

Doctoral Dissertation
Doctoral Program in ELECTRICAL, ELECTRONICS AND COMMUNICATIONS ENGINEERING (32.th cycle)

Development of CMOS Integrated Circuits for SiPM Readout

Weishuai Cheng
* * * * * *

Supervisors
Prof.Amedeo Staiano, Supervisor
Prof.Michele Goano , Co-supervisor

Doctoral Examination Committee:
Prof.Gianluca Traversi, Referee,Università degli Studi di Bergamo
Prof.Piero Giubilato, Referee, Università di Padova

Politecnico di Torino
To be define, 2020

I hereby declare that, the contents and organisation of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

. . . . . .

Weishuai Cheng

Turin, To be define, 2020

Summary

Silicon photo-multiplier (SiPM) is a new generation of solid-state photon sensors, invented by Russian scientists during the 1990s, which working principle is based on the operation of photo-diode arrays in Geiger mode. Similarly to the conventional Photon-Multiplier Tube (PMT), a SiPM features a high gain in the same range, high Photon Detect Efficiency (PDE) and excellent time response. On top of that, SiPMs provide important advantages such as low operating voltage, compactness and insensitivity to magnetic fields. Initial drawbacks in terms of high cross-talk, dark count and generally high correlated noise, has been significantly reduced with constant improvements in manufacturing technology. As a consequence, SiPMs became an attractive choice for an increased number of applications in nuclear medicine, high energy and astroparticle physics experiments, automotive (LIDAR) and homeland security.

ASICs (Application Specified Integrated Circuits) have been used in radiation sensor readout for more than 40 years, providing specific solutions in terms of signal amplification, shaping, and digitization. The development of custom integrated electronics for the readout of highly segmented SiPM matrices allowed for the construction of compact photoelectronic modules for detectors implementing aggressive spatial resolution and channel density.

This thesis exploits the development of innovative CMOS integrated electronics for SiPM readout. The versatility of the architecture of the ASICs herein described should allow for the use of these electronics in system-grade detectors for medical imaging, particle and nuclear physics, using both room temperature inorganic (crystal) scintillators or the fast scintillation of compton scattering in novel noble-liquid based PET detectors operating at cryogenic temperatures.

This manuscript provides detailed description, analytical depiction and silicon characterisation results of the front-end CMOS circuitry to readout SiPM sensors at liquid nitrogen temperature and room temperature. In order to study and verify the functionalities of the basic circuit modules for a chip-level integration.

Then, a 32-pixel mixed-signal ASIC for the readout of SiPM matrices at cryogenic temperature is developed. This chip arranges its channels in 4×8 arrays, expected to readout large area SiPMs. The readout channel consists of high bandwidth and low input-impedance pre-amplifier, leading edge discriminators, low

power TDCs and logic control and data transmission blocks, providing the time based SiPM readout up tp 1 cm^2 and maximum event rate of 5 MHz per channel, with the power consumption less than 10 mW per channel and time bin 50ps(25ps).

A 64-channel mixed-signal ASIC for highly capacitive silicon detectors is also described.

Acknowledgements

And I would like to acknowledge ...
To be written...

I would like to dedicate this thesis to my loving family

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Silicon photon-multiplier (SiPM) is a novel solid-state photon sensor with single-photon detection capability. A SiPM consists of an array of Single Photon Avalanche Diodes (SPADs) working in Geiger mode and summing the output in one collection electrode. The technology and design of the single cell defines the sensor's characteristics in terms of photon sensitivity (quantum efficiency), gain and operation voltage, whereas the number of cells in the array impact the dynamic range of the device.

Compared to more conventional photon sensors like PMTs, SiPMs have additional advantages in terms of compactness and robustness, insensitivity to magnetic fields, low bias voltage, which make SiPMs up-and-coming candidates substitute these conventional sensors and also to develop new applications in different fields like nuclear and high energy physics experiments, nuclear medicine, distance measurement.

In this chapter, the fundamental physics of the SiPM is firstly introduced, along with the comparison with other families of photon sensors. Thereafter, the motivation of each dedicated project is introduced, and the outline of this thesis is depicted.

1.1 SiPM Principles and Characteristics

1.1.1 SiPM Operating Principles

The Silicon Photo-multiplier (SiPM, also multi-pixel photon counter, MPPC) is a solid-state photon sensor consisting of a parallel array of independent cells, each of them is equal and made of a PN junction operating in Geiger mode, in series with a resistor (called quenching resistor, R_q); all micro cells are connected to a common output.

Figure 1.1 is a simplified illustration of the SiPM structure. Each microcell is sensitive to incident photons and the photon detection efficiency (PDE) peaks at

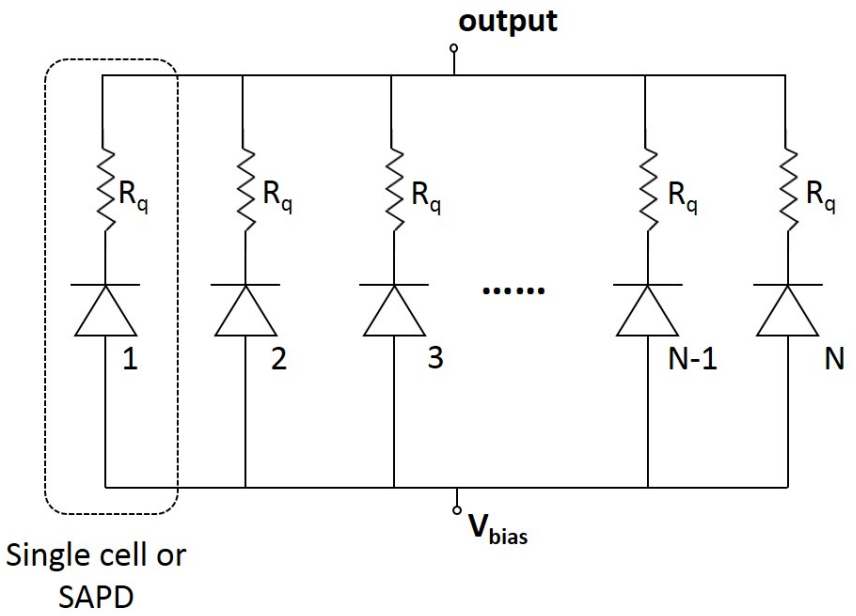


Figure 1.1: SiPM Structure.

a given wavelength, which is a strong function of technology parameters and the SPAD design. The incident photon creates an electron-hole pair and the geiger mode operation causes an avalanche pulse which is only mitigated by the series connection of a quenching resistor, which causes a dynamic drop of the biasing voltage.

In order to have a better understanding of the working principle of each cell, figure 1.2 shows the main types of diodes in parallel: PN junction, PIN (P: p-type, I: intrinsic, N: n-type) diode, Avalanche Photon Diode (APD).

A PN junction (figure 1.2a) is the simplest diode and ranks also amongst the most basic embodiments of a semiconductor detector. It is formed by two homogeneous regions of p- and n-type semiconductor material [76]. The gradient of electrons and holes in two types of material causes the diffusion process (electrons from n to p, holes from p to n), which creates the depletion region.

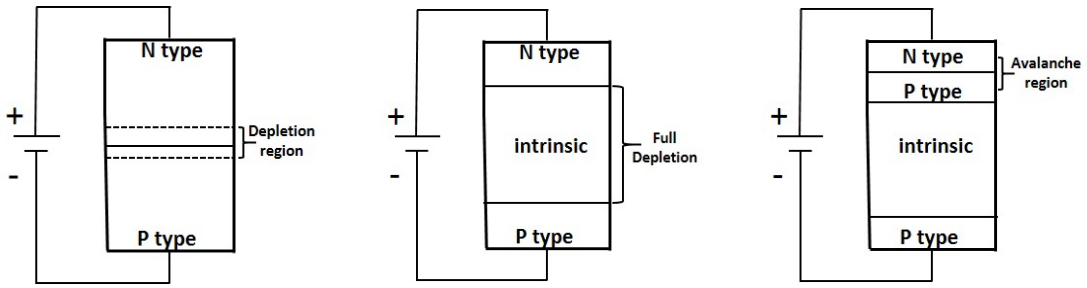
The depletion region is an attractive medium for the photon detection and it can be enlarged by applying the reversely biased voltage. The interaction with an incident photon creates electron-hole pair in medium volume, under the mechanisms of photoelectric absorption, Compton scattering or pair production [55]. The free charge drifts under the electric field of the depletion region and is collected by the electrodes, where the instantaneous current signal induced by moving charged particles can be calculated by Shockley–Ramo theorem [48].

The two drawbacks of PN juction for photon detection are limited depletion

volume and unity gain.

In a PIN diode structure, shown in figure 1.2b, a much thicker layer of intrinsic semiconductor is added between the two layers of n- and p-type material, resulting in a much wider depletion region (sensitive volume). But the PIN diode is still not able to detection the low flux light because there is no amplification process.

Compared with the PIN, an APD contains an additional p-type layer between the n-type and intrinsic layer, providing a high doping gradient at this p-n junction. The bias voltage, therefore, creates a strong electrical field in this region. Electrons or holes generated by absorption of incident photons drift to this region where they could gain enough energy to make secondary multiplication.



(a) PN junction (b) PIN diode (c) APD

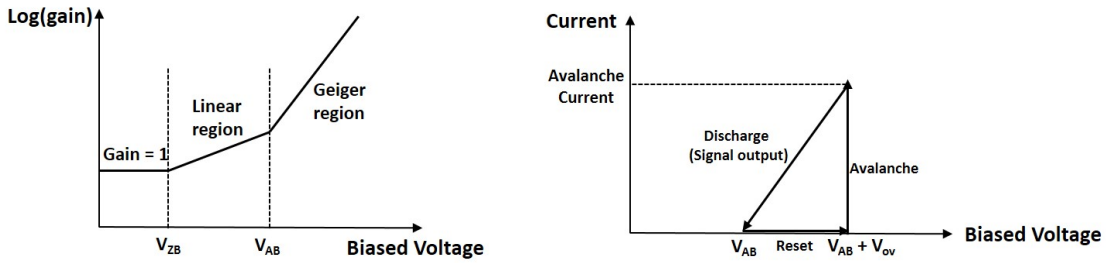
Figure 1.2: Simplified schematic structures of different types of photon diodes.

Several working regimes exist that vary depending on the applied bias voltage on APD. The relation of Bias Voltage and the multiplication process (gain) depicts in figure 1.3a [61]:

- At low electric field, the bias voltage less than "Zener Breakdown" voltage (V_{ZB}), no secondary carriers are created, similar to PIN diode.
- When the bias voltage is over than "Zener Breakdown" voltage but less than the "Avalanche Breakdown" voltage (V_{AB}), only the electrons can gain enough energy and generate secondary electron-hole pairs, thus this avalanche process stops when electrons leave the avalanche region, creating an amplified signal that will be proportional to the primary ionization signal. The gain can typically reach values in the order of 10^3.
- After the bias voltage exceeds the "Avalanche Breakdown" (Geiger) voltage, both the electrons and holes can create avalanche multiplication, resulting in the self-sustaining avalanche process. The gain can be as high as 10^5 - 10^6, providing the ability to detect single photon.

The micro cell of the SiPM is composed of Geiger mode APD in series with a passive resistor (R_q), as shown in figure 1.1, and figure 1.3b shows the three main processes of operating principle for single photon detection:

- Avalanche: the SPAD is biased with the voltage higher than V_{AB}, usually several Volts higher than V_{AB}, called Over Voltage (V_{OV}). In this condition, the SPAD is very sensitive to the incident single photon, which could generate a primary electron-hole pair and then start the avalanche process. It should be mentioned that the avalanche current value is a constant value, which is not related to the initial amount of charge, and consequently the total charge of the single SPAD is not proportional to the number of contemporary incident photons.
- Discharge: This process can be also called as "Quenching", the avalanche current flows across the quenching resistor and creates a sufficiently large drop on the bias voltage of the diode, consequently stopping the avalanche.
- Reset: during this "Recharge" phase, the external bias supply will recharge the junction capacitance through the quenching resistor. The diode is ready again to detect a new incident photon.



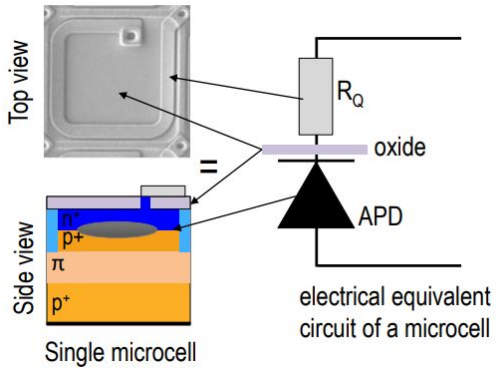
(a) Bias regions of APD

(b) Process of Avalanche in Geiger mode

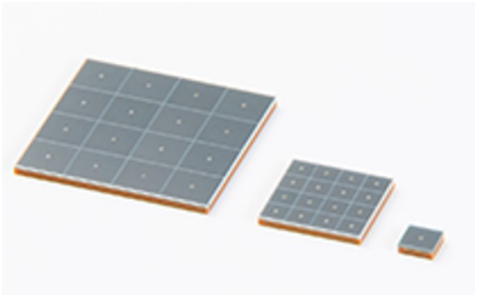
Figure 1.3: APD Bias regions and avalanche in Geiger mode.

A SiPM integrates up to hundreds to thousands of SPADs (each one in series with a quenching resistor), where all the cells are summed in a common output. More detailed structure views is shown in figure 1.4a. The analogue amplitude of the superimposed signal from multiple fired cells can provide the information on the number of detected photons: the dynamic range is defined by the number of sub-cells in the SiPM. Figure 1.4b shows an example of a SiPM by Hamamatsu company which is optimized for PET imaging, with the single pitch of 50 μm and the optional number of cells.

The operation of the SiPM in Geiger mode gives it single-photon sensitivity and fast timing response. Table 1.1 shows the comparison among different photon sensors.



(a) Structure views of SiPM cell [72]



(b) SiPM product example [47]

Figure 1.4: 2D structure of SiPM and the product example.

	PIN	APD	SiPM	PMT
Gain	1	10^2	10^6	10^6
Operation voltage	5 V	100 - 500 V	20 - 60 V	1000 V
Large area	No	No	Yes (scalable)	Yes
Noise	Low	Medium	Medium	Low
Uniformity	Excellent	Good	Excellent	Good
Time response	Fast	Fast	Very fast	Fast
Energy resolution	High	Medium	High	High
Temperature sensitivity	Low	High	Medium	Low
Magnetic resist	Yes	Yes	Yes	No
Compact	Yes	Yes	Yes	No

Table 1.1: Properties comparison of different Photon sensors.

1.1.2 SiPM Properties at Room and Cryogenic temperature

As the latest advance of solid-state photon sensors, SiPMs have been replacing conventional PMTs in many fields. Indeed, the advantage of SiPM stems both from its working principle and also due to the fact that initial drawbacks like high dark current and cross talk are now substantially reduced with rapid technology development. In this section, the basic properties of SiPMs will be introduced. In respect to the state of the art knowledge of SiPM technology and applications at commercial temperature range (-40 up to 125 °C), we extend the analysis to the operation of SiPMs at cryogenic temperatures.

Gain

As described before, the charge released in the avalanche process of the SPAD is acting on the PN junction. Thus the gain of single-photon detection can be written as:

$$Gain = \frac{C_d \cdot (V_{Bias} - V_{AB})}{e} = \frac{C_d \cdot V_{OV}}{e} \tag{1.1}$$

where C_d is the capacitance of the diode depletion region, e is the elementary charge, and V_{Bias}, V_{AB}, and V_{OV} are bias voltage on the diode, avalanche breakdown voltage, and over-voltage respectively. The gain of SiPM single cell is typically in the order of 10^6, while process and statistical local variations can impact the uniformity of the gain within the pixel matrix.

From equation (1.1), we can obtain the equation below to study the temperature dependence if we assume a constant bias voltage.

$$\frac{\partial Gain}{\partial T} = \frac{\partial C_d}{\partial T} - \frac{\partial V_{AB}}{\partial T} \tag{1.2}$$

Actually, the breakdown voltage is more sensitive to temperature than junction capacitance. In paper [70], the authors characterized three products manufactured respectively by FBK, Hamamatsu and SensL. When the junction temperature spans from - 40 to + 40 °C, the breakdown voltage variation is about 8% while the junction capacitance only about 1%. Consequently, a temperature-compensated high-voltage power supply needs to be adopted for large temperature range applications.

In paper [5], SiPM properties were measured with the temperature ranges from 40 K to 300 K, with similar trend on the breakdown voltage.

Timing Response

The good time resolution of SiPMs makes them a good candidate for fast timing systems where a very good time measurement is needed, like PET and LIDAR. The timing jitter of the SiPM arises from the process of photon-volume interaction and avalanche. The main contributions to this effect can be categorized into two parts [93]:

- Signal Fluctuations: Finite Light Pulse Width Effect, Avalanche Buildup Process, Minor Carrier Diffusion for Non-peaking Wavelength in PDE, Noise Source in Passive Quenching Elements, Thermal Noise Pile-up Effects, Detector Leakage Current. These effects can be summarized as the stochastic parameters in each process and the noise sources, which can be written as: σ_{fluc}.
- Signal Slope: Avalanche Propagation Process, Parasitic Capacitor, Pixel Uniformity, can be written as: $Slope$.

Whenever the time stamp of the photodetector module is obtained with a front-end electronics implementing a leading-edge threshold, the time resolution can be determined by: $\sigma_{fluc}/Slope$. Thereby, since the signal slope is determined also by the number of detected photon in the same time, the timing characteristics of the SiPM are usually defined in terms of single-photon time resolution. An example output waveform for a single photon event (S14160-1310PS product from Hamamatsu) is shown in figure 1.5. The typical rising time of single-photon signal is in order of nanosecond, and the falling time constants is O(20-100 ns).

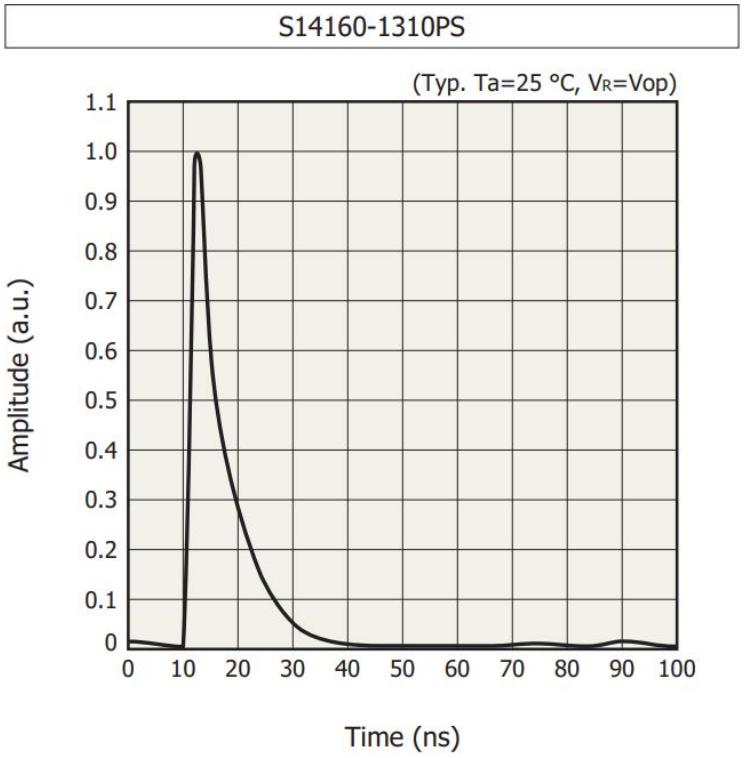


Figure 1.5: SiPM output example [47].

