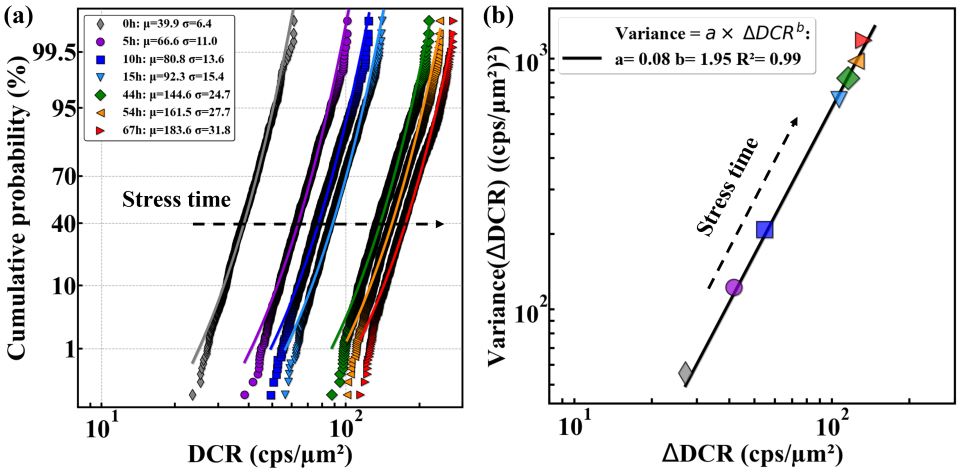


Figure 4.19: (a) Cumulative probability (Henry plot) for reference 32x32 SPAD array as a function of ΔDCR measured for different stress times (symbols). (b) Device-to-device ΔDCR variance $\sigma^2_{\Delta DCR}$ as a function of its mean $\overline{\Delta DCR}$ for different stress times (symbols). μ represents the mean of the normal distribution and σ the standard deviation. A quadratic relationship is found between $\sigma^2_{\Delta DCR}$ and $\overline{\Delta DCR}$.

Additional 24-hour stresses are performed on eight reference SPADs for each stress condition reported in **Table 4.1** (**Fig. 4.20 (a)**) and at T_s=353K, V_s=3V and 0.32W/m² for different ion implantation or interface processes (**Fig. 4.20 (c)-(d)**). Ten junction architectures, each consisting of 64 SPADs, are also stressed at T_s=353K, V_s=3V and 0.02W/m² for 67 hours (**Fig. 4.20 (b)**). These results confirmed the quadratic trend, regardless of the stress conditions, architectures, implantation scheme, and interface processes.

6.2.2 Simulation of post-stress defect localization

After identifying the degradation mechanisms in **Section 5**, the experiment and simulation will now focus on using temperature extraction to investigate the localization of defects induce by stress. Based on the simulation methodology described in **Chapter 3 Section 3**, including NRM and BTBT for the carrier generation rate, together with the avalanche breakdown probability P_t calculated from the drift-diffusion Monte-Carlo (see **Chapter 2 Section 2.4**), the DCR is simulated for each defect site. The defects are set as neutral with a single energy level at mid-gap and the cross-section is fixed at 10^{-15}cm² as for pre-stress DCR simulation. The defect concentration is chosen to reflect the defects induced by stress.

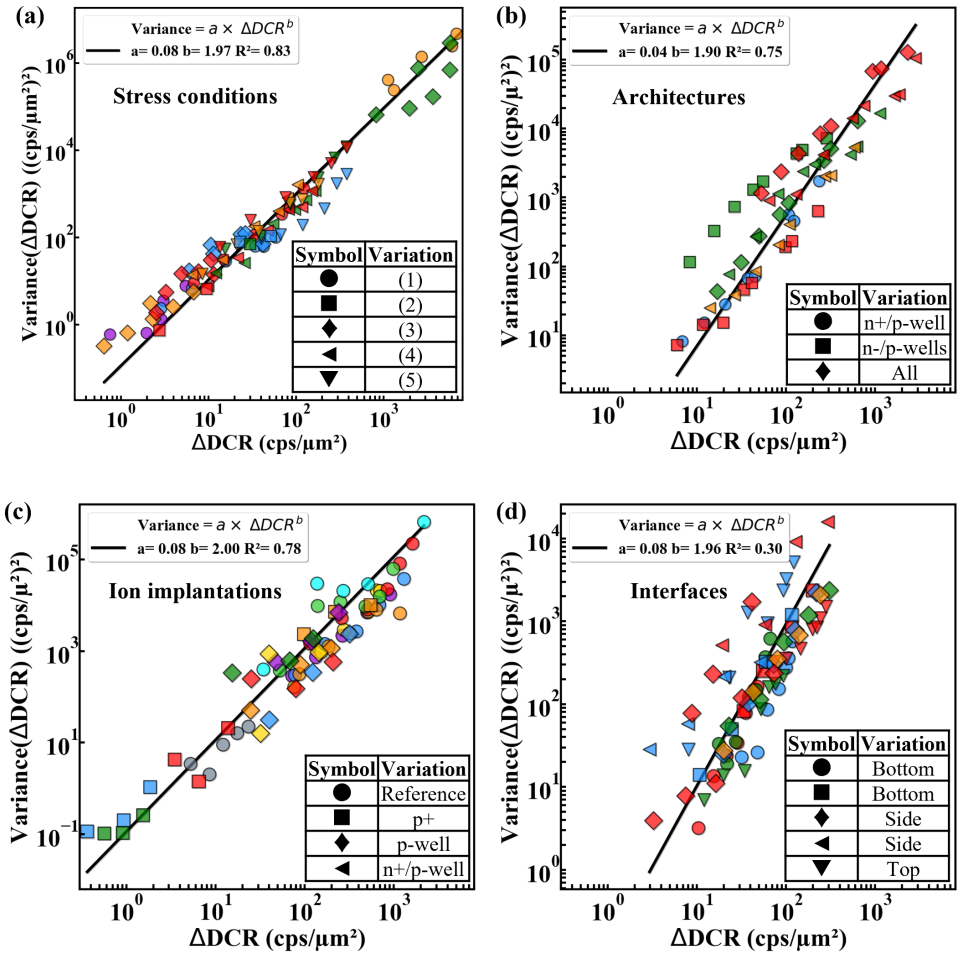


Figure 4.20: Device-to-device ΔDCR variance $\sigma^2_{\Delta DCR}$ as a function of its mean $\overline{\Delta DCR}$ along stress time. Each symbol represents (a) a different stress condition reported in **Table 4.1**, (b) a different junction architecture such as lateral extension or overlap of multiplication region and guard rings, (c) a different ion implantation scheme varying implantation energy and/or dose of multiplication region and guard rings and (d) a different interface process such as passivation and/or annealing budget. Each color among symbols represents a variant among a variation such as different temperatures at a given voltage. A quadratic relationship is found between $\sigma^2_{\Delta DCR}$ and $\overline{\Delta DCR}$ regardless of the stress conditions. The coefficient of determination R^2 might be low due to data dispersion from each variation.

From this procedure, the potential stress-induced defect positions and their associated carriers generation mechanisms are identified through voltage and temperature simulations. The Arrhenius activation energy E_a of DCR is extracted from measurements shown in **Fig. 4.21 (a)** performed before and after stress at T_c=300K, 333K, and 353K for the different stress conditions (**Table 4.1**). As a reminder, the carrier concentration is dependent on the Fermi level E_F which relies on the biasing as explained in **Chapter 2 Section 3.1**. In an electric field-free doped region, dopants n and p will be greater than the intrinsic concentration n_i^2 ($n \cdot p \gg n_i^2$). Charges are accumulating and defects act as recombination center. Minority carriers are sensitive to diffusion process ($E_a \approx E_g$). In an electric field doped region, n and p will be smaller than n_i^2 ($n \cdot p \ll n_i^2$). Charges are depleting and defects act as generation center ($E_a \approx E_t \approx E_g$

in the case of a mid-gap defect as explained in **Chapter 2 Section 3.1.4**). The extracted activation energies E_a have the same values regardless of the stress conditions, which indicates identical degradation localization. In order to simulate the average thermal dependence shown in **Fig. 4.21 (b)**, defects are placed at the depleted top interface located at the boundary of the n+ and n-well, as illustrated in **Fig. 1.10**. The decrease of E_a at high voltage after stress infers the pre-dominance of generation mechanism from defects at the depleted interface over diffusion from an increasing number of defects in this depleted region.

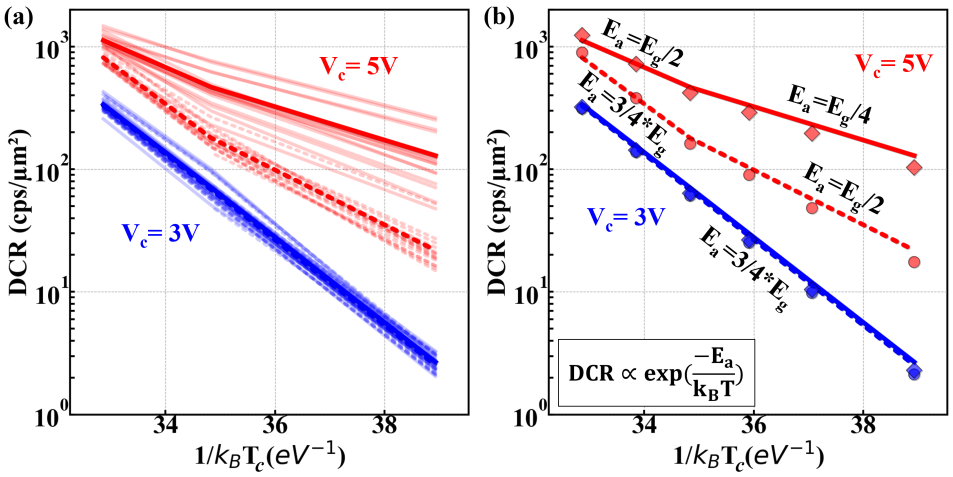


Figure 4.21: (a) DCR as a function of $1/k_BT_c$ measured at $V_c = 3V$ (blue) and 5V (red) for pre-stress (dashed lines) and post-stress (solid lines) values at different stress conditions (**Table 4.1**). DCR medians of stress conditions are bold lines and are reported on (b). (b) Median DCR as a function of $1/k_BT_c$ measured (lines) and simulated (symbols) at $V_c = 3V$ and 5V. The activation energy E_a is equivalent to the slope of the line following the Arrhenius law ln(DCR) vs $1/k_BT_c$. The extracted activation energies E_a are consistent across all stress conditions, suggesting that the degradation localizations are identical.

6.2.3 Model for predicting device-to-device variability

After identifying the potential defect localizations, the ΔDCR device-to-device variance $\sigma^2_{\Delta DCR}$ is studied by uniformly drawing defect positions [144][145][146][147] within a nanometer-thick box at the depleted top interface as depicted in **Fig. 4.22**. The defect concentration between each device is drawn from a Poisson distribution [151] with an average defect number defined from ten to one million by multiple of ten (N_0 to N_5). **Fig. 4.22** illustrates the various components of the simulation methodology proposed to model the device-to-device variance $\sigma^2_{\Delta DCR}$. Convolution between a uniform distribution of random defect positions (**Fig. 4.22 (a)**) associated with different DCR levels (**Fig. 4.22 (b)**) and an increasing number of defects Poisson-distributed in quantity (**Fig. 4.22 (c)**) results in the simulated cumulative plot (**Fig. 4.23 (a)**). The simulated device-to-device variance as a function of DCR (**Fig. 4.23 (b)**) reproduces the experimental quadratic law shown in **Fig. 4.19 (b)**. Attempts to change the distributions of position and defect number have been unsuccessful in capturing this dependence. In addition, the defect energy distribution [153] is considered and did not affect the result. This provides

further confirmation that the quadratic law results from the uniform distribution of defects at the top interface.

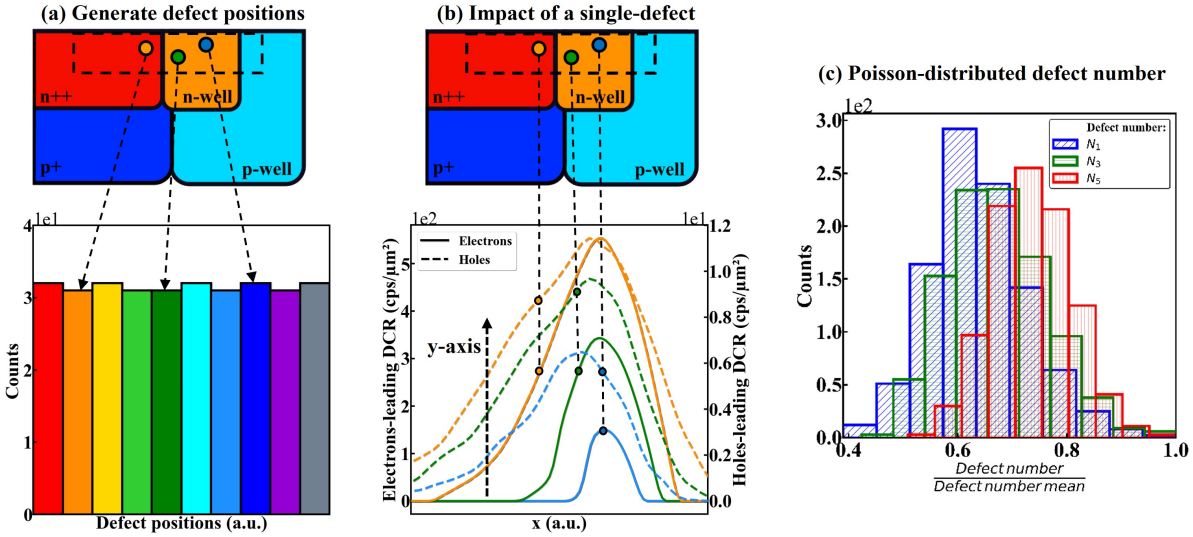


Figure 4.22: Suggested model of DCR device-to-device variability convoluting (a) the uniformly distributed defect positions (each bin represents 30 potential positions) at the depleted top interface (boundary between the n+ and n-well regions in **Fig. 1.10**), (b) the impact of defect position on DCR (each color refer to the defect position reported on (a)) and (c) the defect number Poisson-distributed from device-to-device (each color refers to defect quantity on **Fig. 4.23**). The corresponding defect impact as a function of the position in the device with the position histogram are shown in the lower part of the figure. The upper part represents the cross section of the junction with a zoom on the nanometer-thick box at the depleted top interface where defects are drawn.

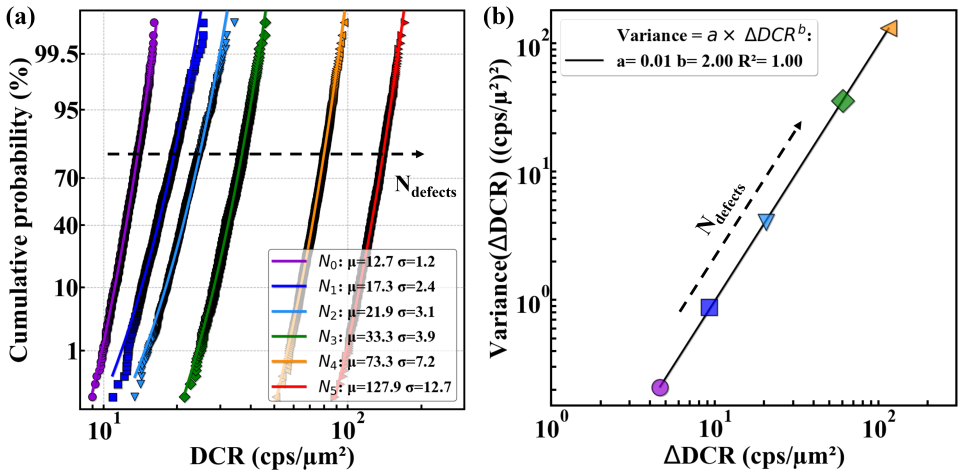


Figure 4.23: Cumulative probability (Henry plot) for a thousand reference devices as a function of ΔDCR for an increasing number of defects simulated (symbols). (b) Device-to-device ΔDCR variance $\sigma^2_{\Delta DCR}$ as a function of its mean $\overline{\Delta DCR}$ for an increasing defect number. A quadratic relationship is found between $\sigma^2_{\Delta DCR}$ and $\overline{\Delta DCR}$. $N_{defects}$ represents the average defect number of the Poisson distribution defined from ten to one million by multiple of ten (N_0 to N_5). A quadratic law between ΔDCR device-to-device variance $\sigma^2_{\Delta DCR}$ and its mean $\overline{\Delta DCR}$ is simulated.

Conclusions

Measurements and Monte-Carlo simulations were used to capture the variability of ΔDCR during stress. The stress results revealed a resilient quadratic relationship between the device-to-device variance $\sigma^2_{\Delta DCR}$ of ΔDCR and its mean $\overline{\Delta DCR}$, which was found to be independent of architectures, stress conditions, and processes. To understand the origin of this quadratic law, potential defect localizations were first identified based on temperature measurements taken before and after stress. The defects were then set at the depleted top interface (n+/n-well) to reproduce the extracted thermal dependency of DCR, accounting for the drift-diffusion carrier motion in avalanche breakdown probability simulation and a NRM-assisted carrier capture/emission rate for the carrier generation from defects. Temperature measurements and simulations before and after stress showed that the activated defects were located at the upper interface, regardless of the stress conditions. The device-to-device variability during stress was then studied using this simulation approach, coupled with defect position and number distribution. The experimental quadratic relationship between the device-to-device variance $\sigma^2_{\Delta DCR}$ of ΔDCR and its mean $\overline{\Delta DCR}$ was generated by convolving a uniform distribution of random defect positions, each with an assigned DCR value, with an increasing number of defects Poisson-distributed in quantity.

7 Conclusions

1

Metastable materials tend to degrade spontaneously in order to reach a state of lower Gibbs free energy, which takes into account the energy and disorder of the system. This degradation process is driven by a change in Gibbs free energy ΔG, and the rate of degradation is determined by the free energy of activation ΔG^*. Similar to a chemical reaction, it is possible to define a degradation rate constant κ to determine the activation energy ΔG^* for the degradation process. The theoretical framework for the occurrence of degradation enables the comprehension of SPAD aging over the duration of stress.

2

Borderless nitride, which is a hydrogenated amorphous silicon nitride film ($a-SiN_x:H$), can adopt several crystallographic configurations depending on the N/Si ratio. In the presence of hydrogen, $Si-H$ and $N-H$ bonds tend to replace $Si-Si$ and $Si-N$ bonds. The main defects in silicon nitride are dangling bonds, which can be either silicon or nitrogen. When three nitrogen atoms bond to one silicon atom, it is generally referred to as a K center, and nitrogen dangling bonds are referred to as N centers when the nitrogen atom is bonded to two silicon atoms. Both K and N centers are amphoteric traps, either acceptor or donor.

The Si/SiO_2 system experiences a lattice mismatch between Si and SiO_2, resulting in defects at the interface and in the oxide. These defects are commonly referred to as P_{b0} and P_{b1}, and they require passivation with hydrogen. Border traps are found within a 3nm range on both sides of the interface, which is caused by oxygen sub-stoichiometry during the gradual transition from $Si-Si_4$ to $Si-O_4$. This leads to the formation of X and Y centers, as well as E'_h centers that are located close to the interfacial layer. Additionally, there are Oxygen Vacancy (OV) and E' centers that are located further from the interfacial layer. These defects act as R-G centers for carriers.

3

From MOSFET perspective, the key parameters are V_{ds}, which represents carrier energy, and V_g, which represents carrier density and oxide field. The permanent component of BTI involves hole trapping in $Si-H$, $Si-H$ bond breakage, and hydrogen transfer into $Si-O$, which is facilitated by holes. On the other hand, the recoverable component of BTI includes hole capture/emission from oxygen vacancies, leading to the formation of E$'$ centers, as well as hole capture/emission from K centers and hydrogen retro diffusion, which can passivate the interface.

The physical mechanism underlying HCD can be broken down into three main steps. Firstly, electron defect generation occurs at the Si/SiO_2 interface through $Si-H$ bond breakage. Secondly, electron trapping takes place in the oxide, leading to bulk defect generation. Finally, charge trapping occurs in the oxide/spacer. The field dependency of HCD is driven by two fundamental components: carrier density and carrier energy, which will be further explained in **Chapter 5 Section 1**.

4

The objective of the SPAD reliability study is to identify potential degradation mechanisms and their locations. Based on the set of stress conditions, three distinct aging processes are suggested, one related to voltage, another to dark- and photo-generated current, and the last to temperature. It appears that the first two degradation mechanisms are connected through the quantity of hot-carriers present. Based on the lack of correlation observed between device-to-device avalanches at stress and different light aperture sizes, the triggered avalanches are not considered as stress factors. One of the identified degradation mechanisms is the accumulation at the hydrogenated amorphous silicon nitride film ($a - SiN_x : H$) during temperature stress in darkness without any applied voltage. On the other hand, ΔDCR showed a linear dependence on irradiance and an exponential dependence on the inverse of V_s, while fresh current measurements exhibited a similar dependence on voltage and irradiance. Hot-Carrier Degradation (HCD) is proposed as the degradation mechanism related to the growth of carrier multiplications, based on the number of carriers that increase linearly with irradiance and exponentially with the inverse of V_s. The aging mechanism is found to be consistent across stresses at various irradiances and voltages at the same temperature, as evidenced by the same time power. The manufacturing process is analyzed through a Design of Experiment (DoE), which revealed the importance of the upper interface in the degradation process.

5

Measurements and Monte-Carlo simulations are used to capture the variability of ΔDCR during stress. The stress results revealed a resilient quadratic relationship between the device-to-device variance $\sigma^2_{\Delta DCR}$ of ΔDCR and its mean $\overline{\Delta DCR}$, which is found to be independent of architectures, stress conditions, and processes. To understand the origin of this quadratic law, potential defect localizations are first identified based on temperature measurements taken before and after stress. The defects are then set at the depleted top interface (n+/n-well) to reproduce the extracted thermal dependency of DCR, accounting for the drift-diffusion carrier motion in avalanche breakdown probability simulation and a NRM-assisted carrier capture/emission rate for the carrier generation from defects. Temperature measurements and simulations before and after stress showed that the activated defects are located at the upper interface, regardless of the stress conditions. The device-to-device variability during stress is then studied using this simulation approach, coupled with defect position and number distribution. The experimental quadratic relationship between the device-to-device variance $\sigma^2_{\Delta DCR}$ of ΔDCR and its mean $\overline{\Delta DCR}$ is generated by convolving a uniform distribution of random defect positions, each with an assigned DCR value, with an increasing number of defects Poisson-distributed in quantity.

Chapter 5

Modeling of Dark Count Rate drift in SPADs

Objectives

The chapter discusses the state-of-the-art of Hot-Carrier Degradation (HCD) in MOSFETs. This involves the contributions of carrier energy distribution and disorder-induced variations in Si-H bond-dissociation energies at the SiO_2 interface to the degradation process. In the previous chapter, ΔDCR was studied at a fixed stress duration to extract the stress condition dependencies (namely temperature, irradiance and voltage) and to identify degradation mechanisms in the SPAD architecture. This chapter aims to explain the time dependence of ΔDCR under device stress to reflect the physics of the interface defect formation process and to predict device reliability based on the existing literature of HCD in MOSFETs. A Full-Band Monte-Carlo (FBMC) is used to compute the carrier energy distribution in order to determine if hot carriers have the threshold energy to dissociate Si-H bonds. The relationship between the degradation of ΔDCR in SPADs and the electric current without avalanches is studied to confirm the role of hot-carriers quantity in the degradation. Furthermore, the currents generated in darkness and under irradiance, along with the impact-ionization multiplication of carriers, are modeled to be incorporated into the degradation rate. The defect creation kinetics is then modeled using the carrier energy distribution against the $Si-H$ bond dissociation distribution to determine the probability of impact-ionization leading to bond dissociation. The objective is to provide an explanation for the increase in ΔDCR over time under various characterization and stress conditions.

Contents

1 Theoretical background: Modeling approaches of Hot-Carrier Degradation in MOSFETs

> This section summarizes the analysis of Hot-Carrier Degradation (HCD) models presented in S. E. Tyaginov, “Physics-Based Modeling of Hot-Carrier Degradation,” in "Hot Carrier Degradation in Semiconductor Devices", T. Grasser, Ed., Springer International Publishing, 2015, pp. 105–150 [154].

Dangling bonds at the disordered Si/SiO_2 interface are electrically active and can capture/emit carriers. The role of these dangling bonds in the degradation magnitude of MOS transistors was evidenced by the "giant isotope effect" [155]. Deuterium was used instead of hydrogen to anneal the dangling bonds at the Si/SiO_2 interface. The silicon-hydrogen/deuterium ($Si-H/D$) bonds at the interface break under hot carrier interactions and leads to hydrogen/deuterium diffusion. This induces a change in the interface defect profile increasing the total net carrier generation. The bond dissociation energy of silicon-deuterium bonds is higher that of silicon-hydrogen bonds leading to lower bond-breakage rate at the interface. For instance, ΔDCR in the studied SPAD was divided by two by changing the chamber gas during the annealing step from forming Forming Gas Annealing (FGA) to High-Pressure Deuterium annealing (HPD). Therefore, it is important to explain and model the degradation of MOS transistors under stress conditions, specifically temperature and applied voltage, while considering the impact of dangling bonds. This approach will be then applied to the degradation of SPAD.

1.1 Analysis based on carrier energy distribution

The mechanisms behind Hot-Carrier Degradation (HCD) kinetics in Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) have been described by Hess *et al.* [156][157][158] using both single- and multiple-carrier processes. Long-channel MOSFETs were operated at high voltages, resulting in the presence of carriers with energies exceeding the bond dissociation reaction threshold. The large number of these energetic carriers makes it possible for a single-particle collision to initiate the $Si-H$ bond dissociation process as illustrated in **Fig. 5.1**. The desorption rate of the Single Process (SP) is [156]:

$$R_{SP} \sim \int_{E_{th}}^{\infty} I(E)P(E)\sigma_{II}(E)dE \tag{5.1}$$

$I(E)$ is the carrier flux depending on energy, $\sigma_{II}(E)$ the Keldysh-like impact-ionization reaction cross section and $P(E)$ the desorption probability, while the integration starts from the threshold energy E_{th} for bond-dissociation process. In this case, the temperature increase results in a decrease of HCD [159]. The scattering mechanism rates increase with temperature, and therefore depopulate the high-energy fraction of the carrier ensemble.

As the device dimensions in short-channel MOSFETs decrease, the mean energy of the carrier energy distribution reduces and the bond dissociation can be initiated by a sequence of colder carriers that collide with the interface, inducing either phonon absorption or emission

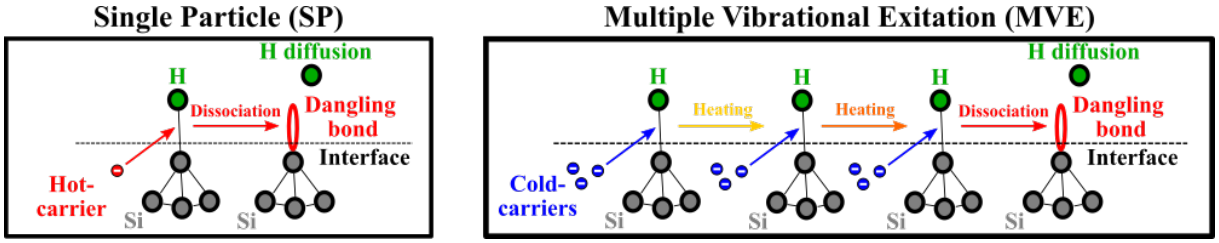


Figure 5.1: Diagram depicting HCD and its two processes: Single Process (SP) and Multiple Vibrational Excitation (MVE). In the SP process, a single high-energy carrier dissociates the $Si-H$ bond, resulting in the release of hydrogen. In contrast, the MVE process involves a series of lower-energy carriers that cause the bond to vibrate and eventually break down.

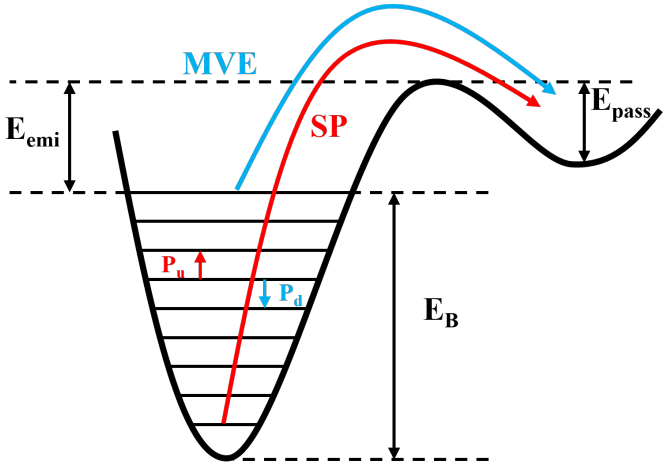


Figure 5.2: The truncated harmonic oscillator model provides a schematic representation of the Si-H bond. Bond dissociation can occur through either a single high-energy carrier or a series of cold carriers, which are referred to as Single Process (SP) and Multiple Vibrational Excitation (MVE) processes, respectively.

as illustrated in **Fig. 5.1**. In this process, the $Si-H$ bond can be modeled as a quantum well with eigenstates that are filled by these bond excitation/deexcitation processes. When the energy acquired by the bond reaches the highest bonded level through Multiple Vibrational Excitation (MVE), the bond dissociation occurs **Fig. 5.2**. This process is sometimes called a multiple-particle mechanism. The phonon excitation rate P_u is given by:

$$P_u \sim \int_{E_{th}}^{\infty} I(E)\sigma_{emi}(E)[1 - f_{ph}(E + \hbar\omega)]dE \tag{5.2}$$

and the phonon decay rate P_d by:

$$P_d \sim \int_{E_{th}}^{\infty} I(E)\sigma_{abs}(E)[1 - f_{ph}(E - \hbar\omega)]dE \tag{5.3}$$

The reaction cross sections for phonon absorption and emission are represented by σ_{abs} and σ_{emi}, respectively, with $\hbar\omega$ denoting phonon energy and $f_{ph}(E)$ including the phonon occupation numbers, which obeys Bose–Einstein statistics. The bond-breakage rate of the MVE process is [156]:

$$R_{MP} = (\frac{E_B}{\hbar\omega} + 1)[P_d + exp(-\frac{\hbar\omega}{k_B T})][\frac{P_u + \omega_e}{P_d + exp(-\frac{\hbar\omega}{k_B T})}]^{-\frac{E_B}{\hbar\omega}} \tag{5.4}$$

with E_B being the energy of the last bonded level in the quantum well, the phonon reciprocal life-time ω_e and T is the lattice temperature. In the Rauch–La Rosa paradigm [160][161][162][163], HCD is exacerbated at higher temperatures in contrast to long-channel MOSFETs. The energy deposited by carriers drives HCD instead of the maximal electric field, which was in the lucky electron model [164]. The lucky electron model describes the probability of carrier impact-ionization using **eq. 2.2**, which depends on the strength of the maximum electric field. When a lucky electrons gains enough kinetic energy to be injected in the oxide layer, it creates electron-hole pairs. The holes generated by this process can then become trapped in the oxide layer, leading to a buildup of positive charge (see **Chapter 4 Section 2.2**). In the Rauch–La Rosa energy paradigm, Electron-Electron scattering (EES) can populate the high-energy tail of the carrier ensemble as shown in **Fig. 5.3 (a)** strengthening the role of SP mechanism. The rate of both impact ionization and hot-carrier induced interface state generation can be expressed in terms of integrals of the form:

$$R = \int f(E)S(E)dE \tag{5.5}$$

where $f(E)$ represents the carrier distribution function, while $S(E)$ represents the reaction cross-section. $f(E)$ decreases rapidly as energy increases, whereas $S(E)$ increases in a power-law manner with energy as shown in **Fig. 5.3 (b)**. The interplay between the carrier distribution function and the reaction cross-section leads to a peak in the rate at a specific energy. This peak is determined by the following criterion which represents a trade-off between these two factors:

$$\frac{dln(n(E))}{dE} = -\frac{dln(S(E))}{dE} \tag{5.6}$$

The energies that have the greatest influence are determined by the knee points, which correspond to areas of high curvature in either $ln(n(E))$ or $ln(S(E))$.

Bravaix *et al.* [143][166][167] incorporates the interaction between the SP and MVE mechanisms present in the Hess model. Additionally, the carrier energy distribution function is substituted by operation/stress condition factors related to the knee energies in the impact ionization and bond-breakage rates following the Rauch–La Rosa approach. The three basic modes of HCD are associated with the SP mechanism, EES, and the MVE process, which are considered independent. The SP mechanism is characterized by high average carrier energies and is explained by the lucky electron model. The MVE process is associated with high carrier flux but low carrier energies and is described by the truncated harmonic oscillator model for the $Si - H$ bond. The intermediate regime, which is characterized by moderate current densities and moderate carrier energies, is governed by EES. The degradation regime in HCD undergoes a gradual shift from the EES mode to the MVE mechanism, with a mixed-mode regime in between, during which the relative contributions of each mechanism change progressively. To account for the competitive nature of EES, SP, and MVE mechanisms in HCD, the contributions

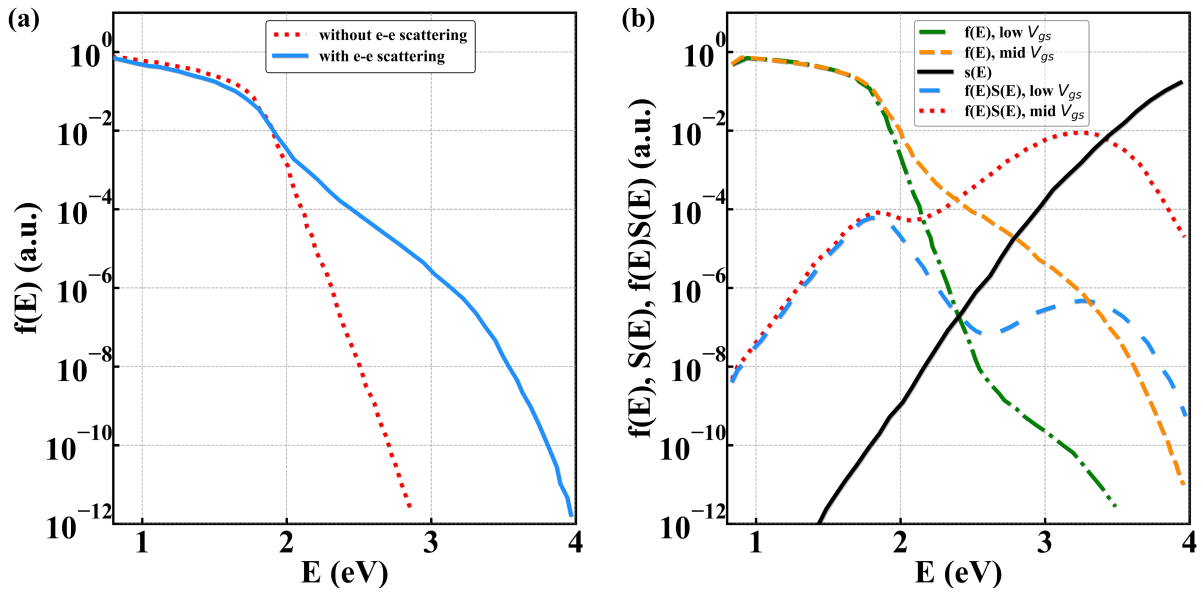


Figure 5.3: (a) Electron-electron scattering affects the shape of the carrier energy distribution function, resulting in the appearance of an extra peak in the high-energy tail of the distribution function. (b) The energy-driven paradigm is illustrated in the diagram, which shows that the knee energies shift depending on the voltage applied. Data reported are from [165].

of each process are weighted using empirical prefactors that reflect their respective probabilities of occurrence. The Arrhenius-type terms E_{emi} depicted in **Fig. 5.2** that describes the release of hydrogen from the last bonded state to the transport mode is also taken into consideration.

In their work, Tyaginov *et al.* [168] incorporate the concept of carrier Acceleration Integral (AI) into the Bravaix model framework. The AI is defined as the cumulative ability of the carrier ensemble to dissociate bonds, and can be expressed as:

$$I_{SP,MP}^{(e/h)} = \int_{E_{th}}^{\infty} f^{(e/h)}(E) g^{(e/h)}(E) \sigma_{SP,MP}^{(e/h)}(E) \nu(E) dE \tag{5.7}$$

where $f^{(e/h)}(E)$ is the carrier density function for electrons/holes, either calculated by a Full-Band Monte-Carlo (FBMC) (MONJU [169]) or by a deterministic Boltzmann transport equation (BTE) solver that employs the Spherical Harmonics Expansion (SHE) of the carrier density function (ViennaSHE [170]). $g^{(e/h)}(E)$ is the density of states, $\nu(E)$ is the group velocity, while $\sigma_{SP,MP}^{(e/h)}(E)$ is the Keldysh-like reaction cross section for the SP and MVE mechanisms triggered by electrons or holes:

$$\sigma_{SP,MP}^{(e/h)}(E) = \sigma_{0,SP,MP}^{(e/h)}(E - E_{th,SP,MP})^{\rho_{II}} \tag{5.8}$$

with the attempt rate $\sigma_{0,SP,MP}^{(e/h)}$ and the energy exponent of the impact-ionization dynamics $\rho_{II} = 11$ [171]. The threshold energy is E_{th}=1.5eV for each process. The rate at which bonds are broken can be expressed as the product of the attempt frequency $\nu_{SP,MP}^{(e/h)}$ and the AI weighted by each mechanism's probability $p_{SP,MP}$, where SP and MVE are the two mechanisms being considered.

Conclusions

The SP mechanism can induce damage such as Si-H bond break and oxide border trap due to high-energetic carriers. The mix mode mechanism involves two electrons creating one defect, and can also result in $Si-H$ bond break, thus in depends on both carrier energy and flux. For the MVE process, the electron energy is much lower than the energy required to break $Si-H$, so only the flux is important. This mechanism is only observable on thin oxide. The impact is more severe at low temperatures when considering carrier energy. Conversely, when considering carrier flux, the impact is more severe at high temperatures.

1.2 Analysis based on Si-H bond dissociation energy distribution

For an accurate degradation kinetics model of HCD, it is essential to consider not only the carrier energy distribution function but also the distribution of $Si-H$ bonds. The Penzin model [172] describes the bond rupture process by a kinetic equation for the passivated bond concentration N_{hb} with a depassivation and passivation reaction rate defined as:

$$\frac{dN_{hb}}{dt} = -\kappa N_{hb} + \gamma(N_{total} - N_{hb}) \tag{5.9}$$

where κ is the forward (depassivation) reaction rate while γ is the backward (passivation) rate. N_{total} is the initial concentration of $Si-H$ bond precursors plus the number of interfacial states. As described in **Chapter 4 Section 1**, the forward reaction rate has the following structure:

$$\kappa = \kappa_0 exp(\frac{E_a}{k_B T})\kappa_H \tag{5.10}$$

with the attempt frequency κ_0 and κ_H being the hot-carrier acceleration factor. In the context of the reaction rate, the $Si-H$ activation energy E_a is the parameter that explains the temperature-dependent behavior of the Arrhenius process. This term is controlled by the local hot carrier current I_{HC}:

$$\kappa_H = 1 + \delta_{HC}|I_{HC}|^{\rho_{HC}} \tag{5.11}$$

where δ_{HC} and ρ_{HC} are fitting parameters. The experimental data cannot be fully described by the exponential kinetics provided by **eq. 5.9**, as a constant activation energy would result in a power law with a power of one. This contradicts the sub-linear time dependence of HCD mechanism as observed in [173]. To overcome this limitation, the activation energy of bond dissociation depends on the hydrogen density and the electric field. The Si/SiO_2 interface, along with its surrounding area, is commonly referred to as a capacitor. In this context, the hydrogen that is released is believed to carry a charge, along with the remaining dangling bonds. Consequently, an extra electric field is introduced into the system due to these charges. This field acts as a barrier that prevents further hydrogen ions from leaving the system, thereby increasing the potential barrier that separates the bonded and transport states:

$$E_a = E_a^0 + \delta|F|^\rho + \beta k_B T ln(\frac{N_{tot} - N_{hb}}{N_{tot} - N_{hb}^0}) \tag{5.12}$$

$$\beta = 1 + \beta_\perp F_\perp \tag{5.13}$$

with E_a^0 being the activation energy in the absence of mobile hydrogen and N_{hb}^0 the initial concentration of $Si-H$ bond precursors. In summary, the system being treated as a capacitor means that discharging the capacitor requires additional energy to compensate for the change in the electric field, which is directly proportional to the normal component of electric field, represented as $\beta_\perp F_\perp$. Furthermore, the external electric field F can alter the bond by stretching or compressing it, which in turn modifies the activation energy governed by the term $\delta|F|^\rho$.

Returning the Hess model, it is important to note that it also includes the assumption that the activation energy E_a for the rate at which $Si-H$ bonds break is subject to statistical distribution, which was supported by *ab initio* studies [174][175]. The variation in E_a was observed through charge-pumping measurements, as shown in **Fig. 5.4**. This variability results in different power-law slopes observed during degradation [156][176][177]. This distribution of E_a induced by disorder at the Si/SiO_2 interface is described by the following Fermi derivative distribution [178]:

$$g(E_a) = \frac{1}{\sigma_{E_a}} \frac{exp(\frac{<E_a> - E_a}{\sigma_{E_a}})}{[1 + exp(\frac{<E_a> - E_a}{\sigma_{E_a}})]^2} \tag{5.14}$$

In the above equation, $< E_a >$ represents the average activation energy, while σ_{E_a} denotes the Half Width at Half Maximum (HWHM) of the activation energy distribution.

The Penzin model can also be derived from the disorder-induced local variations among the $Si-H$ bond energy at the Si/SiO_2 interface which gives the $Si-H$ bond concentration N_{hb}:

$$N_{hb}(t) = N_{hb}^0 \int_0^\infty g(E_a) exp(\kappa(E_a)t) dE_a \tag{5.15}$$

Under certain assumptions (which are explained in detail in **Appendix D**), the analytic solution of **eq. 5.15** can be expressed in the form of a power law:

$$N_{hb}(t) = \frac{N_{hb}^0}{1 + (\kappa t)^\alpha} \tag{5.16}$$

The equation shows that the number of hydrogen bonds decreases with time, and the rate of decrease is determined by the value of κ and the exponent α which is equal to $\frac{k_B T}{\sigma_{E_a}}$. As time goes on, the denominator of the equation becomes larger, causing $N_{hb}(t)$ to decrease.

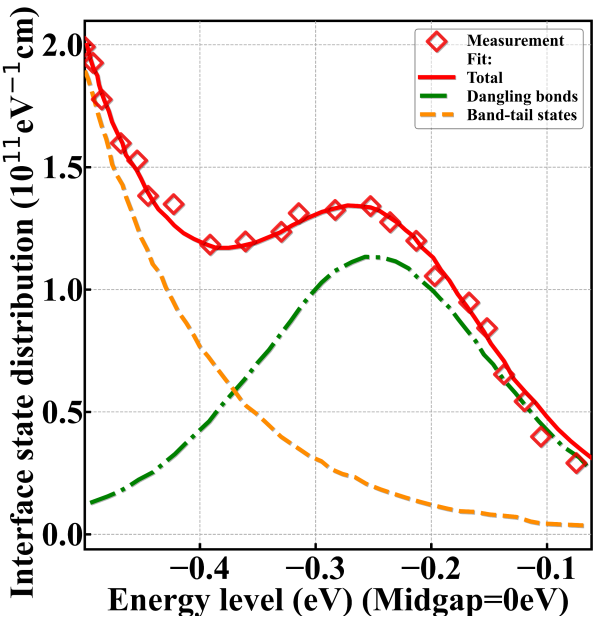


Figure 5.4: Charge-pumping measurements reveal a Fermi-derivative distribution of electronic state energies that is broadened due to the desorption of hydrogen at the interface. The distribution of electronic energy levels of defects at the interface has a half-width of approximately 0.1eV. Data reported are from [176].

Conclusions

For an accurate model of the HCD mechanism, it is necessary to consider the carrier energy distribution function, the distribution of $Si-H$ bonds, and the dissociation process. The SP mechanism can induce damage such as Si-H bond break and oxide border trap due to high-energetic carriers. The mix mode mechanism involves two electrons creating one defect, and can also result in $Si-H$ bond break, thus in depends on both carrier energy and flux. For the MVE process, the electron energy is much lower than the energy required to break $Si-H$, so only the carrier flux is important. The impact is more severe at low temperatures when considering carrier energy. Conversely, when considering carrier flux, the impact is more severe at high temperatures. The bond-dissociation energy can be affected by the hydrogen density and the transversal component of the electric field, which can modulate the local electrostatic potential and stretch the bond. Another approach is to consider the disorder-induced local variations in $Si-H$ bond energy at the Si/SiO_2 interface, which can match the observed sub-linear time dependence of the degradation.

2 Modeling of Dark Count Rate drift in SPADs under Hot-Carrier Degradation

The main results of this section have been published in "Hot-Carrier Degradation modeling of DCR drift in SPADs," ESSDERC 2023 - IEEE 53rd European Solid-State Device Research

Conference (ESSDERC), 2023, pp. 61-64 [179]. Additionally, an upcoming publication titled "Modeling of Hot-Carrier Degradation driven by silicon-hydrogen bond dissociation in SPADs" will provide further insights into the research.

In micrometer-large SPADs, the device operating voltages are high and carriers with energies above the threshold energy of the bond dissociation reaction E_{th} are likely present in substantial density. Such energetic carriers likely trigger the bond dissociation process in a SP collision as explained in the previous section. This results in a stress-time dependent change in DCR drift (ΔDCR) as shown in **Chapter 4 Section 5**. This assumption will be verified hereafter based on Full-Band Monte-Carlo (FBMC) simulations which compute the carrier distribution as a function of energy considering band structures and scattering mechanisms.

