



HAL
open science

Imageur CMOS ultra haute sensibilité

Rabih Kazma

► **To cite this version:**

Rabih Kazma. Imageur CMOS ultra haute sensibilité. Micro et nanotechnologies/Microélectronique. Université Grenoble Alpes, 2016. Français. NNT : 2016GREAT105 . tel-01582601

HAL Id: tel-01582601

<https://theses.hal.science/tel-01582601v1>

Submitted on 6 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : **NanoElectronique et NanoTechnologies**

Arrêté ministériel : 7 août 2006

Présentée par

Rabih KAZMA

Thèse dirigée par **Olivier ROSSETTO** et
codirigée par **Gilles SIACRD**

préparée au sein du **Laboratoire LPSC**
dans **l'École Doctorale Electronique, Electrotechnique,
Automatique et Traitement du Signal, EEATS**

Imageurs CMOS ultra haute sensibilité

Thèse soutenue publiquement le 01 décembre 2016,
devant le jury composé de :

M. Dominique GINHAC

Professeur à l'université de Dijon, Président

M. Francis CALMON

Professeur à INSA LYON, Rapporteur

M. Wilfried UHRING

Professeur à l'université de Strasbourg, Rapporteur

M. Olivier ROSSETTO

Maître de conférences à l'université Grenoble Alpes, Directeur de thèse

M. Gilles SICARD

Ingénieur au CEA/LETI docteur, Co-directeur de thèse



A mes parents et grâce à eux
A mon frère Thab et grâce à lui

Avant-Propos

Ce manuscrit de thèse est le fruit de trois ans de travail de thèse mené au sein des laboratoires LPSC et TIMA sous la direction de Mr. Olivier Rossetto et Mr. Gilles Sicard. Mes remerciements leurs reviennent, naturellement en premier. J'apprécie énormément toutes les discussions que nous avons eues ensemble au fil de ces années, les conseils qu'ils m'ont toujours offerts et leurs supports aux moments difficiles.

Je tiens à remercier Mr. Arnaud Lucotte, le directeur du laboratoire LPSC, de m'avoir accueilli dans le laboratoire pendant toute la période de ma thèse.

Mes remerciements vont bien sûr également à mes collègues du service électronique du laboratoire LPSC. Je n'oublierai jamais tous ces moments de bonheur qu'on a partagé ensemble pendant ces trois ans. Je tiens à remercier particulièrement : Joel Bouvier, Jean-Pierre Scordilis, Damier Tourres, Cong Li, Jean-Luc Bouly et Ludovic Eraud avec lesquels, j'ai partagé les bons repas à la cantine (H2) durant ces années.

Je continue en remerciant tous les membres du laboratoire LPSC et TIMA pour tous les moments que nous avons partagés ensemble.

Je remercie l'équipe d'enseignement du Master 2 Nanoélectronique et Nanotechnologie pour la bonne formation acquise au sein de l'Université Joseph Fourier (Université Grenoble Alpes actuellement).

Je tiens à remercier tous les membres du jury pour l'intérêt qu'ils ont porté à mon travail en acceptant de faire partie de ce jury.

Evidement une grosse partie de mes remerciements revient à mes parents et mon frère. Ils m'ont soutenu depuis le tout jeune âge, ils m'ont toujours encouragé à aller plus loin et me surpasser.

Résumé

Il existe une multitude d'applications où la détection de très faibles niveaux lumineux (jusqu'à quelques photons, voir le photon unique) est une partie clé de la mesure dans des domaines tels que l'imagerie médicale, l'astrophysique, la vision 3D, la spectroscopie, etc... Jusqu'à un passé récent, le seul composant permettant de détecter des niveaux aussi bas que le photon unique était le photomultiplicateur. Ce composant est volumineux, fragile, nécessite l'utilisation de hautes tensions et possède une résolution spatiale faible.

Depuis quelques années, la faisabilité de photodétecteurs à photons uniques connus sous l'acronyme de SPAD (pour Single-Photon Avalanche Diode) dans un procédé de fabrication CMOS standard a été démontrée. Ces avancées permettent de réaliser des imageurs de haute résolution et grande sensibilité.

Le travail dans cette thèse démarre par la compréhension du comportement de l'élément photosensible (le SPAD) pour finir par la conception d'un circuit de lecture associé au SPAD. Le comportement du SPAD est simulé en utilisant l'outil Silvaco qui va permettre la prédiction du comportement de la photodiode face au mécanisme d'avalanche.

Le deuxième axe de travail proposé dans cette thèse est un modèle haut niveau du SPAD suivi par son circuit de lecture. Cette modélisation nous permet de construire un modèle temporel des instants d'arrivées des impulsions au SPAD tout en distinguant leurs origines.

Le troisième axe propose une nouvelle architecture de lecture. Un des défis importants dans la conception des circuits de lecture des SPADs est d'atteindre une grande dynamique de lecture. Pour cela on propose une nouvelle méthode de lecture qui nous permet d'atteindre une très grande dynamique. Cette nouvelle méthode de lecture repose sur une nouvelle architecture du pixel contenant, entre autres, le SPAD et un compteur analogique. En s'inspirant des techniques de lectures existantes, et grâce à cette nouvelle architecture, une nette amélioration de la dynamique est alors obtenue. Le principal avantage qu'offre cette architecture est qu'elle améliore la dynamique sans ajout de matériel.

Finalement, le dernier chapitre de ce manuscrit de thèse montre les simulations des circuits de lecture qui existent déjà dans la littérature. On s'intéresse aux circuits de lecture basé sur un compteur analogique pour différentes raisons principalement la surface. Un nouveau circuit de lecture est finalement proposé pour pixel de type SPAD. Ce circuit de lecture qui est inspiré des architectures simulées, va nous permettre de réaliser notre fonction de comptage analogique avec les meilleures performances possibles.

Abstract

A number of vision applications (medical imaging, astrophysics, ...) require the detection of very low light levels (up to a few photons, or the single photon). Until recently, only one component could detect levels as low as the single photon: the photomultiplier. This component is bulky, fragile, requires the use of high voltage and has a low spatial resolution.

Since few years, it has been demonstrated the feasibility of designing avalanche diode using CMOS standard process. These advances allow to achieve high-resolution imaging and high sensitivity.

This PhD work starts by the study of the behavior of the photon sensing element (SPAD) and finish by the design of a new readout circuit of single photon avalanche photodiode. The behavior of the SPAD is simulated using Silvaco which allows the prediction of SPAD behavior once the avalanche mechanism is detected.

The second line of the work proposed in this thesis is a new high-level model of SPAD followed by its readout circuit based on an analog counter. This model takes into account the natural sources of photons and dedicated noise and allows us to build a temporal construction of the arrival instant of pulses to the SPAD.

The third line offers a novel readout method to enhance the readout dynamic range of pixel based on single photon avalanche photodiode. This method overcomes the limitation of existing methods by using a double readout technique based on both time and amplitude measurements without decreasing the fill factor. Another advantage of the proposed method is that improvement on the dynamic range is achieved without adding extra electronics.

The last chapter of this thesis shows the simulation of readout circuits that already exist in the literature. We are interested in readout circuit based on an analog counter for various reasons mainly the surface. A new readout circuit is finally proposed for SPAD based pixel. This readout circuit which is inspired from the simulated architectures will allow us to achieve our analog counter with the best performance.

Table des matières:

Intoduction générale	1
Chapitre 1 Généralité sur les capteurs de vision à photon unique	5
1.1 Introduction	5
1.2 Photon : propriété.....	6
1.3 Photodétecteur à photon unique	6
1.3.1 Photomultiplicateurs : PMTs.....	6
1.3.2 EMCCD.....	8
1.3.3 Capteur CMOS APS.....	9
1.3.4 Photodiodes à avalanche en mode linéaire	10
1.3.5 Photodiodes à avalanche en mode Geiger	11
1.3.6 SiPM	12
1.3.7 Superconducting Nanowire Single-Photon Detector (SNSPD).....	13
1.4 Conclusion	15
Chapitre 2 Les photodiodes à avalanche à photon unique ou SPAD	16
2.1 Introduction	16
2.2 Geiger mode Avalanche photodiode	17
2.3 Zone de déplétion	17
2.4 Fabrication des SPAD.....	18
2.4.1 Reachthrough SPAD.....	19
2.4.2 Les SPAD planaires	20
2.5 Facteurs de mérites.....	24
2.5.1 Probabilité de détection de photon (PDP).....	24
2.5.2 Temps de résolution (le Jitter en temps).....	25
2.5.3 Temps mort (Dead Time).....	26
2.5.4 Taux de comptage dans le noir (DCR).....	27
2.5.5 Probabilité de diaphonie (Crosstalk Probability)	28
2.6 Circuits d'étouffement ou de « quenching ».....	29
2.6.1 Circuit d'étouffement passif.....	30
2.6.2 Circuit d'étouffement actif	32
2.6.3 Circuit d'étouffement mixte	34
2.6.4 Circuit d'étouffement avec comparateur de courant	35
2.7 Conclusion	37
Chapitre 3 Etat de l'art du SPAD et des circuits de lecture	38

3.1	Introduction	38
3.2	Evolution des structures des SPADs	38
3.2.1	SPAD CMOS en technologie haute tension (High Voltage HV)	39
3.2.2	STI-Bounded CMOS SPAD 180 nm	40
3.2.3	SPAD CMOS en technologie 130 nm	41
3.2.4	SPAD CMOS en technologie 90 nm	42
3.2.5	SPAD CMOS en technologie 65 nm	42
3.3	Les mécanismes de lectures des SPADs	44
3.3.1	Event driven readout	44
3.3.2	Sequential readout (lecture séquentielle)	45
3.4	Etat de l'art des circuits de lecture	45
3.4.1	Compteur analogique : source de courant et capacité	46
3.4.2	Compteur analogique : Transfert de charge	48
3.5	Conclusion	50
Chapitre 4	Modèle physique et haut niveau des SPADs	51
4.1	Introduction	51
4.2	Modèle physique et simulation électrique	51
4.2.1	Présentation de l'outil Silvaco	51
4.2.2	Structure PN sans anneau de garde	53
4.2.3	Structure pn avec anneau de garde	54
4.3	Modèle haut niveau du SPAD	55
4.3.1	Le processus de Poisson	56
4.3.2	Processus de Poisson : Propriété	57
4.3.3	Processus de Poisson : algorithmes de simulation	57
4.3.4	Source lumineuse	59
4.3.5	Taux de comptage e noir	62
4.3.6	Modélisation du temps mort	63
4.3.7	Modélisation de l'afterpulsing	64
4.3.8	Modèle complet	65
4.4	Modèle du circuit de lecture	66
4.5	Conclusion	68
Chapitre 5	Du SPAD au pixel	69
5.1	Introduction	69
5.2	Introduction aux techniques de lecture dans un SPAD	69
5.2.1	Méthode basée sur le temps d'intégration fixe	70
5.2.2	Méthode basée sur la tension de référence fixe	71

5.3	Architecture proposée à dynamique élevée.....	72
5.3.1	Concept général.....	73
5.3.2	Pixels à fort taux de comptage.....	74
5.3.3	Pixels à faible taux de comptage.....	75
5.3.4	Amélioration de la dynamique : impact de l'architecture.....	76
5.4	Impact de la Non linéarité (NL).....	79
5.5	Réduction de l'erreur de quantification en temps.....	82
5.5.1	Résolution en temps uniforme.....	83
5.5.2	Système de mesure en temps multi-résolution.....	85
5.6	Modèle en langage VHDL-AMS.....	88
5.6.1	Modèle du pixel en VHDL-AMS.....	88
5.6.2	Modèle vhdl_ams : Multiples temps de résolution.....	92
5.6.3	Modèle VHDL-AMS : Mesure de l'amplitude de tension.....	94
5.7	Conclusion.....	96
Chapitre 6	Implémentation CMOS du circuit de lecture du SPAD.....	97
6.1	Introduction.....	97
6.2	Simulation de différents compteurs existants.....	98
6.2.1	Compteur analogique à courant commuté.....	98
6.2.2	Compteur analogique : intégrateur à base de capacité commutée.....	101
6.2.3	Compteur analogique par transfert de charge.....	105
6.3	Circuit électrique proposé.....	108
6.3.1	Vue globale du circuit proposé.....	108
6.3.2	Compteur analogique.....	109
6.3.3	Comparateur.....	112
6.3.4	Compteur complet à deux étages.....	114
6.4	Conclusion.....	116
Chapitre 7	Conclusion générale et perspectives futures.....	117
7.1	Conclusion générale.....	117
7.2	Perspectives.....	119
	Bibliographie.....	120
	Liste des publications.....	124
	Résumé.....	125
	Abstract.....	125

Table des Figures:

Figure 1.1 : Schéma du principe de fonctionnement d'un photomultiplicateur (wikipedia)	7
Figure 1.2 : Schéma du principe de fonctionnement des EMCCDs	9
Figure 1.3 : Schéma typique d'un APS (a), avec son diagramme de fonctionnement (b)	10
Figure 1.4 : Caractéristique typique d'une jonction pn	11
Figure 1.5 : Schéma électrique du SiPM	13
Figure 1.6 : Structure du SNSPD	14
Figure 1.7 : Structure serpentine SNSPD	14
Figure 2.1 : Reachthrough SPAD McIntyre [Dautet_93]	19
Figure 2.2 : Rechthrough SPAD [Zappa_04]	19
Figure 2.3 : SPAD planaire de Haitz [Goetzberger_63]	21
Figure 2.4 : Cova SPAD basé sur une double couche épitaxiales [Cova_04]	22
Figure 2.5 : CMOS SPAD HV de Rochas	22
Figure 2.6 : STI-Bounded SPAD CMOS 130nm entouré par une couche de passivation (p-type) [Gersbach_08]	23
Figure 2.7 : Variation de la PDE en fonction de longueur de la longueur onde [Cova_96]	25
Figure 2.8 : Temps de résolution (FWMH) [Cheng_12]	26
Figure 2.9 : Temps mort dans un SPAD	27
Figure 2.10 : La probabilité d'une diaphonie	28
Figure 2.11 : Circuit d'étouffement passif [Cova_96]	30
Figure 2.12 : Modèle électrique PQC [Cova_96]	31
Figure 2.13 : Circuit d'étouffement actifs [Tisa_08a]	33
Figure 2.14 : Circuit d'étouffement mixte [Tisa_08b]	34
Figure 2.15 : Circuit d'étouffement avec comparateur de courant [Mita_08]	35
Figure 3.1 : Structure du SPAD p-/n- en technologie 0,35 μ m CMOS HV	39
Figure 3.2 : STI-Bounded CMOS SPAD en technologie 180 nm [Finkelstein_06]	40
Figure 3.3 : LV CMOS SPAD en technologie 130 nm [Niclass_07]	41
Figure 3.4 : SPAD CMOS en technologie 90 nm [Richardson_12]	42
Figure 3.5 : Structure CMOS SPAD en technologie 65 nm (en bas) et le layout du SPAD CMOS 65 nm (en haut) [Charbon_13]	43
Figure 3.6 : Le principe d'un Event driven readout [Niclass_06]	44
Figure 3.7 : Lecture séquentielle : lecture par pixel (a) lecture par colonne (b) [Kumar_12]	45
Figure 3.8 : Compteur analogique basé sur la recharge d'une capacité	46
Figure 3.9 : Compteur analogique double étage [Palubiak_11]	47
Figure 3.10 : Principe du compteur à transfert de charges	48
Figure 3.11 : schéma électrique du circuit de lecture basé sur un transfert de charge [Ekaterina_14]	48
Figure 4.1 : Organigramme du logiciel SILVACO	52
Figure 4.2 : distribution du champ électrique dans la structure n+/p_substrat	53

Figure 4.3 : Caractéristiques I(V) de la jonction n+/p_substrat sans anneau de garde	54
Figure 4.4 : distribution du champ électrique dans la structure n+/p_substrat avec anneau de garde ..	54
Figure 4.5 : Caractéristiques I(V) de la jonction n+/p_substrat avec anneau de garde.....	55
Figure 4.6 : Structure du modèle développé	56
Figure 4.7 : exemple d'un trajectoire d'un processus de Poisson.....	57
Figure 4.8 : la distribution temporelle des impulsions provenant de la source de photon.	60
Figure 4.9 : Histogramme du nombre de photons/seconde pour 10000 échantillons	60
Figure 4.10 : variation de la PDP en fonction de la longueur d'onde pour les technologies 0.35 et 0.8 μm [Niclass_06].....	61
Figure 4.11 : le taux de comptage dans le noir en fonction de l'excès de la tension de polarisation pour différentes températures [Niclass_06].....	62
Figure 4.12 : λ_{DCR} Histogramme pour un SPAD de $36 \mu\text{m}^2$	63
Figure 4.13 : le modèle du temps mort	64
Figure 4.14 : la variation de la probabilité d'Afterpulsing en fonction du temps morts [Niclass_06]..	65
Figure 4.15 : Modèle complet haut niveau du SPAD.....	66
Figure 4.16 : modèle haut niveau du SPAD suivi de son circuit de lecture	66
Figure 5.1 : Méthode basée sur un temps d'intégration fixe	70
Figure 5.2 : schéma bloc du principe de la méthode basé sur une tension de référence fixe.....	71
Figure 5.3 : Méthode basée sur un tension de référence fixe	72
Figure 5.4 : Schéma bloc de l'architecture du pixel proposée.....	73
Figure 5.5 : Concept général de la nouvelle architecture de lecture.....	74
Figure 5.6 : Traitement des faibles taux de comptages : rouge comptage, Blue décomptage	75
Figure 5.7 : Fonction réciproque qui lie le taux de comptage au temps	77
Figure 5.8 : Relation entre $N\gamma$ et $T1$	79
Figure 5.9 : Simulation de l'architecture propose en présence de la non linéarité	81
Figure 5.10 : teste de robustesse de l'architecture.....	81
Figure 5.11 : la mesure du temps dans l'architecture proposée.....	83
Figure 5.12 : Sortie numérique qui dépend du nombre de photons $N\gamma$	84
Figure 5.13 : Distribution de l'erreur de quantification pour 7 bits de sorties numérique.....	84
Figure 5.14 : Sortie numérique en fonction du nombre de photons $N\gamma$ en utilisant trois résolution temporelles différentes	85
Figure 5.15 : Principe du système multi-résolution et le lien avec le compteur système	86
Figure 5.16 : Erreur de quantification pour un registre pixel 7 bits (multiples temps de résolution) ...	87
Figure 5.17 : Erreur de quantification pour un registre pixel 8 bits (multiples temps de résolution) ...	87
Figure 5.18 : Modèle VHDL-AMS du pixel proposé.....	88
Figure 5.19 : Sortie numérique pour la mesure en temps (simple fréquence système).....	89
Figure 5.20 : Simulation mesure en temps (simple temps de résolution).....	90
Figure 5.21 : Simulation de la mesure en temps en présence de la non linéarité (NL).....	91
Figure 5.22 : Simulation paramétrique pour les valeurs de $N\gamma$ entre 127 et 1023	91
Figure 5.23 : Modèle VHDL-AMS du pixel implémenté pour simuler les deux cas de mesures.....	93

Figure 5.24 : Simulation de la mesure en temps du modèle VHDL-AMS (méthode multi-résolution)	93
Figure 5.25 : Simulation de la mesure en amplitude du modèle du pixel en vhdl-ams	93
Figure 6.1 : Réalisation électrique de l'architecture [Palubiak_11] simple étage	98
Figure 6.2 : Simulation transitoire de l'architecture de [Palubiak_11] simple étage.....	99
Figure 6.3 : Simulation Monte Carlo pour 10 simulations.....	100
Figure 6.4 : Circuits à capacités commutes (Intégrateur) [Auvray]	101
Figure 6.5 : : Intégrateur formée avec des capacités commutées	102
Figure 6.6 : Simulation paramétrique en variant le gain de l'amplificateur opérationnel (100,150, 1000 et 1500).....	103
Figure 6.7 : Intégrateur en utilisant un amplificateur opérationnel pmos 5 transistors.....	103
Figure 6.8 : Simulation Monte Carlo du circuit à capacités commutes (intégrateur)	104
Figure 6.9 : Principe de transfert de charge.....	105
Figure 6.10 : Schéma électrique compteur analogique par transfert de charge [Ekaterina_14]	106
Figure 6.11 : Simulation transitoire pour le schéma électrique de [Ekaterina_14] avec un rapport K = 128	107
Figure 6.12 : Simulation MC en mode mismatch (200 simulations).....	107
Figure 6.13 : Architecture du modèle électrique proposée.....	109
Figure 6.14 : Schéma du compteur analogique à base de transfert de charge	110
Figure 6.15 : Simulation transitoire pour identifier l'impact du coefficient K.....	110
Figure 6.16 : Simulation Monte Carlo (100 Simulations).....	111
Figure 6.17 : Zoom Figure 6.12 pour voir la marge entre la 10ème et la 11ème impulsions	112
Figure 6.18 : Schéma du comparateur.....	113
Figure 6.19 : Simulation Monte Carlo en continue comparateur	113
Figure 6.20 : Schéma électrique du modèle complet	114
Figure 6.21 : Simulation transitoire des sorties compteur 1 et comparateur (droite) et compteur 2 (gauche)	114

Introduction générale

La lumière a des propriétés intéressantes qui attirent les chercheurs en informatique et en électronique depuis longtemps. Sa nature ondulatoire et donc sans masse fait d'elle une candidate parfaite pour le transfert de donnée à longue distance. Cependant une nouvelle tendance peut être observée. Les chercheurs se tournent vers la photonique, la science de la lumière pour communiquer également à courte distance. En effet, le photon, contrairement à l'électron, ne subit pas d'effets résistifs, capacitifs ou inductifs. La nature ondulatoire permet aux concepteurs de surmonter des problèmes tels que l'injection de bruit ou de diaphonie. C'est vrai qu'actuellement, on ne peut pas remplacer entièrement l'électronique traditionnelle par la photonique mais une combinaison des deux sera très certainement adoptée afin de profiter des deux mondes.

Les technologies qui utilisent la détection photonique sont en plein essor. Ces technologies prennent une place de plus en plus importante dans notre vie quotidienne. La détection photonique touche à des applications comme la cartographie, la météorologie, l'automobile, l'imagerie médicale ou les téléphones intelligents et tablette. Par exemple, le domaine médical bénéficie grandement des technologies de détections photoniques et l'imagerie 3D. Une des applications médicales est la tomographie optique diffuse (TOD). Elle utilise la lumière avec des longueurs d'onde du rouge et des infrarouges pour caractériser en trois dimensions des tissus biologiques en profondeur. Cette technique est utile dans le dépistage pour le cancer du sein ou encore dans l'étude de l'activité cérébrale, mais devient moins performante pour la détection de tumeurs cancéreuse à l'intérieur d'un patient.

La photodétection est alors un sujet de recherche très actif de nos jours et concevoir des photodétecteurs qui fonctionnent avec une grande sensibilité et une grande densité d'intégration est devenu un défi réalisable. Pour ces raisons, les photodiodes à avalanche polarisée en mode Geiger (connu sous le nom de Single Photon Avalanche Photodiode, ou SPAD) sont capables de détecter les photons uniques. Pour cela, elles suscitent beaucoup d'intérêt depuis quelques années.

Il y a quelques années, les capteurs de vision CMOS ont proposé une alternative aux capteurs (CCD ou Charge Coupled Device) qui ont dominé le marché pendant une longue période. Le capteur CMOS constitue un axe de recherche et de développement très important sur lequel de nombreuses équipes de recherche sont focalisées. L'évolution rapide des imageurs CMOS et la capacité d'intégrer des millions de transistors sur une même puce a permis d'espérer des performances inimaginables de capteurs de visions, ce qui a servi dans des applications qui nécessitent une grande sensibilité, une forte résolution et des dynamiques élevées.

Depuis la réalisation du premier SPAD proposée par Haitz en 1960, les architectures n'ont cessé d'évoluer. L'évolution rapide des imageurs CMOS a fait en sorte que les SPAD en CMOS ont évolué également. Cette évolution de SPAD se traduit en termes de performances qui touchent les facteurs de mérites comme la résolution en temps, la probabilité de détection de photon ou le temps mort.

Le sujet que nous traitons dans cette thèse concerne principalement l'amélioration de la dynamique de lecture. Actuellement, plusieurs architectures de circuit de lecture existent, ces architectures sont limitées par la dynamique qui constitue un paramètre affectant leurs performances. Pour cela, nous abordons dans ce mémoire de thèse, les principaux axes suivants :

- Compréhension et modélisation de l'élément photosensible,
- Inclusion de l'élément photosensible dans un pixel actif,
- Proposition d'une nouvelle architecture de lecture qui vise l'amélioration de la dynamique de lecture,
- Réalisation d'un circuit électrique.

Ce manuscrit de thèse est structuré de la façon suivante : le **chapitre 1** est une généralité sur les principaux capteurs de visions qui existent, en allant du Photomultiplicateurs, qui a constitué le premier photodétecteur capable de la détection de photon unique, pour arriver aux SiPM.

Le **chapitre 2** est une introduction aux photodiodes à avalanches polarisées en mode Geiger. Dans ce chapitre, nous détaillerons l'effet d'avalanche dans ce type de photodiodes et nous expliquerons les différentes structures de fabrication existantes actuellement permettant la production des SPADs. Ensuite, nous évoquerons le sujet des facteurs de mérites qui vont servir à caractériser les performances des SPADs. Nous citerons encore les différentes architectures de circuits d'extinction pour étouffer l'avalanche.

Le **chapitre 3** revient sur les différentes structures de SPAD en technologie CMOS allant de 0.8μ à 65 nm tout en analysant leurs atouts et leurs limitations. Dans ce chapitre également, on propose une étude de l'état de l'art des circuits de lecture de SPAD.

Le **chapitre 4**, en utilisant le logiciel Silvaco, présente des simulations physiques et électriques de la photodiode à avalanche. Un modèle haut niveau de la photodiode à avalanche suivi par son circuit de lecture à dynamique limitée est présenté. Ce nouveau modèle, prend en considération les différents facteurs de mérites qui caractérisent les performances du SPAD.

Dans le **chapitre 5**, nous proposons une nouvelle architecture de lecture du SPAD. Cette architecture est une combinaison des architectures de lectures existantes qui vise à améliorer la dynamique de lecture. Des simulations au niveau système confirmeront la validité de cette architecture. On propose également un modèle vhdl-ams de l'architecture du pixel.

Enfin, le **chapitre 6** propose une réalisation électrique de l'architecture à grande dynamique présentée au **chapitre 5**.

Une conclusion synthétique finalise ce manuscrit et ouvre l'horizon à d'autres perspectives pouvant être menées dans le futur.

Chapitre 1 *Généralité sur les capteurs de vision à photon unique*

1.1 Introduction

La détection de lumière de très faible intensité existe depuis un siècle. En 1913, Elster et Geiter ont inventé le tube photoélectrique quelques années après la découverte de l'effet photoélectrique par Albert Einstein. Il a fallu attendre plus de 20 ans pour voir l'apparition du premier photomultiplicateur (en anglais Photo Multiplier Tube (PMT)) inventé dans les laboratoires RCA, et qui est devenu en 1936 un produit commercial. Depuis cette date, les photons uniques sont détectables.

Avant de décrire les détecteurs de photon unique, il est important de définir le terme "photon unique". Techniquement, tous les capteurs de vision sont des détecteurs de photon mais tous les détecteurs ne sont pas capables de détecter un photon unique (dans ce cas, le signal généré par un seul photon est trop faible pour être détecté et se trouve noyé dans le bruit). Un détecteur de photon unique est capable de mesurer une ou plusieurs caractéristiques de la présence d'un photon unique sans avoir besoin de la présence de plusieurs.

Les détecteurs de photon unique ont attiré l'attention en raison de leurs rôles critique dans beaucoup d'applications importantes telles que l'imagerie médicale, la vision 3D, le temps de vol (TOF), la spectroscopie, l'imagerie quantique ou encore la cryptographie. Les détecteurs de photon unique forment un groupe de dispositifs spécifiquement conçus et construits pour détecter les signaux lumineux extrêmement faibles. En général, un détecteur de photon unique doit avoir les propriétés suivantes : une haute efficacité de détection, une bonne résolution temporelle et un faible bruit.

Au cours des dernières années les chercheurs ont développé différents types de détecteurs de photon unique, pour différentes applications, basés sur différentes technolo-

gies ou mécanismes physiques. Nous allons dans ce chapitre détailler l'état de l'art des principaux détecteurs de photon unique existants et évaluer leurs avantages et leurs inconvénients.

1.2 Photon : propriété

La lumière possède à la fois les propriétés d'une onde et celle d'une particule (dualité onde-corpuscule). Le photon est la particule élémentaire, qui possède un quantum d'énergie associée à une onde lumineuse (onde électromagnétique). Un photon a une masse nulle et ne porte pas de charge électrique. Il est caractérisé soit par sa longueur d'onde λ , soit par sa fréquence ν ou par son énergie E exprimé en électron-volt (eV). L'énergie d'un photon associée à une onde lumineuse est donnée par l'équation (1.1).

$$E = h \cdot \nu = \frac{h \cdot c}{\lambda} \quad (1.1)$$

Où c représente la vitesse de la lumière et h la constante de Planck.

Les propriétés des photons et leurs interactions avec d'autres particules sont appréhendées théoriquement par la mécanique quantique et par de nombreuses expériences comme l'électrodynamique qui décrit les interactions des photons avec la matière. C'est donc cet aspect corpusculaire de la lumière qui permet de décrire les interactions qui sont mises en jeu dans les détecteurs de lumière, ou détecteurs de photons.

1.3 Photodétecteur à photon unique

Comme cité précédemment, l'aspect corpusculaire de la lumière est le phénomène responsable de la détection. Concrètement, cette détection est l'interaction du rayonnement avec la matière qui permettra la détection d'un photon.

Le choix d'un photodétecteur sera dicté par divers paramètres, le plus important dans le cas de la détection des photons uniques étant la sensibilité. Dans cette section, nous citerons les divers types de détecteurs courants particulièrement dédiées à détecter les faibles flux lumineux, d'autant que dans ce secteur, on observe une forte évolution technologique ces dernières années.

1.3.1 Photomultiplicateurs : PMTs

Le photomultiplicateur, connu sous le nom de Photo Multiplier Tube ou PMT, était jusque dans les années 80, le seul détecteur de flux lumineux de très faible intensité existant sur le marché. Il a été inventé en 1936 dans le laboratoire RCA en Russie. Le principe de fonctionnement des PMTs est basé sur deux phénomènes physiques fondamentaux : la

photoémission et l'émission secondaire. La Figure 1.1 explique le principe de fonctionnement des photomultiplicateurs [Marmonier_02].

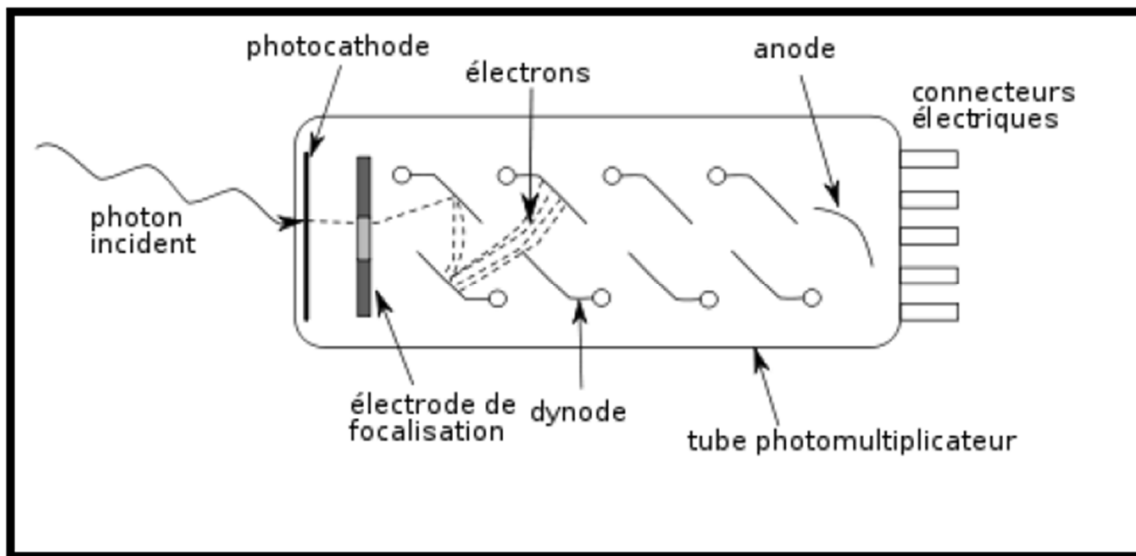


Figure 1.1 : Schéma du principe de fonctionnement d'un photomultiplicateur (wikipedia)

Le photomultiplicateur est composé de :

- Une photocathode, qui par effet photoélectrique convertit le flux lumineux en flux d'électron ;
- Un système électro optique qui concentre et accélère le flux ;
- Un multiplicateur d'électrons formé par une série d'électrodes (dynodes) à émission secondaire ;
- Une anode qui recueille le flux d'électrons et fournit le signal de sortie.

Le principe de fonctionnement est le suivant : lorsqu'un photon heurte la photocathode, des électrons sont alors produits par effet photoélectrique. L'électrode de focalisation permet de diriger les électrons créés vers le multiplicateur d'électron, celui-ci étant composé de plusieurs dynodes. Chaque dynode est maintenue à une valeur de potentiel plus importante que la précédente. Les électrons incidents sont accélérés par le champ électrique ainsi produit et vont frapper la première dynode. Ce choc va créer d'autres électrons, moins énergétique mais plus nombreux. Tous ces électrons vont à nouveau être accélérés et vont frapper la seconde dynode, arrachant de nouveaux électrons à cette dernière, et ainsi de suite jusqu'à atteindre l'anode. L'arrivée de ces électrons sur la dernière dynode ou anode crée une brève impulsion de courant signalant ainsi la détection d'un photon [Pellion_08].

Les principales caractéristiques d'un photomultiplicateur sont :

- Le gain très élevé ($\sim 10^6$ électrons créés par photon incident),
- Le temps de transit qui représente le temps entre l'illumination et le moment de l'impulsion sur l'anode donc traduit une bonne résolution temporelle,
- Le très faible taux de comptage dans le noir grâce au faible bruit (parasite) [Jradi_10].

Malgré les avantages que présente le photomultiplicateur, il souffre de plusieurs inconvénients principalement une tension d'alimentation élevée qui va de quelques centaines de volts pour atteindre plusieurs milliers de volts. Le PMT est un détecteur lourd, fragile et très sensible au champ magnétique. Pour surmonter les inconvénients du photomultiplicateur, les chercheurs ont commencé dès les années 80 à tester de nouveaux dispositifs de détection de photon pour différentes applications.

1.3.2 EMCCD

L'EMCCD ou (*Electron Multipliering Charged Coupled Device*) représente une importante évolution dans la détection de la lumière au cours de la dernière décennie. Ce photodétecteur a révolutionné le domaine de la spectroscopie et de l'imagerie à faible niveaux lumineux ce qui n'était pas envisageable avant l'invention de l'EMCCD. La multiplication des électrons (EM) sur la puce du CCD permet l'amplification des électrons capturés du signal avant qu'il ne passe à la phase de lecture ce qui entraîne une réduction du bruit relatif du circuit de lecture.

Le principe des EMCCDs est le suivant : le photodétecteur possède deux registres de lecture supplémentaire par rapport à la technologie CCD classique, un registre normal ou registre conventionnel et le registre (EM) consacré à la multiplication des électrons. Les tensions utilisées sur le registre (EM) sont plus élevées que celles utilisées sur le registre conventionnel. On applique des tensions élevées sur le registre (EM) afin que les électrons acquièrent une énergie suffisante pour que des électrons supplémentaires soient libérés par le processus d'impact d'ionisation. Le paquet de charges créé par impact d'ionisation est stocké dans le pixel suivant tout au long de la chaîne. Cette étape de gain permet de multiplier le signal sans ajouter de bruit supplémentaire. Le gain dans le nombre d'électrons pourra atteindre des facteurs de l'ordre de 1000. La Figure 1.2 montre le principe de la multiplication des électrons.

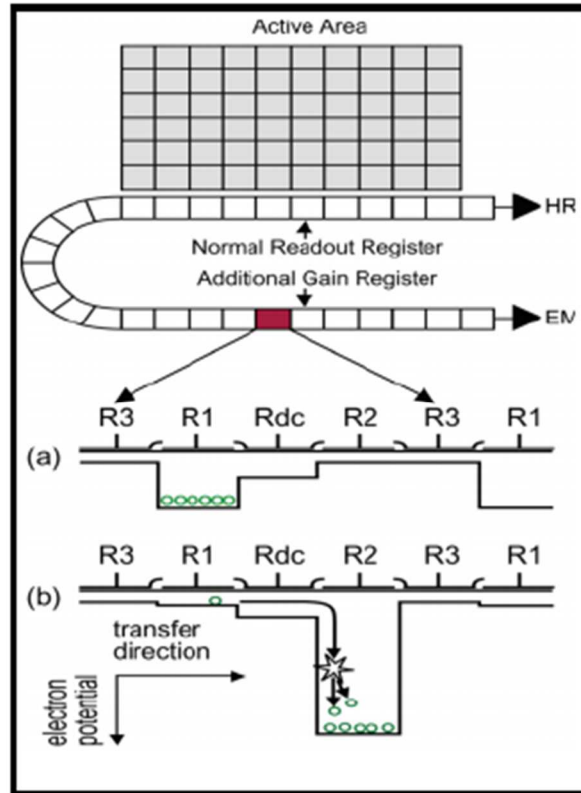


Figure 1.2 : Schéma du principe de fonctionnement des EMCCDs

Les principaux avantages des capteurs EMCCDs sont : une forte efficacité quantique et une bonne sensibilité au photon unique. Le principal inconvénient par rapport à la technologie CCD est la forte consommation surtout pour les faibles niveaux d'intensité [Andor_technology_05].

1.3.3 Capteur CMOS APS

Les capteurs dits à pixels actifs (APS, Active Pixel Sensor) utilisent des éléments actifs dans le pixel (des transistors) pour amplifier le signal acquis avant qu'il soit converti en un signal numérique. Ce signal est généré par une jonction P-N ou une jonction P-I-N (avec une zone non-dopée, dite intrinsèque I, intercalée entre deux zones dopées P et N) polarisé dans la zone inverse de la photodiode. De nos jours, le terme APS est lié au capteur de vision CMOS. Peu coûteux, ce type de capteur a eu une ascension fulgurante dans les années 2000 à 2010 suite à la forte demande des clients essentiellement dans le secteur grand public (appareils photos, téléphones ...).

Récemment, les capteurs de visions CMOS APS ont été réalisés pour détecter les flux lumineux de faibles niveaux. Il est probable que ces détecteurs seront conformes à la définition de la détection du "photon unique" dans les prochaines années. Les capteurs CMOS APS comme les capteurs EMCCDs, ne disposent pas actuellement d'une résolution suffisante dans le domaine temporel pour fonctionner en mode photon unique. Cependant, ces capteurs ne nécessitent pas de refroidissement et devraient être beaucoup moins coûteux

par rapport aux capteurs CCDs [Fishburn_12]. La Figure 1.3 montre le schéma type d'un capteur APS avec trois transistors et son diagramme de fonctionnement.

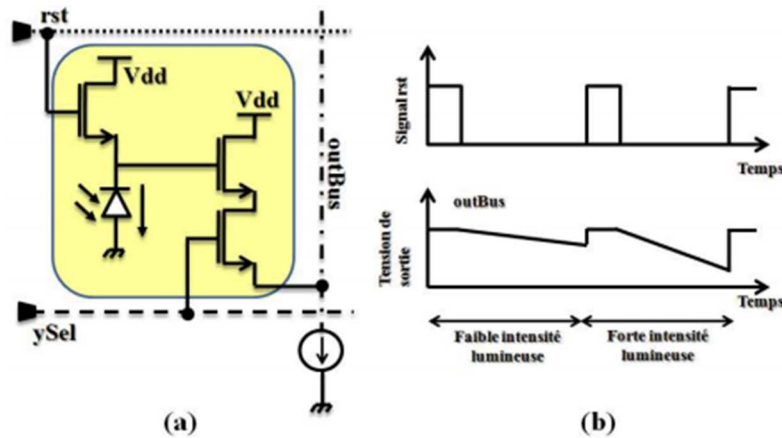


Figure 1.3 : Schéma typique d'un APS (a), avec son diagramme de fonctionnement (b)

Toutefois, le bruit n'est pas totalement éliminé avec ce type de circuit puisque le procédé de fabrication comporte de légères variations menant à la génération des transistors quelque peu différents. Ce bruit spatial connu par le terme (FPN pour Fixed Pattern Noise) se manifeste par la variation des valeurs de sorties des pixels sous illumination constante [Amhaz_12].

1.3.4 Photodiodes à avalanche en mode linéaire

Un nouveau concept permettant la détection de très faible flux lumineux est né dans les années 90, basé sur une caractéristique élémentaire des jonctions à semi-conducteurs. Les photodiodes à avalanche (PDAs) sont des photodétecteurs à semi-conducteur qui utilise le principe du gain interne d'une jonction en appliquant une forte tension de polarisation en inverse. Grâce à ce gain interne, les PDAs permettent la détection de faibles niveaux de lumière et sont utilisés dans une grande variété d'applications. En général, les PDAs sont des structures basées sur une jonction p-n ou p-i-n.

Lorsque la photodiode est utilisée en mode linéaire, c'est-à-dire avec une tension de polarisation en inverse assez élevée, mais légèrement en dessous de la tension de claquage, elle peut être utilisée pour obtenir une réponse linéaire en fonction du flux lumineux incident (cas des capteurs APS). Lorsqu'un photon est absorbé dans sa zone de déplétion, il crée une paire d'électron trou. Les porteurs photogénérés sont accélérés par le champ électrique appliqué aux bornes de la photodiode et créent à leur tour d'autres paires d'électron trou par ionisation par impact qui vont à leur tour en former d'autres. Ce mécanisme présente un gain fixe qui dépend de la tension de polarisation et selon les caractéristiques de la photodiode.

En général, pour une paire d'électron trou, créée par photon incident, une multiplication de paires allant de quelques dizaines à quelques centaines est obtenue. Comme les PDAs produisent un signal relativement faible, ce signal doit être amplifié par un préamplificateur de charges avec un très faible bruit afin de ne pas détériorer son rapport signal sur bruit (SNR). Ce photodétecteur présente deux désavantages quant à la gigue temporelle (résolution temporelle), le premier est dû au fait que la PDA requiert quelques photons pour que son avalanche soit détectable, posant une incertitude sur le temps d'arrivée du photon, en considérant que l'arrivée des photons suit une distribution de Poisson. Le deuxième gros inconvénient provient du faible rapport signal sur bruit de la photodiode accentuant cette incertitude [Boisvert_14].

1.3.5 Photodiodes à avalanche en mode Geiger

Les photodiodes à avalanches offrent une seconde option puisqu'elles sont conçues pour opérer plus particulièrement au-delà de la tension de claquage. Lorsqu'elles sont polarisées dans ces conditions, elles opèrent en mode Geiger, plus communément appelées selon l'acronyme anglais SPAD (Single Photon Avalanche Photodiode). Le champ électrique très élevé appliqué entre ses bornes permet à la photodiode de passer dans un état métastable où un seul photon absorbé dans la région photosensible suffit à déclencher une avalanche par impact ionisant dans le réseau cristallin. La Figure 1.4 montre les deux zones de fonctionnement des photodiodes à avalanches.

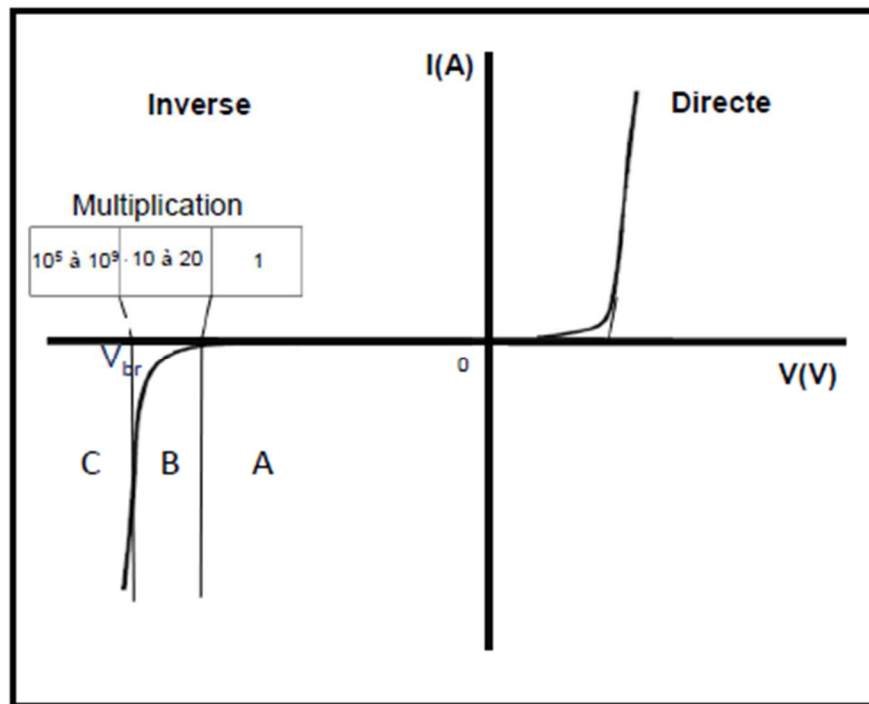


Figure 1.4 : Caractéristique typique d'une jonction pn

La Figure 1.4 présente la caractéristique électrique $I(V)$ d'une jonction p-n avec :

- M : le coefficient de multiplication.
- A : zone de fonctionnement sans amplification : $M = 1$.
- B : zone de fonctionnement linéaire : $M = 10$ à 200 .
- C : zone de fonctionnement Geiger : $M = 10^5$ à 10^9 .