In the avalanche process, the diode discharges after receiving a photon, its potential has a drop of over-voltage (V_{OV}) and releases the amount of charge defined by equation (1.1). This process has a time constant calculated by equation (1.3) which impacts the intrinsic time resolution of the SPAD.

$$t_{rise} = (C_d + C_q) \cdot (R_q || R_d) \tag{1.3}$$

Where C_d and R_d are the capacitance and resistance of the photon diode, and C_q and R_q are parasitic capacitance and resistance of quenching resistor, respectively.

The timing resolution of SiPM is also related to the number of cells, since the passive SPADs have the loading effects to the fired ones.

The conventional SiPMs have the single-photon time resolution of 100 ps - 200 ps [24].

This value is related to the number of sub-cells in SiPM and size of the matrix. In one hand, a higher number of cells typically increases the total resistance of the connection path between the quenching resistor and the low-impedance collection electrode. On the other hand, the size of the matrix is proportional to the total parasitic capacitance in parallel with the fired cell.

The Hamamatsu MPPC s10632-11 (1x1 mm2) series has an intrinsic single-photon resolution of about 160ps FWHM, while the FWHM of the MPPC S10632-33 (3x3 mm2) quotes more than 600ps [47]. Besides, the single-cell (SPAD) resolution has the theoretical performance of the order of 10 ps [95] and was reported a resolution of about 40 ps in experimental result [81].

Photon Detection Efficiency

The photon detection efficiency (PDE) is the statistical probability that a signal is generated by an incoming photon. The relation can be written as:

$$PDE = QE * P_{av} * Fillfactor \tag{1.4}$$

Where:

- QE: Quantum efficiency, function of photon wavelength (λ), expresses the probability for a photon with given energy ($E = hc/\lambda$) to create an electron-hole pair inside the pixel. This efficiency is related to the interaction mechanism between photon and silicon [55].
- P_{av}: Avalanche initiation probability, the probability of the photon-created carrier to trigger the avalanche process, which is a function of the bias voltage.
- Fill factor: the ratio of photon sensitive area to the total area. The fill factor is generally range from 25% to 75% depending on sub-cell and pitch layout: insensitive area could result from the passive quenching resistance, signal routing lines and guard ring structures. A larger size of the pixel could increase the fill factor but will also decrease the pixel density.

The PDE of a SiPM at peak sensitivity wavelength usually ranges from 20% to 50%. Commercial SiPM manufacturers could provide products with different properties. PDE is a key parameter to take into account, and this parameter can be optimized for specific applications. As an example, figure 1.6 shows the product S14520 series made by Hamamatsu featuring high UV sensitivity suitable for Cherenkov light detection.

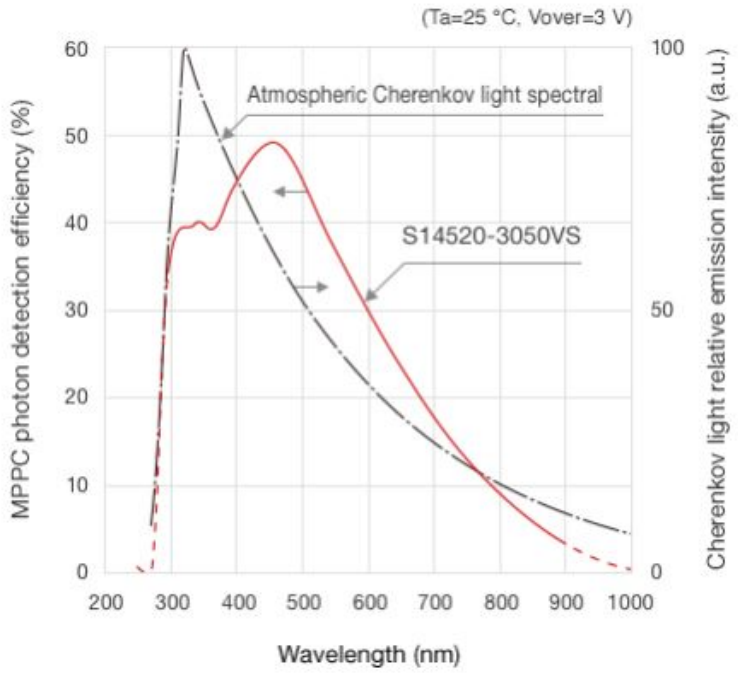


Figure 1.6: PDE optimized for Cherenkov telescope array [47].

Dark Count Rate

Dark Count Rate (DCR) evaluates the probability of avalanche process generated by thermal generation or tunnelling effect instead of a real photon event. The dark count rate is the main limitation for SiPM in low flux light application, especially for single-photon detection. Signal outputs happen even if the SiPM operates in complete dark, most of them corresponding to single-photon amplitude, thus called dark count. DCR is defined by the dark count number divided by the total area, usually with the unit of Hz/mm^2. In operation, DRC mainly depends on temperature and breakdown voltage. The temperature dependence can be expressed as [71]:

$$DCR \propto T^2 \cdot exp(-\frac{E_c - E_T}{kT}) \tag{1.5}$$

DCR is also related to crystal parameter like lattice defect, which could intermediate energy levels and thus enhance the excitation [98]. With the development of wafer and material process, the DCR provided by commercial companies has been sufficiently reduced for the last few years. To date, S13360 series from Hamamastu have a typical DRC about 55 kHz/mm^2 [47].

Figure 1.7 shows an example of DCR as a function of the temperature for

different values of bias voltage. The NUV-HD-LF SiPM produced by FBK show a DCR that can be as low as $0.01 Hz/mm^2$ [5] at cryogenic temperatures.

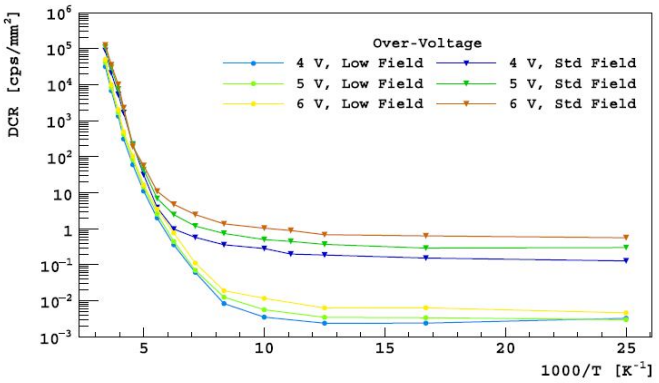


Figure 1.7: Dark count rate as a function of temperature for different over voltages [5].

Cross Talk and After pulse

Similarly to DCR, both the cross-talk and after pulse can be considered as "correlated noise". Figure 1.8 shows the output signal waveform of the single photon and associated crosstalk and afterpulse events.

Cross talk is the effect that avalanche process causes to adjacent pixels and can be categorized by optical cross-talk and electrical cross-talk.

Optical cross talk origins the fact that every 10^5 carriers can generate about three photons with the energy over the silicon bandgap (1.14 eV) during the avalanche process. These secondary photons can travel to the neighbouring cell and produce a new avalanche.

Electrical cross talk is produced with the reason that avalanche carriers have the probability of going through the boundary of the fired pixel and reaching a neighbouring SPAD, causing new avalanche processes [53].

Depending on the principle, the cross talk signal has, therefore, a correlated time to the initial signal with a delay (order of few ns) which acts as the signal superposition. Besides, cross-talk is related to overvoltage, and pixel size since the overvoltage directly impacts the gain of APD corresponding to the number of avalanche carriers, and the smaller size of the pixel can enhance the ability of photons or carriers to reach the neighbouring pixels. Cross-talk can be efficiently reduced improving the manufacturing process, like implementing the optical trenches around each SiPM pixel and block the photons travelling to close pixels [46].

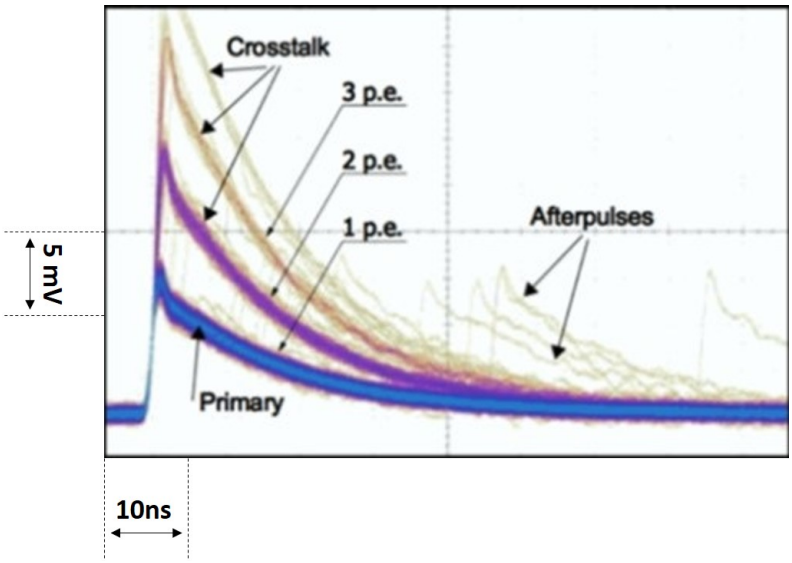


Figure 1.8: The cross talk and after pulse.

Cross talk doesn't show a strong dependence on temperature. The experimental result is shown in figure 1.9, where the NUV-HD-LF SiPM fabricated by FBK is characterized at the temperature range between 40K and 300K.

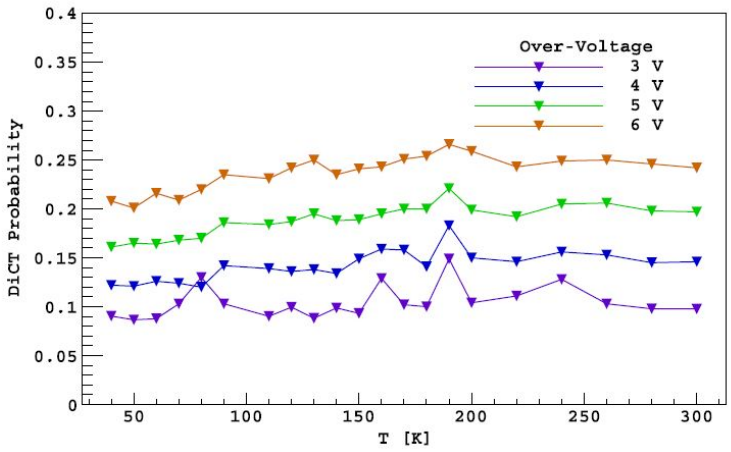


Figure 1.9: Cross talk characterization of SiPM by FBK, from 300 K to 40 K[5].

After pulse is caused by carriers trapped by the lattice during an avalanche and with some probability to be released after a particular time, usually during the recovery time. Therefore, it can generate pulses (second avalanche) after the peak of the real event, but smaller in amplitude since the electric field of the pixel is not

fully recovered.

After pulse depends on both the overvoltage and the temperature. A higher overvoltage could increase the number of avalanche carriers (gain) and thus enhance the probability of a carrier to be trapped.

Temperature dependence of After pulse is due to two effects: the trapping constant increases at low temperature and thus enhance the probability of release of carrier to cause the After-Pulse event; Quenching resistance varies with the temperature, and a longer recharge time (caused by growing quenching resistance) can suppress the After-Pulse effect [73].

Figure 1.10 shows an experimental result:

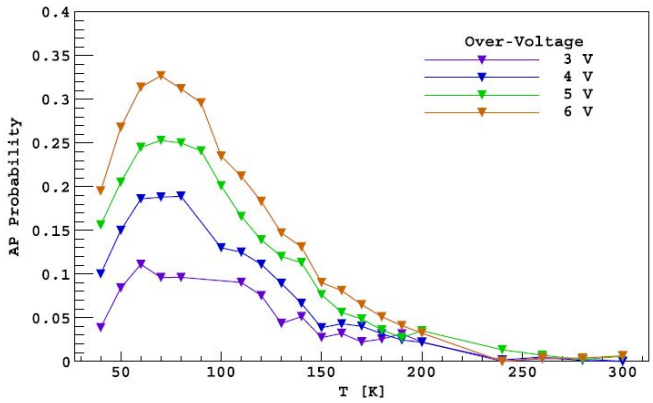


Figure 1.10: After pulse as a function of temperature and over voltage. [5]

Dynamic Range

As mentioned before, microcell (SPAD) of SiPM works like a digital counter since the pixel can not have a response to a new photon during the avalanche process. Therefore, the number of photons can only be reflected by the number of fired pixels. The number of fired pixels (N_{fired}) is related to the number of pixels (N_{pixel}), number of incident photons ($N_{incident}$) and PDE, with the following equation [15]:

$$N_{fired} = N_{pixel} \cdot (1 - exp\frac{N_{incident} \cdot PDE}{N_{pixel}}) \tag{1.6}$$

This is, however, valid only if the following condition is observed:

$$PDE \cdot N_{incident} << N_{pixel}$$

The equation can be approximated as:

$$N_{fired} = N_{incident} \cdot PDE \tag{1.7}$$

The dynamic range of SiPM is limited by the number of sub-cells, and the worse inlinearity occurs when the number of incident photon approximates the total number of sub-cells.

Figure 1.11 shows the example of S12572-015C from Hamamatsu [46].

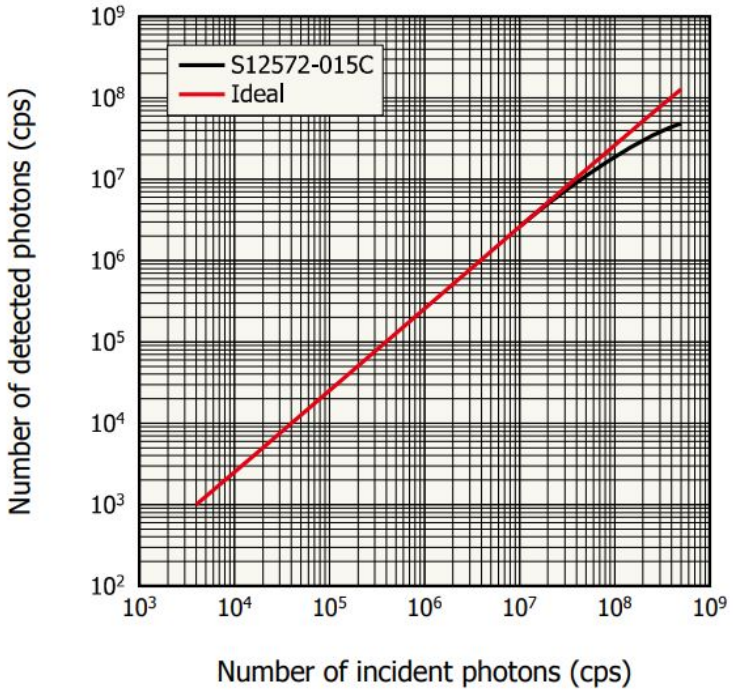


Figure 1.11: Dynamic range measurement compared with ideal response.

1.1.3 State of the art: SiPM products

Owing both to its working principle and relevant progress in the manufacturing technology, SiPMs present nowadays excellent performances for light detection and they have been applied in many different fields.

In general, different properties of SiPM can show a trade-off relation. Such as improving the overvoltage have the advantages of larger gain and better PDE and time response but could worsen the dark count rate, cross-talk and after pulse. Similarly, a large number of total pixels improves the dynamic range but increases dark count (DCR·area) and degrades the time response. Therefore, commercial companies typically provide a wide range of products for different applications.

Products like MPPC S14160 series of Hamamatsu, MICROFC Series of SensL are optimized for low dark, high timing resolution, suitable for the applications such like PET or radiation monitors; MPPC S13720 series of Hamamatsu and MICROR series of SensL are optimized for high PDE in near-infrared, which can be used to measure distance (LiDAR). MPPC S13370 series of Hamamatsu are featured with high efficiency in detecting vacuum ultraviolet (VUV), expecting for detecting scintillation light of liquid xenon (175 nm, PDE = 25%) and liquid argon (120 nm, PDE = 15%). MPPC S14520 series have good UV sensitivity suitable for Cherenkov light detection (mentioned in the last section); NUV-HD–Cryo SiPM from FBK is optimized for cryogenic temperature operations. [47][91][64]

Table 1.2 shows the parameters of some off-the-shelf SiPM products from commercial companies.

Label	Area [mm^2]	Pitch [μm]	Gain	DCR [kHz/mm^2]	Cross -talk	Peak λ and PDE	$V_{br} + V_{ov}$ [V]
[1]	3 x 3	50	$2.5 \cdot 10^6$	167	7%	50%/450 nm	38 + 2.7
[2]	3 x 3	50	$1.7 \cdot 10^6$	56	3%	40%/450 nm	53 + 3
[3]	3 x 3	50	$3.6 \cdot 10^6$	178	5%	40%/600 nm	42 + 5
[4]	3 x 3	25	$1.1 \cdot 10^6$	296	5%	22%/660 nm, 7%/905 nm	57 + 7
[5]	1 x 1	35	$1.7 \cdot 10^6$	3800	43%	9.1%/905 nm	25 + 7
[6]	3 x 3	20	$1 \cdot 10^6$	33	3%	24%/420 nm	24.5 + 2.5
[7]	6 x 6	35	$1 \cdot 10^6$	150	25%	50%/420 nm	24.5 + 6
[8]	4 x 4	40	$3 \cdot 10^6$	100	25%	43%/420 nm	26 + 6

Table 1.2: List of SiPM products.

With the manufacturers and Series:
[1] Hamamatsu MPPC S14160-3050HS
[2] Hamamatsu MPPC S13360-3050PE
[3] Hamamatsu MPPC S14420-3050MG
[4] Hamamatsu MPPC S13720
[5] SensL MICRORB-10035
[6] SensL MICROFC-30020
[7] SensL MICROFJ-60035
[8] FBK ASD-NUV4S-P

1.2 SiPM applications

1.2.1 SiPM in particle and nuclear physics

Calorimetry is the dominant field of the SiPMs application in particle physics. The pioneer is CALICE Analog Hadron Calorimeter (AHCAL) prototype [7], completed in the year 2007, which is aimed to study hadron showers and test the concept of particle flow, and also expected to study the performance and reliability of SiPMs readout in a large scale. In the AHCAL project, in total 7608 SiPMs, made by MEPhI/PULSAR, each works with a scintillator tile and provide necessarily high granularity detection modules for events reconstruction [94]. The Tokai-to-Kamioka (T2K) [4] is a baseline neutrino oscillation experiment and the first physics experiment to adopt large scale of SiPMs.

Silicon photomultipliers have been proposed for a wide range of future detectors for dark matter and neutrino experiments using noble elements. Table 1.3 lists the status and main characteristics of the current and planned developments in this field.