2.1 Simulation of the carrier energy distribution

A Full-Band Monte-Carlo (FBMC) method [180][181] was used to demonstrate SP degradation process in SPADs through computation of carrier energy distribution $n(E)$ which is refers as $f(E)$ in **Section 1**. This approach stands on silicon band structures computed with Tight-Binding method by Linear Combination of Atomic Orbitals (LCAO) [182]. Phonon scattering, ionized impurity scattering, and carrier-carrier scattering rates are calculated consistently with the band structure. To simulate charge transport in semiconductors using the Monte Carlo method, the semi-classical Boltzmann transport equation (BTE) must be solved. This equation describes the changes in the carrier energy distribution function $f(k,r,t)$ over time in both real and momentum space as:

$$\frac{\partial f(k,r,t)}{\partial t} + \nu \cdot \nabla_r f(k,r,t) + \frac{\Im}{\hbar} \bigtriangledown_k f(k,r,t) = \frac{\partial f(k,r,t)}{\partial t}\Big|_{scattering} + s(k,r,t) \tag{5.17}$$

The first term of the left-hand side of the equation describes the changes in the distribution function over time, while the second term accounts for deviations from the equilibrium state caused by variations in real space. The third term takes into account momentum space variations due to Lorentz force $\Im = -e(F + \nu \cdot B)$. The right-hand side of the equation considers the impact of scattering, and generation-recombination mechanisms represented by $s(k,r,t)$.

Two processes are involved in the collision term: (1) carriers with a wave vector k' can be scattered to k with a transition rate $S(k',k)$, thus increasing $f(k,r,t)$ and (2) Carriers with a wave vector k can be scattered to k' with a transition rate $S(k,k')$ decreasing $f(k,r,t)$. The competition of these two actions leads to the following expression:

$$\frac{\partial f(k,r,t)}{\partial t}\Big|_{scattering} = \int f(k',t)(1-f(k,t))S(k',k)dk' - \int f(k,t)(1-f(k',t))S(k',k)dk' \tag{5.18}$$

Assuming a homogeneous material, constant in space, this implies that the gradient in real space is zero ($\nabla_r f(k,r,t) = 0$). In steady-state, constant in time, this implies that the time

derivative is zero ($\frac{\partial f(k,r,t)}{\partial t} = 0$). The generation-recombination mechanisms are not considered ($s(k,r,t) = 0$). Assuming no magnetic field, the equation **eq. 2.1** is simplified to:

$$\frac{\partial f(k,r,t)}{\partial t}\Big|_{scattering} = \frac{-eF}{\hbar}\nabla_k f(k,r,t) \tag{5.19}$$

Since only the collision and diffusion terms remain, the emphasis is on creating a band structure $E_n(k)$ (with n the band index) and utilizing the Fermi golden rule for the transition rate $S(k,k')$ and $S(k',k)$ to determine the diffusion terms.

With the theoretical background having been introduced, the simulation methodology will now be described. During the initialization stage, a rejection sampling in carrier energy density, which is the occupation probability at a given energy multiplied by the state density at this energy, is performed to determine the energy of each carrier. The wave vector $\overrightarrow{k}$ of each carrier is randomized on the resultant energy iso-surface at each time step. The carrier acceleration is computed according to the integration of the second Newton's law which leads to:

$$\overrightarrow{\Delta k} = \frac{1}{\hbar} q \overrightarrow{F} \Delta t_{ff} \tag{5.20}$$

The carrier free-flight time Δt_{ff} is computed by a rejection sampling in the interaction probabilities with the different scattering mechanisms. The scattering is either elastic (carrier-impurity) saving carrier energy, lowly inelastic (carrier-phonon or carrier-carrier) or highly inelastic (impact ionization or Auger recombination) leading to meV-exchange or to eV-exchange respectively. The impact ionization rate is modeled by the Keldysh formula $P_{II} = P_{II}^0(E-E_0)^{\alpha_0}$ [54] with E and E_0 the carrier energy and the threshold impact-ionization energy, P_{II}^0 is the attempt rate and α_0 the energy exponent of the impact-ionization dynamics [58][59]. To simulate the post-ionization state, it is assumed that the carrier kinetic energy lost by the incoming particle during the impact is constant and equal to the band gap E_g [62][63] to fulfill the laws of energy conservation and momentum conservation as explained **Chapter 2 Section 2.1**.

The top interface was previously identified as degradation localization in the SPAD architecture (see **Chapter 4 Section 5** and **Section 6.2**) thus a focus on $n(E)$ in its vicinity is realized. One hundred of electron-hole pairs are positioned either close to the n++/p+ or n-/p-well region at the start of the simulation. The electrons move along the upper interface toward the cathode (not shown here), while the holes travel in opposite direction as previously demonstrated with the drift-diffusion Monte-Carlo simulations (see **Fig. 2.16**). $n(E)$ is extracted at the top interface either in n++/p+ or n-/p-well region and normalized by the total carrier distribution obtained in the simulation to compare hot carrier populations between each condition. As shown in **Fig. 5.5 (a)-(b)**, some energy-tail electrons reach the mean threshold dissociation energy of $Si-H$ bond (E_{th}=1.5eV [143][175]) depending on stress voltage V_s (**Fig. 5.5 (c)-(d)**) and stress temperature T_s (**Fig. 5.5 (e)-(f)**). The positive thermal dependence of carrier energy is a consequence of carrier multiplication [42] as previously explain for P_t in **Chapter 2 Section 2.2**. The carrier-phonon mean free path decreases with increasing lattice temperature which results in a lower impact-ionization rate (**eq. 2.19**) under the same electric field. The

same thermal and voltage dependencies have been observed on the overall population. There are two potential degradation dynamics according to the localization either in n++/p+ or n-/p-well region. These two different degradation regions can be stimulated independently depending on the characterization conditions, as illustrated by **Fig. 2.8**. The simulated carrier energy distributions $n(E)$ along V_s and T_s are integrated over the bond dissociation energy to account for the bond-dissociation probability by hot carrier interaction at each carrier energy.

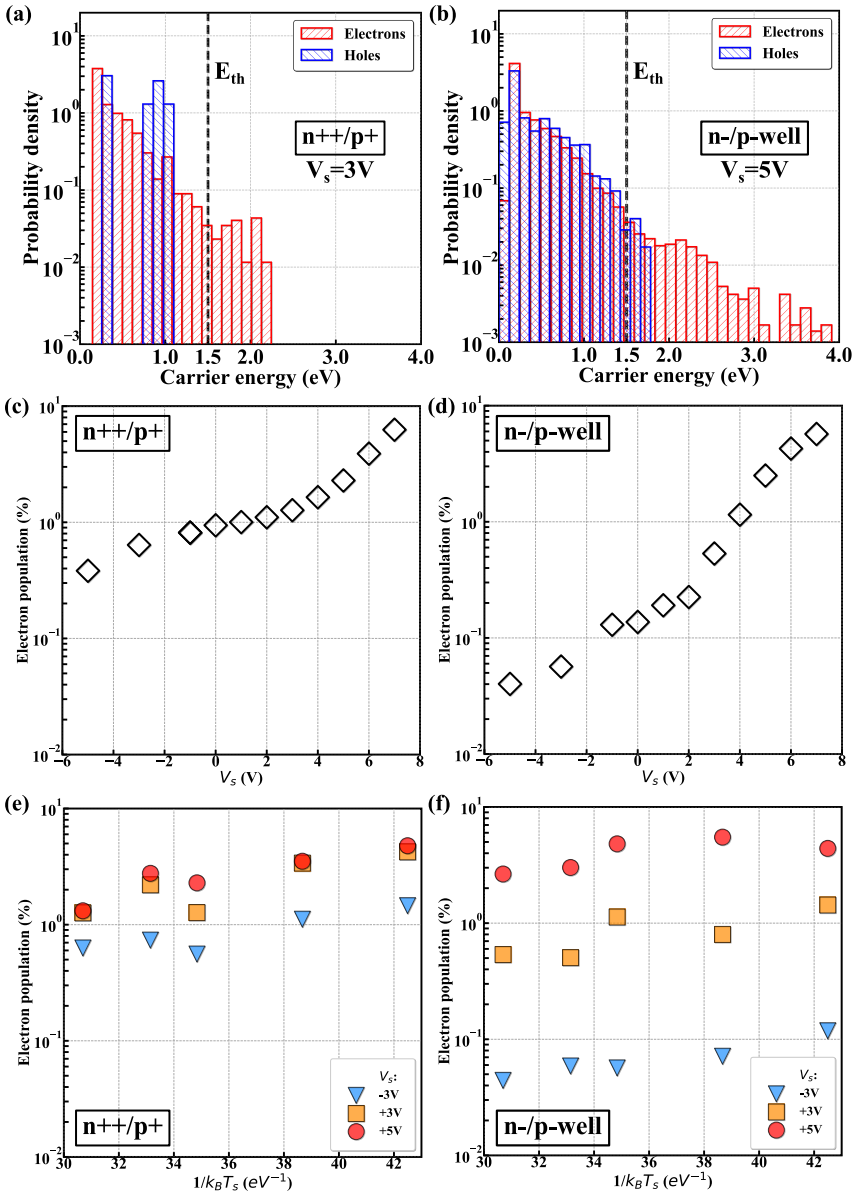


Figure 5.5: Probability density of electron (red) and hole (blue) energy simulated by a Full-Band Monte-Carlo (FBMC) [180][181] at T_s=333K for (a) V_s=3V in n++/p+ region (LB) and (b) V_s=5V in n-/p-well region (HB). Carriers acquiring energy above Si-H bond dissociation (1.5eV) are extracted and reported in (c) and (d) as a function of V_s at T_s=333K and in (e) and (f) as a function of $1/k_BT_s$ at V_s=-3V (blue triangles), 3V (orange squares), 5V (red circles).

Conclusions

Full-Band Monte-Carlo (FBMC) allows to extract the carrier energy distribution $n(E)$ within the device accounting for the scattering mechanisms and the carrier acceleration. Some energy-tail electrons reach the mean threshold dissociation energy of $Si-H$ bond (E_{th}=1.5eV) depending on V_s and T_s. There are two potential degradation dynamics that can occur depending on the localization of the degradation, either in the n++/p+ region or in the n-/p-well region. The two distinct degradation regions can be selectively stimulated depending on the conditions of characterization.

2.2 Degradation induced by current without avalanches

As discussed in **Chapter 4 Section 5.4**, the current density has a significant impact on the degradation as it increases the number of hot carriers that can reach the Si-H dissociation as demonstrated in the previous section. Further verification of the impact of current density on degradation will be conducted through stresses and measurements. Measurements and stresses are performed on eight SPADs stressed for each 24-hours stress conditions reported in **Table 5.1**. **Fig. 5.6** displays the time chronograph for the initialization and stress sequence procedure. During initialization, the electrical current I below the avalanche breakdown voltage V_{BD} is measured before the stress sequences (green color in **Fig. 5.6**) to extract the current density as a function of temperature and voltage. DCR is measured before and after each stress period (blue color in **Fig. 5.6**). The absolute ΔDCR corresponding to the difference between the DCR at a given stress time and the initial DCR is then computed according to **eq. 4.10**. DCR measurements are performed at T_c and V_c above V_{BD} applied while the diode is in reverse bias condition. To mitigate the impact of avalanche on degradation, the stress voltage V_s is applied either above or below V_{BD}. The extrinsic devices showing higher DCR levels, resulting likely from implant defects in the junction as explained in **Chapter 3 Section 3**, are not considered to ensure these do not affect the average behavior of the intrinsic population.

Table 5.1: Stress conditions to model DCR drift

T_s (K)	V_s (V)	Irradiance (W/m²)
333	-3/-1/1/2/3/4/5	0.0
333	-9/-6/-3/-1/1/3/4/5	0.3
300/333/373	3	0.0/0.3
353	-3/4/5	0.0
423	5	0.0
353/373/423	-3	0.33
353	-3	0.01
353	4	0.1/0.33/0.43
353	5	0.08/0.33/0.42
373	5	0.22

Our previous study has shown that DCR degradation could occur even without avalanches (**Chapter 4 Section 5.2**). Avalanches may result in large amounts of hot-carriers but they

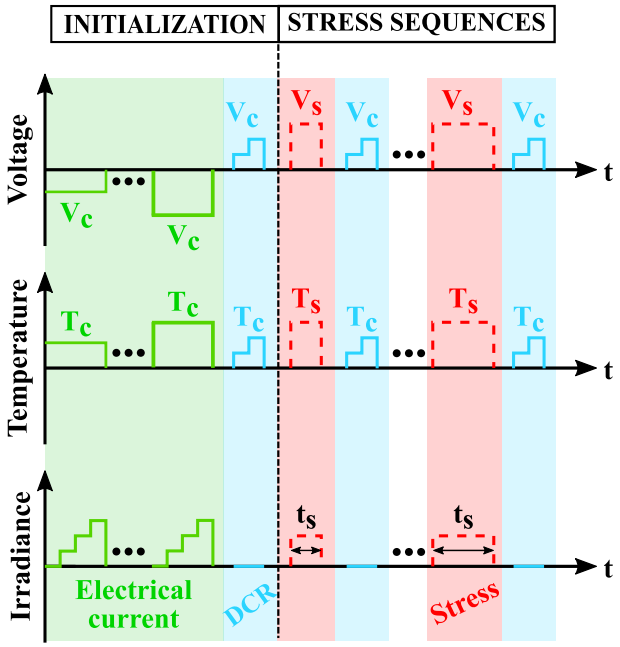


Figure 5.6: Time chronograph of initialization and stress sequences. Bold lines are measurement and dashed lines are stress. Electrical current is measured at different V_c, T_c and irradiance before stress (green box). DCR measurements are performed before stress and after each stress phase at different V_c and T_c (blue box). Devices are stressed (red box) for an increasing stress time t_s at V_s, T_s and irradiance.

cannot flow near the interface due to electrostatic potential drop [88] as shown in **Fig. 4.17**. These hot carriers generated during avalanches are unable to reach and cause degradation to the upper interface. Typical $DCR-$ and $I-V$ curves are shown in **Fig. 5.7 (a)-(b)**. The dark- and photo-generated current, not triggering avalanches, cannot be distinguished from the avalanche current when the diode is biased above V_{BD}. As an example, the Charges Per Pulse (CPP) (which is defined as the number of charges per avalanche, as explained in **Chapter 1 Section 5**) can be obtained by dividing the electrical current I by the DCR value. In **Fig. 5.7 (c)**, it can be observed that the CPP value remains constant throughout the stress duration. This suggests that the increase in dark current over time is still significantly smaller than the avalanche current. In fact, the charges per avalanche are expected to remain relatively constant, with the exception of the statistical dispersion caused by the multiplication process. If the dark current were to become dominant over the avalanche current, this value would change, but as shown in **Fig. 5.7 (c)**, this is not the case.

However, the current above V_{BD} without avalanches can be assumed that it is proportional to the current measured below V_{BD}, in darkness or under illumination, accounting for the multiplication rate. The DCR drift after stress is therefore compared to the current measured below V_{BD} before stress and reported in **Fig. 5.8** for both characterization LB and HB regions. With (V_s=4V and 5V) or without (V_s=-3V) avalanche triggering, ΔDCR versus current measured without avalanches follows a linear relationship with the irradiance. The same aging mechanisms can be assumed to relate to the increasing number of hot carriers that enhance $Si-H$

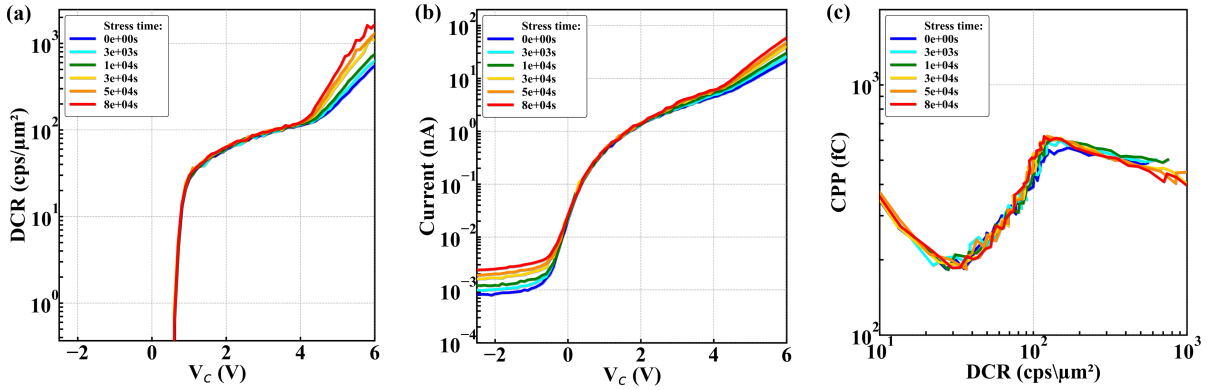


Figure 5.7: (a) DCR and (b) electrical current I measured as a function of V_c for an increasing stress time. By dividing the electrical current I by the DCR value, one can obtain (c) Charges Per Pulse (CPP) as a function of DCR along stress duration. It has been observed that the CPP remains constant over stress time, indicating that the increase in dark current ($0 > V_c$) is much smaller than the avalanche current ($0 < V_c$).

bond dissociation. According to these results, the current density without avalanches will be extrapolated above V_{BD} to compute the degradation rate based on the number of available hot carriers.

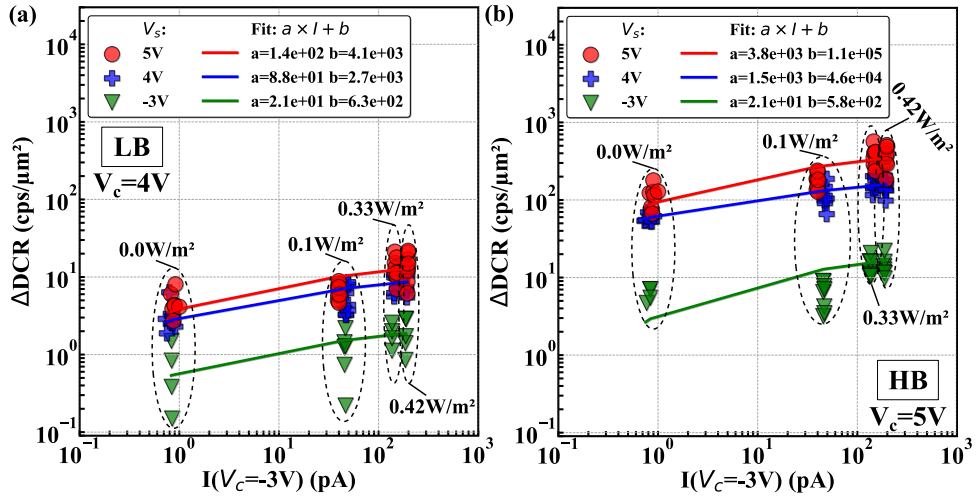


Figure 5.8: ΔDCR as a function of the electrical current I (without avalanches) measured (symbols) and fitted (lines) by a linear relationship at V_c=-3V and at T_c=353K for several irradiance levels from 0.0W/m² (left) to 0.42W/m² (right). ΔDCR was obtained from 24-hours stresses at different V_s and at T_s=353K for the same irradiance levels and was measured at T_c =333K and at (a) V_c =4V (LB) and (b) 5V (HB).

Conclusions

Due to the large avalanche current, it is not possible to measure the current above V_{BD} without avalanches. The constant value of the CPP along the stress duration confirms this fact. It is assumed that the current above V_{BD} without avalanches is proportional to the current measured below V_{BD}, taking into account the multiplication rate. The degradation rate is dependent on the current without avalanches, as there is a resilient relationship between the change in ΔDCR and the current measured without avalanches. This relationship appears to be related to the increasing number of hot carriers that enhance $Si-H$ bond dissociation.

2.3 Modeling of electrical current components

To improve readability of the different modules of the reliability model developed, some equations from **Chapter 2** are restated along with their corresponding parameter values in **Table 5.2**.

2.3.1 Modeling of the multiplication current

After simulating the carrier energy distribution for a given number of electron-hole pairs, the total carrier density flowing in the device must be known to model the degradation rate based on the hot-carrier quantity. As previously explained, the current density above V_{BD} is modeled by extrapolating the current density below V_{BD} to get rid of the avalanche current. The multiplication rate is equal to:

$$G_\alpha = \alpha_n \frac{j_n^{d,ph}}{q} + \alpha_p \frac{j_p^{d,ph}}{q} \tag{5.21}$$

α_n and α_p are the ionization rates which are defined as the mean number of charges generated by distance unit traveled by an electron or a hole. The impact-ionization coefficient are defined according to the Chynotew law [68] which states that:

$$\begin{cases} \alpha_n = \gamma a_n exp(\frac{-\gamma b_n}{E}) \\ \alpha_p = \gamma a_p exp(\frac{-\gamma b_p}{E}) \end{cases} \tag{5.22}$$

$$\gamma = \frac{tanh(\frac{\hbar\omega_{op}}{2k_B T_0})}{tanh(\frac{\hbar\omega_{op}}{2k_B T})} \tag{5.23}$$

The coefficients $a_{n,p}$, $b_{n,p}$, and $\hbar\omega_{op}$ were measured by van Overstraeten and de Man [67]. T_0 and T are the reference and the absolute temperature, and ω_{op} is the angular frequency of optical phonons. γ and $\hbar\omega_{op}$ express the temperature dependence of the optical-phonon gas against which carriers are accelerated. The multiplication current is equal to G_α multiplied by q and by the multiplication-sensitive area which is defined as the space charge region S_{SCR}. The multiplication current contribution is coupled with the dark- and photo-generated ones considering the simulated electric field F (see **Fig. 2.8**). The contribution of carrier multiplication is indicated

by the second term in **eq. 5.42** for the dark current and in **eq. 5.34** for the photo-generated current.

Table 5.2: DCR drift model parameters

Quantity	Units	Value	Ref./Cmt.
a_n	cm^{-1}	$7.03 \cdot 10^5$	
a_p	cm^{-1}	$1.582 \cdot 10^6$ [a] / $6.71 \cdot 10^5$ [b]	
b_n	$V \cdot cm^{-1}$	$1.231 \cdot 10^6$	[67]
b_p	$V \cdot cm^{-1}$	$2.036 \cdot 10^6$ [a] / $1.693 \cdot 10^6$ [b]	
T_0	K	300	
$\hbar\omega_{op}$	eV	0.063	
N_t^0	cm^{-2}	10^{12}	[91]
σ_{t^0}	cm^{-1}	10^{-15}	[74]
E_t	eV	$E_g/2$	
m_0 [c]	kg	$9.11 \cdot 10^{-31}$	
$\frac{m_n}{m_0}$	1	0.2 [d]	[51]
$\frac{m_p}{m_0}$	1	0.49 [e]	
E_g^0	eV	$4.73 \cdot 10^{-4}$	
α_g	$eV \cdot K^{-1}$	$4.73 \cdot 10^{-4}$	[93]
β_g	K	636	
E_p	eV	0.015	[183]
N_{hb}^0	cm^{-2}	$3 \cdot 10^{12}$	[184]
E_{th}	eV	1.5	[143][175]
σ_{bd}	eV	0.1	[185]
σ_{II}^0	$eV^{-\rho_0} \cdot cm^2$	10^{-20}	Arbitrarily set
ρ_0	1	11	[185]

[a] electric field range: 0.175-0.4 MV/cm
[b] electric field range: 0.4-0.6 MV/cm
[c] m_0 is the free electron rest mass
[d] considering conduction band minimum at k=0
[e] considering heavy hole valence band maximum at E=k=0

2.3.2 Modeling of the dark current

The model for the dark current involves the use of the Shockley-Read-Hall (SRH) Recombination-Generation (R-G) rate of carriers [186], which is described in **Chapter 2 Section 3.1**. This model is based on previous results presented in **Chapter 3**, which demonstrated the presence of defects at the top depleted interface where n_i^2 significantly exceeds n and p. The generation rate, G_d, can be simplified under the assumption of the worst-case scenario, where the defect profile for capturing electrons and holes is the same, and defects are amphoteric [187]. This scenario is explained in **Chapter 4 Section 2.1**, where a single initial interface defect density, N_t^0, is defined at a single-energy mid-gap level, E_t. The simplified equation for G_d is given by:

$$G_d = \frac{N_t^0 n_i \sigma_t v_{th}^n v_{th}^p}{v_{th}^n exp(\frac{E_t - E_F^i}{k_B T}) + v_{th}^p exp(\frac{E_F^i - E_t}{k_B T})} \tag{5.24}$$

where E_i is the intrinsic Fermi level, and E_v and E_c are valence and conduction band energy. The temperature-dependent intrinsic density n_i can be expressed as a function of temperature using the following equation:

$$n_i = \sqrt{N_c N_v} exp(\frac{-E_g}{2k_B T}) \tag{5.25}$$

where the effective densities of states in the conduction band N_c and in the valence band N_v are:

$$N_{c,v} = 2(\frac{2\pi m_{n,p} k_B T}{h^2})^{\frac{3}{2}} \tag{5.26}$$

with $m_{n,p}$ the effective mass of electron and hole and h the Planck's constant. The band gap follows [93]:

$$E_g = E_g^0 - \frac{\alpha_g \cdot T^2}{\beta_g + T} \tag{5.27}$$

where E_g^0 is the bandgap at 0K and α_g and β_g some experimental parameters depending on the semiconductor material. The Fermi level for an intrinsic semiconductor E_i is:

$$E_i = \frac{E_g}{2} + \frac{3k_B T}{4} ln(\frac{m_p}{m_n}) \tag{5.28}$$

The thermal velocity v_{th} is equal to:

$$v_{th} = \sqrt{\frac{3k_B T}{m_{n,p}}} \tag{5.29}$$

The electric field-enhancement of generation rate through Trap-Assisted Tunneling (TAT) and Poole-Frenkel effect (PF) (see **Chapter 2 Section 3.2.2** and **Section 3.2.1** respectively) is considered through the emission/capture cross section defined as:

$$\sigma_t = \sigma_t^0 \cdot (1 + \Gamma_{TAT,PF}) \tag{5.30}$$

where $\Gamma_{TAT,PF}$ is a field-enhancement factor defined with the default parameters specified in [47]. Considering **eqs. 5.21-5.30** with parameters reported in **Table 5.2**, the dark current I_d is equal to G_d multiplied by q and the area populated by defects V_d along with the dark current multiplied:

$$I_d = q \cdot (G_d \cdot V_d + \alpha_{n,p} \cdot G_d \cdot \frac{V_d}{S_{SCR}}) \tag{5.31}$$

The dark current modeling reproduces experimental trend as shown along voltage in **Fig. 5.9 (a)** and temperature in **Fig. 5.9 (b)**. As previously mentioned, the strong avalanche effect makes it impossible to measure current above V_{BD} without avalanche. This is supported by dark current modeling results, as illustrated in **Fig. 5.10**, where the current without avalanche is significantly lower than the current with avalanche. The dynamics of the voltage can be explained by the critical role played by multiplication caused by the impact-ionization of carriers.

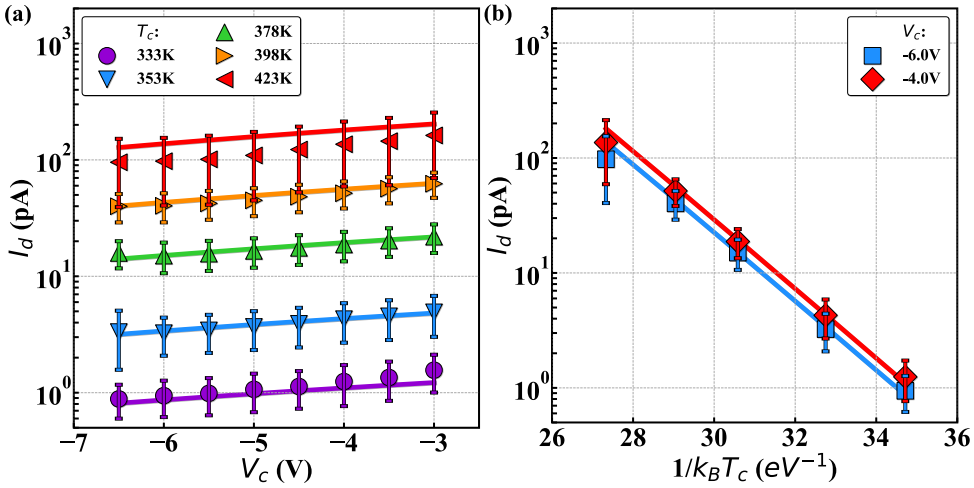


Figure 5.9: (a) Dark current as a function of V_c below V_{BD} measured (symbols) and modeled (lines) at different T_c. (b) Current as a function of $1/k_BT_c$ measured (symbols) and modeled (lines) at different V_c. Error bars represent the standard deviation between each device. The thermal dependence of dark current can be explained by the SRH R-G rate of carriers, while the voltage dependence is attributed to the multiplication caused by the impact-ionization of carriers.

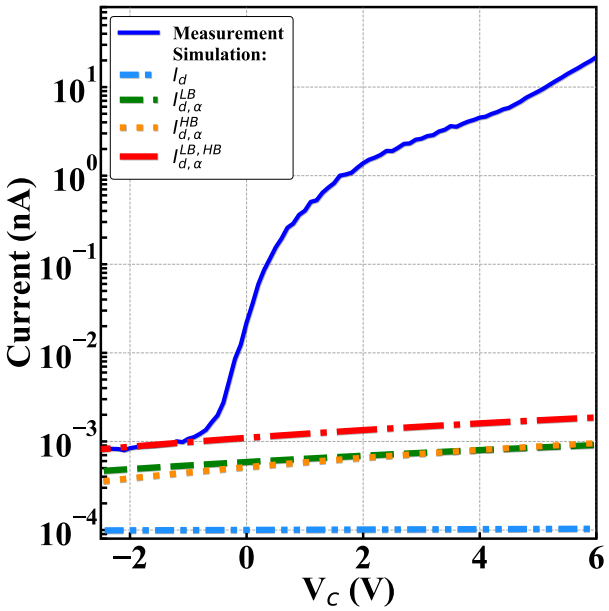


Figure 5.10: Dark current as a function of V_c below or above V_{BD} measured (bold lines) and modeled (dashed/dotted lines) considering dark current without or with carrier impact-ionization within LB, HB and both regions. Measuring current above V_{BD} without avalanche is impossible due to the strong avalanche effect.

2.3.3 Modeling of the photo-generated current

The photo-generation rate is modeled according to the Beer-Lambert law [188] and assuming all photo-generated carriers are collected in the depletion region which leads to:

$$G_{ph} = qF_0(e^{-\alpha y_0} - e^{-\alpha y_1}) \tag{5.32}$$

where α is the absorption coefficient per unit length, F_0 is the incident monochromatic photon flux (which is the irradiance reported in **Table 4.1** converted in number of photons per second and surface unit), and y_0 and y_1 are the absorption depth at the surface and at the edge of the depleted volume respectively. The absorption coefficient α in indirect band gap semiconductor is proportional to [183]:

$$\alpha \propto \frac{(h\nu - E_g + E_p)^2}{exp(\frac{E_p}{k_B T}) - 1} + \frac{(h\nu - E_g - E_p)^2}{1 - exp(\frac{-E_p}{k_B T})} \tag{5.33}$$

The first term relates to phonon absorption and the second one to phonon emission with E_p the phonon energy and ν the light frequency considering the light wavelength defined at 940nm. Considering **eqs. 5.32-5.33** and **eqs. 5.21-5.23** with parameters reported in **Table 5.2**, the photo-generated current is equal to G_{ph} multiplied by q and the effective area of the space charge region (V_{SCR}) plus the photo-induced multiplication current:

$$I_{ph} = q \cdot (G_{ph} \cdot V_{SCR} + \alpha_{n,p} \cdot G_{ph} \cdot \frac{V_{SCR}}{S_{SCR}}) \tag{5.34}$$

The photo-generated modeling shows good agreement with experiment as shown along voltage in **Fig. 5.11 (a)** and temperature in **Fig. 5.11 (b)**. A distinction in current magnitude is observed above V_c=-5V and despite identical irradiance level which relates to carrier multiplication in the depletion region as expected from electric field simulations (red line and orange box in **Fig. 2.8**a).

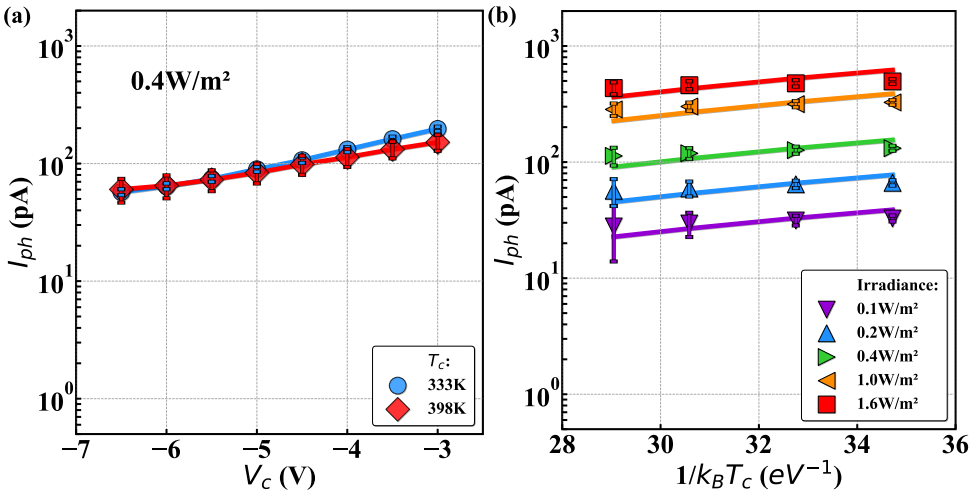


Figure 5.11: (a) Photo-generated current as a function of V_s below V_{BD} (without avalanches) measured (symbols) and modeled (lines) at different T_s and at 0.4W/m². (b) Current as a function of $1/k_BT_s$ measured (symbols) and modeled (lines) at V_s=-4V for several irradiance levels. Error bars represent the standard deviation between each device.

Conclusions

The defect profile at the top interface, which was selected for DCR, can also be used to capture the dark current, taking into account the impact-ionization multiplication of carriers. The results obtained from the modeling approach have provided additional confirmation that the current without avalanches is completely overshadowed by the presence of avalanches. The behavior of photo-generated carriers highlights the significance of carrier multiplication as a function of the applied voltage. The modeling approaches for both dark- and photo-generated current exhibit good agreement with experimental data. These approaches will be employed to determine the quantity of hot-carriers and model the degradation rate.

2.4 Selected degradation kinetics along stress for the creation of defects

The defect creation dynamics must be explained to capture ΔDCR along stress time accounting for the $Si-H$ bond dissociation energy distributions modulated by hot-carrier collisions with bonds. The bond dissociation energy E_{hb} depends on the local variation of electrical and chemical potential as explained in **Section 1**. Disorder-induced variations among these $Si-H$ bond dissociation energies E_{hb} at the Si/SiO_2 interface have been shown [189][190] to be a plausible degradation source of the sub-linear time dependence of the defect creation process. Assuming that concentration of $Si-H$ bond precursors $N_{hb}(t)$ follows a first-order rate equation during stress neglecting the backward (passivation) rate (see **eq. 2.4**) and that defect creation events are independent yields:

$$\frac{dN_{hb}}{dt} = -\kappa N_{hb} \tag{5.35}$$

The analytical solution of the above equation gives:

$$N_{hb}(t) = N_{hb}^{0} exp(-\kappa(t)t)) \tag{5.36}$$

N_{hb}^{0} is the initial concentration of Si-H bond precursors. The corresponding portion of the interface defect density N_t(t) along stress time is found to be:

$$N_t(t) = N_{hb}^{0} - N_{hb}(t) \tag{5.37}$$

Including **eq. 5.36** gives:

$$N_t(t) = N_{hb}^{0}(1 - exp(-\kappa(t)t)) \tag{5.38}$$

To accurately model the number of dangling bond defects resulting from the breaking of Si-H bonds, it is necessary to incorporate the distribution of Si-H bond-dissociation energies. The distribution of disorder-induced variations of Si-H bond-dissociation energies is assumed to follow a Fermi-derivative distribution [178][177]:

$$\tilde{g}(E_{hb},t) = \frac{N_{hb}(t)}{\sigma_{hb}} \frac{exp(\frac{E_{th}-E_{hb}}{\sigma_{hb}})}{[1+exp(\frac{E_{th}-E_{hb}}{\sigma_{hb}})]^2} \tag{5.39}$$

E_{th} and σ_{hb} are the mean and the Half Width at Half Maximum (HWHM) of the bond dissociation energy distribution respectively. The required bond-dissociation energy E_{hb} to break a $Si-H$ bond does not correspond to the defect energy level E_t resulting from this process. The change in $\tilde{g}(E_{hb},t)$ along stress time is considered by the concentration of $Si-H$ bond precursors $N_{hb}(t)$ that reduces proportionally with the creation of dangling bond defects $N_t(t)$ as illustrated by **eq. 5.37**. As explained in **Section 1**, the bond dissociation rate constant κ must account for the cumulative ability of the carrier ensemble to dissociate bonds by impact-ionization and for the Arrhenius process expressing that reactions proceed exponentially faster as the temperature is raised which results in:

$$\begin{aligned}\kappa(t) = \int_0^{+\infty} \int_{E_{hb}}^{+\infty} \frac{1}{q} \cdot (j_{d,\alpha}^{total}(t) + j_{ph,\alpha}^{total}) \cdot \frac{n(E)}{n_{total}} \\ \cdot \sigma_{II}(E,E_{hb}) \frac{\tilde{g}(E_{hb},t)}{\tilde{g}_{total}(t)} exp(\frac{-E_{hb}}{k_b T}) dE dE_{hb}\end{aligned} \tag{5.40}$$

where the integration is performed over the bond energy E_{hb} and over the carrier energy E starting from E_{hb}. $n(E)$ is integrated over the degradation area. n_{total} stands for the total carrier energy distribution and $\tilde{g}_{total}(t)$ for the total $Si-H$ bond dissociation energy distribution along stress duration. The Keldysh-like impact-ionization reaction cross section $\sigma_{II} = \sigma_{II}^0 (E-E_{hb})^{\rho_0}$ is used for the SP mechanism triggered by hot carriers with σ_{II} being the attempt rate and ρ_0 the energy exponent of the impact-ionization reaction dynamics fitted on HCD in NMOS [171]. The attempt rate σ_{II}^0 drives the degradation temporal dynamics and the number of defect precursors N_{hb} the degradation magnitude. This model gives a comprehensive explanation of the sub-linear time-dependence of the degradation driven by the $Si-H$ bond dissociation energy distribution downsizing by the hot-carriers impact-ionizing the bonds. The model predicts the saturation of defect creation when time tends to infinity due to the exhaustion of defect precursors N_{hb} [191].

In **Fig. 5.12**, σ_{II} is plotted as a function of both the $Si-H$ bond energy E_{hb} and the carrier energy E, with the data ordered from low to high values of each parameter. As the $Si-H$ bond energy increases, the probability of dissociating a $Si-H$ bond by impact-ionization decreases, while it increases with increasing carrier energy. As a result, the distribution of $Si-H$ bond-dissociation energy states becomes exhausted at low E_{hb}, as shown in **Fig. 5.13 (a)**. Additionally, there is an increase in the creation of defects at a given stress time, which corresponds to the bond-dissociation energy states that are available for impact-ionization by hot carriers, as depicted in **Fig. 5.13 (b)**. Consequently, the bond dissociation rate constant κ undergoes a change over the duration of the stress, as depicted in **Fig. 5.13 (c)**.

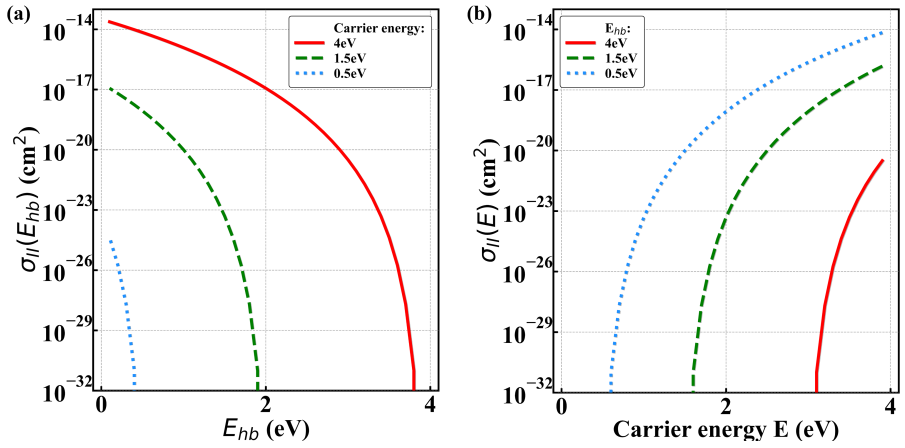


Figure 5.12: σ_{II} as a function of (a) the $Si-H$ bond energy E_{hb} and (b) the carrier energy E from low to high E and E_{hb} respectively. The impact-ionization probability of dissociating a $Si-H$ bond decreases with increasing $Si-H$ bond energy but increases with increasing carrier energy. The increase in the creation of defects at a given stress time corresponds to the bond-dissociation energy states that are available for impact-ionization by hot carriers.

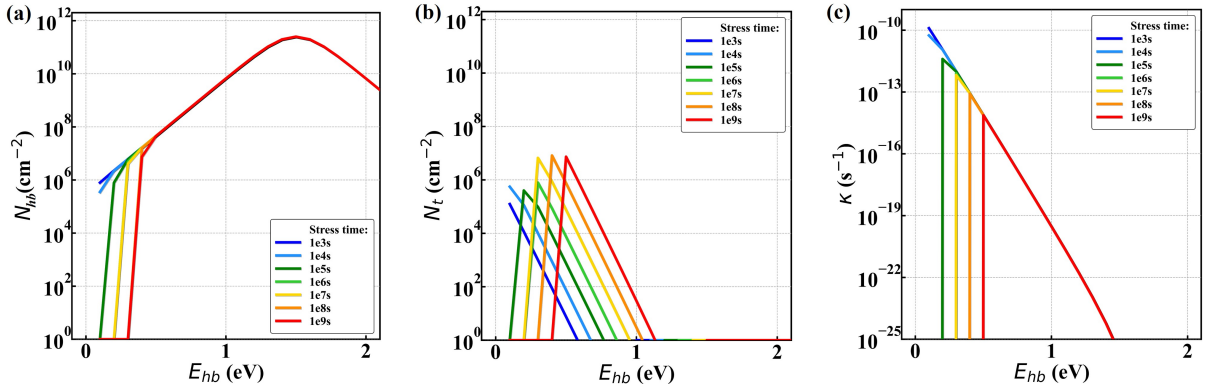


Figure 5.13: (a) $Si-H$ bond-dissociation energy distribution N_{hb},(b) defect distribution N_t and (c) bond dissociation rate constant κ as a function of the $Si-H$ bond energy E_{hb} along stress duration. The impact-ionization probability results in the depletion of $Si-H$ bond-dissociation energy states at low E_{hb}.

Conclusions

A first-order rate equation in the concentration of Si-H bond precursors $N_{hb}(t)$ is used to model the degradation and extract the number of defects created during stress. The bond dissociation rate constant κ takes into account the distribution of disorder-induced variations of Si-H bond-dissociation energies against the carrier energy distribution, which is used to calculate the bond distribution exhaustion during stress. Additionally, the interaction between carriers and bonds is considered by incorporating the current density and the bond dissociation probability through the impact-ionization reaction cross section σ_{II}. The definition of σ_{II} that takes into account both the $Si-H$ bond energy E_{hb} and the carrier energy E predicts a more significant depletion of $Si-H$ bond-dissociation energy states at low E_{hb}, resulting in higher defect creation.

2.5 Modeling of the Dark Count Rate drift

Once the kinetics of defect formation have been identified, calculating the generation rate from stress-induced defects becomes necessary in order to determine the extent of Dark Count Rate (DCR) drift over the duration of the stress. The same previous hypothesis stands for the defect creation regarding the initial defect generation rate of carriers (**eq. 5.24**) at the top depleted interface and at a single-energy mid-gap level E_t. At the considered defect density N_t, the presence of traps inside the device does not significantly alters the device electrostatic profile thus P_t is assumed to be constant along stress time. According to these assumptions, the DCR created along stress time can be rewritten in a simple form by replacing N_t^0 by $N_t(t)$ in **eq. 5.24** as follows:

$$\begin{aligned} DCR(t) =& P_t \cdot G_d(t) = P_t \cdot N_{hb}^0(1 - exp(-\kappa(t)t)) \\ & \cdot \frac{n_i \sigma_t v_{th}^n v_{th}^p}{v_{th}^n exp(\frac{E_t - E_F^i}{k_B T}) + v_{th}^p exp(\frac{E_F^i - E_t}{k_B T})} \end{aligned} \tag{5.41}$$

The modeling methodology is summarized in **Fig. 5.14**. The characterization effect is interpreted from the first and the third term of **eq. 5.41** which are the avalanche breakdown probability P_t and the carrier generation rate from created defects along stress duration divided by the number of created defects. P_t is simulated along V_c and T_c (see **Fig. 2.12**). The intrinsic concentration n_i is calculated from **eqs. 2.86-2.89** and the thermal velocity v_{th} from **eq. 2.101** as a function of T_c. The capture cross section σ_t is computed from **eq. 5.30** as a function of V_c depending on the simulated electric field F (**Fig. 2.8 (a)**). n_i, v_{th} and σ_t are injected into **eq. 5.24** to calculate the generation rate of carriers from stress-induced defects.

The stress impact is captured by the second term of **eq. 5.41** which is the number of defects N_t(t) created along stress time from **eq. 5.38**. The time-dependent bond dissociation rate constant $\kappa(t)$ is computed from **eq. 5.40**. The carrier energy distribution $n(E)$ is simulated by the Full-Band Monte-Carlo (described in **Section 1.12**) along V_s and T_s range. The dark- and photo-generated current densities are modeled from **eq. 5.24** (**Fig. 5.9**) and **eq. 5.32** (**Fig. 5.11**) respectively coupled with the multiplication current from **eq. 5.21** considering the simulated electric field F (**Fig. 2.8 (a)**). The degradation rate is adjusted based on the modeled dark current. The Si-H bond dissociation energy distribution $\tilde{g}(E_{hb}, t)$ is calculated from **eq. 5.39** at each iteration to account for the exhaustion of defect precursors by dangling-bond defects.