Une fois déclenchée, l'avalanche s'emballe et cause l'autodestruction de la photodiode si cette dernière n'est pas rapidement stoppée par un circuit capable de diminuer sa tension sous la tension de claquage. Par sa vitesse de réaction, son emballement dans l'avalanche et son gain très important, la photodiode à avalanche en mode Geiger ne nécessite qu'un seul photon pour fournir un signal.

Les photodiodes à avalanche en mode Geiger peuvent être fabriquées dans divers procédés comme le procédé optoélectronique ou le procédé CMOS standard. Chacun présente ses avantages et ses inconvénients. Actuellement la majorité est fabriquée en utilisant le process CMOS standard qui possède un coût de production inférieur à celui des autres process dû à sa simplicité et ses volumes de productions importants. De plus le process CMOS permet d'intégrer l'électronique de lecture de la photodiode juste à côté de cette dernière, sur la même puce. Le principal inconvénient des SPADs est le faible facteur de remplissage du fait que la photodiode et son circuit de lecture sont intégrés justement sur une seule puce [Boisvert_14].

1.3.6 SiPM

Les Silicon photomultipliers ou SiPM ont été conçus pour être une alternative aux PMTs. Les SiPM sont des dispositifs de silicium composés généralement de 100 à 1000 photodiodes avalanche (PDAs) en parallèle sur une seule puce comme le montre la Figure 1.5.

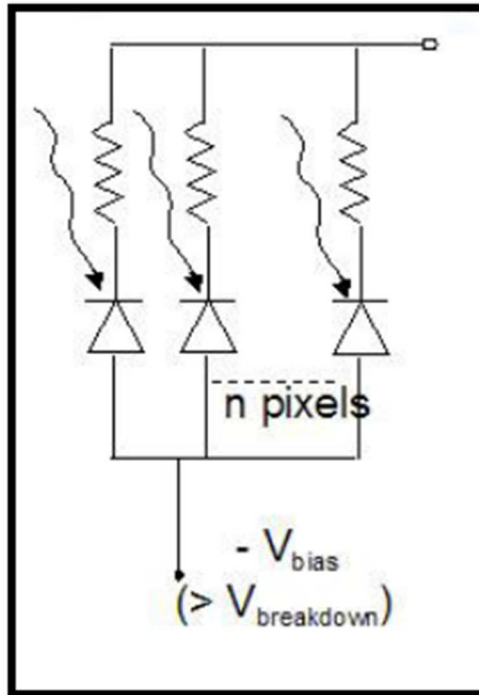


Figure 1.5 : Schéma électrique du SiPM

L'avalanche qui se produit dans une des photodiodes de la matrice se traduit par une impulsion de courant en sortie. Si plusieurs événements d'avalanche se produisent simultanément sur deux pixels différents, la sortie sera une sommation de ces impulsions, qui correspond au nombre de photons incidents qui touchent le photodétecteur.

Les SiPMs ont montrés leurs utilités pour des applications comme la médecine nucléaire. Grâce à un gain élevé, un faible bruit, une insensibilité au champ magnétique et sa grande surface, le SiPM est devenu le premier choix de détecteurs pour les système PETs. En raison de la haute résolution en temps (< 100 ps de gigue) ce type de photodétecteur est utilisé en masse pour les applications basées sur le temps de vol (TOF). Cependant, le signal de lecture rapide reste un problème en raison de l'absence de circuit intégré sur la puce, ce qui nécessite une électronique de lecture à hautes performances [Hsu_10].

1.3.7 Superconducting Nanowire Single-Photon Detector (SNSPD)

Les superconducting nanowire single-photon detector (SNSPD) sont des photodétecteurs très sensibles au photon unique pour des longueurs d'ondes dans le domaine du visible jusqu'à l'infrarouge proche. Ils ont un faible taux de comptage dans le noir et une faible résolution en temps [Dorenbos_11].

En effet, la structure du SNSPD est composée d'une couche ultra fine (environ 5 nm) et ultra étroite (environ 100 nm) formant un fil supraconducteur de Nitrure de Niobium (NbN) comme le montre la Figure 1.6.

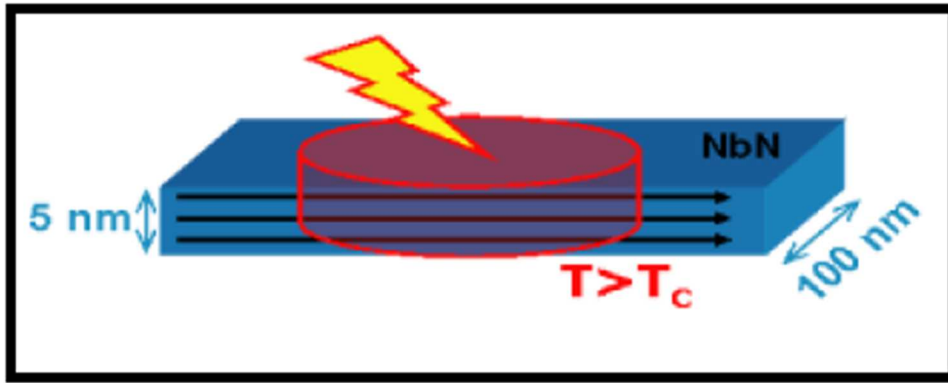


Figure 1.6 : Structure du SNSPD

Le principe de fonctionnement du SNSPD est le suivant : le nano fil est maintenu au-dessous de sa température critique supraconductrice et polarisé avec un courant continu proche mais inférieur à son courant critique supraconducteur. Lorsqu'un photon est absorbé, l'énergie correspondante excite la population des électrons conduisant à la formation d'un point chaud résistif à travers la largeur du fil. Une impulsion de tension est ainsi créée dans le circuit externe. Ce nano-fil peut être réalisé en forme de serpentín afin de couvrir une superficie de plusieurs micromètres carrés pour améliorer la détection lumineuse.

La Figure 1.7 montre une structure en serpentín du superconducting nanowire single-photon detector.

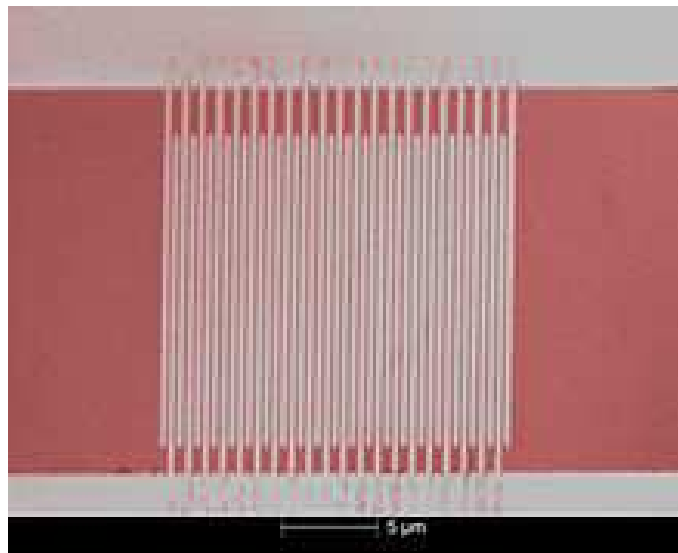


Figure 1.7 : Structure serpentín SNSPD

Les nano-fils reposent sur la supraconductivité, ce qui exige des faibles températures. Ce type de photodétecteur n'a pas eu un succès commercial à grande échelle et il est toujours en phase de recherche.

1.4 Conclusion

Dans ce chapitre, nous avons détaillé l'état de l'art des différents dispositifs utilisés pour détecter les très faibles niveaux lumineux. Les PMTs ont été le premier photodétecteur apte à détecter les photons uniques. Plus récemment, un nombre important de photodétecteurs à photon unique ont été mis au point comme nous l'avons montré tout au long de ce chapitre. Selon la nature de l'application, on choisit le détecteur approprié.

Dans le chapitre suivant, on va s'intéresser aux photodiodes à avalanche qui fonctionnent en mode Geiger et qui seront le cadre de ce travail de thèse. L'état de l'art des photodiodes à avalanche sera détaillé depuis leur apparition jusqu'aux évolutions les plus récentes.

Chapitre 2 *Les photodiodes à avalanche à photon unique ou SPAD*

2.1 Introduction

Au cours des dernières années, les dispositifs de comptage de photons ont été largement utilisés dans une large gamme d'applications : imagerie médicales, vision 3D, temps de vol (TOF), spectroscopie, imagerie quantique, cryptographie.

Un tourbillon d'activité a conduit au développement d'une famille de semi-conducteurs détecteurs de photon unique appelés SPAD (Single-Photon Avalanche Diode) dans un certain nombre de technologie CMOS pour différentes applications. Les CMOS SPADs ont montré leurs utilités dans des applications où le faible bruit, la haute résolution et une plage de dynamique élevée sont importantes. De plus grâce à la miniaturisation des procédés de fabrication CMOS, l'intégration de millions de pixels avec la capacité de détection du photon unique sur une seule puce devient possible.

Nous avons vu qu'une jonction pn, polarisée dans le mode Geiger, c'est-à-dire au-delà de sa tension de claquage, permet de réaliser la « détection du photon unique » grâce à des niveaux d'amplification bien supérieurs au mode classique. Il est important d'expliquer, comment il est possible de polariser une photodiode au-delà de sa tension de claquage, et de réaliser des taux de multiplication très élevés sans causer de dégâts à la jonction.

Dans ce chapitre, nous détaillerons l'effet d'avalanche dans les photodiodes à avalanches polarisées en mode Geiger et nous expliquerons les différentes structures de fabrication existantes actuellement permettant la production des SPADs. Ensuite, nous évoquerons le sujet des facteurs de mérites qui vont servir à caractériser les performances des SPADs. Pour finir ce chapitre, nous présenterons les différents types de circuits d'extinction (quenching circuit) qui vont éteindre l'avalanche et qui vont servir ensuite à restaurer

la polarisation de la photodiode pour qu'elle soit en mesure de détecter un nouveau photon.

2.2 Geiger mode Avalanche photodiode

Les SPADs sont essentiellement une jonction p-n polarisée en inverse au-dessus (en valeur absolue) de la tension de claquage. A cette polarisation, le champ électrique est très important (plusieurs centaines de KV/cm). Lorsqu'une photodiode à avalanche polarisée en mode Geiger est exposée à la lumière, elle absorbe un photon dans la région de multiplication, qui correspond à la zone de déplétion de la jonction. Une paire électron-trou est alors créée, séparée, puis accélérée à une très grande vitesse par le champ électrique important de la jonction. Une très grande quantité de nouvelles paires électron-trou est alors créée par impact ionisant dans le réseau cristallin de silicium et le courant d'avalanche augmente rapidement (\sim mA). Cette avalanche génère une impulsion de sortie qui permet de faire la détection d'un seul photon.

Le courant continue à circuler jusqu'à ce que l'avalanche soit arrêtée par abaissement de la tension de polarisation au-dessous de la tension de claquage. Le champ électrique devient alors plus faible et incapable d'accélérer les porteurs. Enfin la tension de polarisation doit être restaurée pour être en mesure de détecter un autre photon.

Le phénomène d'avalanche n'est pas un phénomène destructif pour la jonction, si la puissance dissipée : $P=V_R \cdot i$, est raisonnablement faible. Le coefficient de multiplication M est défini par la formule empirique de Miller donné par l'équation (2.1) où V_{br} est la tension de claquage et V est la tension de polarisation appliquée à la photodiode

$$M = \frac{1}{1 - \left(\frac{V}{V_{br}}\right)^m} \quad (2.1)$$

Comme on peut le constater d'après l'équation (2.1), lorsque V approche de la tension limite V_{br} (tension de claquage), le coefficient M devient très élevé [Pellion_08].

2.3 Zone de déplétion

La zone de déplétion aussi appelée zone de charge d'espace (ZCE) peut être séparée en deux zones (pour un SPAD) :

1. Zone de dérive (Drift region) : où les générations des porteurs prévus sont négligeables.
2. Zone de multiplication : la région où presque tout l'impact ionisant aura lieu.

Dans un SPAD où le champ électrique est uniforme à travers toute la zone de déplétion, la zone de multiplication occupe la quasi-totalité de la zone de déplétion. En revanche, dans le cas où le champ électrique est non uniforme, cette zone de multiplication constitue une portion de cette zone de déplétion. La distinction entre zone de multiplication et zone de dérive est importante pour différentes raisons. Pendant l'avalanche, le flux de porteurs qui passe à travers la zone de dérive crée une zone résistante de charge d'espace qui réagit comme une faible résistance. La propagation de l'avalanche ne se produira pas dans la zone de dérive ce qui donne une importance à la zone de multiplication qui détermine comment l'avalanche se propage. Il est important de signaler qu'une deuxième zone de déplétion existe entre la jonction N_well/p_substrat, mais dans cette zone les porteurs sont absorbés à l'intérieur du substrat et ne vont pas causer une avalanche [Fishburn_12].

2.4 Fabrication des SPAD

Les SPADs sont généralement conçus pour optimiser les paramètres qui seront détaillés dans la suite de ce chapitre :

- La probabilité de détection : qui représente la probabilité qu'un photon incident déclenche une avalanche.
- Bruit : la réduction de bruit principalement lié au taux de comptage dans le noir et aux porteurs piégés dans la zone de multiplication pendant l'avalanche (afterpulsing).
- Jitter : Correspond à l'incertitude en temps pour déterminer l'instant d'arrivée des photons.
- Taux de comptage : pour maximiser le nombre de photons que l'on sera en mesure de compter par unité de temps.
- Facteur de remplissage : pour maximiser la surface de la zone photosensible du pixel (surface de la photodiode/surface totale du pixel).
- Surface de la zone active : Certaines applications nécessitent une grande zone photosensible pour faciliter l'alignement de l'optique, d'autres exigent la miniaturisation des pixels pour l'intégration de la matrice.

Dans les sections suivantes, nous allons décrire les différentes implémentations pour fabriquer un SPAD. Les SPADs sont généralement de deux types : Reachthrough (structure épaisse) ou Planar SPAD (structure planaire). Dans la suite de cette section, on détaillera les avantages et les inconvénients de chaque structure [Finkelstein_14].

2.4.1 Reachthrough SPAD

Dans les structures Reachthrough, la majorité de l'épaisseur du substrat constitue la région d'absorption maximisant ainsi la probabilité d'absorption. Dans la première structure Reachthrough SPAD proposée par McIntyre [Dautet_93], une technologie personnalisée a été utilisée avec un substrat silicium p ultra-faiblement dopé spécialement développé pour générer la photodiode à avalanche Reachthrough comme le montre la Figure 2.1.

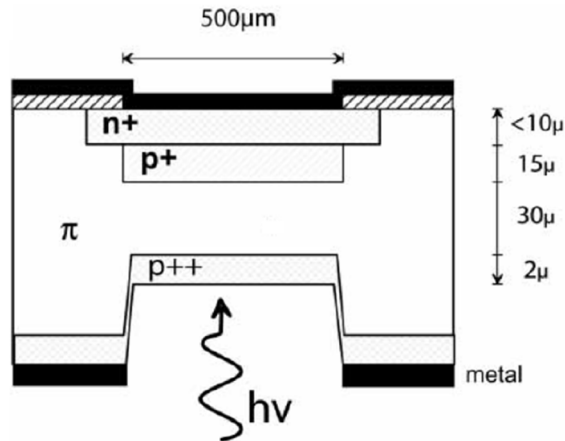


Figure 2.1 : Reachthrough SPAD McIntyre [Dautet_93]

Dans cette structure la région de déplétion varie de 20 à 100 μm avec une tension de claquage élevée (de 100 à 500V). Les effets de bords ont été évités en utilisant une couche p+ à enrichissement et en réduisant l'épaisseur du silicium sur les bords par gravure.

Une deuxième structure du SPAD Reachthrough a été proposée par [Zappa_04], visant l'optimisation de la première structure proposée par McIntyre. Ce dispositif est également fabriqué en utilisant un processus optimisé dédié à l'obtention d'une large zone active de 200 μm. Cette zone active est définie par une couche p+ et une diffusion profonde au centre comme le montre la Figure 2.2.

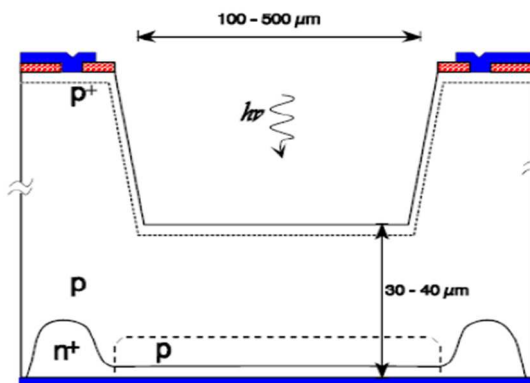


Figure 2.2 : Reachthrough SPAD [Zappa_04]

Grâce à sa large région de déplétion, cette structure montre une efficacité quantique de 30% à 500 nm et de 70% à 700 nm. Cette région de déplétion épaisse induit un Jitter supérieur à 300 ps ce qui est inacceptable pour plusieurs applications. En plus, ce dispositif exige une tension de polarisation entre 300 à 400 V ce qui nécessite des circuits spécialisés pour dissiper l'excès de puissance ce qui fait que la technologie dédiée à la fabrication de ce dispositif est coûteuse [Finkelstein_14].

Ensuite, plusieurs structures de SPAD Reachthrough ont été développées, mais la barrière de la forte tension de polarisation, le coût et la complexité des technologies dédiées ont incité les scientifiques à développer des nouvelles structures pour surmonter les limites des SPAD Reachthrough.

2.4.2 Les SPAD planaires

La principale différence entre la structure planaire et la structure Reachthrough est la réduction de la surface de la zone absorbante donc la zone de multiplication (diamètre et largeur). Les SPADs planaires ont été fabriqués depuis 1960 par Haitz [Haitz_63]. Ils utilisent des étapes de traitement qui sont semblables à celles utilisées pour la fabrication des structures à semi-conducteurs classiques, tel que le transistor métal-oxyde-semiconducteur (MOS). De part de leur topologie, les SPADs planaires permettent d'obtenir des caractéristiques différentes :

- Réduction du Jitter (la gigue temporelle) due aux diamètres réduits de la jonction.
- Réduction du temps mort due aux petites capacités de la jonction.
- Réduction de la consommation d'énergie due à des tensions de polarisation plus faibles.
- Réduction de la probabilité de détection due aux régions d'absorption étroites.
- Possibilité d'intégration à grande échelle avec des circuits de traitement sur la même puce.

La mise en œuvre de la première photodiode à avalanche proposée par Haitz utilise un anneau de garde comme le montre la Figure 2.3.

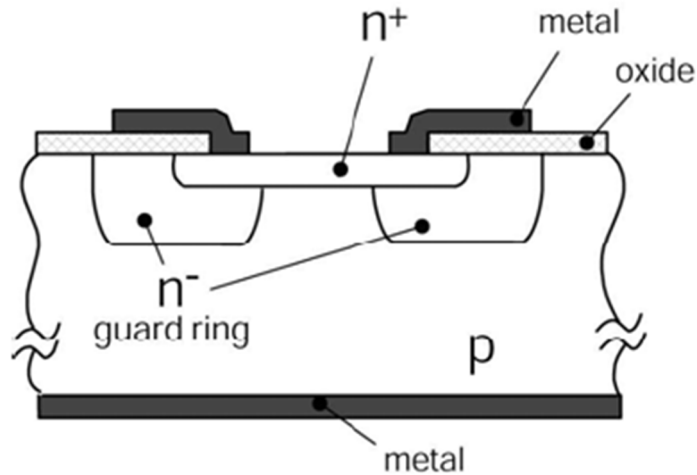


Figure 2.3 : SPAD planaire de Hartz [Goetzberger_63]

La structure est basée sur une jonction n^+/p substrat. Pour éviter le claquage prématuré lié aux effets de bord, une diffusion n^- est réalisée sur la périphérie de la jonction. Étant donné que la zone de déplétion résultante est plus étroite que celle des structures Reachthrough (épaisses), la tension de polarisation nécessaire est beaucoup plus basse (environ 32 V).

En 1989, Cova a introduit une nouvelle structure du SPAD planaire basée sur une couche p -épitaxiale avec un Jitter réduit comme le montre la Figure 2.4. Les photons incidents qui touchent la surface supérieure sont absorbés dans la couche épitaxiale de type p , Ils sont attirés par un faible champ électrique venant de la jonction n^+/p , où ils seront multipliés. Ce qui différencie la structure de Cova par rapport à celle de Hartz est l'utilisation d'une structure à double couches épitaxiales, formant ainsi deux diodes. La diode inférieure a pour rôle d'empêcher les électrons photogénérés à partir du substrat de diffuser dans la jonction formant le SPAD provoquant une augmentation de l'incertitude de synchronisation. La structure épitaxiale supérieure a été utilisée pour former une jonction ultra propre pour le SPAD. Afin de surmonter la résistivité de cette couche, une couche p^+ a été implantée en dessous de la couche épitaxiale supérieure. Enfin, une implantation d'une couche p^+ au centre de cette structure a pour rôle d'éviter le claquage prématuré de la structure [Cova_04].

La Figure 2.4 montre la structure produite par Cova.

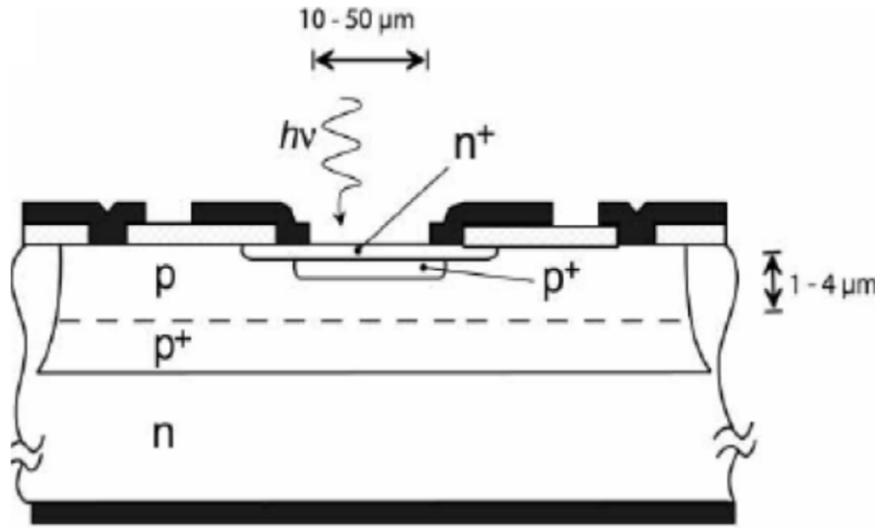


Figure 2.4 : Cova SPAD basé sur une double couche épitaxiales [Cova_04]

Cette structure présente une tension de claquage de 50 V, un Jitter de 55 ps et un taux de comptage meilleur que celui de la structure de Haitz.

La structure la plus utilisée jusqu'à nos jours est celle implémentée par Rochas en 1992. Cette structure a été implémentée en utilisant une technologie CMOS High Voltage. La Figure 2.5 montre une telle structure.

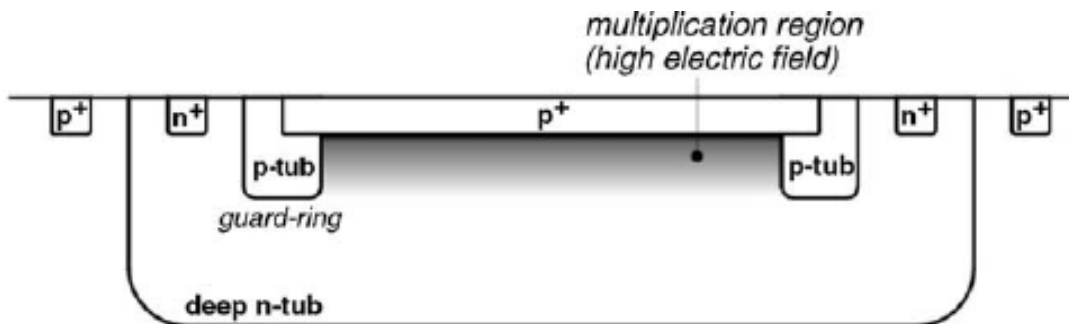


Figure 2.5 : CMOS SPAD HV de Rochas

Cette structure consiste en :

- Un SPAD circulaire.
- Une structure fabriquée avec une technologie HV-0.8μ CMOS processus.
- Une double jonction p+/deep n-tube/p-substrat.
- L'utilisation d'un substrat p et une tension d'alimentation allant de 2,5 à 50 V.
- Le claquage prématuré aux bords (PEB) a été évité en utilisant un anneau de garde par diffusion de deux p-tube relativement faiblement dopées qui vont éliminer les effets des coins en réduisant le gradient du champ électrique à la périphérie de la diode pour qu'on obtienne un champ électrique plus intense au centre et non pas sur les bords de la jonction p+/deep n-tub.
- La tension du claquage de la jonction p+/deep n-tube est de 25,5 V.

- Il y a une région de déplétion deep n-tub/p-substrat mais, dans cette région, les porteurs entrants sont balayés vers le substrat et ne causeront pas une ionisation [Niclass_05].

Une caractéristique des générations CMOS au-dessous de 250 nm est l'existence d'une couche n-buried (n-enterré) formée par implantation d'ion à haute énergie. Cette couche a un profil de dopage rétrograde de telle sorte que la surface a un faible dopage et que ce dernier augmente avec la profondeur. La structure de la Figure 2.6 montre le SPAD dans une technologie LV-CMOS 130nm. Cette technique utilise, pour réduire le taux de comptage dans le noir (DCR), un anneau de garde formé par une couche (shallow trench isolation STI) et une couche de passivation p-type qui vont créer un efficace anneau de garde contre les (PEBs). La couche p-type a pour rôle de minimiser le champ électrique entre le (STI) et la couche (n-well) diminuant ainsi la probabilité des (PEBs) [Gersbach_08].

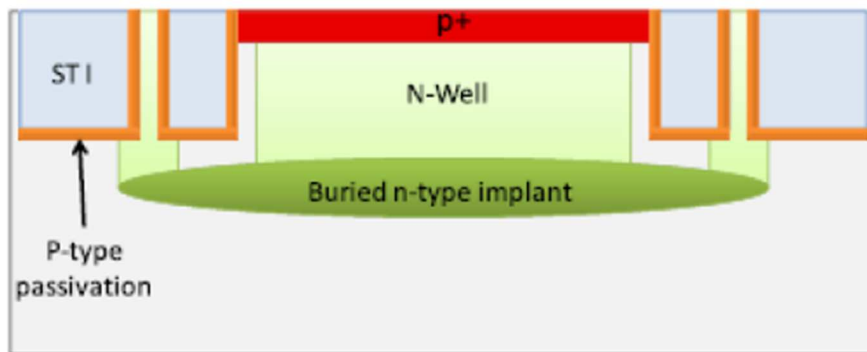


Figure 2.6 : STI-Bounded SPAD CMOS 130nm entouré par une couche de passivation (p-type) [Gersbach_08]

A cause des contraintes de design, seulement les SPADs octogonaux sont réalisables dans cette technologie. L'utilisation de cet anneau de garde réduit la probabilité pour les porteurs minoritaires d'entrer dans la région de multiplication ce qui réduit le taux de comptage dans le noir (DCR).

Le tableau (1.1) résume les principales différences entre la structure Reachthrough et la structure planaire

Table 1.1 : comparaison entre la structure épaisse et planaire [Finkelstein_14]

	Structure Reachthrough	Structure planaire
Taux de comptage dans le noir	>1000 / sec	~1000 / sec
Efficacité quantique	70%	40%
Tension de fonctionnement	(250-400) V	(10-40) V
Taille pixel	100 μm (sans le circuit)	50 μm (avec le circuit)

	d'extinction)	d'extinction)
Précision en temps	Faible (150 ps)	Bonne (<50 ps)
Taux de comptage maximal	10 Mcps	40 Mcps
Temps mort	40 ns	10 ns

2.5 Facteurs de mérites

Pour caractériser les performances des SPADs, plusieurs facteurs de mérites existent. Les principaux facteurs de mérites sont :

- La probabilité de détection de photon (PDP),
- Le taux de comptages dans le noir (DCR),
- La probabilité d'impulsion ultérieure (Afterpulsing probability),
- Le temps mort,
- Le Jitter en temps (temps de résolution),
- La probabilité de diaphonie,

Dans la suite de ce chapitre on détaille l'influence de ces facteurs sur l'évaluation des performances des SPADs.

2.5.1 Probabilité de détection de photon (PDP)

Pour détecter un photon, non seulement il doit être absorbé dans la région active du détecteur et interagir de façon à générer une paire électron-trou, mais il est également nécessaire que les porteurs réussissent à déclencher une avalanche.

La probabilité de détection de photon (PDP) est la probabilité qu'un photon incident heurte la partie active de la photodiode et déclenche une avalanche. En raison de la nature statistique de l'impact d'ionisation et la pénétration de la lumière dans le silicium, cette probabilité est toujours inférieure à un. Cette probabilité dépend fortement de l'efficacité quantique (QE) et la longueur d'onde (λ).

L'efficacité de détection de photon (PDE) est l'efficacité qu'un photon interagissant dans la partie active du SPAD déclenche une avalanche. L'équation (2.2) définit la PDE

$$PDE = QE.FF.TP \quad (2.2)$$

Avec :

Efficacité Quantique (QE) = probabilité qu'un photon absorbé produise une paire électron trou.

Fill factor (FF) = le facteur de remplissage donné par l'équation (2.3)

$$FF = \frac{\text{surface de la zone active}}{\text{surface du pixel}} \quad (2.3)$$

Triggering Probability (TP) = probabilité de déclenchement.

La Figure 2.7 montre la variation du PDE en fonction de la longueur d'onde pour un SPAD CMOS en prenant en compte l'excès de la tension de polarisation.

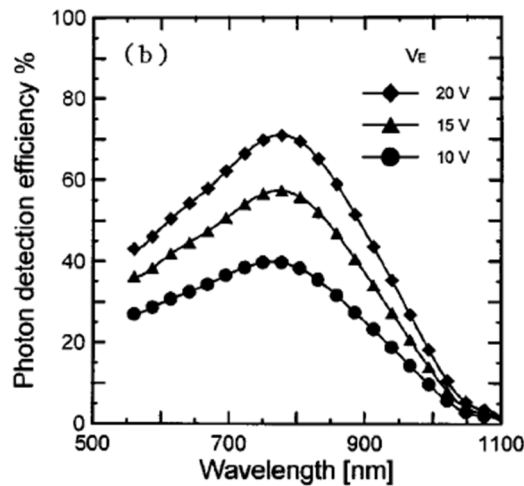


Figure 2.7 : Variation de la PDE en fonction de longueur de la longueur onde [Cova_96]

Comme le montre la Figure 2.7, la PDE augmente en fonction de la longueur d'onde pour atteindre un pic à 800 nm, ensuite cette PDE diminue progressivement pour atteindre un minimum à 1100 nm. On constate également qu'avec l'excès de la tension de polarisation V_E , la PDE augmente. Il faut signaler que la variation de la PDE donnée dans la Figure 2.7 est mesurée dans le cas du silicium, ce qui justifie ce pic à 800 nm.

2.5.2 Temps de résolution (le Jitter en temps)

Le temps de résolution représente la fluctuation statistique de l'intervalle de temps entre l'arrivée des photons sur le SPAD et l'impulsion de sortie. Ce Jitter est mesuré en utilisant la largeur à mi-hauteur (Full width at half maximum (FWHM)) donnée par la différence entre les deux valeurs extrêmes du temps d'arrivée des impulsions détectées en

sortie. En général, le temps de résolution est le facteur limitant la résolution temporelle du système de mesure.

Le Jitter dans un SPAD dépend principalement de deux paramètres : les variations de la profondeur des porteurs photogénérés (profondeur d'interaction). En effet, la position réelle d'absorption du photon dans le matériau suit une fonction exponentielle, et donc il y a des variations dans le temps pour lequel les porteurs dérivent hors de la zone d'absorption. Le Jitter dépend également d'un deuxième paramètre, c'est le processus d'accumulation de l'avalanche. Ce processus est généralement le mécanisme prépondérant, de l'ordre de la centaine de picosecondes, si par exemple, on dispose d'un champ électrique élevée et d'une zone de multiplication mince inférieure aux rapports des coefficients d'ionisations α/β (α pour les électrons et β pour les trous) donnés par les équations (2.4) et (2.5). Ainsi le Jitter est réduit en réduisant le bruit d'avalanche [Osaka_86] [Harrison_99].

$$\alpha(E) = \alpha_0 \cdot \exp(-[c_n/E]^{m_n}) \quad (2.4)$$

$$\beta(E) = \beta_0 \cdot \exp(-[c_p/E]^{m_p}) \quad (2.5)$$

La Figure 2.8 montre la détermination du temps de résolution en utilisant le FWHM

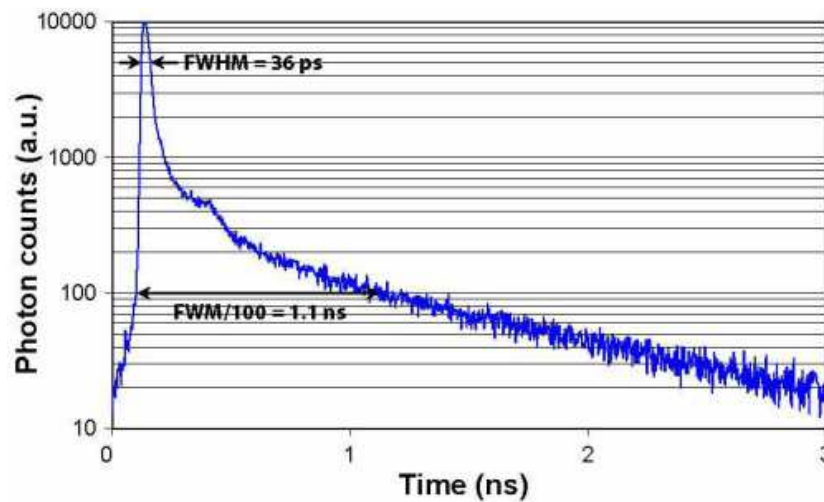


Figure 2.8 : Temps de résolution (FWMH) [Cheng_12]

2.5.3 Temps mort (Dead Time)

À la suite d'une avalanche, et pendant un temps court, le SPAD est incapable de détecter un autre photon. Ce temps mort est le temps où le SPAD est maintenu avec une polarisation en dessous (en valeur absolue) de la tension d'avalanche de la jonction, ainsi il est incapable de détecter un nouveau photon. Pour un bon fonctionnement du SPAD, il faut que le temps mort soit le plus petit possible. La Figure 2.9 représente le temps mort

dans un SPAD. Ce temps mort sera donc aussi grandement dépendant de l'architecture du circuit de quenching, circuit qui sera détaillé par la suite.

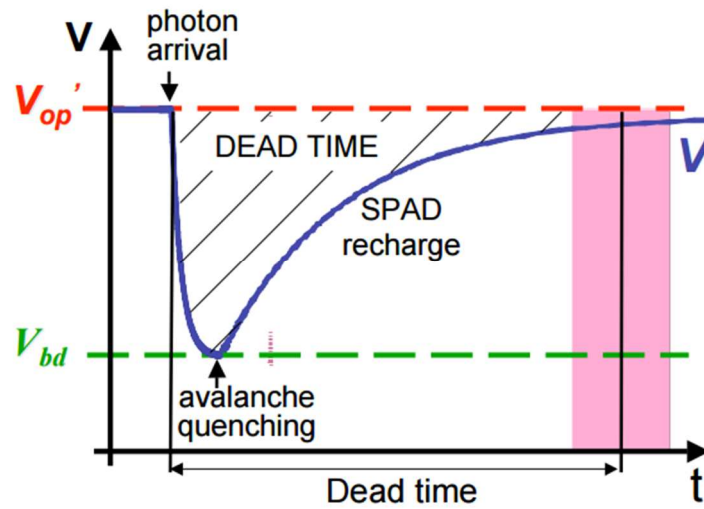


Figure 2.9 : Temps mort dans un SPAD

Comme on peut le constater d'après la Figure 2.9, il s'agit principalement du temps d'étouffement de l'avalanche et le temps de recharge de la photodiode pour qu'elle soit en mesure de détecter un nouveau photon.

2.5.4 Taux de comptage dans le noir (DCR)

Le comptage dans le noir représente les faux événements détectés par l'élément photosensible en absence de la lumière. Même en l'absence de photon, différents mécanismes génèrent des paires électron-trou. Ces paires ainsi générées produisent une avalanche parfaitement identique à une avalanche issue de la génération d'une paire par photon incident. D'un point de vue macroscopique, le comptage dans le noir représente une fuite dans la jonction ou un comptage de signal de lumière parasite.

Les principales sources de comptage dans le noir sont :

- La génération thermique : Dans les semi-conducteurs, la génération thermique des porteurs est attribuée à des transitions directe bande à bande (bande de valence vers bande de conduction) où via la théorie de Shockley-Read-Hall (SRH) [Hsu_10]. En général, une transition directe est peu probable, en particulier dans la bande interdite des matériaux tels que le silicium avec une largeur de bande de 1.12eV, même à des températures élevées. La théorie (SRH) illustre les pièges associés à la génération thermique dans quatre processus (capture d'électron/émission d'électron/capture du trou/émission du trou). La génération thermique de porteurs augmente avec la température.

- L'effet tunnel : la région de multiplication mince souffre d'un effet de tunnel élevé à cause du champ électrique élevé nécessaire à l'impact d'ionisation. Une région de multiplication épaisse limite l'effet tunnel et diminue le champ électrique exigé pour le claquage. Les inconvénients d'une telle région épaisse est le temps de résolution dégradé et le bruit excessif.
- Afterpulsing : les pièges à l'intérieur de la région de multiplication ont tendance à capturer des porteurs photogénérés lors d'une avalanche qui seront libérés aléatoirement plus tard et peuvent potentiellement re-déclencher une avalanche corrélée dans le temps avec l'impulsion originale. Ce phénomène est appelé afterpulsing (impulsion ultérieure). Le paramètre caractérisant cet effet est connu sous le nom d'afterpulsing probability (la probabilité d'Impulsion ultérieure), il est fonction de nombres de porteurs impliqués dans une avalanche, qui dépend à son tour de la capacité parasite du SPAD.

Le tableau (2.2) résume les dépendances de bruit dans le SPAD

Table 2.2 : Dépendance de bruit dans le SPAD

	Effet tunnel	Shockley-Read-Hall (SRH)	Afterpulsing
Corrélation (Avec une avalanche précédente)	Non	Non	Oui
Distribution	Poisson	Poisson	Exponentielle
Dépendance en tension	Très forte dépendance	Forte dépendance	Forte dépendance
Dépendance en température	Faible dépendance	Forte dépendance	Dépendance inverse

2.5.5 Probabilité de diaphonie (Crosstalk Probability)

Lorsque les SPADs se trouvent dans une matrice de pixels, ils peuvent être soumis à une diaphonie ce qu'on appelle "Crosstalk" comme le montre la Figure 2.10.

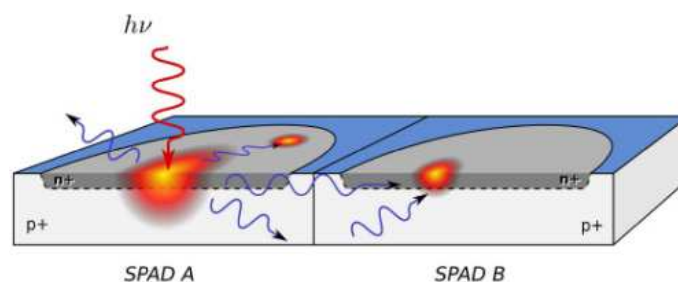


Figure 2.10 : La probabilité d'une diaphonie

Cette diaphonie peut être de nature optique ou électrique :

- Diaphonie optique : Durant l'avalanche certains photons peuvent être libérés en raison de l'électroluminescence, ces photons peuvent causer une avalanche dans les pixels voisins générant ainsi des impulsions parasites non liées à une excitation directe du pixel. Cette probabilité est beaucoup plus faible dans les matrices de SPADs intégrés par rapport aux matrices hybrides. Ceci est dû au fait que la capacité parasite de la diode dans la version intégrée est un ordre de grandeur plus petite que celle dans la version hybride, réduisant ainsi l'énergie dissipée lors d'un événement Geiger.
- Diaphonie électrique : au contraire, c'est une diaphonie produite par le fait que les photons absorbés au-delà de la jonction p-n en profondeur dans le substrat, génèrent des porteurs qui peuvent diffuser dans les pixels voisins.

2.6 Circuits d'étouffement ou de « quenching »

Comme on l'a vu précédemment, pour arrêter l'avalanche, il faut abaisser la tension de polarisation au-dessous de la tension de claquage dans une zone où le champ électrique devient incapable d'initier une avalanche. Le circuit qui effectue une telle opération est désignée par "quenching circuit" (circuit d'étouffement). Cette opération nécessite de l'électronique appropriée capable de :

1. Détecter l'augmentation du courant d'avalanche ;
2. Générer à la sortie une impulsion synchrone avec le changement de courant ;
3. Eteindre l'avalanche en abaissant la polarisation au-dessous de la tension de claquage ;
4. Remettre la polarisation au-delà de la tension de claquage pour qu'elle redevenue opérationnelle.

Pour le design du circuit d'étouffement, il faut prendre en considération plusieurs caractéristiques afin d'optimiser l'efficacité de ce circuit. Premièrement, une bonne conception du circuit d'étouffement est primordiale car elle influence directement le Jitter du SPAD. En plus, il faut que le circuit soit de la plus petite surface possible pour maximiser le taux de remplissage du pixel. En réduisant la surface du circuit d'étouffement, la consommation en puissance, la dissipation de chaleur, la diaphonie optique et le nombre total de charges de l'avalanche seront réduits à leur plus petite valeur [Boisvert_14]. L'afterpulsing est un paramètre fortement lié au circuit d'étouffement. Ce circuit contrôle le temps mort de la photodiode, c'est-à-dire la période où le SPAD est intentionnellement sous polarisé. Durant cette période les porteurs pris dans les pièges profonds ont une pro-

babilité plus élevée d’êtres relâchés avant que la photodiode soit de nouveau polarisée au-dessus de la tension de claquage [Bonanno_05].

Ils existent différents types de circuit d’étouffement, de type passif ou actif, ainsi qu’une recharge passive ou active. En se basant sur les différents paramètres cités dans le paragraphe précédent, il devient plus facile de comparer les avantages et les inconvénients des différentes architectures de circuits d’étouffement entre elles. Parmi les diverses architectures de circuits d’étouffement, on retrouve le circuit d’étouffement passif, le circuit d’étouffement actif, le circuit d’étouffement mixte et le circuit d’étouffement avec comparateur de courant.

2.6.1 Circuit d’étouffement passif

La façon la plus simple pour éteindre l’avalanche est d’utiliser un circuit d’étouffement passif (Passive Quenching Circuit, PQC). L’étouffement s’effectue principalement en utilisant une résistance (R_L) élevée en série avec le SPAD. Une fois l’avalanche détectée, le circuit d’extinction passif développe une chute de tension aux bornes de la résistance (R_L) qui va permettre l’extinction de cette avalanche comme le montre la Figure 2.11.

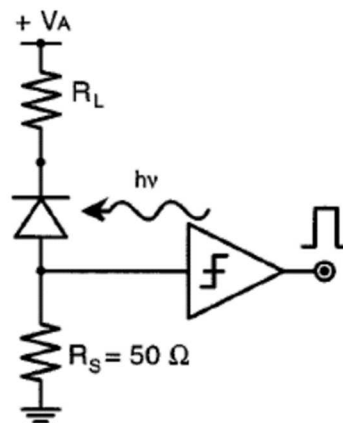


Figure 2.11 : Circuit d’étouffement passif [Cova_96]

Pour mieux comprendre le fonctionnement du PQC, on l’intègre dans un circuit combinant un modèle électrique de la SPAD donné par la Figure 2.12 et qui est constituée par :

- Une source tension fixe : $V = V_{br}$ représentant la tension de claquage,

- Une résistance R_d qui représente la résistance de charge d'espace de la diode, dépend de la structure du semi-conducteur,
- Un interrupteur en série avec R_d et V_{br} ,
- C_d : capacité de la jonction,
- C_s : capacité parasite,
- V_A : la tension de polarisation,
- R_L : Résistance d'étouffement,
- R_s : Résistance en série du SPAD.

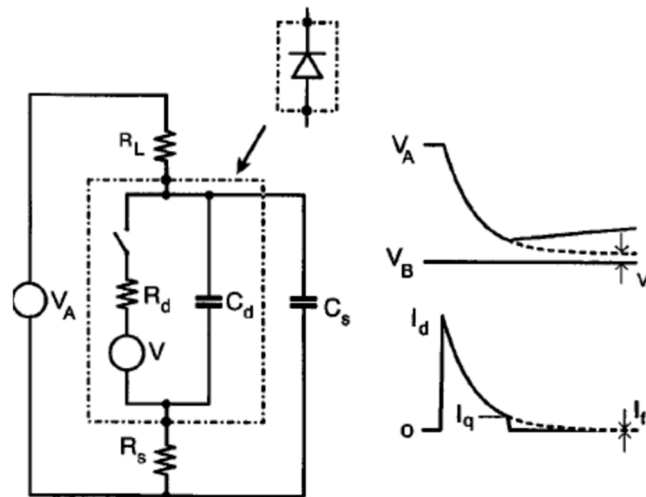


Figure 2.12 : Modèle électrique PQC [Cova_96]

Une polarisation V_A est appliquée, réglée à une valeur $V_A > V_{br}$ typiquement comme le montre l'équation (2.6) :

$$V_A = V_{br} + V_{ex} \quad (2.6)$$

Où V_{ex} représente l'excès à la tension de claquage V_{br} .