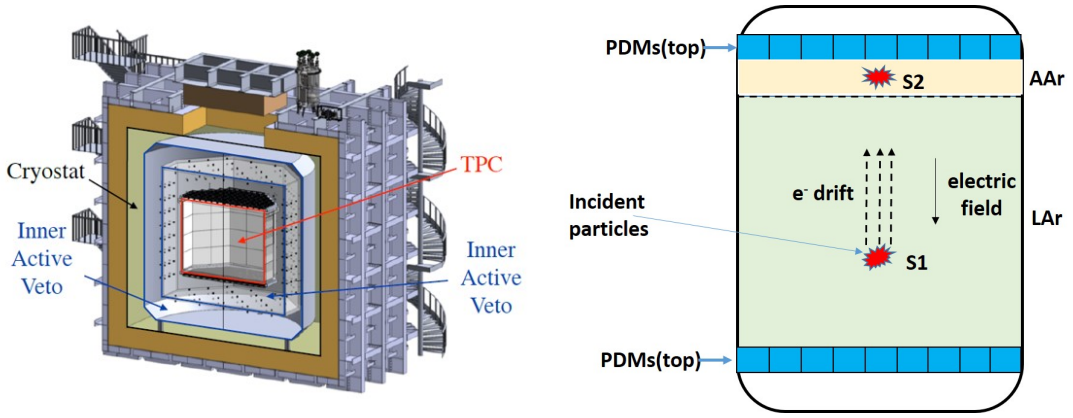
Project	Application	Status	Fiducial Volume	SiPM area
DarkSide-20k [3]	DM	Design	LAr 20 t	20 m^2
DARWIN [1]	DM	Design	LXe, 50 t	8 m^2
Proto-DUNE-SP [14]	ν oscillation	Operation	LAr, 77 t	
DUNE [37]	ν oscillation	Concept	LAr, 17,000 t	
nEXO [36]	$\beta\ \beta$ decay	Design	LXe 5 t	5 m^2
MEG-2 [69]	$\mu \rightarrow e\gamma$	Construction	LXe	0.6 m^2

Table 1.3: SiPM applied in liquid Xenon (LXe) or Argon (LAr) detectors.

Figure 1.12 shows the basic information of LAr TPC detection equipment applied in Darkside-20k project for WIMP (Weakly Interacting Massive Particles, one possibility of dark matter) search, and it shares the similar working principle with detectors summarized in table 1.3.

As shown in figure 1.12b, the incident particles make interactions with the liquid Argon atoms, creating the scintillation light (called S1 signal) with a wavelength of 128 nm, and ionizing free electrons. The S1 signal will be detected by the Photon Detection Module (abbreviation with PDM, applying SiPM as the photon sensor) mounted in both top and bottom of the TPC. The electrons will drift under the applied electric field, arrive at the air argon region and create the second scintillation light called S2 signal. The S2 signal will be detected by SiPMs on the top (anode) of the TPC.

S1 signal can be used for energy information extraction and Pulse Shape Discrimination (PSD). S2 signal is proportional to the initial ionization thus can be



(a) 3D view of LAr TPC.

(b) Working principle of LAr TPC.

Figure 1.12: LAr TPC equipment and working principle in Darkside-20k project.

used for energy extraction of the event. The distribution of S2 at top PDMs reveals the horizontal coordinate of the initial event, and the drift time of S2 (time difference between S1 and S2) can be used to exact the vertical position. Furthermore, accurate time information of S1 from both top and bottom PDMs can be also used for reconstruction of 3D position of the initial event.

1.2.2 SiPM in medical imaging

SiPM is an excellent candidate for nuclear medical techniques, mainly because of its high gain, fast response, low operating voltage, insensitivity to magnetic field and compactness. The applications of SiPM in nuclear medical, mainly PET (Positron Emission Tomography) and SPECT (Single-photon emission computed tomography), are reviewed in the following contents.

PET

PET is able to provide metabolic and functional information by detecting gamma rays emitted by radio-tracers with the radioactive atom emitting positrons. Figure 1.13 illustrates the working principle and a simplified design of a PET detector ring.

In a typical embodiment, each photodetector module of the detector ring comprises the front-end readout electronics, the photosensor (e.g. SiPM) and a scintillating crystal that converts the 511 keV photon into visible light, detectable by the SiPM. The simplified representation in figure 1.14 is called scintillator detector, which consists of scintillator material, optical grease and the photon sensor. The scintillator material converts incident particles or radiation into a large number of

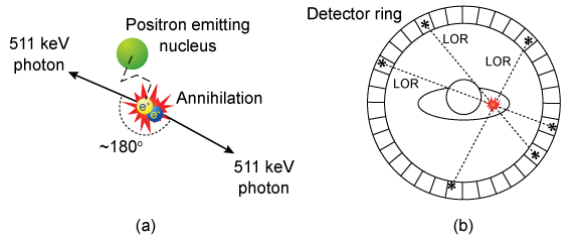


Figure 1.13: Positron Emission Tomography working principle [40].

low-energy photons, the photon sensor converts the optical signal into electrical signal, and the optical grease works to improve the coupling between the scintillator material and the photon sensor (for solid state scintillator).

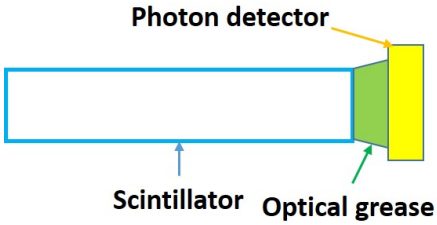


Figure 1.14: Structure of the typical scintillator detector.

The nuclear reaction shown by equation 1.8 can happen in proton-rich nuclei: one proton decays to a neutron, a positron and a neutrino, which is so-called β^+ decay.

$$p \rightarrow n + e^+ + v_e \tag{1.8}$$

The radioisotopes like ^{11}C, ^{31}N, ^{15}O, and ^{18}F can work as the radio-tracers and be injected to the patient body. The PET scanner detects and records the positron emission events and reconstructs the accumulation information of the radio-pharmaceutical in order to reveal the properties of the living tissues.

The released positron will soon reach approximately its thermal energy and annihilates with an electron, creating a pair of gamma rays with an energy of 511 keV each in the opposite direction. The PET sensor can detect the two events and build a Line of Response (LOR), which indicates the line along which the annihilation occurred. A given amount of LORs is used to reconstruct the positron emission zone. Whenever made possible by the photodetector and associated electronics, time of flight (ToF) information of each event is recorded to discard the background and noise, which improves sufficiently the reconstructed resolution. The principle of resolution improved by ToF information is shown in figure 1.15.

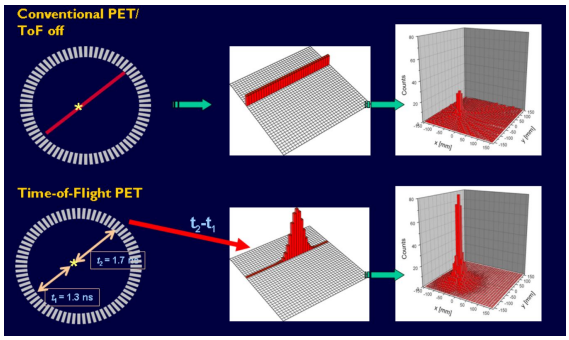


Figure 1.15: Time of Flight information to boost the reconstruction quality [92].

In a ToF measurement, Coincidence Time Resolution (CTR) in FWHM is the crucial parameter to evaluate the ToF-PET system. To date, clinical ToF-PET systems have a CTR below 600 ps FWHM [100]. More accurate detection is able to build a high-quality image with fewer radio events which could reduce the radiation dose to the patients or, equivalently, reduce the examination time.

SiPM is firstly used to replace the PMT in the hybrid system PET/MRI [104] in order to combine the functional information (PET) and the excellent anatomical information (MRI), called multi-modal diagnostic imaging, because of its insensitivity to the magnetic field. PMT is reported that its performance starts to show significant distortion with a magnetic field of only 3 mT [102], while the clinical MRI devices work at magnetic field strength in the order of several T. Besides, compared with PMT, SiPM also has advantages of compactness, relatively low cost and easy integration with electronics, which makes it an ideal replacement of PMT in next-generation of ToF-PET systems. The major players in molecular imaging, like GE Healthcare, Philips and Siemens Healthineers, have chosen SiPM to implement their more advanced ToF-PET systems (combined with CT or MRI) [50][80][38], corresponding to the CTR FWHM of 200 ps to 400 ps.

Some research PET system, like EndoTOFPET-US project, can achieve a time resolution of 200 ps [105], which builds a novel multi-modal tool integrated PET and ultrasound imaging aimed for prostate and pancreatic tumours diagnosis. The PET system includes a head extension for a commercial US endoscope and a PET plate outside the body in coincidence with the PET head, and the photon detection is implemented by scintillator based digital SiPMs and analogue SiPMs respectively.

In the laboratory, the capability of near 100 ps FWHM was demonstrated by measuring the time coincidence of two scintillators coupled SiPMs with a laser source [43], which showed the intrinsic time resolution of these detection modules were limited by scintillator instead of SiPM. Further effort is towards developing new scintillating materials and data acquisition systems in order to provide a minimal degradation of the intrinsic timing performance of the photo-detector.

In PETALO project [88], a novel concept for PET scanner based on liquid Xenon with TOF measurement is proposed. Liquid Xenon used for scintillating offers the features of high scintillation yield, fast decay time, and uniform response. The Monte Carlo investigation shows an instinct time resolution of 70ps and a time resolution of 30-50 ps can be obtained by using Cherenkov light.

The 3Dπ project is aimed to develop three dimensional positron identification with liquid Argon total-body TOF-PET (conceptional structure shown in figure 1.16), benefits directly from the technologies of Darkside-20k project like cryogenic SiPM and readout electronics. In principle, this proposed work can make it possible for direct 3D position imaging (the conventional approach combine multiple 2D projections obtained from different angles), achieve the clinical imaging in mm order resolution, and reduce the radiation dose to the patients by two orders of magnitude (10 mSv to 0.1 mSv) in each clinical exam [2].

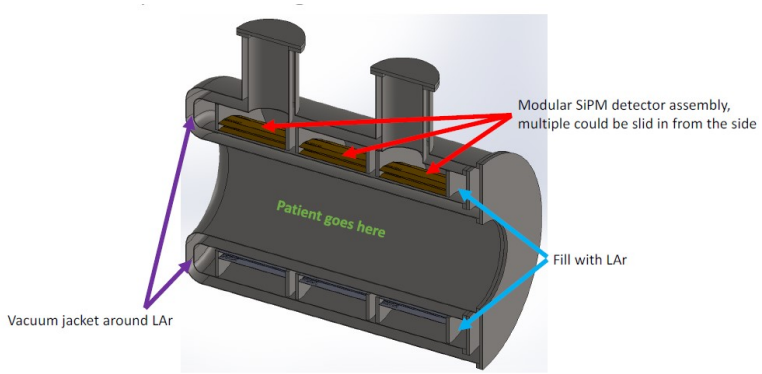


Figure 1.16: Conceptional structure of 3Dπ equipment [78].

PET systems with such performance at the system level are not yet available. Further research in new scintillating materials and data acquisition systems able to provide a minimal degradation of the intrinsic timing performance of the photo-detector plus scintillator is necessary.

SPECT

In addition to PET, the other widely applied nuclear medicine technique of functional imaging is SPECT (Single-photon emission computed tomography). The single photon is emitted by radiotracer injected to patients' body and detected by the photon sensor. Different with PET technique, radioisotopes used in SPECT do not release back-to-back photons but only single gamma photon at about 100 keV to 300 keV, therefore, the photon sensors have to work with collimators in order to know the direction of incoming photons. Figure 1.17 shows the schematic of SPECT system. Although SPECT has less resolution than PET technique depending on

its working principle, more radio-tracers can be used to label the different organs: different radio-nucleus may be linked to specific pharmacological medicines that interact with different biological processes. Thus a single technique is capable of distinguishing more pathologies simultaneously.

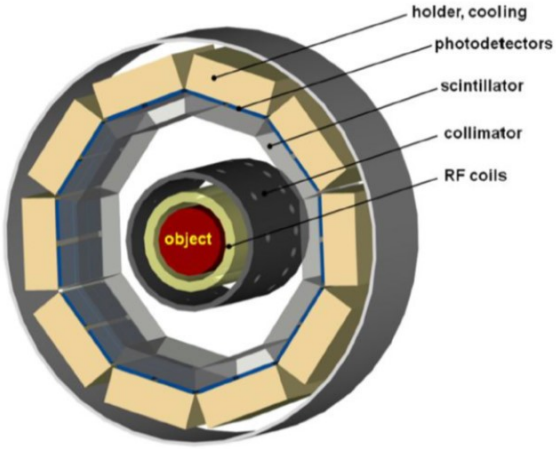


Figure 1.17: Basic schematic of SPECT system [97].

SiPM is employed in SPECT system mainly because it is insensitive to the magnetic field. Since the multi-modal diagnostic imaging is the mainstream, SiPM based gamma camera makes it possible to combine SPECT with MRI. The INSERT (INtegrated SPECT/MRI for Enhanced Stratification in RadiochemoTherapy) project, funded by the European Community developed the first clinical SPECT/MRI system [52]. The SPECT detector is implemented by continuous CsI scintillator of 8 mm thickness, coupled to an array of 144 SiPMs produced by FBK, merging in 8 mm x 8 mm pixel.

SiPM is not so commonly used in SPECT compared to PET since the collimator dominates the spatial resolution and the photon sensor is expected to provide better energy resolution where the CdTe and CZT semiconductor detectors are more feasible by direct gamma-ray energy conversion scheme compared with the scintillator coupled with photon sensor. While in the before-mentioned INSERT project, SiPMs are adopted since they are cheap and simplify the electronic system [97].

Chapter 2

Review of SiPM readout ASICs

With the wide application of SiPMs in different fields, several ASICs have been developed to readout SiPMs with corresponding requirements. These requirements are related to the sensors and systems, such like: event rate, power consumption, readout density (channels in parallel), measurements (Timing or Energy), resolution and working environment (temperature, radiation hardness). In this chapter, some representative ASICs developed for different projects and applications are reviewed.

2.1 TOFPET

The TOFPET (Time Of Flight for Positron Emission Tomography) ASIC was developed in the framework of the EndoTOFPET-US collaboration, a multi-modal imaging technique for endoscopic exams of the pancreas or the prostate. The design targets the readout of L(Y)SO scintillator coupled SiPMs of 3 mm x 3mm area in ToF PET system, and a system level CTR resolution of 200 ps FWHM [87].

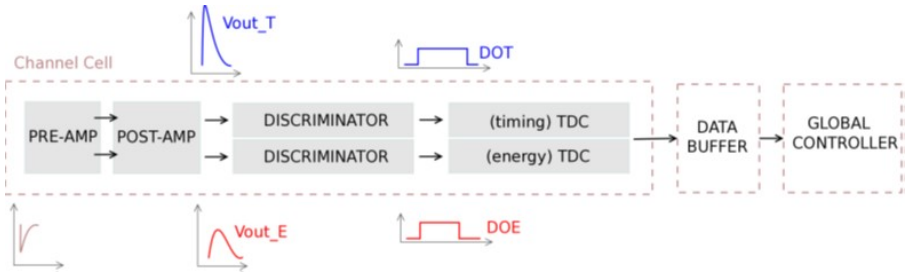


Figure 2.1: Basic architecture of TOFPET single channel[87]

The TOFPET ASIC had its first silicon results in 2013, integrating 64 channels and implemented in 8-metal IBM CMOS 130nm technology. The basic architecture of one channel is shown in figure 2.1. Each channel consists of two independent

input stages for two polarities readout (electrons or holes), then the signal is divided into two branches, the faster branch is optimized for better timing resolution, and the slower one is used to measure the energy information with Time over Threshold (ToT) method. The front-end part is shown in figure 2.2.

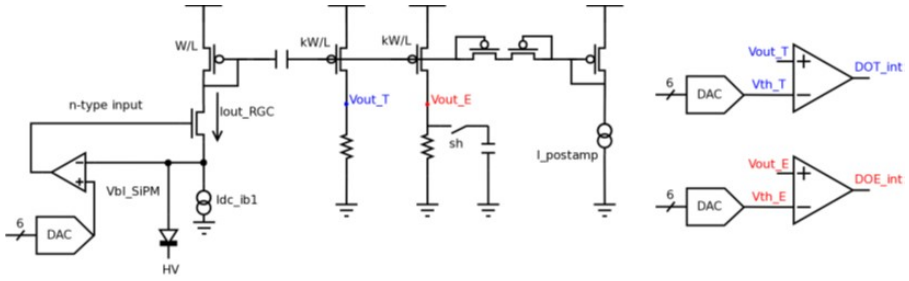


Figure 2.2: Basic architecture of TOFPET front-end amplifier (shown for n-input circuit)[87]

The main feature of front-end amplifier is the topology: Regulated Common Gate (RCG) or Regulated Cascode (RGC) structure, which is a common gate amplifier with differential feedback between the gate and source and works as a very low impedance current buffer, providing a fast response with low power consumption. Two independent types of RCGs implemented with complementary transistors to have more flexibility. Then a fast branch and slow branch optimized for timing and energy measurement respectively, the discriminator of fast branch provides the arrival time of each event, and the discriminator in slow branch configures a higher threshold to validate the event with energy larger than single photon, which can be used for dark count rejection.

Figure 2.3 describes the scheme for dark count rejection. With the lower threshold in timing branch, the trigger will have more accurate time arrival information, and the higher threshold (higher than single-photon amplitude) in energy branch provides the validation for a given event.

In TOFPET ASIC the Time to Digital Converter (TDC) is based on analogue interpolation, a technique that is very suitable for very low-power operation and low-to-moderate event rate capability.

The schematic of the working principle is shown in figure 2.4. The trigger signal from the fast branch will start a discharge with a constant current (I_{TAC}) on a capacitor (C_{TAC}) and stop when the next clock cycle arrives. Then the voltage value will be copied to a new but four times larger capacitor (C_{TDC}), with 32 times smaller current (I_{TDC}) recharging to the reference voltage. Consequently, the binning of the timing measurement becomes a function of the clock period divided by the interpolation factor, which is defined by the ratio of the charge/discharge

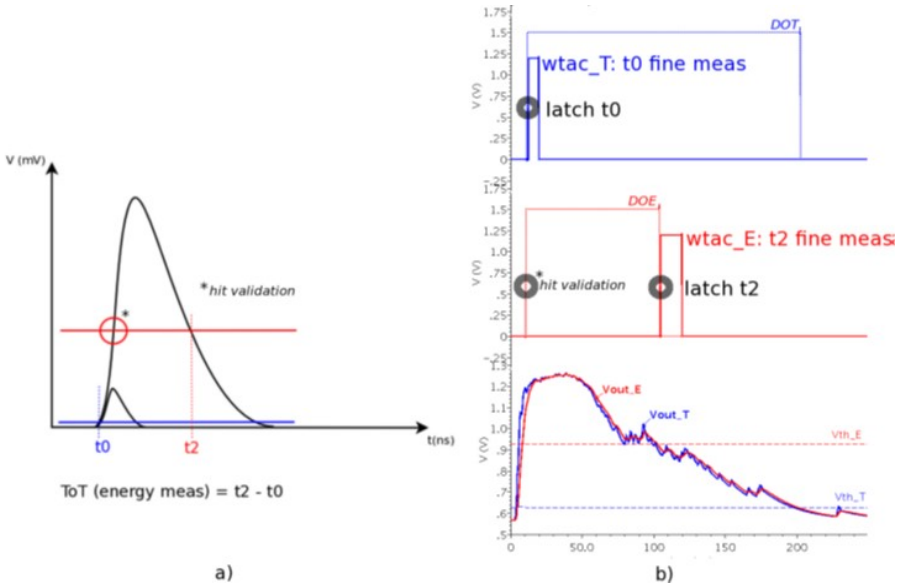


Figure 2.3: Dual thresholds for dark count rejection[87]

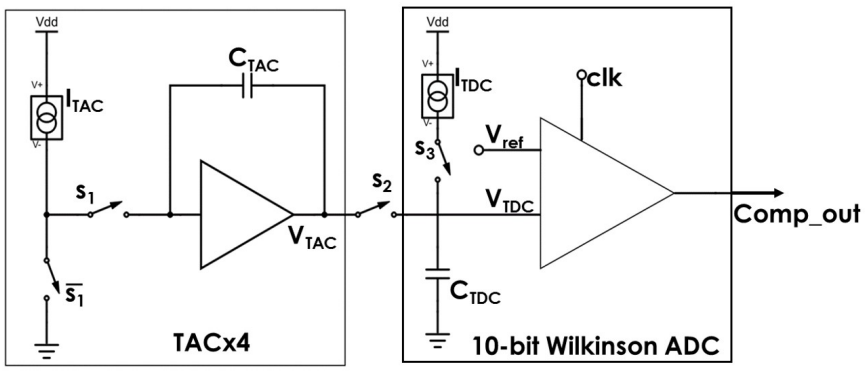


Figure 2.4: Working principle of analogue interpolation TDC [20]

time constants:

$$T_{bin} = \frac{T_{cycle}}{4 \times 32} \tag{2.1}$$

where T_{cycle} is the operating clock cycle. For the nominal value of 160 MHz in TOFPET, the time bin of 50 ps can be achieved with about 1 mW of power consumption and the maximum event rate of 100 kHz per channel is estimated in TOFPET ASIC. The same back-end part is also shared in the work [83] [20].