This methodology is now confronted to experimental results. The degradation model is firstly compared to various stress voltage and temperature conditions across different characterization voltage and temperature conditions. For stresses in darkness and under illumination at different stress voltages and a constant stress temperature (first and second rows in **Table 5.1**, a good correlation is obtained between the proposed model and the experiment as shown in **Fig. 5.15 (a)-(b)**. From characterization perspective, the voltage dependence of avalanche breakdown probability enables the replication of the characterization trend in both

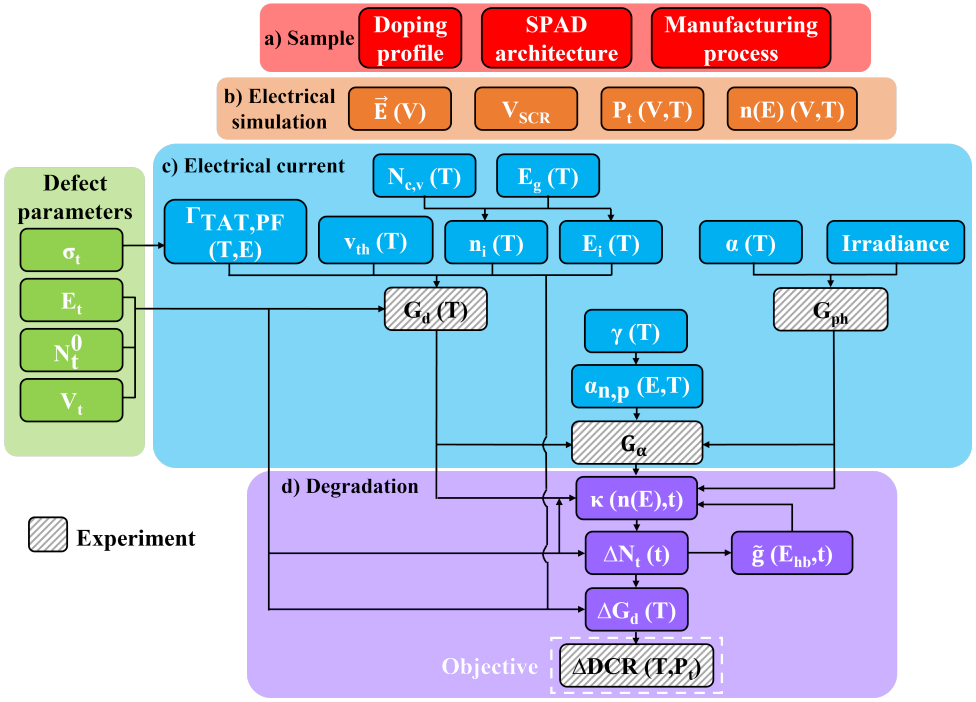


Figure 5.14: Proposed methodology to predict ΔDCR based on: (a) SPAD architecture and process to simulate (b) electric field F, volume of space charge region V_{SCR}, avalanche breakdown probability P_t and carrier energy distribution $n(E)$ integrated with (c) dark-, photo- and multiplication rates $G_{d,ph\alpha}$ to calculate (d) the bond dissociation rate constant κ. See parameter description in the text (also shown in **Table 5.2**). Only first order dependencies are shown in parenthesis such as the bond dissociation rate constant κ which depends on the energy carrier distribution $n(E)$ (first order) which depends on V and T (second order). Defect parameters are arbitrarily fixed. Variables in dashed grey boxes are fitted with experimental data.

the LB (V_c=3V) and HB (V_c=5V) regions. From stress perspective, the voltage dependence of hot-carrier multiplication enables the replication of the experimental trend in both dark and irradiance conditions.

For stresses in darkness and under illumination at different stress temperatures and a constant stress voltage, a good correlation is obtained between the proposed model and the experiment as shown in **Fig. 5.15 (c)-(d)**. The characterization trend in the LB (V_c=3V) and HB (V_c=5V) regions is captured by the exponential thermal dependence of SRH generation rate. Disregarding temperature characterization effect due to the SRH generation rate of carriers which doubles each 10K in the temperature difference and the avalanche breakdown probability P_t each 40K, the number of defects $N_t(t)$ created along stress time doubles each 40K and 30K in darkness and under light respectively. The degradation increases as the temperature increases which is contrary to the behavior expected for generic SP in long-channel MOSFETs [159]. The degradation rates depend on the carrier energy and density. The carrier energy loss dynamics by scattering mechanisms increasing with temperature are the same from MOSFETs to SPADs as shown in **Fig. 5.5 (d)-(e)** against **Fig. 5.3**. However in MOSFET transistors, the carrier density is proportional to the gate voltage minus the threshold voltage (which is

approximately linear with temperature) whereas in SPADs is following the diode reverse current (which increases exponential with temperature through the SRH generation rate) as shown in **Fig. 5.9 (b)**.

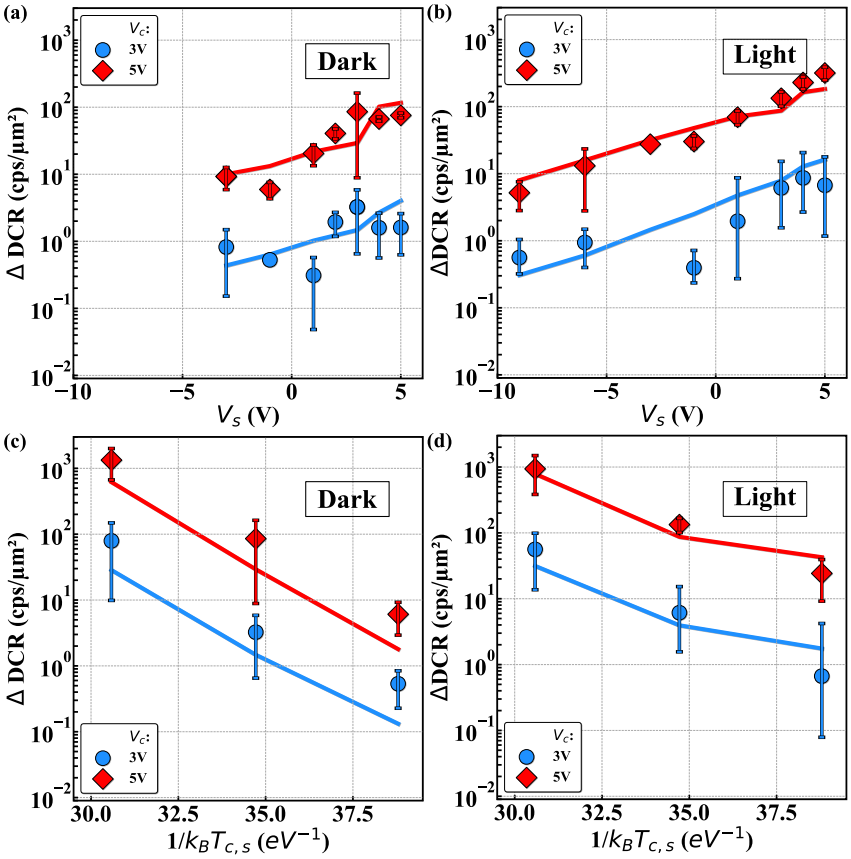


Figure 5.15: ΔDCR as a function of V_s after stresses (a) in darkness and (b) at 0.3W/m² and at T_s=333K measured (symbols), and modeled (lines) at T_c=333K and at V_c=3V (LB - blue) and 5V (HB - blue). ΔDCR as a function of $1/k_BT_s$ for stresses (c) in darkness and (d) at 0.3W/m² for V_s=3V measured (symbols), and modeled (lines) at T_c=T_s and at V_c=3V (LB - blue) and 5V (HB - red). Error bars represent the standard deviation between each device for the same stress condition.

Furthermore, the correlation plots of LB and HB region shown in Fig. **Fig. 5.16** enlighten the appropriate relationship achieved for all stress conditions reported in **Table 5.1**.

In the existing literature, HCD have been extensively studied under short stress time to asymptotically approximate the solution of the first order kinetic equation by a power law [185] which is to said that created defect do not modify significantly the bond dissociation rate constant κ. In this study, the stress time range is extended up to 10^6s to probe long-term device aging behavior. The modeled stress temporal dynamics characterized in both region is in good agreement with the experiment as shown in **Fig. 5.17**. The time-dependent degradation is effectively captured by considering the time dependence of the first-order kinetics equation for $Si-H$ bond dissociation, which incorporates the modulation of the bond dissociation rate constant by the exhaustion of $Si-H$ defect precursors. Nevertheless, a slight discrepancy is noticed at low stress times below 10^4s may arise from the use of a single $Si-H$ bond distribution

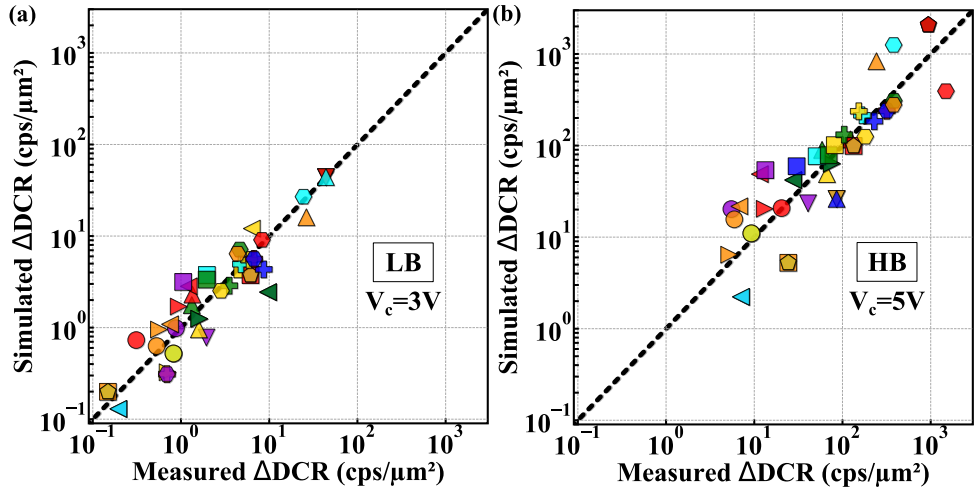


Figure 5.16: ΔDCR correlation plots between experiment and modeling measured at (a) V_c=3V (LB) and (b) 5V (HB) for each stress condition reported in Table 5.2. Each symbol corresponds to a different stress conditions averaged on eight devices. The dashed line is the y=x line to underline the correlation between experiment and modeling.

to model degradation kinetics, which may not be accurate for the micrometer squared area under investigation. Additionally, measuring low count rate drift can be difficult due to the statistical dispersion from device-to-device, where the ΔDCR mean follows a quadratic law with its variance as explained in **Chapter 4 Section 6.2**.

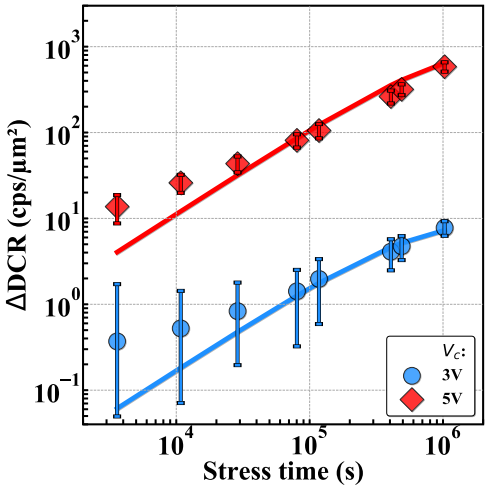


Figure 5.17: ΔDCR as a function of the stress time for stresses on 64 SPADs in darkness at V_s=3V, V_s=353K and at 0.3W/m² measured (symbols) and modeled (lines) at V_c=3V (LB - blue) and 5V (HB - red) for T_c=333K.

2.6 Modeling of the dark current drift

The determination of the generation rate from stress-induced defects can also aid in assessing the drift in dark current over the duration of the stress. This dark current along stress time is established by substituting N_t^0 with $N_t(t)$ in equation **eq. 5.24**, while keeping all other

parameters constant, together with **eq. 5.42**. The hypothesis previously stated regarding defect creation also applies to the initial generation rate of carriers (**eq. 5.24**) at the top depleted interface and at a single-energy mid-gap level E_t. These assumptions yield to:

$$I_d(t) = q \cdot N_{hb}^0(1 - exp(-\kappa(t)t)) \cdot \frac{n_i \sigma_t v_{th}^n v_{th}^p}{v_{th}^n exp(\frac{E_t - E_F^i}{k_B T}) + v_{th}^p exp(\frac{E_F^i - E_t}{k_B T})}(V_d + \alpha_{n,p} \cdot \frac{V_d}{S_{SCR}}) \quad (5.42)$$

It is observed that the dark current increased proportionally with the number of defects created over time, as demonstrated in **Fig. 5.18**. This implies that the defect profile and localization remain unchanged before and after stress, similar to ΔDCR. These findings suggest that the inclusion of dark current drift in the calculation of κ using equation **eq. 5.40** is necessary.

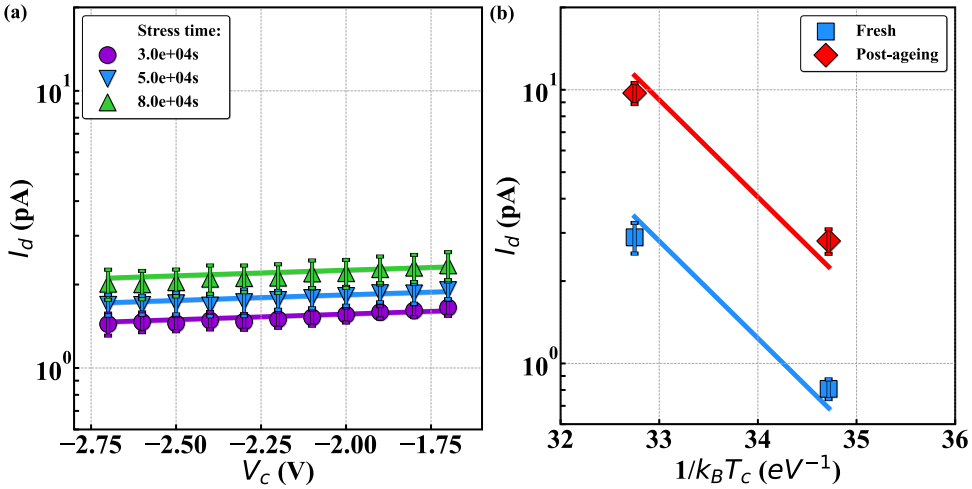


Figure 5.18: (a) Dark current as a function of V_c below V_{BD} measured (symbols) and modeled (lines) at different T_c for an increasing stress duration. (b) Dark current as a function of $1/k_BT_c$ measured (symbols) and modeled (lines) at V_c=-2V before (rectangles) and after (diamonds) stress. The stress was performed at stress voltage V_s=5V, irradiance of 0.3W/m² at 940nm and temperature T_s=373K. Error bars represent the standard deviation between each device. The assumption is made that the thermal dynamics are identical to the ones depicted in **Fig. 5.9 (b)**. The proposed degradation model suggests that as more defects are created over time, there is a proportional rise in the dark current density.

Conclusions

The developed model is capable of accurately describing the hot-carrier degradation that induces ΔDCR in SPADs across a wide range of operating conditions. The number of available hot carriers is estimated by analyzing the current density in both darkness and under irradiance, using a full-band Monte-Carlo simulation to determine the carrier energy distribution. The main contributors to the total defect creation are identified as hot electrons. The sub-linear time dependence of the defect creation process is captured by the exhaustion of the bond dissociation energy distribution due to impact-ionizing hot carriers. The two different slopes observed in hot-carrier degradation are explained by the degradation occurring in two different regions of the device architecture. In contrast to long-channel MOSFETs, the carrier density in SPADs increases exponentially with the applied voltage, leading to an increase in degradation as a function of temperature. The model presented is capable of accurately predicting the increase in dark current that occurs during stress. This is achieved by incorporating the dark current drift into the calculation of κ.

3 Conclusions

1

For an accurate model of the HCD mechanism, it is necessary to consider the carrier energy distribution function, the distribution of $Si-H$ bonds, and the dissociation process. The SP mechanism can induce damage such as Si-H bond break and oxide border trap due to high-energetic carriers. The mix mode mechanism involves two electrons creating one defect, and can also result in $Si-H$ bond break, thus in depends on both carrier energy and flux. For the MVE process, the electron energy is much lower than the energy required to break $Si-H$, so only the carrier flux is important. The impact is more severe at low temperatures when considering carrier energy. Conversely, when considering carrier flux, the impact is more severe at high temperatures. The bond-dissociation energy can be affected by the hydrogen density and the transversal component of the electric field, which can modulate the local electrostatic potential and stretch the bond. Another approach is to consider the disorder-induced local variations in $Si-H$ bond energy at the Si/SiO_2 interface, which can match the observed sub-linear time dependence of the degradation.

2

Full-Band Monte-Carlo (FBMC) allows to extract the carrier energy distribution $n(E)$ within the device accounting for the scattering mechanisms and the carrier acceleration. Some energy-tail electrons reach the mean threshold dissociation energy of $Si-H$ bond (E_{th}=1.5eV) depending on V_s and T_s. There are two potential degradation dynamics that can occur depending on the localization of the degradation, either in the n++/p+ region or in the n-/p-well region. The two distinct degradation regions can be selectively stimulated depending on the conditions of characterization.

3

Due to the large avalanche current, it is not possible to measure the current above V_{BD} without avalanches. The constant value of the CPP along the stress duration confirms this fact. It is assumed that the current above V_{BD} without avalanches is proportional to the current measured below V_{BD}, taking into account the multiplication rate. The degradation rate is dependent on the current without avalanches, as there is a resilient relationship between the change in ΔDCR and the current measured without avalanches. This relationship appears to be related to the increasing number of hot carriers that enhance $Si-H$ bond dissociation.

4 The defect profile at the top interface, which was selected for DCR, can also be used to capture the dark current, taking into account the impact-ionization multiplication of carriers. The results obtained from the modeling approach have provided additional confirmation that the current without avalanches is completely overshadowed by the presence of avalanches. The behavior of photo-generated carriers highlights the significance of carrier multiplication as a function of the applied voltage. The modeling approaches for both dark- and photo-generated current exhibit good agreement with experimental data. These approaches will be employed to determine the quantity of hot-carriers and model the degradation rate.

5 A first-order rate equation in the concentration of Si-H bond precursors $N_{hb}(t)$ is used to model the degradation and extract the number of defects created during stress. The bond dissociation rate constant κ takes into account the distribution of disorder-induced variations of Si-H bond-dissociation energies against the carrier energy distribution, which is used to calculate the bond distribution exhaustion during stress. Additionally, the interaction between carriers and bonds is considered by incorporating the current density and the bond dissociation probability through the impact-ionization reaction cross section σ_{II}. The definition of σ_{II} that takes into account both the $Si-H$ bond energy E_{hb} and the carrier energy E predicts a more significant depletion of $Si-H$ bond-dissociation energy states at low E_{hb}, resulting in higher defect creation.

6 The developed model is capable of accurately describing the hot-carrier degradation that induces ΔDCR in SPADs across a wide range of operating conditions. The number of available hot carriers is estimated by analyzing the current density in both darkness and under irradiance, using a full-band Monte-Carlo simulation to determine the carrier energy distribution. The main contributors to the total defect creation are identified as hot electrons. The sub-linear time dependence of the defect creation process is captured by the exhaustion of the bond dissociation energy distribution due to impact-ionizing hot carriers. The two different slopes observed in hot-carrier degradation are explained by the degradation occurring in two different regions of the device architecture. In contrast to long-channel MOSFETs, the carrier density in SPADs increases exponentially with the applied voltage, leading to an increase in degradation as a function of temperature. The model presented is capable of accurately predicting the increase in dark current that occurs during stress. This is achieved by incorporating the dark current drift into the calculation of κ.

Chapter 6

Assessing the trade-offs in direct Time-of-Flight Measurements

Objectives

The objective of this section is to establish a clear relationship between the parasitic noises, including both DCR and ambient light, and the photon signal reflected back from the target. This will be accomplished by considering factors such as target reflectance and the distance between the sensor and the target. The section will also clarify the attenuation of light intensity and the statistical fluctuations of light. To achieve this, the section will investigate various mechanisms that can disable signal detection at different distances from the target, such as optical crosstalk or pile-up. Schematics and simulations will be utilized to describe these mechanisms, while disregarding the external circuitry or photonic algorithm used to process the signal.

Contents

1 Understanding the optical system: components and principles

In order to comprehend the distance measured between the sensor and the target based on the emitted photons, it is necessary to introduce the entire system along with its surrounding. Previously, only a single SPAD device was taken into account, but now the entire pixel array will be considered. A typical illustration of the optical system is described in **Fig. 6.1**. The system is composed of:

1. a Vertical-Cavity Surface-Emitting Laser (VCSEL) which emits light vertically from the surface
2. a reference SPAD array which provides a reliable reference signal for time zero for the return SPAD array
3. a cover glass which protects the underlying components from damage and contamination by providing a barrier against dust or moisture and which improves the optical performance of the system by reducing reflections and glare
4. a return SPAD array which detects the photons that are reflected back to the detector from the target allowing to measure the distance to the target

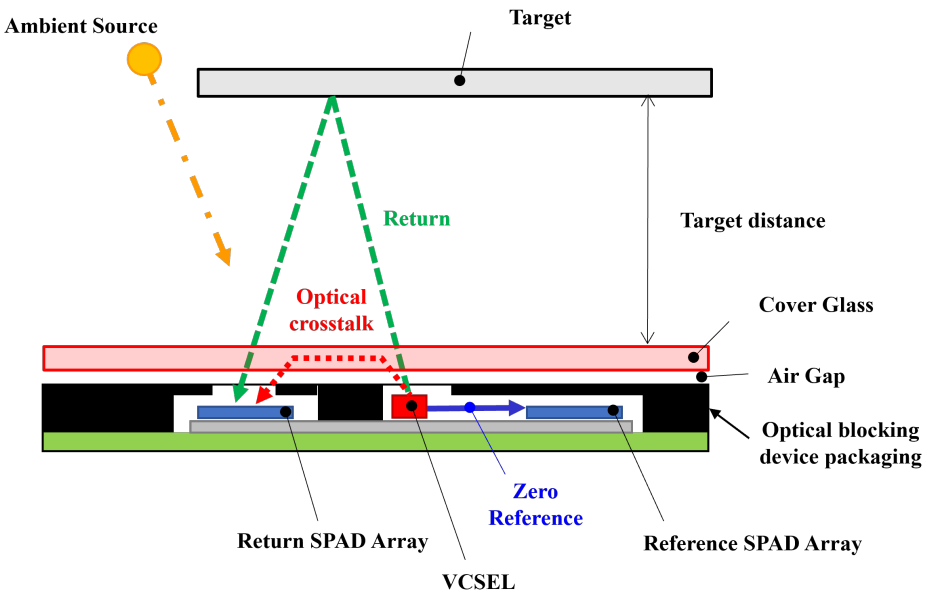


Figure 6.1: Schematic of the optical system including: VCSEL, reference and return SPADs, cover glass and target. The reference signal is depicted by the blue bold line, the returning signal is represented by the green dashed while the ambient light and optical crosstalk are represented by the orange dashed-dotted and red dotted lines, respectively. The lines depicted do not accurately represent the optical path of the light.

The working principle of a direct Time of Flight (dToF) measurement for a single SPAD is as follows. A laser pulse is emitted from the VCSEL towards a target, and the time taken for

the pulse to reflect back to the sensor is measured. The distance is calculated by multiplying the time taken for the pulse to travel to the target and back by the speed of light in the medium through which the pulse is traveling. To adjust for the round-trip travel of light to the target and back, this product is halved. The pixel array is then utilized to construct the histogram of the photon arrival time, which is achieved by considering each readout. This counting histogram is then used to reconstruct the depth map of the scene. The time between two VCSEL pulses, also known as the pulse repetition period, determines the maximum range that can be measured. It must be long enough to allow the sensor to detect the returning photons before the next pulse is emitted. If the pulse repetition period is too short, the returning photons from the previous pulse may still be arriving at the sensor when the next pulse is emitted, resulting in wrap-around errors in the range measurement.

Under ideal conditions, a linear relationship between the measured distance $d_{measured}$ and the distance from the target d_{target} exists when considering only the photons emitted from the VCSEL and reflecting back to the sensor from the target as illustrated in **Fig. 6.2**. As depicted in **Fig. 6.1**, the system comprises various components, including the cover glass, which can reflect some of the photons emitted from the VCSEL directly back to the sensor. This reflection results in optical crosstalk, which is illustrated by a second histogram appearance in the right-hand side corner of **Fig. 6.2**, and reduces the maximum distance that can be measured to the target. Additionally, depending on the distance to the target d_{target} and its reflectance R_{target}, multiple photons can arrive at the detector within the same time interval, making it impossible for the detector to distinguish between them. This leads to the piling up of the system, as illustrated in the left-hand side corner of **Fig. 6.2**, and distorts the counting histogram, thereby reducing the minimum distance that can be measured to the target. The optical crosstalk and pile-up mechanisms with other phenomenon will be described in details in the following.

Nevertheless, the distance range may contain some level of uncertainty as illustrated in **Fig. 6.2** because it is an average obtained from the counting histogram. This uncertainty is caused by the statistical characteristics of light. The accuracy of distance measurement can be improved by repeating the light pulse. The histogram accumulates the received events from each pulse, rather than averaging separate histograms from each pulse. To comprehend how light is impacted by photon statistical dispersion and distance, it is crucial to have a clear understanding of these concepts.

1.1 Intensity of light as a function of target distance

The reason why noise goes with the square of the distance is due to the way that energy dissipates over distance. The irradiance can be expressed as:

$$I_L = \frac{P}{A} \tag{6.1}$$

where P is the power of the source. The circular spot of the laser is characterized by its area A, which is given by:

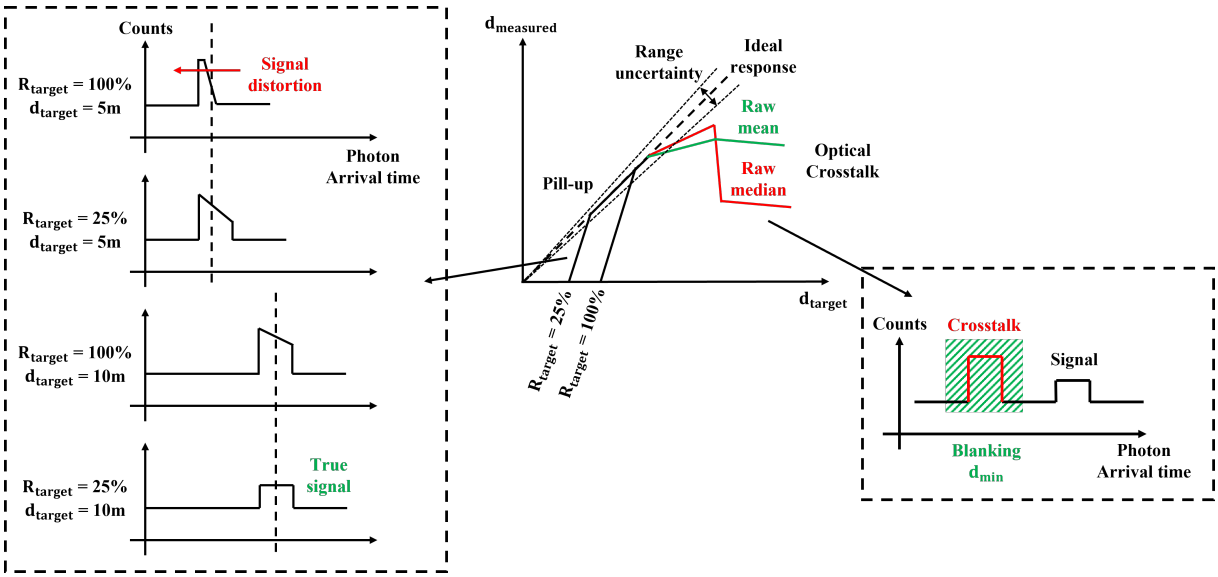


Figure 6.2: Maximum distance measured d_{max} as a function of the distance to the target d_{target} considering pile-up and optical crosstalk mechanisms. On the left-hand side, the accumulation of multiple photons (pile-up) can reach the detector simultaneously, making it difficult for the detector to differentiate between them, decreasing the minimum measurable distance. One of the primary causes of this mechanism is the reflectance of the target R_{target}. On the right-hand side, the optical crosstalk caused by photons emitted from the VCSEL that reflect directly back to the sensor decreases the maximum measurable distance.

$$A = \pi w^2 \tag{6.2}$$

where w is the radius of the spot. As per Gaussian beam theory, the spot radius of the laser increases linearly with distance beyond the Rayleigh range, while within the Rayleigh range, the spot size still increases, but at a sublinear rate. The Rayleigh range Z_R is the distance over which the beam waist size remains approximately constant before it starts to diverge as depicted in **Fig. 6.3**. The Rayleigh range Z_R can be determined using the following expression:

$$Z_R = \frac{\pi w_0^2}{\lambda} \tag{6.3}$$

where w_0 is the beam waist size which is the narrowest spot size along the beam path, and λ is the wavelength of the laser beam.

The half-angle of the linear divergence after the Rayleigh range is given by:

$$\frac{\theta}{2} = \theta = \frac{\lambda}{\pi w_0} \tag{6.4}$$

The beam radius w is given by:

$$w = r \cdot tan(\theta) \tag{6.5}$$

So the area of the laser spot is:

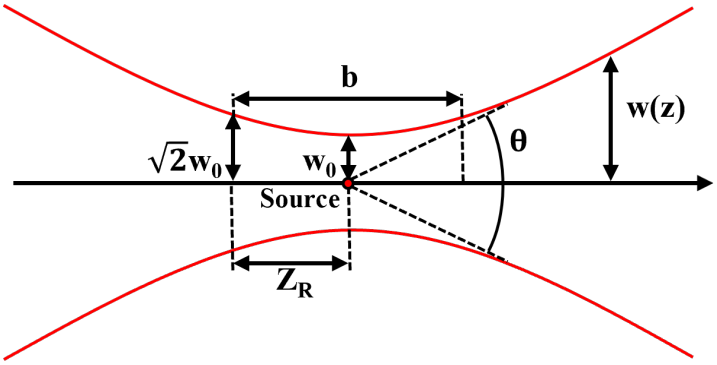


Figure 6.3: Diagram of Gaussian beam waist parameters. The Rayleigh range, denoted by Z_R, is the distance over which the size of the beam waist remains relatively constant before it begins to diverge.

$$A = \pi w^2 = \pi(rtan(\theta))^2 \tag{6.6}$$

Inserting **eq. 6.6** into **eq. 6.1**, the irradiance is given by:

$$I_L = \frac{P}{A} = \frac{P}{\pi(rtan(\theta))^2} = \frac{1}{tan^2(\theta)} \cdot \frac{P}{\pi r^2} \tag{6.7}$$

The equation shows that the light intensity is directly proportional to the power of the light source and inversely proportional to the area of the target. The equation also demonstrates that the light intensity decreases as the angle of incidence increases, following a square law with the tangent of the angle. Notice that the laser irradiance has exactly the same form as for a point source at a distance r:

$$I_P = \frac{P}{4\pi r^2} = \frac{1}{4} \cdot \frac{P}{\pi r^2} \tag{6.8}$$

The only difference is the constant multiplier, $\frac{1}{4}$ versus $\frac{1}{tan^2(\theta)}$, resulting in more intense light for the same amount of power.

1.2 Photon statistical fluctuations

After clarifying the attenuation of light over distance, the focus now shifts to the statistical fluctuations of photon arrival times. It occurs because light is made up of discrete particles called photons, which arrive at the detector at random intervals. This randomness in the arrival time of photons results in fluctuations in the intensity of the light signal, which is referred to as light shot noise. The probability of a certain number of events occurring within a fixed time interval can be described by the Poisson distribution, provided that the events occur independently and at a constant rate. The events correspond to the emission and detection of photons, and the rate corresponds to the average number of photons emitted or detected per unit time. This can be written by the probability p of a certain number of arrivals i for a specified constant rate of arrivals s:

$$p(i) = exp(-s) \cdot \frac{s^i}{i!} \tag{6.9}$$

The irradiance, indicated by s, represents a specific number of photons per unit area per unit time. When integrated over a given period, such as the exposure time, s denotes the number of photons per unit area, which is referred to as exposure. A fundamental characteristic of the Poisson distribution is that, for a given mean rate (signal s), the standard deviation of random arrivals (shot noise σ_s) is proportional to the square root of the signal:

$$\sigma_s = \sqrt{s} \tag{6.10}$$

For example, suppose a mean arrival rate of 1000 photons per second. The shot noise standard deviation would be equal to the square root of the mean, which is 32 photons per second. This means that the actual number of photons detected per second would fluctuate around the mean by approximately 32 photons, which is a relatively insignificant deviation from the mean. Now, let's consider a mean arrival rate of 10 photons per second. In this case, the shot noise standard deviation would be equal to the square root of the mean, which is 3 photons per second. This means that the actual number of photons detected per second would fluctuate around the mean by approximately 3 photons, which is a significant deviation from the mean. These examples demonstrate how light shot noise is particularly important in low light conditions, where the number of photons detected is small, and the statistical fluctuations can significantly affect the accuracy and precision of optical measurements. It is worth nothing to remind that accuracy refers to how close a measurement is to the true value. And precision, on the other hand, refers to how consistent or reproducible a measurement is.

1.3 Optical cross talk

After discussing the attenuation and statistical dispersion of light, the focus will now shift towards optical crosstalk. Optical crosstalk when the cover glass reflects the emitted signal from a source back to the return SPAD array, as illustrated by the red dotted lines in **Fig. 6.1**. The short optical path results in the optical crosstalk pulse appearing almost immediately with minimal time delay, which is demonstrated by the second histogram in the right-hand side corner of **Fig. 6.2**. The level of optical crosstalk present depends on various factors, such as the optical setup (e.g. Field of View (FoV), VCSEL-return array separation and VCSEL pulse width), cover-glass geometry, and properties. Additionally, optical crosstalk may vary over time due to external factors such as scratches or dirt on the cover-glass, which can result in smudging. As a consequence, there is a change in the linearity behavior between d_{max} and d_{target} as shown in the **Fig. 6.2**, particularly over long distances from the target, where the attenuation of light is significant.

In order to enhance the understanding of the impact of optical crosstalk on distance ranging, the simulation methodology is introduced beforehand. To simulate the effects of signal and optical crosstalk, the number of photons emitted by the VCSEL, target reflectance, distance between the sensor and target, and integration time of the sensor are defined. A Poisson distribution of

photons emitted by the VCSEL is then generated to represent the true signal that the sensor should detect. The amount of optical crosstalk that the sensor will receive by reflecting off the cover glass and returning to the sensor is calculated based on the cover glass reflectance and the distance between the sensor and the cover glass. Similarly, the amount of light reflecting off the target and returning to the sensor depends on the reflectance of the target and the distance between the sensor and the target. These signals are combined to create a simulated signal that represents what the sensor will detect. The integration time of the sensor is then applied to the combined signal to simulate the sensor's response. The simulation is repeated multiple times to generate statistical data on the accuracy of the sensor's range estimation.

To identify the point at which optical crosstalk becomes significant, one approach is to compare the mean and median values of the signal histogram, as illustrated in **Appendix E Fig. E.1** for increasing distances. At shorter target distances, both the mean and median values increase linearly as the time taken for photons to travel back is less than the time taken for photons to reflect off the covering glass. As the distance between the sensor and target increases, the pulse shape becomes non-rectangular, resulting in a slight offset between the mean and median values of the signal histogram. At distances where photons have enough time to reflect off the covering glass before returning to the sensor, the median value of the signal histogram shifts to the optical crosstalk pulse from the target pulse. In contrast, the mean value still remains between the optical crosstalk and target pulses. This shift in the median value causes the range estimation to change from an over-estimate to an under-estimate as compared to the mean range estimation.

There are various methods available to compensate for optical crosstalk and obtain the true signal. At first glance, one might assume that by subtracting the optical crosstalk histogram from the combined histogram of both optical crosstalk and target returns, only the target returns would remain in the resulting histogram. However, implementing this concept in practice can be challenging. One of the main difficulties is determining the accuracy of the optical crosstalk signal at the time of subtraction. The level of optical crosstalk can vary between different parts and over time due to factors such as scratches and smudges on the cover-glass. Variations in the temporal position of the optical crosstalk pulse can occur due to changes in temperature and applied voltage, which alter the Photo-Detection Efficiency (PDE). Despite subtracting the optical crosstalk from the target signal, the shot noise present in the system cannot be entirely eliminated, which means that the optical crosstalk still persists to some extent.

To achieve optical crosstalk immunity, the target must be positioned beyond a specific distance that enables the separation of the target and crosstalk signals. The blanking of the signal is achieved by temporarily disabling the sensor's detection circuitry during the anticipated time of the light pulse's reflection off the covering glass and return to the sensor. This ensures that only the true signal reflecting from the target is detected, preventing the detection of the optical crosstalk signal. The maximum measurable distance is reduced by this method as it defines the maximum time that can be measured.

1.4 Field-of-View

The level of optical crosstalk is influenced by the Field of View (FoV), which refers to the extent of the observable area that the optical system can capture at a given time, as illustrated in **Fig. 6.4**. It is typically measured in degrees which is represented by θ in **Fig. 6.4**. Expanding the FoV can provide a wider range of observation. However, it can also result in an increase in optical crosstalk.

Furthermore, to accurately measure the distance to a target, it is necessary to convert radial measurements into their perpendicular equivalents. This conversion is important to ensure that the measured distance is not affected by the angle at which the sensor captures the target, which can be influenced by the FoV. This is important when dealing with perpendicular targets, as it is more useful to report the perpendicular ranges for each zone rather than the radial distances. The success of the distance conversion depends on the FoV and the number of zones being taken into account on the SPAD array. For low angles, the radial distance R_i is almost equal to the perpendicular distance H_i, and a minor range gain error can be calculated by taking the weighted average across one zone. However, at high angles, the radial distance R_i becomes greater than the perpendicular distance H_i, leading to a significant difference in range between the corner zones and the center zones when dealing with a perpendicular. This can result in signal distortion due to a longer time delay in radial distance and a weaker signal due to increased distance. To overcome this issue, a scaling factor can be used to accurately reproduce the actual distance.

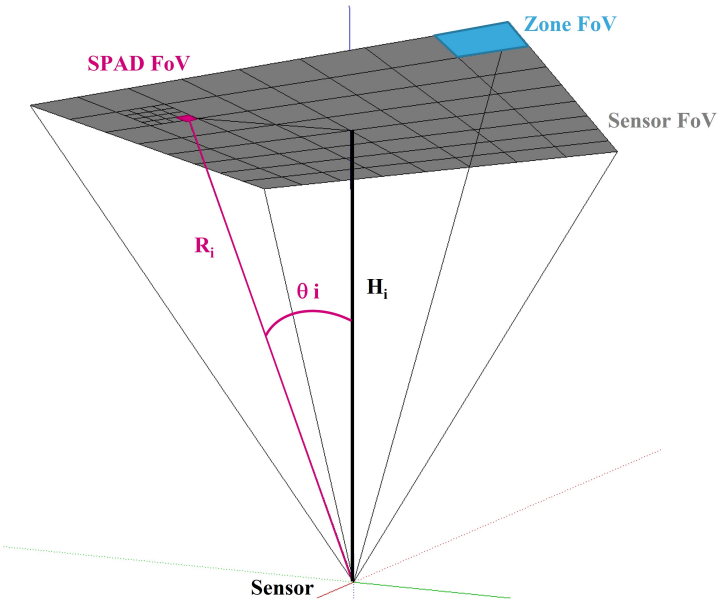


Figure 6.4: The Field of View (FoV) is illustrated in the figure for a SPAD, a zone, and the sensor. The perpendicular distance is denoted as H_i, the radial distance as R_i, and the FoV angle as θ. When the angle is low, the radial distance is nearly equal to the perpendicular distance. However, when the angle is high, the radial distance becomes greater than the perpendicular distance. This can result in a significant difference in range between the corner zones and the center zones when measuring a perpendicular distance.

The proposal is to calculate the average scaling factors for each SPAD in a given zone Z:

$$H_Z = R_Z \cdot SF_Z \tag{6.11}$$

This can be achieved by averaging the cosine of the SPAD angle for each individual SPAD:

$$SF_Z = \frac{1}{|Z|} \sum_{i \in Z} cos(\theta_i) = Median_{i \in Z}(cos(\theta_i)) \tag{6.12}$$

Another option is to calculate the average angle for all SPADs in the zone and then take the cosine of this average angle:

$$SF_Z = cos(\frac{1}{|Z|} \sum_{i \in Z} \theta_i) = cos(Median_{i \in Z} \theta_i)) \tag{6.13}$$

The formulas presented can be used to estimate radial measurements into their perpendicular equivalents, which is important for accurately measuring the distance to a target. The accuracy of the estimation depends on the FoV of the optical system, which is represented by θ.

1.5 Pile-up effect

A trade-off exists between the number of available SPADs for detection and the number of detectable photons. As illustrated on the left-hand side of **Fig. 6.5** for a low reflectance target located 10 meters away from the sensor, the restricted number of returning photons causes in distinct SPAD triggering. The estimated distance provided by the counting histogram is in close proximity to the actual distance. As depicted in the right-hand side of **Fig. 6.5** for a high reflectance target located 10 meters away from the sensor, the high number of coincident photons cause all SPADs to trigger at the front edge of the VCSEL pulse, then none or very few of them would be available for detecting photons that arrive later as they are already occupied. This causes a shift in the average or median position of detected pulses in the counting histogram. The occurrence of pile-up, where the detector cannot differentiate between multiple photons arriving within the same time interval, due to event congestion, can result in underestimation of distance range measurement. For example, a 66ps shift in the detected pulses is sufficient to cause a 1cm shift in the range estimate.

1.6 Wrap-around

Wrap-around is a phenomenon that can occur when the distance to a target exceeds the maximum measurable range. In this situation, the returning photons from the target arrive at the sensor after the next laser pulse has been emitted. As a result, the sensor detects the returning photons as if they had traveled a shorter distance than they actually did, causing an error in the range measurement. The reference SPAD array is utilized to define the time zero by detecting the time interval between each VCSEL pulse.

In the scenario depicted on the left-hand side of **Fig. 6.6**, the photon covers a distance of 200mm within a time span of 1332ps, which is significantly shorter than the VCSEL period

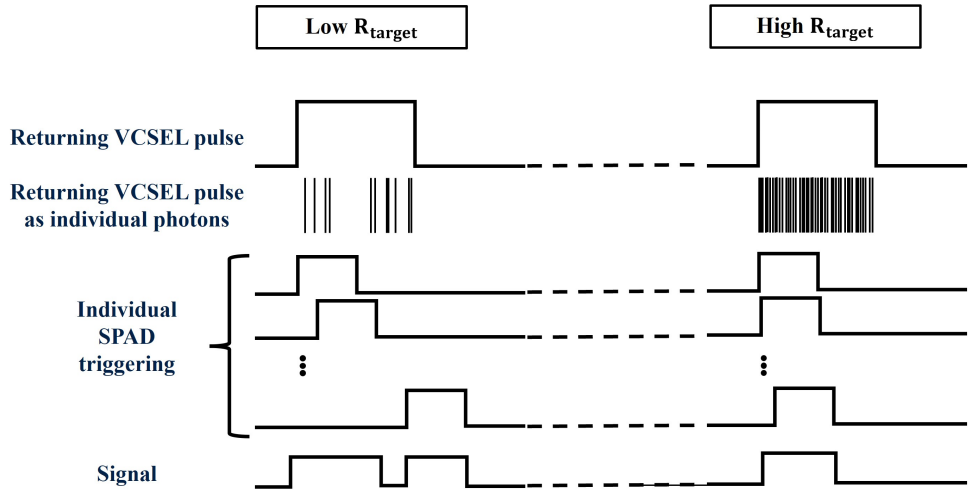


Figure 6.5: The pile-up phenomenon occurs when multiple photons arrive at the detector within the same time interval, making it impossible for the detector to distinguish between the individual photons.

pulse. In the scenario depicted on the right-hand side of **Fig. 6.6**, when the total distance traveled by the photon exceeds 3000mm, the time taken for the photon to travel this distance is greater than the VCSEL period pulse. This is longer than the VCSEL and system repetition rate. As a result, return photons may be incorrectly associated with subsequent VCSEL pulses, leading to a reported short target distance. Wrap around can be mitigated by using padding techniques which add extra values to the edges of signal.

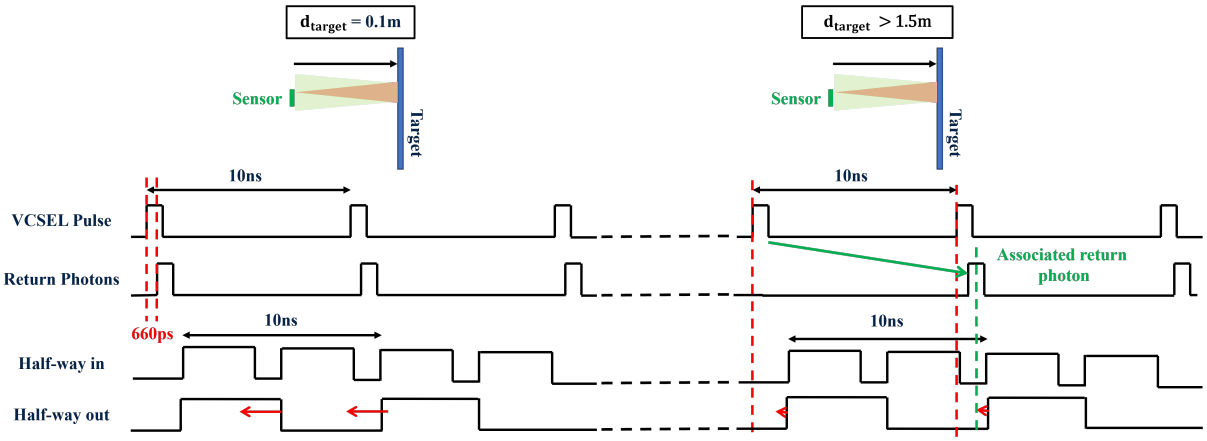


Figure 6.6: The wrap-around phenomenon happens when the returning photons from the target arrive at the sensor after the next laser pulse has been emitted, resulting in an error in the distance range measurement.