Dans cette configuration de polarisation, un seul photon créant une paire électron-trou suffit pour déclencher l'avalanche. Cela correspond à la fermeture de l'interrupteur dans le circuit équivalent de la diode, permettant ainsi au courant de circuler librement à travers les résistances R_L , R_D et R_s jusqu'à la masse et de l'accumulation des charges dans les capacités C_D et C_s qui veulent se décharger. Le courant d'avalanche crée une chute de tension aux bornes de la résistance R_L , ce qui diminue la tension aux bornes de la photodiode jusqu'à descendre en dessous de sa tension de claquage. Le rôle de la petite résis-

tance R_s est de créer une tension à ses bornes suffisante pour dépasser la tension de seuil et déclencher le comparateur de tension [Niclass_05].

On considère que la valeur de la résistance R_s est négligeable par rapport à R_L et que la résistance $R_L \gg R_d$. Le temps d'étouffement (T_q : quenching time) est donné par l'équation (2.7).

$$T_q = (C_d + C_s) \frac{R_d R_L}{R_d + R_L} \cong (C_d + C_s) R_d \quad (2.7)$$

D'après l'équation (2.7), ce temps dépend fortement des deux capacités $C_d + C_s$.

Une fois que la photodiode se retrouve sous polarisée (c'est-à-dire en dessous de sa tension d'avalanche), le courant d'avalanche disparaît, la tension aux bornes des résistances redevient nulle, repolarisant ainsi la photodiode au-delà de sa tension de claquage et la réinitialisant dans conditions initiales donc capable de détecter à nouveau un photon [Hsu_10]. L'ouverture de l'interrupteur correspond à la recharge du SPAD, le temps de recharge T_r est donné par l'équation (2.8)

$$T_r = R_L (C_d + C_s) \quad (2.8)$$

Ce circuit montre plusieurs avantages. Ses principaux atouts sont sa simplicité de conception et son faible nombre de composants ce qui signifie une petite surface occupée. En utilisant des composants passifs, ce type de circuit commence à agir dès le déclenchement de l'avalanche comparativement par rapport aux autres types de circuit qui peuvent nécessiter un délai. Les désavantages d'un tel circuit sont principalement le temps d'étouffement et le temps de recharge qui sont lents. Finalement, le temps mort de la photodiode n'est pas contrôlable avec cette architecture, il dépend de la valeur de la résistance R_L . Le circuit est de ce fait susceptible de redéclencher une avalanche, dû à un porteur se libérant d'un piège profond dans la photodiode (Aterpulsing), étant donné que le bruit post impulsion est plus élevé dans les instants qui suivent une avalanche [Tisa_08a].

2.6.2 Circuit d'étouffement actif

Afin de surmonter les inconvénients des circuits d'étouffement passifs (PQC) principalement le temps de recharge long et le temps mort qui n'était pas bien défini, un autre type de circuits a été introduit basé sur les composants actifs, il est connu sous le nom de circuit d'étouffement actif (Active-Quenching Circuits, (AQC)). Contrairement au PQC, l'électronique du circuit d'étouffement actif est plus élaborée et sa conception est plus complexe. La Figure 2.13 montre un exemple du circuit d'étouffement actif [Boisvert_2014].

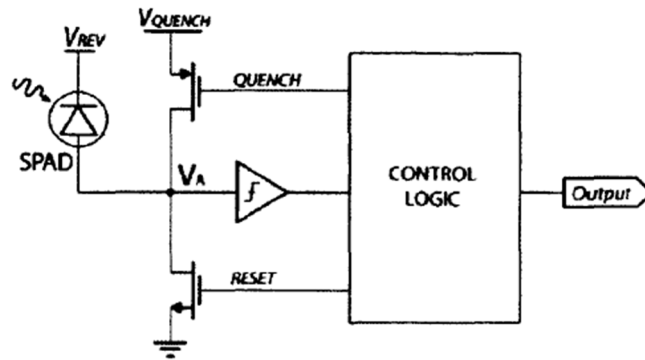


Figure 2.13 : Circuit d'étouffement actifs [Tisa_08a]

Le circuit d'étouffement actif de la Figure 2.13 est constitué d'un transistor PMOS pour l'étouffement de la photodiode et d'un transistor NMOS pour la réinitialisation et d'un comparateur de tension pour la détection d'un événement, en plus de toute la logique de contrôle. A l'état initial, la tension V_A à l'anode de la photodiode est de 0 V. la tension V_{REV} appliquée aux bornes de la photodiode est équivalente à la tension de claquage de la diode plus la tension d'excès désirée pour être en mode Geiger.

Lorsqu'un photon est absorbé dans la zone de déplétion, il déclenche une avalanche, la tension au nœud V_A augmente. Lorsque la tension V_A atteint le niveau fixé (déterminer par la tension de seuil du comparateur), le comparateur bascule et active le circuit de contrôle logique qui active à son tour le transistor d'étouffement par le signal Quench et impose la tension V_{QUENCH} au nœud V_A . La tension V_{QUENCH} , supérieure à celle de l'excès de tension appliquée à la photodiode, va faire en sorte de diminuer cette tension sous la tension de claquage pour étouffer l'avalanche. Suite à cet étouffement, le transistor d'étouffement reste activé jusqu'à ce que le temps mort soit établi par le circuit de contrôle logique. Après ce temps mort, le transistor d'étouffement est désactivé en même temps que le transistor de recharge est activé. Finalement, la tension au nœud V_A est remise à la masse repolarisant de nouveau la photodiode à une tension supérieure à celle de claquage.

Les principaux avantages des circuits d'étouffement actifs par rapport aux circuits passifs sont la rapidité d'étouffement de l'avalanche et la rapidité de la recharge de la photodiode. Un avantage supplémentaire du circuit d'étouffement actif par rapport au circuit d'étouffement passif est son contrôle du temps mort et l'ajustement de sa durée [Tisa_08a]. L'inconvénient de ce type de circuits est qu'ils nécessitent un laps de temps pour réagir au courant d'avalanche, pendant ce court délai de réaction, le risque d'avoir un afterpulsing est très élevé. Les circuits d'étouffement actifs sont beaucoup plus complexes que ceux des circuits passifs et exigent un nombre élevé de composants donc une surface occupée beaucoup plus importante.

2.6.3 Circuit d'étouffement mixte

Comme on a vu dans les deux sections précédentes, les circuits d'étouffement passifs et actifs présentent plusieurs avantages. Malgré ces avantages, ces circuits souffrent de nombreux inconvénients. Pour surmonter les inconvénients des circuits d'étouffement passifs et actifs, une nouvelle architecture mixte active et passive, qui a pour objectif de combiner les avantages des deux circuits, a été développée. Certaines architectures sont basées sur un étouffement passif et une recharge active, d'autres vont choisir un étouffement mixte passif et actif et il peut en être de même pour la recharge. Un exemple d'un circuit d'étouffement mixte est celui donnée par la Figure 2.14. Dans cet exemple, l'étouffement est mixte passif et actif combinés avec la recharge active.

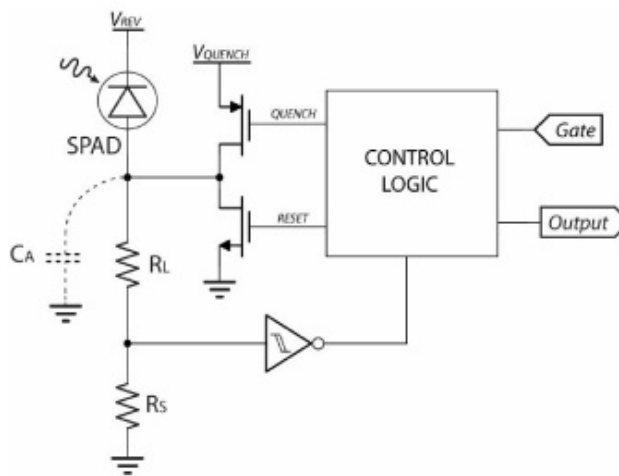


Figure 2.14 : Circuit d'étouffement mixte [Tisa_08b]

Lorsque l'avalanche est déclenchée, les résistances R_L et R_S démarrent la phase de l'étouffement comme les circuits d'étouffement passifs. Lorsque la tension aux bornes de la résistance R_S devient suffisante pour déclencher le comparateur, un signal est envoyé à l'étage logique de contrôle qui va activer le transistor PMOS d'étouffement. La tension V_{QUENCH} est alors imposée à l'anode du SPAD terminant l'étouffement comme dans les circuits actifs. La logique de contrôle désactive le transistor d'étouffement et active le NMOS de recharge une fois que le temps mort désiré est écoulé. A l'issue de la phase de recharge, le NMOS est désactivé et la photodiode est restaurée dans les conditions initiales en attente d'un nouvel événement.

L'avantage d'un circuit d'étouffement mixte est qu'il possède à la fois les avantages du circuit d'étouffement passif et ceux actif. Cette combinaison minimise le nombre de charges impliquées dans l'avalanche, la diaphonie et la puissance dissipée. En procédant à la recharge active comme dans les circuits AQC, ce type de circuit possède une recharge très rapide et un temps mort contrôlable. L'inconvénient d'un tel circuit est qu'il mélange

deux types de circuits ce qui augmente le nombre de composants nécessaires au fonctionnement, ce qui exige une surface importante occupée et rajoute une complexité dans sa conception.

2.6.4 Circuit d'étouffement avec comparateur de courant

Plusieurs architectures d'un circuit d'étouffement avec comparateur de courant ont été proposées pour remplacer les circuits d'étouffement actif. Une des architectures de ce type de circuit avec comparateur de courant est donnée par la Figure 2.15.

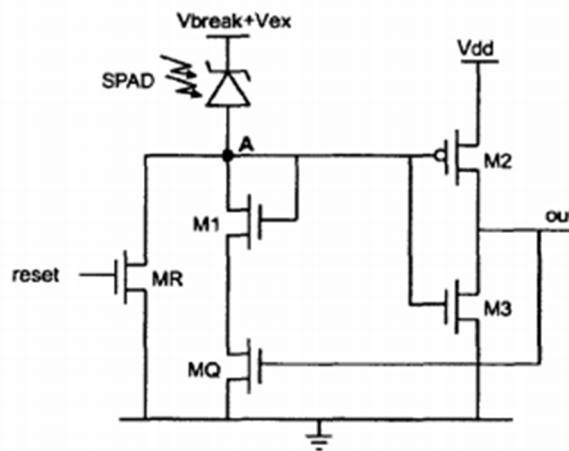


Figure 2.15 : Circuit d'étouffement avec comparateur de courant [Mita_08]

Ce circuit est composé d'un miroir de courant formé par les transistors M₁ et M₃. Le miroir de courant a pour rôle de contrôler le transistor M_Q en série avec le SPAD, afin d'augmenter sa résistance et d'étouffer la photodiode [Mita_08]. En même temps, un signal est envoyé à la logique du circuit pour indiquer l'arrivée de l'événement. La recharge se fait comme dans les circuits d'étouffement actifs.

En utilisant un comparateur de courant, le courant d'avalanche est détecté beaucoup plus rapidement qu'avec un comparateur de tension. En revanche, un nombre important de porteurs de charges circulent à travers la zone de déplétion à cause de la faible résistance initiale en série avec la photodiode et augmente lentement pendant l'avalanche. Dans ce type de circuit d'étouffement, le temps d'étouffement est plus long et le bruit plus conséquent par rapports aux autres architectures.

Pour conclure cette partie qui porte sur les circuits d'étouffements, il est important de préciser que le choix du circuit d'étouffement dépend fortement de l'application et de son cahier des charges. Le tableau (2.3) récapitule les avantages et les inconvénients des circuits d'étouffement cités précédemment.

Table 2.3: Avantages/inconvénients des circuits d'éteuffements

Circuit d'éteuffement	PQC	AQC	Mixte	Avec comparateur de courant
Temps d'éteuffement	+	-	+	-
Temps de re-charge	-	+	+	-
Charge d'avalanche	+	-	+	-
Bruit	+	+	+	-
Surface occupé	+	-	-	-
Complexité	+	-	-	+

2.7 Conclusion

Au début de ce chapitre, nous avons expliqué le mode de fonctionnement des photodiodes à avalanches en expliquant ce qui se passe dans la zone de déplétion de la photodiode. Plusieurs solutions permettent de fabriquer des photodiodes à avalanche. Dans ce chapitre, on a présenté les deux structures les plus courantes qui sont le SPAD Reachthrough et le SPAD planaire.

Pour caractériser les performances du SPAD, de nombreux facteurs de mérites existent. Il s'agit principalement de : la probabilité de détection de photon (PDP), le taux de comptage dans le noir (DCR), la probabilité d'impulsion ultérieure, le Jitter en temps, le temps mort, et la probabilité de diaphonie.

Pour étouffer l'avalanche dans le SPAD, il faut utiliser un circuit d'étouffement qui permet d'abaisser la tension de polarisation au-dessous de la tension de claquage. Le choix du circuit d'étouffement est important puisqu'il affecte le fonctionnement du SPAD. Dans ce cadre, plusieurs architectures ont été présentées dans ce chapitre principalement les circuits d'étouffement passifs, actifs, mixte ou avec comparateur de courant.

Le chapitre suivant décrira l'évolution des designs du SPAD dans les différentes technologies CMOS allant de 0.8 μm à 65 nm. Il introduit également les différentes architectures des circuits de lecture de SPAD.

Chapitre 3 *Etat de l'art du SPAD et des circuits de lecture*

3.1 Introduction

Depuis la réalisation du premier SPAD proposée par Haitz en 1960, les architectures n'ont cessé d'évoluer. L'évolution des designs est basée sur l'optimisation d'un ou de plusieurs facteurs de mérites décrits précédemment dans le chapitre 2. Dans ce chapitre, on va détailler plusieurs architectures de SPAD dans différentes technologies CMOS allant de $0,8 \mu$ à 65 nm tout en analysant leurs atouts et leurs limitations.

Pour exploiter l'information provenant du SPAD, il faut associer à la photodiode à avalanche un système de lecture capable de « lire » cette information. La particularité des circuits de lecture d'un SPAD est qu'ils doivent gérer des signaux impulsionnels, ce qui impose un cahier des charges qui les différencie des circuits de lecture des pixels conventionnels. En général, plusieurs paramètres permettent l'évaluation des performances des circuits de lecture. Il s'agit principalement de la consommation, la surface occupée, la linéarité, la résolution en temps, la dynamique, etc... Dans ce chapitre, on va détailler les différentes architectures de circuits de lecture qui existent en précisant leurs points forts et faibles, pour ensuite trouver le meilleur compromis afin de proposer notre circuit de lecture qui permettra, comme nous le verrons, d'optimiser la dynamique de lecture.

3.2 Evolution des structures des SPADs

Dans cette section, nous décrivons l'évolution des designs des SPADs dans les différentes technologies CMOS. A chaque design de SPAD proposé, certains paramètres ont été améliorés. Cette amélioration sur un des paramètres du SPAD affecte en général un autre paramètre, ce qui fait que chacun de ces designs présente des avantages, mais en même temps pour certain paramètre présente également des inconvénients. Dans cette partie, on

détaillera les différentes architectures existantes sans juger de leurs performances et optimisation qui sont fortement liées au cahier des charges.

3.2.1 SPAD CMOS en technologie haute tension (High Voltage HV)

La structure standard du SPAD en technologie CMOS High Voltage a été proposée et implémentée par Edoardo Charbon et son équipe à l'École Polytechnique Fédérale de Lausanne [Niclass_05]. Elle correspond à la structure développée à la section (2.4.2) basée sur une jonction p⁺/deep_n tub entourée d'un anneau de garde constitué d'une diffusion "p-tub".

Malgré les avantages qu'offre la structure conventionnelle du SPAD en CMOS HV, cette structure a montré de nombreuses limites. Les principales limitations de la structure conventionnelle sont :

- PDE : L'efficacité de détection de photon est relativement faible vu l'étroite région de déplétion et de multiplication,
- Le bruit : Il constitue une limitation de cette structure concernant les zones fortement dopées qui sont susceptibles d'avoir des centres de piègeages et de génération, ce qui va augmenter l'effet tunnel et donc le comptage dans le noir,
- L'Afterpulsing et le DCR seront aggravés en utilisant les couches p⁺ et n⁺.

Pour ces raisons l'équipe de Charbon a proposé une nouvelle structure de SPAD toujours dans une technologie CMOS HV et toujours en 0,35 μm . La Figure 3.1 montre la coupe de la nouvelle structure [Xiao_07].

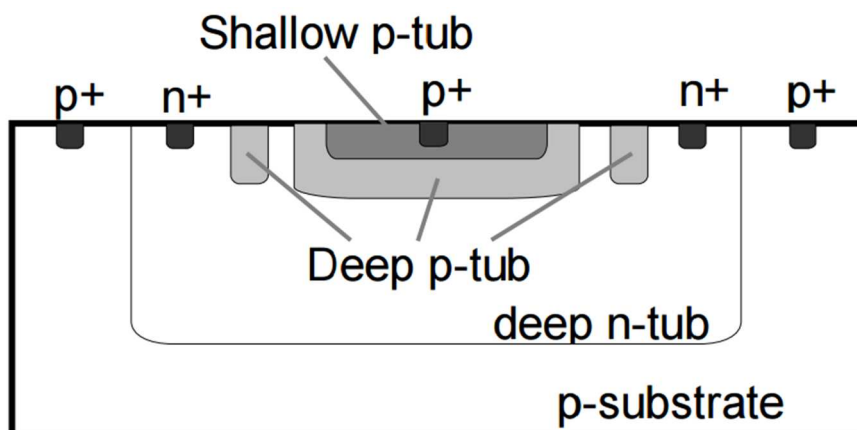


Figure 3.1 : Structure du SPAD p-/n- en technologie 0,35 μm CMOS HV

Cette nouvelle structure propose une architecture basée sur une jonction p-/n- à la place de p⁺/n- de la structure précédente. Grâce à cela, on augmente la profondeur de la jonction et la largeur de la zone de déplétion ce qui va augmenter le PDE.

3.2.4 SPAD CMOS en technologie 90 nm

Cette structure du SPAD réalisée en technologie CMOS en 90 nm est donnée sur la Figure 3.4.

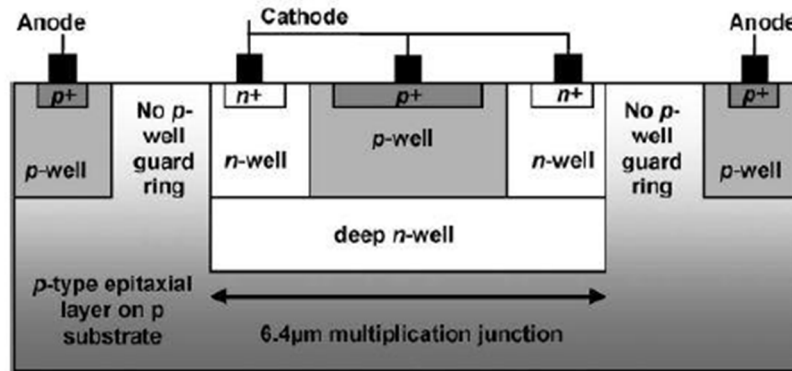


Figure 3.4 : SPAD CMOS en technologie 90 nm [Richardson_12]

Cette structure apporte une énorme amélioration au niveau de la détection proche de l'infrarouge par rapport aux anciennes structures de SPADs. Cette structure de SPAD CMOS 90 nm présente les caractéristiques suivantes :

- La région de multiplication est la jonction deep n-well/ p-type epitaxial,
- L'anneau de garde est formé par une implantation rétrograde de la couche deep n-well pour créer un fort gradient de dopage à la jonction plane avec la couche p-type epitaxial,
- La couche p-well est placée au-dessus de la partie active pour détecter les faibles longueurs d'ondes,
- L'efficacité de l'anneau de garde est dans sa capacité à réduire le champ électrique autour de la périphérie du dispositif [Richardson_12].

3.2.5 SPAD CMOS en technologie 65 nm

La dernière structure de SPAD qu'on présente dans le cadre de ce mémoire de thèse est le SPAD CMOS fabriqué en technologie 65 nm comme le montre la Figure 3.5 [Charbon_13].

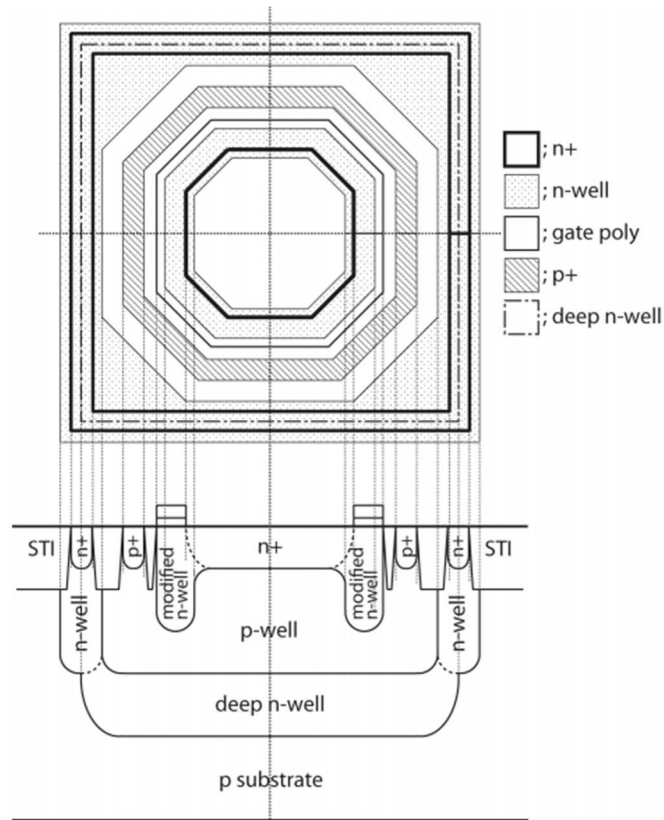


Figure 3.5 : Structure CMOS SPAD en technologie 65 nm (en bas) et le layout du SPAD CMOS 65 nm (en haut) [Charbon_13]

Le dispositif de la Figure 3.5 est un SPAD octogonal basé sur une jonction n+/p-well. Pour éviter le claquage prématuré dans cette structure, la jonction est entourée par un anneau de garde constitué d'une couche n-tube. Dans cette structure, l'avantage est qu'on peut boucler le circuit en entier sur l'anneau de garde grâce à la technologie CMOS 65 nm.

Les tests sur la structure SPAD CMOS 65 nm ont montrés les caractéristiques suivantes :

- Le taux de comptage dans le noir (DCR) est de $1,5 \text{ KHz}/\mu\text{m}^2$ à un excès de tension de polarisation $V_{ex} = 200 \text{ mV}$,
- L'afterpulsing est inférieure à 1 % pour un temps mort supérieure à $5 \mu\text{s}$,
- Le Jitter en temps est mieux que 235 ps .
- Le pic de la probabilité de détection de photon (PDP) est à 450 nm à un excès de tension de polarisation $V_{ex} = 200 \text{ mV}$.

Pour toutes ces caractéristiques, ce dispositif est très performant pour les applications exigeant une faible consommation et une très haute résolution comme l'imagerie 3D ou la spectroscopie [Charbon_13].

3.3 Les mécanismes de lectures des SPADs

Quand le SPAD détecte un photon et que ce photon déclenche une avalanche, la circuiterie associée au SPAD réagit pour lire cette information. Les mécanismes de lecture les plus répandues dans le cadre d'un SPAD sont « l'Event driven readout » et le « Sequential readout ». Dans la suite de cette section, on détaillera le principe de fonctionnement de chaque mode de lecture.

3.3.1 Event driven readout

Dans l'architecture Event driven readout, le pixel est lu uniquement lorsqu'un événement est survenu, en l'occurrence, la détection d'un photon. Un exemple d'une lecture en mode Event driven readout est donné par la Figure 3.6.

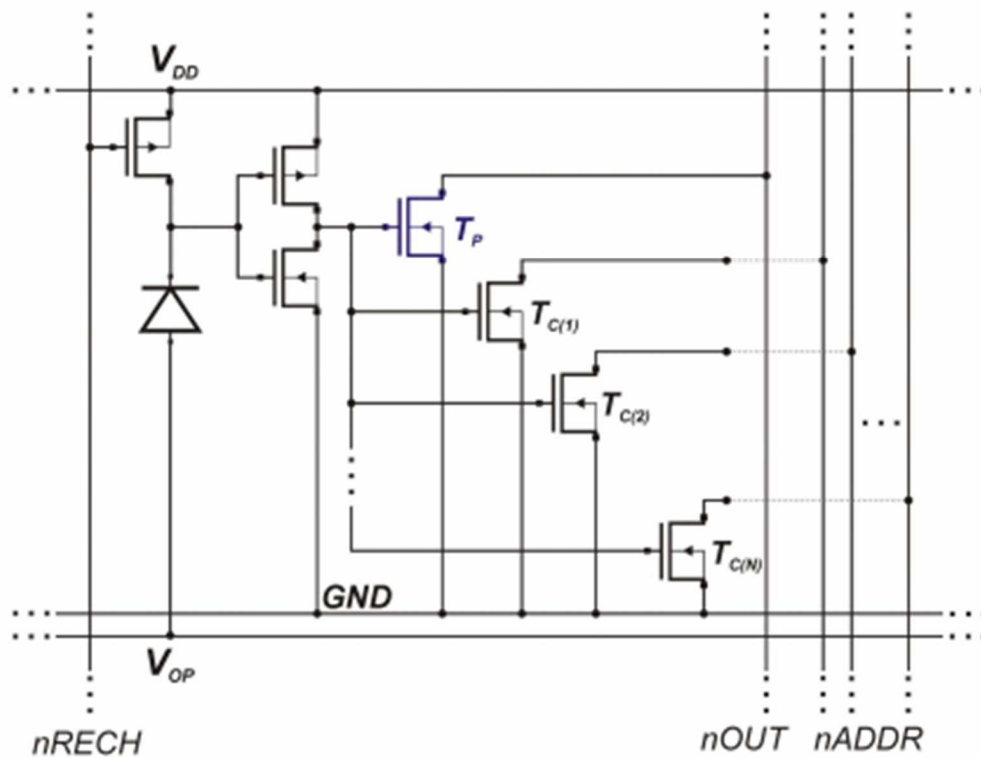


Figure 3.6 : Le principe d'un Event driven readout [Niclass_06]

Le mécanisme utilisé dans cette architecture est un mécanisme de signalisation par lequel le pixel communique l'apparition de l'événement à l'unité de commande.

C'est une architecture à faible puissance puisque le pixel est lu seulement quand il a une nouvelle donnée suite à la détection du photon. Une telle architecture exige un matériel capable d'intégrer un mécanisme de signalisation rapide, donc une résolution élevée [Lodha_12]. Une des techniques qui utilise le mode Event driven readout est le temps de vol (TOF). En effet, cette technique est basée sur un traitement individuel des photons qui

permet de déterminer instant par instant l'impact de chaque photon. On utilise souvent le TOF quand on dispose d'une application à faible taux de comptage.

3.3.2 Sequential readout (lecture séquentielle)

Ce mécanisme de lecture est basé sur une lecture en série de chaque pixel de la matrice. La Figure 3.7 montre les deux architectures possibles pour la lecture séquentielle : la lecture par pixel et la lecture par colonne.

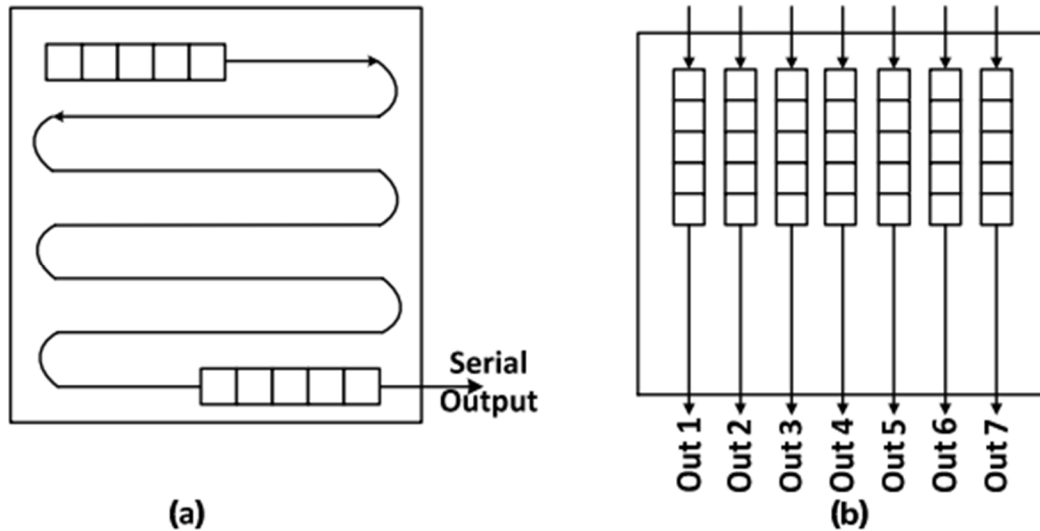


Figure 3.7 : Lecture séquentielle : lecture par pixel (a) lecture par colonne (b) [Kumar_12]

Le problème de la lecture par pixel est qu'elle est lente, mais son avantage réside dans son faible coût du fait qu'elle ne nécessite que d'une seule sortie pour effectuer la lecture. La lecture par colonne nécessite un nombre de sorties beaucoup plus élevé que celui de la lecture par pixel. L'avantage d'une telle lecture est le parallélisme qui conduit à une lecture plus rapide que la lecture par pixel.

Il faut préciser que la consommation en énergie pour les deux architectures de lecture séquentielle est plus élevée que celle pour la lecture Event driven, étant donné que la lecture de tous les pixels est indépendante de l'arrivée d'un événement [Kumar_12].

3.4 Etat de l'art des circuits de lecture

Plusieurs modes de lectures des événements existent pour les SPADs. On a cité à la section (3.3.1) les techniques de temps de vol (TOF). Dans le cadre de ce travail, la technique retenue est basée sur l'intégration (comptage de photon). Comme dans les cas des pixels APS, une fois que le temps d'intégration fixé s'est écoulé, le nombre total de photon ayant touché un pixel sera enregistré.

La plupart des circuits de lecture de SPAD sont conçus pour réaliser des fonctions de comptage de photons. La réalisation d'une fonction de comptage dans un intervalle de temps donnée est simplement réalisable avec des circuits de lecture basés sur un compteur numérique car ils fournissent un signal de sortie numérique. Bien que la solution entièrement numérique réponde aux exigences de robustesse et de facilité de mise en œuvre, la zone occupée dans le pixel par ce système devient extrêmement grande, ce qui réduit considérablement le facteur de remplissage. Pour cette raison, on s'intéresse, dans le cadre de cette thèse, aux architectures analogiques.

Dans cette section, on étudiera les différentes architectures analogiques de circuits de lectures de SPAD existantes, pour ensuite proposer notre circuit de lecture basé sur un compteur analogique dans le chapitre 6.

3.4.1 Compteur analogique : source de courant et capacité

Le premier circuit de lecture basé sur un compteur analogique présenté dans cette section est formé d'une capacité et d'une source de courant comme le montre la Figure 3.8.

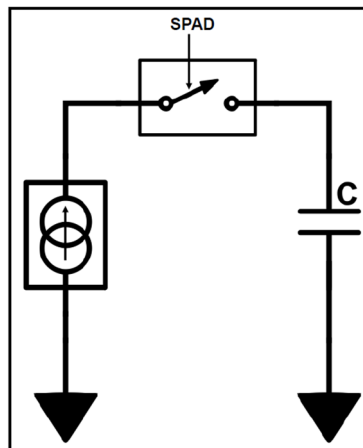


Figure 3.8 : Compteur analogique basé sur la recharge d'une capacité

L'impulsion provenant du SPAD, qui a une durée fixée qui représente le temps mort, contrôle la source de courant commutée qui décharge la capacité C. Ce qui fait qu'à chaque fois que le SPAD envoie une impulsion, la tension aux bornes de la capacité C augmente selon l'équation (3.8).

$$V_c(n) = \frac{n \cdot I \cdot \tau_{SPAD}}{C} + V_0 \quad (3.8)$$

Où V_c est la tension aux bornes de la capacité qui est une fonction du nombre de comptage n , I est le courant de la source de courant contrôlable, τ_{SPAD} est le temps mort imposé par le circuit du SPAD et V_0 est la tension initiale aux bornes de la capacité C.

Un des problèmes récurrents dans le design électrique des compteurs analogiques est de conserver une linéarité à travers une large gamme de comptage. En théorie, un compteur analogique comme celui de la Figure 3.8 ne souffre pas de problème de non linéarité car il est capable de fournir des pas de tension constante. En réalité, à chaque fois que le SPAD produit une impulsion, la tension aux bornes de la capacité augmente d'un pas de tension. A cause du bruit, dès que le comptage augmente, la taille du pas de tension commence à se réduire. Ainsi, le taux de comptage maximal que le compteur analogique peut atteindre est limité par la non linéarité maximale tolérable (saturation du générateur du courant). La dynamique du compteur sera limitée par ces phénomènes de saturation.

Pour surmonter le problème de la faible dynamique du compteur analogique à simple étage, une architecture à double étage a été proposée par [Palubiak_11] comme le montre la Figure 3.9.

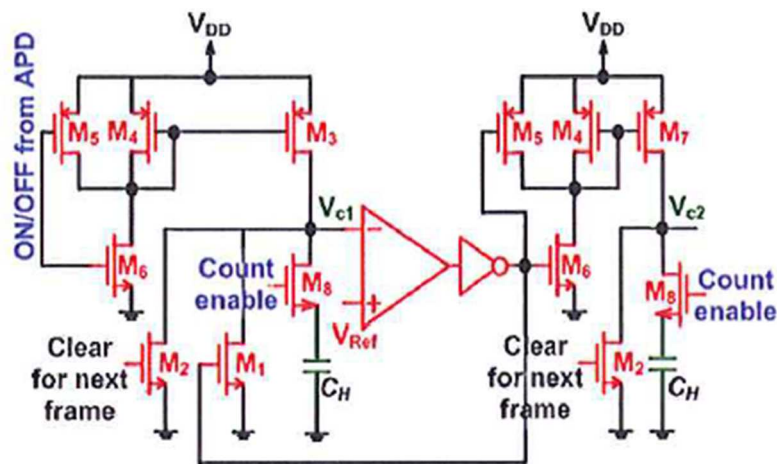


Figure 3.9 : Compteur analogique double étage [Palubiak_11]

Dans ce design, la source de courant a été remplacée par le design de la Figure 3.9. Les transistors M_4 , M_6 et M_3 composent la source de courant du premier étage. La source de courant du deuxième étage est composée par les transistors M_4 , M_6 et M_7 . Le transistor M_6 contrôlé par l'impulsion du SPAD, permet l'incrément du compteur. Les transistors M_5 ont été ajoutés pour assurer le blocage complet des transistors M_3 et M_7 . Lorsque le 1^{er} étage du compteur a atteint sa valeur maximale, le comparateur autorise l'incrément du 2^{ème} étage du compteur tout en réinitialisant le 1^{er} étage. Avec ce double étage, la dynamique du compteur est donnée par l'équation (3.9).

$$N_{tot} = N1.N2 \tag{3.9}$$

Où $N1$ et $N2$ représentent la dynamique de comptage du premier et du deuxième compteur respectivement. Les auteurs qui ont proposé cette solution affirment atteindre une dynamique de 900, soit une dynamique de 30 par étage.

3.4.2 Compteur analogique : Transfert de charge

Une autre solution pour intégrer un compteur analogique est d'utiliser un circuit basé sur un transfert de charge entre 2 capacités. La première capacité C est préchargée à une tension V_e . Une impulsion issue du SPAD ferme l'interrupteur $S1$ qui permet de charger la capacité KC avec la charge stockée dans la capacité C . Ce circuit sera détaillé dans le chapitre 5, en particulier la non linéarité inhérente à cette architecture ainsi que la dynamique espérée qui sont fortement dépendantes du rapport des capacités K . La Figure 3.10 montre le schéma de principe du comptage analogique à transfert de charge.

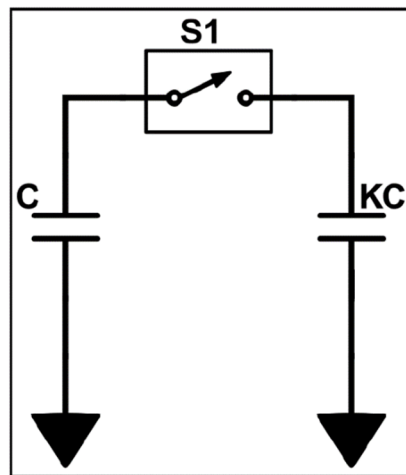


Figure 3.10 : Principe du compteur à transfert de charges

Pour réaliser électriquement le modèle du compteur analogique de la Figure 3.10, nous verrons que le rapport K entre les deux capacités doit être relativement important. Une solution proposée est donnée en Figure 3.11 où la première capacité est composée de la capacité parasite C_p et une capacité NMOS $CM12$ importante. La Figure 3.11 montre le schéma électrique du compteur analogique.

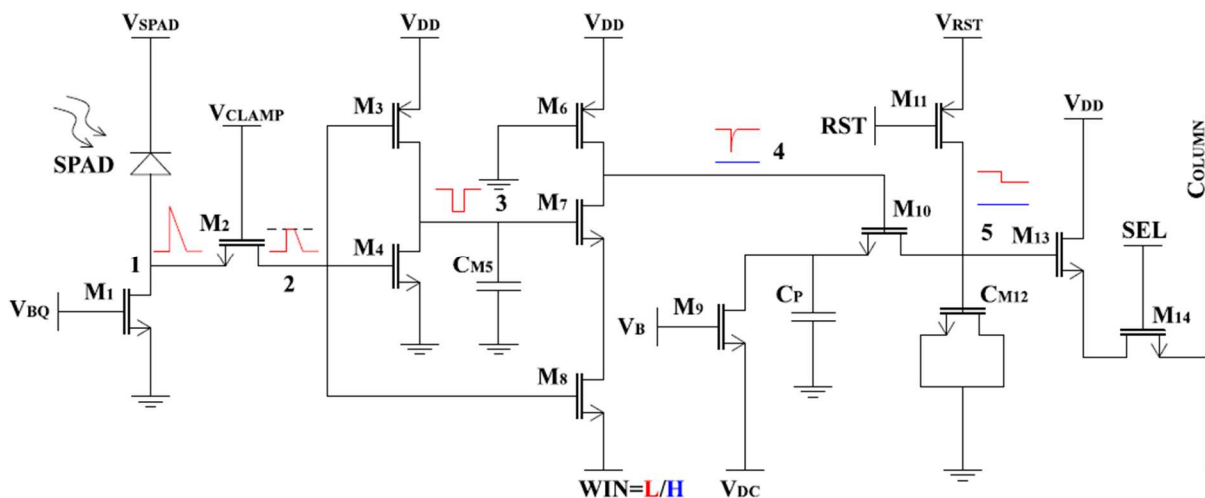


Figure 3.11 : schéma électrique du circuit de lecture basé sur un transfert de charge [Ekaterina_14]

Dans le circuit de la Figure 3.11, l'avalanche est éteinte passivement. Au début du cycle de fonctionnement, la capacité C_{M12} est remise à la tension V_{RST} . Le transistor M_9 fournit un courant faible pour charger la capacité C_P à la tension de référence V_B . La capacité C_P est complètement déchargée dans C_{M12} pendant le temps mort du SPAD.

L'interrupteur MOS M_{10} est exploité pour permettre le transfert de charge entre les capacités C_{M12} et C_P , ce qui prend approximativement une centaine de picoseconde. L'impulsion arrivant au nœud 4 est suffisamment large pour permettre la fermeture du transistor M_{10} dans le temps exigé a un transfert de charge complet.

A chaque fois qu'une avalanche est déclenchée, l'interrupteur MOS composé par le transistor M_{10} se ferme. La tension au nœud 5 diminue d'un ΔV selon l'équation (3.10).

$$\Delta V = \frac{C_p}{C_p + C_{M12}} (V_{out} - V_B) \quad (3.10)$$

D'après l'équation (3.10), le ratio entre les capacités C_P et C_{M12} déterminera le nombre maximal d'impulsion que le circuit peut compter. Le rapport K entre C_P et C_{M12} sera choisi de façon à avoir un ΔV suffisant pour ne pas être gêné par le bruit, et d'autre part suffisamment petit pour garder une dynamique importante. Un étalonnage est exigé pour estimer le nombre actuel du compteur venant du pixel, et ceci est à cause de la non linéarité dont la tension de sortie dépend fortement.

Toutes les simulations qui ont été faites dans le cadre de ce compteur analogique ont porté essentiellement sur l'influence de la capacité C_{M12} sur la largeur du pas de tension ΔV et sur la non linéarité du circuit. Les simulations ont montré que pour obtenir un ΔV de 8 à 18 mV, la valeur optimale de la capacité C_{M12} varie de 200 à 400 fF alors que la capacité C_P possède une valeur de l'ordre du fF, ce qui fait un rapport K d'environ 200 [Ekaterina_14]. Nous proposerons, aussi dans le chapitre 5, une architecture de lecture qui s'affranchi de la non linéarité du compteur analogique tant que ce dernier présente une caractéristique monotone.

3.5 Conclusion

Ce chapitre était consacré à l'étude de l'état de l'art du SPAD et des circuits de lecture associée au SPAD. Dans la première partie, on a décrit au fil du temps les évolutions que les SPADs ont subis et qui ont touché leurs designs dans différentes technologies CMOS. Les SPADs en technologie 0,35 μm ont un design différent que ceux des technologies plus récentes. Les modifications sur le design ont pour objectif d'améliorer un ou plusieurs facteurs de mérites détaillés, quant à eux, dans le chapitre 2. Actuellement, les technologies CMOS utilisent des architectures circulaires ou orthogonales pour éviter le claquage prématuré de la jonction.

La deuxième partie était consacrée à l'étude des architectures des systèmes de lecture des SPADs. La première architecture pour lire un SPAD est connue sous le nom de Event driven readout dans laquelle, le pixel est lu uniquement lorsqu'un événement est survenu. La deuxième lecture possible pour un SPAD, est une lecture séquentielle qui nous propose soit une lecture par pixel ou une lecture par colonne.

A la fin de ce chapitre, nous avons détaillé deux designs électriques de circuit de lecture du SPAD basé sur des compteurs analogiques pour lesquelles on a présenté les performances annoncées par les auteurs. Ces designs ont permis de nous positionner par rapport à notre cahier de charges et constituent un point de départ afin de proposer une architecture électrique qui respecte le plus ce cahier de charge.

L'étude de l'état de l'art du SPAD et des circuits de lectures constitue une introduction au chapitre suivant, dans lequel on propose une simulation physique de la structure du SPAD de Hartz proposée par la technologie CMOS HV en 0,35 μm , et un modèle haut niveau du pixel comptant le SPAD et son circuit de lecture.

Chapitre 4 *Modèle physique et haut niveau des SPADs*

4.1 Introduction

La conception des nouvelles générations de détecteurs de photons basées sur une jonction PN polarisée en inverse au-dessus de la tension d'avalanche (en valeur absolue) constitue une phase importante pour les systèmes nécessitant une très grande sensibilité de détection. Les outils de C.A.O (Conception Assistée par Ordinateur) offrent la possibilité de modéliser physiquement ces jonctions et ensuite de les simuler électriquement afin de mieux comprendre le mécanisme d'avalanche. Cette modélisation physique constitue une étape cruciale dans la prédiction du comportement du système et donc dans la conception du dispositif.

Dans ce chapitre nous présenterons l'outil de modélisation physique et électrique Silvaco, grâce à qui nous modéliserons physiquement la jonction PN et ensuite nous la simulerons électriquement. Nous proposerons dans ce chapitre également un modèle haut niveau de la photodiode. Ce modèle prend en considérations tous les facteurs de mérites que nous avons détaillé dans le chapitre précédent et représente un premier modèle haut niveau existant de la photodiode à avalanche polarisée en mode Geiger. L'intérêt de ce modèle est qu'il permet la simulation de la photodiode et du circuit de lecture au niveau système. Ce circuit de lecture sera le sujet principal des chapitres suivants.

4.2 Modèle physique et simulation électrique

4.2.1 Présentation de l'outil Silvaco

Silicon Valley Company (Silvaco) est le premier fournisseur mondial de logiciels TCAD (Technology Computer Aided Design) qui est une branche de la CAO électronique visant à modéliser la fabrication des dispositifs à semi-conducteurs et leur fonction-

nement. Silvaco offre à ces utilisateurs le Virtual Wafer Fab (VWF) qui représente un ensemble d'outils de simulation et d'outils interactifs permettant la conception et l'analyse de la plupart des dispositifs semi-conducteurs.

Les modules de VWF sont principalement des outils interactifs et des outils de simulations.

Les outils interactifs sont :

- DECKBUILD : interface avec les simulateurs Athena et Atlas,
- TONYPLOT : outil d'analyse graphique 1D et 2D des résultats des simulations,
- DEVEDIT : outil permettant de créer une structure de composant à partir de données fournis par le fondeur (dimensions, dopages...).