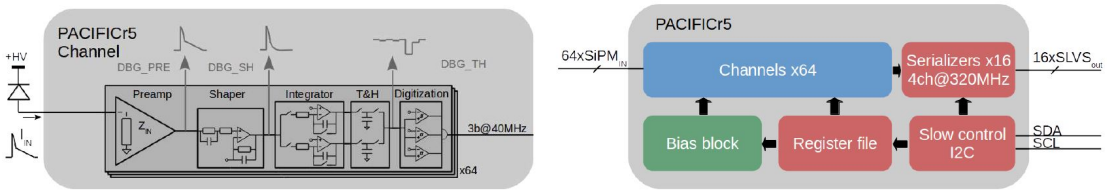
A second version of this ASIC TOFPET2 [35] was submitted in 2015, mainly with the optimizations: adding a new linear charge measurement block in each

channel implemented by a charge integrator and ADC; improved event rate capability by increasing the clock frequency.

In summary, the TOFPET ASICs family is one of the most representative readout ASICs for SiPMs in time of flight application. The main features are low power high bandwidth pre-amplifier, dual thresholds for dark count rejection and low power interpolation TDC. It should be emphasized that the analogue interpolation TDCs are widely adopted in most of ASICs for SiPMs readout since their good time resolution potential (tens of ps compared with the intrinsic time resolution in the order of 100 ps in scintillator coupled SiPM system), good linearity, simple structure and low power consumption.

2.2 PACIFIC

PACIFIC (Low Power ASIC for the SCIntillating FIbre TraCker readout) was developed to readout SiPMs in SciFi Tracker of the LHCb detector, where the SiPMs are used to readout scintillator fibre [25][23].



(a) Analog channel signal processing

(b) Control and data output

Figure 2.5: PACIFIC analog chain and control blocks [25]

Figure 2.5 shows the latest version architecture of PACIFIC. This chip is submitted in 2017, 64-channel and fabricated in TSMC 130 nm process [25].

The schematic of PACIFIC input stage is shown in figure 2.6, which works as a current conveyor and reads out the anode signal.

The current conveyor provides constant input impedance and input voltage implemented by two feedback loops, aimed for low impedance and high bandwidth readout from SiPM (60 $\mu m \times 60$ μm each micro-cell, 104 cells per channel). The following Trans-impedance amplifier (TIA) is used to transfer the current signal to the voltage. Then a pole-zero cancellation circuit and fast shaper stage manage to remove the slow component from recharge time of SiPM as well as the shorter time component, associated with parasitic capacitance and the amplifier input impedance. They achieve to integrate the pulse signal within 10 ns which is required to overcome the fluctuations of the scintillation signal shape due to the low photostatistics [27]. Two time-interleaved gated integrators are employed to minimize the dead time and followed by a dual passive track and hold structure.

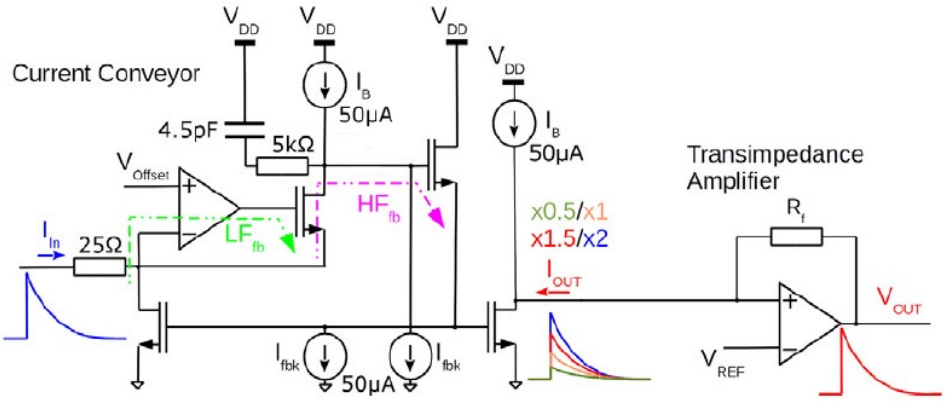


Figure 2.6: PACIFIC input stage [25]

A 2-bit non-linear flash ADC operating at 40 MHz is used to digitize the tracked and held voltage. Finally, every four adjacent channels are summed by a serializer which transmits the data off-chip with differential SLVS.

The channel architecture allows the trigger-less readout from scintillator fibre coupled SiPM, managing to complete all the processes of readout, shaping, integration, track and hold and digitization (2-bit) within one clock cycle (25 ns). This ASIC has a power consumption below 10 mW per channel and is optimized for radiation tolerance because of its working environment.

2.3 ANGUS

ANGUS, a 36-channel ASIC in CMOS 0.35 μm technology, was developed in the framework of the INSERT project. This collaboration proposed the development of a novel Single Photon Emission Computed Tomography (SPECT) system compatible with a magnetic resonance (MR) apparatus, where the SiPMs are able to show their high immunity to strong magnetic field environment. This ASIC targets the achievement of high spectroscopy performances at the low gamma energies (in the 100 to 300 keV range) used in SPECT [97].

Figure 2.7 shows the channel architecture of ANGUS, and each channel is expected to readout an 8 x 8 mm^2 SiPM. The input stage is a current conveyor to have a low input impedance. Then the current signal is contained with a shaper, to transfer the current to voltage signal and improve the signal to noise (SNR) ratio by the RC filter with a time constant programmable from 200 ns up to 10 μs. A Baseline Holder (BLH) structure is used to hold the baseline of output in the expected value. Finally, the signal is divided into two branches: one is followed with a discriminator for timing measurement; the other is a peak stretcher to capture

the peak of shaper output, which is proportional to the event energy. Both the outputs are sent off-chip for the digitization process.

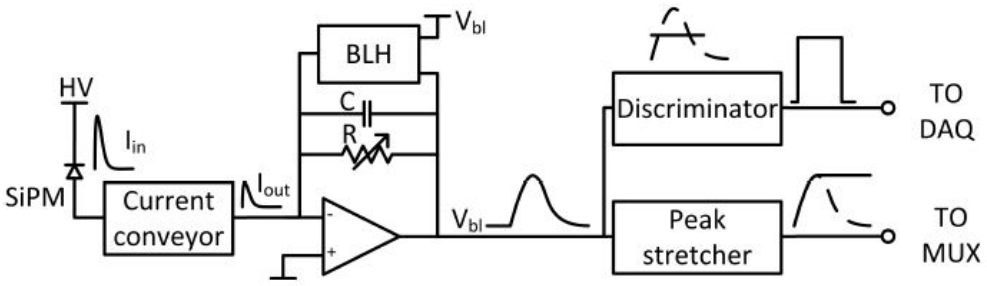


Figure 2.7: Architecture of the ANGUS ASIC channel[97]

The input stage is a low input impedance current conveyor and its working principle is described in figure 2.8. A negative signal from SiPM causes the rise of source voltage of transistor M1, then this variation is mirrored by M3 and M4. With the relation of $(W/L)_3/(W/L)_4 = (W/L)_1/(W/L)_2$, M2 is able to make a perfect compensation to the gate of M1, resulting a null input impedance in the ideal condition. In the real case, limited by the finite drain resistance of transistors and the asymmetry, the input impedance is not zero and an approximate relation is given by equation 2.2, and in ANGUS, its value is about dozens of Ω.

$$R_{in} \approx \frac{1}{g_{m1}}[\frac{g_{d1}+g_{d3}}{g_{m3}}\frac{g_{d2}+g_{d4}}{g_{m2}}] \tag{2.2}$$

2.4 VATA64-HDR16

VATA64-HDR16, fabricated by 0.35 um CMOS process and developed by GM-IDEAS (Norway), is a 64 channel front-end ASIC for SiPM readout [45]. The dedicated SiPMs are used in the detection of Cherenkov light, a field that may require the capability to deal with very low light levels. This ASIC provides accurate time arrival information in addition to the energy measurement, in order to take advantage of the synchronous emission of photons from Cherenkov cone and distinguished with SiPM dark count events.

The input stage is a charge sensitive amplifier, which integrated the input charge and sent out a voltage waveform proportional to the event energy. Then the signal is divided into two branches, one is fed to a fast shaper followed by a discriminator. An on-chip Time to analogue converter (TAC) and a Sample-and-Hold (SH) circuit are implemented; The other branch has a slow shaper followed by a peak-hold circuit. These outputs of full chip are sent outside via two multiplexers and the digitization is provided off-chip.

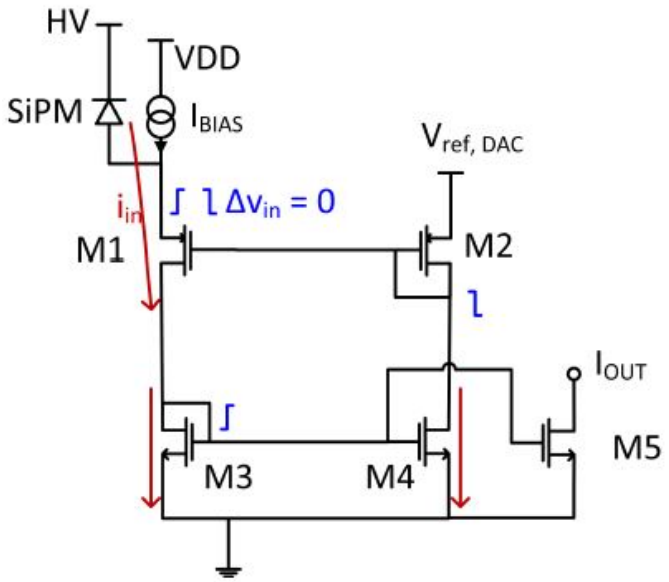


Figure 2.8: Working principle of input stage in ANGUS ASIC [97]

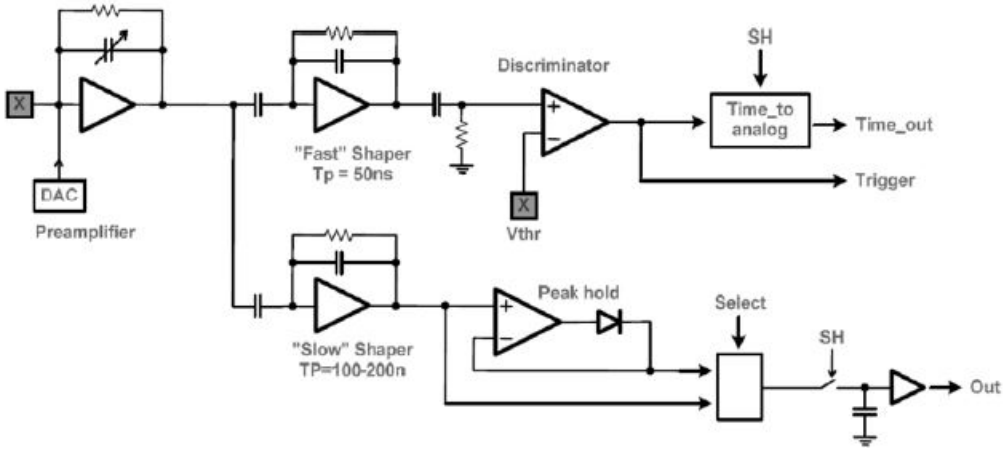


Figure 2.9: Architecture of the VATA64-HDR16 ASIC channel[10]

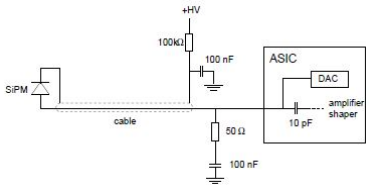
2.5 SPIROC

SPIROC (Silicon Photomultiplier Integrated Read Out Chip) is a 36-channel ASIC implemented by 0.35 um SiGe technology and originally developed to read out SiPMs for a prototype hadronic calorimeter at the proposed International Linear Collider of Desy [13] [58].

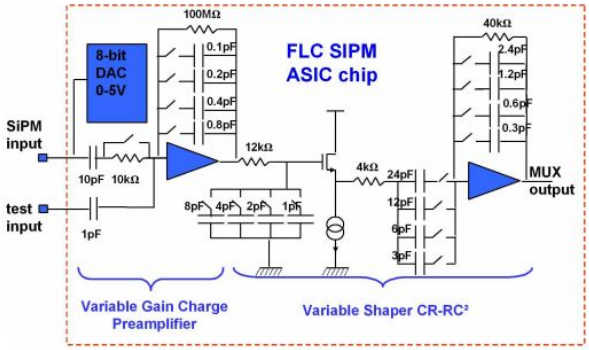
As shown in figure 2.10a, a 50 Ω load resistor is used to connect the SiPM and

converts the SiPM current to voltage, providing an impedance matching with the transmission line and limiting the voltage variation at the input node. One 100 nF capacitor is in series to the 50 Ω resistor, which works as short circuit for very short signal delivered by SiPM.

In the ASIC, a $CR-(RC)^2$ filter is adopted to filter the noise with programmable shaping times from 40 ns to 180 ns. Another fast shaper (not shown in the figure) is used to generate the trigger signal. Then, a peak stretcher works to sample and hold the peak amplitude of the shaper output.



(a) Connection scheme of SiPM to ASIC chip



(b) Overview of the channel architecture in SPIROC

Figure 2.10: Readout scheme of SPIROC ASIC

2.6 MD-SiPMs

According to the working principle of SiPM, a given number of SPADs are summed to a common output and then readout by electronics. These SPADs can also be read out individually with electronics where SPADs are integrated into the same substrate with the electronics, called digital SiPM or SPAD array. In principle, this embodiment has a high potential to detect the photons since it is able to collect the maximum information and the intrinsic time resolution of each SPAD. This solution requires the implementation of the photosensor and CMOS readout in the same substrate, which poses severe constraints to the optimisation of the sensor, in particular concerning correlated noise.

In the previously mentioned EndoTOFPET-US project, a multi-digital SiPM (MD-SiPM) system is developed for TOF-PET detection inside the endoscopic ultrasound probe, where an extremely miniaturized PET detector is implemented [16].

As depicted in figure 2.11, the detection module consists of 9 x 18 MD-SiPM arrays and each MD-SiPM has a 16 x 26 pixel array in total area of 800 x 780 μm^2,

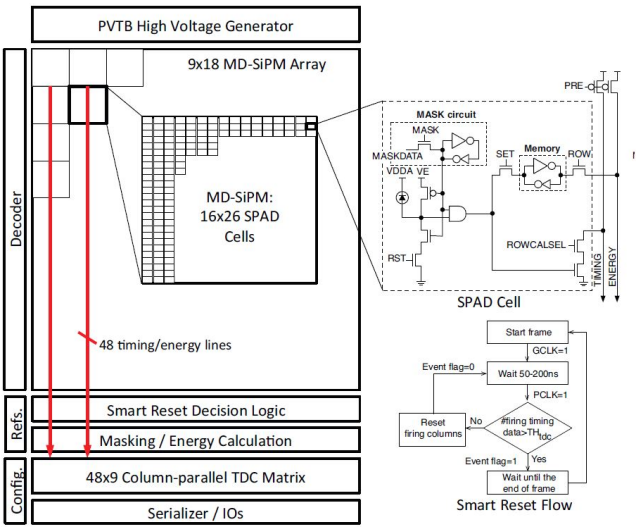


Figure 2.11: Diagram structure of DS-SiPM system in EndoTOFPET-US project [16]

50 x 30 μm^2 of single size. Each pixel consists of a SPAD and a 1-bit counter for pixel firing registration.

Each section of 416 pixels in each MD-SiPM share 48 TDCs, and the timestamp is triggered by the 1-bit counter and digitized by TDC. This endoscopic sensor can detect 67,392 photons and the first 432 time-of-arrival events in every detection frame of 6.4 μs [16].

A detailed Monte Carlo simulation is reported in [44], which aims to study the potential time resolution between analogue SiPM and MD-SiPM scheme. One of the results is shown in figure 2.12, where both simulate the result of readout from a crystal with dimensions of $2 \times 2 \times 20\ mm^3$. In the simulations, dark count and cross talk of the SiPM, as well as the electronic noise, are neglected.

The conclusion can be summarized as the analogue SiPM, and MD-SiPM readout from scintillator can have comparable time resolution when the latter scheme makes use of maximum likelihood estimation, and the digital scheme has more robustness. In contrast, the analogue scheme requires more critical threshold values.

2.7 Conclusion

Table 2.1 summarized the specifications of reviewed SiPM readout ASICs. Compared to digital SiPM, the analogue SiPM has a better compromise on low cost and excellent resolution. In general, multi-channel ASICs used for SiPMs readout need to cope with the high capacitance of this kind of sensor (typically in the order

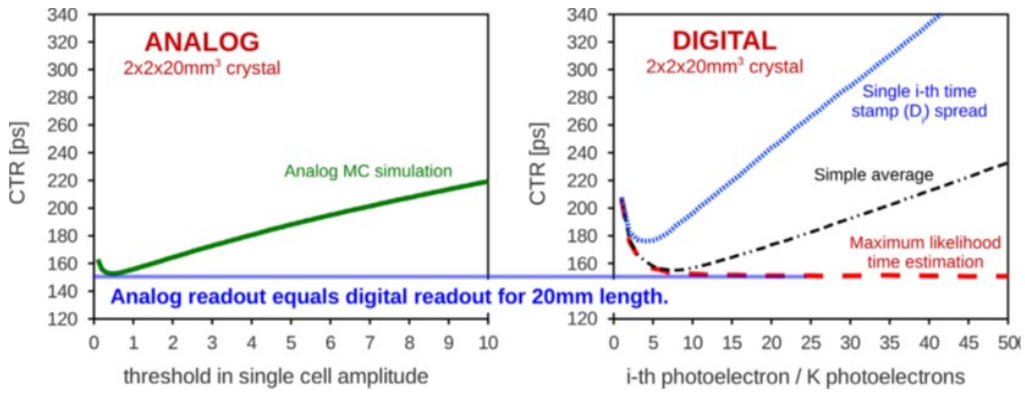


Figure 2.12: comparison between Analog and MD-SiPM simulations [44]

of 50 pF/mm^2), while providing a time-based readout with sub-nanosecond timing resolution. Due to the common stringent system requirement in terms of high channel density and limited budget for heat dissipation, most of the dedicated ASICs need an input stage with high bandwidth but limited power consumption. Analogue interpolation TDC is preferred in on-chip digitization since its lower power and suitable time resolution in time of flight application of SiPMs. Almost all the SiPM ASICs are developed with the application in room temperature or cooled down to about 0 $°C$.

The growing interest on the use of SiPMs on detectors operating at cryogenic temperatures would call for the design of dedicated ASICs. At the time of this work, there are no solutions proposed in terms of mixed-signal ASICs for SiPM readout that are able to operate in a temperature range down to 87K (LAr boiling temperature). Indeed, such a development would allow for the construction of cryogenic photon-detection modules capable of operating inside the detector cryostat, providing amplification, signal conditioning and digitization of each event. This thesis introduces the challenges and circuit level design of a multi-channel mixed-signal CMOS readout integrated circuit for SiPM readout at cryogenic temperature.