Conclusions

An optical system has been introduced, consisting of a VCSEL, reference and return SPAD arrays, and a cover glass to protect the system from dust. Under ideal conditions, a linear relationship between the measured distance $d_{measured}$ and the distance from the target d_{target} exists when considering only the photons emitted from the VCSEL and reflecting back to the sensor from the target. The light emitted from the VCSEL attenuates with the square of the distance to the target, and the shot noise of light due to the Poisson nature of light can cause statistical dispersion in the photon packet, which can be detrimental at low photon numbers. The cover glass introduces optical crosstalk when photons reflect back from the cover glass to the sensor, which limits the distance ranging at long distances. Subtraction of crosstalk is challenging due to shot noise and other factors such as scratches and smudges on the cover glass. To achieve optical crosstalk immunity, the signal is temporarily disabled (blanking) during the time when the emitted light pulse is expected to reflect off the covering glass and return to the sensor. The FoV determines the observable area that the optical system can capture at a given time. A wider range of observation can lead to an increase in optical crosstalk. Scale factors are used to estimate radial measurements into their perpendicular equivalents, which is crucial for accurately measuring the distance to a target. The pile-up phenomenon occurs when multiple photons arrive at the detector within the same time interval, disabling the detector from distinguishing between the individual photons. The pile-up phenomenon can limit the minimum measurable distance. Wrap-around occurs when the returning photons from the target arrive at the sensor after the next laser pulse has been emitted. Therefore, the pulse repetition period must be carefully designed to avoid this error, as a too-short period with respect to the returning photons from the previous pulse can result in measurement errors.

2 Dark Count Rate vs maximum distance ranging

This section aims to determine the maximum distance range d_{max} that the device should be able to measure in the prevailing environmental conditions. To identify a histogram bin as containing a signal, it is necessary for the signal to be distinguishable from the ambient light and DCR noise floor. The histogram processing requires the signal to be greater than the noise level E_{noise} plus a confidence factor multiplied by the standard deviation of the noise, as illustrated in **Fig. 6.7**. E_{noise} will refer to the sum of the ambient light noise and the DCR contribution. Additionally, there is shot noise on the signal, which necessitates a further 2 standard deviation margin for 94% accurate ranges. Therefore, the threshold for detecting a target is established based on the noise level E_{noise} and the shot noise of the signal.

In order to establish a distinguish between noise and signal levels, it is first necessary to determine the impact of reflectance and target distance on the scaling of distance for signal strength which can be summarized in a three-step process, which are illustrated in **Fig. 6.8**, as follows:

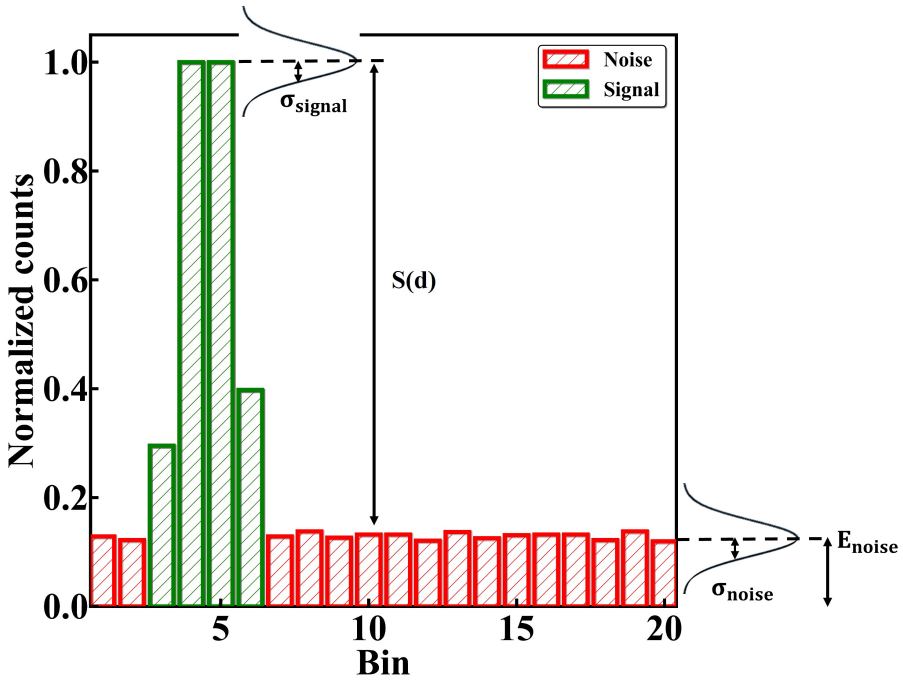


Figure 6.7: Typical simulated signal histogram, which includes noise, is depicted along with the Poisson distribution of light and their respective standard deviations overlaid on both the signal and ambient light.

1. Calibrate the signal per SPAD S_{cal} to account for a specific target reflectance and distance.
2. Determine the relationship between scaling and target reflectance R_{cal} to predict the signal strength for a target with varying reflectance at a constant distance.
3. Adjust the distance scaling d_{cal} based on the signal strength that decreases proportionally to the square of the distance.

A pulse segmenter, that separates the continuous stream of photon pulses into n individual segments based on timing criteria, attempts to detect the relevant pulse signal $S(d)$ at a distance d, as depicted in **Fig. 6.7**. This signal is proportional to the peak rate per SPAD S_{cal} obtained from the calibration measurement, which attenuate with the square of the distance:

$$S(d) = \alpha S_{cal} \cdot \frac{R}{R_{cal}} \cdot (\frac{d_{cal}}{d})^2 \tag{6.14}$$

The pulse segmenter locates the highest point of the average signal within a specific range of adjacent bins. Since this only represents a portion of the entire pulse signal, a coefficient α is applied to account for this partial measurement.

At the maximum distance achievable d_{max}, the pertinent pulse signal $S(d_{max})$ can be expressed as:

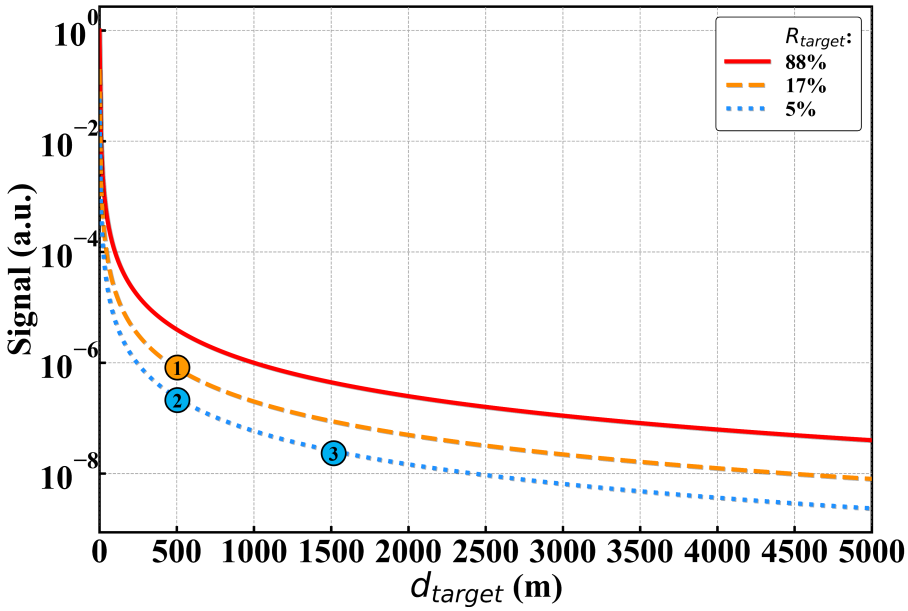


Figure 6.8: Signal as a function of the target distance d_{target} for different target reflectance R_{target}. The impact of reflectance and target distance on signal strength is calibrated for: (1) a specific target reflectance and distance, (2) different reflectances at constant distance, and (3) different distances at constant target reflectance.

$$S(d_{max}) = \alpha S_{cal} \cdot \frac{R}{R_{cal}} \cdot (\frac{d_{cal}}{d_{max}})^b \tag{6.15}$$

The minimum detection rate for distances below d_{max} is determined by S_{conf}, as illustrated in **Fig. 6.9**. This is expressed as:

$$\frac{1}{2} + \frac{1}{2} erf(\frac{S_{conf}}{\sqrt{2}}) \tag{6.16}$$

For instance, if S_{conf} is set to 2, the detection rate for targets below d_{max} will be 97.72%. Alternatively, if S_{conf} is set to 3, the detection rate will increase to 99.87% for targets below d_{max}.

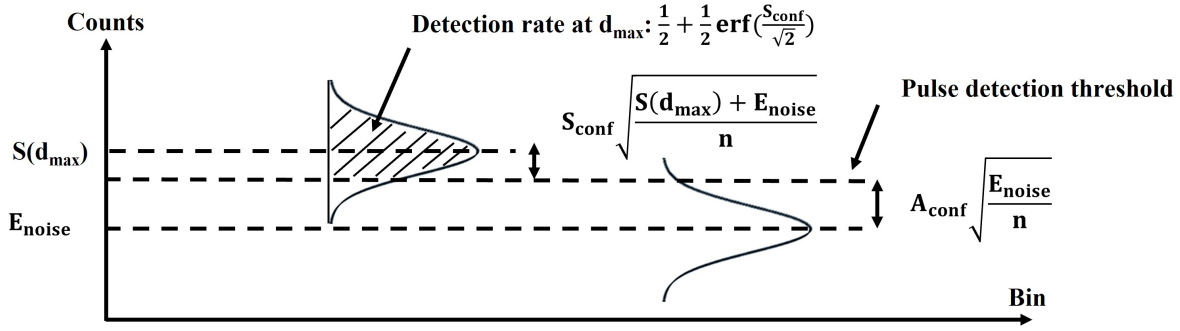


Figure 6.9: Worst-case scenario is when there is a significant overlap in the standard deviation of the signal and the ambient light, which results in a pulse detection threshold.

The threshold for signal detection in the presence of shot noise, as depicted in **Fig. 6.9**, can

be expressed as:

$$S_{conf}\sqrt{\frac{S(d_{max})+E_{noise}}{n}} \tag{6.17}$$

The threshold for signal detection in presence of the ambient light and DCR noises, as illustrated in **Fig. 6.9**, can be expressed as:

$$A_{conf}\sqrt{\frac{E_{noise}}{\sqrt{2}}} \tag{6.18}$$

Considering both shot noise from the signal as well as ambient light and DCR noises, the threshold for the pulse segmenter, as represented in **Fig. 6.9**, is given by:

$$S(d_{max})+E_{noise}-S_{conf}\sqrt{\frac{S(d_{max})+E_{noise}}{n}}=E_{noise}+A_{conf}\sqrt{\frac{E_{noise}}{n}} \tag{6.19}$$

The pulse segmenter threshold is established by A_{conf}, which determines whether the pulse segmenter identifies a target. The solution to this quadratic equation yields the maximum detectable distance:

$$d_{max}=d_{cal}\left(\frac{\alpha S_{cal}\cdot\frac{R}{R_{cal}}}{\frac{1}{4}\left(\frac{S_{conf}}{\sqrt{n}}+\sqrt{\left(\frac{S_{conf}}{\sqrt{n}}\right)^2+4\cdot\left(E_{noise}+\frac{A_{conf}}{\sqrt{n}}\sqrt{E_{noise}}\right)^2}-E_{noise}\right)}\right)^{\frac{1}{2}} \tag{6.20}$$

As a reminder, the calibration distance is denoted as d_{cal}, the peak rate per SPAD during calibration is denoted as S_{cal}, and the target reflectance obtained from calibration measurement is denoted as R_{cal}. A coefficient, denoted as α, is used to account for the entire pulse signal. The target reflectance under consideration is denoted as R, and the number of bins in the pulse segmenter is denoted as n. A confidence parameter, denoted as S_{conf}, is used to consider the desired percentage of detection rate. The noise floor, which accounts for both DCR and ambient light, is denoted as E_{noise}. Finally, a confidence parameter, denoted as A_{conf}, is used to set the pulse segmenter threshold.

The tuning parameter for the system is solely based on confidence (S_{conf} and A_{conf}). The shot noise causes variation in the number of ambient events per bin in each bin of the histogram. The pulse detection threshold for each bin is determined by multiplying the shot noise with a factor that depends on the ambient level in that bin.

Table 6.1 provides information on both outdoor light levels and material reflectances. The plots in **Fig. 6.10**, generated using the analytical formula **eq. 6.20**, illustrates that the influence of noise on the DCR becomes apparent in the presence of low ambient light, as demonstrated in **Fig. 6.10 (a)**. Furthermore, objects with low reflectivity are more vulnerable to DCR due to signal attenuation as illustrated in **Fig. 6.10 (b)**.

In **Chapter 5**, a modeling approach is proposed to estimate ΔDCR, which is reported in **Fig. 6.11** for both low-bias (V_c=3V) and high-bias (V_c=5V) characterization voltage regions.

Condition	Illumination (lux)	Material	Reflectance (%)
Sunlight	107527	Water	7 - 9
Full daylight	10752	Forest	7-14
Overcast day	1075	Concrete rough	20-30
Very dark day	107	Marble polished	30-70
Twilight	10.8	Plaster light	40-45
Deep twilight	1.08	Nickel highly polished	50-60
Full moon	0.108	Chrome polished	60-70
Quarter moon	0.0108	Copper highly polished	70-75
Starlight	0.0011	Aluminum pure highly polished	80-87
Overcast night	0.0001	Silver highly polished	90 - 92

Table 6.1: Outdoor light levels and material reflectances from [192].

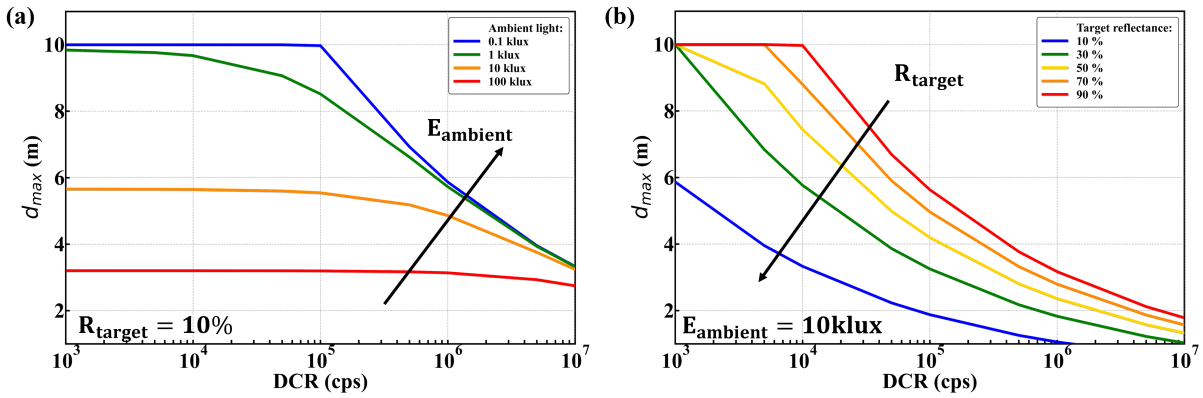


Figure 6.10: Maximum distance ranging d_{max} as a function of DCR for (a) different ambient light irradiance at a reflectance target R_{target} of 20% and (b) different reflectance target R_{target} at the ambient light irradiance $E_{ambient}$=0.1klux.

The worst-case scenario of low reflectance in a forest is considered. When the level of ΔDCR over stress duration is increasing under low irradiance levels of an overcast night, the maximum detectable distance d_{max} started to decrease, as shown in **Fig. 6.11 (a)**. Similarly, when the level of ΔDCR over stress duration is increasing under the high irradiance level of full daylight, d_{max} started to decrease, as illustrated in **Fig. 6.11 (b)**. By comparing both plots in **Fig. 6.11**, it can be observed that the maximum detectable distance d_{max} starts from a plateau that is limited either by ΔDCR or ambient light levels.

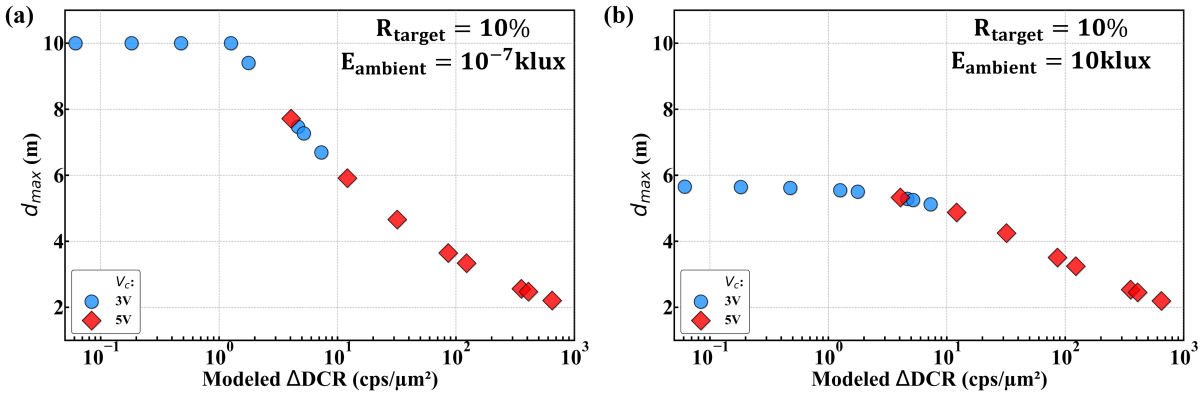


Figure 6.11: Maximum distance ranging d_{max} as a function of modeled ΔDCR for a reflectance target R_{target} of 10% with at the ambient light irradiance E_{noise} of (a) 10^{-7}klux and (b) 10klux.

Conclusions

To produce reliable distance ranging, the accumulated photon signal reflecting back from the target to the sensor must be dominant over other random noise sources. It is essential to distinguish a signal from the ambient noise and the DCR floor to identify a bin as containing a signal. Calibrating the model involves determining how reflectance and target distance impact the scaling of distance for signal strength. A pulse segmenter is used to detect the relevant pulse signal by separating returning photons from different laser pulses. An analytical formula can be derived to estimate the maximum distance ranging d_{max} using the maximum signal strength $S(d_{max})$, ambient noise level E_{noise}, and confidence levels S_{conf} and A_{conf}. The confidence level is the only tuning parameter for the system. Simulation results show that the contribution of noise to the DCR becomes significant in the presence of high ambient light, and low-reflective targets are particularly susceptible to DCR due to signal attenuation.

3 Conclusions

1

An optical system has been introduced, consisting of a VCSEL, reference and return SPAD arrays, and a cover glass to protect the system from dust. Under ideal conditions, a linear relationship between the measured distance $d_{measured}$ and the distance from the target d_{target} exists when considering only the photons emitted from the VCSEL and reflecting back to the sensor from the target. The light emitted from the VCSEL attenuates with the square of the distance to the target, and the shot noise of light due to the Poisson nature of light can cause statistical dispersion in the photon packet, which can be detrimental at low photon numbers. The cover glass introduces optical crosstalk when photons reflect back from the cover glass to the sensor, which limits the distance ranging at long distances. Subtraction of crosstalk is challenging due to shot noise and other factors such as scratches and smudges on the cover glass. To achieve optical crosstalk immunity, the signal is temporarily disabled (blanking) during the time when the emitted light pulse is expected to reflect off the covering glass and return to the sensor. The FoV determines the observable area that the optical system can capture at a given time. A wider range of observation can lead to an increase in optical crosstalk. Scale factors are used to estimate radial measurements into their perpendicular equivalents, which is crucial for accurately measuring the distance to a target. The pile-up phenomenon occurs when multiple photons arrive at the detector within the same time interval, disabling the detector from distinguishing between the individual photons. The pile-up phenomenon can limit the minimum measurable distance. Wrap-around occurs when the returning photons from the target arrive at the sensor after the next laser pulse has been emitted. Therefore, the pulse repetition period must be carefully designed to avoid this error, as a too-short period with respect to the returning photons from the previous pulse can result in measurement errors.

2

To produce reliable distance ranging, the accumulated photon signal reflecting back from the target to the sensor must be dominant over other random noise sources. It is essential to distinguish a signal from the ambient noise and the DCR floor to identify a bin as containing a signal. Calibrating the model involves determining how reflectance and target distance impact the scaling of distance for signal strength. A pulse segmenter is used to detect the relevant pulse signal by separating returning photons from different laser pulses. An analytical formula can be derived to estimate the maximum distance ranging d_{max} using the maximum signal strength $S(d_{max})$, ambient noise level E_{noise}, and confidence levels S_{conf} and A_{conf}. The confidence level is the only tuning parameter for the system. Simulation results show that the contribution of DCR becomes significant in the presence of low ambient light, and low-reflective targets are particularly susceptible to DCR due to signal attenuation.

Conclusion and perspectives

Conclusions

This thesis is centered around the development of models for simulating and predicting the performance parameters of Single-Photon Avalanche Diode (SPAD) for distance ranging applications. These models have been implemented in the Sentaurus tool [47], and they provide scientific contributions to the following areas:

A new modeling methodology for simulating and predicting the avalanche breakdown probability, avalanche breakdown voltage, Photo-Detection Efficiency and jitter

An electric-field based approach

The non-linear McIntyre model is utilized to predict the probability of avalanche breakdown in a SPAD device, considering impact-ionization coefficients and electric field streamlines. The P_t maps are influenced by temperature, voltage, carrier type, dopant profile, and device dimensions, and can be explored through different physical fields to define streamlines. The model's accuracy is confirmed by comparing the simulated P_t with experimental results using the PDE equation, covering the entire range of temperature and voltage. The breakdown voltage V_{BD} is extracted from both the PDE and current using the model and Sentaurus workbench, with good agreement. The analytical model in the sdevice module also reveals the activity of both regions, although with a lower magnitude when integrated across the entire device.

A drift-diffusion Monte-Carlo approach

The Monte-Carlo method is employed to capture individual carrier paths in the SPAD device, considering both drift and diffusion resulting from Brownian motion. To achieve the avalanche breakdown condition, the multiplication gain factor, which depends on the impact-ionization probability, must tend towards infinity. However, the deterministic McIntyre model computation along electric field streamlines does not capture the activation of avalanche-leading regions by diffusing carriers that overcome electrostatic potential barriers in the device. By analyzing each mesh point within the device, it is possible to differentiate between central and peripheral

avalanche contributions and study the statistical behavior of each carrier path. For instance, the shape of jitter can be explained by the time it takes to evacuate carriers from depleted to undepleted regions.

A new modeling methodology for simulating and predicting the pre-stress Dark Count Rate

A mean-based approach

To simulate the average behavior of DCR, temperature and electric field effects are considered in both low- and high-bias regions. In the low-bias region, BTBT has an impact at low temperatures, while thermal SRH generation from the neutral region becomes dominant at higher temperatures. In the high-bias region, SRH-TAT from the peripheral n-/p-well region is the main driving force. However, modeling the minority carrier lifetime through dopant distribution is insufficient for capturing the high-bias regime and required the inclusion of a trap profile within the device. Using DoE, the most crucial factors and their interactions that impact DCR can be identified. Architecture constraints and ion implantation conditions provide insight into the contribution of avalanche breakdown probability to DCR. Interface manufacturing highlights the role of interfaces in the origin of the low-bias regime.

A Monte-Carlo approach

The Monte-Carlo approach is effective in capturing both intrinsic and extrinsic populations of DCR. In the low-bias region, the statistical dispersion of the intrinsic population is considered by defect clusters randomly placed along the side interface, in the near-contact region, and within the multiplication region. In the high-bias region, this trend is only observed in defects randomly placed at the edge of the wells and rings, where the defect concentration is expected to be high due to multiple implant steps. In contrast, the extrinsic population is modeled by defects randomly placed within the wells and rings.

A new modeling methodology for simulating and predicting the post-stress Dark Count Rate

A mean-based approach

The SPAD reliability study aimed to identify potential degradation mechanisms and their locations under different stress conditions. Three distinct aging processes are suggested, one related to voltage, another to dark- and photo-generated current, and the last to temperature. Triggered avalanches are not considered as stress factors due to the lack of correlation observed between device-to-device avalanches at stress and different light aperture sizes. One of the identified degradation mechanisms is the accumulation at the hydrogenated amorphous silicon nitride film ($a - SiN_x : H$) during temperature stress in darkness without any applied voltage. ΔDCR and

current measurements showed a linear dependence on irradiance and an exponential dependence on the inverse of V_s. Hot-Carrier Degradation (HCD) is proposed as the degradation mechanism related to the growth of carrier multiplications, based on the number of carriers that increase linearly with irradiance and exponentially with the inverse of V_s. The same aging mechanism is observed for stresses at various irradiances and voltages at the same temperature, as indicated by the same time power. The manufacturing process is analyzed through a DoE, revealing the significance of the upper interface in the degradation process.

A Monte-Carlo approach

To capture the variability of ΔDCR during stress, measurements and Monte-Carlo simulations are conducted. The stress results showed a resilient quadratic relationship between the device-to-device variance $\sigma^2_{\Delta DCR}$ of ΔDCR and its mean $\overline{\Delta DCR}$, which is independent of architectures, stress conditions, and processes. To understand the origin of this quadratic law, potential defect localizations are identified based on temperature measurements taken before and after stress. The defects are then set at the depleted top interface to reproduce the extracted thermal dependency of DCR, accounting for the drift-diffusion carrier motion in avalanche breakdown probability simulation and a NRM-assisted carrier capture/emission rate for the carrier generation from defects. Temperature measurements and simulations before and after stress showed that the activated defects are located at the upper interface, regardless of the stress conditions. Using this simulation approach, coupled with defect position and number distribution, the device-to-device variability during stress is studied. The experimental quadratic relationship between the device-to-device variance $\sigma^2_{\Delta DCR}$ of ΔDCR and its mean $\overline{\Delta DCR}$ is generated by convolving a uniform distribution of random defect positions, each with an assigned DCR value, with an increasing number of defects Poisson-distributed in quantity.

A cutting-edge Hot-Carrier Degradation modeling methodology for simulating and predicting the Dark Count Rate drift along stress duration

A first-order rate equation in the concentration of Si-H bond precursors $N_{hb}(t)$ is utilized to model the degradation and determine the number of defects formed during stress. The bond dissociation rate constant κ considers the distribution of disorder-induced variations of Si-H bond-dissociation energies against the carrier energy distribution, incorporating the impact-ionization reaction cross section σ_{II} to account for current density and bond dissociation probability. By considering both the $Si-H$ bond energy E_{hb} and the carrier energy E, the definition of σ_{II} leads to a more significant depletion of $Si-H$ bond-dissociation energy states at low E_{hb}, resulting in higher defect creation. The developed model accurately describes hot-carrier degradation that induces ΔDCR in SPADs, covering a broad range of operating conditions. The primary contributors to total defect creation are identified as hot electrons, and the sub-linear time dependence of the defect creation process is attributed to the exhaustion of the bond disso-

ciation energy distribution due to impact-ionizing hot carriers. The two different slopes observed in hot-carrier degradation are explained by the degradation occurring in two different regions of the device architecture. The presented model accurately predicts the increase in dark current during stress by incorporating the dark current drift into the calculation of κ.

A new modeling methodology for simulating and predicting the distance range in direct Time of Flight measurement as a function of Dark Count Rate

Accurate distance ranging requires distinguishing the signal from ambient noise and the DCR floor, and ensuring the target's accumulated photon signal dominates over other random noise sources. This involves calibrating the model to determine how target distance and reflectance affect the scaling of distance for signal strength. The pulse segmenter separates returning photons from different laser pulses to detect the relevant pulse signal. An analytical formula can estimate the maximum distance ranging d_{max} using the maximum signal strength $S(d_{max})$, ambient noise level E_{noise}, and confidence levels S_{conf} and A_{conf}, with the confidence level being the only tuning parameter for the system. Simulation results show that DCR impact becomes significant in low ambient light, and low-reflective targets are more susceptible to DCR due to signal attenuation.

Perspectives

The research conducted has paved the way for several potential research directions, with the main ones being outlined in this section.

A transfer of the modeling methodology of post-stress Dark Count Rate for industrial electric-field based modeling

The modeling of Dark Count Rate (DCR) drift in the SPAD architecture under Hot-Carrier Degradation (HCD) requires the carrier energy distribution function. However, the traditional method of using full-band Monte-Carlo is time-consuming and demands significant computational resources. Instead, simplified approaches like Drift Diffusion (DD) and Energy Transport (ET) schemes [193][194] for solving the Boltzmann transport equation (BTE) can be used to substitute the Monte-Carlo method of carrier transport treatment. Previous studies [193] have shown that DD and ET methods are applicable for transport in MOSFETs with channel lengths not shorter than 0.1 µm. Thus, it is possible that these schemes can also be used for HCD modeling in long-channel devices. The Single Process (SP) mechanism dominates the degradation in the studied SPAD architecture, which is controlled by the high-energy tails of the carrier energy distribution function. Therefore, the choice of transport scheme used to evaluate the carrier energy distribution function can significantly impact the model results.

In the ET-based approach, the average energy from the Monte-Carlo solution is used to emulate the solution of an ET model. The carrier energy distribution function is then evaluated at each position at the interface using the average carrier energy $< E > (x)$ obtained from the ET scheme. On the other hand, in the DD-based model, only the lateral electric field profile calculated with the Monte-Carlo approach is retained. This field profile is then converted into the average carrier energy using the equation $< E >= \frac{3}{2}k_BT + q\tau_e\mu F^2$ where q is the electron charge modulus, E is the energy relaxation time, μ is the carrier mobility, and F is the electric field. The carrier energy distribution function can then be calculated using the equation $f(E) = Aexp(-\frac{E}{<E>})$.

Disregarding stress effect, the electric field can be used to extrapolate the avalanche breakdown probability from voltage measurement for characterization purposes. In addition, the generation rate can then be extracted from thermal measurement, assuming that thermal carrier generation dominates over tunneling, which is the case in the studied SPAD architecture.

In summary, the carrier energy distribution function can be determined once the electric field profile is extracted, allowing for the computation of the degradation rate based on the $Si - H$ bond dissociation energy distribution function and current density. Additionally, stress-induced defects can be characterized using the electric field-dependent avalanche breakdown probability and the thermal dependence of the carrier generation rate.

List of publications

M. Sicre, M. Agnew, C. Buj, J. Coignus, D. Golanski, R. Helleboid, B. Mamdy, I. Nicholson, S. Pellegrini, D. Rideau, D. Roy, and F. Calmon, "Dark Count Rate in Single-Photon Avalanche Diodes: Characterization and Modeling study," in *ESSCIRC 2021 - IEEE 47th European Solid State Circuits Conference (ESSCIRC)*, 2021, pp. 143–146. DOI: 10.1109/ESSCIRC53450.2021.9567806

R. Helleboid, D. Rideau, I. Nicholson, N. Moussy, O. Saxod, M. Basset, J. Grebot, A. Zimmerman, B. Mamdy, D. Golanski, M. Agnew, S. Pellegrini, and **M. Sicre**, "Comprehensive modeling and characterization of Photon Detection Efficiency and Jitter in advanced SPAD devices," in *ESSDERC 2021 - IEEE 51st European Solid-State Device Research Conference (ESSDERC)*, 2021, pp. 271–274. DOI: 10.1109/ESSDERC53440.2021.9631801

M. Sicre, M. Agnew, C. Buj, J. Coignus, D. Golanski, R. Helleboid, B. Mamdy, I. Nicholson, S. Pellegrini, D. Rideau, D. Roy, and F. Calmon, "Dark Count Rate in Single-Photon Avalanche Diodes: Characterization and Modeling study," in *ESSDERC 2021 - IEEE 51st European Solid-State Device Research Conference (ESSDERC)*, 2021, pp. 143–146. DOI: 10.1109/ESSDERC53440.2021.9631797

R. Helleboid, D. Rideau, J. Grebot, I. Nicholson, N. Moussy, O. Saxod, M. Basset, A. Zimmer, B. Mamdy, D. Golanski, M. Agnew, S. Pellegrini, **M. Sicre**, C. Buj, G. Marchand, J. Saint-Martin, M. Pala, and P. Dollfus, "Comprehensive Modeling and Characterization of Photon Detection Efficiency and Jitter Tail in Advanced SPAD Devices," *IEEE Journal of the Electron Devices Society*, vol. 10, pp. 584–592, 2022. DOI: 10.1109/JEDS.2022.3168365

M. Sicre, M. Agnew, C. Buj, C. Coutier, D. Golanski, R. Helleboid, B. Mamdy, I. Nicholson, S. Pellegrini, D. Rideau, D. Roy, and F. Calmon, "Statistical measurements and Monte-Carlo simulations of DCR in SPADs," in *ESSCIRC 2022- IEEE 48th European Solid State Circuits Conference (ESSCIRC)*, 2022, pp. 193–196. DOI: 10.1109/ESSCIRC55480.2022.9911519

M. Sicre, X. Federspiel, D. Roy, B. Mamby, C. Coutier, and F. Calmon, "Identification of stress factors and degradation mechanisms inducing DCR drift in SPADs," in *2022 IEEE International Integrated Reliability Workshop (IIRW)*, 2022, pp. 1–5. DOI: 10.1109/IIRW56459.2022.10032759

A. Jouni, V. Malherbe, B. Mamdy, T. Thery, **M. Sicre**, D. Soussan, V. Lorquet, S. D. Paoli, J.-M. Belloir, V. Lalucaa, C. Virmontois, G. Gasiot, and V. Goiffon, "Study of Proton-Induced Defects in 40-nm CMOS SPADs," *IEEE Transactions on Nuclear Science*, vol. 70, no. 8, pp. 1680–1686, 2023. DOI: 10.1109/TNS.2023.3257740

A. Jouni, **M. Sicre**, V. Malherbe, B. Mamdy, T. Thery, J.-M. Belloir, D. Soussan, S. De Paoli, V. Lorquet, V. Lalucaa, C. Virmontois, G. Gasiot, and V. Goiffon, "Proton-Induced Displacement Damages in 2-D and Stacked CMOS SPADs: Study of Dark Count Rate Degradation," *IEEE Transactions on Nuclear Science*, vol. 70, no. 4, pp. 515–522, 2023. DOI: 10.1109/TNS.2023.3248521

T. Garba-Seybou, X. Federspiel, F. Monsieur, **M. Sicre**, F. Cacho, J. Hai, and A. Bravaix, "Location of Oxide Breakdown Events under Off-state TDDB in 28nm N-MOSFETs dedicated to RF applications," in *2023 IEEE International Reliability Physics Symposium (IRPS)*, 2023, pp. 1–8. DOI: 10.1109/IRPS48203.2023.10117725

M. Sicre, X. Federspiel, B. Mamdy, D. Roy, and F. Calmon, "Characterization and modeling of DCR and DCR drift variability in SPADs," in *2023 IEEE International Reliability Physics Symposium (IRPS)*, 2023, pp. 1–5. DOI: 10.1109/IRPS48203.2023.10117921

M. Sicre, D. Roy, and F. Calmon, "Hot-Carrier Degradation modeling of DCR drift in SPADs," in *ESSDERC 2023 - IEEE 53rd European Solid-State Device Research Conference (ESSDERC)*, 2023, pp. 61–64. DOI: 10.1109/ESSDERC59256.2023.10268534

M. Sicre, X. Federspiel, B. Mamdy, D. Roy, and F. Calmon, "Modeling of Hot-Carrier Degradation driven by silicon-hydrogen bond dissociation in SPADs," *unpublished*, 2023

Appendix A

Variations of overall avalanche breakdown probability map

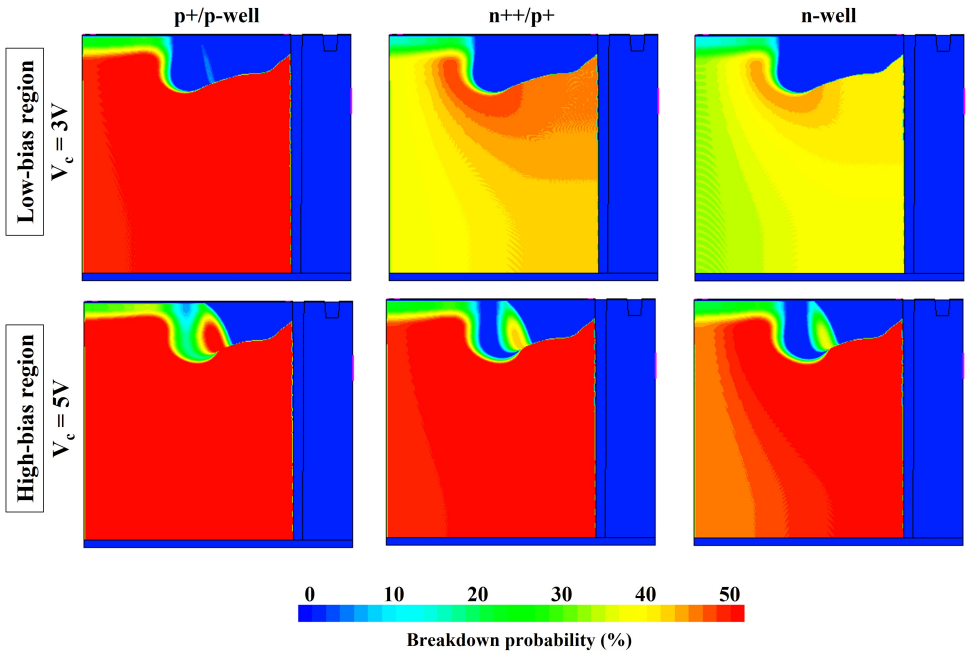


Figure A.1: Overall avalanche breakdown probability map simulated at T_c=333K and V_c=4V for different main junction architectures. The shape of the breakdown probability is altered by the extension and overlap of different regions, which affects the electric field distribution.

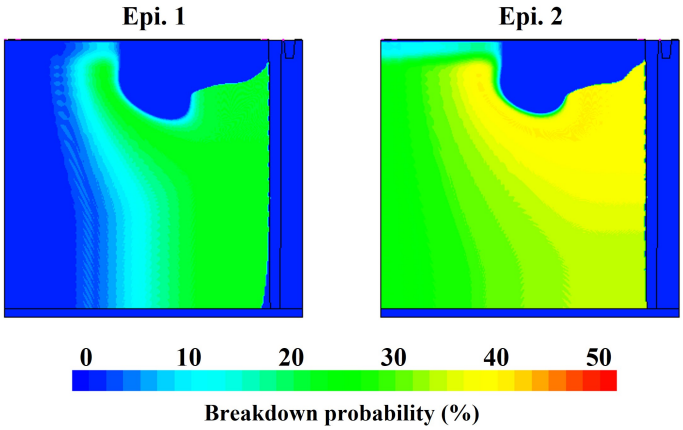


Figure A.2: Overall avalanche breakdown probability map simulated at T_c=333K and V_c=3V for different epitaxial growth. The silicon epitaxial growth has been processed with two different methods. The former by keeping constant the gas flow inside the main chamber (Epi. 1) and the latter by changing the flux at specific times during the manufacturing (Epi. 2). The electrostatic landscape is modulated by the dopant distribution within the volume, leading to significant changes in the probability of avalanche breakdown.

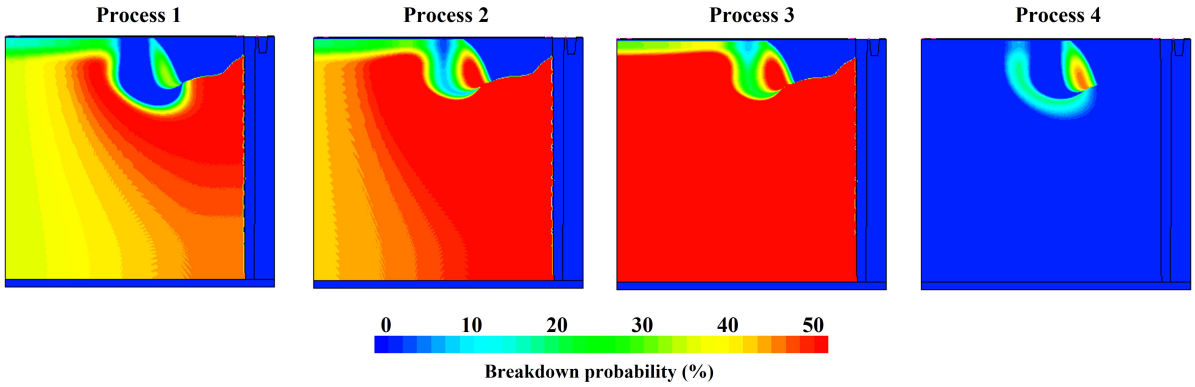


Figure A.3: Overall avalanche breakdown probability map simulated at T_c=333K and V_c=4V for different ion implantation schemes. The electric field profile of the multiplication region is altered by the dopant distribution in the main junction, resulting in changes to the probability of avalanche breakdown.

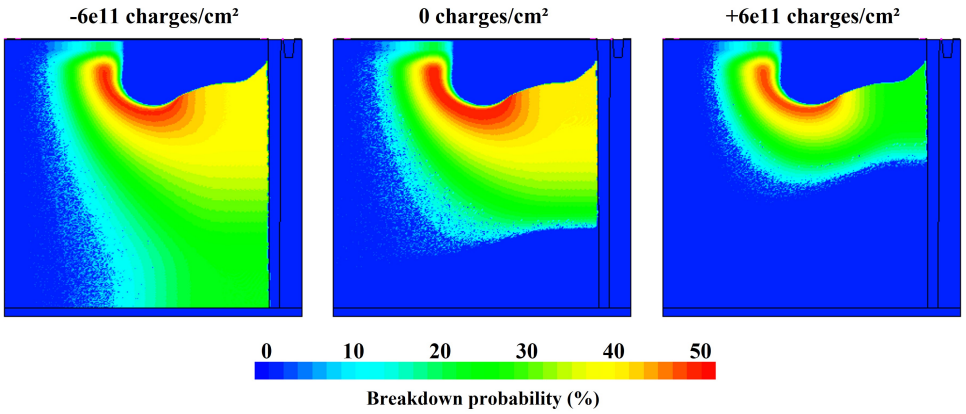


Figure A.4: Overall avalanche breakdown probability map simulated at T_c=333K and V_c=3V for different level of fixed charges in the bottom oxide. The magnitude of avalanche breakdown probability is influenced by the strength of the electric field, as it is calculated along the streamlines of the field.

Appendix B

Tunneling framework: Reflection and transmission at barriers

In classical physics, a particle with insufficient energy to overcome a potential barrier would be reflected back by the barrier. The classical reflection coefficient for a barrier of this nature is one. In quantum physics, a particle have a probability to tunnel through even though its energy is less than the height of the barrier. This is due to the wave-like nature of particles. The particle is described by a wave function $\psi(x)$ impinging the potential barrier which follows the time–independent Schrodinger equation [207]:

$$\frac{d^2\psi(x)}{dx^2} + \kappa(x)^2\psi(x) = 0 \tag{B.1}$$

where:

$$\kappa(x) = \begin{cases} (\frac{2m}{\hbar^2}(E - V(x))^{1/2}, & E > V(x) \\ i(\frac{2m}{\hbar^2}(V(x) - E))^{1/2}, & E < V(x) \\ \frac{2}{\lambda_t(x)}, & E = \frac{\hbar^2 k^2}{2m} \end{cases} \tag{B.2}$$

where $\lambda_t(x) = (2\kappa(x))^{-1} = \frac{\hbar}{2(2m(V(x)-E)^{1/2}}$ is the tunneling length which represents the distance over which the wave function of the particle decays exponentially as it tunnels through the barrier.

For an arbitrary potential under the Wentzel-Kramers-Brillouin (WKB) approximation of a slowly varying potential ($|k'| \ll k^2$), the solutions of **eq. B.1**:

$$\psi(x) = \begin{cases} exp(ikx) - Rexp(-ikx) = Aexp(ikx) + Bexp(-ikx), & x < 0 \\ Texp(ikx) = Fexp(ikx), & x > a \\ A'exp(\kappa x) - Bexp(-\kappa x) = \frac{C}{\sqrt{\kappa(x)}}exp(\int_0^x \kappa(t)dt) + \frac{D}{\sqrt{\kappa(x)}}exp(-\int_0^x \kappa(t)dt), & 0 < x < a \end{cases} \tag{B.3}$$

Assuming the continuity of $\psi(x)$ and $\frac{d\psi}{dx}$ at the two interfaces as boundary conditions, the

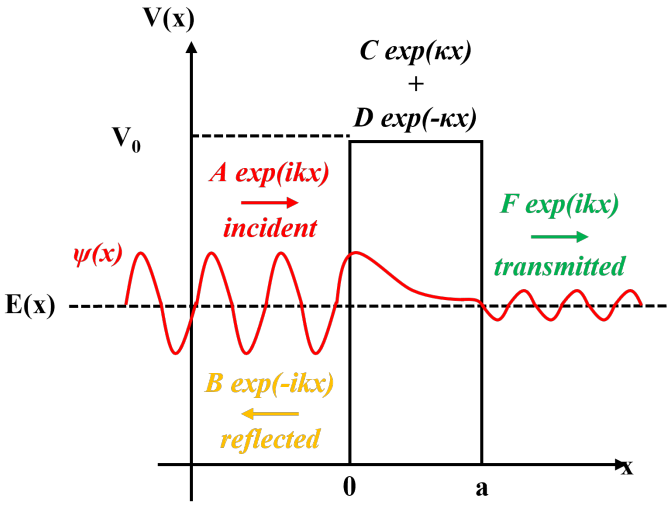


Figure B.1: Representation of a wave function $\psi(x)$ of a particle tunneling across a potential barrier. In the first region where x<0, an incident wave propagates through a potential-free zone and coexists with a reflected wave. in the second region, a portion of the incident wave packet that has not been reflected at x=0 propagates through a zone with a constant potential and tunnels through to the third region at x=L. In the third region where x>L, the transmitted wave that has tunneled through the potential barrier, moves in a potential-free zone. The energy of the incident particle $E(x)$ (dashed line) determines the behavior of the wave in each region.

transmission coefficient is derived as:

$$|T|^2 = \frac{(2k\kappa)^2}{(k^2+\kappa^2)^2 sinh^2(2\kappa a) + (2k\kappa)^2} \tag{B.4}$$

Assuming a wide and tall barrier $\kappa a \gg 1$

$$|T|^2 \approx (\frac{2k\kappa}{k^2+\kappa^2})^2 exp(-4\kappa a) \tag{B.5}$$

And for low tunneling probability $T \ll 1$

$$|T|^2 \approx (\frac{2k\kappa}{k^2+\kappa^2})^2 exp(-2\kappa a) \tag{B.6}$$

with $\kappa a = [\frac{2ma^2}{\hbar^2}(V_0 - E)]^{1/2}$. Considering the logarithm of the tunneling probability in this case:

$$log|T|^2 \approx -2\kappa(2a) + 2log\frac{2(ka)(\kappa a)}{(ka)^2(\kappa a)^2} \tag{B.7}$$

The first term dominated the second term which states that the transmission probability decays exponentially with barrier width (beyond the tunneling length) and exponentially with the square root of the energy. By assuming that the barrier as juxtaposition of N square barriers and their transmission coefficients are multiplicative (most reflected transmission are independent improbable event) this leads to:

$$
\begin{aligned}
log|T|^2 &\approx \sum_n log|T_n|^2 \\
&\approx -2\sum_n \Delta x < \kappa_n > \\
&\approx -2\int dx\sqrt{(\frac{2m}{\hbar^2})[V(x)-E]}
\end{aligned}
\tag{B.8}
$$

which results in:

$$
|T|^2 \approx exp(-2\int dx\sqrt{(\frac{2m}{\hbar^2})[V(x)-E]}) \tag{B.9}
$$

However the WKB approximation fails at turning point where V(x) = E, the particle would experience a sudden change in potential energy. One expands $V(x) = E + V'(0) + O(x^2)$ around x=0, assuming a linear potential and derives the Schrodinger equation to capture this discontinuity as follows:

$$
\frac{d^2\psi(x)}{dz^2} - z\psi(x) = 0 \tag{B.10}
$$

with $z = \alpha x$ and $\alpha = (\frac{2mV'(0)}{\hbar^2})^{1/3}$. The Airy functions are introduced:

$$
\psi(z) = aA_i(z) + bB_i(z) \tag{B.11}
$$

$$
A_i(z) \sim \begin{cases} \frac{1}{\sqrt{\pi}(-z)^{1/4}} sin[\frac{2}{3}(-z)^{3/2} + \frac{\pi}{4}], & z \ll 0 \\ \frac{1}{\sqrt{\pi}z^{1/4}} exp(-2z^{3/2}/3), & z \gg 0 \end{cases} \tag{B.12}
$$

$$
B_i(z) \sim \begin{cases} \frac{1}{\sqrt{\pi}(-z)^{1/4}} cos[\frac{2}{3}(-z)^{3/2} + \frac{\pi}{4}], & z \ll 0 \\ \frac{1}{\sqrt{\pi}z^{1/4}} exp(2z^{3/2}/3), & z \gg 0 \end{cases} \tag{B.13}
$$

The tunneling mechanisms must be considered when wave-like particles striking potential barrier especially in semiconductor devices where a potential exist in-between the conduction and valence bands as well as from a trap state to the bands.