Les outils de simulation sont :

- ATHENA : simulateur de procédés technologiques qui permet de simuler les différentes étapes effectuées en salles blanches et ainsi d'obtenir la structure du composant (couches constitutives, dimensions, géométrie) avec son profil de dopage.
- ATLAS : simulateur de dispositifs semi-conducteurs qui permet d'obtenir leurs caractéristiques Courant - Tension et autres paramètres électriques (statiques ou dynamiques).

La Figure 4.1 donne un organigramme du logiciel SILVACO.

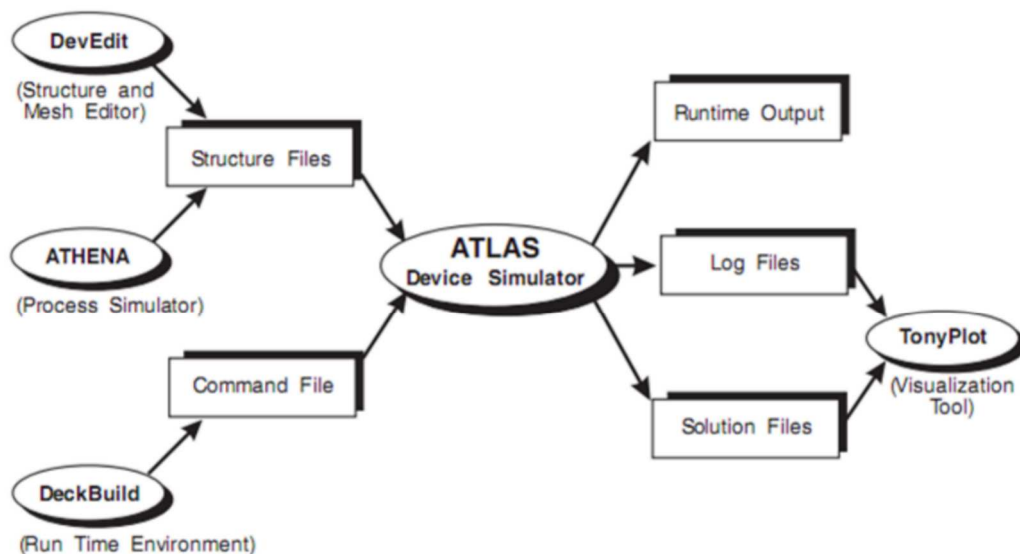


Figure 4.1 : Organigramme du logiciel SILVACO

Les entrées et sorties sont des fichiers. Il faut définir dans un premier temps la structure du composant, ensuite il faut préciser la façon dont on souhaite le simuler. Une fois que la simulation est accomplie, le simulateur génère un fichier technique qui est un compte rendu du déroulement de la simulation, un fichier résultat pour les courants ou tensions aux électrodes, un ou plusieurs fichiers résultats où sont consignées les valeurs des différentes grandeurs physiques aux différents points de la structure. Comme on peut le constater d'après la Figure 4.1, le cœur de Silvaco est le simulateur Atlas qui va permettre de réaliser les simulations électriques.

4.2.2 Structure PN sans anneau de garde

Comme cité précédemment, la modélisation physique constitue une étape importante dans la prédiction du comportement du dispositif. L'objectif de cette section est de modéliser physiquement la photodiode à avalanche puis de la simuler électriquement pour voir l'intérêt de l'utilisation d'un anneau de garde.

On a testé dans un premier temps la structure basée sur une jonction n+/p_substrat (la structure proposée par [Haitz_63]). L'objectif est d'éviter l'anneau de garde qui occupe une place importante (10% en moyenne pour une photodiode de 30 μm de diamètre) [Jradi_10]. La Figure 4.2 montre la distribution du champ électrique dans toute la structure en polarisant la jonction en inverse avec une polarisation supérieure à la tension de claquage ($V \gg V_{br}$) pour qu'on soit sûr que la photodiode soit dans le mode Geiger.

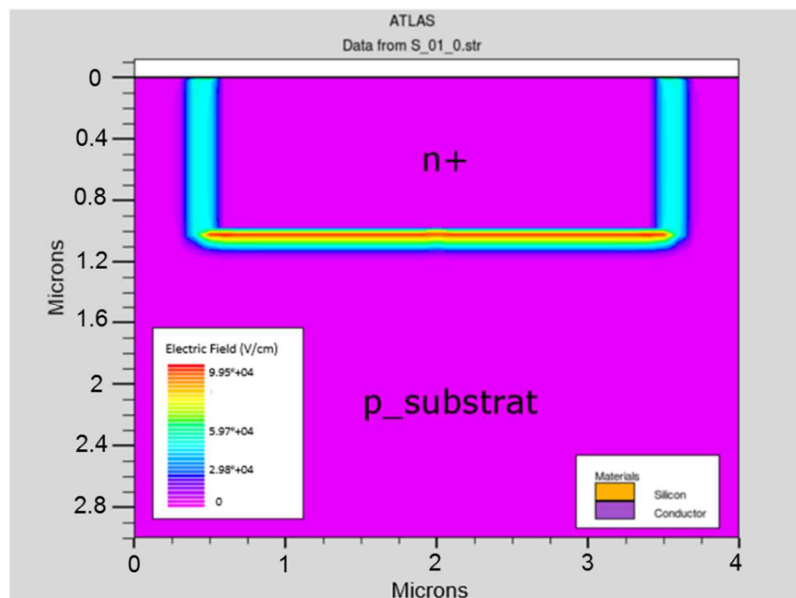


Figure 4.2 : distribution du champ électrique dans la structure n+/p_substrat

Comme on peut le constater d'après la Figure 4.2, le champ électrique élevé se concentre à la surface de contact de la jonction n+/p_substrat et aux coins. La simulation élec-

trique en utilisant l’outil de simulation ATLAS nous permet d’obtenir les caractéristiques courant-tension de cette jonction n+/p_substrat. Cette courbe I(V) est alors donnée par la Figure 4.3.

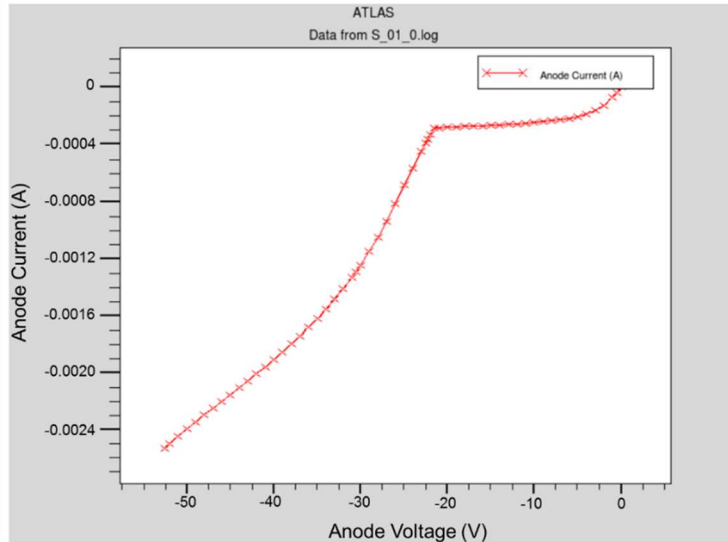


Figure 4.3 : Caractéristiques I(V) de la jonction n+/p_substrat sans anneau de garde

D’après la Figure 4.3, on remarque que la caractéristique I(V) de la jonction n+/p_substrat montre un claquage prématuré et ne ressemble pas à une caractéristique typique d’une jonction pn et c’est à cause du pic du champ électrique sur les bords de la jonction.

4.2.3 Structure pn avec anneau de garde

Pour éviter le claquage prématuré, une nouvelle structure comportant un anneau de garde est alors simulée comme le montre la Figure 4.4.

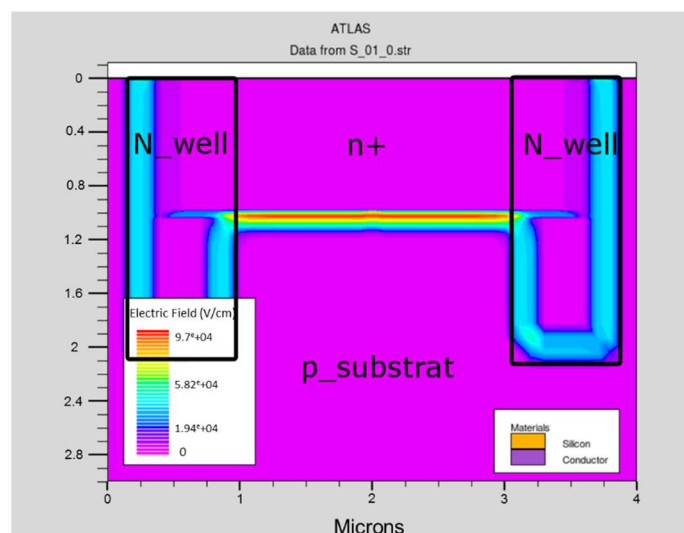


Figure 4.4 : distribution du champ électrique dans la structure n+/p_substrat avec anneau de garde

Dans cette structure, l'anneau de garde est réalisé avec la couche N_{well} du process qui présente l'avantage d'être faiblement dopée. Grâce à l'ajout de cette structure, le pic du champ électrique est évité aux coins de la jonction n⁺/p_{substrat}, cela se traduit par la caractéristique I(V) donnée par la Figure 4.5.

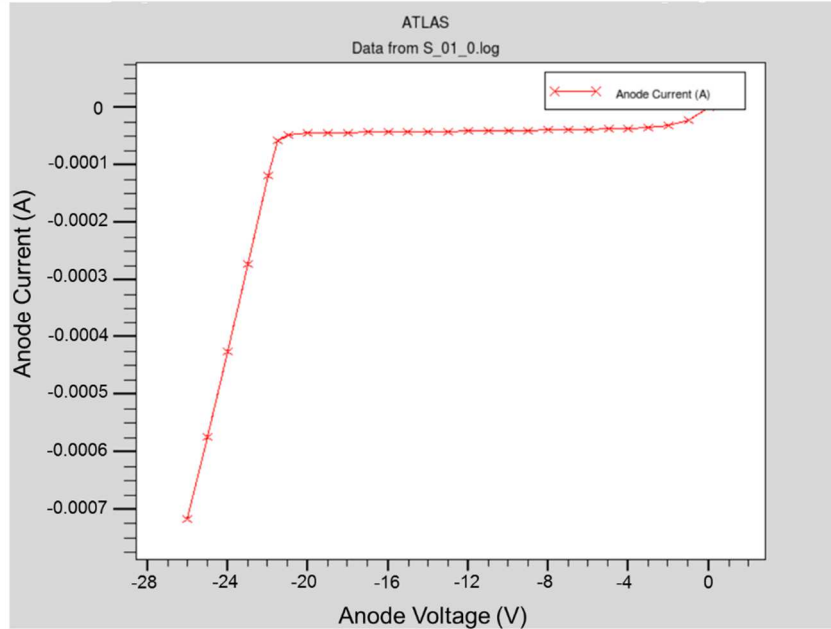


Figure 4.5 : Caractéristiques I(V) de la jonction n⁺/p_{substrat} avec anneau de garde

La caractéristiques I(V) de la structure avec anneau de garde montre que le claquage prématuré a été évité et que la courbe I(V) est typique d'une jonction PN. La tension de claquage est de 22 V (en valeur absolue) ce qui n'est pas loin des résultats réels de la structure de Haitz [Goetzberger_63].

Pour conclure cette étude dans la technologie visée, les simulations en utilisant l'outil Silvaco confirment que la présence de l'anneau de garde dans la structure de la photodiode est indispensable.

4.3 Modèle haut niveau du SPAD

L'objectif de cette section est de décrire le modèle haut niveau que nous avons développé. Ce modèle doit prendre en compte les sources de photons naturelles (sources lumineuses) ainsi que le bruit qui constitue une source de génération d'impulsions importante dans un SPAD.

Les facteurs de mérites vont caractériser les performances du SPAD, il s'agit principalement des caractéristiques détaillées dans le chapitre 2 : la probabilité de détection de photon (PDP), le taux de comptages dans le noir (DCR), la probabilité d'impulsion ultérieure (Afterpulsing) et le temps mort (DT). Il faut signaler que la diaphonie n'est pas prise en compte puisqu'il s'agit d'une étude d'un seul pixel, donc la probabilité de l'avoir est

inexistante dans le modèle proposé. L'objectif est de construire un modèle temporel des instants d'arrivés des photons et impulsions. La Figure 4.6 montre une vue globale de notre modèle.

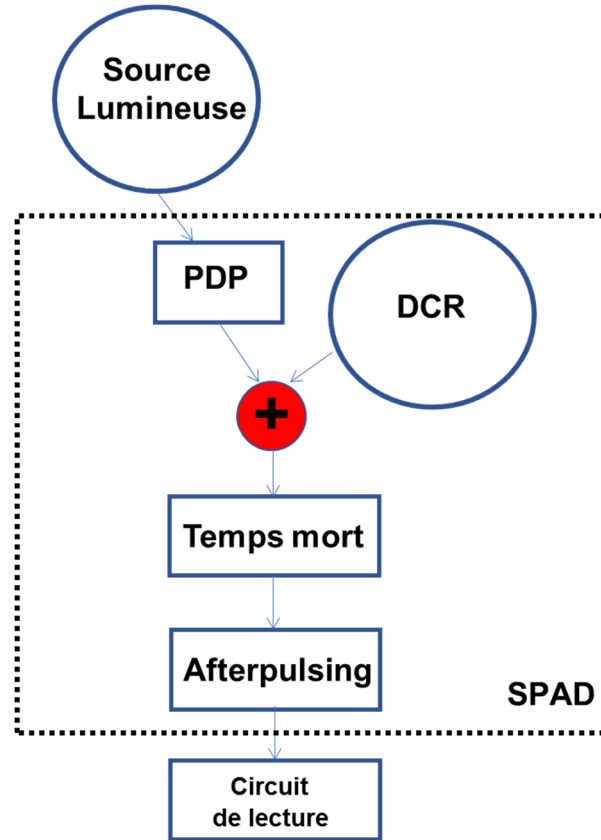


Figure 4.6 : Structure du modèle développé

4.3.1 Le processus de Poisson

Si on s'intéresse à la production de photons par une source lumineuse, pour la majorité des sources lumineuses, la statistique temporelle de production de photon va suivre une distribution qui suit la loi de Poisson (distribution Poissonienne). Les sources de bruit dans un SPAD qui produisent des impulsions électriques suivent également une distribution Poissonienne [Repich_09]. Dans cette section, on va détailler le processus de Poisson afin de comprendre la distribution temporelle de toutes ces impulsions venant des différentes sources pour ensuite construire le modèle temporel des impulsions électriques générées par le SPAD.

Un processus de Poisson est un processus stochastique qui compte le nombre d'événements et les instants auxquels ces événements se produisent dans un intervalle de temps donné. C'est donc un modèle probabiliste des situations qui voient un flux d'événements se produire les uns à la suite des autres de façon aléatoire dans le temps.

On considère des évènements qui se produisent à des dates aléatoires, et on s'intéresse, pour tout $t > 0$ au nombre d'évènements qui se sont produits au cours de l'intervalle $]0, t]$, que l'on note $N(t)$.

On définit alors le processus de Poisson d'intensité $\lambda > 0$ comme un processus $N(t)$ satisfaisant les conditions Ci suivantes :

- C1 : le processus est à accroissements indépendants, ie $\forall 0 \leq t_1 \leq t_2 \dots \leq t_n$, les variables aléatoires $N(t_i) - N(t_{i-1})$, $i = 1, \dots, n$ sont globalement indépendants ;
- C2 : le processus est à accroissements stationnaires, ie $\forall t, h > 0$, la loi de $N(t + h) - N(t)$ ne dépend que de h ;
- C3 : $P(N(t + h) - N(t) \geq 1) = \lambda h + o(h)$ et $P(N(t + h) - N(t) \geq 2) = o(h)$.

La condition C3 signifie que sur un petit intervalle de temps, on observe 0 ou 1 évènement, et que la probabilité d'observer un évènement est proportionnelle au temps écoulé. On peut remarquer que $N(0) = 0$. Comme le montre la Figure 4.7, les trajectoires sont par définition croissantes, continues à droite avec une limite à gauche et elles croissent par sauts d'une unité.

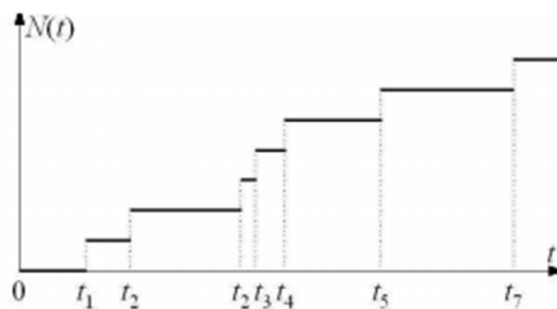


Figure 4.7 : exemple d'un trajectoire d'un processus de Poisson

4.3.2 Processus de Poisson : Propriété

Pour tout $t > 0$, la variable $N(t)$ qui suit une loi de Poisson de paramètre λt , est donné par l'équation (4.1) :

$$P(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \tag{4.1}$$

Dans ce cas, la moyenne $E(N) = \lambda t$ et la variance $Var(N) = \lambda t$ représente la loi de poisson de paramètre λt .

4.3.3 Processus de Poisson : algorithmes de simulation

Dans un processus de poisson, la théorie probabiliste permet de déterminer :

- La loi du temps d'attente X_1 du premier événement A,
- La loi du temps d'attente X_r du r-ème événement A,
- La loi du nombre N d'événements A qui se sont produits dans une durée $[0, t]$.

Le temps d'attente d'occurrence moyen est donné par l'équation (4.2). C'est la loi exponentielle de paramètre λ .

$$E(X_1) = \frac{1}{\lambda} \quad (4.2)$$

Le nombre moyen d'événements A qui se produisent dans une durée t est donné par l'équation (4.3).

$$E(N) = \lambda t \quad (4.3)$$

Il y a donc en moyenne λ événements A par unité de temps (λ est la cadence du phénomène).

Ils existent plusieurs algorithmes de simulation qui permettent de simuler la loi Poissonienne. On peut retenir principalement deux méthodes :

- La première méthode est la méthode par inversion applicable pour des variables aléatoires (V.A.) à fonction de répartition simple dont le théorème est : soit U une variable de loi $U(0,1)$ et F une fonction de répartition quelconque. La variable aléatoire X définie par l'équation (4.4) a pour fonction de répartition F .

$$X = F^{-1}(U) \quad (4.4)$$

- La deuxième méthode est basée sur un changement de variable dont le théorème est le suivant : soit X une variable aléatoire réelle de densité f_x et soit ϕ une application dérivable strictement croissante. Alors la variable aléatoire donnée par l'équation (4.5) [Guillet_11] est :

$$Y = \phi(X) \quad (4.5)$$

Admet pour densité la fonction f_Y définie par l'équation (4.6).

$$f_Y(y) = \frac{1}{\phi'(\phi^{-1}(y))} f_X(\phi^{-1}(y)) \quad (4.6)$$

Un processus de Poisson sur \mathbb{R}^+ d'intensité λ est un processus de renouvellement N_t tel que l'équation (4.7) :

$$N_t = \sum_{n \in \mathbb{N}} 1_{t_n \leq t} \quad (4.7)$$

Où les t_i sont des V.A. indépendantes de la loi exponentielle de paramètre λ . Les t_i sont appelés les instants de saut du processus.

Pour simuler un processus de Poisson, on peut simuler des V.A. de loi exponentielle de paramètre λ dont la densité est $f_\lambda(x) = \lambda e^{-\lambda x}$, en utilisant la méthode d'inversion donnée par l'équation (4.8).

$$t = \left(\frac{-1}{\lambda}\right) \ln(1 - Y) \quad (4.8)$$

Où Y est une variable aléatoire qui a une distribution uniforme sur $(0,1)$.

4.3.4 Source lumineuse

Comme cité précédemment, on propose une construction du modèle temporel des instants de génération des impulsions électriques en sortie du SPAD. C'est un modèle de simulation, implémenté en utilisant MATLAB, d'un pixel contenant l'élément photosensible (la photodiode) suivi de son circuit de lecture. Le modèle proposé nous permet de simuler le comportement du SPAD avec précision et de montrer les imperfections du système qui peuvent affecter le résultat.

La première étape dans la construction de ce modèle est de déterminer les instants d'arrivés des photons provenant de la source lumineuse qui représente l'entrée principale du système. Pour simuler cette source, nous avons modélisé une source de photons monochromatique afin de déterminer la moyenne du nombre de photons par seconde qui rentre en collision avec l'élément photosensible. Ce nombre de photon par seconde est fortement liée à la surface de la photodiode, la puissance de la source lumineuse et la longueur d'onde de la source de lumière monochromatique. Etant donné que la source lumineuse suit la distribution Poissonienne, λ est alors le nombre de photon par seconde qui représente l'intensité du processus de Poisson donné par l'équation (4.9).

$$\lambda = \frac{P \cdot A \cdot \lambda_{source}}{h \cdot c} \quad (4.9)$$

Où P est la puissance de la source lumineuse, λ_{source} est la longueur d'onde, h est la constante de Planck, c la vitesse de la lumière et A la surface de la photodiode.

La simulation de la loi exponentielle de paramètre λ donnée par l'équation (4.8) nous a permis de déduire l'information temporelle *instant_event* qui nous donne les instants d'arrivés des photons provenant de la source lumineuse. Grâce à la simulation MATLAB, nous avons pu construire la distribution temporelle des photons comme le montre la Figure 4.8, pour un $\lambda_{source} = 570$ nm, $P = 0.0004$ W/m² et $A = 36$ μm².

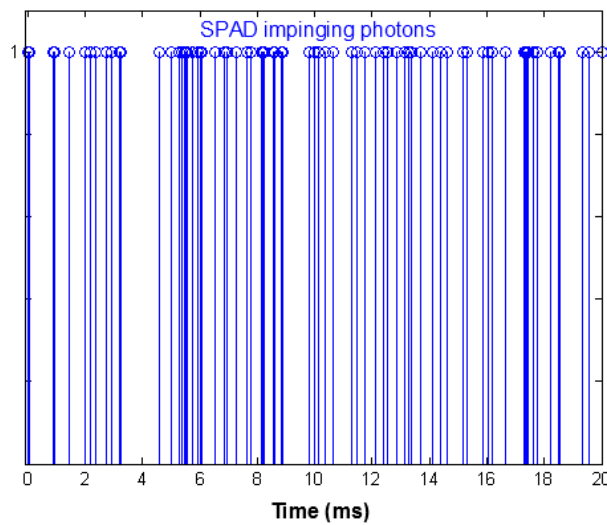


Figure 4.8 : la distribution temporelle des impulsions provenant de la source de photon.

L'histogramme de la Figure 4.9 vérifie que la distribution des impulsions provenant de la source lumineuse suit le processus de Poisson.

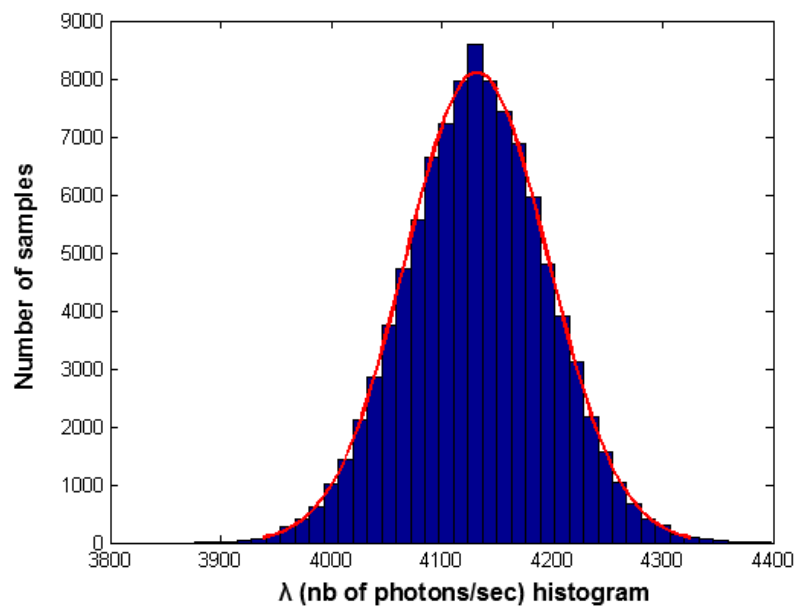


Figure 4.9 : Histogramme du nombre de photons/seconde pour 10000 échantillons

En effet la simulation de 10000 échantillons est faite en utilisant un modèle de source monochromatique de puissance 0.0004 W/m^2 , la longueur d'onde est de 570 nm pour un SPAD de $36 \mu\text{m}^2$. L'utilisation de l'équation (3.9) nous permet de calculer la valeur théorique de $\lambda(\lambda_{théorique})$, la moyenne et l'écart type. Dans notre cas $\lambda_{théorique} = 4,3119 \cdot 10^3 \text{ ph/sec}$, donc pour vérifier que la distribution des photons suit le processus de Poisson il faut que λ soit égale à la moyenne ce qui est le cas de la Figure 4.9 [KAZ-MA_15a].

Pour que le modèle soit réaliste et précis, il faut que la probabilité de détection de photon (PDP) soit prise en compte. Cette probabilité dépend fortement de l'efficacité quantique (QE) et de la longueur d'onde (λ). La Figure 4.10 montre la variation de la PDP en fonction de la longueur d'onde et l'excès de la tension de polarisation pour les technologies CMOS de 0.35 et $0.8 \mu\text{m}$.

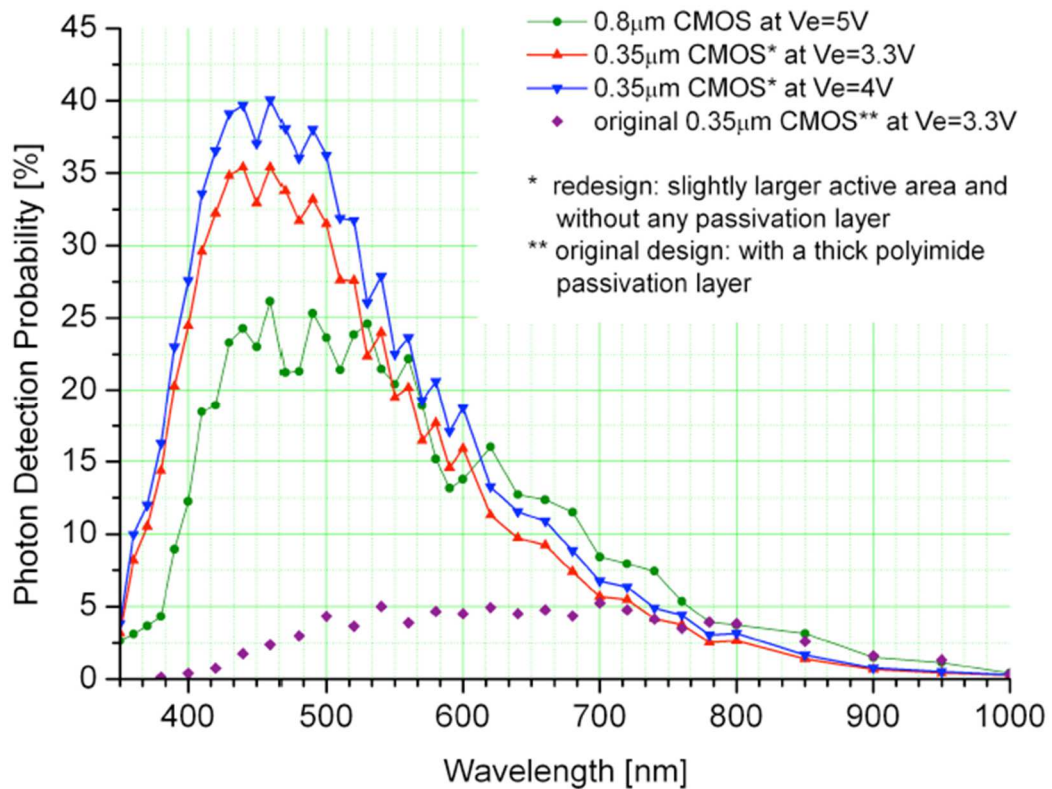


Figure 4.10 : variation de la PDP en fonction de la longueur d'onde pour les technologies 0.35 et $0.8 \mu\text{m}$ [Niclass_06]

En tenant compte du fait que la probabilité (PDP) n'est pas de 100% , la distribution temporelle des photons sortants de la source n'est plus la même. Dans notre modèle, chaque photon va subir un test pour savoir s'il a été absorbé ou pas. Après avoir déterminé les instants d'arrivés des photons *instant_event*, on procède sur MATLAB à un test pour ces instants d'arrivés en prenant en compte les paramètres de la technologie $0.35 \mu\text{m}$ (en termes de QE) avec un excès de tension de polarisation $V_{ex} = 0$ et une longueur d'onde de 570 nm . Pour ces paramètres on a une Probabilité PDP = 40% , ce qui va générer une nou-

velle distribution des photons. En appliquant la PDP au *instant_event*, on va retenir que 40 % des photons de la source lumineuse (40 % des impulsions de la Figure 4.8 sont gardées) ce qu'on appelle les *instant_event_PDP*.

4.3.5 Taux de comptage dans le noir

Le taux de comptage dans le noir (DCR) représente une source de bruit interne. Cet événement qui cause l'avalanche a pour origine plusieurs phénomènes :

- La génération thermique,
- L'effet tunnel,
- La diffusion des porteurs des régions neutres.

Pour les technologies CMOS SPADs, le taux de comptage dans le noir (DCR) dépend fortement de la température et de l'excès de tension de polarisation comme le montre la Figure 4.11.

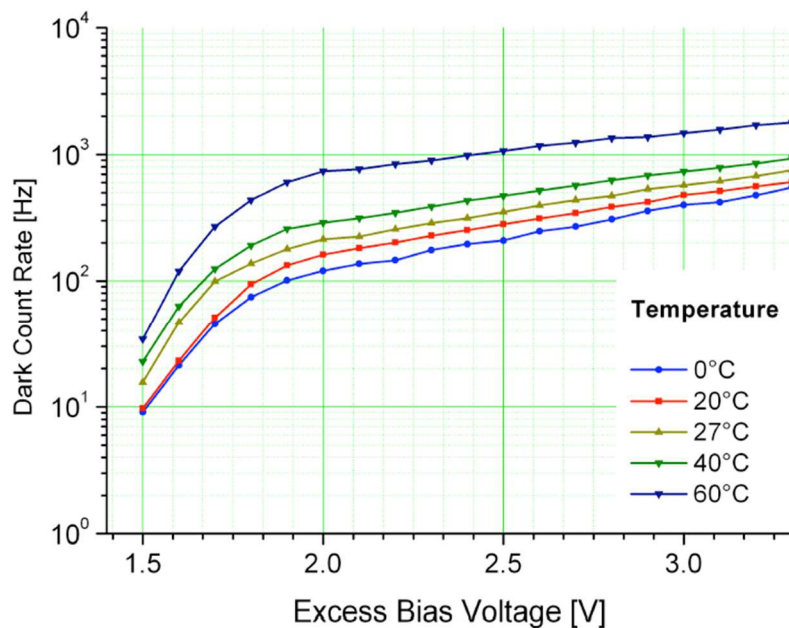


Figure 4.11 : le taux de comptage dans le noir en fonction de l'excès de la tension de polarisation pour différentes températures [Niclass_06]

Le taux de comptage dans le noir est une source de bruit interne qui suit une distribution Poissonienne. Il est modélisé dans le modèle haut niveau par une source d'impulsion qui génère des impulsions qui suivent la distribution Poissonienne. En procédant comme dans le cas de la source lumineuse, donc en utilisant l'équation (3.8), les instants d'arrivés des impulsions provenant du taux de comptage dans le noir (DCR) sont calculés *instant_event_DCR*.

Dans ce modèle, pour une photodiode de $6 \times 6 \mu\text{m}^2$, le DCR est égale à $10 \text{ Hz}/\mu\text{m}^2$. En effet, ce paramètre varie de 1 à $10 \text{ Hz}/\mu\text{m}^2$ dans la technologie $0.35 \mu\text{m}$, donc on a choisi le pire cas, ce qui fait que $\lambda_{DCR} = 360$. Pour vérifier que le modèle de cette source suit le processus de Poisson, on procède comme dans le cas de la source lumineuse en simulant l'histogramme de 10000 échantillons pour voir si $\lambda_{DCR} = \text{moyenne_DCR}$.

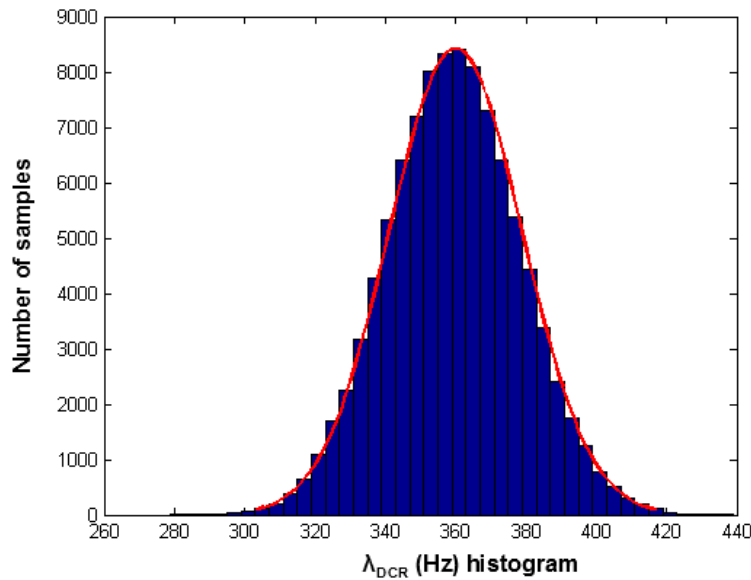


Figure 4.12 : λ_{DCR} Histogramme pour un SPAD de $36 \mu\text{m}^2$

La Figure 4.12 montre bien que la distribution de DCR est une distribution Poissonienne.

4.3.6 Modélisation du temps mort

Une des figures de mérites principale qui caractérise le SPAD est le temps mort, et sa prise en compte est nécessaire pour rendre le modèle réaliste. Comme on a vu dans le chapitre 2, quand le SPAD détecte un photon pendant un temps donné, il est incapable de détecter un nouveau photon. Le temps morts (DT) est le temps où le SPAD reste sous les conditions de polarisation, ainsi incapable de détecter un nouveau photon. Pour un bon fonctionnement du SPAD, il faut que le temps mort soit le plus petit possible.

Dans les travaux effectués sur MATLAB, nous avons développé une méthode numérique qui permet de prendre en compte le temps mort du SPAD. Cette méthode est basée sur un test entre deux impulsions successives. Grâce à ce test on va identifier la distance temporelle entre ces deux impulsions dans un premier temps. Si cette distance est inférieure à celle du temps mort, l'impulsion est rejetée et l'indicateur passe à l'impulsion suivante qui subira le même test, sinon l'impulsion est conservée.

Les impulsions provenant de la source lumineuse et les impulsions provenant du taux de comptage dans le noir sont étalées sur le même espace temporel et subiront ce test comme le montre la Figure 4.13.

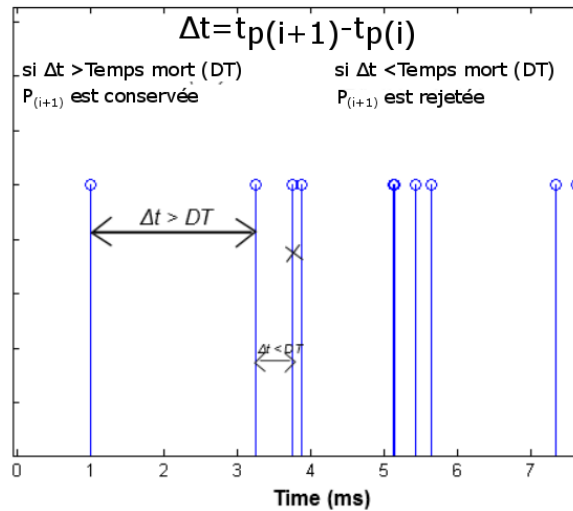


Figure 4.13 : le modèle du temps mort

A l'issue de ce test que toute les impulsions ont subi, une partie des impulsions qui respecte le modèle du temps mort seront gardées et celles qui ne respectent pas ce modèle seront rejetées. Il faut préciser que le temps mort minimal pour une technologie 0,35 μm est de 40 ns [Niclass_06]. En appliquant la méthode qu'on a développée précédemment, on arrive à calculer les instants d'arrivées finaux (source lumineuse et DCR) qui respecte le temps mort $instant_event_tot_DT$.

4.3.7 Modélisation de l'afterpulsing

Les pièges à l'intérieur de la région de multiplication ont tendance à capturer des porteurs photo-générés lors d'une avalanche. Ces derniers seront libérés aléatoirement plus tard et peuvent potentiellement re-déclencher une avalanche corrélée dans le temps avec l'avalanche initiale. Ce phénomène est appelé *afterpulsing*. Le paramètre caractérisant cet effet est connu sous le nom d'*afterpulsing probability*, il dépend fortement du temps mort (DT) comme le montre la Figure 4.14.

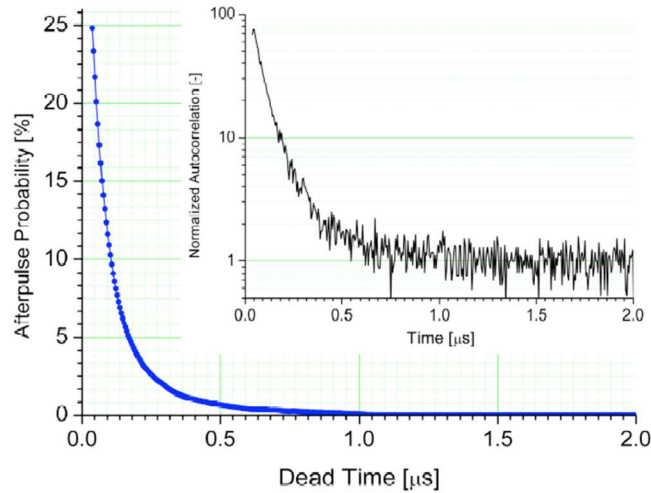


Figure 4.14 : la variation de la probabilité d'Afterpulsing en fonction du temps morts [Niclass_06]

Pour savoir si les impulsions provenant de la source lumineuse et du DCR produisent une impulsion afterpulsing, elles subissent un test qu'on a appelé afterpulsing test. L'afterpulsing test est un test réalisé en utilisant MATLAB qui va prédire selon la probabilité d'afterpulsing si l'impulsion génère ou pas une impulsion afterpulsing.

Pour la technologie $0,35 \mu\text{m}$ la probabilité d'afterpulsing est de 2% pour un temps mort = 40 ns à 25°C . Dans le modèle haut niveau qu'on propose, ces impulsions d'origine afterpulsing apparaîtront aléatoirement, selon la probabilité calculée, juste après le temps mort.

4.3.8 Modèle complet

Une dernière étape manque pour compléter le modèle haut niveau du SPAD : il faut également modéliser le temps de Jitter. Ce temps représente la fluctuation statistique entre l'arrivée du photon au SPAD et l'impulsion de sortie. Pour la technologie CMOS $0.35 \mu\text{m}$, ce temps est de l'ordre de plusieurs centaines de picoseconde, cette valeur est suffisamment faible (devant les autres constantes de temps comme le temps mort par exemple), pour que, dans un premier temps, nous négligions cet effet.

La modélisation des impulsions provenant de la source lumineuse, du taux de comptage dans le noir (DCR) et de l'afterpulsing nous permet de construire un modèle temporel des instants d'arrivées des impulsions et de préciser l'origine de ces impulsions comme le montre la Figure 4.15.

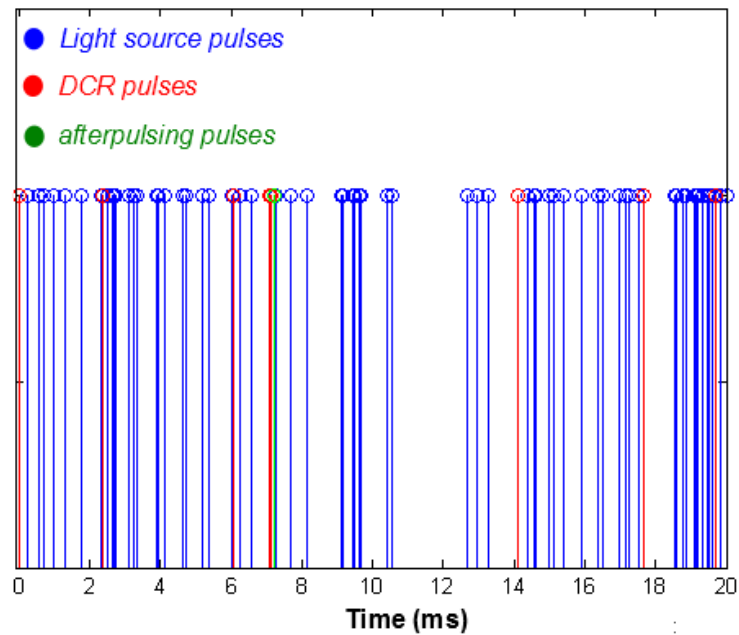


Figure 4.15 : Modèle complet haut niveau du SPAD

4.4 Modèle du circuit de lecture

En sortie du SPAD, nous obtenons une impulsion de tension de courte durée. Cette impulsion est liée à l'interaction d'un photon ou à un événement de type DCR (Dark Count Rate) ou afterpulse. Cette impulsion doit ensuite être traitée, par un circuit de lecture intégré au pixel, comme le montre Figure 4.16.

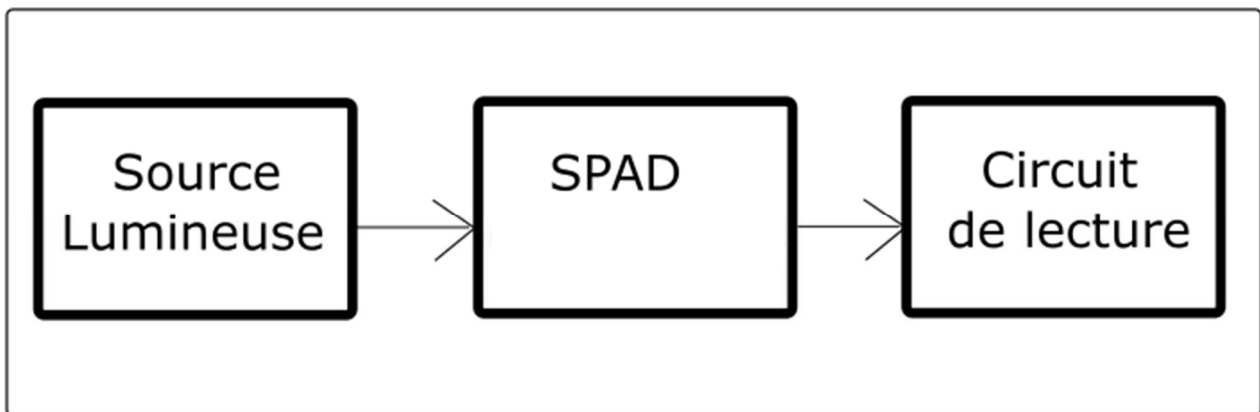


Figure 4.16 : Modèle haut niveau du SPAD suivi de son circuit de lecture

Comme il a été vu au chapitre 3, une des solutions pour enregistrer les événements consiste en l'utilisation d'un compteur analogique. A chaque événement, le compteur s'incrémente d'un quantum de tension. Ce compteur sera limité en termes de dynamique. Dans le modèle haut niveau proposé on s'est limité, dans un premier temps, à une dynamique 200 impulsions (photons, DCR ou Afterpulse). Dès que le compteur atteint sa dy-

namique maximale, on enregistre l'instant et ensuite le compteur est réinitialisé. Nous nous limitons pour l'instant à ce modèle basique. L'objectif de ce travail étant entre autres l'optimisation de la dynamique de lecture. Un modèle plus réaliste du compteur sera détaillé au chapitre 5.

En effet, il existe plusieurs architectures de circuit de lecture associées au SPAD. Dans les prochains chapitres, on va détailler ces architectures tout en évoquant leurs avantages et leurs inconvénients pour au final proposer une nouvelle architecture qui permet de sensiblement accroître la dynamique de lecture. C'est dans ce contexte qu'un modèle plus précis de compteur sera développé.

4.5 Conclusion

Nous avons dans ce chapitre détaillé un modèle complet de SPAD et de son système de lecture. Dans un premier temps, on a modélisé physiquement la photodiode à avalanche en utilisant le logiciel Silvaco, ensuite on l'a simulée électriquement en inverse au-dessus (en valeur absolue) de la tension de claquage pour voir la distribution du champ électrique dans la structure et déduire la courbe caractéristique $I(V)$. Pour éviter le claquage prématuré, la simulation électrique a montré que la présence de l'anneau de garde dans la structure de la photodiode est cruciale.

Dans la deuxième partie de ce chapitre, nous avons proposé un modèle haut niveau du SPAD suivi par son circuit de lecture. Ce modèle représente une construction temporelle des instants d'arrivés des photons (ou impulsions) provenant de toutes les sources naturelles ou sources de bruit existantes. L'avantage principal de ce modèle est qu'il prend en considération tous les facteurs de mérites qui caractérisent les performances du SPAD (Afterpulsing, temps mort). La modélisation haut niveau du circuit de lecture basé sur un compteur de photons et d'impulsion a été présentée pour savoir à quel moment, le compteur atteint sa dynamique maximale, incapable ensuite de recevoir des nouvelles impulsions.

La modélisation du circuit de lecture avec une dynamique limitée constitue une introduction au prochain chapitre dont l'objet principal est d'améliorer cette dynamique de lecture en proposant une nouvelle architecture de lecture.