Entries	TOFPET2	PACIFICr5	ANGUS	VATA64	SPIROC2c
number of ch	64	64	36	64	36
Process Node	CMOS 110 nm	CMOS 130 nm	CMOS 350 nm	CMOS 350 nm	SiGe 350 nm
Readout type	current	current	current	charge	voltage
Application	ToF-PET	SciFi Tracker	SPECT	astroparticle	calorimeter
Measurement	T&E	T&E	T&E	T&E	T&E
Energy extract	ToT/Peak	Peak	Peak	Peak	Peak
Digitization	on-chip	on-chip	off-chip	off-chip	off-chip
Power supply	1.2 V	1.2 V	3.3 V	3.3 V	3.3 V
Power Consumption	10 mW /channel	10 mW /channel	8 mW /channel	no /channel	25 uW /channel
Clock	200 MHz	40 MHz	none	none	5 MHz
Event rate	600 kHz/ch				
Dynamic range	2500 p.e			12 pC	2000 p.e

Table 2.1: Summary of reviewed ASICs for SiPM readout

Chapter 3

Test chips studies for SiPM readout at cryogenic temperature

This chapter introduces the study of the front end circuit for SiPM readout, which are expected to read out large area of SiPMs and work at cryogenic temperatures.

Firstly, the SiPM electrical model and this is fundamental to perform the accurate simulation; Then the effects on CMOS devices of working at cryogenic temperatures are reviewed, in order to understand the main issues of designing the cold CMOS circuits; After these, the test structures mainly consists of the front-end circuit are designed and tested in liquid nitrogen environment, expecting to study the circuits parameters in cryogenics temperature. The study of test chips creates the basis to develop a mixed-signal full chip in next step.

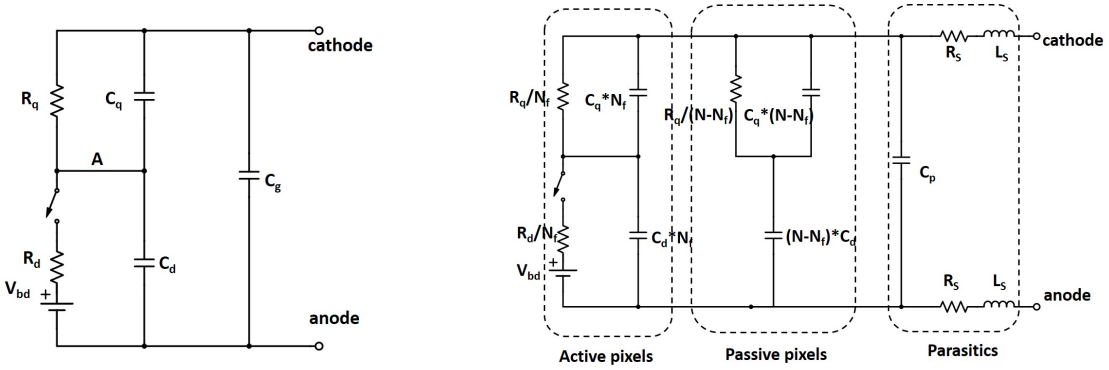
3.1 SiPM electrical model

The input stage is the first key part for SiPM readout ASIC since it is directly connected with the sensor, thereby defining the bandwidth, achievable event rate, power consumption, timing performance and signal-to-noise ratio. In order to start the design of an optimised input stage for SiPM readout, the electrical model of the sensor should be effectively characterized.

SiPM consists of an array of Geiger mode APDs in series with quenching resistor and common output (as shown in figrue 1.1). Replacing with electrical resistors and capacitors, the electrical model of fired SPAD is shown in 3.1a.

where:

- R_q is the quenching resistance, the value is from 100 $k\Omega$ to several $M\Omega$.
- C_q is the parasitic capacitance of quenching resistor, in general, its value is only a few fF



(a) SPAD electrical model[6]

(b) SiPM model used for simulation [6]

Figure 3.1: SiPM electrical model for single pixel fired

- R_d is the resistance of diode during the avalanche process, generally in the order of 1 $k\Omega$
- C_d is diode capacitance, which is mainly related to the micro-cell size, range from 10 fF to 100 fF
- C_g is the contribution of parasitic capacitance between the substrate of device and the contact of quenching resistor, can be called grid capacitance with the value about several fF of each cell.
- N is the total number of micro-cells and N_f is the number of cells which detects photons in the same time.

In this electrical model, the photon event can be modelled by closing the switch. At this point, the current through R_d starts to discharge the node A shown in figure 3.1a with capacitors C_d and C_q and resistor R_q. Assuming the anode potential as zero, the potential of A (V_A) equals to the difference between cathode and anode reverse bias (V_{bias}) before the avalanche process happens. During the avalanche process, the switch being closed, the value of V_A will drop to:

$$V_A = V_{bias} - \frac{V_{bias} - V_{bd}}{R_d + R_q} \cdot R_q \cdot \approx V_{bd} \tag{3.1}$$

where V_{bd} is the break down voltage. The corresponding time constant is:

$$t_{rise} = (C_d + C_q) \cdot (R_q || R_d) \approx (C_d + C_q) \cdot R_q \tag{3.2}$$

Thus this value also sets the intrinsic limit of the signal rise time, in the order of tens of pico-second, while in real-world applications this rise time is usually slowed down by the limited bandwidth of readout input stage.

For the electrical model, the avalanche occurrence switch will be open when the current drops below a threshold current (I_{th}) and the avalanche can not be self-sustained. This I_{th} is a little bit larger than the latching current, $I_{latching}$, with the equation [28]:

$$I_{latching} = \frac{V_{bias} - V_{bd}}{R_d + R_q} = \frac{V_{ov}}{R_d + R_q} \tag{3.3}$$

After the switch is opened, C_d will be recharged by R_q and C_q back to V_{bias}, with the time constant:

$$t_{fall} = (C_d + C_q) \cdot R_q \tag{3.4}$$

The initially avalanche current can be as high as several milliamperes and is defined by over voltage (V_{ov}, the difference of V_{bias} and V_{bd}) divided by R_d [28].

Consequently, the electrical dynamic behaviour of the SPAD can be expressed in figure 3.2. Using the voltage controlled switch to define a start time (t_0) and time duration (T), to model a detected photon. The avalanche current reaches the peak value when the switch is closed, then reduces exponentially and stops when approximates the threshold current. The area formed by I_{R_d} and time axis is equal to the charge released in avalanche process by the diode, thus the value of T can be calculate by I_{peak}, I_{th} and Q. The current waveform through R_d shown in figure 3.2 can be approximated as a triangle shape, considering the relation $I_{latching} << I_{peak}$. Consequently, the avalanche current can also be modelled as a current pulse (replace the voltage control switch, V_{bd}, and R_d), and the pulse shape is referred to waveform of I_{R_d} shown in figure 3.2.

We can assume the $R_q = 300k\Omega$, $R_d = 1k\Omega$, $C_d = 50fF$, $C_q = 5fF$, $V_{ov} = 5V$, we have:

- $t_{rise} = R_d \cdot (C_d + C_q) = 55ps$
- $t_{fall} = R_q \cdot (C_d + C_q) = 16.5ns$
- $I_{peak} = \frac{V_{ov}}{R_d} = 5mA$
- $I_{latching} = \frac{Vov}{R_q + R_d} = 16.7uA$
- $Q = V_{ov} \cdot C_d = 250fC$
- $T \approx \frac{2Q}{I_{peak}} = 100ps$

Arranging the SPADs in parallel, the SiPM model is shown by figure 3.1b, Where N_f is the number of simultaneously fired cells and N is the total number of cells of one SiPM, C_p is considered to count the lumped contribution of C_g. The series resistance and inductance in interface are modelled as R_S and L_S.

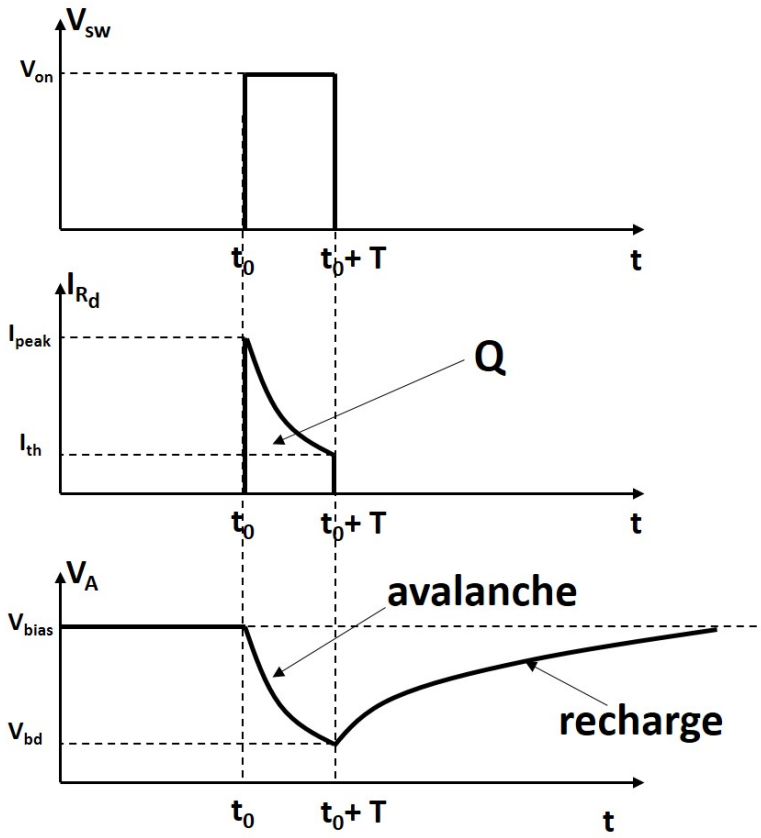


Figure 3.2: Transient current, voltage waveforms of SPAD model, realized with voltage control switch, simulates a detected event

In the SiPM, the equivalent loading effect of passive cells is described by:

$$Z_{eq} = (\frac{R_q}{N - N_f} || \frac{1}{j\omega C_q \cdot (N - N_f)}) + \frac{1}{j\omega (N - N_f) \cdot C_d} \tag{3.5}$$

If we still use the the value of $R_q = 300k\Omega$ and $C_d = 50fF$, for a 3 x 3 mm^2 SiPM with cell size of 30 x 30 μm^2, thus the total number of cells equals to 10,000. Assuming $N_f << N$, the value of resistive loading effect equals to:

$$\frac{R_q}{N} = \frac{300k\Omega}{10000} = 30\Omega \tag{3.6}$$

Which is a quite small value, and we have $C_g << N \cdot C_d$, then the loading effect can be approximated as:

$$Z_{eq} \approx C_{eq} = \frac{1}{j\omega N \cdot C_d} = \frac{1}{j\omega \cdot 5nF} \tag{3.7}$$

This is a very large capacitance for the readout circuit, especially when a very accurate timing measurement is required. In figure 3.3, in order to study the capacitive load's effect on signal readout, the small signal equivalent circuit is described. The anode signal of SiPM is read out with the circuit which has the input impedance of R_{in}, and the input voltage is V_{in}. The series inductance L_s is ignored to simplify the calculation, and the diode current signal is expressed by I_d.

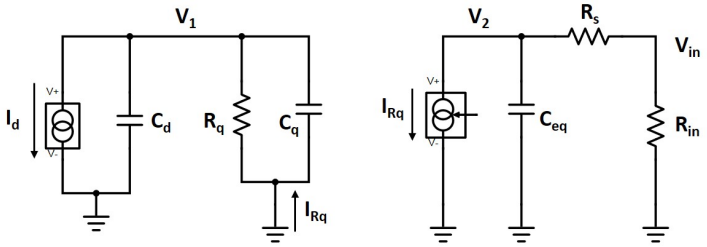


Figure 3.3: Small signal equivalent circuit of SiPM readout

According to the Kirchhoff's Current Law (KCL), there are the two equations when analyzing the two parts in figure 3.3:

$$I_d + \frac{V_1}{j\omega C_d} = I_{R_q} \tag{3.8}$$

$$V_1 = -\frac{I_d}{j\omega(C_d + C_q) + \frac{1}{R_q}} \tag{3.9}$$

$$V_2 = -\frac{I_{R_q}}{j\omega C_{eq} + \frac{1}{R_s + R_{in}}} \tag{3.10}$$

$$I_{R_q} + \frac{V_2}{j\omega C_{eq}} + \frac{V_2}{R_s + R_{in}} = 0 \tag{3.11}$$

$$\frac{V_2}{R_s + R_{in}} = \frac{V_{in}}{R_{in}} \tag{3.12}$$

Solving the above equation and transferring the result to time domain, the expression of V_{in} is:

$$V_{in} = -\frac{Q \cdot R_{in}}{t_{fall} - t_{in}}\left(\frac{t_q - t_{in}}{t_{in}} e^{-\frac{t}{t_{in}}} + \frac{t_{fall} - t_q}{t_{fall}} e^{-\frac{t}{t_{fall}}}\right) \tag{3.13}$$

In equation 3.13, diode current signal is a delta like pulse to simplify the calculation and $I_d = Q\delta(t)$. t_{fall} is mentioned above and equals to $R_q \cdot (C_d + C_q)$, $t_q = R_q \cdot C_q$, $t_{in} = (R_s + R_{in}) \cdot C_{eq}$

An intuitive simulation shown by figure 3.4, the readout circuit implemented by a voltage amplifier, can give some basic information about the effects of the readout circuit.

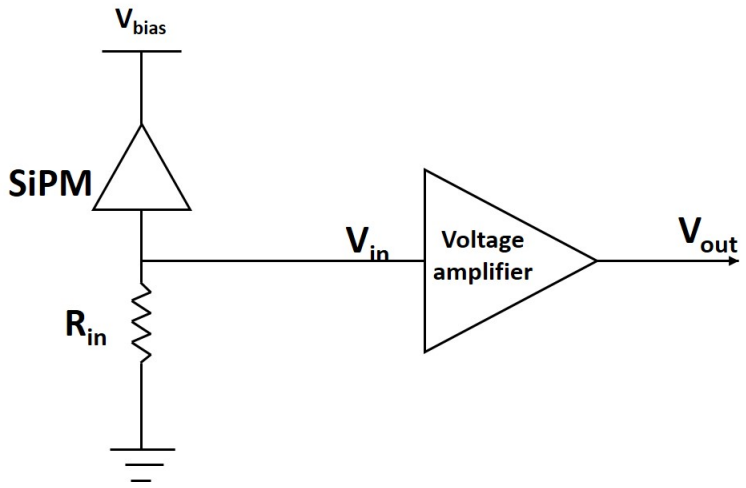


Figure 3.4: Simplified simulation schematic of SiPM and readout circuit

Figure 3.5 shows the effect of input resistance of the readout circuit (lower input resistance brings a shorter readout time since the input current waveform has a larger amplitude). Figure 3.6 shows that the high enough bandwidth is critical for high resolution timing measurement.

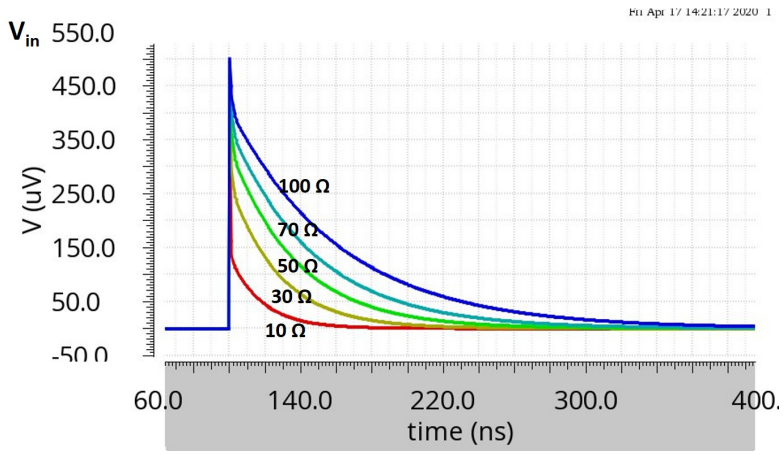


Figure 3.5: Input voltage and input resistance of readout circuit

In this work, the specific parameters are chosen from the Near Ultra Violet High Density Low Field (NUV-HD-LF) SiPM manufactured by FBK with a single-pixel

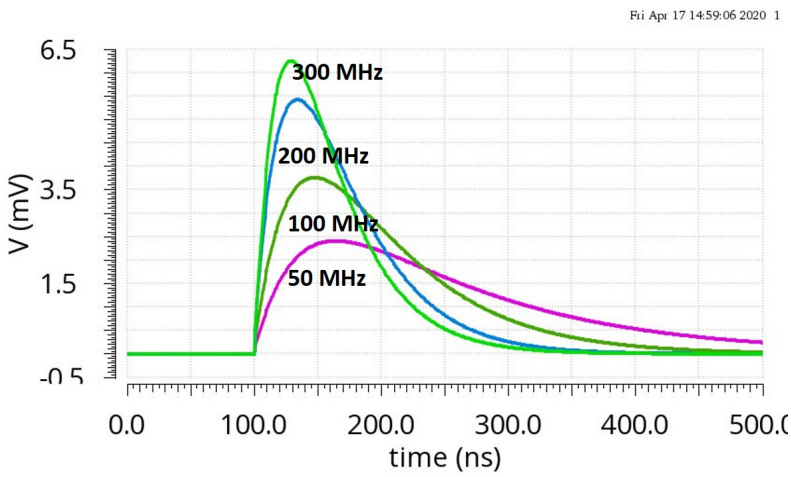


Figure 3.6: Output voltage and bandwidth of readout circuit, the gain of voltage amplifier is 20, $R_{in} = 50\Omega$

size of $30\mu m \times 30\mu m$, which is developed for Darkside-20k project, optimized and verified for cryogenic operation (87 K)[41]. The dedicated parameters are shown in table 3.1.

Parameter	Value@77 K	Value@300 K
Gain	1.9 $\times 10^6$	1.65 $\times 10^6$
R_q	6.5 MΩ	1.5 MΩ
C_q	5 fF	5 fF
C_d	61 fF	53 fF
R_d	1 kΩ	1 kΩ
C_p	100 pF	100 pF

Table 3.1: List of SiPM parameters used for simulation

3.2 Cold CMOS electronics design review

3.2.1 Transistor properties and temperature

CMOS electronics operating at cryogenic temperature bring significant improvements such like low power consumption, ultra-low noise and fast speed. As a consequence, recently there is a considerable attention for the use of cryo-CMOS in fields like quantum computer [101], space applications [82], neutrino and dark matter experiments [18] [3]. Neutrino and dark matter experiments usually use large volume

liquid Argon or Xenon based Time Project Chambers operating underground. The front-end electronics have to operate together with the detectors mounted closely with the TPC to minimize the noise.

Physics properties of semiconductor materials usually show strong temperature dependence, and there are lots of studies that have been carried out to reveal the transistor parameters in cryogenic temperature. Among them, two main conditions are in liquid Holmium (4.2 K) and liquid Nitrogen (77 K), and we focus on the latter one considering our application, in which the transistor behaviours are simpler to predict because of a substantial increase of carrier-out effect under 77 K [74]. A summary is given below:

Mobility increase

Mobility characterizes the drift velocity of the carrier under the electric field, expressed with the equation 3.14, where μ is the mobility:

$$v_d = \mu \cdot \mathbf{E} \tag{3.14}$$

In transistors, the drain current and trans-conductance is directly depending on the mobility. The mobility of electrons is about three time than the holes. Take NMOS as the example, the drain current in linear region is expressed with the equation 3.15, and the equation for saturated region is shown in 3.16 [79]. Where these two equations are actually derived from 3.14.

$$I_{DS} = \mu_n Cox \frac{W}{L}[(V_{GS} - V_{TH})V_{DS} - \frac{1}{2}V_{DS}^2] \tag{3.15}$$

$$I_{DS} = \frac{1}{2}\mu_n Cox \frac{W}{L}(V_{GS} - V_{TH})^2 \tag{3.16}$$

Table 3.2 describes the variables used to define the drain current.