Appendix C

Nonradiative Multiphonon-assisted trapping modeling

1 Normal coordinates in a two-level system

The transition probabilities $W_{c/e}$ are determined based on non-radiative multiphonon theory, which describes the lattice relaxation that occurs during the charge trapping/detrapping process. Nonradiative recombination occurs when an electron and a hole combine to release energy without the emission of a photon. This process typically involves the transfer of energy to lattice vibrations, resulting in the generation of multiple phonons. When a carrier (electron or hole) is captured by a defect level, it moves from the free state (conduction or valence band) to the defect level, resulting in a change in the electric charge of the defect. This change then causes the equilibrium coordinate of the system to shift from Q_a to Q_b, as illustrated in **Fig. C.1**. The process follows the full adiabatic one-electron Schrödinguer equation including the linear electron-lattice interaction. When considering a system that is limited to two states, namely states a and b that correspond to an unoccupied and occupied defect, respectively, the electron-lattice coupling term causes a displacement of the lattice oscillator, represented by $Q \rightarrow (Q - Q_a)$. This displacement shifts the total energy, denoted as E_{na}, which can be expressed in a single vibrational mode, as described in [60]:

$$[\frac{1}{2}(\frac{\partial Q}{\partial t})^2 + \frac{1}{2}\hbar\omega(Q - Q_b)^2]X_{nb}(Q) = (E_{nb} - W_b^0 + \frac{1}{2}\hbar\omega Q_b^2)X_{nb}(Q) \tag{C.1}$$

where $X_{na,nb}$ are the harmonic oscillator wavefunctions. The left-hand side of the equation contains the kinetic energy term, which is proportional to the square of the time derivative of the normal coordinate Q, and the potential energy term, which is proportional to the square of the displacement of the oscillator from its equilibrium position $Q_{a,b}$. The right-hand side of the equation contains the total energy of the system, which is the sum of the energy of the nth energy level $E_{na,nb}$, the electronic energy in state a and b in the absence of perturbation, $W_{a,b}^0$, and the potential energy due to the displacement of the oscillator from its equilibrium position $Q_{a,b}$. The energies of the ground and excited states, shown in **Fig. C.1**, can be calculated by neglecting the energy related to the lattice vibrations. The resulting equation for the energy of

state a as a function of the equilibrium coordinate Q is given by:

$$E_a(Q) \approx W_a^0 + \frac{1}{2}\hbar\omega(Q - Q_a)^2 \tag{C.2}$$

$$E_b(Q) \approx W_b^0 + \frac{1}{2}\hbar\omega(Q - Q_b)^2 \tag{C.3}$$

For the total energy one has to add the lattice vibration: $E_{na}(Q) = E_a(Q) + \frac{1}{2}\hbar\omega(n + \frac{1}{2})$. Therefore, normalizing these relation by $\hbar\omega$ one can defined the number of phonons involved in the transition as: $n_T = \frac{\Delta E}{\hbar\omega}$ where $\Delta E = E_a(Q) - E_b(Q)$. One notes that it corresponds to the change of energy in the normalized configuration coordinate associated with $\Delta Q = Q_b - Q_a$. The Huang-Rhys factor defined as $S = \frac{1}{2}(Q_b - Q_a)^2$ gives in the same way the number of phonons involved in the transition.

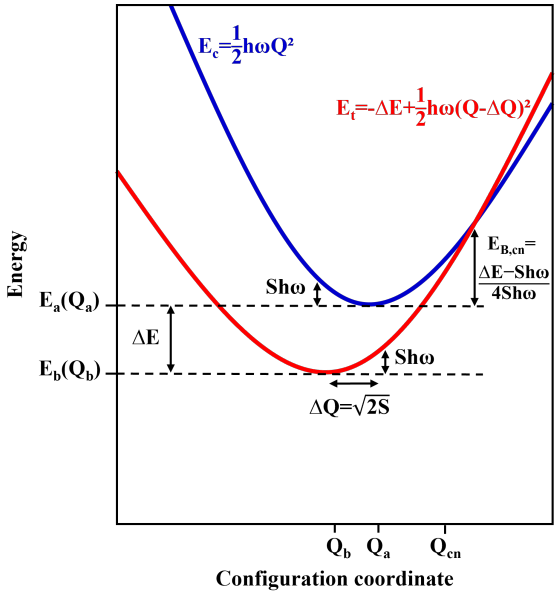


Figure C.1: Configuration coordinate diagrams for capture/emission process of electrons in conduction band. $\hbar\omega$ is associated to an effective distortion mode, ΔE the energy difference between the conduction band and the trapped state energy, and S the Huang Rhys factor.

Using **Fig. C.1**, one can find the point Q_{cn} where $E_a(Q)$ crosses $E_b(Q)$:

$$n_T + \frac{1}{2}Q_{cn}^2 = \frac{1}{2}(Q_{cn} + \sqrt{2S})^2 \tag{C.4}$$

$$2n_T = 2Q_{cn}\sqrt{2S} + 2S \tag{C.5}$$

$$Q_{cn} = \frac{n_T - S}{\sqrt{2S}} \tag{C.6}$$

This point is important for the transition a $\rightarrow$ b as it defines an activation energy $E_{B,cn}$ to overcome in order for the transition to occur. For the a $\rightarrow$ b transition, the barrier to overcome

is given by:

$$E_{B,cn} = \frac{1}{2}\hbar\omega Q_{cn}^2 = \frac{1}{2}\hbar\omega(\frac{n_T - S}{\sqrt{2S}})^2 = \frac{(\Delta E - S\hbar\omega)^2}{4\hbar\omega} \tag{C.7}$$

Indeed the process of transition from a to b can be divided into the two processes: first the electron-phonon-assisted electronic transition from the free electron to the capture state leading to a transition from Q_a to Q_{cn}; followed by a lattice relaxation through phonon-phonon interaction bringing Q_{cn} to Q_b. The factor S is very important in the calculation of the rate of transition and can be better understood by considering the difference in energy at Q_a and Q_b. The energy difference between these two vertical transition differs from a quantity $2S\hbar\omega$ that characterizes the lattice relatxation between the two states na and n'b.

2 Non-radiative transition formalism

The Non-Radiative Multiphonon-assisted (NRM) model is a theoretical framework used to transition probabilities $W_{c/e}$ by describing the interaction between electrons and lattice vibrations (phonons) as introduced in the previous section. The transition from a delocalized bulk state to a localized defect state can occur in the first order of electron-phonon coupling due to significant local lattice relaxation associated with the change in the defect's charge state. The Fermi golden rule is employed to derive the probability (W) per unit time for a free electron (or hole) in a delocalized state to be captured in a localized defect state. Under the Born-Oppenheimer and static approximation, using the second quantization framework, the transition rate is derived as [208][209]:

$$W_{ab}(E) = \frac{2\pi}{\hbar} A_{vn} \sum_{n'} |\sum_s A_s \left\langle X_{n'b}|a_s + a_s^\dagger|X_{na}\right\rangle|^2 \delta(E_{n'b} - E_{na} - E) \tag{C.8}$$

with $A_s = \sqrt{(\hbar/2M_s\omega_s)}\left\langle \Phi_b^0|\omega_s(r)|\Phi_a^0\right\rangle$ representing the wave-function overlap from the free state Φ_A to the bound state Φ_B through the perturbation ω_s calculated from the overlapping of the defect and trap wavefunctions and A_{vn} denoting the thermal averaging. δ(x) is the Dirac delta function. The energy dissipated through phonon mode s is $(n_s - n_s')\hbar\omega_s$ and the difference in energy between state a and b writes $E_{na} - E_{n'b} = E - E_t + \sum_s (n_s - n_s')\hbar\omega_s$. a_s and $a_s^\dagger$ are the creation and annihilation operators of phonons respectively.

Assuming that all phonon mode fequencies ω_s reduce to an average single frequency ω and using the properties of the Bessel functions one derived [209]:

$$W_c(E) = \frac{2\pi}{\hbar} R(\Delta E)[|V_\Lambda|^2(1 - \frac{\Delta E}{\hbar\omega S})^2 + |V|^2(1 - \frac{|V_\Lambda|^2}{S|V|^2})\sqrt{(\frac{\Delta E}{\hbar\omega S})^2 + 4\tilde{n}(\tilde{n}+1)}] \tag{C.9}$$

with the notations $|V_\Lambda|^2 = |\sum_s V_s \cdot \Lambda_s|^2$, $|V|^2 = |\sum_s V_s|^2$ and $S^2 = |\sum_s \Lambda_s|^2$ where the N-dimensional vector $\Lambda_s = \frac{1}{\sqrt{2}}(Q_2 - Q_1)_s$ is introduced representing the shift of the equilibrium positions of the lattice during the transition from state a to state b by mean of vibration modes.

The normal coordinates are defined for electron-phonon coupling as $Q_s = \sqrt{\frac{\hbar}{2M_s\omega_s}}(a_s + a_s^\dagger)$. Defining the ratio $r = |V_\Lambda|^2/(|V|^2 S)$ the transition rate writes:

$$W_c(E) = \frac{2\pi}{\hbar} R(\Delta E)|V|^2[rS(1 - \frac{\Delta E}{\hbar\omega S})^2 + (1-r)\sqrt{(\frac{\Delta E}{\hbar\omega S})^2 + 4\tilde{n}(\tilde{n}+1)}] \tag{C.10}$$

In which $\Delta E = E - E_t$ represents the difference between the energy of the free carrier and the trap energy below E_c, and $\tilde{n}$ is given by the Bose-Einstein distribution of phonons as $\tilde{n} = [exp(\frac{\hbar\omega}{k_B T}) - 1]^{-1}$. The term R(ΔE) is expressed as:

$$R(\Delta E) = \frac{1}{\hbar\omega} exp[-(2\tilde{n}+1)S + \frac{\Delta E}{2k_B T}] I_{m=\frac{E}{\hbar\omega}}(2S\sqrt{\tilde{n}(\tilde{n}+1)}) \tag{C.11}$$

being I_m the reduced Bessel function of order m. In order to evaluate the transition rate W_c one has to calculate the matrix elements S, $|V|^2$ and $|V_\Lambda|^2$:

$$S = \sum_s |\frac{1}{\hbar\omega\sqrt{2}} \left\langle \Phi_b^0|\omega_s|\Phi_b^0 \right\rangle - \left\langle \Phi_a^0|\omega_s|\Phi_a^0 \right\rangle|^2 \tag{C.12}$$

$$|V|^2 = \sum_s |\left\langle \Phi_b^0|\omega_s|\Phi_a^0 \right\rangle|^2 \tag{C.13}$$

$$|V \cdot \Delta|^2 = |\sum_s \left\langle \Phi_b^0|\omega_s|\Phi_a^0 \right\rangle [\left\langle \Phi_b^0|\omega_s|\Phi_b^0 \right\rangle - \left\langle \Phi_a^0|\omega_s|\Phi_a^0 \right\rangle]|^2 \tag{C.14}$$

In general there is a local change of the atomic positions near the defect between the ground state (defect occupied) and the first excited state (e.g. electron released in the conduction band). The wave function of a free electron in the conduction band of a bulk semiconductor is modified by the presence of this defect. A deep bound state is very localized relative to the free state (r → 0). It is assumed that beyond the short-rage attractive potential, there is a long-range Coulomb potential with charge Z. The free-state wave-function for r → 0 is defined as follows:

$$\Phi_a^0 = 1/V_c^{1/2} C_0(\varepsilon_)^{1/2} \tag{C.15}$$

For an attractive center:

$$C_0^{\mu=1}(\varepsilon_) = \frac{2\pi\eta}{1 - exp(-2\pi\eta)} \tag{C.16}$$

for a repulsive center:

$$C_0^{\mu=-1}(\varepsilon_) = \frac{2\pi\eta}{exp(2\pi\eta) - 1} \tag{C.17}$$

and for a neutral center:

$$C_0^{\mu=0}(\varepsilon_) = 1 \tag{C.18}$$

with $\eta = -(\mu/\nu)\sqrt{R_H/\varepsilon}$ which depend on the free-state kinetic energy $\varepsilon = \varepsilon\kappa = \hbar^2k^2/2m$ in the conduction band and $v = \sqrt{R_H/E_t}$ depends on the trap energy E_t with respect to the conduction band, and $R_H = (Z_e^2/4\pi\varepsilon_{Si})^2/2(\hbar^2/m)$ the energy level of a simple Hydrogen-like center of charge $\pm Z_e$ in a Silicon matrix of permitivity ε_{Si}. The coulombic factor $C_0(\epsilon)$ (also called Sommerfield factor) represents the enhancement/reduction of the plane wave at r = 0 caused by a purely Coulombic attractive/repulsive center. The current Sommerfield model gives an estimate of the charge state of the defect in relation to the transition rates.

A Dirac-like-defect trapping potential (i.e. $r \approx 0$ in the core potential of the square well potential) with coulombic tail is given by:

$$\Phi_b^0 = Aexp(-r/\nu a_H)(-r/\nu a_H)^{\mu-1} \tag{C.19}$$

where the Bohr radius $a_H = (\hbar^2/m_0)/(Z_e^2/4\pi\epsilon_S i) = 18.7$ Å in silicon and $A = (4\pi(a_H)^3(\nu/2)^{2\mu+1}$ $\Gamma(2\mu+1))^{-1/2}$ where $\mu = +\nu$ for a positively-charged centre, and $\mu = -\nu$ for a negatively-charged center.

Finally, the small displacements $w_{s,q}(r)$ can be formed based on optical deformation potential coupling like:

$$\omega_{s,q}(r) = D(\mathbf{q}) \cdot exp(i\mathbf{q} \cdot \mathbf{r}) \tag{C.20}$$

where $\mathbf{q}$ the phonon wave vector and $D(\mathbf{q}) = \frac{D^2}{N(M_r\omega/\hbar)}$.

S can be derived from the the quantum-defect wave-function as:

$$S = \int_0^{q_D} \frac{|\langle\Phi_b^0|\omega_s|\Phi_b^0\rangle|^2}{2(\hbar\omega)^2} \cdot \frac{4\pi q^2}{(2\pi)^3} dq \tag{C.21}$$

$$S = S_0 \cdot I(a, 2\mu, \alpha) \tag{C.22}$$

with $I(a,b,c) = \frac{1}{bc}\int_0^1 \frac{x^a sin^2(btan^{-1}(cx))}{[1+(cx)^2]^b}$, $S_0 = \frac{3}{2(\hbar\omega)^2} \cdot \frac{D^2}{M_r\omega/\hbar}$, $\alpha = q_D a_H \nu/2$ and $q_D = (6\pi^2 N/\nu)^2$ is the Debye cutoff wave-vector containing the number of modes N which approximates the Brillouin zone size by a sphere of radius q_D. The first term of $I(a,b,c)$ a is equal to 0 for an optical deformation. The shift in energy of the band edge per unit elastic strain is called the deformation potential. Optical strain exists in materials where each unit cell contains two atoms and is associated with the contrary displacement of the two atoms.

For the electron-phonon coupling defined previously, using the two previously defined states for the electron and for the trap a general expression has been found:

$$|V|^2 = [24\pi 4^\mu C_0(\epsilon)(\nu a_H)^3/V_c]\frac{\Gamma^2(\mu+2)}{\Gamma(2\mu+1)} \cdot \frac{D^2}{M_r\omega/\hbar} \cdot I(a, \mu+1, 2\alpha) \tag{C.23}$$

$$|V_\Lambda|^2 = [36\pi 4^\mu C_0(\epsilon)(\nu a_H)^3/V_c]\frac{\Gamma^2(\mu+2)}{\Gamma(2\mu+1)} \cdot \frac{D^2}{M_r\omega/\hbar} \cdot II^2(a, \mu+1, 2\alpha, 2\mu, \alpha) \tag{C.24}$$

with V_c the volume of the crystal and $II(a, b_1, c_1, b_2, c_2) = \frac{1}{b_1 c_1 b_2 c_2} \int_0^1 x^a \cdot \frac{sin^2(b_1 tan^{-1}(c_1 x))}{[1+(c_1 x)^2]^{b_1/2}} \cdot \frac{sin^2(b_2 tan^{-1}(c_2 x))}{[1+(c_2 x)^2]^{b_2/2}}$. The ratio r then writes:

$$r = \frac{|V_\Lambda|^2}{|V|^2 S} = \frac{II^2(a, \mu+1, 2\alpha, 2\mu, \alpha)}{I(a, \mu+1, 2\alpha) \cdot I(a, 2\mu, \alpha)} \tag{C.25}$$

This ratio being of the order of unity and providing that $(1 - E_t/S\hbar\omega)^2 \gg 1$, the transition rate is dominated by the term $|V_\Lambda|^2$. For instance, by setting $\mu \to 0$:

$$|V_\Lambda|^2 \approx 36\pi(S\hbar\omega)^2 \cdot \frac{C_0(\nu a_H)^3}{V_c} \tag{C.26}$$

and:

$$W_c(x, E, E_T) \approx \frac{2\pi}{\hbar} R(\Delta E)|V_\Lambda|^2 (1 - \frac{\Delta E}{\hbar\omega S})^2 \tag{C.27}$$

The capture/emission rates are derived then as:

$$c_n^{-1}(x, E_t) = \int_{E_c(x)}^{+\infty} N_c(x, E) W_c(x, E, E_T) f_c(x, E) dE \tag{C.28}$$

$$c_p^{-1}(x, E_t) = \int_{-\infty}^{E_v(x)} N_v(x, E) W_v(x, E, E_T)(1 - f_v(x, E)) dE \tag{C.29}$$

The probability to find a state at energy E is given by the Fermi statistics for electron and holes $f_c(x, E) = \left[1 + \exp\left(\frac{E-\mu_n(x)}{k_B T}\right)\right]^{-1}$ and $f_v(x, E) = \left[1 + \exp\left(\frac{E-\mu_p(x)}{k_B T}\right)\right]^{-1}$, respectively.

The impact of a high electric field can be resumed to an effective density of states accounting for tunneling as [210]:

$$N_{c,v}(E, x) = \frac{1}{2\pi} \left(\frac{2m_{e,v}}{\hbar^2}\right)^{3/2} \sqrt{\hbar\Theta_r} F_A(\frac{|E - E_{c,v}(x)|}{\hbar\Theta_r}) \tag{C.30}$$

where $\Theta_r = [q^2F^2/(2m\hbar)]^{1/3}$ is the electrooptical frequency and $F_A(x) = A_i'(x) - xA_i(x)$ with $A_i(x)$ the Airy function and $A_i'(x)$ its derivative.

Similar results are obtained when comparing between Non-Radiative Multiphonon-assisted (NRM) (**Fig. C.2**) and Shockley-Read-Hall (SRH) (**Fig. 2.31**) generation rate models. The concept of capture cross section is replaced by the Huang-Rhys factor which is the the number of phonons involved in the transition. The NRM can be used to overcome the Boltzmann thermal limit which is defined as $E_F = \frac{k_B T}{q} ln(\frac{N_A}{n_i})$ [211].

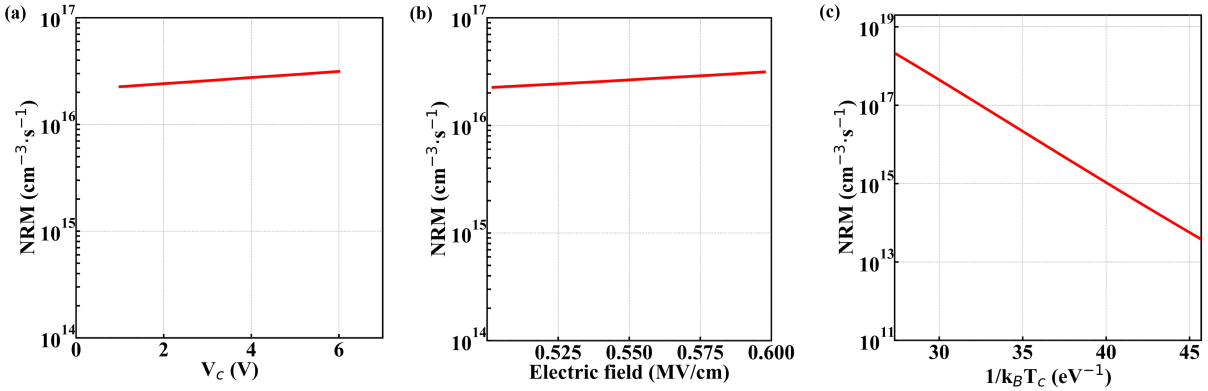


Figure C.2: Non-Radiative Multiphonon-assisted (NRM) emission/capture rate as a function (a) V_c, (b) the electric field and (c) $1/k_BT_c$.

Appendix D

Hot-Carrier Degradation time dependence

> This section reports the development of a solution for addressing the time-dependent characteristics of Hot-Carrier Degradation (HCD), presented in Amr Haggag, "IC-REST(2): A time-to-market driven IC reliability statistics tool," University of Illinois at Urbana-Champaign, 2002 [212].

The defect generation rate, as described by equation **eq. 5.15**, can be solved in the presence of a disorder-induced defect lifetime distribution denoted as $g(\tau_t) = \frac{df(\tau_{it})}{d\tau_{it}}$, where $f(\tau_{it}$ is a Fermi defect activation energy distribution and $\tau_t = 1/\kappa(E_a)$ with $\kappa(E_a)$ is the dissociation rate constant. The hot-carrier-activated defect density is given by:

$$N_t(t) = N_{hb}^0 - N_{hb}(t) \tag{D.1}$$

Including this definition in **eq. 5.15** yields to:

$$N_t(t) = N_{hb}^0 \int_0^\infty (1 - exp(-\frac{t}{\tau_t}))g(\tau_t)dE_a \tag{D.2}$$

The integral of disorder-induced defect lifetime distribution together with the Arrhenius process gives the amount of degradation $F(t)$ as follows:

$$F(t) =< \frac{N_t}{N_{hb}^0} >= \int_0^\infty (1 - exp(-\frac{t}{\tau_t}))g(\tau_{it})d\tau_{it} \tag{D.3}$$

The function $F(t)$ can be expressed as:

$$F(t) = \int_0^\infty g(\tau_{it})d\tau_{it} - \int_0^\infty exp(-\frac{t}{\tau_{it}})g(\tau_{it})d\tau_{it} \tag{D.4}$$

$$F(t) = 1 - \int_0^\infty exp(-\frac{t}{\tau_{it}})\frac{df(\tau_{it})}{d\tau_{it}}d\tau_{it} \tag{D.5}$$

Then integrating by parts:

$$F(t) = 1 - [f(\tau_{it})exp[-\frac{t}{\tau_{it}}]_0^\infty - \int_0^\infty f(\tau_{it})texp(-\frac{t}{\tau_{it}})\frac{d\tau_{it}}{\tau_{it}^2} \tag{D.6}$$

$$F(t) = 1 - [1 - \int_0^\infty f(\tau_{it}) \frac{t}{\tau_{it}^0} exp(-\frac{t}{\tau_{it}}) \frac{\tau_{0,it}}{\tau_{it}^2} d\tau_{it}] \tag{D.7}$$

Then letting $\overline{t} = \frac{t}{\tau_{0,it}}$ and $\overline{s} = \frac{\tau_{0,it}}{\tau_{it}}$ (i.e. $d\overline{s} = -\frac{\tau_{0,it}}{\tau_{it}^2} d\tau_{it}$) one gets:

$$F(t) = \overline{t} \int_0^\infty f(\frac{\tau_{0,it}}{\overline{s}}) exp(-\overline{s}\overline{t}) d\overline{s} \tag{D.8}$$

from which one can obtain the following expression for lifetime distribution f (**eq. 5.14**) as a function of $\overline{s}$:

$$\frac{g(\tau_{0,it}\overline{t})}{\overline{t}} = \int_0^\infty f(\frac{\tau_{0,it}}{\overline{s}}) exp(-\overline{s}\overline{t}) d\overline{s} \tag{D.9}$$

So $g(\tau_{0,it})$ divided by $\overline{t}$ is the Laplace transform of $f(\frac{\tau_{0,it}}{\overline{s}})$ and can be approximated using asymptotic methods (such as the Laplace Method):

$$g(\tau_{0,it}\overline{t}) = \int_C G(\overline{s}) exp(h(\overline{t} \cdot \overline{s})) d\overline{s} \approx 2\pi \frac{G(\overline{s_0}) exp(h(\overline{t} \cdot \overline{s_0}))}{\sqrt{2\pi h''(\overline{t} \cdot \overline{s_0})}} \tag{D.10}$$

where $\overline{s_0}$ lies on the path C and its the saddle point of $h(\overline{t} \cdot \overline{s})$. Applying the approximation with $G(\overline{s}) = \frac{g(\frac{\tau_{it}}{\overline{s}})}{\overline{s}}$ and $h(\overline{t} \cdot \overline{s}) = -\overline{s}\overline{t} - ln(\overline{s})$ yields:

$$F(t) =< \frac{N_t}{D_{hb}^0} >\approx \frac{1}{1 + (\frac{t}{\tau_{0,it}})^{\frac{-k_B T}{\sigma_{it}}}} \approx (\frac{t}{\tau_{0,it}})^{\frac{k_B T}{\sigma_{it}}} \tag{D.11}$$

Appendix E

Optical crosstalk impact on distance measurement

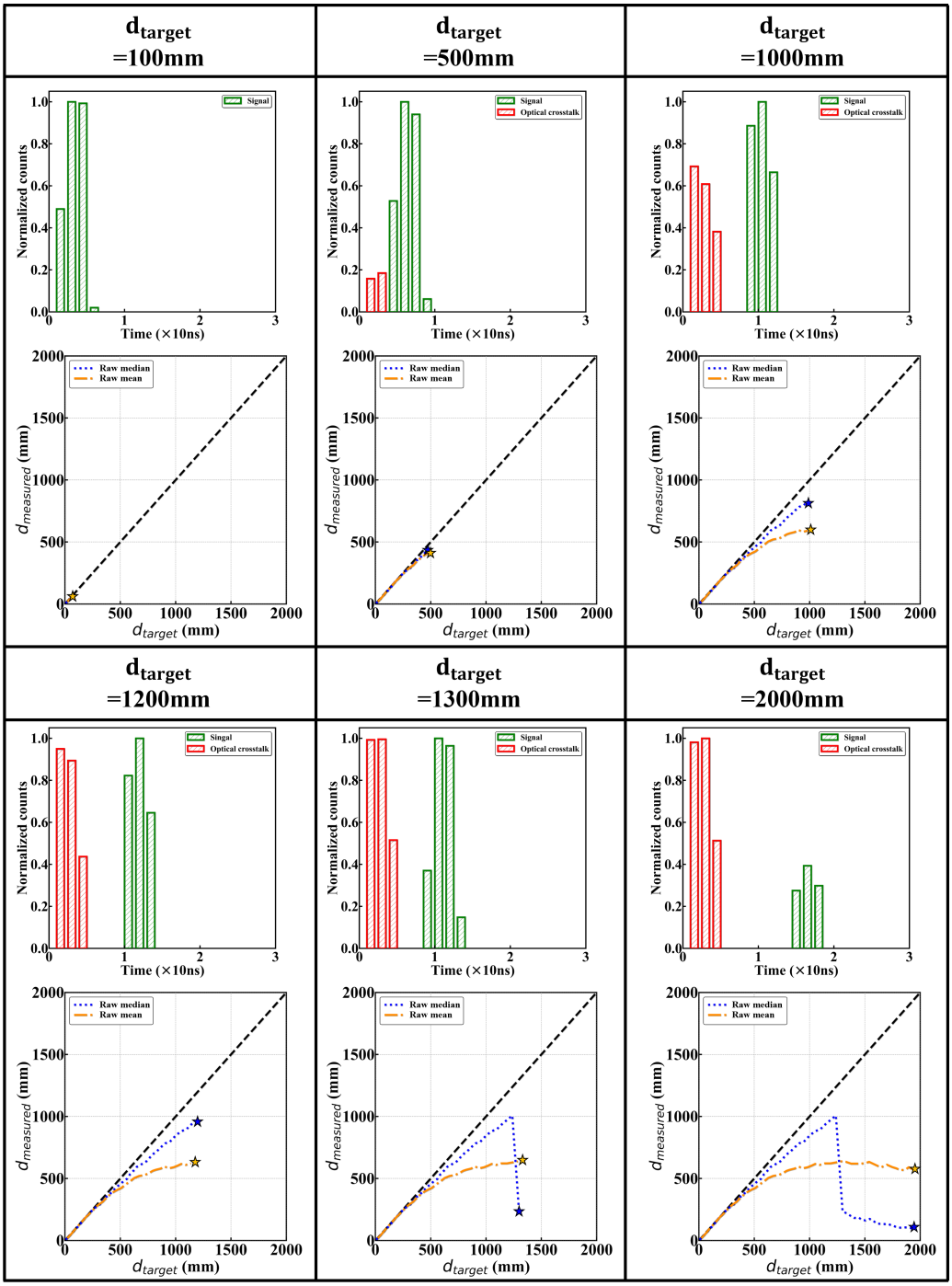


Figure E.1: Simulated normalized count as the function of the photon arrival times as well as maximum distance measurable as function of the distance to the target d_{target} in each box for an increasing d_{target}. The green bins in the histogram represent the true signal, while the red bins represent the optical crosstalk. The correlation plot includes a dashed dark line representing the x-y line, a blue dotted line representing the raw mean, and an orange dashed-dotted line representing the raw median. As the distance to the target increases, the photon reflecting back from the cover glass moves faster than those from the target, resulting in a higher contribution of optical crosstalk relative to the signal. This shift causes the raw median to be lower than the raw mean.

Bibliography

[1] J. C. Maxwell, *A Treatise on Electricity and Magnetism* (Cambridge Library Collection - Physical Sciences). Cambridge University Press, 2010, vol. 1. DOI: 10.1017/CBO9780511709333.

[2] H. Hertz, "Über einen Einfluß des ultravioletten Lichtes auf die electrische Entladung," *Annalen der Physik*, vol. 267, pp. 983–1000, Mar. 2006. DOI: 10.1002/andp.18872670827.

[3] W. Hallwachs, "Ueber den Einfluss des Lichtes auf electrostatisch geladene Körper," *Annalen der Physik*, vol. 269, no. 2, pp. 301–312, 1888. DOI: https://doi.org/10.1002/andp.18882690206.

[4] J. J. T. M. F.R.S., "LVIII. On the masses of the ions in gases at low pressures," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 48, no. 295, pp. 547–567, 1899. DOI: 10.1080/14786449908621447.

[5] P. Lenard, "Erzeugung von Kathodenstrahlen durch ultraviolettes Licht," *Annalen der Physik*, vol. 307, no. 6, pp. 359–375, 1900. DOI: https://doi.org/10.1002/andp.19003070611.

[6] R. A. Millikan, "A Direct Photoelectric Determination of Planck's "h"," *Phys. Rev.*, vol. 7, pp. 355–388, 3 Mar. 1916. DOI: 10.1103/PhysRev.7.355.

[7] J. Elster and H. Geitel, "Verfahren zur Messung von Licht-stärken unter Verwendung einer lichtelectrischen Vacuumzelle - Patentschrift 66969 vom 5. August 1892," *Kaiserl. PatentamtBerlin, and Zeitschr. Instrumentenkund*, vol. 13, pp. 474–475, 1893.

[8] A. Einstein, "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt," *Annalen der Physik*, vol. 322, no. 6, pp. 132–148, 1905. DOI: https://doi.org/10.1002/andp.19053220607.

[9] K. L., "Author's certificate 24040," *priority 4 August 1930*,

[10] V. Zworykin and J. Rajchman, "The Electrostatic Electron Multiplier," *Proceedings of the IRE*, vol. 27, no. 9, pp. 558–566, 1939. DOI: 10.1109/JRPROC.1939.228753.

[11] L. Austin and H. Starke, "Ueber die Reflexion der Kathodenstrahlen und eine damit verbundene neue Erscheinung secundärer Emission," *Annalen der Physik*, vol. 314, no. 10, pp. 271–292, 1902. DOI: https://doi.org/10.1002/andp.19023141003.

[12] E. A. G. Webster, "Single-Photon avalanche diode, theory, simulation, and high performance CMOS integration.," Thèse de doct. Edinburgh University, 2013.

[13] S. Cova, A. Longoni, and A. Andreoni, "Towards picosecond resolution with single-photon avalanche diodes," *Review of Scientific Instruments*, vol. 52, pp. 408–412, Apr. 1981. DOI: 10.1063/1.1136594.

[14] R. McIntyre, "Recent developments in silicon avalanche photodiodes," *Measurement*, vol. 3, no. 4, pp. 146–152, 1985. DOI: https://doi.org/10.1016/0263-2241(85)90024-7.

[15] W. Kindt, "Geiger-mode avalanche photodiode arrays," Thèse de doct. TU Delft. 1999.

[16] A. Rochas, "Single Photon Avalanche Diodes in CMOS Technology," Thèse de doct. Lausanne EPFL : École Polytechnique Fédérale de Lausanne. 2003.

[17] M. A. Karami, M. Gersbach, H.-J. Yoon, and E. Charbon, "A new single-photon avalanche diode in 90nm standard CMOS technology," *Opt. Express*, vol. 18, no. 21, pp. 22 158–22 166, Oct. 2010. DOI: 10.1364/OE.18.022158.

[18] C. Niclass, M. Gersbach, R. Henderson, L. Grant, and E. Charbon, "A Single Photon Avalanche Diode Implemented in 130-nm CMOS Technology," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 13, pp. 863–869, Aug. 2007. DOI: 10.1109/JSTQE.2007.903854.

[19] M. Gersbach, C. Niclass, E. Charbon, J. Richardson, R. Henderson, and L. Grant, "A single photon detector implemented in a 130nm CMOS imaging process," in *ESSDERC 2008 - 38th European Solid-State Device Research Conference*, 2008, pp. 270–273. DOI: 10.1109/ESSDERC.2008.4681750.

[20] A. Gulinatti, I. Rech, F. Panzeri, C. Cammi, P. Maccagnani, M. Ghioni, and S. Cova, "New silicon SPAD technology for enhanced red-sensitivity, high-resolution timing and system integration," *Journal of Modern Optics*, vol. 59, no. 17, pp. 1489–1499, 2012. DOI: 10.1080/09500340.2012.701340.

[21] A. Gulinatti, I. Rech, and M. Ghioni, "SPAD: A Solid-State Device for Detecting Single-Photons," in *Proceedings of the 2014 Cornell NanoScale Science & Technology Facility Annual Meeting*, 2014, pp. 8–8.

[22] J. A. Richardson, 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. DOI: 10.1109/LPT.2009.2022059.

[23] E. A. G. Webster, L. A. Grant, and R. K. Henderson, "A High-Performance Single-Photon Avalanche Diode in 130-nm CMOS Imaging Technology," *IEEE Electron Device Letters*, vol. 33, no. 11, pp. 1589–1591, 2012. DOI: 10.1109/LED.2012.2214760.

[24] R. Henderson, E. Webster, R. Walker, J. Richardson, and L. Grant, "A 3x3, 5 mu m pitch, 3-transistor Single Photon Avalanche Diode Array with Integrated 11V Bias Generation in 90nm CMOS Technology," Jan. 2011, pp. 14.2.1–14.2.4. DOI: 10.1109/IEDM.2010.5703359.

[25] L. Pancheri and D. Stoppa, "Low-Noise CMOS single-photon avalanche diodes with 32 ns dead time," in *ESSDERC 2007 - 37th European Solid State Device Research Conference*, 2007, pp. 362–365. DOI: 10.1109/ESSDERC.2007.4430953.

[26] L. Pancheri and D. Stoppa, "Low-noise single Photon Avalanche Diodes in 0.15 µm CMOS technology," in *2011 Proceedings of the European Solid-State Device Research Conference (ESSDERC)*, 2011, pp. 179–182. DOI: 10.1109/ESSDERC.2011.6044205.

[27] S. Pellegrini, B. Rae, A. Pingault, D. Golanski, S. Jouan, C. Lapeyre, and B. Mamdy, "Industrialised SPAD in 40 nm technology," in *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017, pp. 16.5.1–16.5.4. DOI: 10.1109/IEDM.2017.8268404.

[28] T. C. de Albuquerque, F. Calmon, R. Clerc, P. Pittet, Y. Benhammou, D. Golanski, S. Jouan, D. Rideau, and A. Cathelin, "Integration of SPAD in 28nm FDSOI CMOS technology," in *2018 48th European Solid-State Device Research Conference (ESSDERC)*, 2018, pp. 82–85. DOI: 10.1109/ESSDERC.2018.8486852.

[29] S. Shimada, Y. Otake, S. Yoshida, S. Endo, R. Nakamura, H. Tsugawa, T. Ogita, T. Ogasahara, K. Yokochi, Y. Inoue, K. Takabayashi, H. Maeda, K. Yamamoto, M. Ono, S. Matsumoto, H. Hiyama, and T. Wakano, "A Back Illuminated 6 µm SPAD Pixel Array with High PDE and Timing Jitter Performance," in *2021 IEEE International Electron Devices Meeting (IEDM)*, 2021, pp. 20.1.1–20.1.4. DOI: 10.1109/IEDM19574.2021.9720639.

[30] S. Pellegrini and B. Rae, "A low power, high dynamic range SPAD for sustainable sensing," in *International SPAD Sensor Workshop (ISSW)*, 2022.

[31] K. Morimoto, J. Iwata, M. Shinohara, H. Sekine, A. Abdelghafar, H. Tsuchiya, Y. Kuroda, K. Tojima, W. Endo, Y. Maehashi, Y. Ota, T. Sasago, S. Maekawa, S. Hikosaka, T. Kanou, A. Kato, T. Tezuka, S. Yoshizaki, T. Ogawa, K. Uehira, A. Ehara, F. Inui, Y. Matsuno, K. Sakurai, and T. Ichikawa, "3.2 Megapixel 3D-Stacked Charge Focusing SPAD for Low-Light Imaging and Depth Sensing," in *2021 IEEE International Electron Devices Meeting (IEDM)*, 2021, pp. 20.2.1–20.2.4. DOI: 10.1109/IEDM19574.2021.9720605.

[32] S. Shimada, Y. Otake, S. Yoshida, Y. Jibiki, M. Fujii, S. Endo, R. Nakamura, H. Tsugawa, Y. Fujisaki, K. Yokochi, J. Iwase, K. Takabayashi, H. Maeda, K. Sugihara, K. Yamamoto, M. Ono, K. Ishibashi, S. Matsumoto, H. Hiyama, and T. Wakano, "A SPAD Depth Sensor Robust Against Ambient Light: The Importance of Pixel Scaling and Demonstration of a 2.5µm Pixel with 21.8% PDE at 940nm," in *2022 International Electron Devices Meeting (IEDM)*, 2022, pp. 37.3.1–37.3.4. DOI: 10.1109/IEDM45625.2022.10019414.

[33] M.-J. Lee, A. R. Ximenes, P. Padmanabhan, T. J. Wang, K. C. Huang, Y. Yamashita, D. N. Yaung, and E. Charbon, "A back-illuminated 3D-stacked single-photon avalanche diode in 45nm CMOS technology," in *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017, pp. 16.6.1–16.6.4. DOI: 10.1109/IEDM.2017.8268405.

[34] M.-J. Lee and E. Charbon, "Progress in single-photon avalanche diode image sensors in standard CMOS: From two-dimensional monolithic to three-dimensional-stacked technology," *Japanese Journal of Applied Physics*, vol. 57, 1002A3, Oct. 2018. DOI: 10.7567/JJAP.57.1002A3.

[35] E. Park, W.-Y. Ha, H.-S. Park, D. Eom, H.-S. Choi, D.-H. Ahn, W.-Y. Choi, and M.-J. Lee, "A Back-Illuminated SPAD Fabricated With 40 nm CMOS Image Sensor Technology Achieving Near 40% PDP at 940 nm," *IEEE Journal of Selected Topics in Quantum Electronics*, pp. 1–8, 2023. DOI: 10.1109/JSTQE.2023.3272777.

[36] S.-G. Wuu, H.-L. Chen, H.-C. Chien, P. Enquist, R. M. Guidash, and J. McCarten, "A Review of 3-Dimensional Wafer Level Stacked Backside Illuminated CMOS Image Sensor Process Technologies," *IEEE Transactions on Electron Devices*, vol. 69, no. 6, pp. 2766–2778, 2022. DOI: 10.1109/TED.2022.3152977.

[37] P. Sun, B. Mimoun, E. Charbon, and R. Ishihara, "A flexible ultra-thin-body SOI single-photon avalanche diode," in *2013 IEEE International Electron Devices Meeting*, 2013, pp. 11.1.1–11.1.4. DOI: 10.1109/IEDM.2013.6724606.

[38] M. Hofbauer, B. Steindl, and H. Zimmermann, "Temperature Dependence of Dark Count Rate and After Pulsing of a Single-Photon Avalanche Diode with an Integrated Active Quenching Circuit in 0.35 μm CMOS," *Journal of Sensors*, vol. 2018, pp. 1–7, Jul. 2018. DOI: 10.1155/2018/9585931.

[39] N. Moussy and J. Ouvrier-Buffet, *SPAD-TYPE PHOTODIODE*, U.S. Patent 9,780,247, B2, issued October 3, 2017.

[40] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, "Avalanche photodiodes and quenching circuits for single-photon detection," *Appl. Opt.*, vol. 35, no. 12, pp. 1956–1976, Apr. 1996. DOI: 10.1364/AO.35.001956.

[41] S. Tisa, F. Guerrieri, and F. Zappa, "Variable-load quenching circuit for single-photon avalanche diodes.," *Optics express*, vol. 16 3, pp. 2232–44, 2008.

[42] P. Mars, "Temperature dependence of avalanche breakdown voltage Temperature dependence of avalanche breakdown voltage in p—n junctions†," *International Journal of Electronics*, vol. 32, no. 1, pp. 23–37, 1972. DOI: 10.1080/00207217208938266.

[43] 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. Opt. Express 16, 8381-8394," *Optics express*, vol. 16, pp. 8381–94, Jul. 2008. DOI: 10.1364/OE.16.008381.

[44] *3D Imaging and Sensing 2023: Market and Technology Report*. l' Yole Intellegicence 2023, YINTR23329.

[45] "Sentaurus Process User Guide," *Version P-2019.03*, Mar. 2019.

[46] "Calibre DESIGNrev Reference Manual," *Version 2019.4*,

[47] "Sentaurus Device User Guide," *Version P-2019.03*, Mar. 2019.

[48] "Sentaurus Visual User Guide," *Version P-2019.03*, Mar. 2019.

[49] P. A. Wolff, "Theory of Electron Multiplication in Silicon and Germanium," *Phys. Rev.*, vol. 95, pp. 1415–1420, 6 Sep. 1954. DOI: 10.1103/PhysRev.95.1415.

[50] M. Levinshtein, J. Kostamovaara, and S. Vainshtein, "BREAKDOWN PHENOMENA IN SEMICONDUCTORS AND SEMICONDUCTOR DEVICES," *International Journal of High Speed Electronics and Systems*, vol. 14, pp. 921–1128, Dec. 2004. DOI: 10.1142/S012915640400306X.

[51] B. V. Zeghbroeck, *Principles of Semiconductor Devices*. 2004.

[52] C. L. Anderson and C. R. Crowell, "Threshold Energies for Electron-Hole Pair Production by Impact Ionization in Semiconductors," *Phys. Rev. B*, vol. 5, pp. 2267–2272, 6 Mar. 1972. DOI: 10.1103/PhysRevB.5.2267.

[53] S. Ahmad and W. Khokley, "Transition probability of impact ionization in silicon," *Journal of Physics and Chemistry of Solids*, vol. 28, no. 12, pp. 2499–2507, 1967. DOI: https://doi.org/10.1016/0022-3697(67)90037-6.

[54] L. V. Keldysh, "Concerning the Theory of Impact Ionization in Semiconductors," 1965.