Chapitre 5 *Du SPAD au pixel*

5.1 Introduction

Dans le chapitre précédent, nous avons présenté un modèle haut niveau de l'élément photosensible (le SPAD) suivi de son circuit de lecture. C'est une construction temporelle des instants d'occurrence des impulsions, et il permet également la simulation au niveau système du pixel en prenant en compte tous les facteurs de mérite. Ce modèle offre même la possibilité de trier les impulsions selon leurs sources de génération. Pour compléter le modèle haut niveau, un modèle basique du circuit de lecture basé sur un compteur analogique ayant une dynamique limitée a été présenté.

Le modèle de compteur analogique proposé dans le chapitre 4 modélise un compteur idéal. En réalité, le compteur analogique souffre de nombreux problèmes comme la non linéarité, la dynamique, les dispersions qui affecte le pas de tension ΔV , etc.

Dans ce chapitre, une nouvelle technique de lecture visant l'amélioration de la dynamique de lecture du SPAD est présentée. Grâce à cette nouvelle architecture, notre système pourra atteindre des dynamiques importantes tout en limitant l'ajout de circuiterie. Dans ce chapitre également, une méthode de lecture permettant l'amélioration de la résolution en temps dans les mesures sera présentée.

Pour finir ce chapitre, un modèle VHDL-AMS de la nouvelle architecture sera proposé et ensuite simulé. Cette simulation électrique confirme la validité de cette nouvelle technique et son apport sur l'amélioration de la dynamique de lecture.

5.2 Introduction aux techniques de lecture dans un SPAD

Pour lire l'information provenant du SPAD, plusieurs techniques de lecture existent. Dans un premier temps, on présente les deux principales techniques de lecture basées sur des compteurs analogiques existants dans la littérature. Dans la suite de ce chapitre, on

détaillera l'architecture que nous proposons et qui vise à améliorer la dynamique de lecture.

Quand on parle des applications qui reposent sur du comptage de photon, deux techniques de lecture, basées sur un compteur analogique, sont utilisées. La première est basée sur la mesure du temps nécessaire pour atteindre une tension de référence donnée. La deuxième méthode mesure la valeur de tension atteinte par le compteur analogique pour un temps d'intégration donné.

La dynamique du compteur est déterminée par le nombre maximal de pas de tension de ce dernier. Le taux de comptage est donné par l'équation (5.1).

$$\tau(i, j) = \frac{V_c(i, j)}{t(i, j)} \quad (5.1)$$

Où $V_c(i, j)$ est la tension en sortie du compteur analogique et $t(i, j)$ est le temps du comptage du pixel (i, j) . En effet, le taux de comptage moyen est le rapport (N/t) avec N le nombre de photons compté. Ce nombre est associé à une tension V_c qui représente une mesure de N (tension du compteur après comptage de N coups).

5.2.1 Méthode basée sur le temps d'intégration fixe

Comme cité précédemment, la première méthode pour compter le nombre d'évènements ayant déclenché une impulsion en sortie du SPAD (photon ou bruit) utilise un temps d'intégration fixe. Le principe est simple, dans cette méthode on mesure l'amplitude en tension du compteur analogique V_{measured} une fois le temps d'intégration préfixé t_{int} atteint. La Figure 5.1 montre le principe de cette méthode.

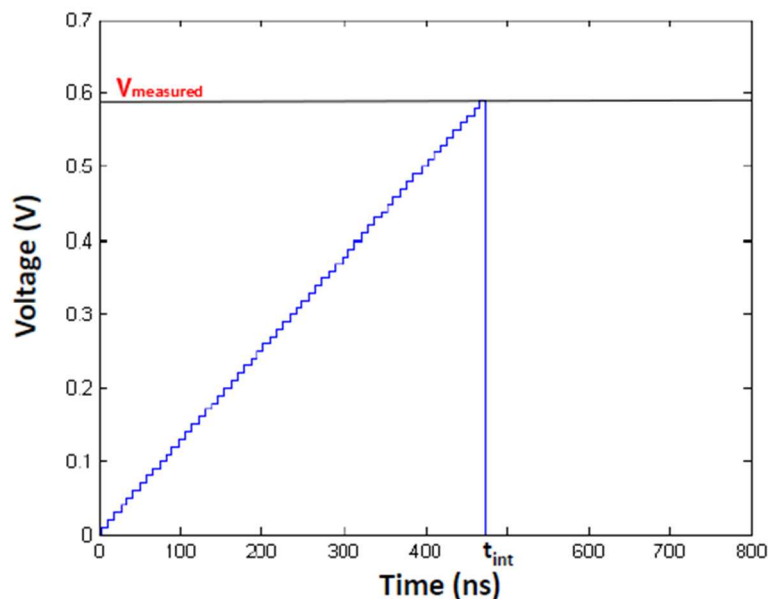


Figure 5.1 : Méthode basée sur un temps d'intégration fixe

Malgré la simplicité de cette méthode, elle souffre d'une faible résolution à cause de la résolution limitée du compteur analogique. Avec un compteur analogique simple, on peut difficilement espérer dépasser une dynamique d'environ 100. Les simulations électriques que nous présenterons au chapitre suivant montrent même que cette valeur est très optimiste. Notons aussi que dans ce cas, l'information de sortie est une grandeur analogique qui est la tension V_{measured} de la Figure 5.1.

5.2.2 Méthode basée sur la tension de référence fixe

La deuxième méthode pour compter les photons incidents au niveau d'un SPAD, est basée sur une tension de référence fixe. Le principe est similaire (dual même) à celui d'un temps d'intégration fixe, mais dans ce cas, on mesure le temps nécessaire pour que le compteur analogique atteigne une tension de référence fixée.

Cette technique nécessite la présence d'un compteur numérique global à tous les pixels (qui sert à compter le temps écoulé). Chaque pixel intègre un registre numérique et un compteur analogique. Une fois que le compteur analogique du pixel a atteint la tension de référence V_{ref} , le temps t_{measured} du compteur numérique est enregistré dans le registre pixel. La Figure 5.2 montre le schéma bloc du principe de cette mesure.

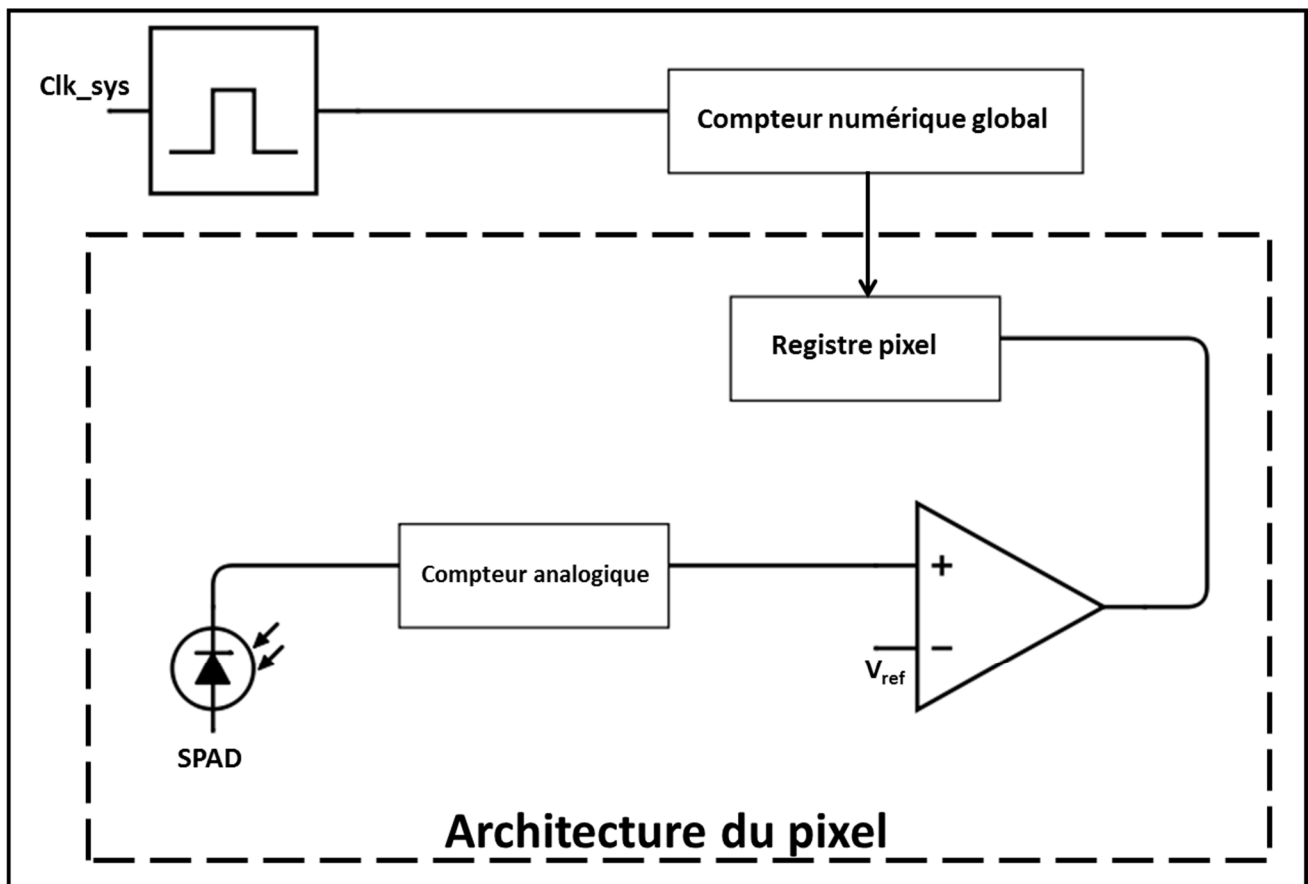


Figure 5.2 : schéma bloc du principe de la méthode basé sur une tension de référence fixe

La Figure 5.3 montre la simulation de la mesure en temps pour un pixel donné.

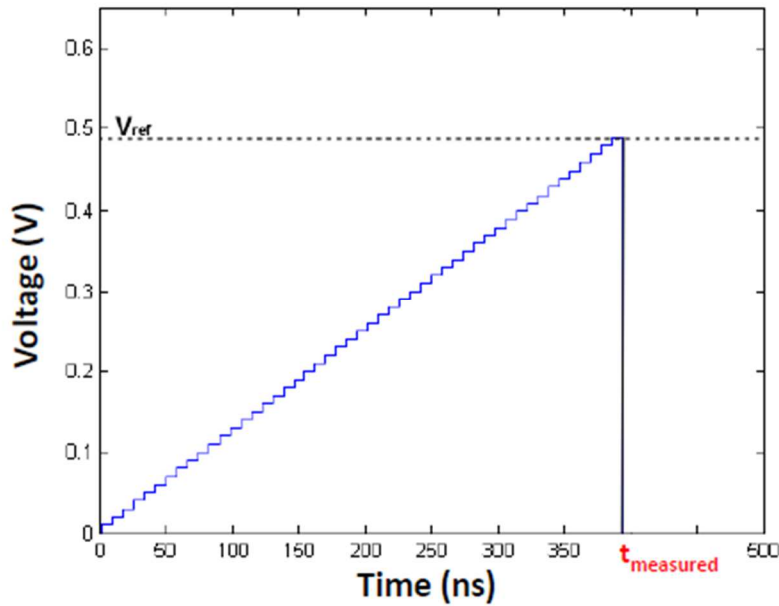


Figure 5.3 : Méthode basée sur un tension de référence fixe

Cette méthode numérique de mesure du temps pour les applications utilisant le comptage de photons est dérivé des techniques de conversion temps-numérique (Time To Digital Converter ou TDC). Le principal avantage de cette technique est qu'on a une sortie numérique directe, mais cette technique est limitée par la résolution du compteur numérique global. Cela signifie que les entrées qui ne parviennent pas à atteindre V_{ref} à la durée maximale en temps du compteur numérique seront perdues. Cette technique souffre d'un inconvénient supplémentaire, c'est le facteur de remplissage. En effet, l'utilisation d'une surface sensible importante de la puce dédiée à l'électronique de résolution en temps affecte le facteur de remplissage et le rend relativement faible [Charbon_13b].

5.3 Architecture proposée à dynamique élevée

Pour les applications qui sont basées sur un comptage de photon, on a vu qu'il existe principalement deux architectures de lecture pour un circuit de lecture formé par un compteur analogique. Si on devait choisir la meilleure technique de lecture, le choix dépendrait fortement de la nature de l'application et de son cahier de charges.

Dans cette section, une nouvelle architecture de lecture est proposée. Cette architecture présente une nouvelle méthodologie visant l'amélioration de la dynamique de lecture. Cette nouvelle technique de lecture pour le circuit de lecture du SPAD combine globalement les deux techniques de mesures présentées précédemment : mesure de l'amplitude et mesure du temps. La Figure 5.4 montre le schéma bloc de la nouvelle architecture de lecture.

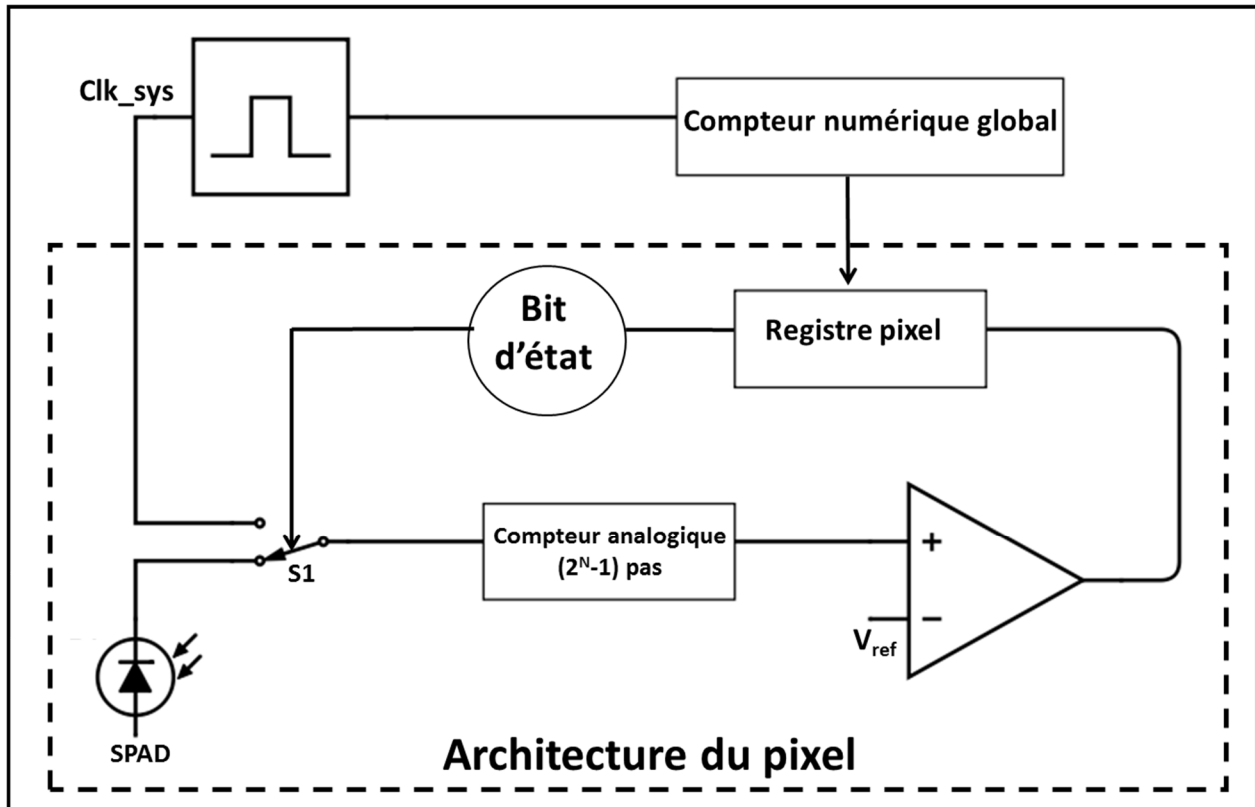


Figure 5.4 : Schéma bloc de l'architecture du pixel proposée

Cette nouvelle architecture du pixel ressemble beaucoup à celle utilisée pour la méthode basée sur une tension de référence fixe. Cependant, un bit d'état et un interrupteur (S1) sont ajoutés. Il faut préciser que le compteur numérique global est commun à tous les pixels de la matrice.

5.3.1 Concept général

Le concept général de la nouvelle architecture de lecture proposée pour améliorer la dynamique est le suivant : A chaque fois qu'un événement survient (un photon, un DCR ou un afterpulse), le compteur analogique du pixel concerné est incrémenté d'un pas de tension ΔV comme le montre l'équation (5.2).

$$V_c(t) = V_c(t - 1) + \Delta V \quad (5.2)$$

Comme cité précédemment, dans cette architecture on va combiner les deux méthodes de mesure d'amplitude et de temps. Pour appliquer cette technique, dans un premier temps, et selon le taux de comptage moyen dans la matrice, on fixe un temps d'intégration t_{int} . Egalement, on fixe une tension de référence V_{ref} .

Quand le SPAD est touché par un évènement, le compteur analogique du pixel concerné est incrémenté de ΔV . Après un nombre M d'évènements, les pixels qui ont un taux

de comptage élevé (High Counting Rate) atteignent la tension de référence V_{ref} avant l'instant t_{int} c'est le cas a. En revanche, les pixels ayant un faible taux de comptage (Low Counting Rate) n'atteignent pas la tension de référence V_{ref} à l'instant t_{int} ($V < V_{ref}$) c'est le cas b. La Figure 5.5 montre les deux cas de figure de cette nouvelle architecture.

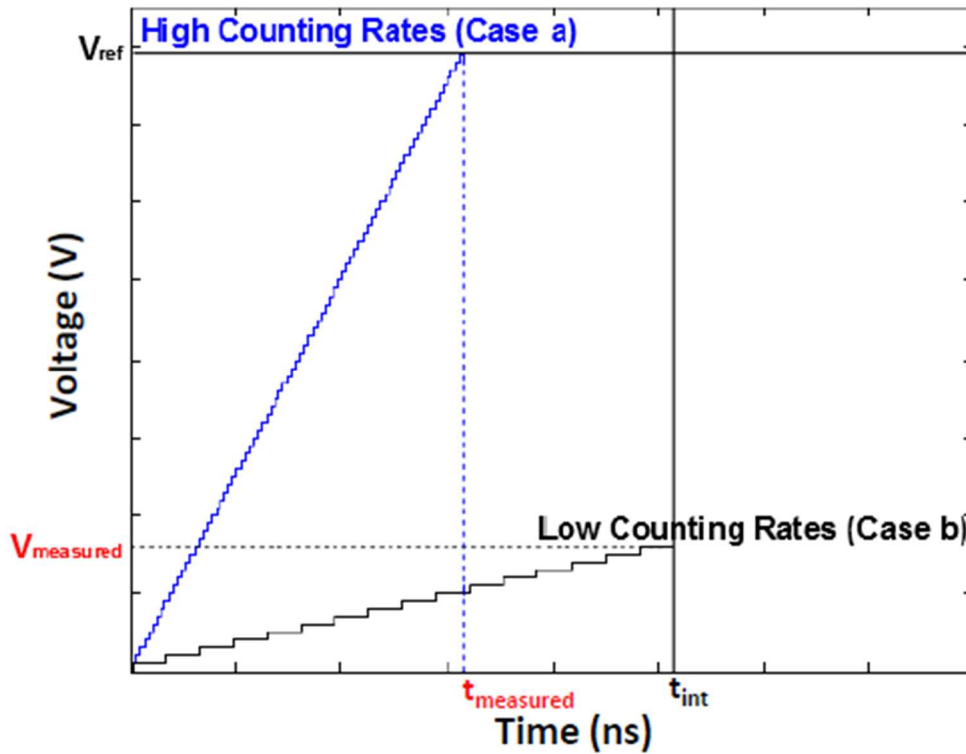


Figure 5.5 : Concept général de la nouvelle architecture de lecture

A l'issue de ce tri des pixels basé sur leurs taux de comptage, les pixels à fort taux de comptage subiront un traitement différent à ceux à faibles taux de comptage. Dans la suite de ce chapitre, on détaillera le traitement approprié de chaque cas.

5.3.2 Pixels à fort taux de comptage

Dans l'architecture de lecture qu'on propose dans le cadre de ce chapitre, les pixels sont traités par rapport à leur taux de comptage. Un traitement approprié et différent aura lieu selon la luminosité du pixel.

Pour les pixels à forts taux de comptage (cas a de la Figure 5.5), au moment où le compteur analogique du pixel concerné atteint la tension de référence V_{ref} , le temps mesuré $T_{measured}$ est stocké dans le registre numérique du pixel. Le bit d'état affiche la valeur de 0, indiquant ainsi que le pixel concerné fait partie de ceux qui ont un fort taux de comptage. Donc pour ce type de pixel, le traitement est similaire à celui de la méthode basée sur la tension de référence de la section (5.2.2).

5.3.3 Pixels à faible taux de comptage

Dans cette section, on présente le traitement des pixels à faible luminosité. Comme cité précédemment, le traitement des pixels à faible intensité lumineuse est différent que celui des pixels à fort intensité lumineuse. Dans ce cas (le cas b de la Figure 5.5), à l'instant t_{int} déterminé par le taux de comptage moyen, la tension du compteur analogique du pixel est inférieure à la tension de référence fixe V_{ref} .

Le but de ce traitement consacré aux pixels à faible taux de comptage consiste à numériser directement la valeur de la tension atteinte ($V_{measured}$) par le compteur analogique à l'instant d'intégration fixe t_{int} . En fait, Cette méthode de mesure numérique de tension pour les applications qui demande du comptage de photon est une technique inspirée du Convertisseur analogique-numérique à rampe (rampe ADC). L'inconvénient de cette technique est que l'échantillonnage du signal analogique impose un temps mort au niveau des pixels plus important, mais cependant, ces pixels ne sont maintenant plus perdus.

Contrairement au traitement des pixels à fort taux de comptage, à l'instant t_{int} pour les pixels à faibles taux de comptage, le comptage de photon en utilisant le compteur analogique du pixel s'arrête. La tension en sortie du compteur analogique interne au pixel est donnée par l'équation (5.3) (dans l'hypothèse d'un compteur analogique idéal)

$$V_c(t_{int}) = N_\gamma \cdot \Delta V \tag{5.3}$$

Où N_γ est le nombre total des photons comptés à l'instant t_{int} . L'objectif est de délivrer directement une valeur numérique qui sera l'image de N_γ . Connaissant V_c , nous avons deux façons de procéder pour remonter à N_γ pour les pixels à faibles taux de comptages. La Figure 5.6 montre les deux méthodes développées pour le traitement des pixels à faible luminosité.

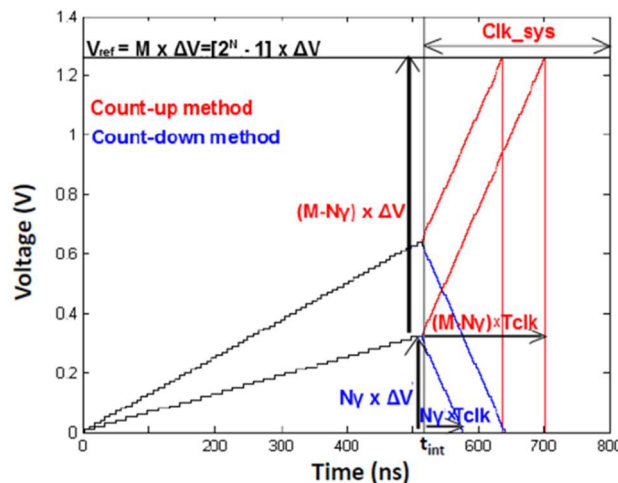


Figure 5.6 : Traitement des faibles taux de comptages : rouge comptage, Blue décomptage

La première méthode dédiée au traitement des pixels à faibles intensité lumineuse est une méthode de comptage (la courbe rouge de la Figure 5.6). Comme cité précédemment, à l'instant t_{int} , on arrête le comptage des impulsions issues du SPAD. A cet instant, le compteur numérique global à tous les pixels est remis à zéro. Comme on peut le voir sur la Figure 5.4, l'horloge globale du système (Clk_sys) du compteur numérique système est appliquée au compteur analogique du pixel grâce à l'interrupteur (S1). Le comptage continue jusqu'à ce que le compteur analogique atteigne la tension de référence V_{ref} . Comme le montre la Figure 5.6, $(M - N_\gamma)$ pas sont donc nécessaires pour atteindre la tension de référence. Cette valeur $(M - N_\gamma)$ est stockée dans le registre numérique du pixel. Le bit d'état affiche la valeur de 1 pour indiquer la présence du pixel dans une zone à faible taux de comptage.

Une deuxième méthode concernant le traitement des pixels à faible taux de comptage est possible. Cette fois ci, c'est une méthode de décomptage (la courbe bleue de la Figure 5.6). Comme pour la première méthode, à l'instant t_{int} , on bloque le comptage des impulsions issues du SPAD comme précédemment, on applique l'horloge système du compteur numérique au compteur analogique en utilisant l'interrupteur (S1) et en même temps on réinitialise le compteur numérique système comme pour le cas du traitement précédent. Pour ce cas de traitement, au lieu de continuer le comptage afin que la tension du compteur analogique atteigne la tension de référence V_{ref} , on procède à un décomptage du compteur analogique, ce décomptage déclenché à l'instant t_{int} , restera jusqu'à l'instant où le compteur analogique revient à sa tension initiale de départ comme le montre la Figure 5.6. Pour atteindre cette tension initiale au départ du comptage, il faut N_γ pas de tension ΔV . Au niveau pixel dans ce cas, le registre pixel stocke la valeur N_γ et le bit d'état affiche la valeur de 1 pour préciser la position lumineuse du pixel concerné.

Il est important de préciser que la méthode basée sur le décomptage exige un nombre plus important de bloc électronique que la méthode basée sur le comptage. En effet, le compteur analogique devra dans ce cas être conçu avec un mode décomptage, ce qui exigera un surcout de matériel. De plus, comme nous le verrons par la suite, cette dernière solution n'est pas optimale d'un point de vue linéarité.

5.3.4 Amélioration de la dynamique : impact de l'architecture

Comme indiqué au début de la section (5.3), l'objectif de la nouvelle architecture du circuit de lecture du pixel type SPAD est l'amélioration de la dynamique. Après avoir détaillé les deux cas de traitements des pixels, dans cette section, on verra quantitativement comment il a été possible, grâce à cette méthode, d'améliorer la dynamique de lecture.

Combiner les deux méthodes de mesure basées sur une tension de référence fixe et un temps d'intégration fixe aura un impact positif sur la dynamique globale du système. En effet, si le comptage commence à l'instant $t = 0$, à chaque fois qu'un événement survient (photon, DCR ou afterpulse), le compteur analogique du pixel concerné sera incrémenté d'un pas de tension ΔV . La détermination du taux de comptage (τ), dépend du temps selon une loi $\frac{1}{t}$ (équation 5.1). Dans le cas d'une mesure en temps ($N_{\text{fixé}}$), si le temps t tend vers zéro, cela conduit à un taux de comptage infini (dynamique infinie) comme le montre la Figure 5.7. Mais cette dynamique infinie exige une résolution en temps très fine (quantification du temps qui tend vers 0), pour avoir un minimum de précision (horloge à une fréquence infinie et un compteur numérique avec un très grand nombre de bit), donc une limitation au niveau du matériel (surface importante).

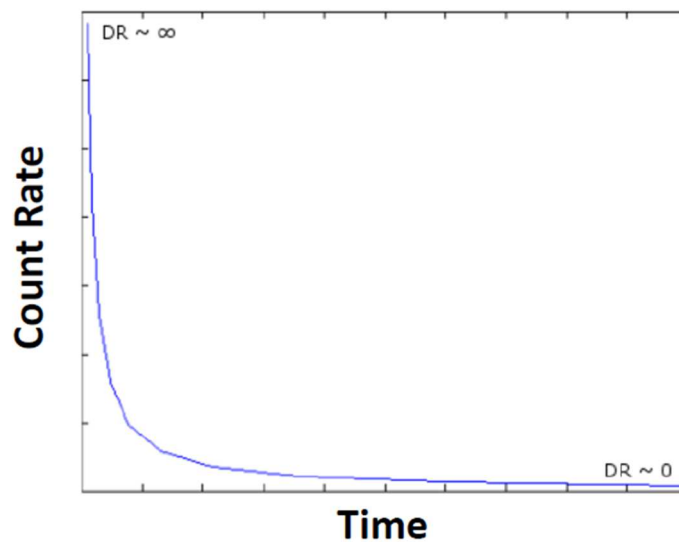


Figure 5.7 : Fonction réciproque qui lie le taux de comptage au temps

Pour limiter le nombre de bits du compteur numérique global, on ne commence pas à mesurer le temps à partir de 0 mais à partir de l'instant t_1 qu'on choisit en fonction de la résolution qu'on veut et en termes de surcout représenté par le nombre de bit du compteur numérique global et du registre pixel. A l'instant t_1 , on démarre le compteur numérique système. Entre les deux instants t_1 et t_{int} , on procède à la mesure du temps en utilisant la méthode basée sur le temps d'intégration fixe. Pour les pixels qui atteignent l'instant T_{int} avec une tension ($V_c(t_{\text{int}}) < V_{\text{ref}}$), l'amplitude de la tension analogique est alors numérisée en appliquant la méthode détaillée dans la section (5.3.3).

Dans la Figure 5.8, la courbe bleue en pointillé montre le taux de comptage maximal que le système proposé peut mesurer. La tension de référence est donnée par l'équation (5.4).

$$V_{ref} = (2^N - 1) \cdot \Delta V \quad (5.4)$$

La dynamique maximale $N_{\gamma max}$ pour (K+N) bits est donnée par l'équation (5.5)

$$N_{\gamma max} = 2^{K+N} - 1 \quad (5.5)$$

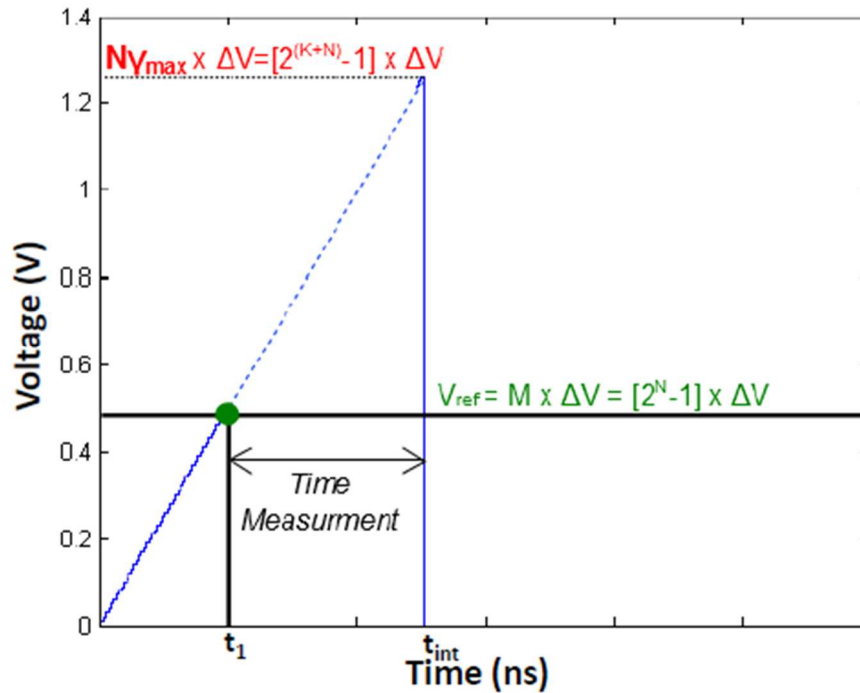
Pour avoir une dynamique maximale $N_{\gamma max}$, l'instant t_1 est calculé grâce aux équations (5.6)(5.7)(5.8).

$$\frac{t_1}{V_{ref}} = \frac{t_{int}}{N_{\gamma max} \cdot \Delta V} \quad (5.6)$$

$$t_1 = \frac{V_{ref} \cdot t_{int}}{N_{\gamma max} \cdot \Delta V} = \frac{(2^N - 1) \cdot \Delta V \cdot t_{int}}{(2^{K+N} - 1) \cdot \Delta V} \quad (5.7)$$

$$t_1 = \frac{(2^N - 1) \cdot t_{int}}{(2^{(K+N)} - 1)} \quad (5.8)$$

D'après la Figure 5.8, il est facile de montrer que si nous commençons la mesure du temps à l'instant t_1 déterminé par l'équation (5.8), la dynamique totale du système sera (N+K) bits, où N est le nombre de bits utilisé pour numériser l'amplitude ($2^N - 1 = M$ est le nombre de pas de tension du compteur analogique). Cela conduit à un système équivalent d'un compteur analogique capable de faire ($2^{N+K} - 1$) pas. Il est important de noter que le pixel qui atteint la tension de référence V_{ref} , avant l'instant t_1 ($t_{measured} < t_1$) est un pixel saturé. Ceci est dû au fait que son taux de comptage τ est supérieure au taux de comptage maximal ($\tau > \tau_{max}$) définie par la dynamique du système.

Figure 5.8 : Relation entre N_γ et T_1

A titre d'exemple, si on fixe $N = 7$ bits et $K = 3$ bits pour des raisons de résolution qu'on évoquera dans la suite de ce chapitre, on peut atteindre une dynamique de lecture en sortie de 10 bits. Les 10 bits de dynamique ne nécessitent que 7 bits effectifs en matériels électroniques. Il faut cependant préciser que coder 10 bits sur 7 effectifs, aura un impact sur la précision. Dans la suite de ce chapitre, on détaillera le choix de la résolution en temps qui résoudra le problème de la précision et qui va nous permettre d'optimiser le choix de N et K en fonction de ce qu'on veut faire [KAZMA_15b].

5.4 Impact de la Non linéarité (NL)

Dans l'architecture de lecture détaillée tout au long de la section (5.3), on a considéré, dans un premier temps, que notre système est idéal. Un des principaux problèmes dans la conception d'un compteur analogique est la difficulté de conserver une bonne linéarité sur une large gamme de comptage. En effet, la présence de la non linéarité a plusieurs sources liées à l'architecture des compteurs analogiques. Les principales sources de non linéarité sont : l'appariement (matching) des capacités, le gain fini des amplificateurs opérationnel, la non linéarité des transistors MOS, etc. La non linéarité peut-être aussi causée par l'architecture retenue du compteur analogique, dans ce cas cette non linéarité est facilement modélisable (c'est le cas comme nous le montrerons plus loin de certaines solutions basées sur des capacités commutées). En théorie, en absence de non linéarité (NL), le compteur analogique assure un incrément de tension ΔV constant quelle que soit la tension de sortie du compteur.

Après la validation de la méthode proposée dans la section (5.3), en absence de non linéarité, nous voulons tester la robustesse de notre architecture face aux non linéarités du compteur analogique. Les imperfections liées à la réalisation matérielle d'un tel compteur sont :

- La réponse transitoire du compteur (RT) : l'ordre de grandeur de la réponse transitoire du compteur est de l'ordre de la nanoseconde, celle du temps mort (DT) est de dizaine de nanosecondes ce qui fait que la réponse transitoire est négligée ($RT \ll DT$).
- Bruit et mismatch de la technologie : c'est un bruit Gaussien. La contribution globale de tous les bruits générés par les différents composants utilisés dans le design du compteur analogique est donnée par l'équation (5.9).

$$\Delta V(t) = \Delta V_0 + \Delta V_{i,j} + V_N(t) \quad (5.9)$$

Où ΔV_0 représente l'incrément en tension idéal et $V_N(t)$ est le bruit aléatoire en tension qui suit une distribution Gaussienne (V_N suit une loi \sqrt{N} où N est le nombre de coups comptés). Le terme $\Delta V_{i,j}$ représente la fluctuation statistique spatiale (mismatch) entre pixels.

- La non linéarité du compteur : l'équation (5.10) donne le modèle du compteur incluant la non linéarité.

$$V_c(t) = V_c(t - 1) + \Delta V[1 - NL \cdot V_c(t - 1)] \quad (5.10)$$

Où NL représente le coefficient de la non linéarité et V_c la tension en sortie du compteur analogique.

On peut constater qu'en insérant une non-linéarité selon l'équation (5.10), lorsque la tension de sortie du compteur analogique V_c augmente, l'incrément ΔV diminue. Ainsi, ce modèle de non linéarité simule de façon réaliste les phénomènes de non linéarités de type saturation. Ce modèle nous a permis de simuler la robustesse de l'architecture proposée. Le but des prochaines simulations est de prouver que l'architecture proposée nous permet de garantir une mesure correcte sur une large gamme de dynamique, même lorsque le compteur analogique présente une forte non-linéarité. Pour cela, la simulation de plusieurs valeurs de non-linéarité a été testée pour des pixels à faible taux de comptage. Ces simulations ont été faites pour deux méthodes de traitement concernant les pixels à faible intensité lumineuse : le comptage et le décomptage.

Afin de montrer l'efficacité de notre architecture, des simulations MATLAB ont été réalisées illustrant ainsi le pixel et son circuit de lecture contenant le compteur analogique. La Figure 5.9 montre l'impact d'une non linéarité de 30% des compteurs analogiques sur les pixels à faibles intensité lumineuse.

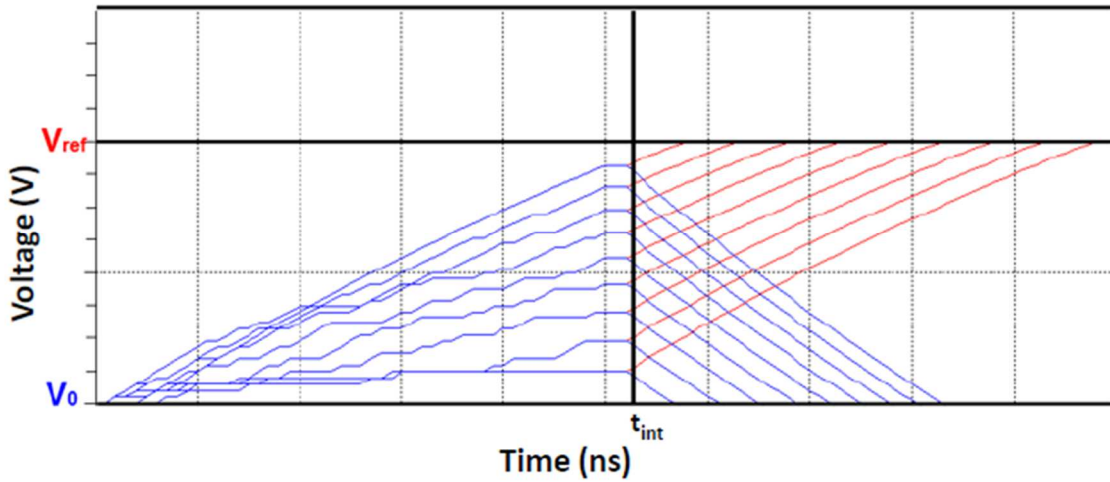


Figure 5.9 : Simulation de l'architecture proposée en présence de la non linéarité

Pour tester la robustesse de l'architecture proposée, il faut voir l'impact de la non linéarité sur le nombre de photon effectivement mesuré en fonction du nombre d'événement réellement générés par le SPAD comme le montre la Figure 5.10 pour différentes valeurs de taux de comptages.

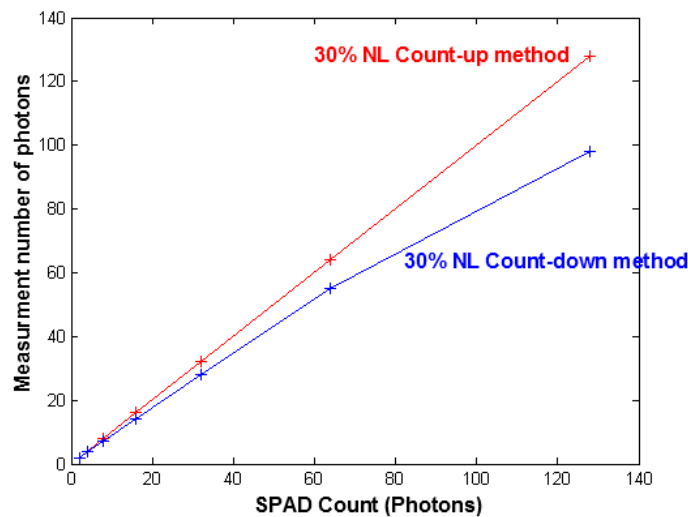


Figure 5.10 : teste de robustesse de l'architecture

Pour interpréter les résultats donnés par la Figure 5.10, on remarque que même en présence d'une non linéarité $NL > 30\%$, la méthode basée sur le comptage résiste face à la

non linéarité du compteur sur une large gamme de dynamique, contrairement à la méthode basée sur le décomptage qui souffre d'une mauvaise immunité en présence de non linéarité pour le même coefficient de non linéarité. La méthode basée sur le comptage assure une bonne linéarité sur une large gamme de dynamique, ce qui fait qu'elle sera la méthode retenue pour la suite de cette étude. Rappelons aussi que cette méthode présente également l'avantage d'être plus simple, puisqu'elle peut se faire sans surcôt en termes de circuiterie.

5.5 Réduction de l'erreur de quantification en temps

L'architecture qu'on propose dans ce travail repose sur une double mesure en temps et en amplitude. Dans un premier temps, on a validé l'architecture en absence du bruit, ensuite la robustesse de l'architecture a été testée face aux non linéarités induites par un compteur analogique.

Comme cela a été présenté précédemment, pour les pixels ayant un fort taux de comptage, on procède à une mesure en temps. Quand on évoque le sujet de la mesure du temps, un des problèmes récurrent est la résolution qui se traduit par une erreur de quantification. Dans cette section, pour optimiser la mesure en temps de l'architecture proposée dans la section (5.3), on a développé une nouvelle méthode de mesure. Grâce à cette méthode, on réduit sensiblement l'erreur de quantification. La relation qui lie le taux de comptage et le temps est donnée par l'équation (5.11).

$$\tau(i, j) = \frac{C}{t(i, j)} \quad (5.11)$$

Où C est la constante qui représente dans notre cas la tension en sortie du compteur analogique (V_c) et $t(i, j)$ est le temps de comptage du pixel (i, j) .

Comme on peut le constater d'après l'équation (5.11), la relation qui lie le taux de comptage et le temps est une fonction réciproque. C'est cette relation qui va nous permettre de définir la plage de dynamique globale et l'erreur de quantification de notre architecture. Lorsque le temps mesuré, t_{measured} , tend vers zéro, cela correspond à un taux de comptage tendant vers l'infini, ce qui permet en théorie d'obtenir une dynamique infinie. En parallèle, dans ces conditions, compte tenu de la quantification de la mesure en temps, l'erreur sur la mesure tendra elle aussi vers l'infini.

Afin de s'affranchir de cet effet (erreur infinie), on va volontairement limiter la dynamique de mesure en temps à un intervalle compris entre t_1 et t_{int} . Dans l'architecture de la section (5.3), en mettant à profit la relation non linéaire qui existe entre le taux de comp-

tage et le temps mesuré, on va pouvoir contrôler la dynamique de notre système. Comme cité précédemment, en utilisant un registre numérique de 7 bits (quantification de 127 pas entre t_1 et t_{int}), la dynamique pourra atteindre les 10 bits. L'objectif maintenant est d'étudier la distribution de l'erreur de quantification et de veiller à que cette erreur soit minimisée pour toute mesure $t_{measured}$ dans la plage comprise entre t_1 et la t_{int} .

5.5.1 Résolution en temps uniforme

Avant de calculer l'erreur de quantification pour la méthode basée sur la mesure du temps, il faut voir dans ce cas la variation du taux de comptage en fonction du temps comme le montre la Figure 5.11.

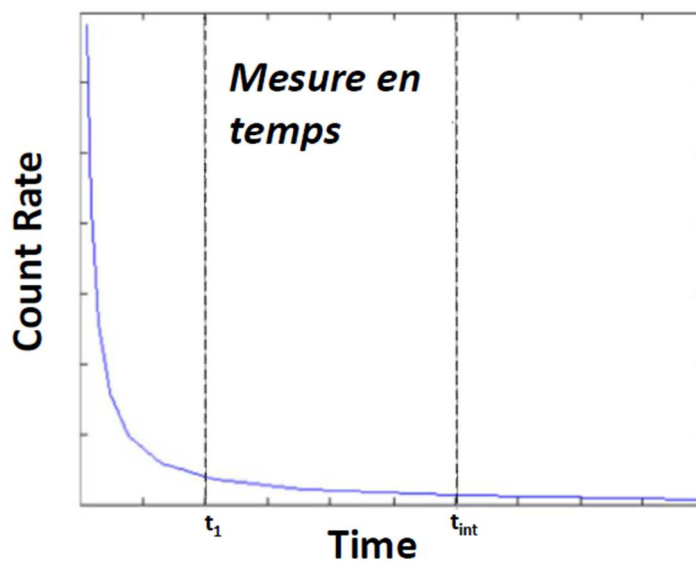


Figure 5.11 : la mesure du temps dans l'architecture proposée

Comme mentionné dans la section (5.3.4), en utilisant N bits pour le registre de pixels et un choix approprié pour t_1 , la dynamique totale du système sera égale à $N+K$ où K est le nombre de bits améliorés qu'on rajoute à la dynamique globale du système. En d'autres termes, en ayant seulement 7 bits de registre de pixels et en fixant $K=3$ (à cause de la contrainte de résolution temporelle), le comptage maximum pourra atteindre 1024 photons, soit une dynamique de 10 bits ($N=7$ et $K=3$). Cette mesure du temps sera codée en utilisant le registre numérique 7 bits du pixel, ce qui permettra de déterminer l'erreur de quantification en fonction de la sortie numérique du compteur numérique système.