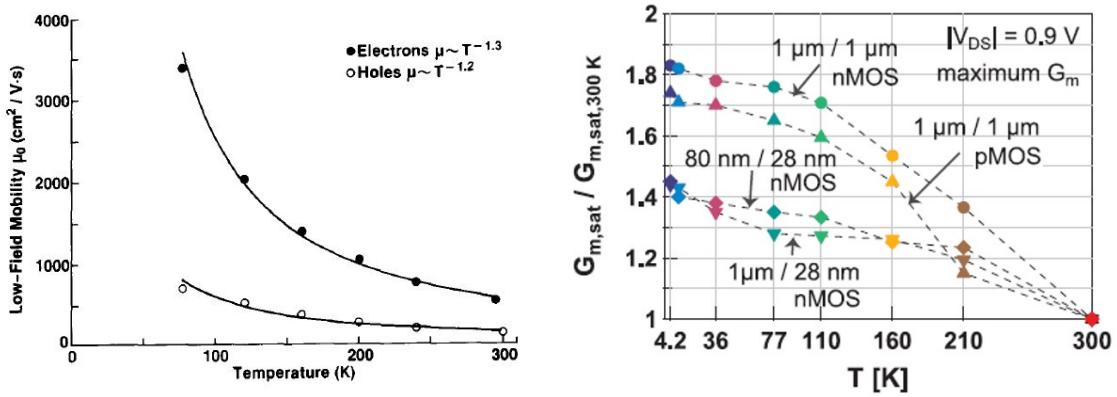
Variables	Description
I_{DS}	Drain current
μ_n	Mobility of electron
C_{ox}	Gate-oxide capacitance per unit area
$\frac{W}{L}$	Width Length ratio
V_{GS}	Gate-to-Source Voltage
V_{DS}	Drain-to-Source Voltage
V_{TH}	Threshold Voltage

Table 3.2: Variables description in drain current equation

And the transconductance (g_m) is defined as the ratio of the change in drain current to the change of the gate to source voltage. The equation of g_m in the saturated region is expressed below:

$$g_m = \frac{\partial I_{DS}}{\partial V_{GS}} = \mu_n C_{ox} \frac{W}{L}(V_{GS} - V_{TH}) \tag{3.17}$$

Experimental results show that the mobility of MOSFET could increase 4-6 times from 300 K to 77 K due to the reduction of carrier scatter result from lattice vibration [21]. Figure 3.7a shows the temperature relation with channel mobility both for electron and hole, which is characterized with a 0.5 μm technology NMOS and PMOS in linear mode. Figure 3.7b gives the measured results in a much smaller technology node (28 nm).



(a) Measured mobility of 0.5 μm node NMOS and PMOS in linear mode [21]

(b) Measured g_m of 28 nm node NMOS and PMOS in saturation mode [11]

Figure 3.7: Increase in mobility as temperature is reduced

Threshold voltage increase

In semiconductor physics, the threshold voltage is defined as the gate-to-source voltage when a conducting path between the source and drain terminals is created. For the NMOS device, this positive gate-to-source voltage could attract enough free electrons to the gate and build the conducting channel where the concentration of free electrons is just equal to the holes of the P-type substrate. The equation is given by 3.18 (without consideration of body effect) [79].

$$V_{TH} = \Phi_{MS} + 2\Phi_F + \frac{\sqrt{4q\epsilon_{si}\Phi_F N_{sub}}}{C_{ox}} \tag{3.18}$$

Where Φ_{MS} is the work function difference between gate and silicon substrate, Φ_F is the Fermi potential of bulk silicon with respect to the intrinsic Fermi level, q

is the charge of electron, ϵ_{si} is the dielectric constant of silicon, N_{sub} is the substrate doping concentration, C_{ox} is gate-oxide capacitance per unit area.

The temperature dependence mainly results from the Fermi potential Φ_F, which can be expressed as:

$$\Phi_F = \frac{kT}{q} ln(\frac{N_{sub}}{n_i}) \tag{3.19}$$

The intrinsic concentration of silicon n_i is very sensitive to temperature, and study carried out in work [67] gives the relation shown in equation 3.20:

$$n_i(T) = 5.29 \times 10^{19} (T/300)^{2.54} exp(-6726/T) \tag{3.20}$$

Besides, freeze-out effect is not severe to reduce the parameter N_{sub} considerably in 77 K and the corresponding plot is shown in figure 3.8 [31].

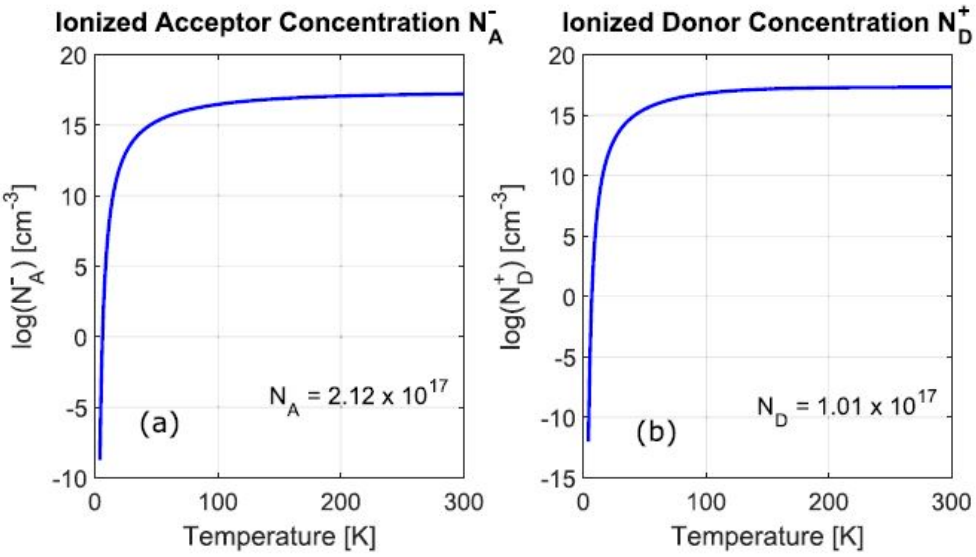
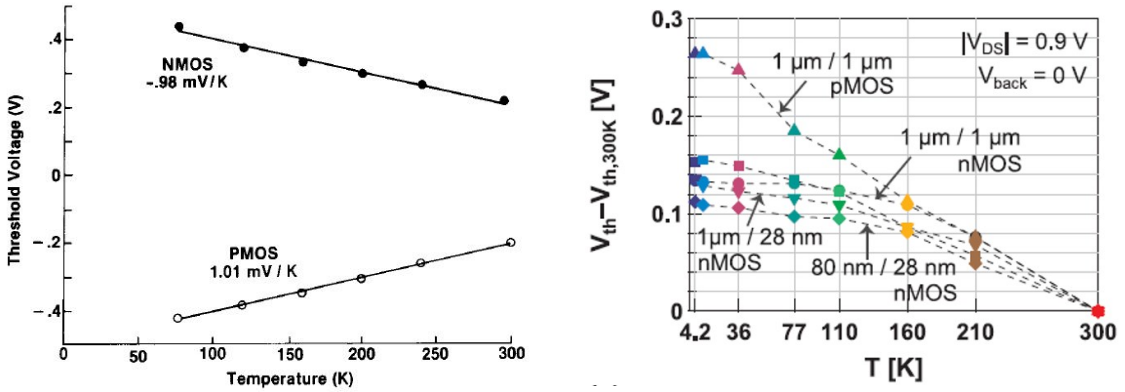


Figure 3.8: Temperature dependence of ionized acceptor and donor concentration [31]

To sum all effects, as temperature is reduced, the magnitude of Fermi potential, and hence, the magnitude of threshold voltage increases, where the magnitude of V_{TH} usually shows a nearly linear relation with temperature, and the magnitude increase about 1 mV/K to 2 mV/K with the temperature [103].

Figure 3.9a reports measured results for both NMOS and PMOS of a 0.5 μm CMOS technology, and figure 3.9b gives the results of a 28 nm CMOS technology.



(a) Measured V_{th} of 0.5 μm node NMOS and PMOS from 77 K to 300 K [21]

(b) Measured V_{th} shift of 28 nm node NMOS and PMOS from 4.2 K to 300 K [11]

Figure 3.9: Increase in mobility as temperature is reduced

Noise reduction

Thermal and shot noise are the two most prominent noise sources in electronic circuits. In a MOS transistor, flicker noise is also an important contribution [84]. Among these, the thermal noise both in MOSFETs and resistors can be reduced with the temperature decrease. Equation 3.21 and 3.22 are the expression of thermal noise density of a MOSFET and passive resistor respectively.

$$I_n^2 = 4kT\gamma g_m \tag{3.21}$$

$$I_{nR}^2 = \frac{4kT}{R} \tag{3.22}$$

Where k is Boltzmann constant, and γ is a factor related to channel length and operating mode.

Furthermore, flicker noise in PMOS is reported to decrease by a half from 300K to 77K [30], thus using PMOS as the input transistor can obtain lower noise.

Hot carrier effect

Hot carrier effect is the mechanism that energetic carriers ("hot carriers") obtain the ability to damage the Si/SiO_2 interface or overcome the barrier and to be trapped in oxide layer, when they flows from source to drain terminals in the channel. This effect results in the degradation of transistor parameters (like threshold voltage, transconductance, and drain current) with time. Cryogenic operation enhanced the hot carrier effect since the reduced scattering with the lattice. In other words, carriers are easier to gain enough energy.

Liquid Argon filled TPCs used for neutrino and dark matter experiments usually require a long term operation about 10 to 20 years or even more, which motivates the lifetime study caused by hot carrier effect especially for 77 K temperature [59] [49] [75].

The lifetime is usually defined as the duration to reach 10% variation on the transistor parameter ((g_m, V_{th} or I_{ds})). This can be measured by exposing the devices to stressing condition to accelerate the degradation. Where stressing usually means applying a voltage that exceed the ordinary value of the process.

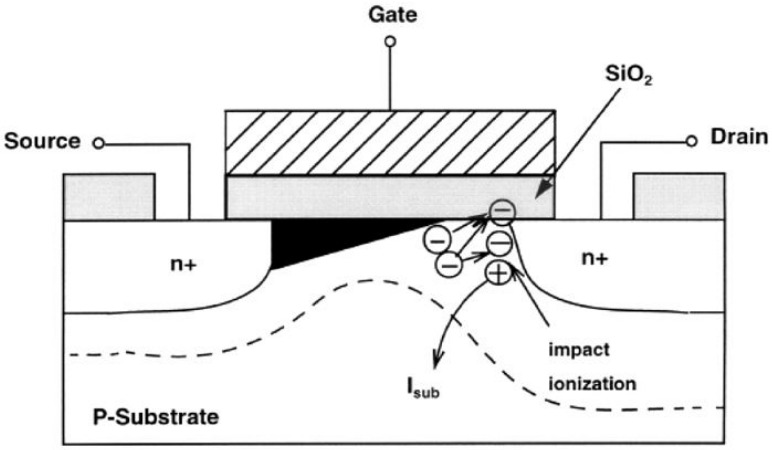


Figure 3.10: Schematic representation of impact ionization by hot electrons in the channel of an NMOS device. The holes produced by impact ionization constitute the substrate current [59]

Figure 3.10 shows the cross-section view of NMOS and the carriers drift. Hot electrons exceeding the energy of about $\phi_i = 1.3$ eV could create electron-hole pairs [96] which called impact ionization. Forced by electric field, the electrons drift to drain, while holes drift to substrate resulting in the substrate current. This current can be expressed as equation 3.23 [51]:

$$I_{sub} = C_1 I_{ds} e^{\phi_i / q\lambda E_m} \tag{3.23}$$

Where C_1 is a constant, I_{ds} is drain current, q is the elementary charge, λ is the mean free path of the electron, and E_m is electric field in the channel. The substrate current is a very useful parameter for monitoring the hot carrier degradation, and there is another equation to calculate the lifetime of device [51] [19]:

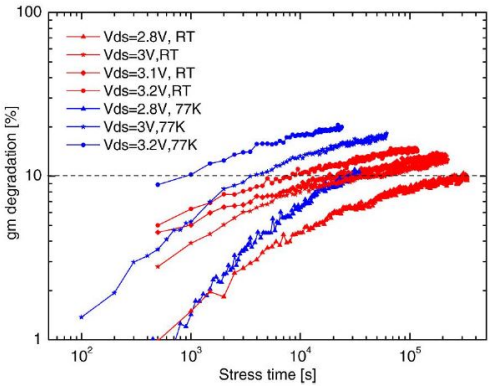
$$\tau = C_2 \frac{W}{I_{ds}} e^{\phi_{it} / q\lambda E_m} \tag{3.24}$$

Where ϕ_{it} is a higher energy level with $\phi_{it} \geq 3.7eV$ (about 4.6 eV for holes), C_2 is a constant and W is channel width. The less fraction hot electrons exceeding to ϕ_{it} are able to create interface state and be trapped in the oxide layer or even generate the gate current.

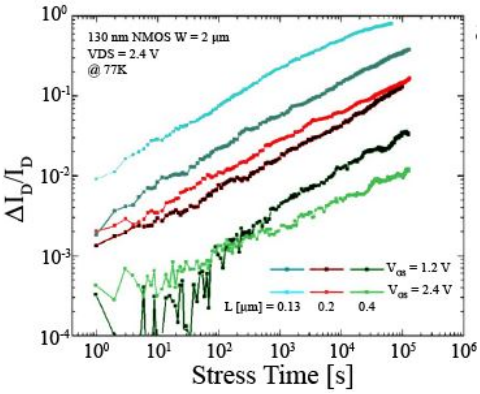
Combining the equation 3.23 and 3.24:

$$\tau = C_3 \frac{1}{I_{ds}/W} \left(\frac{I_{sub}}{I_{ds}}\right)^{-\phi_{it}/\phi_i} \tag{3.25}$$

Here, C_3 is a constant and a function of temperature, channel length and fabrication technology. As the relation is shown in equation 3.25, I_{sub} could be measured to monitor the lifetime.



(a) Measured trans-conductance degradation versus time of an NMOS transistor (L = 180 nm, W = 10 μm) in accelerated stress tests at room temperature (RT) and at 77 K. [59]



(b) Drain current variation versus stress time for different channel lengths [49]

Figure 3.11: Example of experimental results for hot carrier effect test

The experimental results are shown in figure 3.11, and some conclusions can be summarized below:

- Hot carrier effect is an ageing mechanism which does not result in sudden failure but in a decrease of the devices lifetime, and low-temperature operation exacerbates hot carrier degradation.
- Energetic carriers in channel create impact ionization near to the drain terminal, resulting in substrate current to exchange charge and hot carriers trapping into the gate oxide layer to degrade the parameters of g_m, V_{th} and I_{ds}.
- The basic approach of relieving this damage is to reduce the I_{ds}/W or E_m/L.
- PMOS shows a slower degradation than NMOS device.

3.2.2 Conclusion and guidance for cryogenic circuits design

From 300 K to 77 K, almost all characteristics of the transistor have changed sufficiently, which necessarily translates into an added complexity to the circuit design. With circuits operating at cold temperature, enhanced mobility of carriers provides improved transconductance (g_m) of MOSFETs.

The transconductance g_m is the key parameter in analogue circuit design, which is proportional to the gain of an amplifier built by MOSFETs. For a simple common source amplifier, the small-signal DC gain can be expressed by the equation:

$$A = g_m R_{out} \tag{3.26}$$

Where g_m is the transconductance of the input transistor, and R_{out} is the resistance seen from output node.

Besides, the transconductance of the input transistor for the pre-amplifier (PA) could define the bandwidth and noise level, where the PA stage is usually mounted directly with the detector and readout the signal.

g_m is also essential for time performance of the digital network. The switching activity is achieved by charging and discharging the circuit capacitance. This process is mainly related to the g_m, parasitic or load capacitance and power supply voltage. Thus enhanced g_m can help to improve the switching speed in time-sensitive circuits.

The increased threshold in most condition acts as a penalty to modify the operating point of each transistor. For example, the cascode topology may not be reliable because of the sufficiently reduced voltage margin. But it could reduce the leakage current in the cut-off state, consequently reducing the power dissipation in the digital circuit.

Thermal noise reduction at low-temperature helps to improve the signal-to-noise ratio because the thermal noise of the input stage tends to dominant the whole noise level, especially when readout high capacitance detectors.

Hot carrier effect should be considered especially at 77 K. The possible options could be using long channel transistors, low overdrive voltage and trying to choose PMOS as critical transistors instead of NMOS. while the last option may cause trade-off with the circuit performance (holes have lower mobility than electrons).

For some passive parameter, like the metal used for interconnecting the transistors, the parasitic resistance is reduced with the temperature decrease and is reported about one order magnitude decrease from 300 K to 77K [39] [54], while for degenerately doped poly-silicon, the decrease in resistance is not considerable (about 20 %) [39]. The capacitance of the oxide layer is virtually temperature independent [21], and the junction capacitance decreases at lower temperature, thus increasing built-in voltage and also depletion width. The junction capacitance could be further decreased at low temperature since the doping concentration in the MOSFET channel can be reduced to obtain a smaller threshold voltage in 77 K.

Nevertheless, this reduction of junction has only a minimal influence on the overall performance [21].

3.3 Test structures design and study

During the design phase, we need to take into account of the mismatch between the spice model of the MOSFETs and the real characteristic of transistors in 77 K. This is a consequence of the fact that the process design kit (PDK) released by CMOS foundries is only accurate at temperatures down to - 40 °C.

Using the foundry standard PDK, we design a first test chip embedding an input stage for readout of SiPM and basic digital modules for control logic of a mixed-signal design. The circuit level implementation of these IP blocks is described herein.

3.3.1 Schemes of front-end electronics and the test structures

In Darkside-20k project, more than 20 m^2 SiPM is applied to detection the photon signal (two kinds of scintillation lights S1 and S2). To minimize the noise, the front-end electronics need to work closely with the SiPM at the 87K temperature. The cryogenic electronics have to read out 24 cm^2 SiPM in one channel to reduce the background of photon detection module.

The original working scheme of the cryogenic electronics are implemented by the commercial operational amplifiers and the passive devices [3], shown in figure 3.12a. Four TIAs (Trans-Impedance Amplifier) are used to read out 6 cm^2 SiPM each, then the second amplifier sums the 4 branches and give a single channel output to the warm electronics system. The 6 cm^2 SiPM consists of six 1 cm^2 SiPMs organized as three in parallel and two in series to reduce the detector capacitance.

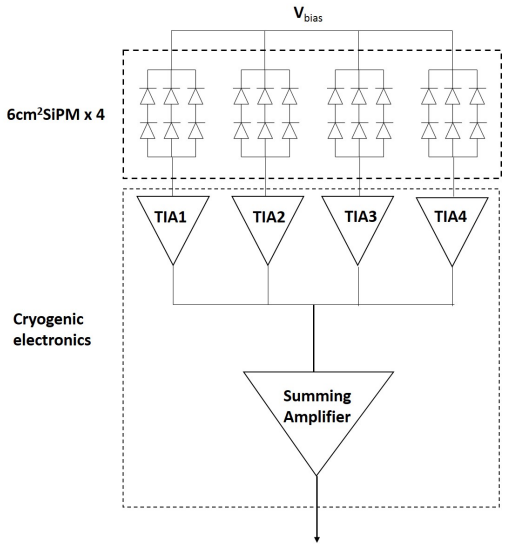
There are several drawbacks of the original scheme:

1. 24 cm^2 SiPM is read out by single channel, and the high detector capacitance hinders the front-end electronics to achieve a good performance. For example, this method can only read out the sensor signal in a 10 ns order timing resolution.