[55] N. Sano and A. Yoshii, "Impact-ionization theory consistent with a realistic band structure of silicon," *Phys. Rev. B*, vol. 45, pp. 4171–4180, 8 Feb. 1992. DOI: 10.1103/PhysRevB.45.4171.

[56] J. Bude, K. Hess, and G. J. Iafrate, "Impact ionization: beyond the Golden Rule," *Semiconductor Science and Technology*, vol. 7, no. 3B, B506, Mar. 1992. DOI: 10.1088/0268-1242/7/3B/132.

[57] N. Sano, T. Aoki, M. Tomizawa, and A. Yoshii, "Electron transport and impact ionization in Si," *Phys. Rev. B*, vol. 41, pp. 12 122–12 128, 17 Jun. 1990. DOI: 10.1103/PhysRevB.41.12122.

[58] Y. Kamakura, R. Fujita, K. Konaga, Y. Ueoka, N. Mori, and T. Kotani, "Full band Monte Carlo simulation of impact ionization in wide bandgap semiconductors based on ab initio calculation," in *2016 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, 2016, pp. 47–51. DOI: 10.1109/SISPAD.2016.7605145.

[59] T. Kunikiyo, M. Takenaka, M. Morifuji, K. Taniguchi, and C. Hamaguchi, "A model of impact ionization due to the primary hole in silicon for a full band Monte Carlo simulation," *Journal of Applied Physics*, vol. 79, no. 10, pp. 7718–7725, May 1996. DOI: 10.1063/1.362375.

[60] B. K. Ridley, *Quantum Processes in Semiconductors*. 2013.

[61] T. Bordelon, X.-L. Wang, C. Maziar, and A. Tasch, "An evaluation of energy transport models for silicon device simulation," *Solid-State Electronics*, vol. 34, no. 6, pp. 617–628, 1991. DOI: https://doi.org/10.1016/0038-1101(91)90136-M.

[62] M. V. Fischetti and S. E. Laux, "Monte carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects," *Phys. Rev. B*, vol. 38, pp. 9721–9745, 14 Nov. 1988. DOI: 10.1103/PhysRevB.38.9721.

[63] T. Yamada and D. Ferry, "Monte Carlo simulation of hole transport in strained Si1 - xGex," *Solid-State Electronics*, vol. 38, no. 4, pp. 881–890, 1995. DOI: https://doi.org/10.1016/0038-1101(94)00123-W.

[64] C. Jacoboni, C. Canali, G. Ottaviani, and A. Alberigi Quaranta, "A review of some charge transport properties of silicon," *Solid-State Electronics*, vol. 20, no. 2, pp. 77–89, 1977. DOI: https://doi.org/10.1016/0038-1101(77)90054-5.

[65] J. Roldan, F. Gamiz, J. López-Villanueva, and J. Carceller, "A Monte Carlo study on the electron-transport properties of high-performance strained-Si on relaxed Si1-xGex channel MOSFETs," *Journal of Applied Physics*, vol. 80, pp. 5121–5128, Dec. 1996. DOI: 10.1063/1.363493.

[66] A. S. Grove, *Physics and Technology of Semiconductor Devices*. 1967.

[67] R. Van Overstraeten and H. De Man, "Measurement of the ionization rates in diffused silicon p-n junctions," *Solid-State Electronics*, vol. 13, no. 5, pp. 583–608, 1970. DOI: https://doi.org/10.1016/0038-1101(70)90139-5.

[68] A. G. Chynoweth, "Ionization Rates for Electrons and Holes in Silicon," *Phys. Rev.*, vol. 109, pp. 1537–1540, 5 Mar. 1958. DOI: 10.1103/PhysRev.109.1537.

[69] G. A. Baraff, "Distribution Functions and Ionization Rates for Hot Electrons in Semiconductors," *Phys. Rev.*, vol. 128, pp. 2507–2517, 6 Dec. 1962. DOI: 10.1103/PhysRev.128.2507.

[70] C. A. Lee, R. A. Logan, R. L. Batdorf, J. J. Kleimack, and W. Wiegmann, "Ionization Rates of Holes and Electrons in Silicon," *Phys. Rev.*, vol. 134, A761–A773, 3A May 1964. DOI: 10.1103/PhysRev.134.A761.

[71] W. Grant, "Electron and hole ionization rates in epitaxial silicon at high electric fields," *Solid-State Electronics*, vol. 16, no. 10, pp. 1189–1203, 1973. DOI: https://doi.org/10.1016/0038-1101(73)90147-0.

[72] D. R. Decker and C. N. Dunn, "Temperature dependence of carrier ionization rates and saturated velocities in silicon," *Journal of Electronic Materials*, vol. 4, pp. 527–547, 1975.

[73] R. Helleboid, D. Rideau, J. Grebot, I. Nicholson, N. Moussy, O. Saxod, M. Basset, A. Zimmer, B. Mamdy, D. Golanski, M. Agnew, S. Pellegrini, M. Sicre, C. Buj, G. Marchand, J. Saint-Martin, M. Pala, and P. Dollfus, "Comprehensive Modeling and Characterization of Photon Detection Efficiency and Jitter Tail in Advanced SPAD Devices," *IEEE Journal of the Electron Devices Society*, vol. 10, pp. 584–592, Apr. 2022. DOI: 10.1109/JEDS.2022.3168365.

[74] M. Sicre, M. Agnew, C. Buj, J. Coignus, D. Golanski, R. Helleboid, B. Mamdy, I. Nicholson, S. Pellegrini, D. Rideau, D. Roy, and F. Calmon, "Dark Count Rate in Single-Photon Avalanche Diodes: Characterization and Modeling study," in *ESSCIRC 2021 - IEEE 47th European Solid State Circuits Conference (ESSCIRC)*, Sep. 2021, pp. 143–146. DOI: 10.1109/ESSCIRC53450.2021.9567806.

[75] W. Oldham, R. Samuelson, and P. Antognetti, "Triggering phenomena in avalanche diodes," *IEEE Transactions on Electron Devices*, vol. 19, no. 9, pp. 1056–1060, 1972. DOI: 10.1109/T-ED.1972.17544.

[76] "Lumerical Inc.,"

[77] R. McIntyre, "On the avalanche initiation probability of avalanche diodes above the breakdown voltage," *IEEE Transactions on Electron Devices*, vol. 20, no. 7, pp. 637–641, 1973. DOI: 10.1109/T-ED.1973.17715.

[78] R. Helleboid, D. Rideau, I. Nicholson, J. Grebot, B. Mamdy, G. Mugny, M. Basset, M. Agnew, D. Golanski, S. Pellegrini, J. Saint-Martin, M. Pala, and P. Dollfus, "A fokker–planck-based monte carlo method for electronic transport and avalanche simulation in single-photon avalanche diodes," *Journal of Physics D: Applied Physics*, vol. 55, no. 50, p. 505 102, Oct. 2022. DOI: 10.1088/1361-6463/ac9b6a.

[79] M. Sicre, M. Agnew, C. Buj, C. Coutier, D. Golanski, R. Helleboid, B. Mamdy, I. Nicholson, S. Pellegrini, D. Rideau, D. Roy, and F. Calmon, "Statistical measurements and Monte-Carlo simulations of DCR in SPADs," in *ESSCIRC 2022- IEEE 48th European Solid State Circuits Conference (ESSCIRC)*, Sep. 2022, pp. 193–196. DOI: 10.1109/ESSCIRC55480.2022.9911519.

[80] C. Jacoboni, "Theory of Electron Transport in Semiconductors," 2010.

[81] G. A. Pavliotis, *Stochastic Processes and Applications : Diffusion Processes, the Fokker-Planck and Langevin Equations*. Nov. 2014.

[82] G. E. P. Box and M. E. Muller, "A Note on the Generation of Random Normal Deviates," *Annals of Mathematical Statistics*, vol. 29, pp. 610–611, 1958.

[83] N. Arora, J. Hauser, and D. Roulston, "Electron and hole mobilities in silicon as a function of concentration and temperature," *IEEE Transactions on Electron Devices*, vol. 29, no. 2, pp. 292–295, 1982. DOI: 10.1109/T-ED.1982.20698.

[84] C. Canali, G. Majni, R. Minder, and G. Ottaviani, "Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature," *IEEE Transactions on Electron Devices*, vol. 22, no. 11, pp. 1045–1047, 1975. DOI: 10.1109/T-ED.1975.18267.

[85] O. Triebl, "Reliability issues in high-voltage semiconductor devices," reposiTUm, Dissertation, Technische Universität Wien, 2012.

[86] R. J. Mcintyre, "A new look at impact ionization-Part I: A theory of gain, noise, breakdown probability, and frequency response," *IEEE Transactions on Electron Devices*, vol. 46, pp. 1623–1631, 1999.

[87] P. Yuan, K. A. Anselm, C. C. Hu, H. Nie, C. Lenox, A. L. J. Holmes, B. G. Streetman, J. C. Campbell, and R. J. Mcintyre, "A new look at impact ionization-Part II: Gain and noise in short avalanche photodiodes," *IEEE Transactions on Electron Devices*, vol. 46, pp. 1632–1639, 1999.

[88] A. Ingargiola, M. Assanelli, A. Gallivanoni, I. Rech, M. Ghioni, and S. Cova, "Avalanche buildup and propagation effects on photon-timing jitter in Si-SPAD with non-uniform electric field," *Proc SPIE*, vol. 7320, May 2009. DOI: 10.1117/12.818521.

[89] G. D. Watkins, "Intrinsic defects in silicon," *Materials Science in Semiconductor Processing*, vol. 3, no. 4, pp. 227–235, 2000. DOI: https://doi.org/10.1016/S1369-8001(00)00037-8.

[90] J. B. et al., *Point Defects in Semiconductors II. Experimental Aspects.* 1983.

[91] T. Nguyen, K. Keunen, V. Afanas'ev, and A. Stesmans, "Interface state energy distribution and Pb defects at Si(110)/SiO2 interfaces: Comparison to (111) and (100) silicon orientations," *Journal of Applied Physics*, vol. 109, pp. 013 710–013 710, Jan. 2011. DOI: 10.1063/1.3527909.

[92] Y. Fedorenko, L. Truong, V. Afanas'ev, A. Stesmans, Z. Zhang, and S. Campbell, "Impact of nitrogen incorporation on interface states in (100)Si/HfO2," *Journal of Applied Physics*, vol. 98, pp. 123 703–123 703, Dec. 2005. DOI: 10.1063/1.2140871.

[93] Y. Varshni, "Temperature dependence of the energy gap in semiconductors," *Physica*, vol. 34, no. 1, pp. 149–154, 1967. DOI: https://doi.org/10.1016/0031-8914(67)90062-6.

[94] M. Tyagi and R. Van Overstraeten, "Minority carrier recombination in heavily-doped silicon," *Solid-State Electronics*, vol. 26, no. 6, pp. 577–597, 1983. DOI: https://doi.org/10.1016/0038-1101(83)90174-0.

[95] J. A. del Alamo and R. M. Swanson, "Modelling of minority-carrier transport in heavily doped silicon emitters," *Solid-State Electronics*, vol. 30, no. 11, pp. 1127–1136, 1987. DOI: https://doi.org/10.1016/0038-1101(87)90077-3.

[96] G. A. M. Hurkx, D. B. M. Klaassen, and M. Knuvers, "A new recombination model for device simulation including tunneling," *IEEE Transactions on Electron Devices*, vol. 39, pp. 331–338, 1992.

[97] "Spectroscopie du courant d'obscurité induit par les effets de déplacement atomique des radiations spatiales et nucléaires dans les capteurs d'images cmos à photodiode pincée," Ph.D. dissertation.

[98] B. Yamina, B. Bouabdallah, B. Zebentout, M. Debab, and Z. Benamara, "Numerical Analysis of Short Channel Poly-Si TFTs Under Off Conditions," *Sensor Letters*, vol. 9, pp. 2360–2363, Dec. 2011. DOI: 10.1166/sl.2011.1767.

[99] T. N. Kevin Lemiere Christophe Inguimbert, "Comparison of 1D and 3D Electric Field Enhancement Analytical Models to Calculate the Dark Current Non Uniformity," *NSREC, SAN ANTONIO, United States*, Jul. 2019.

[100] A. Schenk, "Rigorous theory and simplified model of the band-to-band tunneling in silicon," *Solid-State Electronics*, vol. 36, no. 1, pp. 19–34, 1993. DOI: https://doi.org/10.1016/0038-1101(93)90065-X.

[101] R. Kubo, "Statistical-Mechanical Theory of Irreversible Processes. I. General Theory and Simple Applications to Magnetic and Conduction Problems," *Journal of the Physical Society of Japan*, vol. 12, no. 6, pp. 570–586, 1957. DOI: 10.1143/JPSJ.12.570.

[102] R. Enderlein and K. Peuker, "On the theory of the electric conductivity of solids in a strong electric field," *physica status solidi (b)*, vol. 48, no. 1, pp. 231–241, 1971. DOI: https://doi.org/10.1002/pssb.2220480122.

[103] A. Heigl, A. Schenk, and G. Wachutka, "Correction to the Schenk Model of Band-to-Band Tunneling in Silicon Applied to the Simulation of Nanowire Tunneling Transistors," Jun. 2009, pp. 1–4. DOI: 10.1109/IWCE.2009.5091099.

[104] E. Kane, "Zener tunneling in semiconductors," *Journal of Physics and Chemistry of Solids*, vol. 12, no. 2, pp. 181–188, 1960. DOI: https://doi.org/10.1016/0022-3697(60)90035-4.

[105] W. Brauer and H.-W. Streitwolf, *Theoretische Grundlagen der Halbleiterphysik*. Akademie-Verlag, Berlin, 1977.

[106] C. Canali, C. Jacoboni, F. Nava, G. Ottaviani, and A. Alberigi-Quaranta, "Electron drift velocity in silicon," *Phys. Rev. B*, vol. 12, pp. 2265–2284, 6 Sep. 1975. DOI: 10.1103/PhysRevB.12.2265.

[107] R. M. Martin, "Dielectric Screening Model for Lattice Vibrations of Diamond-Structure Crystals," *Phys. Rev.*, vol. 186, pp. 871–884, 3 Oct. 1969. DOI: 10.1103/PhysRev.186.871.

[108] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. National Bureau of Standards, Appl. Math. Ser. Nr. 55, 1964.

[109] Y. Kang, H. Lu, Y.-H. Lo, D. Bethune, and W. Risk, "Dark count probability and quantum efficiency of avalanche photodiodes for single-photon detection," *Applied Physics Letters*, vol. 83, pp. 2955–2957, Oct. 2003. DOI: 10.1063/1.1616666.

[110] W. Kindt and H. Zeijl, "Modelling and Fabrication of Geiger mode Avalanche Photodiodes," *Nuclear Science, IEEE Transactions on*, vol. 45, pp. 715–719, Jul. 1998. DOI: 10.1109/23.682621.

[111] "TSUPREM-4 twodimensional process simulation program," *Technology Modeling Associates, Inc.*,

[112] Y. Xu, P. Xiang, X. Xie, and Y. Huang, "A new modeling and simulation method for important statistical performance prediction of single photon avalanche diode detectors," *Semiconductor Science and Technology*, vol. 31, 2016.

[113] T. Knezevic, L. Nanver, and T. Suligoj, "2D dark-count-rate modeling of PureB single-photon avalanche diodes in a TCAD environment," Feb. 2018, p. 55. DOI: 10.1117/12.2290757.

[114] A. Panglosse, P. Martin-Gonthier, O. Marcelot, C. Virmontois, O. Saint-Pe, and P. Magnan, "Dark Count Rate Modeling in Single-Photon Avalanche Diodes," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. PP, pp. 1–9, Feb. 2020. DOI: 10.1109/TCSI.2020.2971108.

[115] J. McPherson, "Stress Dependent Activation Energy," in *24th International Reliability Physics Symposium*, 1986, pp. 12–18. DOI: 10.1109/IRPS.1986.362105.

[116] J. Robertson, "Defects and hydrogen in amorphous silicon nitride," *Philosophical Magazine B*, vol. 69, no. 2, pp. 307–326, 1994. DOI: 10.1080/01418639408240111.

[117] J. Robertson, W. L. Warren, and J. Kanicki, "Nature of the Si and N dangling bonds in silicon nitride," *Journal of Non Crystalline Solids*, vol. 187, pp. 297–300, Jul. 1995. DOI: 10.1016/0022-3093(95)00153-0.

[118] W. L. Warren, P. M. Lenahan, and S. E. Curry, "First observation of paramagnetic nitrogen dangling-bond centers in silicon nitride," *Phys. Rev. Lett.*, vol. 65, pp. 207–210, 2 Jul. 1990. DOI: 10.1103/PhysRevLett.65.207.

[119] D. T. Krick, P. M. Lenahan, and J. Kanicki, "Nature of the dominant deep trap in amorphous silicon nitride," *Phys. Rev. B*, vol. 38, pp. 8226–8229, 12 Oct. 1988. DOI: 10.1103/PhysRevB.38.8226.

[120] W. L. Warren, J. Kanicki, J. Robertson, and P. M. Lenahan, "Energy level of the nitrogen dangling bond in amorphous silicon nitride," *Applied Physics Letters*, vol. 59, no. 14, pp. 1699–1701, Sep. 1991. DOI: 10.1063/1.106222.

[121] J. F. Justo, F. de Brito Mota, and A. Fazzio, "First-principles investigation of $a-SiN_x$: H," *Phys. Rev. B*, vol. 65, p. 073 202, 7 Feb. 2002. DOI: 10.1103/PhysRevB.65.073202.

[122] F. de Brito Mota, J. F. Justo, and A. Fazzio, "Hydrogen role on the properties of amorphous silicon nitride," *Journal of Applied Physics*, vol. 86, no. 4, pp. 1843–1847, Aug. 1999. DOI: 10.1063/1.370977.

[123] J. Robertson, "Electronic structure of silicon nitride," *Philosophical Magazine B*, vol. 63, no. 1, pp. 47–77, 1991. DOI: 10.1080/01418639108224430.

[124] F. J. Grunthaner, P. J. Grunthaner, R. P. Vasquez, B. F. Lewis, J. Maserjian, and A. Madhukar, "High-Resolution X-Ray Photoelectron Spectroscopy as a Probe of Local Atomic Structure: Application to Amorphous SiO_2 and the Si-SiO_2 Interface," *Phys. Rev. Lett.*, vol. 43, pp. 1683–1686, 22 Nov. 1979. DOI: 10.1103/PhysRevLett.43.1683.

[125] A. A. Demkov and O. F. Sankey, "Growth Study and Theoretical Investigation of the Ultrathin Oxide SiO_2 – Si Heterojunction," *Phys. Rev. Lett.*, vol. 83, pp. 2038–2041, 10 Sep. 1999. DOI: 10.1103/PhysRevLett.83.2038.

[126] Y. Nishi, "Study of Silicon-Silicon Dioxide Structure by Electron Spin Resonance I," *Japanese Journal of Applied Physics*, vol. 10, no. 1, p. 52, Jan. 1971. DOI: 10.1143/JJAP.10.52.

[127] A. Stirling, A. Pasquarello, J.-C. Charlier, and R. Car, "Dangling Bond Defects at Si – SiO_2 Interfaces: Atomic Structure of the P_{b1} Center," *Phys. Rev. Lett.*, vol. 85, pp. 2773–2776, 13 Sep. 2000. DOI: 10.1103/PhysRevLett.85.2773.

[128] S. Hofstein, "Stabilization of MOS devices," *Solid-State Electronics*, vol. 10, no. 7, pp. 657–670, 1967. DOI: https://doi.org/10.1016/0038-1101(67)90096-2.

[129] S. Karna, H. Kurtz, W. Shedd, R. Pugh, and B. Singaraju, "New fundamental defects in a-SiO/sub 2/," *IEEE Transactions on Nuclear Science*, vol. 46, no. 6, pp. 1544–1552, 1999. DOI: 10.1109/23.819119.

[130] S. Karna, A. Pineda, R. Pugh, W. Shedd, and T. Oldham, "Electronic structure theory and mechanisms of the oxide trapped hole annealing process," *IEEE Transactions on Nuclear Science*, vol. 47, no. 6, pp. 2316–2321, 2000. DOI: 10.1109/23.903771.

[131] S. Agnello, "Gamma ray induced processes of point defect conversion in silica," Ph.D. dissertation, 2000.

[132] Z.-Y. Lu, C. J. Nicklaw, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, "Structure, Properties, and Dynamics of Oxygen Vacancies in Amorphous SiO_2," *Phys. Rev. Lett.*, vol. 89, p. 285 505, 28 Dec. 2002. DOI: 10.1103/PhysRevLett.89.285505.

[133] A. Lelis, T. Oldham, H. Boesch, and F. McLean, "The nature of the trapped hole annealing process," *IEEE Transactions on Nuclear Science*, vol. 36, no. 6, pp. 1808–1815, 1989. DOI: 10.1109/23.45373.

[134] A. J. Lelis and T. R. Oldham, "Time dependence of switching oxide traps," *IEEE Transactions on Nuclear Science*, vol. 41, pp. 1835–1843, 1994.

[135] M. Denais, V. Huard, C. Parthasarathy, G. Ribes, F. Perrier, N. Revil, and A. Bravaix, "Interface trap generation and hole trapping under NBTI and PBTI in advanced CMOS technology with a 2-nm gate oxide," *IEEE Transactions on Device and Materials Reliability*, vol. 4, no. 4, pp. 715–722, 2004. DOI: 10.1109/TDMR.2004.840856.

[136] B. E. Deal, "Standardized Terminology for Oxide Charges Associated with Thermally Oxidized Silicon," *Journal of The Electrochemical Society*, vol. 127, no. 4, p. 979, Apr. 1980. DOI: 10.1149/1.2129800.

[137] D. Fleetwood, "'Border traps' in MOS devices," *IEEE Transactions on Nuclear Science*, vol. 39, no. 2, pp. 269–271, 1992. DOI: 10.1109/23.277495.

[138] G. Smith, K. McIntosh, J. Donnelly, J. Funk, L. Mahoney, and S. Verghese, "Reliable InP-based Geiger-mode Avalanche Photodiode Arrays," *SPIE*, vol. 7320, May 2009. DOI: 10.1117/12.819126.

[139] M. Sicre, X. Federspiel, D. Roy, B. Mamby, C. Coutier, and F. Calmon, "Identification of stress factors and degradation mechanisms inducing DCR drift in SPADs," in *2022 IEEE International Integrated Reliability Workshop (IIRW)*, 2022, pp. 1–5. DOI: 10.1109/IIRW56459.2022.10032759.

[140] E. Takeda, "Hot-carrier effects in submicrometre MOS VLSIs," *IEE Proceedings: Solid-State Electron Devices*, vol. 131, pp. 153–162, Oct. 1984.

[141] D. Lachenal, Y. Rey-Tauriac, L. Boissonnet, and A. Bravaix, "Reliability investigation of NLDEMOS in 0.13µm SOI CMOS technology," Jan. 2006, pp. 555–558. DOI: 10.1109/ICMEL.2006.1651026.

[142] J.-P. Carrere, S. Place, J.-P. Oddou, D. Benoit, and F. Roy, "CMOS image sensor: Process impact on dark current," Jun. 2014, pp. 3C.1.1–3C.1.6. DOI: 10.1109/IRPS.2014.6860620.

[143] A. Bravaix, C. Guerin, V. Huard, D. Roy, J. Roux, and E. Vincent, "Hot-Carrier acceleration factors for low power management in DC-AC stressed 40nm NMOS node at high temperature," in *2009 IEEE International Reliability Physics Symposium*, 2009, pp. 531–548. DOI: 10.1109/IRPS.2009.5173308.

[144] B. Kaczer, T. Grasser, P. Roussel, J. Franco, R. Degraeve, L.-Å. Ragnarsson, E. Simoen, G. Groeseneken, and H. Reisinger, "Origin of NBTI variability in deeply scaled pFETs," Jun. 2010, pp. 26–32. DOI: 10.1109/IRPS.2010.5488856.

[145] A. Bravaix, D. Angot, V. Huard, L. Rahhal, A. Cros, X. Federspiel, A. Bajolet, Y. Carminati, M. Saliva, E. Pion, and F. Cacho, "BTI variability Fundamental understandings and Impact on digital logic by the use of extensive dataset," Dec. 2013. DOI: 10.1109/IEDM.2013.6724636.

[146] B. Kaczer, J. Franco, M. Cho, T. Grasser, P. J. Roussel, S. Tyaginov, M. Bina, Y. Wimmer, L.-M. Prócel, L. Trojman, F. Crupi, G. Pitner, V. Putcha, P. Weckx, E. Bury, Z. Ji, A. D. Keersgieter, T. Chiarella, N. Horiguchi, G. Groeseneken, and A. V.-Y. Thean, "Origins and implications of increased channel hot carrier variability in nFinFETs," *2015 IEEE International Reliability Physics Symposium*, 3B.5.1–3B.5.6, 2015.

[147] X. Federspiel, C. Diouf, F. Cacho, and E. Vincent, "Comparison of variability of HCI induced drift for SiON and HKMG devices," Apr. 2020, pp. 1–5. DOI: 10.1109/IRPS45951.2020.9128326.

[148] A. Asenov, R. Balasubramaniam, A. Brown, and J. Davies, "RTS amplitudes in decananometer MOSFETs: 3-D simulation study," *Electron Devices, IEEE Transactions on*, vol. 50, pp. 839–845, Apr. 2003. DOI: 10.1109/TED.2003.811418.

[149] K. Takeuchi, T. Nagumo, S. Yokogawa, K. Imai, and Y. Hayashi, "Single-Charge-Based Modeling of Transistor Characteristics Fluctuations Based on Statistical Measurement of RTN Amplitude," Jul. 2009, pp. 54–55.

[150] A. Ghetti, C. Monzio Compagnoni, A. S. Spinelli, and A. Visconti, "Comprehensive Analysis of Random Telegraph Noise Instability and Its Scaling in Deca-Nanometer Flash Memories," *IEEE Transactions on Electron Devices*, vol. 56, no. 8, pp. 1746–1752, Aug. 2009. DOI: 10.1109/TED.2009.2024031.

[151] V. Huard, C. Parthasarathy, C. Guerin, T. Valentin, E. Pion, M. Mammasse, N. Planes, and L. Camus, "NBTI degradation: From transistor to SRAM arrays," Apr. 2008, pp. 289–300. DOI: 10.1109/RELPHY.2008.4558900.

[152] M. Sicre, X. Federspiel, B. Mamdy, D. Roy, and F. Calmon, "Characterization and modeling of DCR and DCR drift variability in SPADs," in *2023 IEEE International Reliability Physics Symposium (IRPS)*, 2023, pp. 1–5. DOI: 10.1109/IRPS48203.2023.10117921.

[153] I. Zbierska, L. Militaru, F. Calmon, F. Sylvain, and G.-N. Lu, "Investigation of Electrical Characteristics of Multi-Gate Bulk nMOSFET," May 2014, pp. 95–98. DOI: 10.1109/MIEL.2014.6842094.

[154] S. E. Tyaginov, "Physics-Based Modeling of Hot-Carrier Degradation," in *Hot Carrier Degradation in Semiconductor Devices*, T. Grasser, Ed., Springer International Publishing, 2015, pp. 105–150. DOI: 10.1007/978-3-319-08994-2{_}4.

[155] J. W. Lyding, K. Hess, and I. C. Kizilyalli, "Reduction of hot electron degradation in metal oxide semiconductor transistors by deuterium processing," *Applied Physics Letters*, vol. 68, no. 18, pp. 2526–2528, 1996. DOI: 10.1063/1.116172.

[156] W. McMahon, A. Haggag, and K. Hess, "Reliability scaling issues for nanoscale devices," *IEEE Transactions on Nanotechnology*, vol. 2, no. 1, pp. 33–38, 2003. DOI: 10.1109/TNANO.2003.808515.

[157] W. McMahon, K. Matsuda, J. Lee, K. Hess, and J. Lyding, "The effects of a multiple carrier model of interface trap generation on lifetime extraction for MOSFETs," English (US), in *2002 International Conference on Modeling and Simulation of Microsystems - MSM 2002*, M. Laudon and B. Romanowicz, Eds., ser. 2002 International Conference on Modeling and Simulation of Microsystems - MSM 2002, 2002 International Conference on Modeling and Simulation of Microsystems - MSM 2002 ; Conference date: 21-04-2002 Through 25-04-2002, 2002, pp. 576–579.

[158] K. Hess, L. Register, B. Tuttle, J. Lyding, and I. Kizilyalli, "Impact of nanostructure research on conventional solid-state electronics: The giant isotope effect in hydrogen desorption and CMOS lifetime," *Physica E: Low-dimensional Systems and Nanostructures*, vol. 3, no. 1, pp. 1–7, 1998. DOI: https://doi.org/10.1016/S1386-9477(98)00211-2.

[159] S. H. Hong, S. M. Nam, B. O. Yun, B. J. Lee, C. G. Yu, and J. T. Park, "Temperature dependence of hot carrier induced MOSFET degradation at low gate bias," *Microelectronics Reliability*, vol. 39, no. 6, pp. 809–814, 1999, European Symposium on Reliability of Electron Devices, Failure Physics and Analysis. DOI: https://doi.org/10.1016/S0026-2714(99)00105-5.

[160] C. Leung, "New mechanism of hot carrier generation in very short channel MOSFETs," English, *Electronics Letters*, vol. 31, 139–141(2), 2 Jan. 1995. DOI: 10.1049/el:19950091.

[161] S. Rauch, G. Rosa, and F. Guarin, "Role of E-E scattering in the enhancement of channel hot carrier degradation of deep sub-micron NMOSFETs at high VGS conditions," vol. 1, Feb. 2001, pp. 399–405. DOI: 10.1109/RELPHY.2001.922933.

[162] P. A. Childs and C. C. C. Leung, "A one-dimensional solution of the Boltzmann transport equation including electron–electron interactions," *Journal of Applied Physics*, vol. 79, no. 1, pp. 222–227, Jan. 1996. DOI: 10.1063/1.360935.

[163] S. Rauch, F. Guarin, and G. LaRosa, "Impact of E-E scattering to the hot carrier degradation of deep submicron NMOSFETs," *IEEE Electron Device Letters*, vol. 19, no. 12, pp. 463–465, 1998. DOI: 10.1109/55.735747.

[164] C. C. Hu, "Lucky-electron model of channel hot electron emission," *1979 International Electron Devices Meeting*, pp. 22–25, 1979.

[165] S. Tyaginov, I. Starkov, H. Enichlmair, J.-M. Park, C. Jungemann, and T. Grasser, "Physics-Based Hot-Carrier Degradation Models," *ECS Transactions*, vol. 35, pp. 321–352, Apr. 2011. DOI: 10.1149/1.3572292.

[166] A. Bravaix, V. Huard, F. Cacho, X. Federspiel, and D. Roy, "Hot-Carrier Degradation in Decananometer CMOS Nodes: From an Energy-Driven to a Unified Current Degradation Modeling by a Multiple-Carrier Degradation Process," pp. 57–103, Oct. 2015. DOI: 10.1007/978-3-319-08994-2_3.

[167] C. Guerin, V. Huard, and A. Bravaix, "General framework about defect creation at the Si/SiO2 interface," *Journal of Applied Physics*, vol. 105, pp. 114 513–114 513, Jul. 2009. DOI: 10.1063/1.3133096.

[168] S. Tyaginov, I. Starkov, O. Triebl, J. Cervenka, C. Jungemann, S. Carniello, J. Park, H. Enichlmair, M. Karner, C. Kernstock, E. Seebacher, R. Minixhofer, H. Ceric, and T. Grasser, "Interface traps density-of-states as a vital component for hot-carrier degradation modeling," *Microelectronics Reliability*, vol. 50, no. 9, pp. 1267–1272, 2010, 21st European Symposium on the Reliability of Electron Devices, Failure Physics and Analysis. DOI: https://doi.org/10.1016/j.microrel.2010.07.030.

[169] C. Jungemann and B. Meinerzhagen, *Hierarchical device simulation: The monte-carlo perspective*. Springer, New York, 2003.

[170] K. Rupp, T. Grasser, and A. Jüngel, "On the feasibility of spherical harmonics expansions of the Boltzmann transport equation for three-dimensional device geometries," in *2011 International Electron Devices Meeting*, 2011, pp. 34.1.1–34.1.4. DOI: 10.1109/IEDM.2011.6131667.

[171] J. Bude and K. Hess, "Thresholds of impact ionization in semiconductors," *Journal of Applied Physics*, vol. 72, no. 8, pp. 3554–3561, 1992. DOI: 10.1063/1.351434.

[172] O. Penzin, A. Haggag, W. McMahon, E. Lyumkis, and K. Hess, "MOSFET degradation kinetics and its simulation," *IEEE Transactions on Electron Devices*, vol. 50, no. 6, pp. 1445–1450, 2003. DOI: 10.1109/TED.2003.813333.

[173] Z. Chen, K. Hess, J. Lee, J. Lyding, E. Rosenbaum, I. Kizilyalli, S. Chetlur, and R. Huang, "On the mechanism for interface trap generation in MOS transistors due to channel hot carrier stressing," *IEEE Electron Device Letters*, vol. 21, no. 1, pp. 24–26, 2000. DOI: 10.1109/55.817441.

[174] B. Persson and P. Avouris, "Local bond breaking via STM-induced excitations: the role of temperature," *Surface Science*, vol. 390, no. 1, pp. 45–54, 1997, Desorption Induced by Electronic Transitions. DOI: https://doi.org/10.1016/S0039-6028(97)00507-4.

[175] B. Tuttle and C. G. Van de Walle, "Structure, energetics, and vibrational properties of Si-H bond dissociation in silicon," *Phys. Rev. B*, vol. 59, pp. 12 884–12 889, 20 May 1999. DOI: 10.1103/PhysRevB.59.12884.

[176] K. Hess, A. Haggag, W. McMahon, K. Cheng, J. Lee, and J. Lyding, "The physics of determining chip reliability," *IEEE Circuits and Devices Magazine*, vol. 17, no. 3, pp. 33–38, 2001. DOI: 10.1109/101.933789.

[177] A. Haggag, W. McMahon, K. Hess, K. Cheng, J. Lee, and J. Lyding, "High-performance chip reliability from short-time-tests-statistical models for optical interconnect and HCI/TDDB/NBTI deep-submicron transistor failures," in *2001 IEEE International Reliability Physics Symposium Proceedings. 39th Annual (Cat. No.00CH37167)*, 2001, pp. 271–279. DOI: 10.1109/RELPHY.2001.922913.

[178] T. Erdogan, V. Mizrahi, P. J. Lemaire, and D. Monroe, "Decay of ultraviolet-induced fiber Bragg gratings," *Journal of Applied Physics*, vol. 76, no. 1, pp. 73–80, 1994. DOI: 10.1063/1.357062.

[179] M. Sicre, D. Roy, and F. Calmon, "Hot-Carrier Degradation modeling of DCR drift in SPADs," in *ESSDERC 2023 - IEEE 53rd European Solid-State Device Research Conference (ESSDERC)*, 2023, pp. 61–64. DOI: 10.1109/ESSDERC59256.2023.10268534.

[180] M. Michaillat, D. Rideau, F. Aniel, C. Tavernier, and H. Jaouen, "Full-Band Monte Carlo investigation of hole mobilities in SiGe, SiC and SiGeC alloys," *Thin Solid Films*, vol. 518, no. 9, pp. 2437–2441, 2010, Proceedings of the EMRS 2009 Spring Meeting Symposium I: Silicon and germanium issues for future CMOS devices. DOI: https://doi.org/10.1016/j.tsf.2009.10.147.

[181] C. Jacoboni and L. Reggiani, "The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials," *Rev. Mod. Phys.*, vol. 55, pp. 645–705, 3 Jul. 1983. DOI: 10.1103/RevModPhys.55.645.

[182] J. C. Slater and G. F. Koster, "Simplified LCAO Method for the Periodic Potential Problem," *Phys. Rev.*, vol. 94, pp. 1498–1524, 6 Jun. 1954. DOI: 10.1103/PhysRev.94.1498.

[183] J. I. Pankove and D. A. Kiewit, "Optical Processes in Semiconductors," *Journal of The Electrochemical Society*, vol. 119, no. 5, 156Ca, May 1972. DOI: 10.1149/1.2404256.

[184] P. M. Lenahan and J. Conley J. F., "What can electron paramagnetic resonance tell us about the Si/SiO2 system?" *Journal of Vacuum Science and Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, vol. 16, no. 4, pp. 2134–2153, Jul. 1998. DOI: 10.1116/1.590301.

[185] A. Haggag, W. Mcmahon, K. Hess, B. Fischer, and L. Register, "Impact of Scaling on CMOS Chip Failure Rate, and Design Rules for Hot Carrier Reliability," *VLSI Design*, vol. 13, Jan. 2001. DOI: 10.1155/2001/90787.

[186] W. Shockley and W. T. Read, "Statistics of the Recombinations of Holes and Electrons," *Phys. Rev.*, vol. 87, pp. 835–842, 5 Sep. 1952. DOI: 10.1103/PhysRev.87.835.

[187] D. Biegelsen, N. Johnson, M. Stutzmann, E. Poindexter, and P. Caplan, "Native defects at the Si/SiO2 interface-amorphous silicon revisited," *Applications of Surface Science*, vol. 22-23, pp. 879–890, 1985. DOI: https://doi.org/10.1016/0378-5963(85)90220-X.

[188] D. F. Swinehart, "The Beer-Lambert Law," *Journal of Chemical Education*, vol. 39, no. 7, p. 333, 1962. DOI: 10.1021/ed039p333.

[189] A. Stesmans, "Passivation of Pb0 and Pb1 interface defects in thermal (100) Si/SiO2 with molecular hydrogen," *Applied Physics Letters*, vol. 68, no. 15, pp. 2076–2078, 1996. DOI: 10.1063/1.116308.

[190] A. Stesmans, "Revision of H2 passivation of P2 interface defects in standard (111)Si/-SiO2," *Applied Physics Letters*, vol. 68, no. 19, pp. 2723–2725, 1996.

[191] C. E. Blat, E. H. Nicollian, and E. H. Poindexter, "Mechanism of negative-bias-temperature instability," *Journal of Applied Physics*, vol. 69, no. 3, pp. 1712–1720, 1991. DOI: 10.1063/1.347217.

[192] *The Engineering ToolBox, Illuminance - Recommended Light Level and Materials - Light Reflecting Factors.* 2012.

[193] T. Grasser, C. Jungemann, H. Kosina, B. Meinerzhagen, and S. Selberherr, "Advanced Transport Models for Sub-Micrometer Devices," in *Simulation of Semiconductor Processes and Devices 2004*, G. Wachutka and G. Schrag, Eds., Vienna: Springer Vienna, 2004, pp. 1–8.

[194] A. Zaka, Q. Rafhay, M. Iellina, P. Palestri, R. Clerc, D. Rideau, D. Garetto, E. Dornel, J. Singer, G. Pananakakis, C. Tavernier, and H. Jaouen, "On the accuracy of current TCAD hot carrier injection models in nanoscale devices," *Solid-State Electronics*, vol. 54, no. 12, pp. 1669–1674, 2010. DOI: https://doi.org/10.1016/j.sse.2010.06.014.

[207] S. Gasiorowicz, *Quantum Mechanics. 3rd Edition,journal=John Wiley, New York*. 2003.

[208] J. H. Zheng, H. S. Tan, and S. C. Ng, "Theory of non-radiative capture of carriers by multiphonon processes for deep centres in semiconductors," *Journal of Physics: Condensed Matter*, vol. 6, no. 9, p. 1695, Feb. 1994. DOI: 10.1088/0953-8984/6/9/012.

[209] D. Goguenheim and M. Lannoo, "Theoretical calculation of the electron-capture cross section due to a dangling bond at the Si(111)-SiO_2 interface," *Phys. Rev. B*, vol. 44, pp. 1724–1733, 4 Jul. 1991. DOI: 10.1103/PhysRevB.44.1724.

[210] A. Schenk, "An improved approach to the Shockley-Read-Hall recombination in inhomogeneous fields of space-charge regions," *Journal of Applied Physics*, vol. 71, pp. 3339–3349, 1992.

[211] A. Beckers and C. Enz, "Theoretical Limit of Low Temperature Subthreshold Swing in Field-Effect Transistors," *IEEE Electron Device Letters*, vol. PP, pp. 1–1, Dec. 2019. DOI: 10.1109/LED.2019.2963379.

[212] A. Haggag, "IC-REST(2): A time-to-market driven IC reliability statistics tool," Dissertation, University of Illinois at Urbana-Champaign, 2002.

Texte de synthèse en français

Contexte général

Imagerie à temps de vol

Les diodes à avalanche à photon unique (SPAD) sont utilisées pour les capteurs de temps de vol (ToF) afin de déterminer la distance d'une cible. Une source laser émet des impulsions de photons, qui sont réfléchies par une cible et la différence entre le temps d'émission et le temps de réception encode la distance comme illustré **Fig. 9.2**.

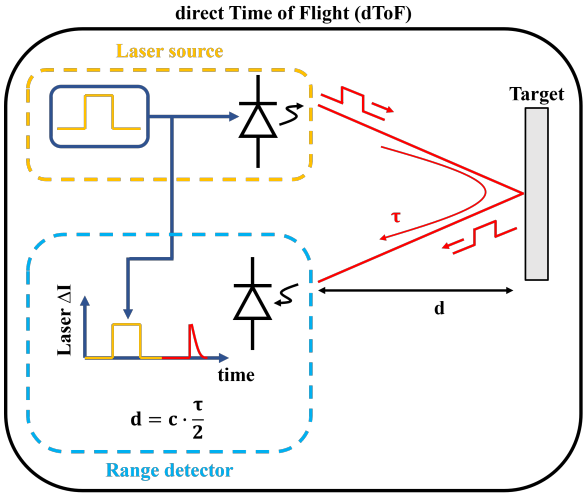


Figure 9.2: Illustration de la technique directe de temps de vol (ToF) utilisée pour déterminer la distance entre le capteur et la cible.

Ces photodétecteurs fonctionnent en déclenchant une avalanche de porteurs de charge lors de l'absorption des photons, ce qui entraîne une amplification substantielle qui peut être détectée. Cependant, ils sont sujets à des déclenchements parasites générés par des porteurs qui peuvent compromettre la précision de la distance mesurée. Ces évènement sont quantifiés sous la forme de taux de comptage dans l'obscurité (DCR). Il est donc crucial d'identifier et de réduire les sources potentielles de DCR.

Architecture de la photodiode à avalanche à photon unique étudiée

Une représentation typique de l'architecture SPAD étudiée [38][39] est représentée dans la **Fig. 9.3**. La couche p-epi agit comme une couche d'absorption qui génère des porteurs lorsqu'elle absorbe des photons. Les électrons générés se déplacent vers la couche n+, tandis que les trous générés se déplacent vers la couche p++. Il est important de confiner le processus de photo-génération à cette région afin de minimiser le bruit excessif qui peut résulter de la génération de porteurs à partir de défauts et de la diffusion de porteurs majoritaires en dehors de cette région. L'application d'une tension suffisamment élevée aux bornes du SPAD entraîne la collecte des électrons par les couches p+ et n++. Ces électrons acquièrent suffisamment d'énergie cinétique pour ioniser par impact certains des électrons de la couche de valence des atomes, générant des paires électron-trou et initiant potentiellement une avalanche auto-soutenue de porteurs. Pour minimiser l'absorption du milieu, ces couches doivent être maintenues minces. La tension de claquage en avalanche du SPAD est principalement déterminée par l'épaisseur de cette couche de multiplication. Des anneaux de garde composés de couches p- et n-well sont mis en œuvre pour augmenter la tension de claquage requise dans la périphérie par rapport à la région centrale, confinant ainsi efficacement l'avalanche au centre. En effet, la réduction du dopage aux bords conduit à une augmentation du champ électrique nécessaire pour l'ionisation par impact. Les couches de contact n++ et p++ garantissent l'établissement de contacts ohmiques avec les électrodes pour fournir un chemin de faible résistance pour l'évacuation du courant.

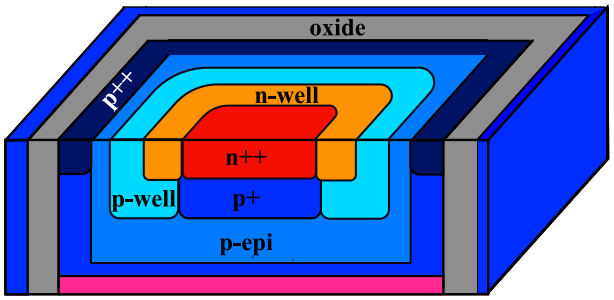


Figure 9.3: Représentation de l'architecture SPAD étudiée.

Circuit d'étouffement utilisé

Lorsqu'une avalanche est déclenchée dans un SPAD, elle génère un courant important qui entraîne une chute de potentiel à l'intérieur du dispositif qui peut être mesurée sous forme d'une impulsion de signal. Cependant, pour éviter une rupture physique, le courant du photodétecteur doit être limité et rapidement réduit. À cette fin, des techniques de circuit d'étouffement dit passif et actif ont été utilisées [40][41] pour supprimer l'avalanche en abaissant la polarisation inverse de la diode en dessous de la tension de claquage assurant ainsi un retour à l'équilibre pour la détection d'autres évènements. Les caractéristiques du circuit de déclenchement passif utilisé sont représentées dans la **Fig. 9.4**. Une combinaison d'un n-MOSFET de commande de faible tension (EN), d'un n-MOSFET en cascade de haute tension et d'une diode de limitation, tous connectés à l'anode du SPAD, permet de désactiver les pixels défectueux présentant un

DCR élevé et les secteurs de la matrice qui ne sont pas adressés ou qui ne sont pas éclairés. Une résistance en polysilicium R_q est enfouie dans une isolation de tranchée profonde (DTI) pour assurer l'étouffement des avalanches. L'ajout d'une capacité de couplage C_MOM filtre le courant continu à haute tension et ne couple que la partie haute fréquence alternative du signal. Une capacité de ballast C_b peut être ajoutée en parallèle pour ajouter plus d'atténuation agissant comme un diviseur de tension capacitif entre la cathode du SPAD et l'entrée de l'inverseur, qui est utilisé pour générer le signal de sortie du pixel. Le nœud flottant à l'entrée de l'inverseur est tiré vers le haut jusqu'à V_{DD} par un p-MOS de mise en forme d'impulsion pour assurer le franchissement du seuil de l'inverseur. La connexion d'un oscilloscope de la série Agilent 6000 à la sortie de l'inverseur donne des signaux typiques qui peuvent être affichés dans différentes fenêtres de temps. Un compteur/timer de fréquence universel Keysight 53131A est branché à la place de l'oscilloscope dans ce qui suit pour moyenner le nombre de comptes sur une échelle de temps plus longue.