La Figure 5.12 montre l'équivalent de 10 bits de dynamique codé sur 7 bits avec une méthode basée sur une résolution en temps unique définie par (T_{clk_sys}) , où $(T_{clk_sys} = \frac{t_{in}-t_1}{127})$. Rappelons aussi que si le nombre de photon compté N_γ est inférieur à 127, alors le compteur analogique n'atteint pas la référence (V_{ref}), et on procède à une mesure en amplitude de la valeur de compteur (cas des pixels avec faible taux de comptage). Ceci ex-

plique que la Figure 5.12, débute à $N_\gamma = 128$. On limite dans ce cas N_γ à 1024, considérant qu'une résolution de 10 bits est visée.

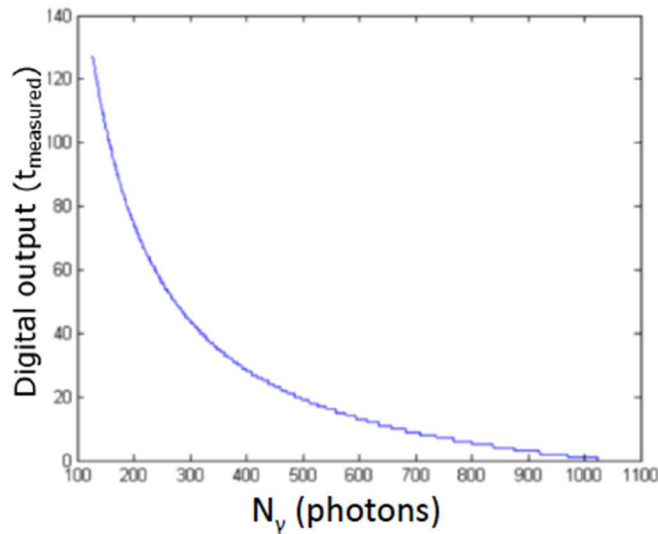


Figure 5.12 : Sortie numérique qui dépend du nombre de photons N_γ

Grâce à la Figure 5.12, l'erreur de quantification de chaque sortie numérique du compteur numérique système est alors calculée et donnée par la Figure 5.13.

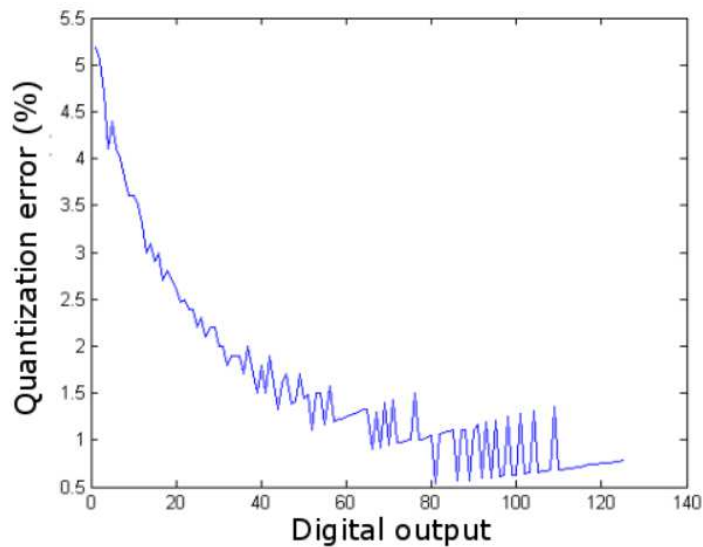


Figure 5.13 : Distribution de l'erreur de quantification pour 7 bits de sorties numérique

Comme on peut le voir sur la Figure 5.13, il y a une erreur de quantification élevée pour les faibles sorties numériques jusqu'à 5.3% pour la valeur numérique de sortie 1 ($N_\gamma = 1024$), ce qui n'est pas le cas pour les sorties numériques avec des valeurs élevées qui gardent une plus faible erreur de quantification. Ce problème nous oblige à développer une nouvelle méthode pour réduire l'erreur de quantification pour les faibles sorties numérique, tout en maintenant la faible erreur pour les sorties numériques à valeurs éle-

vées, et tout en gardant la dynamique élevée de l'architecture proposée. La solution que nous proposons pour optimiser cette erreur est basée sur un système multi résolution qui est présentée dans le paragraphe qui suit.

5.5.2 Système de mesure en temps multi-résolution

Dans la section (5.5.1), en utilisant le même temps de résolution pour toutes les sorties numériques, une erreur de quantification élevée a été trouvée pour les sorties numériques de faibles valeurs, ce qui représente un vrai problème au niveau de la résolution de l'architecture proposée.

En effet, afin de limiter l'erreur importante pour les forts taux de comptage qui provient de la relation en $(1/t)$ entre le taux de comptage et le temps mesuré, il convient de réduire le pas de quantification en temps pour les forts taux de comptage, tout en relâchant la précision de la résolution pour les faibles taux de comptage. Nous proposons donc dans notre cas d'utiliser un système de mesure en temps multi-résolution. En utilisant cette méthode, nous pourrions lisser l'erreur de quantification en modifiant simplement la résolution temporelle pour faire en sorte que l'erreur soit quasiment constante pour toutes les valeurs numériques. Ainsi, à plusieurs instants entre t_1 et l'instant t_{int} , la fréquence d'horloge du compteur est changée, ce qui impliquera un changement dans la résolution en temps comme le montre Figure 5.14.

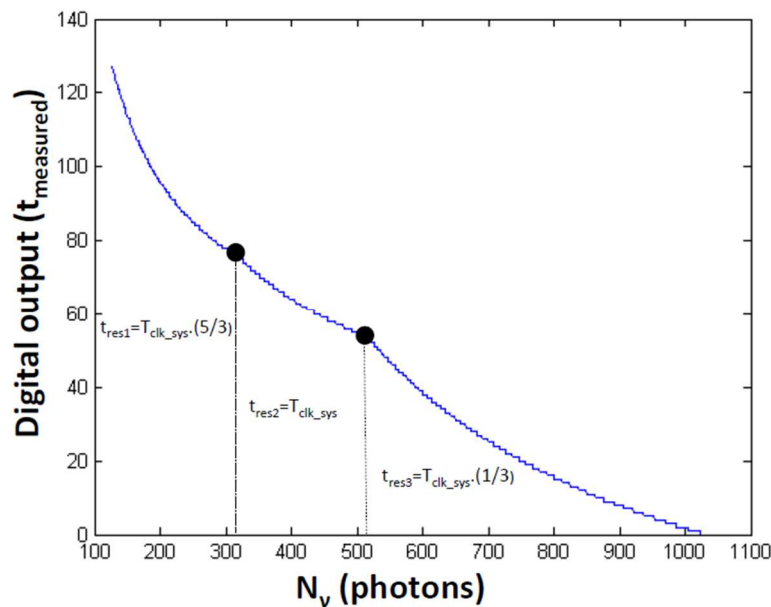


Figure 5.14 : Sortie numérique en fonction du nombre de photons N_γ en utilisant trois résolutions temporelles différentes

Sur la simulation présentée Figure 5.14, les instants t_2 et t_3 ont été déterminés et optimisés par simulation. Ces instants correspondent à un état du compteur système N_2 et N_3 respectivement comme le montre la Figure 5.15.

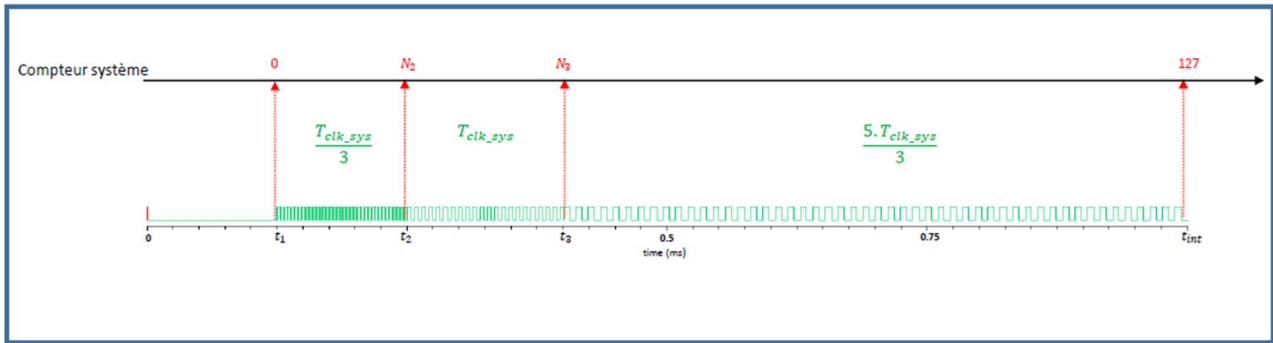


Figure 5.15 : Principe du système multi-résolution et le lien avec le compteur système

Une fois que les optimisations par simulation sont faites pour arriver au système de la Figure 5.15, on se retrouve avec les valeurs suivantes :

1. $\frac{T_{clk_sys}}{3}$: la période entre les instants $t_1 = 150 \mu s < t < t_2 = 248 \mu s$, la sortie du compteur système varie $0 < N_{compteur_système} < N_2$.
2. T_{clk_sys} : la période entre les instants $t_2 = 248 \mu s < t < t_3 = 400 \mu s$, la sortie du compteur système varie $N_2 < N_{compteur_système} < N_3$.
3. $\frac{5.T_{clk_sys}}{3}$: la période entre les instants $t_3 = 400 \mu s < t < t_{int} = 1 ms$, la sortie du compteur système varie $N_3 < N_{compteur_système} < 127$.

Le changement de la fréquence d'horloge aura un impact sur l'erreur de quantification. En effet, appliquer des fréquences différentes se traduit par un changement de la résolution du temps dans le cas de la mesure du temps dans l'architecture proposée, et un lissage de l'erreur de quantification. Ce lissage fait en sorte que l'erreur de quantification soit quasiment identique pour toutes les sorties numériques.

La Figure 5.16 montre l'impact d'une résolution en temps multi-valeurs sur l'erreur de quantification pour les pixels à forts taux de comptage pour trois résolutions différentes. Dans un premier temps, un test en n'utilisant que deux temps de résolution a montré qu'on réduit bien l'erreur de quantification mais de façon insuffisante.

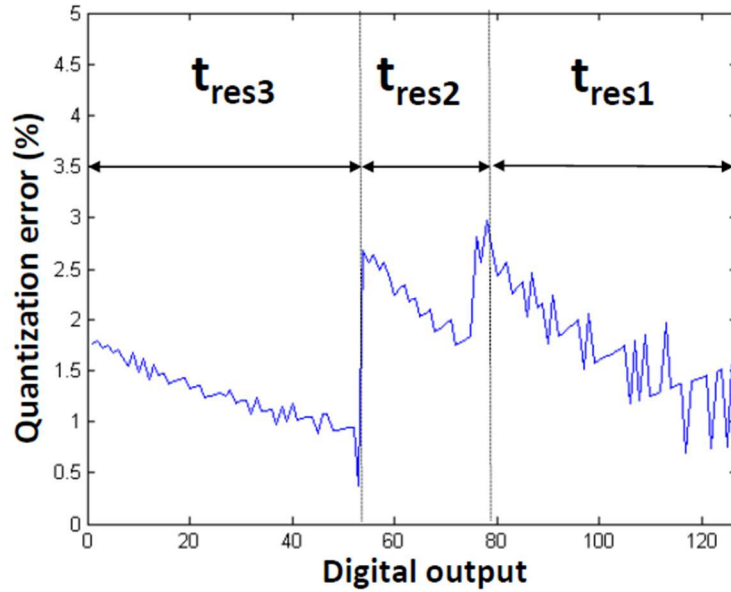


Figure 5.16 : Erreur de quantification pour un registre pixel 7 bits (multiples temps de résolution)

Comme nous pouvons le voir sur la Figure 5.16, grâce à la méthode que nous avons appliquée à la mesure en temps basée sur de multiple temps de résolution, l'erreur de quantification est lissée pour qu'elle soit la plus petite possible sur toutes les sorties numériques. En fixant le temps d'intégration $t_{int} = 1\text{ ms}$ et la période système $T_{clk_sys} = 7\text{ }\mu\text{s}$, l'erreur de quantification sur la mesure du temps est réduite, de 5,5 % en utilisant une fréquence d'horloge unique, à moins de 3 % en utilisant 3 fréquences d'horloge différentes. Si on choisit de dégrader un peu le facteur de remplissage en augmentant d'un bit le registre pixel, pour passer à 8, (échantillonnage en 255 intervalles entre t_1 et t_{int}), une erreur de quantification maximale de 2 % comme le montre la Figure 5.17 est atteinte.

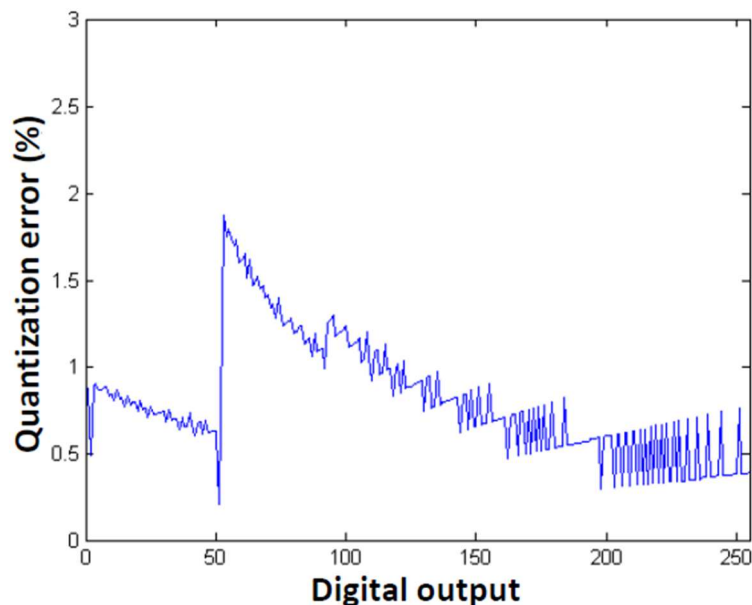


Figure 5.17 : Erreur de quantification pour un registre pixel 8 bits (multiples temps de résolution)

5.6 Modèle en langage VHDL-AMS

La simulation au niveau système en utilisant MATLAB, a montré qu'en utilisant la double mesure temps et amplitude, selon le niveau de la luminosité, la dynamique globale du système est nettement améliorée. Les deux mesures (temps et amplitude) ont été simulées séparément sur MATLAB et la robustesse de l'architecture proposée face à la non linéarité a été montrée également.

Dans cette section, on propose un modèle vhdl-ams du pixel qui implémente l'architecture complète. En effet, le modèle vhdl-ams est plus proche du circuit, il nous permet de modéliser le comportement global du système dans les deux cas de mesures. Les simulations de ce modèle vont confirmer la validité de l'architecture proposée, comme on le verra dans la suite. On présente dans cette partie deux modèles vhdl-ams du pixel qui permettent la simulation des deux mesures en temps : simple et multiple résolution, ce modèle permet également la mesure en amplitude. Ce modèle complètera la validation de notre choix d'architecture, avant de passer à l'implémentation des différents blocs du pixel au niveau transistor.

5.6.1 Modèle du pixel en VHDL-AMS

On propose un modèle vhdl-ams qui comporte l'architecture du pixel proposé. Ce modèle intègre également le compteur numérique système qui réalise l'architecture que nous avons proposé dans la section (5.3). Grâce à l'environnement Cadence qui permet une simulation du langage VHDL-AMS, on a pu réaliser le modèle du pixel et du compteur système comme le montre la Figure 5.18.

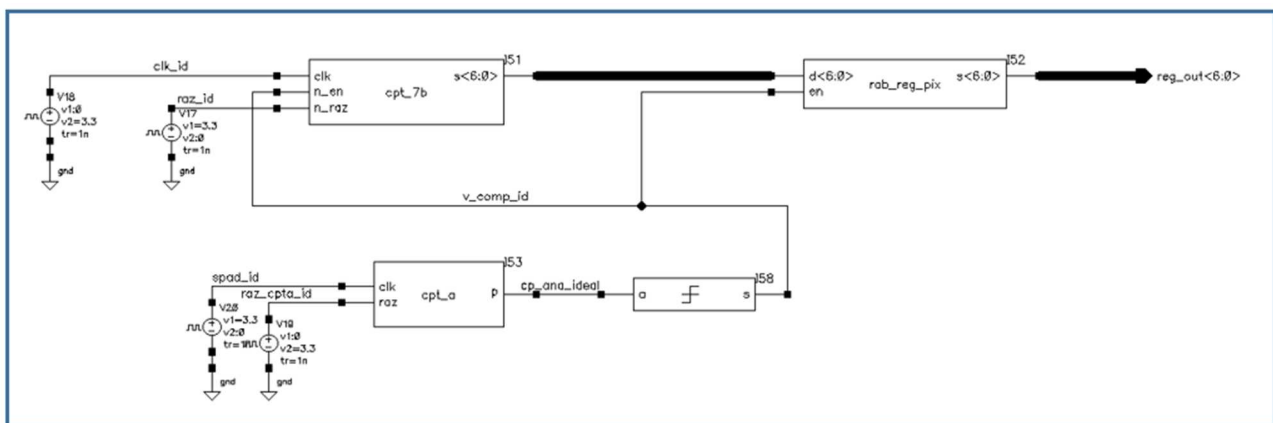


Figure 5.18 : Modèle VHDL-AMS du pixel proposé

Le modèle du pixel en VHDL-AMS comporte les composants suivant :

- Le compteur analogique (cpt_a) : ce compteur analogique est commandé par le signal d'entrée (spad_id). A chaque fois que le signal SPAD envoie une impulsion, le compteur analogique s'incrémente d'un pas de tension ΔV . Une entrée remise à zéro est également présente dans le modèle.

- Le comparateur : il est contrôlé par le signal (cp_ana_ideal), une fois que la tension en sortie du compteur analogique (cpt_a) est supérieure ou égale à la tension de référence V_{ref} , il change l'état de sortie du signal (v_comp_id).
- Le compteur numérique système (cpt_7b) : c'est un compteur numérique de 7 bits. Il a une période système commandé par le signal (clk_id), une entrée remise à zéro (raz_id) et une entrée (v_comp_id) qui commande le compteur numérique, qui va indiquer l'instant où le compteur analogique atteint la tension de référence V_{ref} qui se traduit par le changement d'état du comparateur.
- Le registre numérique pixel (rab_reg_pix) : c'est le composant qui va mémoriser le temps numérique du compteur numérique système. Il est formé principalement de bascules D.

Comme on a vu précédemment, en utilisant qu'une seule période système T_{clk_sys} , la sortie numérique du registre pixel est donnée par la Figure 5.19.

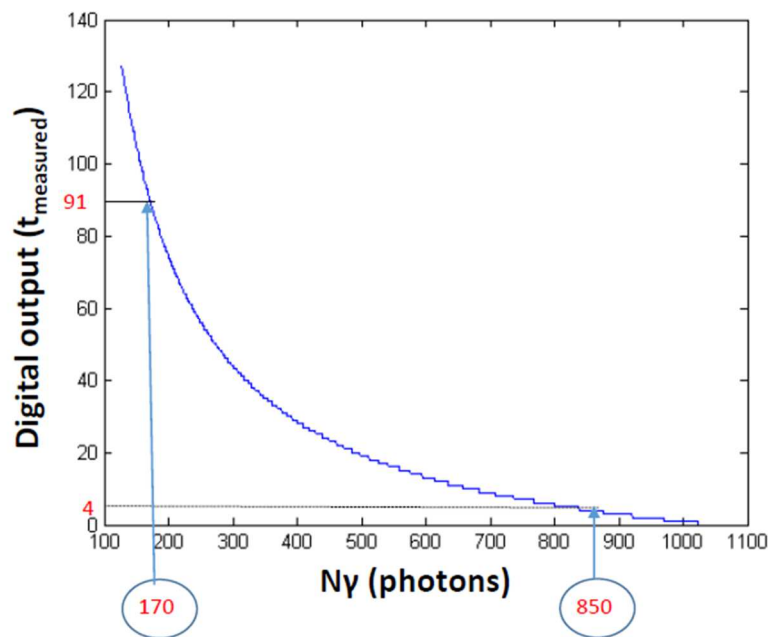


Figure 5.19 : Sortie numérique pour la mesure en temps (simple fréquence système)

Pour vérifier que la mesure en temps fonctionne bien, on a procédé à la simulation du modèle du pixel en vhdl-ams, pour vérifier que les sorties numériques du registre pixel correspondent bien aux valeurs qu'on a eu en simulation de l'architecture au niveau système sur MATLAB. Plusieurs valeurs N_γ sont alors testées, comme on peut le constater sur la Figure 5.19, pour les valeurs 170 et 850 de N_γ , les sorties numériques sont respectivement 91 et 4. Dans un premier temps, on a simulé le pixel en utilisant un compteur analogique idéal qui fournit des pas de tension constant. Les résultats de simulation sont donnés par la Figure 5.20.

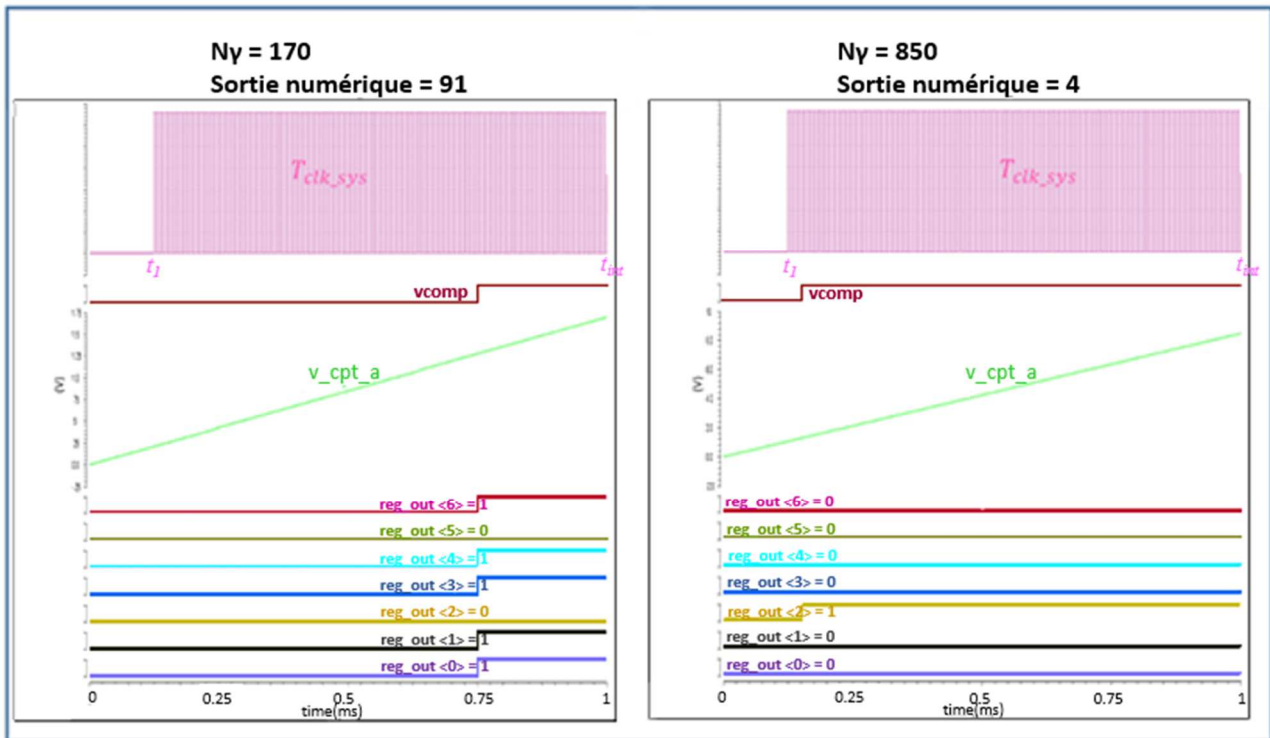


Figure 5.20 : Simulation mesure en temps (simple temps de résolution)

Comme on peut le constater d’après la Figure 5.20, les valeurs de sorties du registre numérique correspondent bien aux valeurs simulées au niveau système avec Matlab (Figure 5.19), ce qui confirme la validité de l’architecture.

Pour tester la robustesse de cette architecture, pour le cas de la mesure en temps, une non linéarité est alors ajoutée au compteur analogique dans le modèle vhdl_ams de la figure Figure 5.18. Le but maintenant est de vérifier si les résultats seront affectés par la présence de la non linéarité.

Le modèle de non linéarité, ajoutée au compteur analogique, est un modèle de non linéarité de type saturation, ce qui fait qu’à chaque fois que le SPAD envoie une impulsion, le compteur analogique s’incrémente d’un pas ΔV . Comme on dispose d’un modèle de non linéarité à saturation, le pas de tension de ΔV , commence à se réduire une fois que le comptage augmente (équation 5.10).

Pour voir l’impact de la non linéarité sur la mesure en temps, on a resimulé le modèle vhdl_ams cette fois ci en présence du modèle de non linéarité pour les mêmes valeurs qu’on a déjà simulé dans le cas où le compteur analogique fournit de pas de tension constant. Les résultats sont donnés par la Figure 5.21.

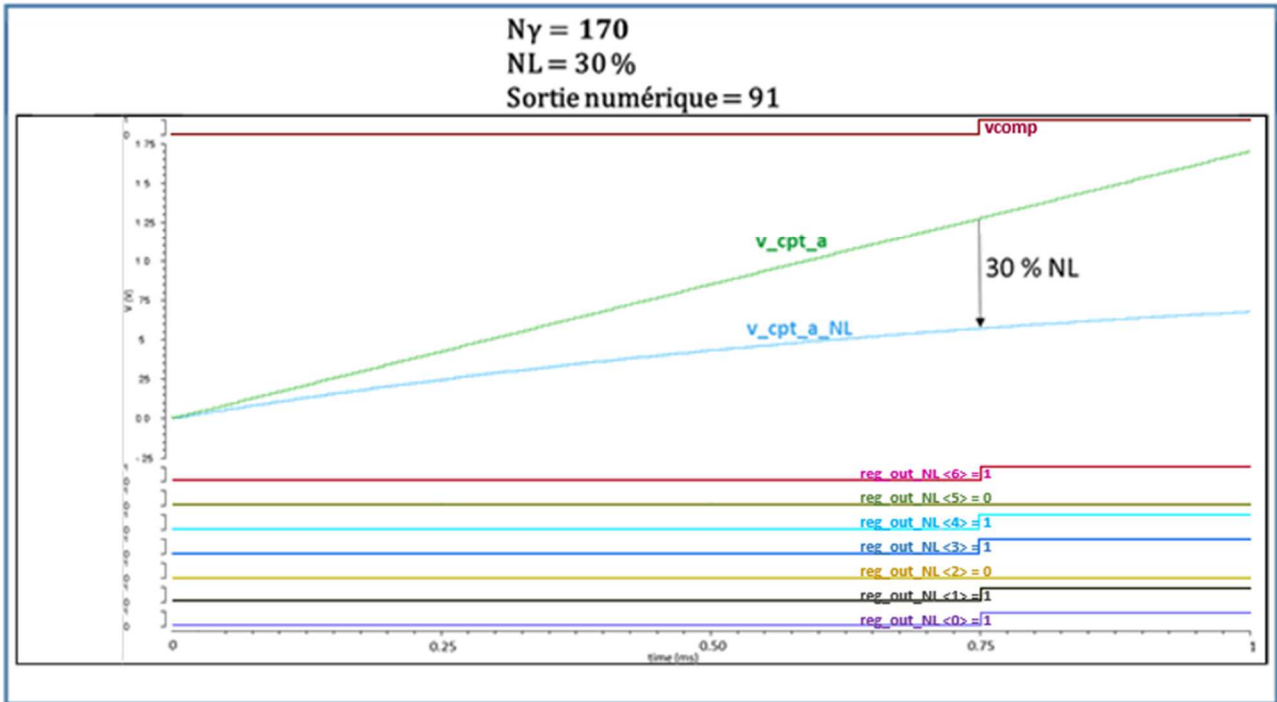


Figure 5.21 : Simulation de la mesure en temps en présence de la non linéarité (NL)

On peut conclure grâce à la Figure 5.21, que cette méthode de mesure en temps est robuste face à la non linéarité. Malgré la présence d'une non linéarité de forte valeur (30 % dans ce cas), la sortie numérique du registre pixel reste intact. Cette mesure a été faite pour différentes valeurs de non linéarité et on peut affirmer qu'on arrive toujours à fournir les 127 pas. Une simulation paramétrique pour toutes les valeurs de N_γ entre 127 et 1023 sous l'environnement cadence est donnée par la Figure 5.22.

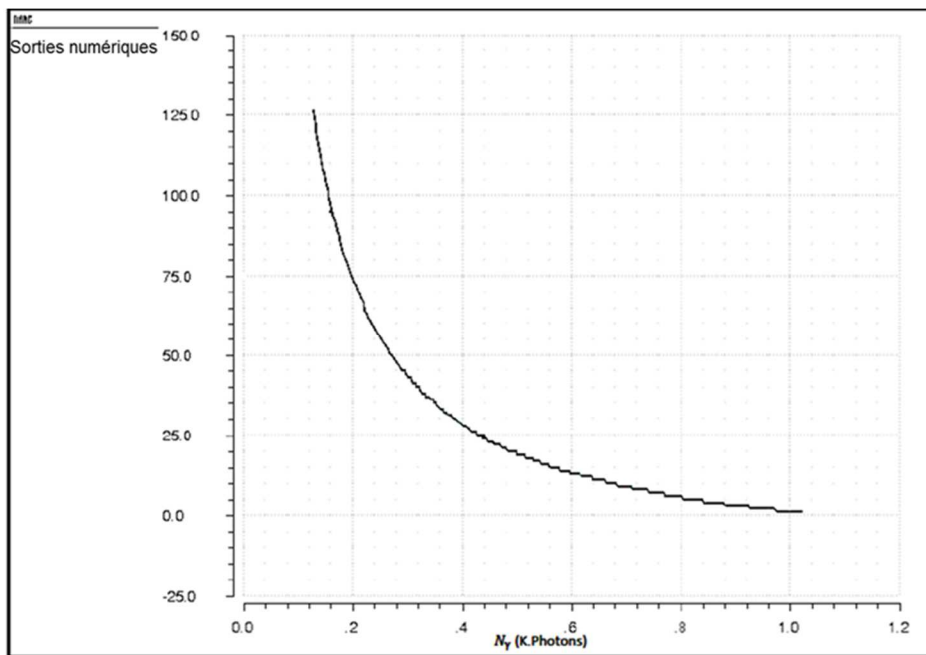


Figure 5.22 : Simulation paramétrique pour les valeurs de N_γ entre 127 et 1023

Comme on peut le constater, en simulant toutes les valeurs de N_γ dans le cas de la mesure en temps, on retombe sur la simulation système Matlab de la Figure 5.12. Cette simulation confirme que le modèle vhdl-ams réalise l'architecture à dynamique élevée du modèle proposé dans la section (5.3).

5.6.2 Modèle vhdl_ams : Multiples temps de résolution

Le premier modèle vhdl-ams qu'on vient de décrire dans la partie précédente évoque le cas de la mesure en temps utilisant une période système unique (T_{clk_sys}). Or dans la section (5.5), on a montré qu'en utilisant un échantillonnage basé sur un simple temps de résolution, l'erreur de quantification pose un problème pour certaines valeurs de sorties numériques. La solution à ce problème était d'utiliser une méthode à multiple temps de résolution, qui fait que l'erreur de quantification est à la fois de plus faible valeur et également lissée sur toute la dynamique.

Dans cette partie, on complètera le modèle vhdl-ams de la section (5.6.1), en intégrant la méthode basée sur le muti temps de résolution afin d'améliorer la résolution du système. Le modèle vhdl-ams intégrera également la mesure en amplitude et, de ce fait, le modèle proposé sera en mesure de simuler le comportement global de l'architecture proposée.

Pour compléter le modèle vhdl-ams qui permet la simulation du comportement global de l'architecture, il a fallu rajouter les composants suivants :

- Le multiplexeur : il a pour rôle de sélectionner la bonne fréquence d'horloge, donc à des instants précis (choisis pour améliorer la résolution), il permettra au pixel de sélectionner la période système optimisée.
- Le bit d'état : qui a pour rôle de préciser la zone de mesure du pixel concerné, 0 pour la mesure en temps et 1 pour la mesure de l'amplitude de tension.

La Figure 5.23, montre l'architecture complète du modèle du pixel en vhdl-ams qui permet de simuler l'architecture proposée.

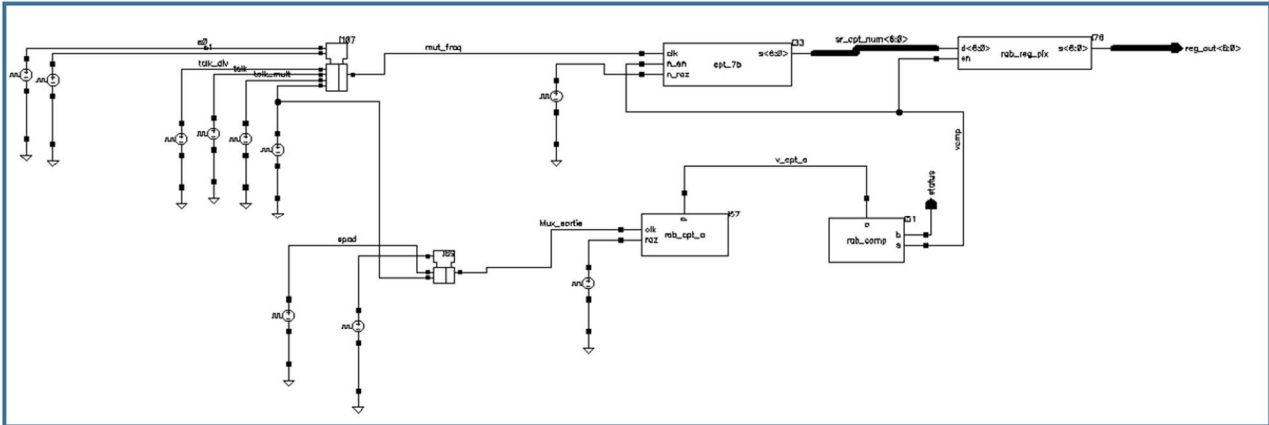


Figure 5.23 : Modèle VHDL-AMS du pixel implémenté pour simuler les deux cas de mesures

Pour vérifier que la simulation de la mesure en temps du modèle du pixel en vhdl-ams basée sur le principe de multiple résolution concorde bien avec les résultats obtenus en simulation en haut niveau, on vérifie les valeurs en sortie du registre pixel. Comme cité précédemment, maintenant à plusieurs instants entre t_1 et t_{int} , la période système change dans le but d'améliorer la résolution en temps du système. La simulation présentée Figure 5.14 montre, après plusieurs essais sur les instants de changement de période système, que en utilisant trois périodes système (équivalent à deux changements de la période système), on arrive à avoir la meilleure distribution de l'erreur de quantification. Dans les simulations du modèle vhdl-ams, on a voulu vérifier que les sorties numériques dans les trois zones de fréquence différente concordent bien avec les valeurs de la simulation au niveau système de la Figure 5.14. La Figure 5.24 montre les résultats de la simulation du modèle vhdl-ams pour la mesure en temps multi-résolution.

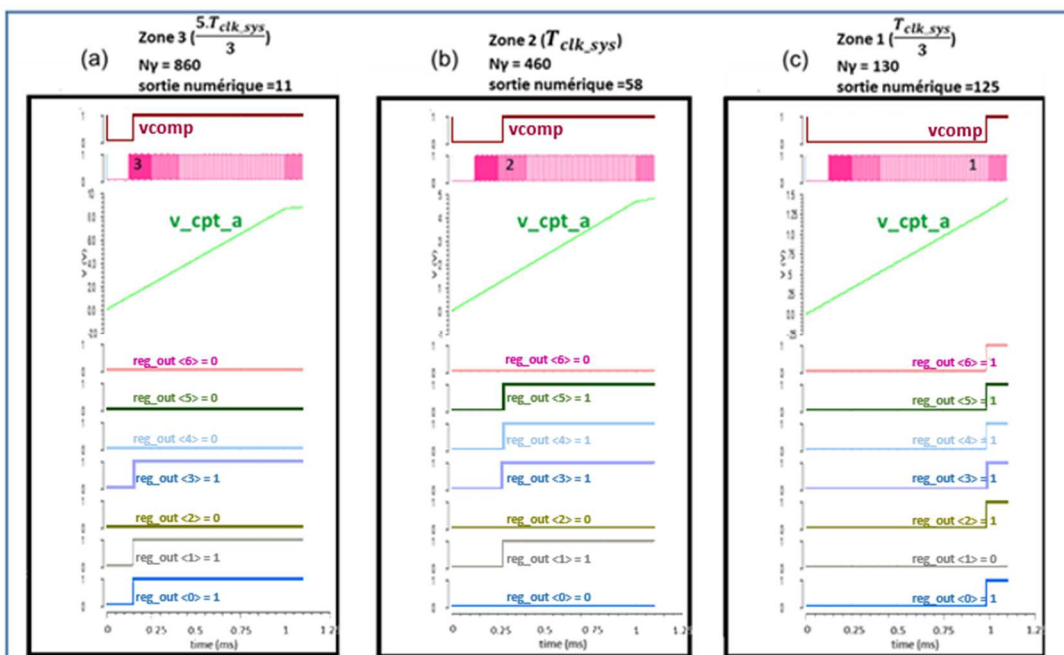


Figure 5.24 : Simulation de la mesure en temps du modèle VHDL-AMS (méthode multi-résolution)

La Figure 5.24 montre la simulation de trois pixels qui sont à fort taux de comptage. Dans ce cas, comme cité précédemment, on procède à la mesure de temps. En appliquant la méthode basée sur le multiple temps de résolution, chacun de ces pixels tombe dans une des zones qui a une période système différente que celle des autres :

- Pixel 1 (Figure 5.24 (a)) : zone 3 avec une période système $T_{clk_sys1} = \frac{5.T_{clk_sys}}{3}$, $N_\gamma = 860$, la sortie numérique $S = 11$, le compteur système bascule à $t = 147 \mu s$.
- Pixel 2 (Figure 5.24 (b)) : zone 2 avec une période système $T_{clk_sys2} = T_{clk_sys}$, $N_\gamma = 460$, la sortie numérique $S = 58$, le compteur système bascule à $t = 258 \mu s$.
- Pixel 3 (Figure 5.24 (c)) : zone 1 avec une période système $T_{clk_sys3} = \frac{T_{clk_sys}}{3}$, $N_\gamma = 130$, la sortie numérique $S = 125$, le compteur système bascule à $t = 980 \mu s$.

On peut constater que les résultats simulés pour les trois valeurs présentées sur la Figure 5.24, on retrouve les mêmes résultats des valeurs numériques de sortie que celles issues des simulations Matlab présentées Figure 5.14 (architecture idéale). Donc l'implémentation du circuit génère des résultats identiques que le modèle théorique.

5.6.3 Modèle VHDL-AMS : Mesure de l'amplitude de tension

Comme cité précédemment, le modèle vhdl-ams de l'architecture du pixel de la Figure 5.23 permet la simulation du comportement global du pixel dans les deux cas de mesures. Pour que le modèle du pixel global soit correct, il a fallu simuler le cas où on a un pixel à faible taux de comptage. Dans ce cas, on ne mesure pas le temps mais l'amplitude de tension.

Dans le cas de la mesure de l'amplitude de tension, comme expliquée dans la section (5.3.3), à l'instant t_{int} la tension en sortie du compteur analogique est inférieure à la tension de référence V_{ref} , le compteur numérique global à tous les pixels est remis à zéro. L'horloge globale du système (Clk_sys) du compteur numérique système est appliquée au compteur analogique. Le comptage continue jusqu'à ce que le compteur analogique atteigne la tension de référence V_{ref} . Finalement, la valeur $(M - N_\gamma)$ est stockée dans le registre numérique du pixel. La Figure 5.25 montre la simulation du modèle vhdl-ams pour le cas d'un pixel à faible taux de comptage.

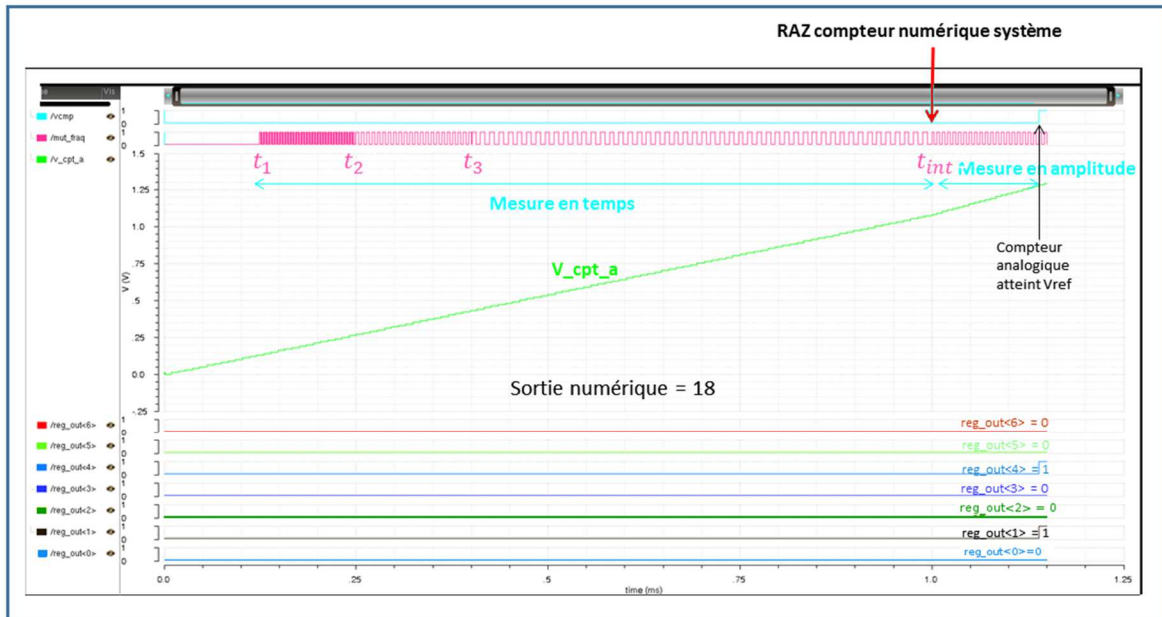


Figure 5.25 : Simulation de la mesure en amplitude du modèle du pixel en vhdl-ams

Cette figure montre que pour une valeur de $N_\gamma = 109$, il va falloir 18 pas pour atteindre la tension de référence V_{ref} . La simulation du modèle du pixel en vhdl-ams pour ce cas montre que la sortie du registre numérique affiche bien la valeur de 18 qui représente l'amplitude de tension nécessaire pour atteindre V_{ref} . Egalement, une série de simulation pour le cas de la mesure en amplitude pour des valeurs de N_γ allant de 1 à 126 a montré qu'on retrouve bien les valeurs numériques $(M - N_\gamma)$ stockées dans le registre numérique pixel, ce qui valide du modèle vhdl-ams pour le cas de la mesure en amplitude.

Cette simulation du dernier cas de fonctionnement du circuit de lecture du SPAD permet de finaliser la validation de notre architecture système, avec une simulation vhdl-ams plus précise que des simulations Matlab, puisque précise au niveau du cycle d'horloge (cycle accurate).

5.7 Conclusion

Nous avons, dans ce chapitre, proposé une nouvelle architecture de lecture de pixel de type SPAD. Cette architecture vise à améliorer de la dynamique, en proposant une nouvelle technique de lecture basée sur une double mesure temps et amplitude.

Un traitement de l'information différent est appliqué au pixel en fonction du niveau de luminosité : une méthode de mesure en temps est utilisée pour les pixels ayant un taux de comptage élevé, alors que les pixels ayant les taux de comptage plus faibles sont traités par une méthode de mesure directe de l'amplitude (qui est directement convertie en valeur numérique au sein même du pixel). On a vu également, qu'un bon choix des instants t_1 et t_{int} , a pour conséquence, une forte amélioration dans la dynamique de lecture tout en gardant un nombre de composants limités.

Pour confirmer l'efficacité de l'architecture, des simulations en insérant de la non linéarité à cette architecture ont montré que cette architecture garde sa robustesse même avec des valeurs de non linéarités élevées.

Pour réduire l'erreur de quantification, une nouvelle méthode basée sur un système multi résolution a été étudié. Grâce à cette méthode, on a pu lisser l'erreur de quantification pour qu'elle soit quasi homogène sur l'ensemble de la dynamique de sortie.

Enfin, un modèle de simulation VHDL-AMS a été présenté et ensuite simulé. Les simulations pour les différentes techniques de mesure en utilisant ce modèle mixte analogique numérique ont confirmé la faisabilité au niveau circuit de l'architecture proposée.

Le travail présenté jusqu'à présent nous a permis de définir de façon précise l'architecture de lecture intégrée au niveau d'un pixel. Dans le chapitre suivant, on présente les solutions matérielles envisageables pour implémenter cette architecture dans une technologie CMOS. Le travail se focalisera en particulier sur le point délicat de cette architecture qui est la réalisation d'un compteur robuste vis-à-vis des dispersions PVT.