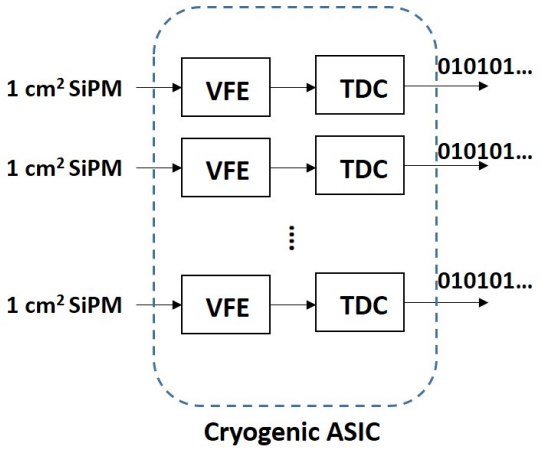
2. The analogue signal has a long transmission to warm electronics system, thus the signal can face the problem of signal to noise ration degradation.

3. The warm electronics system is complicated to analysis the analogue signal of the cryogenic electronics.

The only way to solve the above drawbacks is the digitization at cryogenic condition, and multi-channel readout is needed to read out smaller size SiPM per



(a) The original scheme of Darkside-20k project [29]



(b) New scheme proposed in this work

Figure 3.12: Two approaches to read out out large area SiPM and work at cryogenic temperatures

channel to improve the performance. Therefore, multi-channel ASIC with digitization on chip is the solution. Considering the power consumption budget, the scheme of timing information based measurement and digitization is proposed, as shown in figure 3.12b. In this way, One channel reads out smaller size SiPM (about 1 cm^2) to realize a much better measurement resolution. For instance, subnano second timing resolution is expected which makes it possible to reconstruct the 3D position information directly by the time of flight information of S1 signal (instead of only the drift time of S2). Fully digitized data output can be easily multiplexed and serialized, thus reducing the number of transmission cables. In the meanwhile, the digital output simplified sufficiently the warm electronics system.

According the two schemes described above, two test chips are designed and tested to study the properties of cold CMOS circuits.

TestChip1 integrates the front-end electronics following the scheme shown in figure 3.12a, where the specific circuits are implemented by different structure (TIAs and summing amplifier); TestChip2 consists of the very front end block and the basic logic control and data transmission module (Signal synchronization circuit and LVDS transmitter).

Both the two test chips adopt the same structure of input stage: Regulated Common Gate stage, working as a current buffer or conveyor to provide low input impedance and high bandwidth readout. The following content firstly introduces the analysis of the input stage and then describes the detailed design of the two

test chips. The experimental results are reported at last.

3.3.2 Input stage: the very front end structure

As discussed before, properties of SiPM, specifications of ASICs and guidance of cryogenic CMOS circuits design are reviewed. The input stage plays a key role to cope with the large capacitance sensor. To summarize the main characteristics of this part:

- Low input impedance and enough high bandwidth to improve the intrinsic timing resolution. Besides, low input impedance reduces the voltage variation of the input node during the readout;
- Stability when cope with a large range of detector capacitance, in this condition, from 0 to about 5 nF;
- Low power consumption, typically driven by a limited power budget of the system;
- Both anode and cathode signal readout capabilities;
- Short pulse width of its output, to avoid signal pile-up and increase the event rate ability.

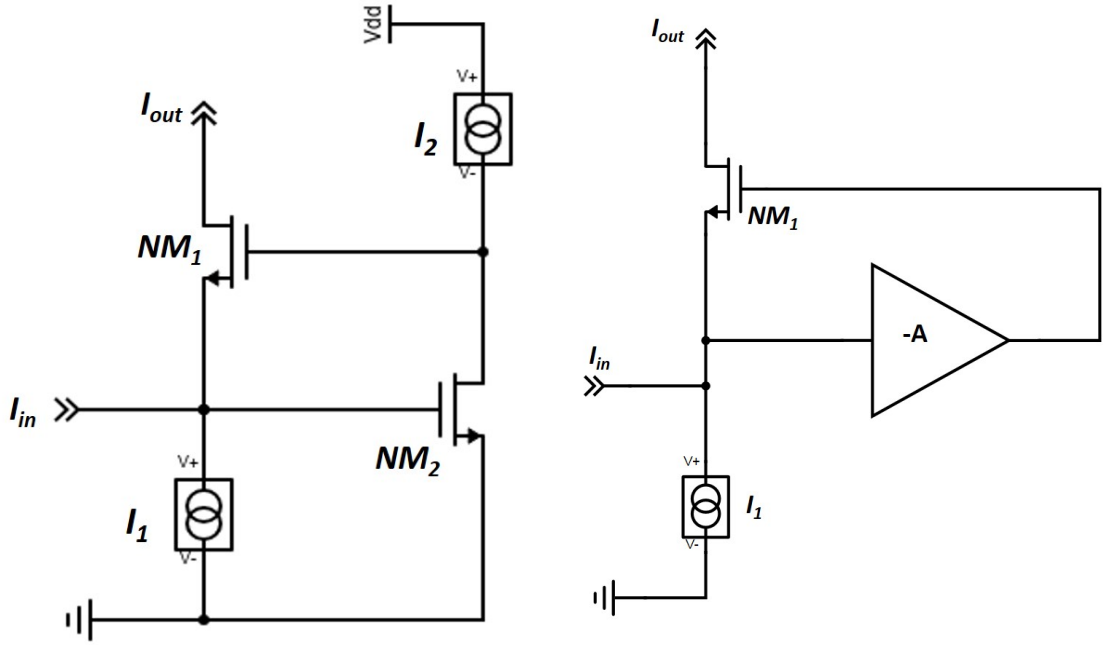
A Regulated Common Gate (RCG) topology is chosen for the input stage structure, which has been adopted in TOFPET ASIC reviewed above. Some theoretical analysis [66] has proved that the RCG structure is able to achieve higher bandwidth with the same power dissipation, in respect to a conventional voltage amplifier with a passive resistor and capacitor feedback based trans-impedance amplifier (TIA).

The input stage is expected to read out the signal from SiPM both anode and cathode, and finally outputs the voltage waveform, which is then fed to a discriminator.

Figure 3.13a shows the topology of RCG current conveyor, which can be described as a common gate stage regulated by a common source amplifier or a differential amplifier. The simplified structure is shown by figure 3.13b, which emphasizes the feedback inside, and the relation can be obtained:

$$I_{in} = g_{m1}(\Delta V_2 + \Delta V_1) = g_{m1}(-A\Delta V_1 + \Delta V_1) = -g_{m1}(A-1)\Delta V_1 \tag{3.27}$$

Where g_m is the trans-conductance of input transistor M_1 and A is the gain provided by common source amplifier (g_m boost stage). Consequently, the input impedance is driven by equation 3.28 and shows that the input impedance magnitude of common



(a) Regulated Common Gate structure works as current buffer

(b) Feedback in Regulated Common Gate

Figure 3.13: Working principle of Regulated Common Gate structure

gate stage is reduced by a factor of A, where the value can be assumed as 100 of a common source amplifier gain implemented by the sub-micron process.

$$Z_{in} = \frac{V_{in}}{I_{in}} = -\frac{1}{g_{m1}(A-1)} \approx -\frac{1}{g_{m1}A} \tag{3.28}$$

The small signal equivalent circuit of a generic regulated common gate amplifier is shown in figure 3.14. Here, using C_{in} to account for the sensor capacitance and the parasitic capacitance of the transistors at the input node which, in the scheme depicted in figure 3.13a, is mostly given by the gate-source capacitance of NM2. C_{gs1} is the capacitance between the gate and source of NM1, the common gate transistor.

The open-loop gain of the common source amplifier is given by equation (3.29):

$$A = g_{m2}(r_{o2}||r_{I_2}) \tag{3.29}$$

where g_{m2} is the transconductance of NM2, and r_{o2}, r_{I_2} are the output resistance of NM2 and current source I_2, respectively.

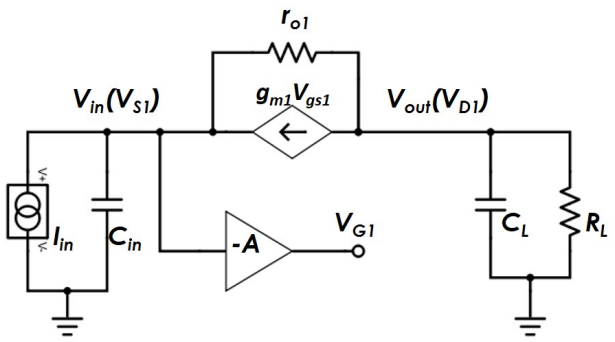


Figure 3.14: Small-signal equivalent of RCG

Likewise, we define g_{m1} and r_{o1} as the transconductance and output resistance of the input transistor NM1.

We can apply the KCL (Kirchhoff current law) at the input and output nodes to obtain the equations (3.30), (3.31) and (3.32):

$$I_{in} + V_{S1}sC_{in} + (V_{S1} - V_{D1})r_{o1}^{-1} + (1 + A)V_{S1}sC_{gs1} = g_{m1}V_{gs1} \tag{3.30}$$

$$V_{D1}Z_L^{-1} + g_{m1}V_{gs1} = (V_{S1} - V_{D1})r_{o1}^{-1} \tag{3.31}$$

and,

$$Z_L = R_L || \frac{1}{sC_L} \tag{3.32}$$

where R_L and C_L are the lump load impedance connected to the drain of NM1. From (3.30) and (3.31) we can use the following approximation:

$$g_{m1}(A + 1) \approx g_{m1}A >> r_{o1}^{-1} \tag{3.33}$$

Assuming high-capacitance detectors, the following approximation is also valid:

$$C_{in} >> (A + 1)C_{gs1} \tag{3.34}$$

As a consequence, we can obtain a simplified transfer function:

$$T_s = V_{out}/I_{in} = -\frac{R_L}{(1 + s\tau_i)(1 + s\tau_L)} \tag{3.35}$$

and the relation that defines the input impedance becomes:

$$Z_{in} = \frac{1}{Ag_{m1}} \tag{3.36}$$

Generally, the two main poles of the g_m-boosted common gate amplifier are defined at the input and output nodes (source and drain of NM1 in figure 3.13a):

$$\begin{cases} \tau_i = \frac{C_{in}}{Ag_{m1}} \\ \tau_L = R_L C_L \end{cases} \tag{3.37}$$

Since the frequency of input pole is A times higher than the ordinary common gate topology, the RCG input-stage is suitable for the readout of sensors with high capacitance.
A third pole τ_R, introduced by the RC time constant seen at the node VG1, defines a frequency dependent gain of the common source stage:

$$A(s) = \frac{A_0}{1 + s\tau_R} \tag{3.38}$$

The equation (3.38) is a revised relation of (3.29), considering (3.38) and also the effect of C_{gs1}. Starting from equations (3.30) and (3.31), one can write the complete transfer function:

$$T_s = -\frac{g_{m1}A_0R_L}{[s^2C_{in}\tau_R + s(C_{in} + A_0C_{gs1}) + g_{m1}A_0](1 + sR_LC_L)} \tag{3.39}$$

The denominator thereby is comprised of a second order polynomial, which may have complex conjugate roots. To avoid the complex conjugate roots, the following relation is necessary:

$$(C_{in} + A_0C_{gs1})^2 > 4g_{m1}A_0C_{in}\tau_R \tag{3.40}$$

which can be rewritten as:

$$C_{in}^2 + (2A_0C_{gs1} - 4g_{m1}A_0\tau_R)C_d + A_0^2C_{gs1}^2 > 0 \tag{3.41}$$

From 3.41 we can define the minimum stability margin, corresponding to the relation:

$$C_{in} = A_0(2g_{m1}\tau_R - C_{gs1}) \tag{3.42}$$

If the sensor capacitance is sufficiently high, with the approximation $C_{in} >> A_0C_{gs1}$, (3.41) becomes:

$$C_{in} > 4g_{m1}A_0\tau_R \tag{3.43}$$

On the other hand, when the sensor capacitance is small, we can obtain:

$$C_{in} < \frac{A_0C_{gs1}^2}{4g_{m1}\tau_R} \tag{3.44}$$

Interestingly enough, according to the (3.43) and (3.44), the RCG circuit is able to avoid the complex conjugate roots when the sensor capacitance is either very small or very high, while a transimpedance amplifier may suffer from instability when the C_d is very large[84]. Therefore, the RCG amplifier is particularly suitable to achieve a fast readout for sensors with very large terminal capacitance. Furthermore, for intermediate values of C_{in}, small values of τ_R and g_{m1} are preferred in order to avoid possible complex conjugate roots in (3.39).

In order to verify this hypothesis, we perform an analysis based on the simulation of the simplified schematic shown in figure 3.15. An ideal current source is employed to provide the common gate current (I_{cg}), and C_l is used to model the load capacitance. The stability simulation of the closed-loop formed by the boosted stage (implemented by a limited bandwidth amplifier) is carried out, where A is defined by equation 3.38 using $A_0 = 50$, $\tau_R = 1k\Omega \times 2pF = 2ns$, $C_{gs1} = 1pF$.

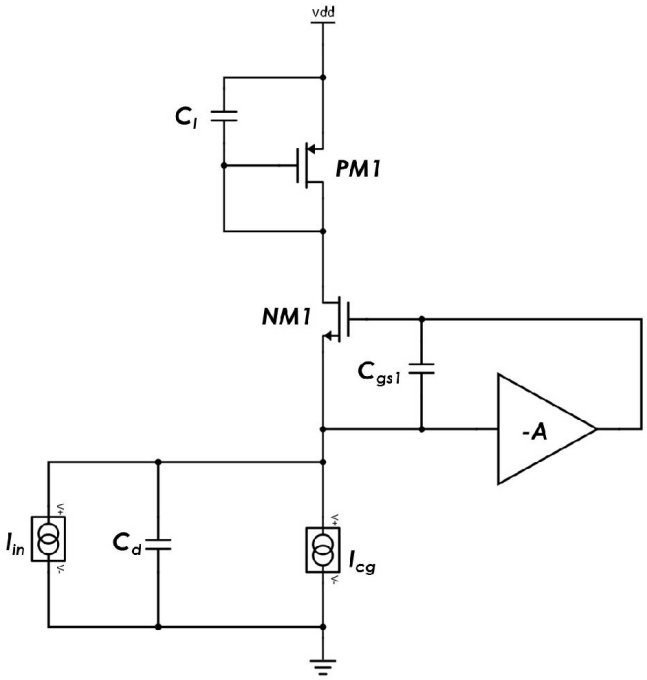


Figure 3.15: Schematic for stability simulation

Table 3.3 summarises the simulation results in terms of the phase margin obtained at different bias conditions of NM1, as a function of the input capacitance. The results are compatible with the analysis above. Setting a smaller common gate current can increase the phase margin for intermediate values of C_{in}.

Phase Margin I_{cg} / C_{in}	10 μA	50 μA
5 pF	92.1 deg	81.3 deg
20 pF	85.4 deg	64.3 deg
50 pF	82.0 deg	58.2 deg
100 pF	82.2 deg	59.0 deg
350 pF	86.3 deg	70.5 deg
500 pF	87.5 deg	74.7 deg

Table 3.3: Stability simulation results

The two dominant sources of electronic noise in MOS devices are flicker and thermal noise. Flicker noise can be modelled with a voltage source series-connected to the gate of the transistor, and expressed as:

$$V_{nf}^2 = \frac{K_f}{C_{ox}WL}\frac{1}{f} \tag{3.45}$$

where K_f is a constant given by the process, C_{ox} is gate oxide capacitance per unit area, WL is the gate area. Its contribution is minimized by choosing a proper area for transistors and decreasing the transconductance of the current sources.

Thermal noise, which spectral density in MOS devices can be represented through a resistor analogy, is given by the general expression of (3.46):

$$I_{nt}^2 = 4kT\gamma g_m \tag{3.46}$$

where k is the Boltzman constant, T is absolute temperature, and γ is a complex function of the basic transistor parameters and bias conditions, with a typical value of 2/3 or higher. The equivalent input node thermal noise of NM2 can be expressed as:

$$V_{n2}^2 = I_{n2}^2/g_{m2}^2$$

where the I_{n2}^2 is the current mode noise defined by equation (3.46). For NM1, since its transconductance is boosted by the factor of A, the equivalent input noise can be expressed as:

$$V_{n1}^2 = I_{n1}^2/A^2g_{m1}^2$$

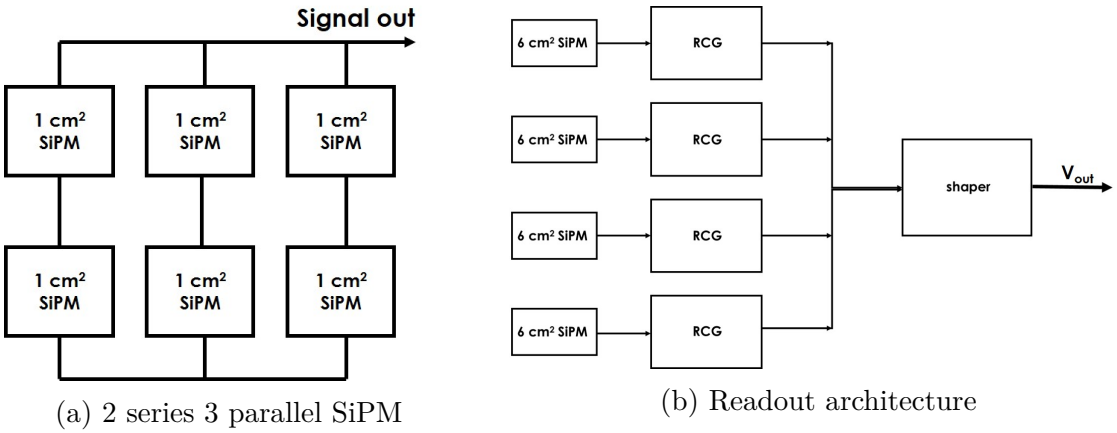
Considering of $A^2g_{m1}^2 >> g_{m2}^2$, NM2 is the dominant source of thermal noise, and its noise contribution to output is:

$$V_{no}^2 = \int_0^\infty V_{n2}^2|T_sZ_{in}|^2df = \frac{R_L^2}{A^2g_{m1}}\frac{I_{n2}^2}{g_{m2}^2}\frac{1}{4(\tau_i+\tau_L)} \tag{3.47}$$

Equation(3.47) indicates that increasing the g_{m2} can considerably decrease the overall noise of RCG circuit, which justifies the need of using a quite large size of common source transistor (NM2).

3.3.3 TestChip1 design

The architecture of TestChip1 is shown 3.16, already mentioned before. In this implementation, each quadrant of 6 cm^2 SiPM is divided in 6 sub-cells each one consisting of 1 cm^2 SiPM, organized with 2 in series and 3 in parallel (2s3p) topology, in order to have a trade-off on reducing the total detector capacitance and the signal gain. In this work, 4 RCGs work as the input stages to read out 6 cm^2 SiPM each. The shaper as the second stage to sum the signals from 4 RCGs and bring the single output which can be probed directly with the oscilloscope.