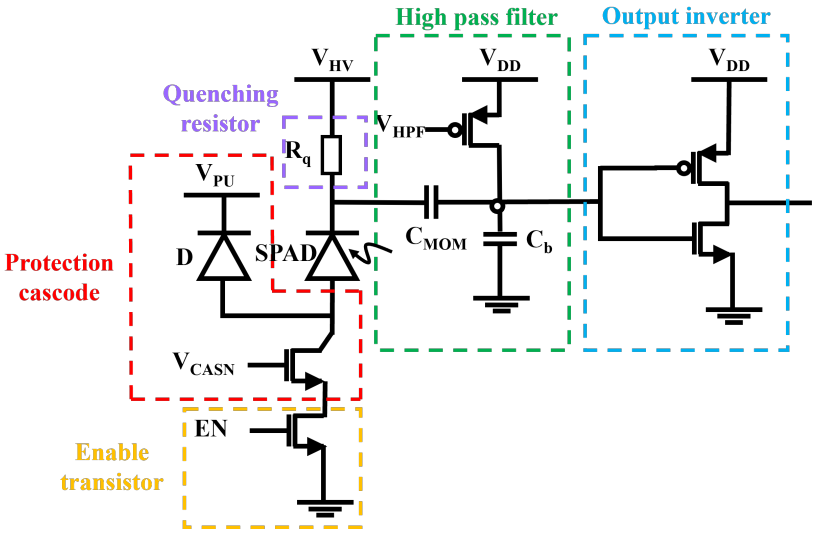


Figure 9.4: Schéma du circuit d'étouffement passif utilisé, qui est composé d'un transistor d'activation, qui allume/éteint le SPAD, d'un transistor de protection en cascade, qui protège le transistor d'activation (EN), d'une résistance de désactivation (R_q), qui garantit le retour à l'état d'équilibre après le déclenchement de l'avalanche, d'un filtre passe-haut, qui façonne le signal, et d'un inverseur de sortie, qui détecte le signal.

Figures de mérite d'un SPAD

Précédemment, le principe de fonctionnement du SPAD a été expliqué avec son circuit d'étouffement. L'attention est maintenant portée sur la description des principales caractéristiques des SPADs, telles qu'illustrées dans la **Fig. 9.5**. La tension de claquage en avalanche V_{BD} représente la tension appliquée pour déclencher les premières avalanches à une température donnée. Le taux de comptage d'obscurité (DCR) est défini comme le taux de comptage sans aucune lumière incidente définissant le signal minimum au-dessus duquel les avalanches sont réellement causés par des photons. Il dépend de la probabilité de claquage en avalanche P_t et du taux de recombinaison-

génération (R-G) des porteurs de charge. Il ne doit pas être confondu avec le courant d'oscurité, qui est le courant résiduel instantané ne déclenchant pas d'avalanches dans la SPAD sans lumière incidente. L'efficacité quantique (QE) est le flux de porteurs de charge généré qui contribue au courant photoélectrique sur le flux de photons incidents. Parfois, l'efficacité de détection de photons (PDE) est préférable lorsque le signal mesuré n'est pas directement corrélé aux photons incidents, par exemple pour les SPADs. Cela permet de corréler le nombre d'impulsions de signal détectées par rapport aux photons incidents. La diaphonie est un courant d'avalanche qui propage des porteurs dans les interconnexions ou les dispositifs voisins. La diaphonie électrique se produit entre des interconnexions parallèles par le biais d'un couplage inductif ou capacitif. La diaphonie optique provient de photons émis pendant une avalanche qui peuvent parcourir des centaines de microns en fonction de leur longueur d'onde avant de générer des paires électron-trou dans les dispositifs voisins. L'intervalle de temps entre l'absorption du photon et le début de l'avalanche présente des fluctuations statistiques liées à la localisation de l'absorption du photon et au déclenchement de l'avalanche. Le jitter est défini comme la largeur à mi-hauteur (FWHM) de cette distribution de temps d'absorption du photon. Les impulsions secondaires retracent leurs origines dans deux mécanismes potentiels. Les phénomènes sont appelés soit le décalage temporel (lag) ou la post-pulsation (afterpulsing), mais les deux sont liés à la libération de porteurs après quelques nanosecondes lorsque la Single-Photon Avalanche Diode (SPAD) récupère, déclenchant une post-pulsation dans le signal électrique. Soit une poche électrostatique capture certains porteurs dans une région neutre ou soit un piège à niveau profond capture et retient les porteurs pendant le processus d'avalanche. Ces effets limitent le temps mort du SPAD qui représente le temps durant lequel le SPAD ne peut détecter aucun autre évènement dû à l'armement ou la recharge du dispositif.

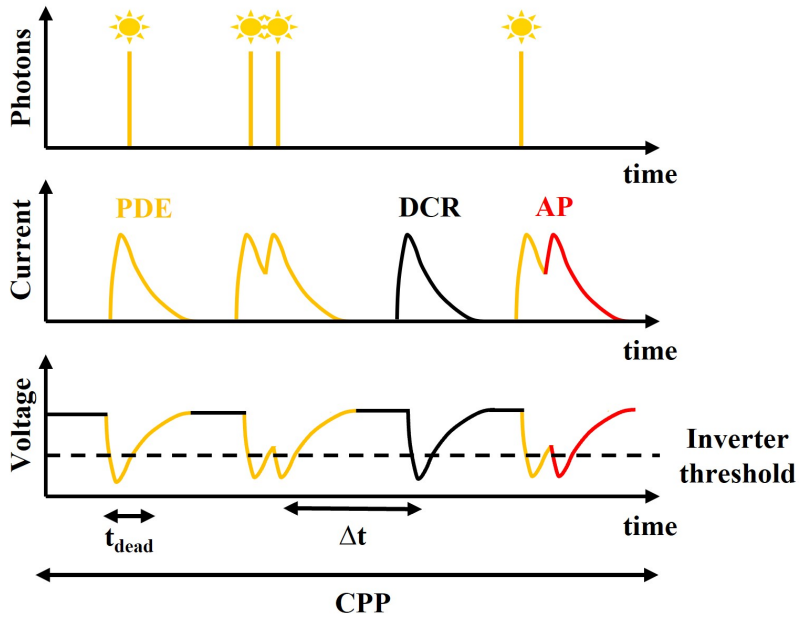


Figure 9.5: Illustration des caractéristiques d'un SPAD en relation avec le flux de photons, le courant et l'impulsion de tension. La détection des photons dépend de l'efficacité de détection de photons (PDE), tandis que les événements qui se produisent en l'absence de lumière sont appelés taux de comptage d'obscurité (DCR). Lorsque des avalanches sont déclenchées immédiatement après une autre avalanche, on parle d'effet d'afterpulsing (AP). Le temps mort, représenté par t_{dead}, fait référence à la durée pendant laquelle le détecteur est en armement ou en recharge. Le nombre de charges par pulsation (CPP), qui indique la consommation d'énergie par avalanche, peut être calculée en divisant le courant mesuré pendant le temps d'intégration par le nombre d'avalanches détectées durant cette période.

Modélisation de la probabilité de claquage en avalanche

Dérivation de la probabilité de claquage en avalanche basée par le champ électrique

Les modèles de prévision de la probabilité de claquage en avalanche se concentrent sur la possibilité pour un porteur d'initier avec succès une avalanche par multiplication de porteurs autosoutenue par ionisation par impact. Dans une simple région de déplétion à une dimension de SPAD, comme illustré dans la **Fig. 9.6**, $P_n(x)$ est la probabilité qu'un électron commençant à x dans la couche de déplétion déclenche une avalanche, $P_p(x)$ la même probabilité pour un trou commençant à x et $P_\alpha(x)$ la probabilité qu'un trou ou un électron déclenche une avalanche à x. La probabilité de claquage en avalanche des électrons et des trous peut être représentée comme une fonction des coefficients d'ionisation et des positions.

Les coefficients de taux d'impact par ionisation sont le seul degré de liberté expérimental du modèle de McIntyre. La probabilité de déclencher une avalanche est contrainte à zéro en raison de valeurs de champ électrique inférieures au seuil pour l'ionisation par impact. Pour prédire l'occurrence de la multiplication, la dynamique du champ électrique par rapport à la tension dans les régions n++/p+ et n-/p-well est rapportée dans la **Fig. 9.7 (a)**. Une fois le seuil

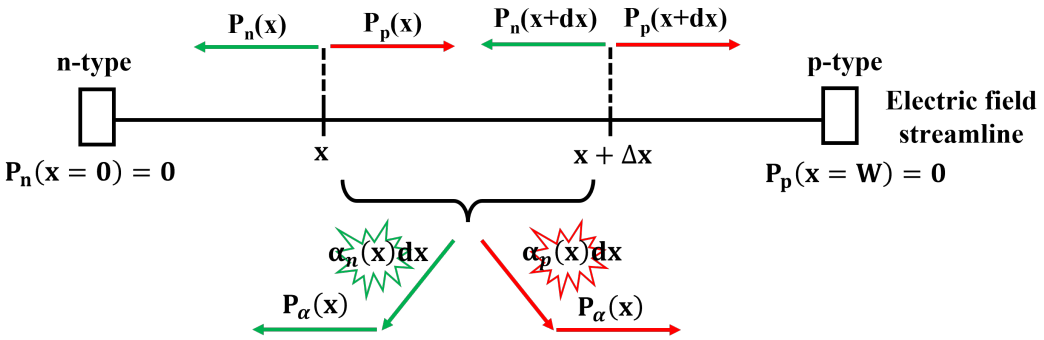


Figure 9.6: Schéma représentant la modélisation de McIntyre de la probabilité de claquage en avalanche pour les électrons (vert) et les trous (rouge) se déplaçant le long d'une ligne de champ électrique. Les porteurs de charge se déplacent soit jusqu'à x, ce qui provoque une avalanche, soit ils provoquent une ionisation d'impact entre x+dx et x. La paire électron-trou résultante provoque alors une avalanche.

de champ électrique pour l'ionisation par impact atteint, des multiplications de porteurs sont attendues dans les régions n++/p+ et n-/p-well.

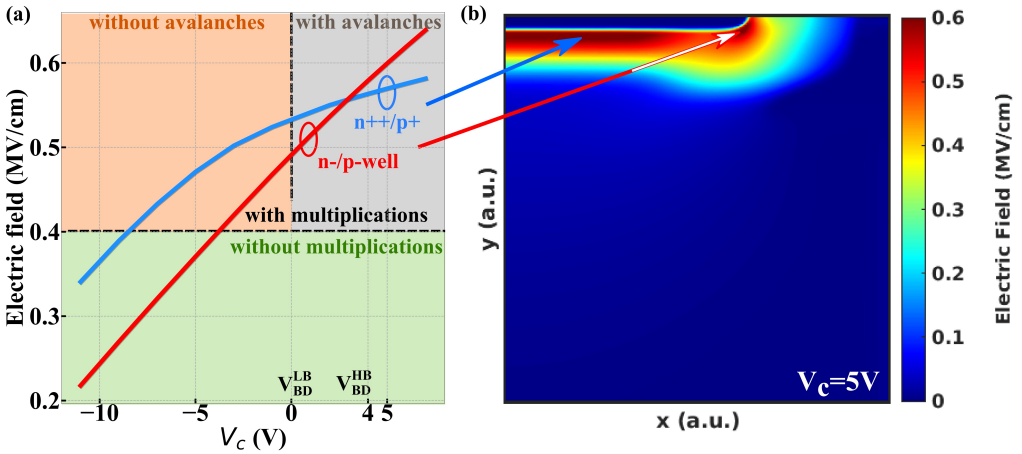


Figure 9.7: (a) Champ électrique en fonction de la tension de caractérisation V_c normalisée par la tension de claquage en avalanche centrale V_{BD}^{LB} dans la région n++/p+ (bleu) et n-/p-well (rouge) simulée à la température de caractérisation T_c=333K. Selon l'amplitude du champ électrique, il y a ou non multiplication soit dans la région n++/p+ soit dans la région n-/p-well (boîte verte et orange) sans avalanches auto-soutenues. Les résultats de simulation de McIntyre présentés dans la figure **Fig. 9.8** indiquent l'apparition d'avalanches dans la région n++/p+ (boîte grise) au-dessus de V_{BD}^{LB}. Les avalanches dans la région n-/p-well sont déclenchées à la tension de claquage en avalanche périphérique V_{BD}^{HB}. (b) Carte de champ électrique simulée à V_c=5V et T_c=333K soulignant l'activité du champ électrique dans les deux régions. Pour plus de clarté, V_{BD} fera référence à V_{BD}^{LB} dans la suite.

La carte simulée de la probabilité de claquage en avalanche est présentée dans la **Fig. 9.8** pour différentes températures à deux tensions d'excès. Différentes régions menant à des avalanches sont activées en fonction de la température, de la polarisation appliquée et du type de porteurs. P_t est ensuite extrait pour les trous en fonction de la tension et de la température dans la

Fig. 9.9 pour mettre en évidence la dynamique thermique et de tension dans les régions centrales n++/p+ et périphériques n-/p-well. La probabilité pour les trous a été sélectionnée en raison de sa proximité plus grande avec l'interface, qui est plus susceptible de présenter des défauts. Deux dynamiques sont observées dans le domaine de tension liées à la force du champ électrique dans les régions n++/p+ et n-/p-well, comme le montre la **Fig. 9.7 (a)**.

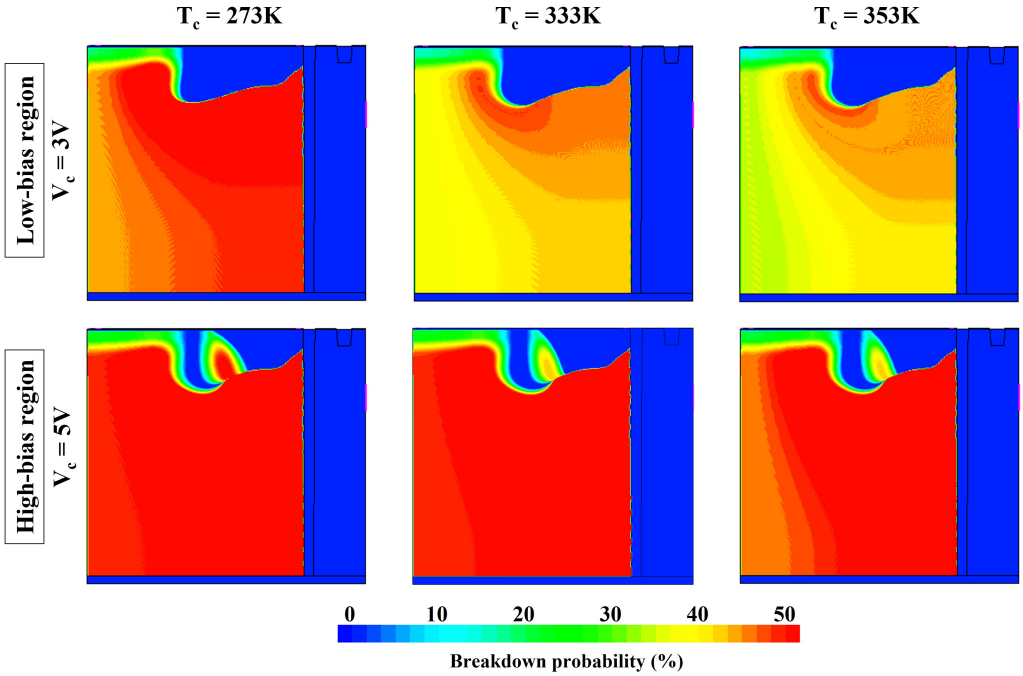


Figure 9.8: Carte simulée de la probabilité totale de claquage en avalanche pour différentes valeurs de T_c à T_c=273K, 333K, 353K et V_c=3V (haut) et 5V (bas). L'extension et l'amplitude de la probabilité de claquage en avalanche diminuent lorsque la température est augmentée.

L'exactitude de la simulation peut être évaluée par rapport aux données expérimentales en utilisant la définition de PDE, qui prend en compte P_t, l'absorption optique (OA) et le facteur de remplissage (FF). La PDE est capturée par le modèle en fonction de la température, comme le montre la **Fig. 9.10 (a)**, et en fonction de la tension, comme le montre la **Fig. 9.10 (b)**. L'activation thermique positive présentée par les coefficients d'ionisation d'impact est responsable de la dépendance thermique positive de PDE par rapport à P_t, plutôt que de la dépendance thermique négative présentée par l'OA.

Un autre avantage de l'utilisation de cette approche de probabilité de claquage en avalanche est la possibilité de déterminer V_{BD} à partir de la courbe PDE en fonction de la tension. Le critère de claquage du dispositif est fixé à un seuil de PDE de 10^{-5} et V_{BD} est identifié comme la tension appliquée minimale à laquelle le seuil est atteint. Pour ce faire, différentes architectures sont comparées par la mesure et la simulation de PDE et V_{BD}, comme illustré dans les **Fig. 9.11 (a)-(b)**.

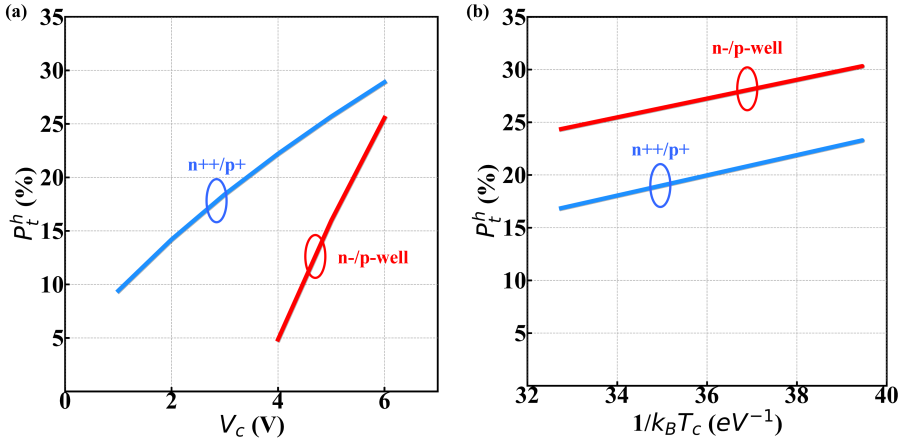


Figure 9.9: (a) Probabilité d'avalanche des trous en fonction de V_c dans les régions n++/p+ (bleu) et n-/p-well (rouge) simulées à T_c=333K. La tendance du champ électrique représentée dans la figure **Fig. 9.7** donne lieu à deux dynamiques en tension qui peuvent être observées dans chaque région. (b) Probabilité d'avalanche des trous en fonction de $1/k_BT_c$ simulée dans la région n++/p+ à V_c=3V et dans la région n-/p-well à V_c=6V. La dépendance de la température des coefficients d'ionisation d'impact est responsable de la dynamique thermique.

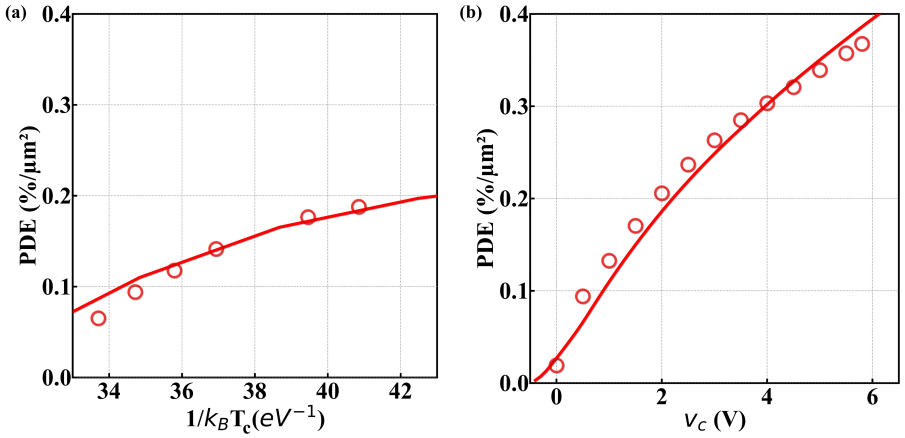


Figure 9.10: PDE mesuré (symboles) et simulé (lignes épaisses) en fonction de (a) $1/k_BT_c$ et (b) V_c pour la diode de référence. Les dépendances thermique et de tension sont conformes aux tendances des coefficients d'ionisation d'impact.

Méthode de Monte-Carlo de dérive-diffusion pour déterminer la probabilité de claquage en avalanche

Une méthode alternative dite de Monte-Carlo fournit une description stochastique des porteurs individuels soumis à la dérive et à la diffusion. Une illustration de la marche aléatoire obtenue est montrée dans **Fig. 9.12 (a)** pour deux points de départ différents et pour différentes simulations. Même lorsque les particules sont initialement placées au même point de départ, elles peuvent présenter des temps de diffusion, des trajectoires et des points d'entrée dans les jonctions distincts, comme illustré dans **Fig. 9.12 (b)**. De plus, les particules peuvent surmonter les barrières de potentiel dans le dispositif en raison de leur énergie thermique. Cela peut conduire

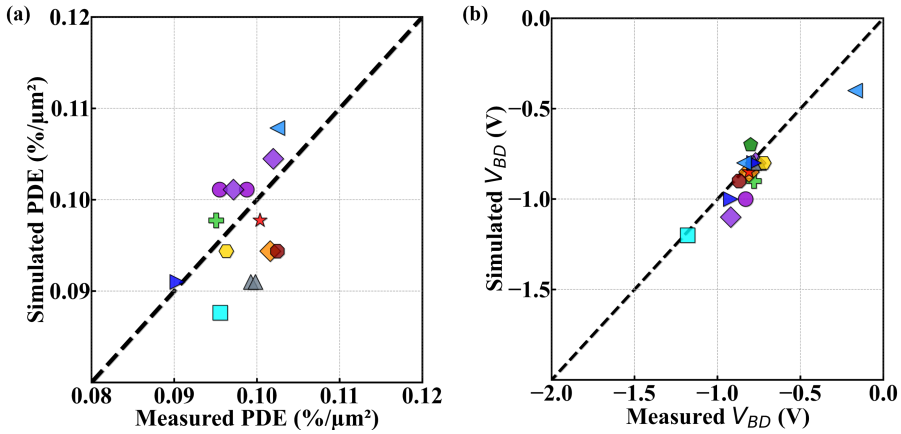


Figure 9.11: Graphiques de corrélation entre les expériences (axe des x) et les simulations (axe des y) pour (a) PDE et (b) V_{BD} ($= V_{BD}^{LB}$) par rapport à une tension de référence à T_c=333K. Chaque symbole correspond à une architecture de dispositif différente. La ligne en pointillé est la ligne y=x, mettant en évidence la corrélation atteinte entre les expériences et les simulations.

à de nouvelles zones actives au claquage en avalanche dans le dispositif, ce qui peut avoir un impact significatif sur les performances du dispositif,tels que de la probabilité de claquage en avalanche et du jitter.

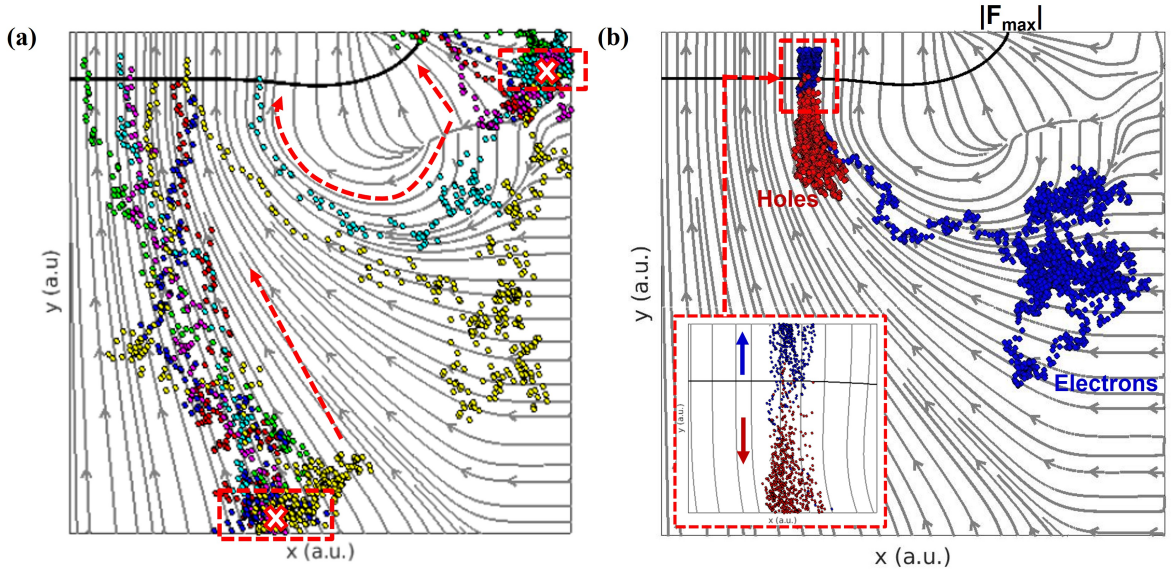


Figure 9.12: (a) Marche aléatoire pour les particules commençant à l'une des deux croix à l'intérieur des boîtes. Chaque couleur représente un chemin de porteur potentiel. Malgré le fait d'avoir le même point de départ, les particules peuvent présenter des variations dans leurs temps de diffusion, leurs itinéraires et leurs points d'entrée dans les jonctions. (b) Résultats de simulation Monte-Carlo de dérive-diffusion pour un seul électron généré à l'interface du dispositif. Les électrons sont représentés en bleu et les trous en rouge. L'encart représente la génération de paires électron-trou dans la région de multiplication. Chaque point correspond à un pas de temps différent. Les lignes grises représentent les lignes de potentiel électrostatique.

La **Fig. 9.13** présente une comparaison des probabilités de claquage en avalanche obtenues

à partir du modèle de McIntyre et de la méthode Monte-Carlo. Les résultats montrent une différence significative dans la région proche du contact, où la méthode Monte-Carlo prédit une probabilité de claquage allant jusqu'à 5%, tandis que le modèle McIntyre prédit une probabilité nulle en raison d'une barrière de potentiel électrostatique. Un électron, venant de la gauche, monte le potentiel. Lorsque le potentiel diminue, il fait face à un champ électrique positif qui le pousse vers la gauche. La barrière coupe la direction des lignes de champ électrique, ce qui empêche les trajectoires de porteurs de suivre les lignes de champ électrique dans le modèle McIntyre, conduisant à l'incapacité de calculer P_t. Cependant, la méthode Monte-Carlo simule le mouvement aléatoire, ce qui permet à la diffusion de porteurs de surmonter la barrière électrostatique et de fournir de nouveaux chemins possibles pour les porteurs. Ces chemins comprennent (A) la diffusion et la collecte vers les contacts sans déclencher une avalanche, conduisant à une réduction de P_t, ou (B) atteindre une région active en avalanche conduisant à une augmentation de P_t. Par conséquent, l'approche Monte-Carlo offre des choix de chemins supplémentaires pour les porteurs générés à l'intérieur du dispositif, ce qui contraste avec la trajectoire de porteur contrainte par les lignes de champ électrique dans le modèle McIntyre.

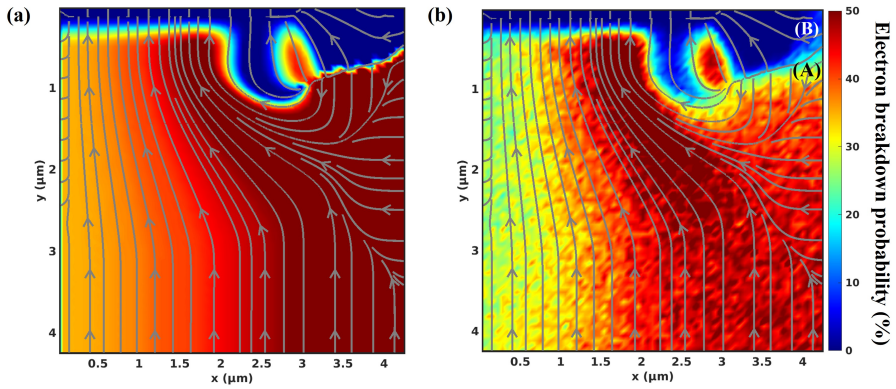


Figure 9.13: Probabilité de rupture des électrons avec (a) le modèle McIntyre comparé à (b) la méthode Monte-Carlo de dérive-diffusion simulée à V_c=5V et T_c=333K. (A) montre une diminution de P_t due à la diffusion de porteurs vers les contacts et (B) une augmentation de P_t en raison de la diffusion de porteurs surmontant la barrière de potentiel électrostatique. Les lignes grises superposées représentent les lignes de potentiel électrostatique.

Taux de génération simulée à partir de la durée de vie limitée par les défauts

Comme illustré dans la **Fig. 9.14**, l'origine du DCR remonte à travers les différents mécanismes R-G suivants:

- Shockley-Read-Hall (SRH): un électron dans la bande de valence peut absorber une vibration du réseau cristallin appelée phonon pour atteindre un état de piège avant d'émettre dans la bande de conduction par interaction avec un autre phonon.

- Poole-Frenkel effect (PF): L'énergie potentielle de Coulomb entre les bandes et les pièges peut être réduite par un champ électrique appliqué, ce qui augmente la probabilité d'être thermiquement excité d'un site de piège à un autre.
- Trap-Assisted Tunneling (TAT): Tout d'abord, l'électron est émis de la bande de valence vers l'état de piège en absorbant un phonon. L'électron peut être partiellement soulevé et tunneliser dans la bande de conduction de manière élastique ou inélastique.
- Band-to-Band Tunneling (BTBT): Un électron dans la bande de valence du semi-conducteur tunnelise à travers la bande interdite vers la bande de conduction sans l'aide de pièges. La bande interdite agit comme une barrière de potentiel que la particule traverse par effet tunnel. Le chemin de tunnelisation le plus probable est à travers la plus petite barrière.

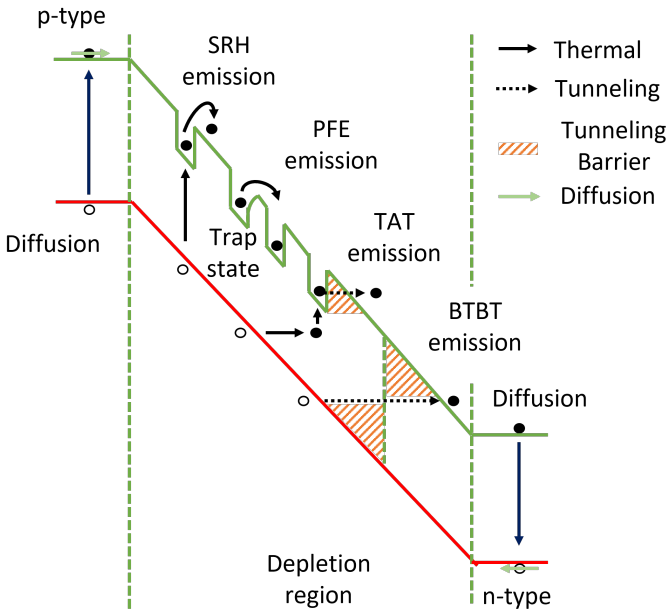


Figure 9.14: Représentation des différents mécanismes de recombinaison-génération (R-G) résultant en DCR lorsqu'ils sont couplés à la probabilité de rupture avalanche P_t

La simulation prend en compte l'état de charge du piège, y compris la localisation des pièges et la définition explicite des pièges par le paramètre de durée de vie des porteurs minoritaires. Cette durée de vie est modélisé comme un produit de la concentration de pièges N_t multiplié par la section efficace de capture $\sigma_{n,p}$ et la vitesse thermique des porteurs v_{th} pour le silicium. Une carte typique de taux de génération simulée à partir du modèle de durée de vie limitée par les défauts SRH est présentée dans la **Fig. 9.15**. La distribution des défauts a été aléatoirisée dans l'espace. La randomisation diffère à chaque fois que le fichier de commande est exécuté. Le nombre spécifié est utilisé comme graine pour le générateur de nombres aléatoires. Comme mentionné précédemment, la définition des pièges permet une investigation dans tout le dispositif, révélant l'impact des défauts d'interface sur le taux de génération de porteurs.

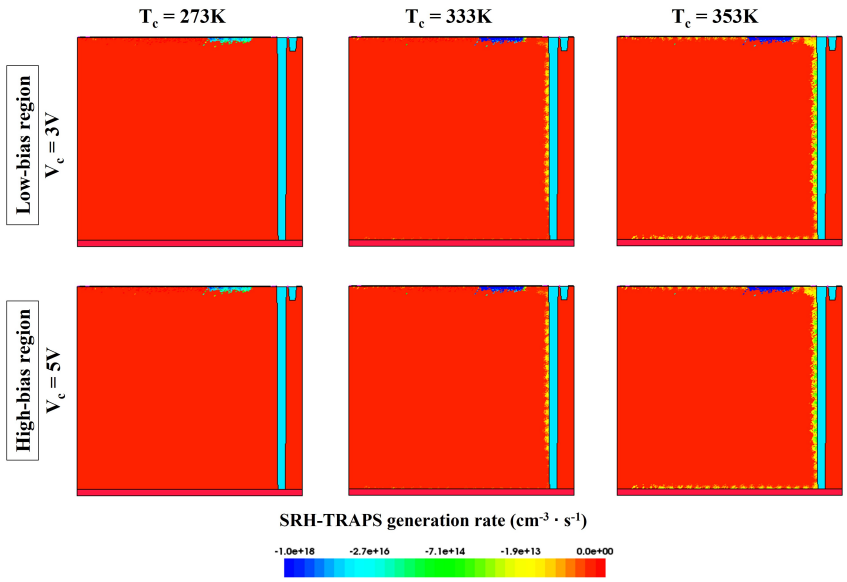


Figure 9.15: Carte du taux de génération de porteurs de charge de SRH limité par les défauts simulée à T_c=273K, 333K et 353K, et V_c=3V et 5V. À mesure que la température augmente, le taux de génération de porteurs de charge diminue.

Étude du DCR sur dispositifs non soumis à la contrainte

Utilisation de la modélisation basée sur la moyenne pour étudier DCR

Les modélisations de la PDE et du DCR sont basée sur la probabilité de claquage en avalanche, mais diffère par les propriétés de l'absorption otpique et des défauts. L'intégration de la multiplication de la probabilité de claquage en avalanche par les processus de génération susmentionnés sur chaque maillage du dispositif permet de calculer un DCR dépendant de la position. Pour la comparaison expérimentale, l'intégration sur l'ensemble du dispositif est effectuée, ce qui donne le DCR moyen en comptes par seconde (cps). De plus, le DCR expérimental et simulé est normalisé par unité de surface photosensible (cps/µm^2). Cette configuration nous permet d'étudier de manière flexible les mécanismes de génération de porteurs qui mènent à l'avalanche à l'intérieur du volume du SPAD ou aux interfaces. Une carte typique du DCR simulée à partir du modèle SRH de durée de vie limitée par défauts est présentée dans la **Fig. 9.16**. Selon la localisation des défauts, la présence de deux points actifs en DCR à l'interface peut être attribuée à la dépendance en tension de la probabilité de claquage par avalanche, comme illustré dans la **Fig. 9.8**. L'encart de l'interface supérieure dans le coin inférieur droit de la figure permet d'observer l'activation des défauts.

Le DCR est simulé en fonction de la tension en considérant des défauts avec un seul niveau d'énergie au milieu de la bande interdite. $\sigma_{n,p}$ est fixé à $10^{-15} cm^2$ et N_t est choisi comme degré de liberté et paramètre représentatif des défauts induits par le procédé (voir légendes dans la figure **Fig. 9.17 (b)**). Lorsque l'état du piège est défini comme neutre, l'utilisation de types d'accepteurs ou de donneurs conduit à des prévisions équivalentes. Cela implique qu'à la densité considérée N_t, la présence de pièges à l'intérieur du dispositif n'altère pas significativement

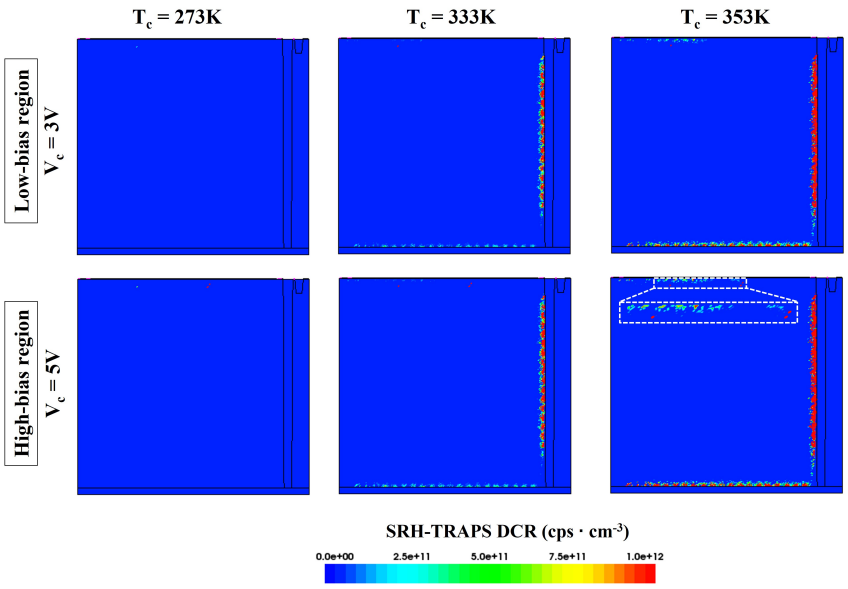


Figure 9.16: Carte de DCR généré par le modèle de SRH de durée de vie limitée par défauts simulée à T_c=273K, 333K et 353K, et V_c=3V et 5V. Le DCR augmente lorsque la température est augmentée, activant à la fois les régions n++/p+ et n-/p-well, comme le montre l'encart à haute température.

le profil électrostatique du dispositif. Les défauts situés à l'intérieur de la zone de l'anneau périphérique et/ou à l'interface supérieure dans cette zone créent l'effet correct requis pour modéliser la région de forte polarisation telle que représenté dans la **Fig. 9.17 (a)**. Avec toutes les contributions incluses, la tendance globale du DCR est correctement reproduite proche des données mesurées avec le bon point d'inflexion lorsqu'elle est tracée en fonction de la tension. De plus, comme représenté dans la **Fig. 9.17 (b)**, la dynamique thermique est reproduite avec précision.

Approche statistique pour modéliser le DCR.

La distribution de défauts définie constante permet de reconstruire les tendances médianes du DCR. Cependant, cette valeur moyenne ne capture pas la population intrinsèque (du 1er au 3ème quartile) et la population extrinsèque (outliers) représentées dans la **Fig. 9.18**. L'étude des fluctuations statistiques locales induites par le procédé de fabrication sera réalisée en utilisant l'approche de Monte-Carlo de dérive-diffusion, ainsi que le taux de génération de BTBT couplé au taux de recombinaison non-radiative assistée par multiphonon (NRM) à partir des défauts, plutôt que dans le cadre du SRH. Les défauts sont aléatoirement positionnés et de tailles variables, allant de défauts uniques à de grands amas allant jusqu'à cent défauts, tous situés dans les zones d'intérêt (RoI) précédemment identifiée dans le dispositif SPAD. D'une part, pour la population intrinsèque à l'intérieur de la région de faible polarisation, des amas de défauts disposés au hasard le long de l'interface latérale ainsi qu'à l'intérieur de la région de contact et de la région de multiplication captent cette dispersion statistique, comme illustré dans **Fig. 9.18 (a)**. Pour la région de forte polarisation, seuls les défauts disposés au hasard au bord des puits et des

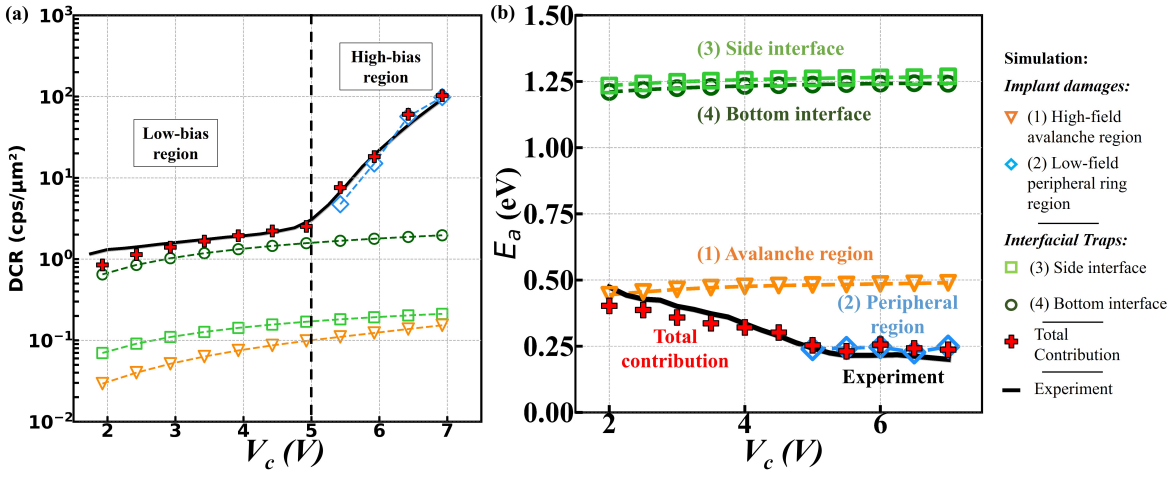


Figure 9.17: (a) DCR en fonction de V_c mesuré (ligne épaisse) et simulé (ligne en pointillé) à T_c=333K avec le modèle SRH-TAT selon l'approche de description de piège pour les différentes contributions de DCR évaluées pour chaque région. De plus, BTBT et la diffusion des porteurs minoritaires ont également été superposées. La concentration de pièges interfaciaux était de $10^9 cm^{-2}$ aux interfaces supérieure, latérale et inférieure. La concentration de pièges était de $10^8 cm^{-3}$ à l'intérieur de la zone centrale et de $10^{11} cm^{-3}$ à l'intérieur des régions périphériques. (b) L'énergie d'activation thermique du DCR E_a en fonction de V_c extraite à 253K, 333K et 353K. L'énergie d'activation thermique est équivalente à la pente de la ligne suivant la loi d'Arrhenius, ln(DCR) vs $1/k_B T_c$. Les lignes pleines représentent les mesures et les lignes en pointillé représentent les données de simulation.

anneaux, où la concentration de défauts est supposée être élevée compte tenu des multiples étapes d'implantation, reproduisent cette tendance, comme illustré dans **Fig. 9.18 (b)**. D'autre part, la population extrinsèque est reproduite avec des défauts disposés au hasard à l'intérieur des puits et des anneaux. La survenue de dommages d'implantation lors de l'étape de traitement d'implantation ionique ultérieure est suspectée [90].

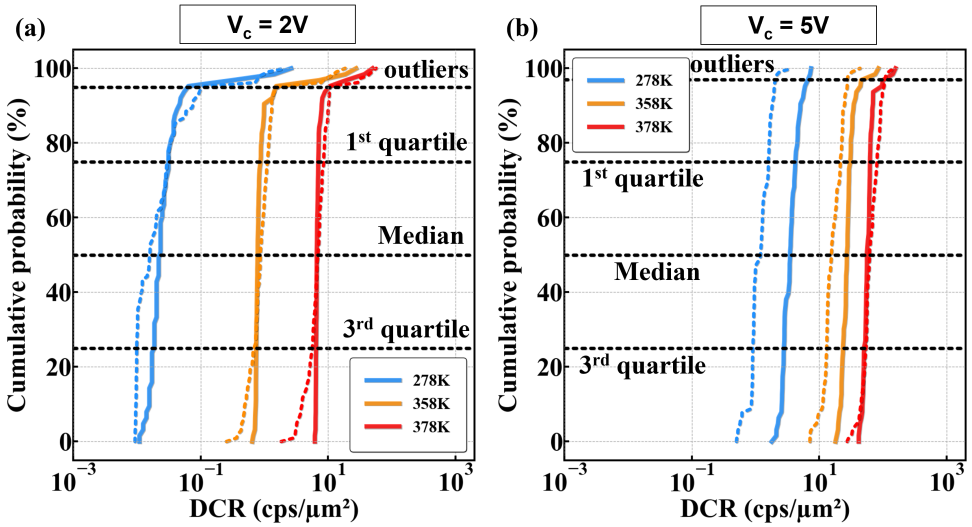


Figure 9.18: Probabilité cumulative en fonction du DCR mesurée et simulée à (a) V_c=2V et (b) V_c=5V. Les simulations sont en pointillé et l'expérience en ligne. Les données mesurées et simulées présentent toutes deux un schéma similaire, avec deux groupes distincts : la population intrinsèque, qui comprend les points de données du 1er au 3ème quartile, et la population extrinsèque, qui contient les valeurs extrinsèques.

Étude du DCR après contrainte

Répartition de l'impact de la contrainte

Fig. 9.19 (a) montre la relation entre le ΔDCR et $1/V_s$, qui présente deux schémas exponentiels: un plus doux à faible polarisation (LB) et un plus net à forte polarisation (HB). Ces dynamiques de tension sont cohérentes avec le DCR de précontrainte observée précédemment, où la multiplication de porteurs se produit dans la région centrale n++ et p+ pour LB, et dans la région périphérique n- et p-well pour HB. Cette dépendance exponentielle de la tension indique une possible dégradation par porteurs chauds (HCD) par ionisation d'impact de paires électron-trou. À un niveau d'irradiance élevé ($1W/m^2$), où le courant provient principalement de la photo-génération, et bien en dessous de V_{BD} (où peu de multiplications de porteurs se produisent), le ΔDCR est largement influencé par le grand volume de porteurs photo-générés, et est donc soumis à une dégradation constante, comme indiqué par les triangles dans **Fig. 9.19 (a)**. Avec (V_s=4V et 5V) ou sans (V_s=-3V) déclenchement d'avalanche, la relation linéaire entre ΔDCR et l'irradiance est maintenue. À partir de ces résultats, il est supposé que les mêmes mécanismes de vieillissement sont liés au nombre de porteurs chauds disponibles, indépendamment du déclenchement d'avalanches. Ces dépendances de dégradation sur le nombre de porteurs avec et sans avalanches seront expliquées dans les paragraphes suivants sur la base de la mesure et de la modélisation du courant électrique.