Chapitre 6 *Implémentation CMOS du circuit de lecture du SPAD*

6.1 Introduction

Dans le chapitre précédent, une nouvelle technique qui vise l'amélioration de la dynamique globale de lecture du SPAD a été présentée. Le cœur de la nouvelle architecture proposée est le compteur analogique. Pour être plus proche du circuit, un modèle vhdl-ams du pixel qui permet de simuler le comportement global de l'architecture proposée dans un environnement de simulation électronique, a été présenté et simulé. La simulation de ce modèle a permis de valider l'apport de cette architecture notamment sur la dynamique et la résolution globale de système.

Dans ce chapitre, le travail sera focalisé sur l'implémentation en CMOS du circuit de lecture du SPAD dont l'architecture a été présentée dans les chapitres précédents. Dans le chapitre 3, un modèle idéal du compteur analogique a été proposé. En réalité, ce compteur présentera des défauts qu'il faut prendre en considération dès la conception comme : la non linéarité, la dynamique, les dispersions du pas de tension ΔV , etc. Ces paramètres définissent les performances du compteur analogique.

Dans ce chapitre, les architectures électriques des compteurs analogiques détaillées dans le chapitre 3 seront simulées. On propose également dans ce chapitre, une nouvelle architecture électrique à double étage qui sera détaillée et simulée. Cette architecture électrique comporte également le design électrique du comparateur. Les performances obtenues par simulations électriques de l'architecture à base d'un compteur analogique à double étage nous incitent à retenir cette architecture pour qu'elle soit la base de l'architecture de lecture détaillée dans le chapitre précédent.

6.2 Simulation de différents compteurs existants

Comme cité dans le chapitre 3, les circuits de lecture de SPAD sont conçus en général pour réaliser une fonction de comptage. Les solutions numériques sont performantes mais elles exigent une surface importante sur la puce ce qui les rendent difficilement utilisable pour certaines applications. On estime qu'un compteur numérique 7 bits nécessite une surface de $3300 \mu\text{m}^2$, tandis qu'un compteur analogique nécessite une surface de $230 \mu\text{m}^2$ en technologie CMOS ams 0,35 μ .

Les circuits de lecture de SPAD à base d'un compteur analogique représentent une alternative à la version numérique. Une difficulté majeure dans la conception des circuits analogiques est que plusieurs paramètres « parasites » interviennent et affectent les performances électriques. Dans cette section, on simulera électriquement les architectures existantes des compteurs analogiques détaillés dans le chapitre 3. Ces simulations nous permettent d'évaluer les atouts et les limitations de ces circuits afin de proposer dans la section suivante notre propre circuit de lecture.

6.2.1 Compteur analogique à courant commuté

La première architecture d'un circuit de lecture à base de compteur analogique à simuler électriquement est celle de [Palubiak_11]. Cette architecture est composée d'une source de courant, une capacité et un interrupteur. Cet interrupteur, commandé par les impulsions provenant du SPAD, va permettre, au moment de sa fermeture, la recharge de la capacité grâce à la source de courant. Le fonctionnement de cette architecture a été détaillé dans la section (3.4.1). La Figure 6.1 montre la réalisation électrique de ce compteur analogique.

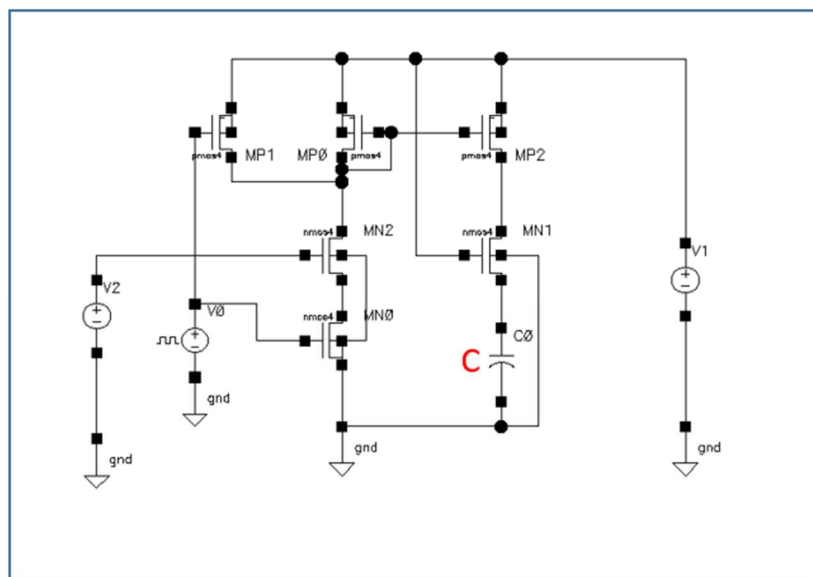


Figure 6.1 : Réalisation électrique de l'architecture [Palubiak_11] simple étage

Comme décrit précédemment dans le chapitre 3, L'impulsion provenant du SPAD, qui a une durée fixe qui représente le temps mort, contrôle la source de courant commutée qui décharge la capacité C . dans le design de la Figure 6.1, la source de courant est remplacée par un miroir de courant. Le transistor (MN2) a pour rôle de contrôler le courant de charge de la capacité c . A chaque impulsion issue du SPAD, le compteur analogique est incrémenté d'une valeur ΔV donnée par l'équation (6.1) .

$$\Delta V = \frac{I \cdot \tau_{SPAD}}{C} \quad (6.1)$$

Où I est le courant de la source de courant contrôlable et τ_{SPAD} est le temps mort imposé par le circuit du SPAD.

Dans un premier temps, une simulation transitoire du circuit pour une valeur de capacité $c = 500 \text{ fF}$ est effectuée. Le but de cette simulation est de vérifier que notre compteur analogique est capable de fournir les 127 pas nécessaire pour atteindre la tension de référence V_{ref} avant qu'il ne rentre en saturation, soit une dynamique équivalente à 7 bits, que nous voulons atteindre. La Figure 6.2 montre la simulation transitoire du compteur à simple étage de [Palubiak_11].

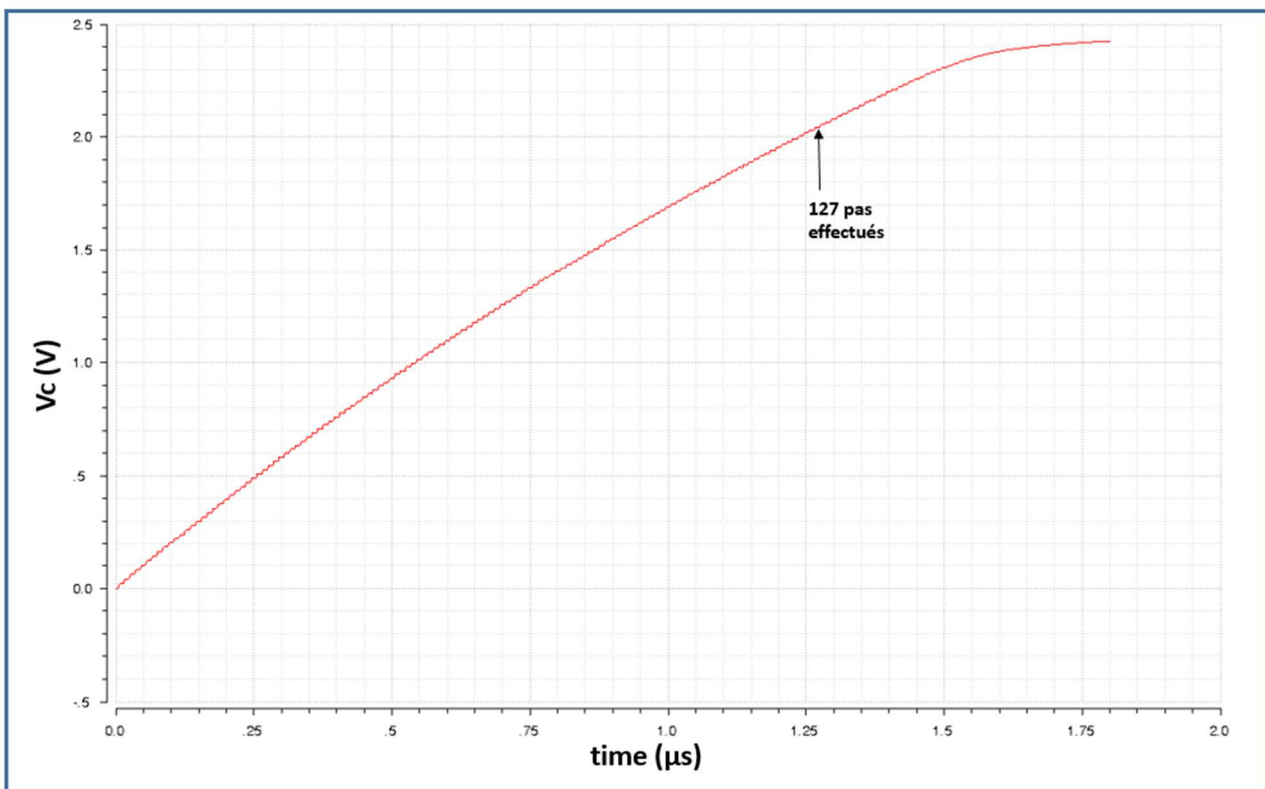


Figure 6.2 : Simulation transitoire de l'architecture de [Palubiak_11] simple étage

Comme on peut le voir sur la Figure 6.2, ce compteur analogique pour une valeur de 500 fF de la capacité c est capable d'effectuer au moins 127 pas nécessaire pour valider

l'architecture proposée dans le chapitre 5. Au-delà de cette valeur, on remarque que le saut entre les pas de tension se réduisent, à cause de la non linéarité des miroirs de courants et le compteur entre en saturation.

Les simulations transitoires ne suffisent pas pour évaluer la robustesse de l'architecture. Pour cela, il faut simuler le compteur en simulation Monte Carlo. En effet, l'objectif de la simulation Monte Carlo est d'émuler la dispersion technologique et donc d'évaluer la robustesse de l'architecture sur un grand nombre (équivalent) de pixel d'une même puce. Cette insensibilité voulue vis-à-vis des dispersions est très importante puisqu'on ne peut pas avoir de réglage individuel des circuits de lecture (réglage de V_{ref}). Pour cela, il faut que tous les circuits de lecture d'un même imageur aient le même comportement. La simulation Monte Carlo en mode mismatch nous permet de vérifier que les dispersions seront suffisamment faibles pour arriver à faire un système de lecture pour l'ensemble des pixels d'un circuit.

Pour tester la robustesse de l'architecture électrique de [Palubiak_11] simple étage face aux dispersions, on a donc procédé à une simulation Monte Carlo. Par exemple, le résultat de 10 simulations Monte Carlo pour une valeur de capacité de 500 fF est montré par la Figure 6.3.

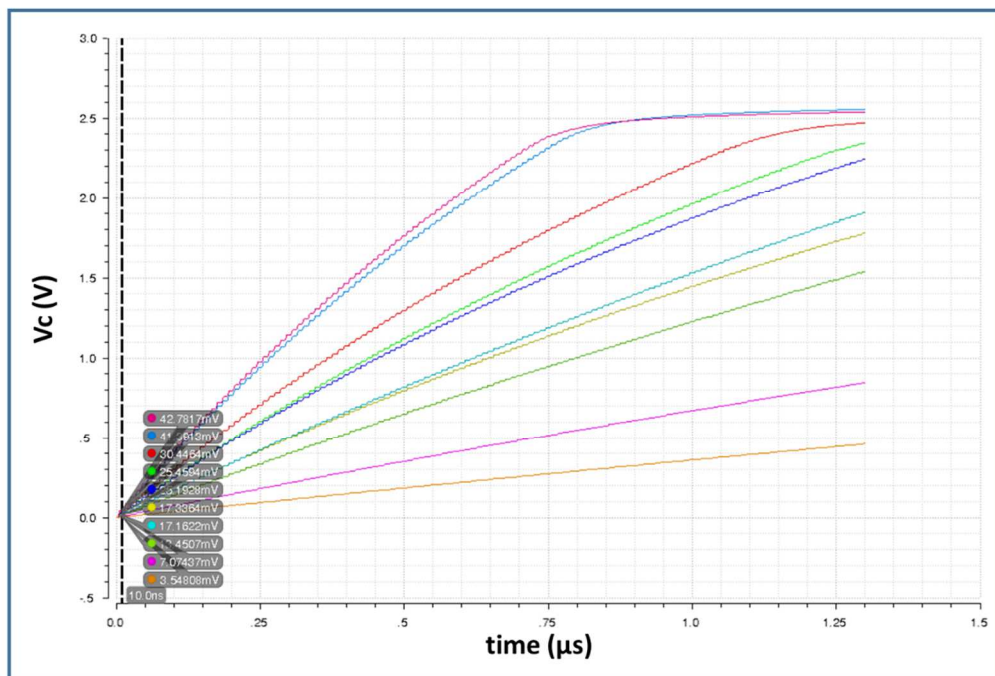


Figure 6.3 : Simulation Monte Carlo pour 10 simulations

La simulation Monte Carlo de l'architecture de [Palubiak_11] simple étage montre les limitations de cette architecture. Cela se traduit par des dispersions au niveau du pas de

tension (de l'incrément ΔV) comme on peut le constater sur la Figure 6.3. Par exemple l'incrément initial varie de 3.4 mV à 42.78 mV. Ces résultats étaient prévisibles, en effet, c'est une architecture composée principalement d'une source de courant et d'une capacité. Un miroir de courant remplace la source de courant dans le circuit électrique. Les défauts de conception liés à la source de courant et à la capacité sont décorrélés et se traduisent par d'énormes dispersions qui se cumulent comme on le constate dans les simulations Monte Carlo. En effet, le pas ΔV dépend aussi de la durée de charge de la capacité c donc ils sont liés au temps mort du SPAD. Ces résultats nous incitent à trouver une architecture différente pour réaliser notre système à haute dynamique.

6.2.2 Compteur analogique : intégrateur à base de capacité commutée

Après avoir essayé l'architecture de [Palubiak_11] à simple étage pour effectuer l'opération de comptage analogique, nous avons eu recours aux circuits à capacités commutées dont on peut espérer une dispersion pixel à pixel limitée en raison de possibilité d'appariement des capacité $C1$ et $C2$ et qui répondent à nos contraintes de dynamique, de sensibilité et de linéarité. Le système est un intégrateur à base de capacités commutées comme le montre la Figure 6.4.

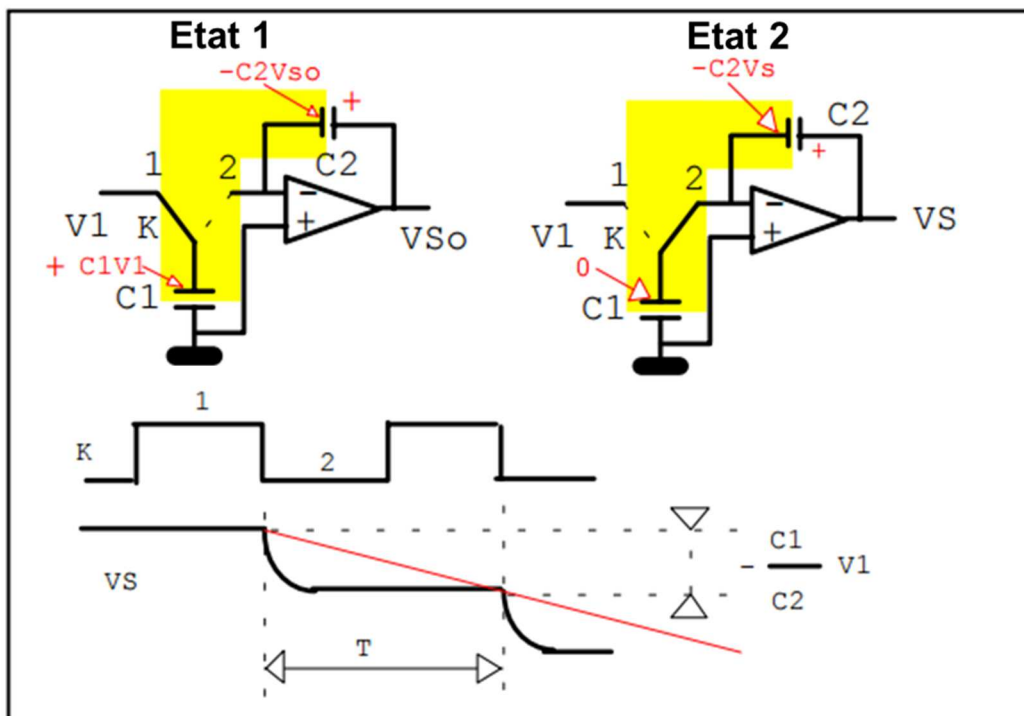


Figure 6.4 : Circuits à capacités commutes (Intégrateur) [Auvray]

D'après cette figure, quand le système est dans l'état 1, la capacité $C1$ est chargée sous une tension $V1$, c'est-à-dire qu'une charge $+C1V1$ est présente sur son armature supérieure. Une charge $-C2Vs0$ est présente sur l'armature gauche de la capacité $C2$. Lorsque l'interrupteur se trouve en position 2, une charge nulle est présente sur l'armature supé-

rière de $C1$ et une charge $-C2V_s$ est présente sur l'armature gauche de la capacité $C2$ [Auvray]. La charge totale dans la zone marquée en jaune est donnée par l'équation (6.2).

$$C1V1 - C2V_s0 = -C2V_s \quad (6.2)$$

On peut conclure la tension de sortie V_s (équation (6.3))

$$V_s = V_s0 - \frac{C1}{C2}V1 \quad (6.3)$$

Sous l'environnement Cadence, on a réalisé le circuit intégrateur comme le montre Figure 6.5.

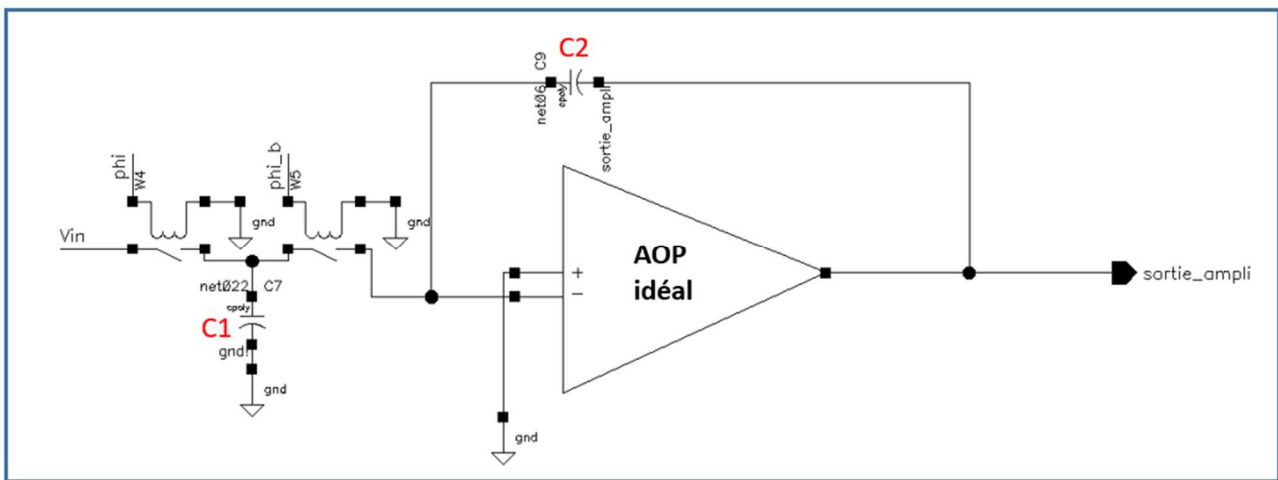


Figure 6.5 : Intégrateur formée avec des capacités commutées

Le schéma électrique de la Figure 6.5 représente le montage inverseur idéal en utilisant un circuit à capacités commutées. Deux capacités de valeurs 58 fF qui remplace les capacités $C1$ et $C2$ de la Figure 6.4. A l'entrée, on a appliqué une tension de -20 mV pour que le système s'incrémente d'un pas $\Delta V = 20$ mV, à chaque fois qu'une impulsion provient du SPAD. Dans un premier temps, une simulation est effectuée avec un amplificateur idéal.

Comme cité précédemment, le gain de l'amplificateur opérationnel constitue un paramètre qui affecte les performances des circuits analogique. Pour cette raison, on a voulu tester l'impact de ce paramètre sur le montage intégrateur de la Figure 6.5, pour cela une simulation paramétrique en modifiant la valeur du gain est effectuée en utilisant l'environnement Cadence. La Figure 6.6 montre les résultats de la simulation paramétrique de la tension en sortie du système.

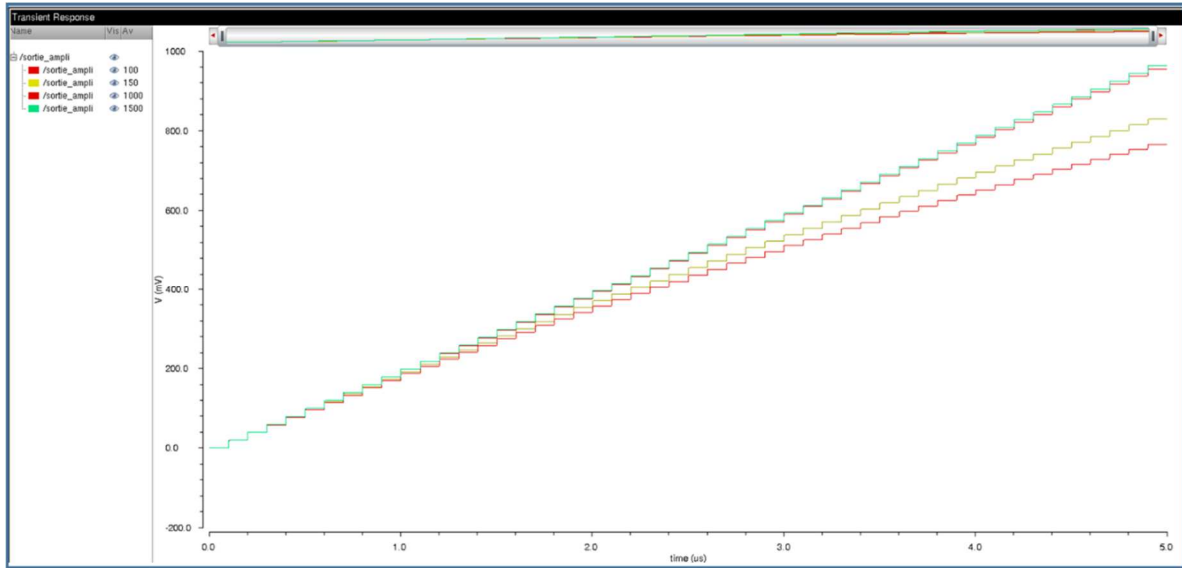


Figure 6.6 : Simulation paramétrique en variant le gain de l’amplificateur opérationnel (100,150, 1000 et 1500)

Le résultat de cette simulation montre que pour les faibles valeurs de gain (100 à 150), la réponse en sortie de l’amplificateur varie énormément ce qui n’est pas le cas pour les valeurs de gain élevées (1000 et 1500) qui montre une réponse quasi identique. Dans le cas d’une intégration CMOS, si on veut s’affranchir des dispersions de la technologie, il faudra idéalement avoir un gain suffisant, au moins supérieur à 1000, ce qui devient difficilement réalisable avec un amplificateur simple étage.

Pour affiner le modèle de l’intégrateur, des simulations avec un modèle au niveau schéma de l’amplificateur opérationnel sont utilisées. Pour cela, on a conçu l’architecture de la Figure 6.5, en intégrant un amplificateur opérationnel à simple étage comme le montre la Figure 6.7.

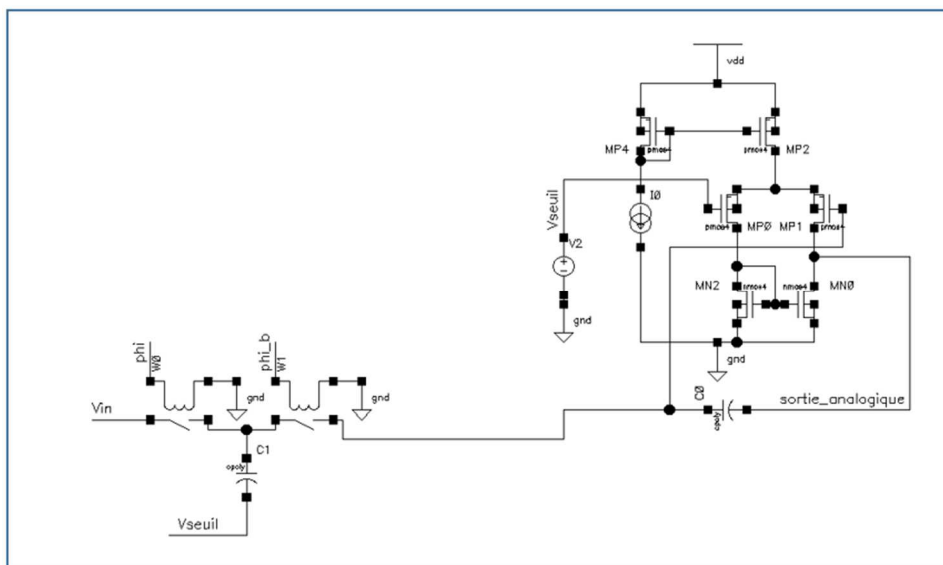


Figure 6.7 : Intégrateur en utilisant un amplificateur opérationnel pmos 5 transistors

Dans cette architecture, un amplificateur opérationnel de type pmos dont le gain est égal à 100 est formé en utilisant 5 transistors. Pour simuler électriquement cet intégrateur, on fixe une tension de seuil (référence), à chaque fois que le SPAD envoie une impulsion, un pas de tension ΔV est décrétement. Par exemple, le résultat de 200 simulations Monte carlo est donné par la Figure 6.8.

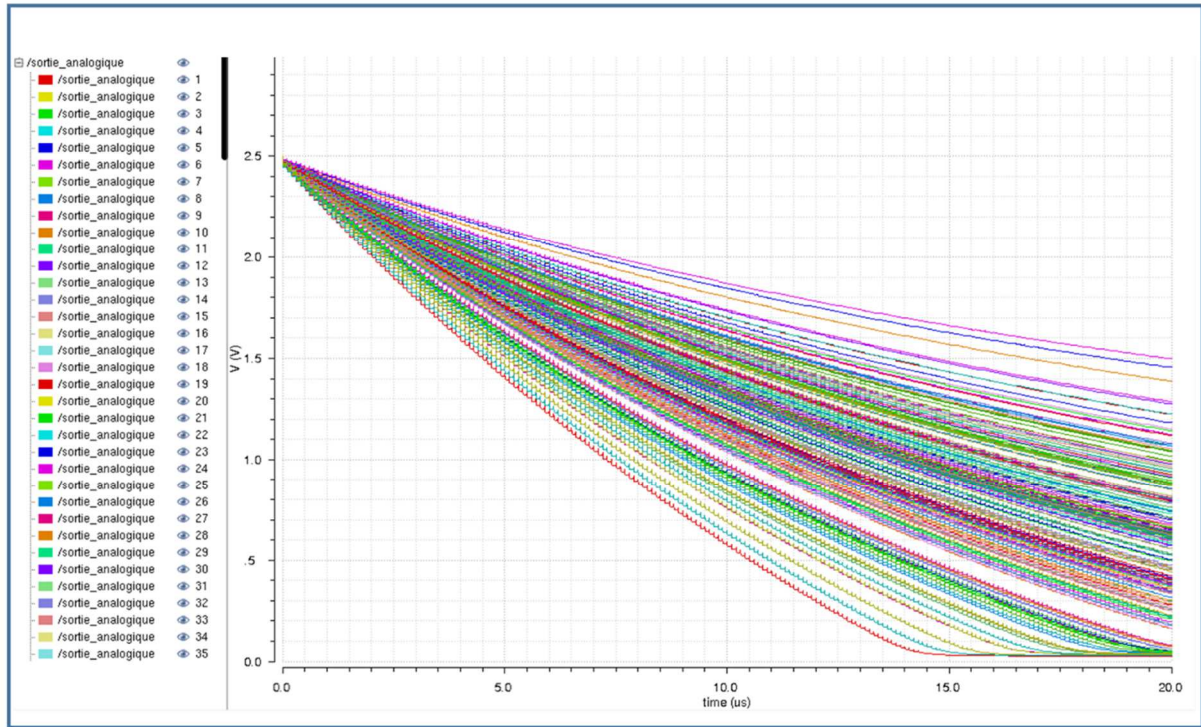


Figure 6.8 : Simulation Monte Carlo du circuit à capacités commutes (intégrateur)

D'après cette figure, on peut constater que cette architecture souffre de problème de dispersions qui affecte le pas de tension ΔV comme dans le cas de l'architecture précédente. En effet, au bout de trois décrétements (dans ce cas un décompteur), l'écart entre les valeurs de tension en sortie du compteur devient trop important, ce qui dégrade énormément la précision et rend cette architecture inadéquate avec le système à dynamique élevée qu'on souhaite réaliser. Le problème dans le design de la Figure 6.7 vient de l'amplificateur opérationnel. En effet, l'utilisation d'un amplificateur opérationnel a engendré l'intégration de l'erreur de l'offset qui est du même ordre de grandeur que notre pas ΔV . Donc cette solution ne peut pas être retenue pour réaliser l'architecture souhaitée. Pour surmonter le problème de l'offset, une solution est d'utiliser un amplificateur opérationnel à auto-zéro, mais dans ce cas, le surcout en circuit serait trop important ce qui nous incite à trouver une architecture moins couteuse qui réalise l'architecture à dynamique élevée.

6.2.3 Compteur analogique par transfert de charge

Les deux architectures détaillées ci-dessus ont montré des très importantes limitations en termes de dispersion, ce qui induit une incapacité à réaliser notre système à dynamique élevée introduit dans le chapitre 5.

Dans le chapitre 3, un modèle électrique de l'architecture basée sur un transfert de charge entre deux capacités a été présenté [Ekaterina_14]. Dans cette partie, on va simuler électriquement le circuit proposé pour voir s'il est en mesure de réaliser notre système souhaité. Le principe est donné par la Figure 6.9.

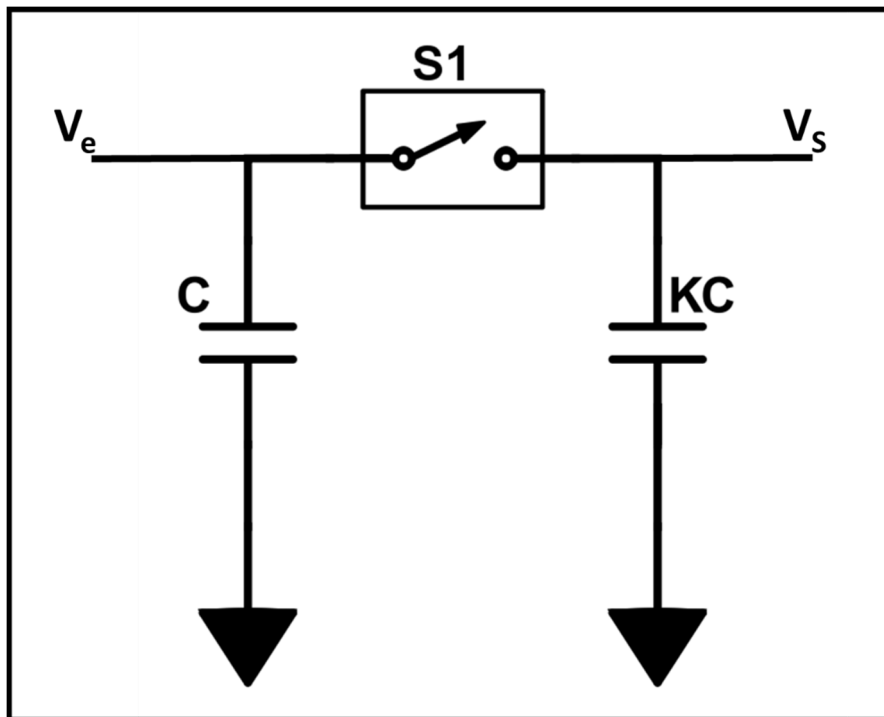


Figure 6.9 : Principe de transfert de charge

Dans cette architecture, le SPAD commande l'ouverture et la fermeture de l'interrupteur $S1$, autorisant ainsi le transfert de charge entre les capacités C et KC . La tension de sortie V_s à chaque impulsion est donnée par l'équation (6.4).

$$V_s(n) = \frac{V_e + K \cdot V_s(n-1)}{K+1} \quad (6.4)$$

Où $V_s(n-1)$ représente la tension en sortie à l'impulsion $(n-1)$ et V_e la tension appliquée à l'entrée.

Comme on peut le constater d'après l'équation (6.4), le coefficient K joue un rôle dans ce transfert de charge. Le rapport $\frac{V_e}{K+1}$ définira le nombre d'incrément maximal que ce

compteur est capable de réaliser pour une tension de référence donnée, ce qui donne une importance cruciale au choix de la valeur du coefficient K.

Dans un premier temps, on a réalisé le schéma électrique utilisé par [Ekaterina_14]. Le transfert de charge dans cette architecture se fait entre une capacité MOS et une capacité parasite c_p . Sous l'environnement cadence, on a réalisé l'architecture de [Ekaterina_14] comme le montre la Figure 6.10.

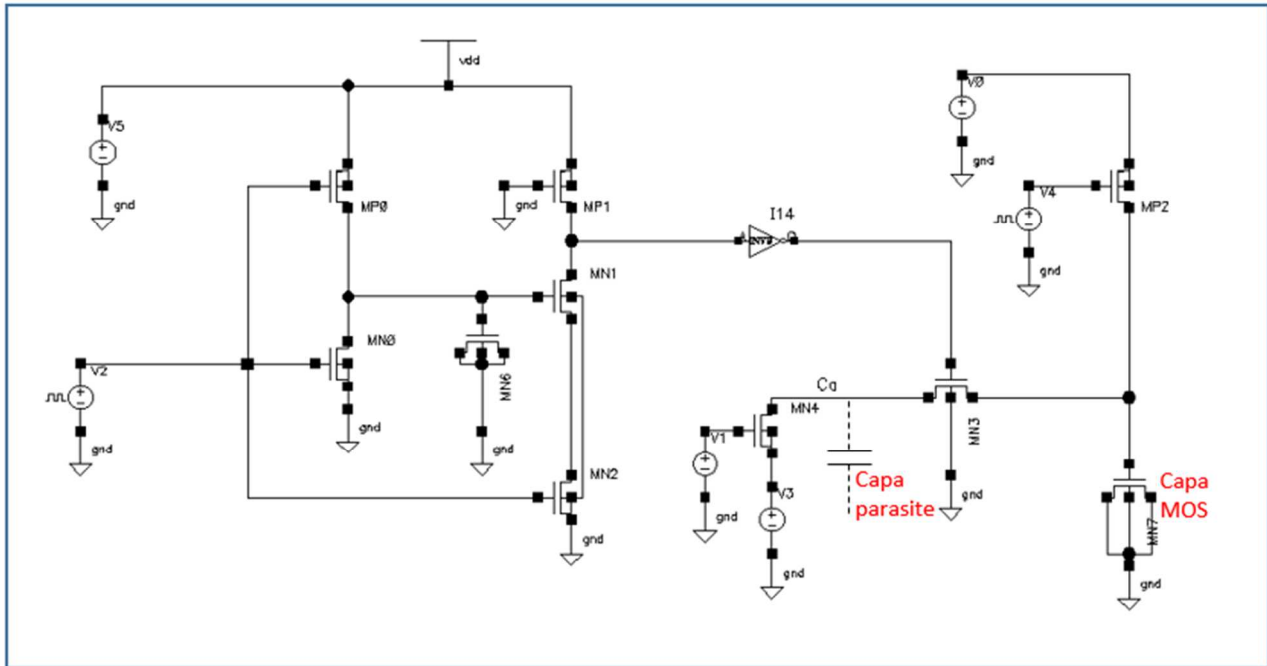


Figure 6.10 : Schéma électrique compteur analogique par transfert de charge [Ekaterina_14]

Dans cette architecture à simple étage, la capacité MOS à une surface de $8 \times 8 \mu\text{m}^2$ ce qui lui donne une valeur de 256 fF et la capacité parasite est égale 2 fF, ce qui fait que le rapport K a une valeur élevée de 128. Pour voir l'impact du coefficient K, on a procédé dans un premier temps à une simulation transitoire. La Figure 6.11 montre le résultat de la simulation transitoire pour un rapport $K = 128$.

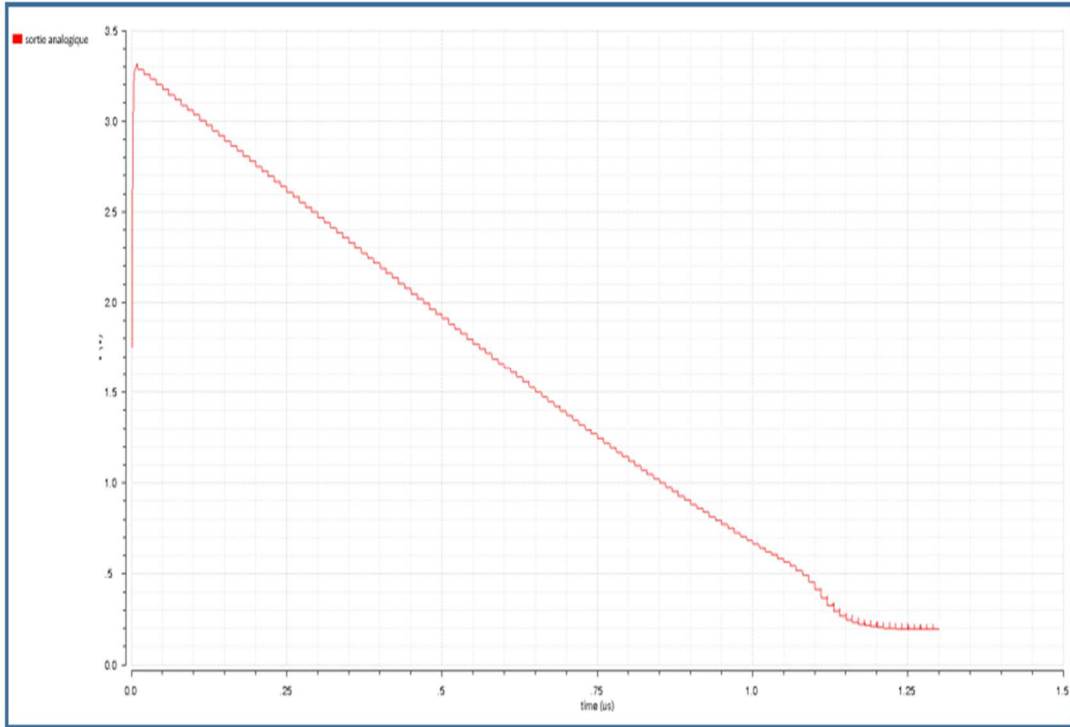


Figure 6.11 : Simulation transitoire pour le schéma électrique de [Ekaterina_14] avec un rapport $K = 128$

Comme on peut le constater sur la Figure 6.11, ce compteur présente une dynamique compatible avec une résolution de 7 bits. Pour tester la robustesse face aux dispersions il a fallu, comme pour les architectures précédentes, simuler le compteur en Monte Carlo. La Figure 6.12 montre le résultat de 200 simulations MC en mode mismatch.

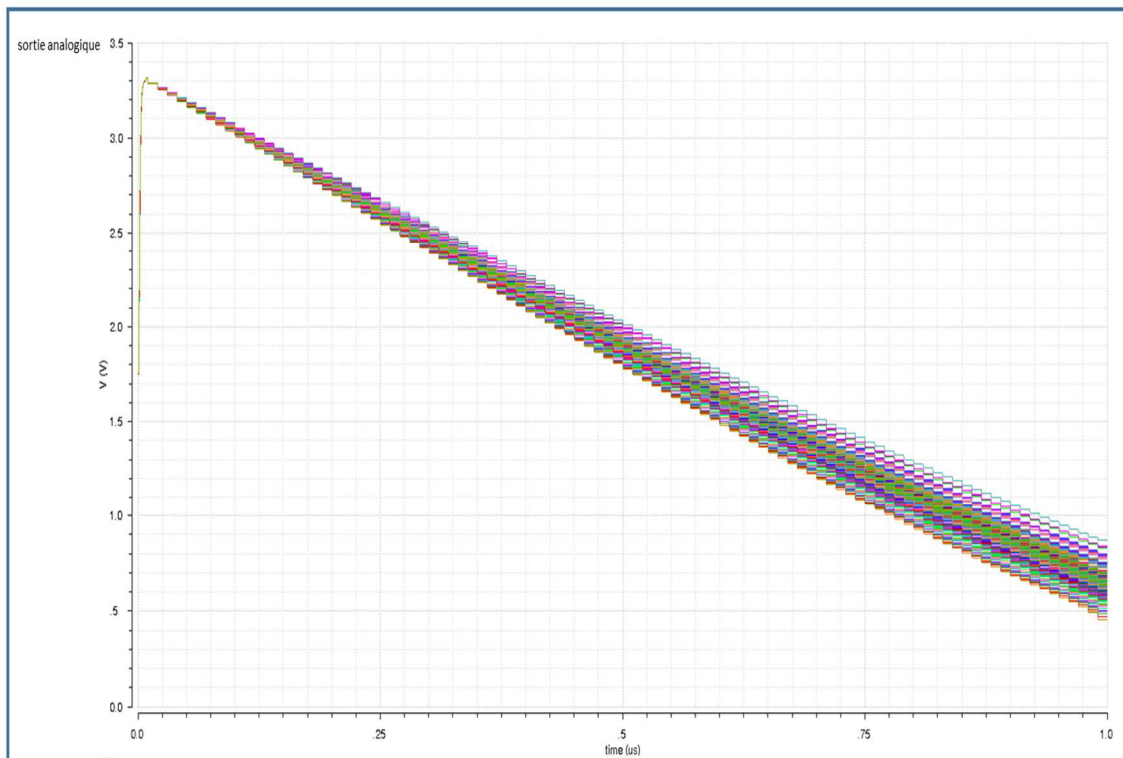


Figure 6.12 : Simulation MC en mode mismatch (200 simulations)

La simulation MC en mode mismatch montre qu'en utilisant le compteur analogique par transfert de charge avec un coefficient K élevé, des problèmes de dispersions perturbent le comportement du circuit. Le résultat de cette simulation confirme la non validité de cette architecture pour réaliser le système à grande dynamique sur une large gamme de pixel. Dans cette architecture, afin d'obtenir un rapport K important, les 2 capacités sont réalisées de façon différente (une capacité parasite et une capacité CMOS), on ne peut donc pas mettre à profit les techniques d'appariement pour limiter l'impact des dispersions technologiques.

6.3 Circuit électrique proposé

Dans la section (6.2), on a détaillé les résultats de simulation des architectures électriques de compteurs analogiques qui existent déjà dans la littérature. Le but de ces simulations était d'analyser la réponse électrique en sortie des compteurs existants afin de déterminer si leurs caractéristiques étaient compatibles avec notre cahier des charges en termes de dynamique, dispersion et robustesse. Ces simulations ont clairement montré que les architectures proposées ne présentent pas des performances suffisantes en termes de dispersion pour être utilisées sans calibration dans un circuit intégrant un grand nombre de pixel. En particulier, il faudrait être en mesure de définir une tension de seuil V_{ref} différente pour chaque pixel tant que les écarts sur le pas en tension de ces compteurs sont importants, donc de faire une calibration du circuit de lecture pixel par pixel. Ceci ne peut être envisagé que si le nombre de pixel intégré est très réduit (et si le facteur de remplissage n'est pas une contrainte).

Dans cette section, on va détailler l'architecture électrique du compteur analogique et du comparateur retenue. Cette architecture inspirée des modèles existants, réalise la fonction qu'on souhaite intégrer pour implémenter l'architecture proposée au niveau système dans le chapitre 5. Les résultats des simulations électriques confirmeront notre choix comme on le verra dans la suite de ce chapitre.

6.3.1 Vue globale du circuit proposé

Comme cité précédemment, l'objectif est de concevoir un compteur analogique capable de fournir 127 pas équivalent aux 7 bits de dynamique tout en respectant un cahier de charge qui prend en considération les paramètres de dynamique, bruit, dispersion et robustesse.

L'objectif est donc d'avoir un compteur analogique ayant une dynamique de 7 bits. Une solution pour atteindre cette dynamique est d'utiliser une architecture de compteur à deux étages. Dans ce cas, pour atteindre une dynamique de 7 bits, il suffit de concevoir un

étage de compteur analogique dont la dynamique est de seulement 11 ou 12 incréments. En effet le principe de ce modèle électrique est similaire à celui utilisée dans le design électrique du compteur de [Palubiak_11] à deux étages et présenté à la section (3.4.1). La Figure 6.13 montre l'architecture du modèle proposée.