Approche statistique pour modéliser la dérive du DCR

L'objectif de la simulation à venir est de reproduire la relation entre la variance de la variation de ΔDCR d'un dispositif à l'autre $\sigma^2_{\Delta DCR}$ et sa moyenne $\overline{\Delta DCR}$ dans les SPAD, en essayant

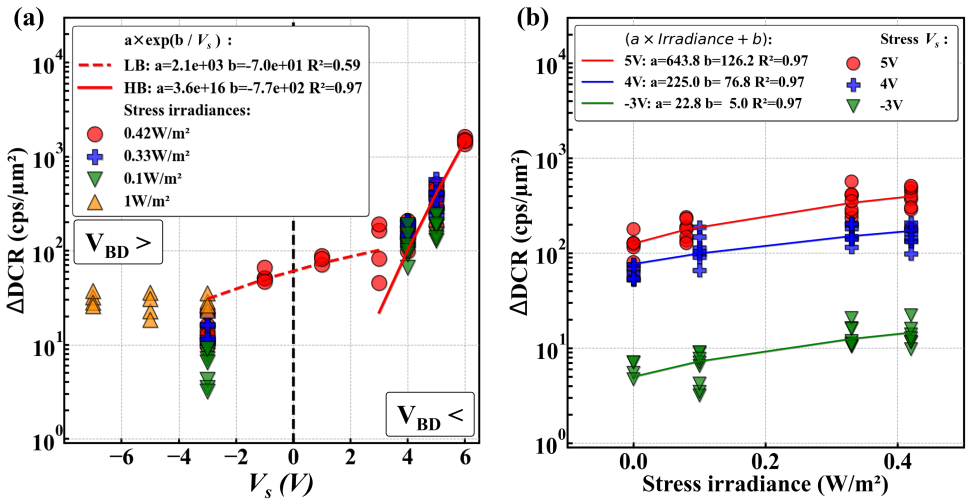


Figure 9.19: (a) ΔDCR mesuré à T_c=333K et V_c=5V en fonction de la tension de contrainte V_s en dessous et au-dessus de V_{BD} pour différentes irradiances de contrainte. (b) ΔDCR en fonction des irradiances de contrainte pour différentes tensions de contrainte V_s. Les dépendances exponentielle de la tension et linéaire de l'irradiance de la dégradation sont liées à la multiplication et à la quantité de porteurs, qui sont associées à la dégradation par porteurs chauds (HCD).

d'utiliser l'approche employée pour la variabilité de la variation de ΔV_{th} d'un dispositif à l'autre dans les MOSFET [144][145][146][147]. Les mesures sur un grand nombre de dispositifs avec un ensemble complet de conditions de contrainte et de processus couplées à une méthodologie de simulation de Monte-Carlo mettent en lumière l'origine potentielle de la relation entre la variance de la variation de ΔDCR d'un dispositif à l'autre $\sigma^2_{\Delta DCR}$ et sa moyenne $\overline{\Delta DCR}$ dans les SPADs. Pour adresser la variabilité de ΔDCR, un réseau de 32x32 SPAD est soumis à une contrainte à T_s=353K, V_s=3V et 0,02W/m² à 940nm pendant 67 heures. Une relation quadratique est trouvée lorsque la variance d'un dispositif à l'autre est tracée en fonction de ΔDCR, comme illustré par **Fig. 9.20 (b)**. La variance ΔDCR d'un dispositif à l'autre, $\sigma^2_{\Delta DCR}$, est étudiée en tirant uniformément les positions des défauts [144][145][146][147] dans une boîte d'un nanomètre d'épaisseur à l'interface supérieure appauvrie en porteurs. La concentration de défauts entre chaque dispositif est tirée d'une distribution de Poisson [151] avec un nombre moyen de défauts défini de dix à un million par multiple de dix (N_0 à N_5). La variance simulée d'un dispositif à l'autre en fonction du DCR reproduit la loi quadratique expérimentale comme reporté dans reportée dans la **Fig. 9.20 (b)** .

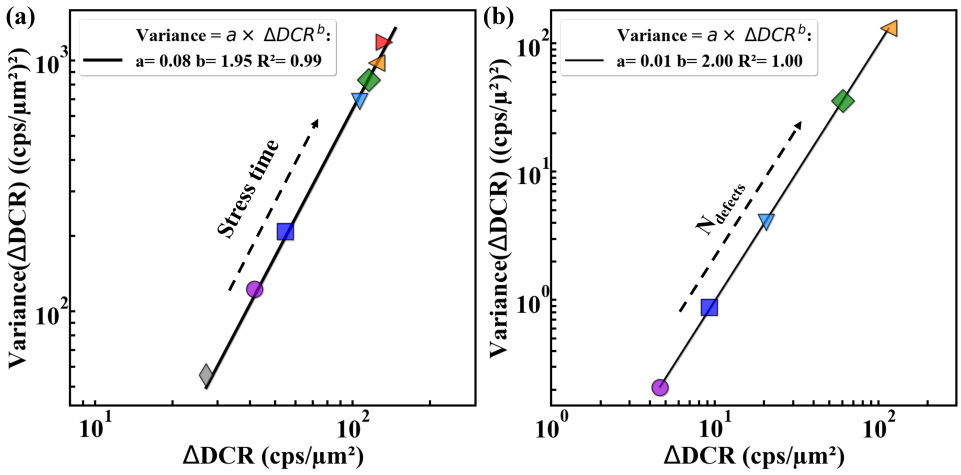


Figure 9.20: (a) Variance de ΔDCR d'un dispositif à l'autre $\sigma^2_{\Delta DCR}$ mesurée en fonction de sa moyenne $\overline{\Delta DCR}$ pour différentes durées de contrainte (symboles). μ représente la moyenne de la distribution normale et σ l'écart-type. Une relation quadratique est trouvée entre $\sigma^2_{\Delta DCR}$ et $\overline{\Delta DCR}$. (a) Variance de ΔDCR d'un dispositif à l'autre $\sigma^2_{\Delta DCR}$ simulée en fonction de sa moyenne $\overline{\Delta DCR}$ pour un nombre croissant de défauts. $N_{defects}$ représente le nombre moyen de défauts de la distribution de Poisson définie de dix à un million par multiple de dix (N_0 à N_5). Une loi quadratique entre la variance de la variation de ΔDCR d'un dispositif à l'autre $\sigma^2_{\Delta DCR}$ et sa moyenne $\overline{\Delta DCR}$ est simulée.

Modélisation de la dérive du DCR

La dérive du DCR (ΔDCR) a été étudiée à une durée de contrainte fixe pour extraire les dépendances des conditions de contrainte (à savoir la température, l'irradiance et la tension) et identifier les mécanismes de dégradation dans l'architecture SPAD étudiée. Maintenant l'objectif est d'expliquer la dépendance temporelle du ΔDCR pour refléter la physique du processus de formation de défauts d'interface et prédire la fiabilité du dispositif sur la base de la littérature existante sur la dégradation par porteurs chauds (HCD) dans les MOSFETs.

Simulation de la distribution d'énergie des porteurs

Une méthode de Monte-Carlo avec structure de bande complète (FBMC) [180][181] a été utilisée pour démontrer le processus de dégradation par porteurs chauds uniques dans les SPADs en calculant la distribution en énergie des porteurs $n(E)$. Cette approche repose sur les structures de bande de silicium calculées avec la méthode de Tight-Binding par LCAO [182]. Les taux de diffusion par phonons, par impuretés ionisées et par porteurs sont calculés de manière cohérente avec la structure de bande complète. L'interface supérieure a été précédemment identifiée comme étant la zone de dégradation dans l'architecture SPAD étudiée, par conséquent une attention particulière est portée sur la distribution en énergie des porteurs dans sa proximité. Les électrons se déplacent le long de l'interface supérieure vers la cathode (non montrée ici), tandis que les trous se déplacent dans la direction opposée. $n(E)$ est extrait à l'interface supérieure soit dans la région n++/p+ soit dans la région n-/p-well et normalisé par la distribution totale de porteurs obtenue dans la simulation pour comparer les populations de porteurs chauds entre chaque

condition de contrainte. Comme le montrent les **Fig. 9.21 (a)-(b)**, certains électrons de la queue de distribution d'énergie atteignent l'énergie de dissociation seuil moyenne de la liaison $Si-H$ (E_{th}=1,5eV [143][175]) en fonction de la tension de contrainte V_s (**Fig. 9.21 (c)-(d)**) et de la température de contrainte T_s (**Fig. 9.21 (e)-(f)**). La dépendance négative de l'énergie des porteurs en température est une conséquence du libre parcours moyen des porteurs-phonons qui diminue avec l'augmentation de la température du réseau cristallin, ce qui entraîne un nombre de collisions plus importants à même champ électrique et donc une perte d'énergie dû à à ces collisions plus fréquentes. Les mêmes dépendances thermiques et de tension ont été observées sur l'ensemble de la population. Il y a deux dynamiques de dégradation potentielles selon la localisation soit dans la région n++/p+ soit dans la région n-/p-well. Ces deux régions de dégradation différentes peuvent être stimulées indépendamment en fonction des conditions de caractérisation, comme illustré par **Fig. 9.7**. Les distributions d'énergie de porteurs simulées $n(E)$ le long de V_s et T_s sont intégrées sur l'énergie de dissociation de liaison $Si-H$ pour prendre en compte la probabilité de dissociation de liaison par interaction de porteurs chauds à chaque énergie de porteur.

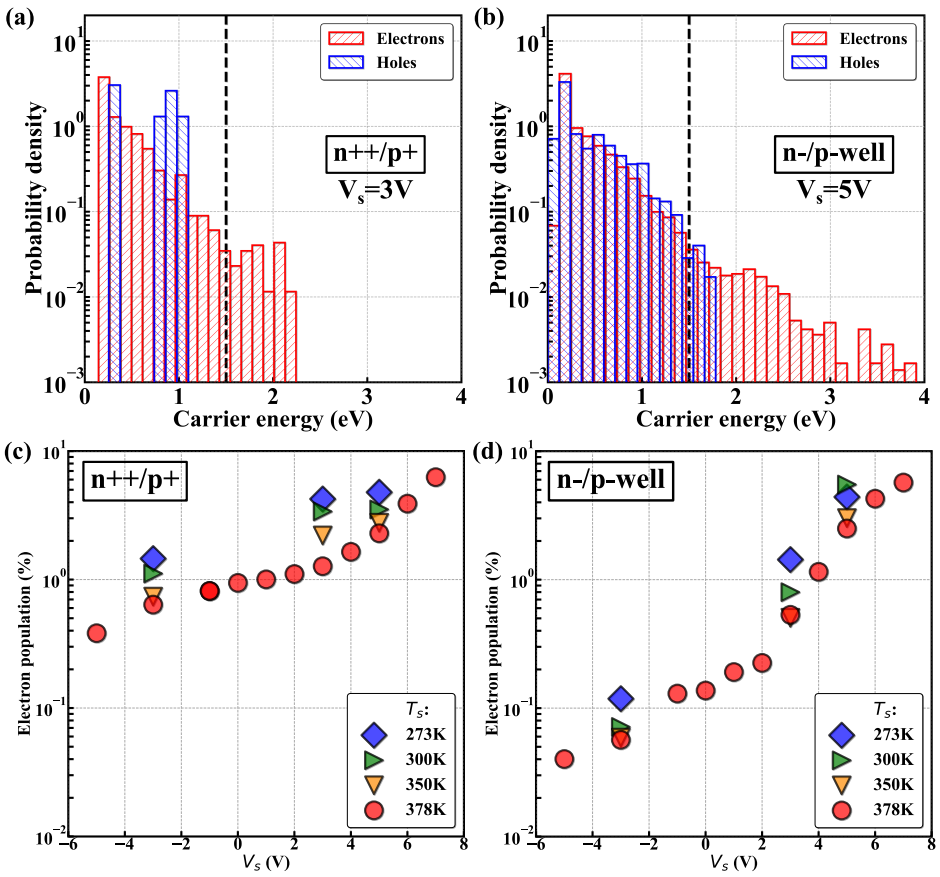


Figure 9.21: Densité de probabilité d'énergie des électrons (rouge) et des trous (bleu) simulées par un Monte-Carlo avec structure de bande complète [180][181] à T_s=333K pour (a) V_s=3V (LB) et (b) V_s=5V (HB). Les porteurs acquérant une énergie supérieure au seuil moyen de dissociation de la liaison $Si-H$ (1,5 eV) sont extraits et rapportés en fonction de V_s à T_s=333K dans (c) et (d).

Dégradation induite par le courant sans avalanches

Notre étude précédente a montré que la dégradation du DCR pouvait se produire même sans avalanches. Les avalanches peuvent entraîner de grandes quantités de porteurs chauds, mais ils ne peuvent pas circuler près de l'interface en raison de la chute de potentiel électrostatique [88]. Ces porteurs chauds générés pendant les avalanches ne peuvent pas atteindre et causer de la dégradation à l'interface supérieure. Le courant d'obscurité et photo-généré, qui ne déclenche pas d'avalanches, ne peut pas être distingué du courant d'avalanche lorsque la diode est polarisée au-dessus de V_{BD}. Cependant, le courant au-dessus de V_{BD} sans avalanches peut être supposé proportionnel au courant mesuré en dessous de V_{BD}, dans l'obscurité ou sous éclairement, en tenant compte du taux de multiplication. La dérive de DCR après contrainte est donc comparée au courant mesuré en dessous de V_{BD} avant contrainte et rapportée dans **Fig. 9.22** pour les deux régions de caractérisation LB et HB. Avec (V_s=4V et 5V) ou sans (V_s=-3V) déclenchement d'avalanches, ΔDCR par rapport au courant mesuré sans avalanches suit une relation linéaire avec l'irradiance. Les mêmes mécanismes de vieillissement peuvent être supposés liés à l'augmentation du nombre de porteurs chauds qui augmentent la dissociation des liaisons $Si-H$. Selon ces résultats, la densité de courant sans avalanches sera extrapolée au-dessus de V_{BD} pour calculer le taux de dégradation en fonction du nombre de porteurs chauds disponibles.

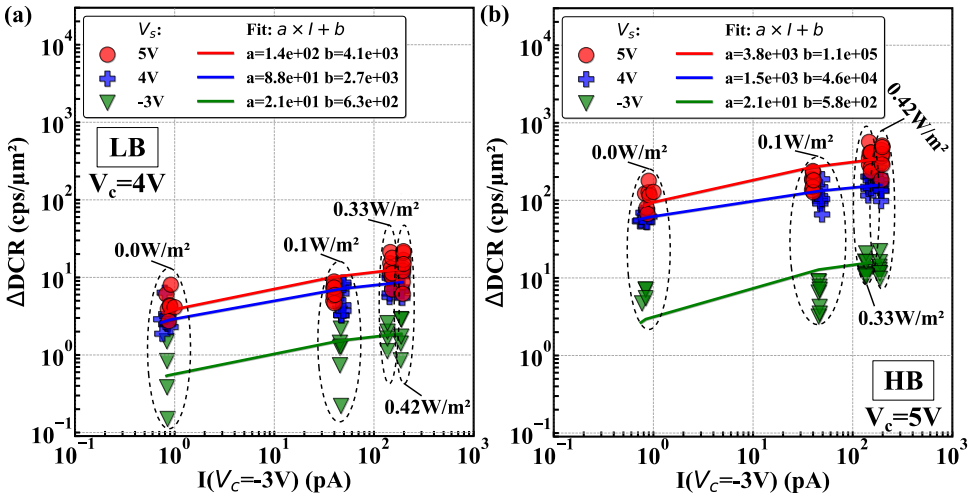


Figure 9.22: ΔDCR en fonction du courant électrique I (sans avalanches) mesuré (symboles) et ajusté (lignes) par une relation linéaire à V_c=-3V et à T_c=353K pour plusieurs niveaux d'irradiance de 0,0 W/m² (gauche) à 0,42 W/m² (droite). ΔDCR a été obtenu à partir de contraintes de 24 heures à différents V_s et à T_s=353K pour les mêmes niveaux d'irradiance et a été mesuré à T_c=333K et à (a) V_c=4V (LB) et (b) 5V (HB).

Formulation d'un modèle pour les composantes du courant électrique

Après avoir simulé la distribution d'énergie des porteurs pour un nombre donné de paires électron-trou, il est nécessaire de connaître la densité totale de porteurs circulant dans le dispositif pour modéliser le taux de dégradation en fonction de la quantité de porteurs chauds. Comme expliqué précédemment, la densité de courant au-dessus de V_{BD} est modélisée en extrapolant la

densité de courant en dessous de V_{BD} pour se débarrasser du courant d'avalanche. Le courant de multiplication est égal au taux de généraion par impact-ionisation multiplié par q et par la zone sensible à la multiplication qui est définie comme la région de charge d'espace S_{SCR}. La contribution du courant de multiplication est couplée avec les courants générés dans le noir et par la lumière, en considérant le champ électrique simulé F (voir **Fig. 9.7**). Le modèle pour le courant d'oscurité implique l'utilisation du taux de recombinaison-génération des porteurs de Shockley-Read-Hall (SRH) [186]. Le taux de génération, G_d, peut être simplifié en supposant le pire scénario, où le profil de défauts pour la capture d'électrons et de trous est le même, et où les défauts sont amphotères [187]. La densité unique de défauts d'interface initiale, N_t^0, est définie à un niveau d'énergie unique au milieu de la bande interdite E_t. La modélisation du courant d'obscurité reproduit la tendance expérimentale, comme illustré en fonction de la tension dans la **Fig. 9.23 (a)** et en fonction de la température dans la **Fig. 9.23 (b)**.

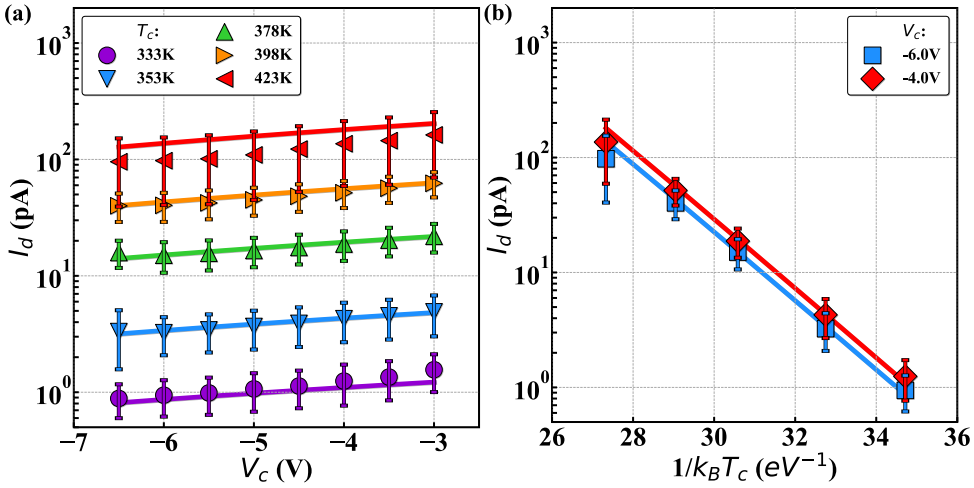


Figure 9.23: (a) Courant d'obscurité en fonction de V_c en dessous de V_{BD} mesuré (symboles) et modélisé (lignes) à différentes températures T_c. (b) Courant en fonction de $1/k_BT_c$ mesuré (symboles) et modélisé (lignes) à différentes tensions V_c. Les barres d'erreur représentent l'écart-type entre chaque dispositif. La dépendance thermique du courant dans l'obscurité peut être expliquée par le taux R-G SRH des porteurs, tandis que la dépendance de la tension est attribuée à la multiplication causée par l'ionisation d'impact des porteurs.

Le taux de photo-génération est modélisé selon la loi de Beer-Lambert [188] en supposant que tous les porteurs photo-générés sont collectés dans la région de déplétion. La modélisation de la photo-génération montre une bonne concordance avec l'expérience, comme illustré en fonction de la tension dans la **Fig. 9.24 (a)** et en fonction de la température dans la **Fig. 9.24 (b)**. Une différence dans l'amplitude du courant est observée au-dessus de V_c=-5V malgré un niveau d'irradiance identique, ce qui est lié à la multiplication des porteurs dans la région de déplétion, comme prévu par les simulations de champ électrique (ligne rouge et boîte orange dans la **Fig. 9.7**a).

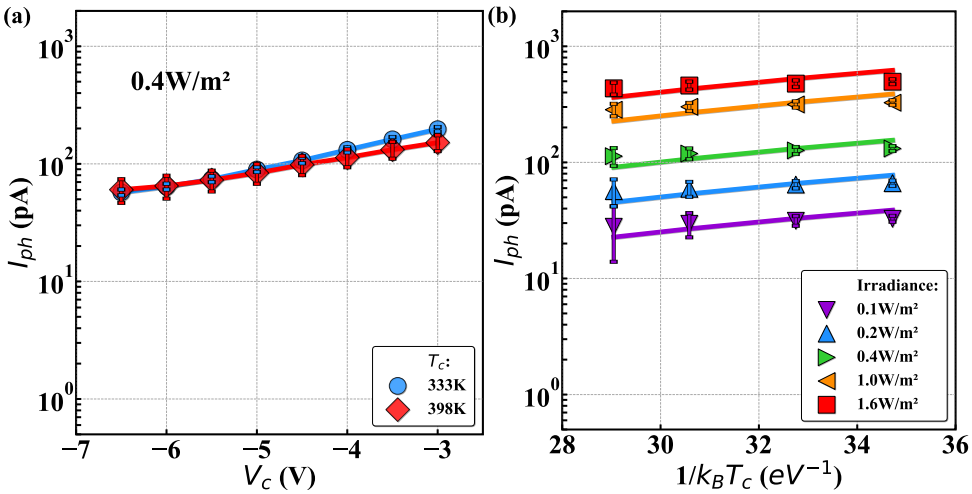


Figure 9.24: (a) Courant photo-généré en fonction de V_s en dessous de V_{BD} (sans avalanches) mesuré (symboles) et modélisé (lignes) à différentes températures T_s et à 0,4W/m². (b) Courant en fonction de $1/k_BT_s$ mesuré (symboles) et modélisé (lignes) à V_s=-4V pour plusieurs niveaux d'irradiance. Les barres d'erreur représentent l'écart-type entre chaque dispositif.

Cinétique de dégradation sélectionnée le long de la contrainte pour la création de défauts

La méthodologie de modélisation est résumée dans **Fig. 9.25**. L'effet de caractérisation est interprété à partir de la probabilité de rupture d'avalanche P_t et du taux de génération de porteurs à partir de défauts créés pendant la durée de la contrainte divisé par le nombre de défauts créés. P_t est simulé le long de V_c et T_c (voir **Fig. 9.9**). La concentration intrinsèque n_i et la vitesse thermique v_{th} sont calculées en fonction de T_c. La section efficace de capture σ_t est calculée en fonction de V_c en fonction du champ électrique simulé F (**Fig. 9.7 (a)**). n_i, v_{th} et σ_t sont injectés dans G_d pour calculer le taux de génération de porteurs à partir de défauts induits par la contrainte.

L'impact de la contrainte est capturé par le nombre de défauts N_t(t) créés pendant la durée de la contrainte. La constante de taux de dissociation de liaison dépendante du temps $\kappa(t)$ est ensuite calculée. La distribution d'énergie des porteurs $n(E)$ est simulée par Monte-Carlo avec structure de bande complète le long de la plage de V_s et T_s. Les densités de courant générées dans l'obscurité et par la lumière sont modélisées et respectivement couplées au courant de multiplication en considérant le champ électrique simulé F (**Fig. 9.7 (a)**). Le taux de dégradation est ajusté en fonction du courant dans l'obscurité modélisé. La distribution d'énergie de dissociation de liaison $Si-H$ $\tilde{g}(E_{hb}, t)$ est calculée à chaque itération pour prendre en compte l'épuisement des précurseurs de défauts par les défauts de liaison pendante.

Modélisation de la dérive du DCR

Cette méthodologie est maintenant confrontée aux résultats expérimentaux. Le modèle de dégradation est d'abord comparé à diverses tensions et températures de contrainte pour différentes conditions de caractérisation. Pour les contraintes en obscurité et sous illumination, une bonne

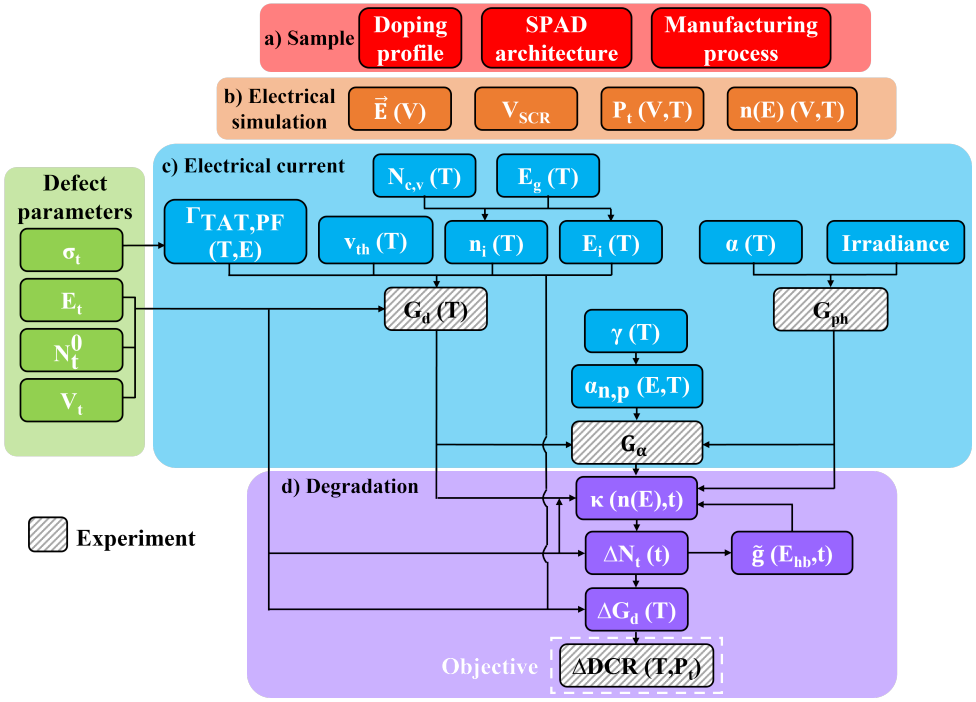


Figure 9.25: Méthodologie proposée pour prédire ΔDCR basée sur : (a) l'architecture SPAD et le processus de simulation, (b) le champ électrique F, le volume de la région de charge d'espace V_{SCR}, la probabilité de claquage en avalanche P_t et la distribution d'énergie des porteurs $n(E)$ intégrée avec (c) les taux de génération dans l'oscurité, sous éclairement et de multiplication $G_{d,ph\alpha}$ pour calculer (d) la constante de vitesse de dissociation de liaison κ. Voir la description des paramètres dans le texte. Seules les dépendances de premier ordre sont indiquées entre parenthèses, telles que la constante de vitesse de dissociation de liaison κ, qui dépend de la distribution d'énergie des porteurs $n(E)$ (premier ordre), qui dépend de V et T (deuxième ordre). Les paramètres de défaut sont fixés arbitrairement. Les variables dans les boîtes grises en pointillés sont ajustées avec des données expérimentales.

corrélation est obtenue entre le modèle proposé et l'expérience, comme le montre la **Fig. 9.26**. Les différentes dynamiques de caractérisation et de contrainte sont reproduites le long de la variation de tension et de température dans les régions LB et HB. Dans la littérature existante, le mécanisme HCD a été largement étudié sous une contrainte de temps court pour approximer asymptotiquement la solution de l'équation cinétique du premier ordre par une loi de puissance [185], ce qui signifie que les défauts créés ne modifient pas significativement la constante de vitesse de dissociation de liaison κ. Dans cette étude, la plage de temps de contrainte est étendue jusqu'à 10^6s pour sonder le comportement de vieillissement à long terme du dispositif. La dynamique temporelle de contrainte modélisée caractérisée dans les deux régions est en bon accord avec l'expérience, comme le montre la **Fig. 9.27**. À des temps de contrainte faibles inférieurs à 10^4s, une légère divergence est observée, peut-être due à l'utilisation d'une seule distribution de liaison Si-H pour modéliser la cinétique de dégradation, qui peut ne pas être précise pour la zone micrométrique étudiée. De plus, la mesure de faibles dérives du DCR peut être difficile en raison de la dispersion statistique d'un dispositif à l'autre, où la moyenne de ΔDCR suit une loi quadratique avec sa variance. En ignorant l'effet de la caractérisation de la tem-

pérature dû au taux de génération SRH des porteurs qui double tous les 10K dans la différence de température et la probabilité de claquage en avalanche P_t tous les 40K, le nombre de défauts $N_t(t)$ créés au fil du temps de contrainte double tous les 40K et 30K dans l'obscurité et sous éclairage respectivement. La dégradation augmente lorsque la température augmente, ce qui est contraire au comportement attendu pour un mécanisme de dégration à porteurs chauds uniques générique dans les MOSFETs à canal long. Les taux de dégradation dépendent de l'énergie et de la densité des porteurs. La dynamique de perte d'énergie des porteurs par des mécanismes de diffusion augmentant avec la température est la même des MOSFETs aux SPADs, comme le montrent les **Fig. 9.21 (d)-(e)**. Cependant, dans le transistor MOSFET, la densité de porteurs est proportionnelle à la tension de grille moins la tension de seuil (qui est approximativement linéaire avec la température), tandis que dans les SPADs, elle suit le courant inverse de la diode (qui augmente exponentiellement avec la température au travers du taux de génération SRH), comme le montre la **Fig. 9.23 (b)**.

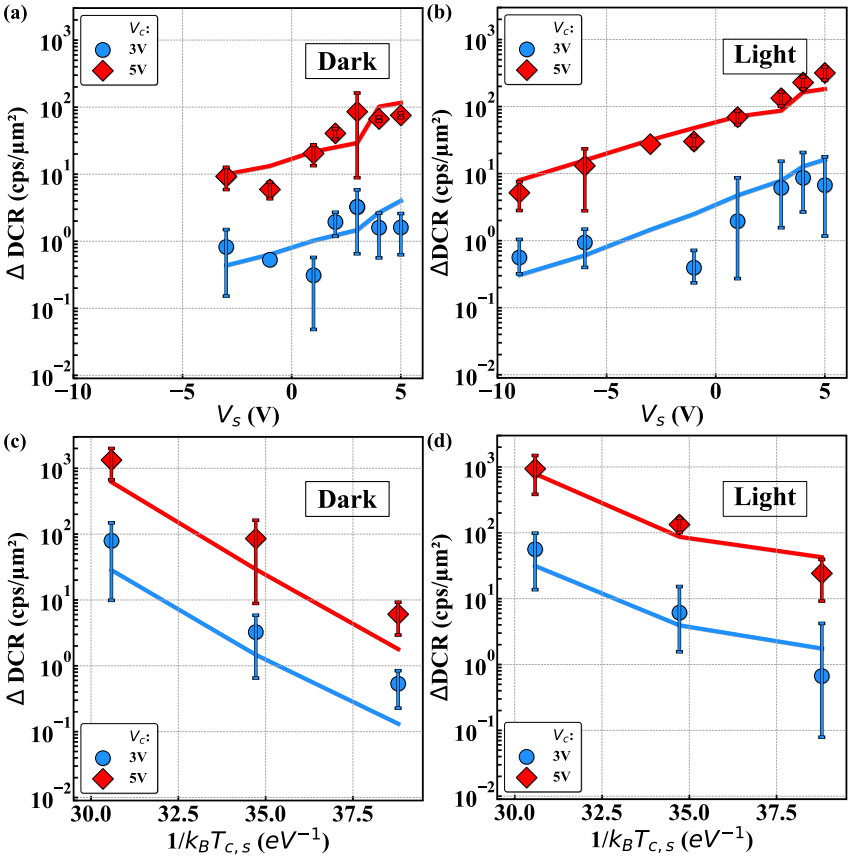


Figure 9.26: ΔDCR en fonction de V_s après des contraintes (a) en obscurité et (b) à 0,3 W/m² et à T_s=333K mesurées (symboles) et modélisées (lignes) à T_c=333K et à V_c=3V (LB - bleu) et 5V (HB - bleu). ΔDCR en fonction de $1/k_BT_s$ pour les contraintes (c) en obscurité et (d) à 0,3 W/m² pour V_s=3V mesurées (symboles), et modélisées (lignes) à T_c=T_s et à V_c=3V (LB - bleu) et 5V (HB - rouge). Les barres d'erreur représentent l'écart-type entre chaque dispositif pour la même condition de contrainte.

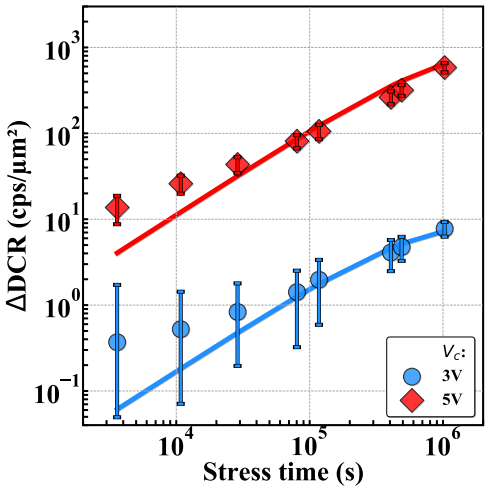


Figure 9.27: ΔDCR en fonction du temps de contrainte pour des contraintes sur 64 SPAD en obscurité à V_s=3V, V_s=353K et à 0,3W/m² mesurées (symboles) et modélisées (lignes) à V_c=3V (LB - bleu) et 5V (HB - rouge) pour T_c=333K.

Évaluation des compromis dans la mesure directe de temps de vol

Compréhension du système optique : composants et principes

Afin de comprendre la distance mesurée entre le capteur et la cible en fonction des photons émis, il est nécessaire de présenter l'ensemble du système ainsi que son environnement. Auparavant, seul un dispositif SPAD était pris en compte, mais maintenant l'ensemble du réseau de pixels sera considéré. Une illustration typique du système optique est décrite dans la **Fig. 9.28**. Le système est composé de :

1. d'un Vertical-Cavity Surface-Emitting Laser (VCSEL) qui émet de la lumière verticalement depuis sa surface

2. d'un réseau de référence de SPAD qui fournit un signal de référence fiable pour le temps zéro pour le réseau de SPADs de retour

3. d'un verre de couverture qui protège les composants sous-jacents contre les dommages et la contamination en fournissant une barrière contre la poussière ou l'humidité et qui améliore les performances optiques du système en réduisant les réflexions et les reflets.

4. d'un réseau de SPADs de retour qui détecte les photons qui sont réfléchis vers le détecteur depuis la cible, permettant de mesurer la distance jusqu'à la cible.

Le principe de fonctionnement d'une mesure dToF pour un seul SPAD est le suivant. Une impulsion laser est émise depuis le VCSEL vers une cible, et le temps nécessaire pour que l'impulsion se réfléchisse vers le capteur est mesuré. La distance est calculée en multipliant le temps nécessaire pour que l'impulsion se déplace vers la cible et revienne par la vitesse de la

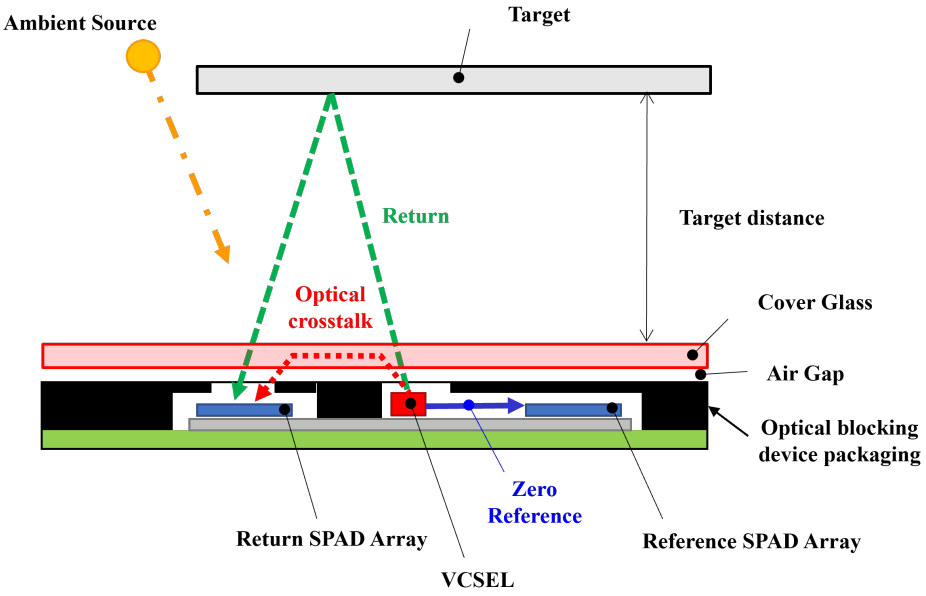


Figure 9.28: Schéma du système optique comprenant : le VCSEL, les SPAD de référence et de retour, le verre de couverture et la cible. Le signal de référence est représenté par la ligne bleue en gras, le signal de retour est représenté par la ligne verte en pointillés tandis que la lumière ambiante et la diaphonie optique sont représentées par les lignes orange en pointillés et rouge en pointillés, respectivement. Les lignes représentées ne représentent pas précisément le trajet optique de la lumière.

lumière dans le milieu à travers lequel l'impulsion se déplace. Pour tenir compte du trajet aller-retour de la lumière vers la cible et retour, ce produit est divisé par deux. Le réseau de pixels est ensuite utilisé pour construire l'histogramme du temps d'arrivée des photons, ce qui est réalisé en considérant chaque lecture. Cet histogramme de comptage est ensuite utilisé pour reconstruire la carte de profondeur de la scène. Le temps entre deux impulsions du VCSEL, également appelé période de répétition d'impulsion, détermine la portée maximale qui peut être mesurée. Il doit être suffisamment long pour permettre au capteur de détecter les photons de retour avant que la prochaine impulsion ne soit émise. Si la période de répétition d'impulsion est trop courte, les photons de retour de l'impulsion précédente peuvent encore arriver au capteur lorsque la prochaine impulsion est émise, ce qui entraîne des erreurs d'enroulement dans la mesure de la portée.

DCR par rapport à la portée maximale de la distance

Pour déterminer la plage de distance maximale d_{max} que le dispositif devrait être capable de mesurer dans les conditions environnementales prévalentes. Pour identifier une plage d'histogramme contenant un signal, il est nécessaire que le signal soit distinguable de la lumière ambiante et du bruit de fond du DCR. Le traitement de l'histogramme nécessite que le signal soit supérieur au niveau de bruit E_{noise} plus un facteur de confiance multiplié par l'écart-type du bruit. Pour le calcul de d_{max}, il est important de rappeler brièvement les différents paramètres qui doivent être pris en compte. La distance d'étalonnage est désignée par d_{cal}, le taux de crête par SPAD

pendant l'étalonnage est désigné par S_{cal}, et la réflectance cible obtenue à partir de la mesure d'étalonnage est désignée par R_{cal}. Un coefficient, désigné par α, est utilisé pour prendre en compte l'ensemble du signal d'impulsion. La réflectance cible considérée est désignée par R, et le nombre de plages dans le segmenteur d'impulsions est désigné par n. Un paramètre de confiance, désigné par S_{conf}, est utilisé pour prendre en compte le pourcentage souhaité de taux de détection. Le bruit de fond, qui prend en compte à la fois le DCR et la lumière ambiante, est désigné par E_{noise}. Enfin, un paramètre de confiance, désigné par A_{conf}, est utilisé pour définir le seuil du segmenteur d'impulsions. Le paramètre de réglage du système est uniquement basé sur la confiance (S_{conf} et A_{conf}). Le bruit de grenaille provoque une variation du nombre d'événements ambiants par plage dans chaque plage de l'histogramme. Le seuil de détection d'impulsion pour chaque plage est déterminé en multipliant le bruit de grenaille par un facteur qui dépend du niveau ambiant dans cette plage.

L'approche de modélisation proposée pour estimer ΔDCR est rapportée dans la **Fig. 9.29** pour les régions de tension de caractérisation à faible polarisation (V_c=3V) et à forte polarisation (V_c=5V). Le pire des scénarios de faible réflectance dans une forêt est considéré. Lorsque le niveau de ΔDCR au cours de la durée de contrainte augmente sous de faibles niveaux d'irradiance d'une nuit couverte, la distance maximale détectable d_{max} commence à diminuer, comme le montre la **Fig. 9.29 (a)**. De même, lorsque le niveau de ΔDCR au cours de la durée de contrainte augmente sous le niveau élevé d'irradiance de la pleine lumière du jour, d_{max} commence à diminuer, comme illustré dans la **Fig. 9.29 (b)**. En comparant les deux graphiques de la **Fig. 9.29**, on peut observer que la distance maximale détectable d_{max} commence à partir d'un plateau qui est limité soit par ΔDCR ou les niveaux de lumière ambiante.

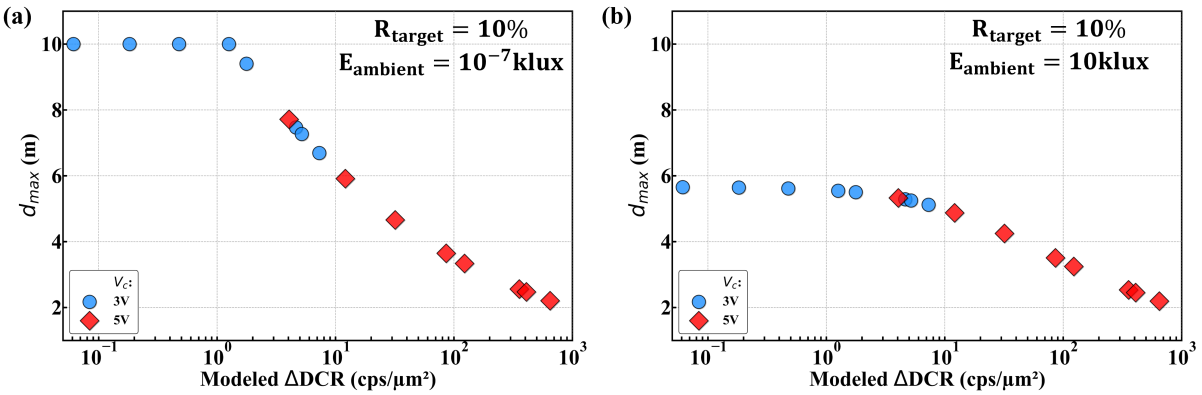


Figure 9.29: Distance maximale de télémétrie d_{max} en fonction de ΔDCR modélisé pour une cible de réflectance R_{target} de 10% à une irradiance de lumière ambiante E_{noise} de (a) 10^{-7}klux et (b) 10klux.

FOLIO ADMINISTRATIF

THESE DE L'INSA LYON, MEMBRE DE L'UNIVERSITE DE LYON

NOM : SICRE DATE de SOUTENANCE : 06/12/2023

Prénoms : Mathieu

TITRE : Study of the noise aging mechanisms in single-photon avalanche photodiode for time-of-flight imaging
Etude des mécanismes de vieillissement du bruit de photodiode à avalanche à photon unique pour l'imagerie par temps de vol

NATURE : Doctorat Numéro d'ordre : 2023ISAL0104

Ecole doctorale : EEA (Électronique, Électrotechnique et Automatique)

Spécialité : Électronique, micro et nanoélectronique, optique et laser

RESUME : Les diodes à avalanche à photon unique (SPAD) sont utilisées pour les capteurs à temps de vol afin de déterminer la distance d'une cible. Cependant, ils sont sujets à des déclenchements parasites par des porteurs de charge générés de manière parasitaire, quantifiés en tant que taux de comptage dans l'obscurité (DCR), ce qui peut compromettre la précision de la distance mesurée. Pour résoudre ce problème, une méthodologie de simulation a été mise en place pour évaluer le DCR. Cela est réalisé en simulant la probabilité de claquage d'avalanche, intégrée avec le taux de génération de porteurs de charge à partir de défauts. Cette méthodologie permet d'identifier les sources potentielles de DCR avant stress. Pour garantir l'intégrité des mesures de distance sur une longue période, il est nécessaire de prédire le niveau de DCR dans diverses conditions d'exploitation. La méthodologie de simulation susmentionnée est utilisée pour identifier les sources potentielles de DCR après stress. Pour un modèle cinétique précis de dégradation de type porteurs chauds (HCD), il est essentiel de considérer non seulement la distribution d'énergie des porteurs, mais également la distribution de l'énergie de dissociation de la liaison Si-H à l'interface Si/SiO_2. La probabilité de dissociation d'ionisation d'impact est utilisée pour modéliser le processus de création de défauts, qui présente une dépendance temporelle sous-linéaire en raison de l'épuisement progressif des précurseurs de défauts. Une mesure précise de la distance nécessite de distinguer le signal du bruit ambiant et du plancher de DCR. L'impact de DCR peut être estimé en considérant la réflectance de la cible et les conditions d'éclairage ambiant. En résumé, ce travail utilise une méthodologie de caractérisation et de simulation approfondie pour prédire le DCR dans les dispositifs de type SPAD le long de sa durée de vie, permettant ainsi d'évaluer son impact sur les mesures de distance.

MOTS-CLÉS : Single-Photon Avalanche Diode (SPAD), Dark Count Rate (DCR), fiabilité, Hot-Carrier Degradation (HCD), Technology Computer Aided Design (TCAD), Monte-Carlo, caractérisation électro-optique.

Laboratoire (s) de recherche : Institut des Nanotechnologies de Lyon (INL) – UMR CNRS 5270

Directeur de thèse : Pr. Francis CALMON

Président de jury :

Composition du jury : Pr. Lucio PANCHERI (rapporteur), Quentin RAFFAY (rapporteur), Pr. Cristell MANEUX (présidente), Pr. Vincent GOIFFON (examinateur), Pr. Francis CALMON, David ROY (invité/co-encadrant), Xavier FEDERSPIEL (invité), Jean COIGNUS (invité)