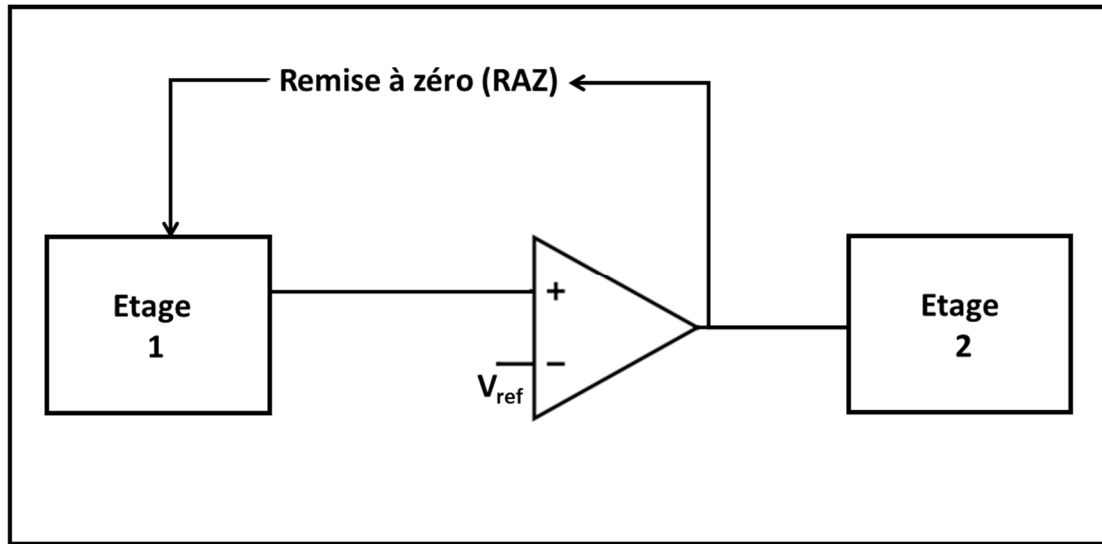


Figure 6.13 : Architecture du modèle électrique proposée

Le fonctionnement du circuit présenté Figure 6.13 du compteur à double étages est le suivant : Lorsque le 1^{er} étage du compteur a atteint sa valeur maximale (la tension de référence), le comparateur autorise l'incrément du 2^{ème} étage du compteur tout en réinitialisant le 1^{er} étage. La dynamique résultante est alors le produit de la dynamique de comptage du premier et du deuxième compteur respectivement.

6.3.2 Compteur analogique

Comme on a vu précédemment, l'architecture à courant commuté et l'architecture à capacités commutées sont limitées par les dispersions de la technologie. On a vu également qu'en utilisant l'architecture par transfert de charge entre une capacité parasite et une capacité MOS, les dispersions affectent le fonctionnement du circuit et le rend inadéquat avec le système souhaité à grande dynamique.

Pour atteindre les performances du système souhaité, notre choix est d'utilisée une architecture basée sur le transfert de charge comme dans l'architecture [Ekaterina_14]. Comme cité précédemment, on dispose d'une architecture à double étage, ce qui veut dire qu'on accepte une dynamique relativement faible par étage (12 incréments max) qui permettra d'atteindre une dynamique globale dans l'ordre de grandeur souhaité (127 pas en tension). Pour cela, on peut se contenter d'un rapport K entre les deux capacités (responsable du transfert de charge) plus faible, donc intégrer dans ce cas, deux vraies capacités C_{poly} , ce qui garantira un meilleur appariement que si on utilise deux capacités de struc-

ture différente. La Figure 6.14 montre le schéma électrique d'un étage du compteur analogique basé sur un transfert de charge entre deux capacités C_{poly} .

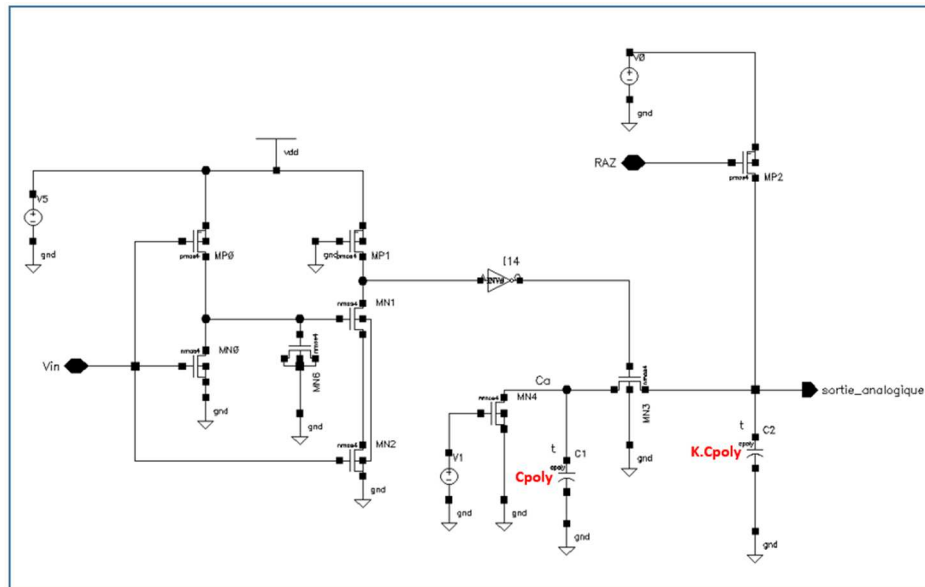


Figure 6.14 : Schéma du compteur analogique à base de transfert de charge

Le transfert de charge dans le schéma électrique de la Figure 6.14 se fait entre deux capacités C_{poly} sous la commande de l'interrupteur formé par le transistor MN3 placé entre les capacités. Des simulations électriques pour différentes valeurs de K (faibles valeurs), vont permettre d'identifier l'impact de ce coefficient comme le montre Figure 6.15.

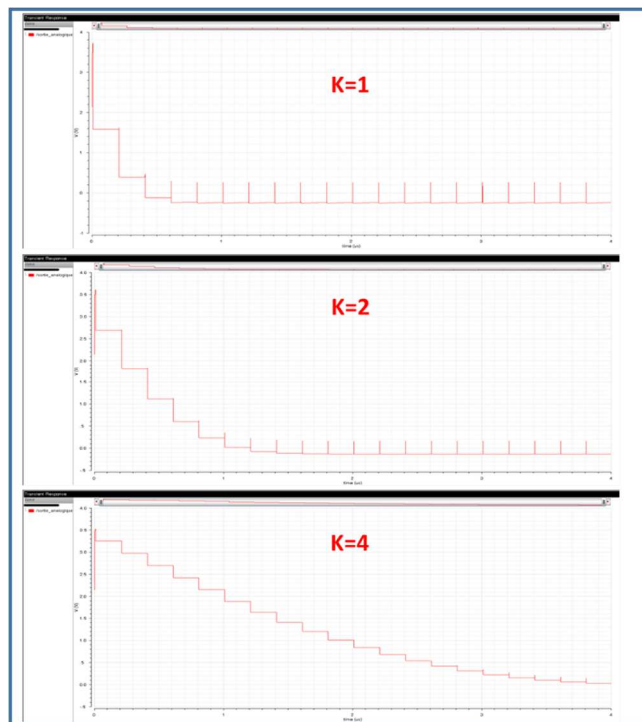


Figure 6.15 : Simulation transitoire pour identifier l'impact du coefficient K

Les résultats de la Figure 6.15 ont confirmé qu'en augmentant le rapport K entre les capacités C_{poly} et $K \cdot C_{poly}$, le nombre de pas de tension ΔV augmente également.

Comme on peut le constater d'après la Figure 6.15, un rapport $K = 4$ suffira pour avoir la dynamique souhaitée par étage. Pour tester la robustesse de cette architecture électrique face aux imperfections de la fabrication, on a procédé à une simulation Monte Carlo en mode mismatch comme le montre la Figure 6.16.

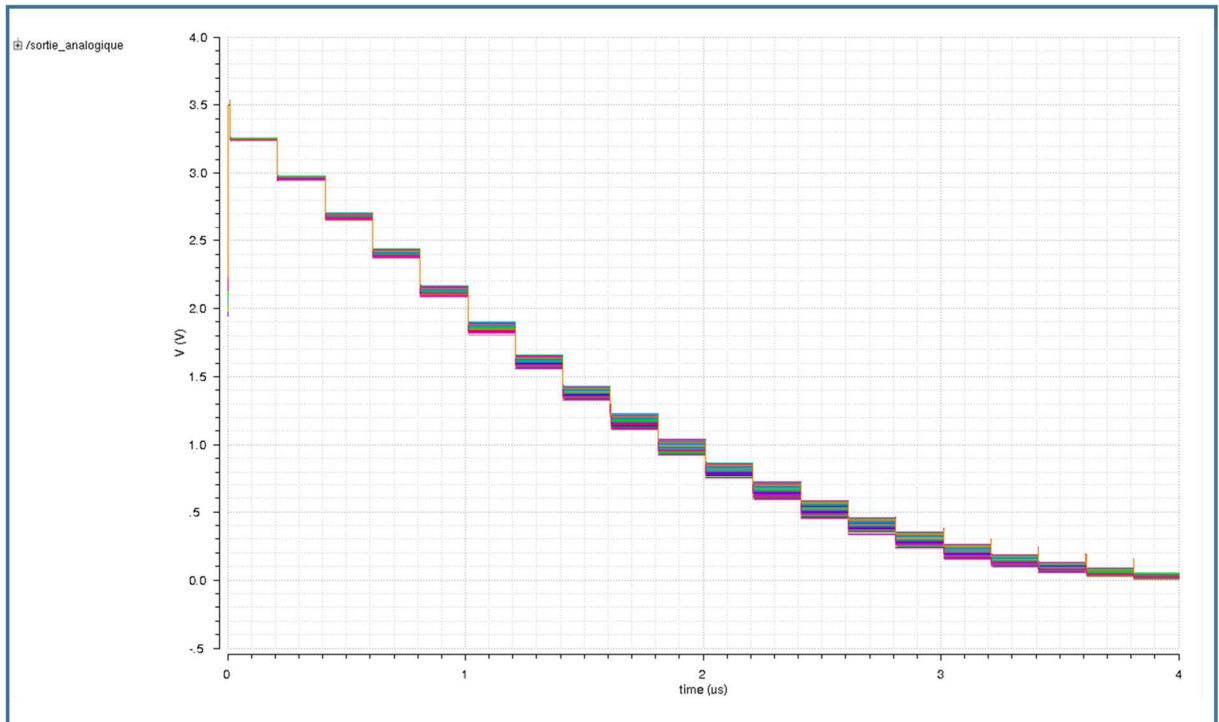


Figure 6.16 : Simulation Monte Carlo (100 Simulations)

Par exemple, la simulation Monte Carlo de 100 itérations de la Figure 6.16, montre que cette architecture basée sur un transfert de charge entre deux capacités appariées résiste bien face aux imperfections liées au design. On a pu grâce à cette simulation compter au moins une dynamique de 12 incrément sans que les dispersions ne dégradent la précision. Pour ces raisons, cette architecture de compteur sera celle retenue pour notre système de lecture de SPAD à grande dynamique.

En zoomant sur la Figure 6.16, on peut voir la marge entre la 10^{ème} et la 11^{ème} impulsion comme le montre la Figure 6.17.

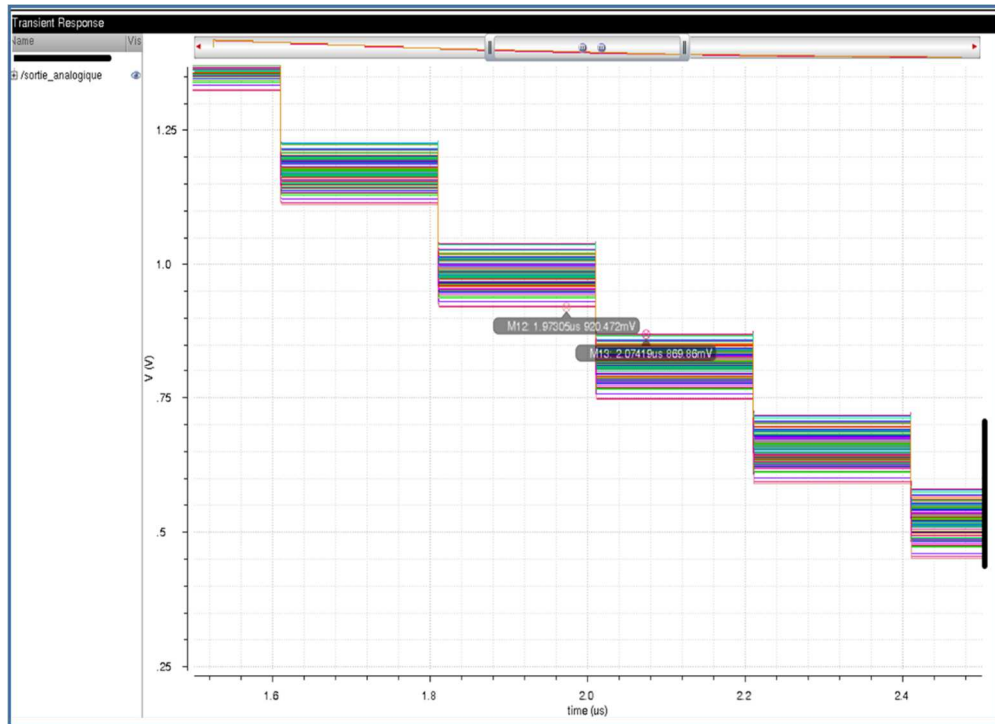


Figure 6.17 : Zoom Figure 6.12 pour voir la marge entre la 10^{ème} et la 11^{ème} impulsions

On déduit de la Figure 6.17 que la marge entre la 10^{ème} et la 11^{ème} impulsion est de 51 mV. C'est cette marge qui va définir le cahier des charges du comparateur comme on le verra dans la suite de cette section.

6.3.3 Comparateur

Après la validation du choix de compteur analogique, on s'intéresse maintenant au comparateur. Ce composant qui a pour rôle d'indiquer à quel moment le compteur analogique 1 atteint sa valeur maximale. Dans ce cas, le signal sortant du comparateur incrémente le compteur analogique 2 tout en réinitialisant le compteur analogique 1. La Figure 6.17 montre la marge entre l'incrément 10 et 11 pour 100 itérations pour le compteur analogique 1.

L'objectif maintenant est de concevoir un comparateur capable, une fois que l'étage 1 passe de l'impulsion 10 à l'impulsion 11, de basculer sa sortie, indiquant ainsi qu'on a franchi la tension de référence autorisant alors la réinitialisation de cet étage et l'incrément de l'étage 2. Dans le design électrique qu'on a développé et comme le montre la Figure 6.17, la marge entre l'impulsion 10 et l'impulsion 11 est de 51 mV. Cette marge est largement suffisante pour garantir un basculement du comparateur au bon instant, même si ce dernier présente un offset important. Il faudra donc fixer la tension de référence V_{ref} à 900 mV pour s'assurer qu'on se positionne bien entre les impulsions 10 et 11. Pour tester le circuit électrique dans un premier temps, on a conçu le comparateur analogique de la Fi-

gure 6.18, avec une entrée qui représente la sortie du compteur analogique et on a fixé $V_{ref} = 900\text{ mV}$.

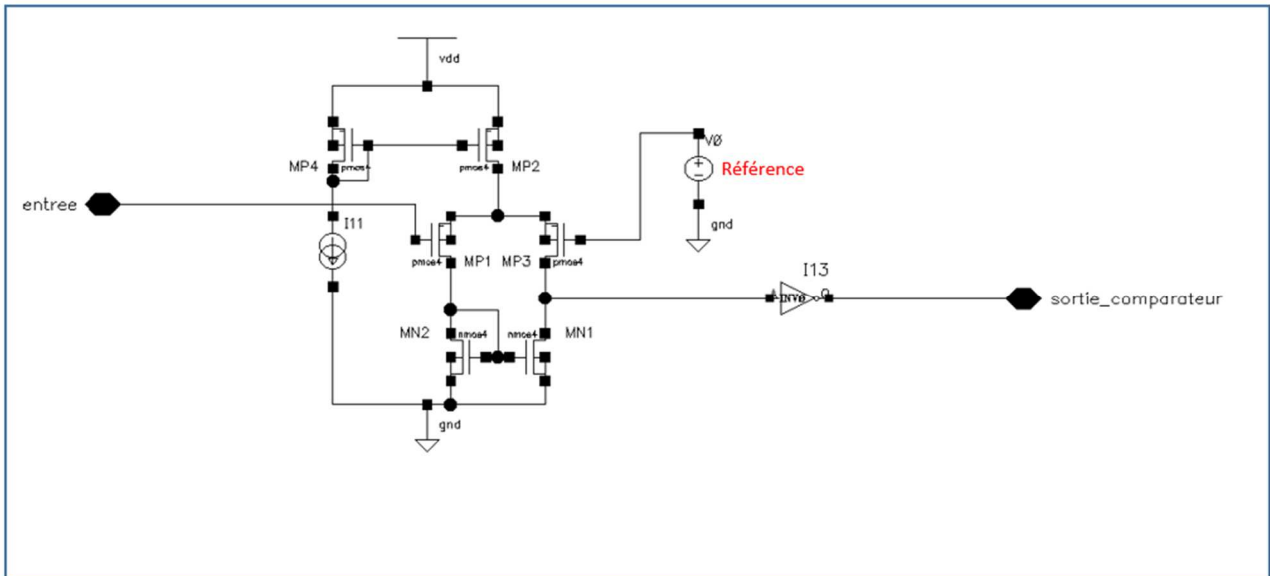


Figure 6.18 : Schéma du comparateur

Dans le design du comparateur, on s'est intéressé surtout à l'offset qui définit la différence entre la valeur maximale et minimale de V_{ref} . En effet, on a vu que la marge entre la 10^{ème} et la 11^{ème} impulsion était de 51 mV, l'objectif est que l'offset du comparateur soit beaucoup moins important que cette marge. 400 simulations Monte Carlo sont données par la Figure 6.19.

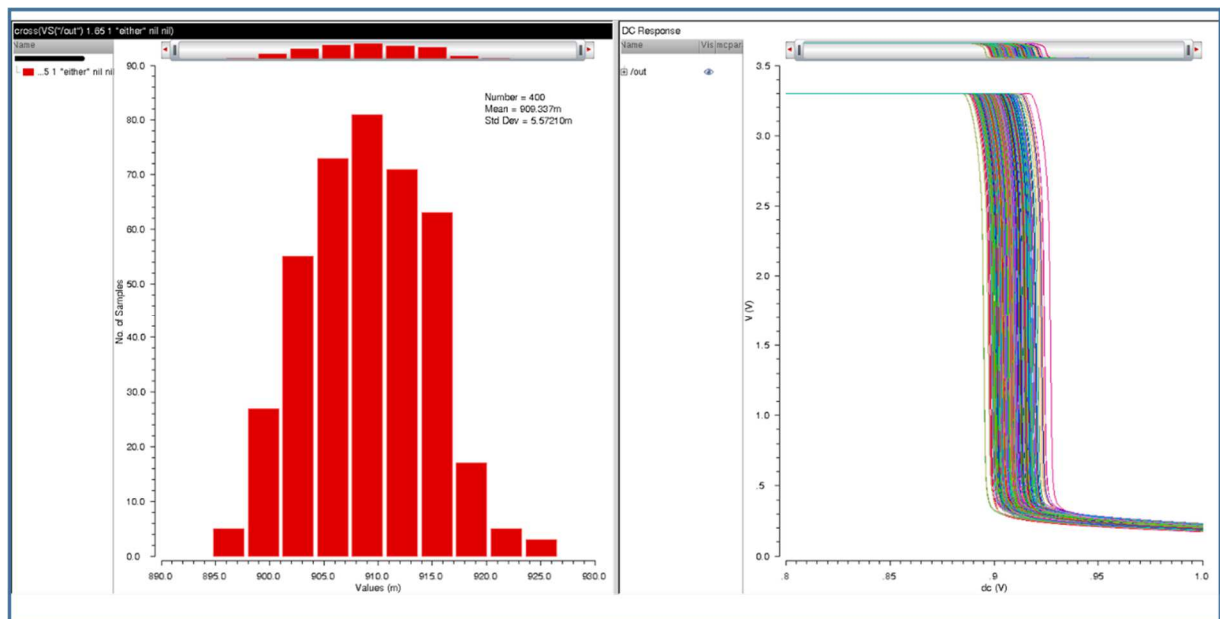


Figure 6.19 : Simulation Monte Carlo en continue comparateur

Les résultats des simulations Monte Carlo mettant en évidence la distribution de l'offset. Ces résultats montrent que l'écart-type de la distribution de l'offset attendu est de

5,5 mV. Avec cette valeur d'offset et pour une marge de 51 mV, on est sûr que le comparateur de tension de la Figure 6.18 est présente un offset suffisamment faible et qu'il commutera bien entre les impulsions 10 et 11 comme on le souhaite.

6.3.4 Compteur complet à deux étages

Afin de réaliser l'architecture du modèle proposé dans le chapitre 5, il faut rajouter le 2^{ème} étage du compteur et le simuler. Comme cité précédemment, dans cette architecture, on dispose d'un compteur à double étage qui, une fois la tension de référence atteinte, incrémente l'étage 2 et réinitialise l'étage 1. La Figure 6.20 montre le schéma électrique du modèle complet.

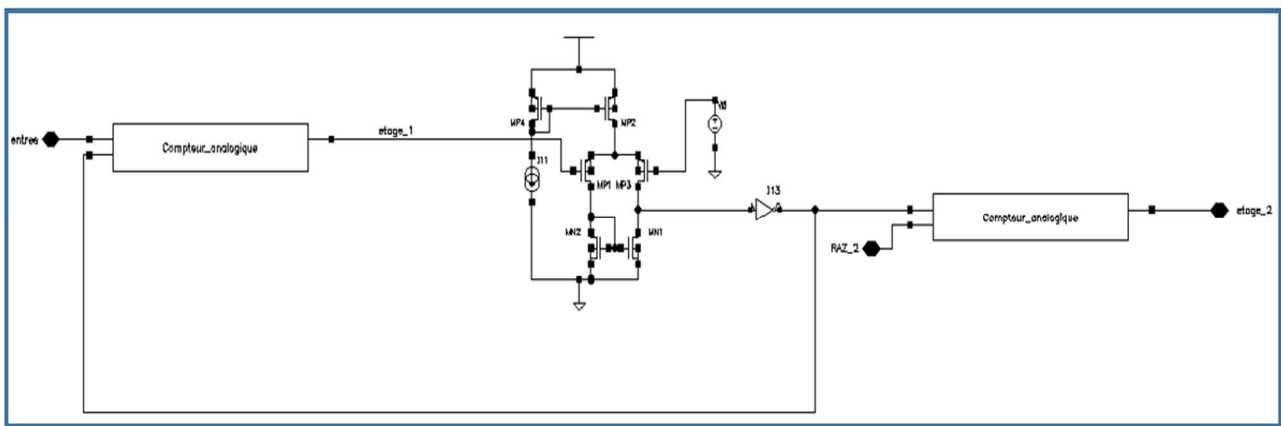


Figure 6.20 : Schéma électrique du modèle complet

La simulation du modèle électrique du compteur à double étage est donnée par la Figure 6.21.

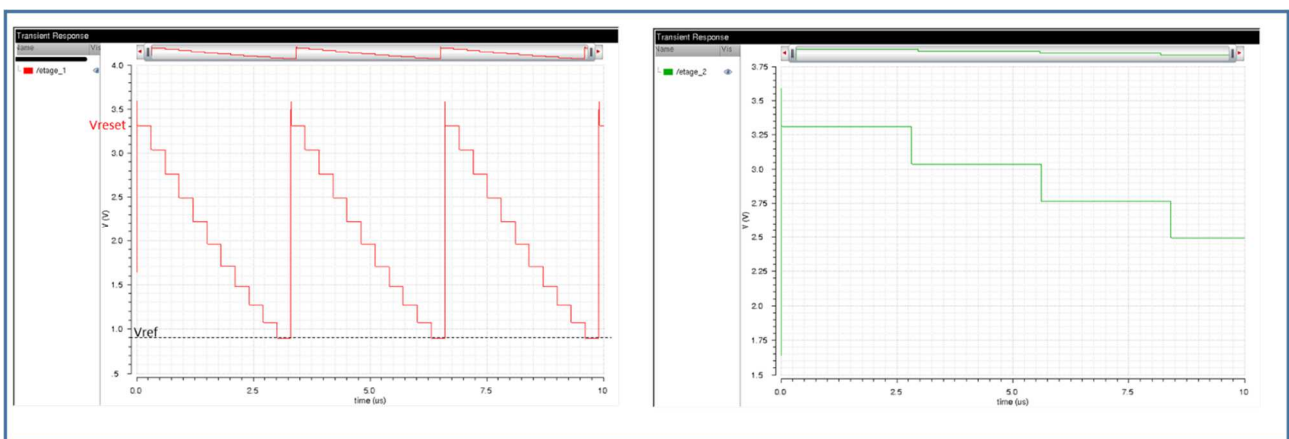


Figure 6.21 : Simulation transitoire des sorties compteur 1 et comparateur (droite) et compteur 2 (gauche)

Sur cette figure, on peut voir le principe du comptage en utilisant cette architecture. En effet, lorsque l'étage 1 du compteur analogique atteint la tension de référence V_{ref} , le comparateur change d'état, ordonnant ainsi la réinitialisation de l'étage 1, et l'incrément de l'étage 2. A chaque fois que l'étage 1 atteint cette tension de référence, on procède à

cette opération, ce qui fait que la dynamique totale du système sera le produit de la dynamique de l'étage 1 et de l'étage 2.

En utilisant l'architecture donnée par la Figure 6.20, notre système sera capable de fournir au moins les 127 pas nécessaire pour réaliser le système à dynamique élevé qu'on a proposé dans le chapitre 5. De plus, ce compteur analogique à double étage est capable de respecter le cahier de charges qu'on a fixé au début de ce chapitre notamment en termes de dispersions. Pour toutes ces raisons, le circuit de la Figure 6.20 est le système retenu pour réaliser l'architecture à grande dynamique.

6.4 Conclusion

Dans ce chapitre, nous avons présenté une réalisation électrique de l'architecture proposée au niveau système du chapitre 5 et qui vise en une amélioration de la dynamique de lecture. En effet dans un premier temps, on a simulé l'architecture à courant commuté de [Palubiak_11] avec un compteur à simple étage. Cette architecture a montré de nombreuses limites notamment dans les résultats des simulations Monte Carlo où les dispersions s'avèrent trop importante, ce qui nous empêche de réaliser notre système souhaité.

La deuxième architecture, utilisant un circuit à capacités commutées, a été étudiée. Malgré sa simplicité, ce circuit a montré ses limites lors des simulations Monte Carlo comme pour le cas du circuit de [Palubiak_11]. C'est la cause pour laquelle cette réalisation n'a pas été retenue.

La troisième architecture simulée est basée sur un transfert de charge [Ekaterina_14]. Dans un premier temps, on a simulé un transfert de charge entre une capacité MOS et une capacité parasite (un rapport des capacités K élevé). Le résultat de la simulation Monte Carlo a montré une légère amélioration face aux dispersions par rapport aux architectures simulées précédemment, mais toujours insuffisante pour réaliser le système souhaité.

Finalement, on a proposé une architecture électrique basée sur un compteur analogique à double étage inspiré du système de [Ekaterina_14], mais cette fois en utilisant un transfert de charge entre deux capacités C_{poly} (et un faible rapport de K). On a montré par les simulations du comparateur et des deux étages que ce système est capable de réaliser l'architecture du système souhaité. Le circuit proposé permet d'atteindre une dynamique de 7 bits pour le compteur analogique et il est suffisamment robuste vis-à-vis des dispersions technologiques pour qu'un imageur complet puisse fonctionner sans avoir besoin de faire une calibration individuelle des circuits de lecture pour chaque pixel.

Chapitre 7 *Conclusion générale et perspectives futures*

7.1 Conclusion générale

Ce manuscrit de thèse synthétise l'étude menée ces dernières années dans le domaine de conception des imageurs CMOS à ultra haute sensibilité. Ces capteurs ont connu des évolutions technologiques importantes lors de ces dernières années, et ont montré leurs utilités en termes d'applications, grâce à leur faible bruit, haute résolution et une plage de dynamique élevée. Parmi les applications qui utilisent les dispositifs de comptage de photon on peut citer : l'imagerie médicale, la vision 3D, la spectroscopie, le temps de vol, la cryptographie, etc.

Les travaux de thèse sont basés principalement sur : l'étude, la conception et la validation de l'architecture de lecture de SPAD.

Le **chapitre 1** a permis de détailler l'état de l'art des différents dispositifs utilisés pour détecter les très faibles niveaux lumineux. En partant du PMT qui a été le premier photodétecteur apte à détecter les photons uniques pour arriver au SPAD. Selon la nature de l'application, on choisit le détecteur approprié.

Dans le **chapitre 2**, nous avons expliqué le mode de fonctionnement des photodiodes à avalanches. On a également expliqué les facteurs de mérites qui caractérisent les SPADs qui sont : la probabilité de détection de photon (PDP), le taux de comptage dans le noir (DCR), la probabilité d'impulsion ultérieure, le Jitter en temps, le temps mort, et la probabilité de diaphonie. La dernière partie de ce chapitre était consacré au circuit d'extinction utilisé pour étouffer l'avalanche, et pour cela, on a détaillé les différentes structures existantes de circuit d'extinctions.

La conception du circuit de lecture de SPAD constitue une phase importante dans le processus de conception d'un imageur. Le but du **chapitre 3** était l'étude de l'état de l'art du SPAD et des circuits de lecture associée au SPAD. Dans la première partie de ce chapitre, on a décrit au fil du temps les évolutions que les SPADs ont subis et leurs designs dans différentes technologies CMOS. Dans la deuxième partie, nous avons présenté une étude des architectures des systèmes de lecture des SPAD. Pour terminer, nous avons détaillé deux designs de circuit de lecture du SPAD basé sur des compteurs analogiques pour lesquels on a présenté les performances annoncées par les auteurs.

Dans le **chapitre 4**, nous avons introduit un modèle haut niveau de l'élément photo-sensible suivi de son circuit de lecture. Ce modèle représente une construction temporelle des instants d'arrivées des impulsions provenant de toutes les sources naturelles ou sources de bruit existantes. L'avantage principal de ce modèle est qu'il prend en considération tous les facteurs de mérites qui caractérisent les performances du SPAD.

Une nouvelle architecture de lecture de pixel de type SPAD a été proposée dans le **chapitre 5**, c'est une technique de lecture basée sur une double mesure temps et amplitude. Grâce à cette architecture, on augmente considérablement la dynamique de lecture globale du système. Des simulations en insérant de la non linéarité à cette architecture ont montré que cette architecture garde sa robustesse même avec des valeurs de non linéarités élevées. En fin de ce chapitre, un modèle vhdl-ams nous a permis de se rapprocher du circuit réel qui a confirmé la validité de cette architecture.

Le dernier chapitre de ce manuscrit de thèse a présenté une réalisation CMOS de l'architecture proposée au niveau système dans le chapitre 5. Dans la première partie, le résultat de la simulation de trois architectures de compteur analogique a été présenté. Les simulations Monte Carlo ont montré l'incapacité de ces solutions à la réalisation du système proposé dans le chapitre 5. Finalement, on a proposé une architecture électrique basée sur un compteur analogique à double étage. On a montré par les simulations du comparateur et des deux étages que ce système est capable de réaliser l'architecture du système souhaité. Le circuit proposé permet d'atteindre une dynamique de 7 bits pour le compteur analogique et il est suffisamment robuste vis-à-vis des dispersions technologiques pour qu'un imageur complet puisse fonctionner sans avoir besoin de faire une calibration individuelle des circuits de lecture pour chaque pixel.

Les travaux de recherche ont contribué à proposer une solution permettant d'améliorer sensiblement la dynamique des circuits de lecture de SPAD par rapport aux solutions proposées dans la littérature.

7.2 Perspectives

Dans le cadre de ce mémoire de thèse, et pour des perspectives à court terme, il reste à valider plusieurs points important du travail qui a été fait :

- Layout : après avoir validé le choix de l'architecture retenu, il faudra passer au layout du circuit proposé pour passer de la validation des simulations électriques à la réalisation du circuit.
- Fabriquer : une fois que le layout est validé, il faudra fabriquer le circuit prototype qui comporte le pixel entier (la photodiode et le circuit de lecture basé sur un compteur analogique à deux étages).
- Tester et caractériser le circuit prototype.

D'ailleurs, en termes de perspectives à moyen terme, l'architecture à dynamique élevée proposée et l'amélioration de la résolution avec le système multi-résolution pourraient s'appliquer sur tous les imageurs plus classiques qui utilisent la notion de la mesure en temps.

A long terme, l'intégration des circuits en 3D dans la conception des imageurs ultra haute sensibilité constituera un axe de recherche important. Ils seront le sujet central de développement pour les années à venir et ouvriront des nouveaux horizons et donneront une importance à la détection des faibles flux lumineux. En effet, l'utilisation de la technologie 3D permettra d'isoler la photodiode (le SPAD) de son circuit de lecture permettant ainsi d'avoir un facteur de remplissage très élevé. On peut par exemple avoir une technologie dédiée au SPAD et en dessous, utiliser une technologie analogique pour le circuit de lecture. Outre le fait d'avoir un facteur de remplissage quasiment égal à 100 %. Cette technologie permettra de diminuer la taille du pixel et donc d'augmenter la résolution.

Bibliographie

[Andor_technology_05]: 'EMCCD for Spectroscopy' URL: <http://www.andor.com/learning-academy/emccds-for-spectroscopy-emccd-technology-in-spectroscopy>.

[Amhaz_12] : H. Amhaz, 'Traitement d'images bas niveau intégré dans un capteur de vision CMOS', Thèse de doct. Université de Grenoble, 2012.

[Auvray] : J. Auvray, 'les amplificateurs intégrés'. URL : http://avrj.cours.pagesperso-orange.fr/Cours/SE_005_Ampli_op.pdf.

[Boisvert_14] : A. Boisvert, 'Conception d'un circuit d'étouffement pour photodiodes à avalanche en mode Geiger pour intégration hétérogène 3D', Thèse de doct. Université de Sherbrooke, 2014.

[Bonanno_05] : G. Bonanno, M. Belluso, S. Billotta and P. Finocchiaro, 'Geiger Avalanche Photodiodes (G-APDs) and Their Characterization', 2005. URL: <http://cdn.intechweb.org/pdfs/17227.pdf>.

[Charbon_13] : E. Charbon, H. Yoon and Y. Maruyama, 'A Geiger Mode APD Fabricated in Standard 65nm CMOS Technology', In IEEE International Electron Devices Meeting (IEDM), 2013.

[Charbon_13b] : E. Charbon, M. Fishburn, R. Walker, R. K. Henderson, C. Niclass, 'SPAD Based Sensors', in TOF Range-Imaging Cameras, Springer-Verlag, pp.11-38, 2013.

[Cheng_12] : J. C. Cheng, 'Self-quenching single photon avalanche photodiodes for near-infrared detection', Thèse de doct. Université de San Diego, 2012.

[Cova_96] : S. Cova, M. Ghioni, A. Lacaita, C. Samori and F. Zappa, 'Avalanche Photodiodes and Quenching Circuits for Single-Photon Detection', Applied Optics, 35.12 (1996), 1956–76 <<http://dx.doi.org/10.1364/AO.35.001956>>.

[Cova_04] : S. Cova, M. Ghioni, A. Lotito, I. Rech and F. Zappa, ' Evolution and prospects for single-photon avalanche diodes and quenching circuits', Journal of Modern Optics, vol. 51, pp. 1267-1288, 2004.

[Dautet_93] : H. Dautet, P. Deschamps, B. Dion, A. D. Macgregor, D. McAcsween, R. J McIntyre, C. Trottier, and P. P. Webb, 'Photon-counting techniques with silicon avalanche photodiodes', Applied Optics, vol. 32, pp. 3894-3900, 1993.

[Dorenbos_11] : S. N. Dorenbos, 'Superconducting Single Photon Detectors', Thèse de doct. Technische Universiteit Delft, 2011.

- [Ekaterina_14] : P. Ekaterina, 'Design and characterization of SPAD based CMOS analog pixels for photon-counting applications', Thèse de doct. University of Trento, 2014.
- [Finkelstein_06] : H. Finkelstein, M. J Hsu and S. Esener, 'An ultrafast Geiger-mode single photon avalanche diode in 0.18 μm CMOS technology . In advanced Photon Counting Techniques, edited by Wolfgang Becker, Proc. of SPIE Vol. 6372, 63720W, (2006).
- [Finkelstein_14] : H. Finkelstein, 'Shallow-trench-isolation bounded single-photon avalanche diodes in commercial deep submicron CMOS technologies', Thèse de doct. Université de San Diego, 2014.
- [Fishburn_12] : M. W. Fishburn, 'Fundamentals of CMOS Single-Photon Avalanche Diodes', Thèse de doct. Technische Universiteit Delft, 2008.
- [Gersbach_08] : M. Gerbash, C. Niclass, E. Charbon, J. Richardson, R. Henderson and L. Grant, 'A Single Photon Detector Implemented in a 130 nm CMOS Imaging Process', 38th European Solid-State Device Research Conference (ESSDERC), 2008.
- [Goetzberger_63] : : A. Goetzberger, R. M. Scarlett, R. H. Haitz, and B. Mcdonald, 'Avalanche effects in silicon p-n junctions .2. Structurally perfect junctions ', Journal of Applied Physics, vol. 34, pp. 1591-1600, 1963.
- [Guillet_11] : T. Guille, 'Rapport de Travaux d'Études et de Recherches Simulation de systèmes de particules en interaction', 2011.
- [Harrison_99] : D. Harrison, R. A. Abram and S. Brand, ' Characteristics of impact ionization rates in direct and indirect gap semiconductors', Journal of Applied Physics volume 85, no. 12 pp. 8186{8192, 1999. doi:10.1063/1.370658.
- [Hsu_10] : M. J. Hsu, 'Development of shallow trench isolation bounded single- photon avalanche detectors for acoustooptic signal enhancement and frequency up-conversion', Thèse de doct. Université de San Diego, 2010.
- [Jradi_10] : K. Jradi,' Etude et Réalisation de photodétecteurs de type APD Geiger pixellisés à très haute sensibilité pour l'astronomie gamma Très Haute Energie', Thèse de doct. Université Paul-Sabatier, 2010.
- [KAZMA_15a] : R. Kazma, O. Rossetto, G. Sicard,' High level model of SPAD based pixel', In Ph.D. Research in Microelectronics and Electronics (PRIME), 2015.
- [KAZMA_15b] : R. Kazma, O. Rossetto, G. Sicard,' High dynamic range readout architecture for SPAD array', Design of Circuits and Integrated Systems (DCIS), 2015.

[Kumar_12] : P. Kumar, 'Time-of-Flight 3D Imaging based on a SPAD-TDC Pixel Array in Standard 65 nm CMOS Technology', Thèse de doct. Technische Universiteit Delft, 2012.

[Lodha_12] : N. Lodha, 'Monolithic Integration of LEDs and SPADs in Standard CMOS Technology for Optical Joystick Application', Thèse de doct. Technische Universiteit Delft, 2012.

[Marmonier_02] : C. Marmonier, 'PHOTOMULTIPLIER TUBES Principles & Applications', 2002.

[Mita_08] : R. Mita and G. Palumbo, 'High-Speed and Compact Quenching Circuit for Single-Photon Avalanche Diodes'. IEEE Transactions on Instrumentation and Measurement, volume 57, numéro 3, p. 543-547, 2008.

[Niclass_05] : C. Niclass, P. Besse and E. Charbon, 'Arrays of Single Photon Avalanche Diodes in CMOS Technology: Picosecond Timing Resolution for Range Imaging (INVITED)'. 1st Range Imaging Research Day, Zurich, Switzerland, 2005.

[Niclass_06] : C. Niclass, M. Sergio and E. Charbon, 'A Single Photon Avalanche Diode Array Fabricated in 0.35 μ m CMOS and based on an Event-Driven Readout for TCSPC Experiments', In SPIE Optics East, Boston, October, 2006.

[Niclass_06b] : C. Niclass, M. Sergio, E. Charbon, 'A Single Photon Avalanche Diode Array Fabricated on Deep-Submicron CMOS Technology', In Design and Test in Europe (DATE), 2006.

[Niclass_07] : C. Niclass, M. Gersbach, R. Henderson, L. Grant and E. Charbon, 'A Single Photon Avalanche Diode Implemented in 130-nm CMOS Technology', In IEEE Journal of selected topics in quantum electronics, vol. 13, no. 4, JULY/AUGUST 2007.

[Osaka_86] : F. Osaka, T. Mikawa and O. Wada, 'Analysis of Impact Ionization Phenomena in InP by Monte Carlo Simulation', Japanese Journal of Applied Physics volume 25, no. Part 1, No. 3 pp. 394-401, 1986. doi:10.1143/JJAP.25.394.

[Palubiak_11] : D. Palubiak, M. M. El-Desouki, O. Marinov, M. J. Deen and Q. Fang, 'High-Speed, Single-Photon Avalanche-Photodiode Imager for Biomedical Applications', In IEEE Sensors Journal (vo: 11, Issue: 10), 2011.

[Pellion_08] : D. Pellion, 'Modélisation, fabrication et évaluation des photodiodes à avalanche polarisées en mode Geiger pour la détection du photon unique dans les applications Astrophysique', Thèse de doct. Université Paul-Sabatier, 2008.

[Repich_09] : M. Repich, D. Stoppa, L. Pancheri and G. F. Betta, 'Simulation modelling for the analysis and the optimal design of SPAD detectors for time-resolved fluorescence

measurements', In Proc. SPIE 7355, Photon Counting Applications, Quantum Optics, and Quantum Information Transfer and Processing II, 2009. doi:10.1117/12.820430.

[Richardson_12] : J. A. Richardson, L. Grant, D. Renshaw and R. Henderson, 'A Single-Photon Avalanche Diode in 90-nm CMOS Imaging Technology With 44% Photon Detection Efficiency at 690 nm', In IEEE Electron Device Letters, vol. 33, Issue: 5, 2012.

[Tisa_08a] : S. Tisa, F. Guerrieri and F. Zappa, 'Monolithic Array Of 32 SPAD Pixels For Single-Photon Imaging At High Frame-Rates', New Developments in Photodetection, 5th International Conférence on New Developments in Photodetection, 2008.

[Tisa_08b] : S. Tisa, F. Guerrieri and F. Zappa, 'Variable-load quenching circuit for single-photon avalanche diodes', Optics Express, volume 16, numéro 3, p. 2232-2244.

[Zappa_04] : F. Zappa, S. Tisa, S. Cova, P. Maccagnani, D. Bonaccini Calia, G. Bonanno, M. Belluso, R. Saletti, and R. Roncella, 'Pushing technologies: single-photon avalanche diode arrays', presented at Advancements in Adaptive Optics, 2004.

[Xiao_07] : Z. Xiao, D. Pantic and R. S. Popovic, 'A new single photon avalanche diode in CMOS high-voltage technology'. In Solid-State Sensors, Actuators and Microsystems Conference, 2007.

Liste des publications

Conférences internationales

- [1] R. Kazma, O. Rossetto, G. Sicard, ' High level model of SPAD based pixel', In Ph.D. Research in Microelectronics and Electronics (PRIME), 2015.
- [2] R. Kazma, O. Rossetto, G. Sicard, ' High dynamic range readout architecture for SPAD array', Design of Circuits and Integrated Systems (DCIS), 2015.
- [3] R. Kazma, O. Rossetto, G. Sicard, 'Reducing Quantization Error in Time Measurement Technique for SPAD Readout Circuit', In IEEE International Conference on Electronics, Circuits and Systems (ICECS), 2016.

Conférence nationale sans comité de lecture :

- [4] R. KAZMA, O. Rossetto, G. Sicard, , "Les photodiodes à avalanches polarisées en mode Geiger (Single Photon Avalanche Diode, SPAD) CMOS à grande dynamique,"18th Journées Nationales du Réseau Doctoral en Micro-nanoélectronique (JNRDM), Bordeaux, Mai 2015.

Résumé

Il existe une multitude d'applications où la détection de très faibles niveaux lumineux (jusqu'à quelques photons, voir le photon unique) est un point clé de la mesure dans des domaines tels que l'imagerie médicale ou l'astrophysique. Jusqu'à un passé récent, le seul composant permettant de détecter des niveaux aussi bas que le photon unique était le photomultiplicateur. Ce composant, volumineux et fragile, nécessite l'utilisation de hautes tensions et possède une résolution spatiale faible. Depuis quelques années, la faisabilité de photodétecteurs à photons uniques connus sous l'acronyme de SPAD (pour Single-Photon Avalanche Diode) dans un procédé de fabrication CMOS standard a été démontrée. Ces avancées permettent de réaliser des imageurs de haute résolution et grande sensibilité. Le travail dans cette thèse démarre par la compréhension du comportement de l'élément photosensible (le SPAD) pour finir par la conception du circuit de lecture associé au SPAD. Le deuxième axe de travail proposé dans cette thèse est un nouveau modèle haut niveau du SPAD suivi par son circuit de lecture. Le troisième axe propose une nouvelle architecture de lecture qui vise l'amélioration de la dynamique globale. Finalement, dans le dernier chapitre on propose une réalisation électrique CMOS du circuit à dynamique élevée.

Mots Clés – *Capteur d'image CMOS, SPAD, Compteur analogique, Modèle haut niveau, Circuit de lecture, Dynamique élevée.*

Abstract

A large number of vision applications (medical imaging, astrophysics, ...) require the detection of very low light levels (up to a few photons, or the single photon). Until recently, only one component could detect levels as low as the single photon: the photomultiplier. However, this component is bulky, fragile, requires the use of high voltages and has a low spatial resolution. Since few years, it has been demonstrated the feasibility of designing avalanche diode using CMOS standard process (SPAD, Single-Photon Avalanche Diode). These advances allow to achieve high-resolution imaging with a high sensitivity. This PhD work starts with the study of the behavior of the photon sensing element and finishes by the design of the readout circuit of SPAD. The second point of the work proposed in this thesis is a new high-level model of SPAD followed by its readout circuit based on an analog counter. The third line offers a novel readout method to enhance the readout dynamic range of pixel based on single photon avalanche photodiode. Finally, this new high dynamic range readout circuit is designed for CMOS SPAD based pixel.

Keywords – *CMOS Image sensor, SPAD, Analog Counter, High Level Model, Readout Circuit, High Dynamic.*