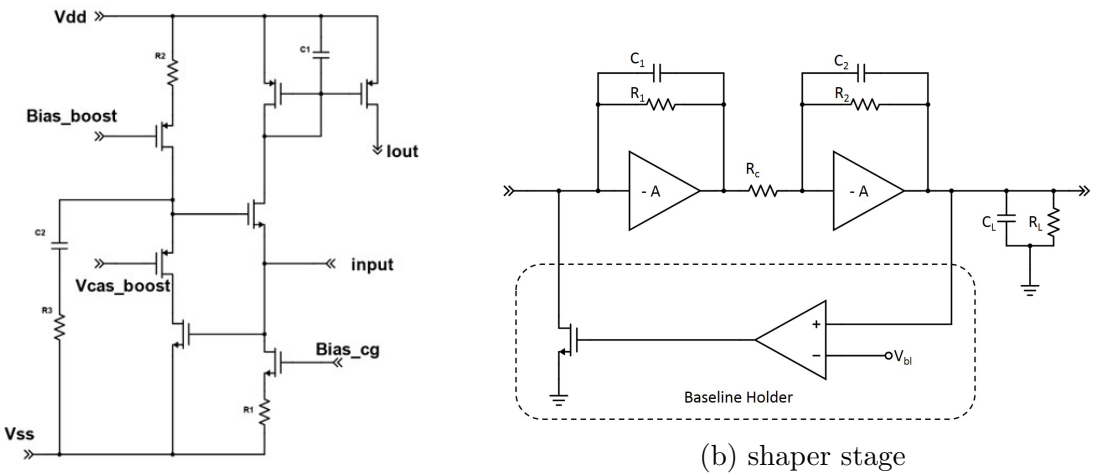
(a) 2 series 3 parallel SiPM

(b) Readout architecture

Figure 3.16: Readout scheme for large area SiPM in one channel

More detailed design is shown in figure 3.17. The RCG stage is optimized for larger area SiPM (6 cm^2) readout, a cascade transistor controlled by Vcas_boost is added in g_m boost stage to improve the ability of g_m enhance, R_2 and R_1 are used to reduce the thermal noise in current source of boost and common gate stage respectively, C_1 is used to improve the SNR of the output signal, and C_2 in series with R_3 are aimed to adjust the poles and zeros in feedback stage to avoid the possible instability.

The summing stage, shown in figure 3.17b, is a conventional CR-RC shaper and works as a TIA. The baseline holder (BLH) structure (transistor-level design in figure 3.19) is used to keep the baseline equal to configurable reference voltage V_{bl}. The front-end circuits are implemented with UMC CMOS process, and the power supplies are ±1.25V. The schematic of shaper-core amplifier is shown in figure 3.18. It is a folded telescopic cascade amplifier with differential inputs and signal output. The class-AB structure is used to enhance the gain and improve the voltage swing.



(a) RCG stage

(b) shaper stage

Figure 3.17: Transistor level schematic design

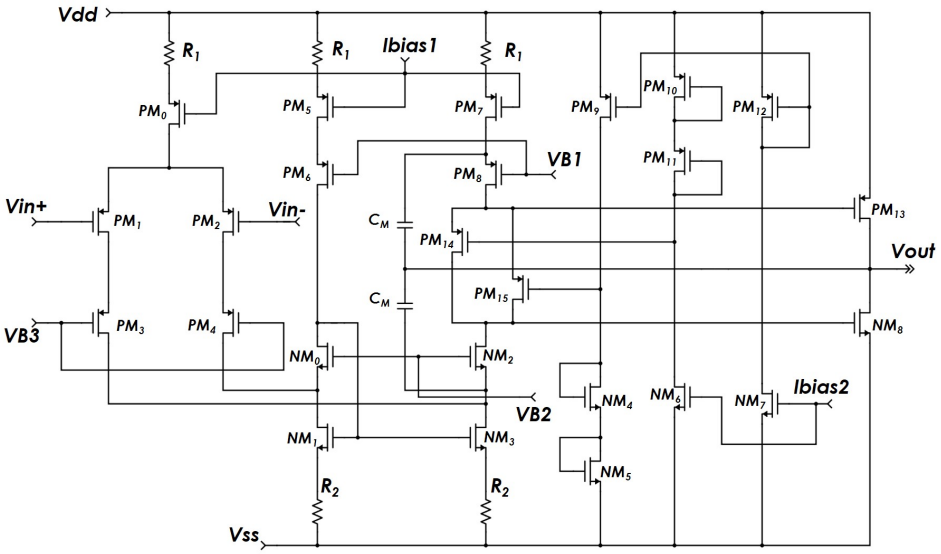


Figure 3.18: Transistor level design of shaper-core amplifier

Table 3.4 summarized the bias values in RCG, shaper and BLH blocks.
Table 3.5 shows the post-layout simulation results.
Figure 3.20 shows the CAD view of the test chip layout.

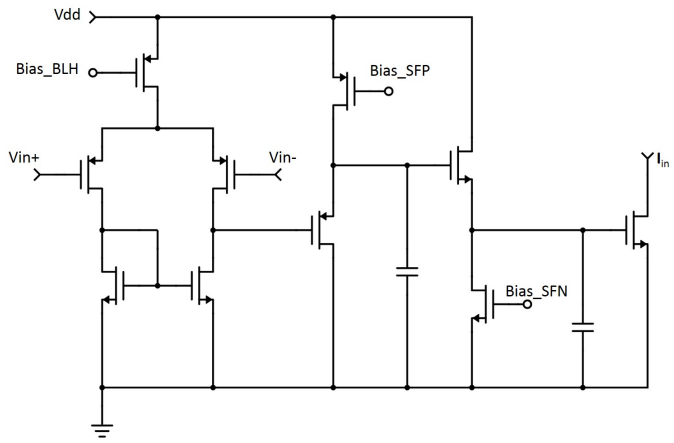


Figure 3.19: Transistor level design of Baseline Holder

Bias name	Position	Type	Nominal value
Bias_cg	RCG	current	100 muA
Bias_boost	RCG	current	6 mA
Vcas_boost	RCG	voltage	500 mV
Ibias1	Shaper	current	2 mA
Ibias2	Shaper	current	250 muA
VB1	Shaper	voltage	150 mV
VB2	Shaper	voltage	250 mV
VB3	Shaper	voltage	500 mV
Bias_BLH	BLH	current	5 nA
Bias_SFP	BLH	current	30 nA
Bias_SFN	BLH	current	10 pA
V_{bl}	BLH	voltage	0 V

Table 3.4: Bias setting of the front-end

Specification	300 K	77 K
Gain	17 mV/pe	12 mV/pe
SNR of 1 pe	8	12
jitter of 1 pe	24 ns	25 ns
Power consumption	78.3 mW	77.8 mW

Table 3.5: Post-layout simulation results

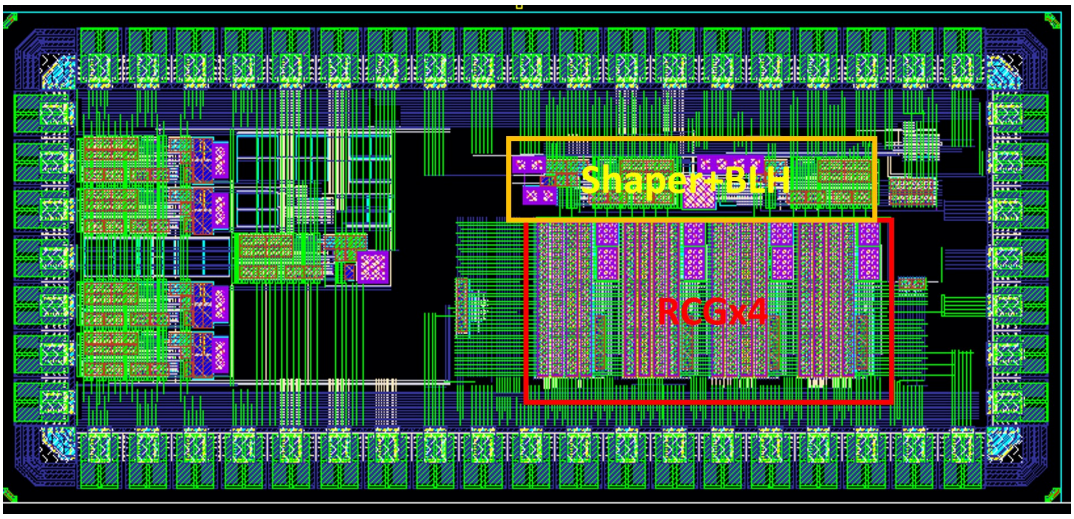
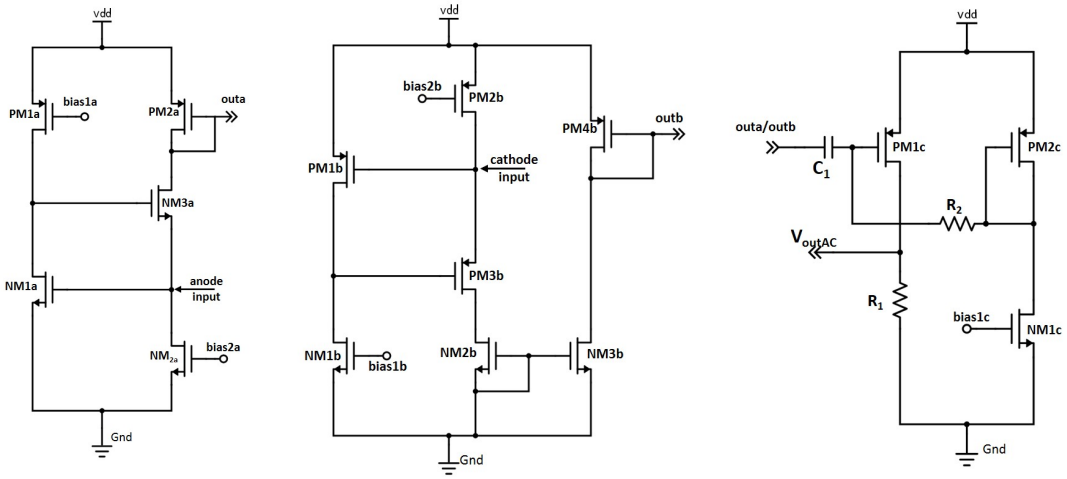


Figure 3.20: Layout view of TestChip1, 2.6 mm x 1.2 mm

3.3.4 TestChip2 design

The new scheme is expected to read out smaller size of SiPM (1 cm^2, with the sensor capacitance of several nF) and digitize the timing information on chip (as shown in figure 3.12b), thereby the test structures include the very front-end circuit and the basic logic control and data transmission circuits.

Figure 3.21 shows the Very front-end with AC coupled method.

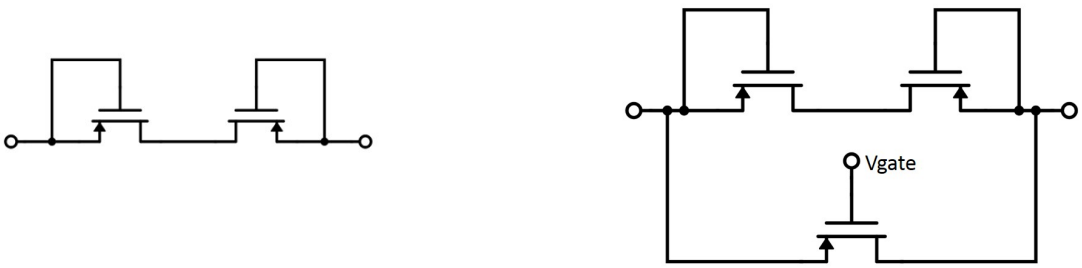


(a) Anode readout (b) Cathode readout (c) AC coupled output

Figure 3.21: Dual polarities readout RCGs with AC coupled scheme

Two complementary RCG structures work for anode (figure 3.21a) or (figure 3.21b) cathode signal readout. The external configuration can be set to choose the working mode and set the anode or cathode RCG signal to the output stage , AC coupled by capacitor C_1 (figure 3.21c). The output stage is a basic common source amplifier with constant resistance load and can be regarded as a trans-impedance amplifier (TIA) to convert the current signal from RCG to the voltage output.

The DC operation point of PM1c is defined by PM2c and NM1c through R_2, where R_2 is in the order of $M\Omega$ to $G\Omega$ to resist the signal loss. The CMOS process can not practically embody such large resistor and two back to back PMOS are in series to implement R_2, shown in figure 3.22a.



(a) back to back MOSFETs

(b) build a path to avoid operating point set failure

Figure 3.22: Large resistance implemented by MOSFETs to set DC operating point

The two PMOS transistors are both diodes connected, and in series with opposite direction, thus at least one PMOS are cut-off to be resistive enough. However, considering the operating at 77 K, the threshold voltage will increase significantly as discussed above, which acts as the "break" of path, causing a considerable voltage drop in DC operating point setting. The simulation result is shown in figure 3.23, the proper value for R_2 is about from $M\Omega$ to $G\Omega$. With the temperature decreasing, the R_2 increasing with exponential speed and can reach $T\Omega$ order which is actually the open path. Furthermore, the gate current is in order of pA. Therefore, the voltage drop can be several tens of mV in 77K.

It should be mentioned that the simulation model is no longer accurate at temperatures below - 40 °C. Nevertheless, possible failure in setting proper DC point need to be considered at cold temperature. Thus an updated version of R_2 shown in 3.22b has been implemented by adding a new PMOS in parallel where a new gate voltage V_{gate} could be set externally to avoid the possible failure.

A better way to deal with the DC point problem at cryogenic temperature is DC-coupled the two RCGs to the output TIA stage. The schematics are shown in 3.24. Compared to the RCGs in AC coupled scheme, the DC coupled one applies a configurable current source in parallel with an output transistor of RCGs in order to adjust the DC operation point, this is, the baseline of voltage output.

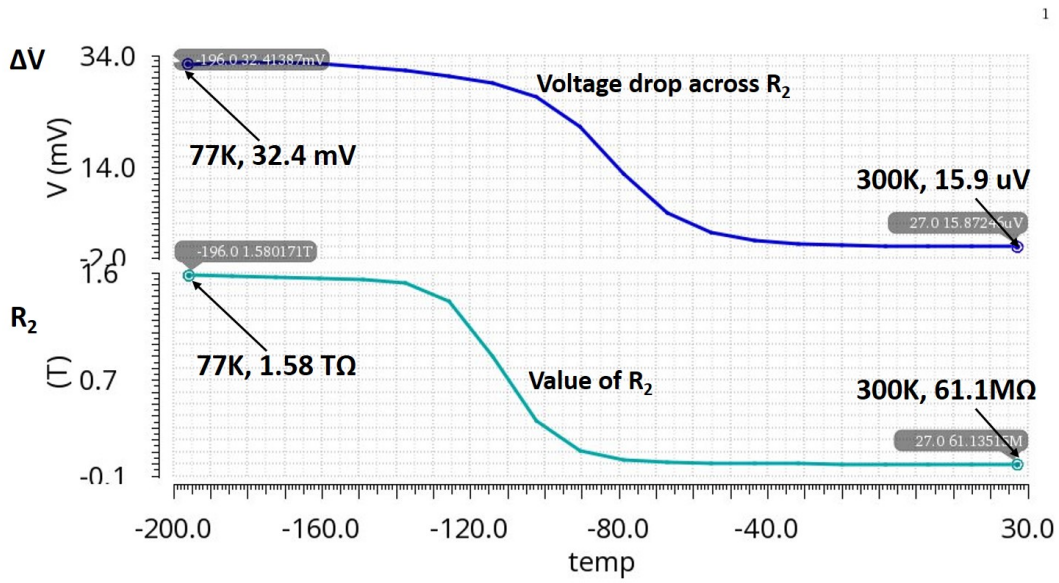
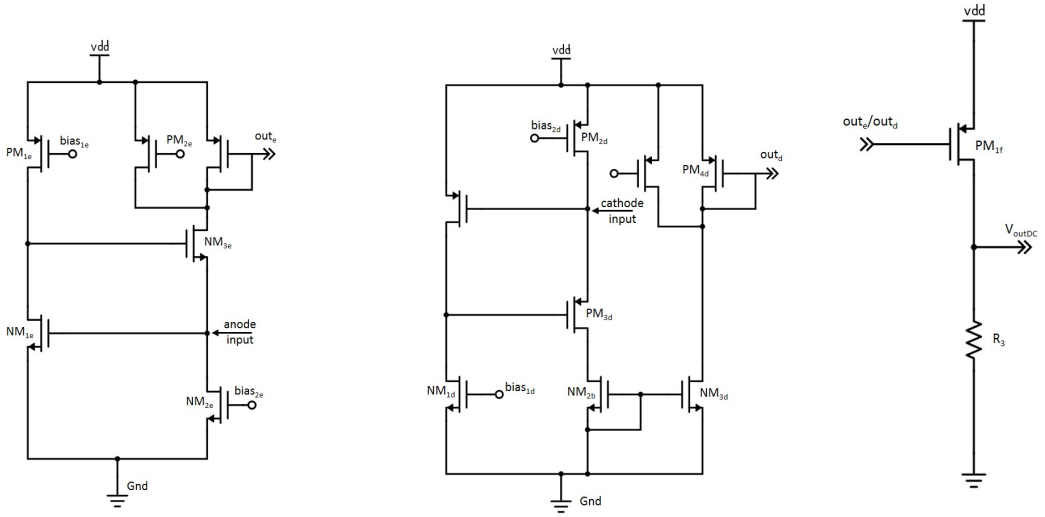


Figure 3.23: Temperature simulation to show the variation of voltage drop (ΔV) and value of R_2



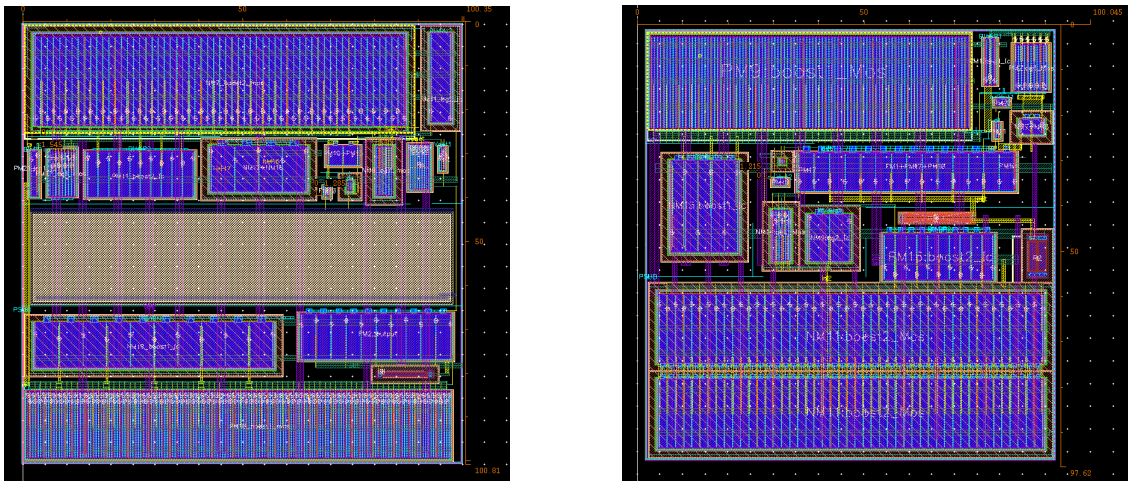
(a) Anode readout

(b) Cathode readout

(c) DC coupled output

Figure 3.24: Dual polarities readout RCGs with DC coupled scheme

The layouts of VFEs shown in figure 5.10a are both defined as the area of 100 x 100 μm^2 to fit the next step integration with the other blocks.



(a) layout of AC coupled VFE

(b) layout of DC coupled VFE

Figure 3.25: Layouts of two VFEs

Figure 3.26 shows the waveforms of very front end circuit, simulated the condition of 1 cm^2 SiPM readout and single photon signal. In 300K, the SNR = 16.8, rms jitter = 3.2 ns; In 77K, SNR = 22.6, rms jitter = 0.6 ns. The simulation shows from room temperature to cryogenic temperature, the noise of very front end is expected to reduce and the bandwidth improve sufficiently.

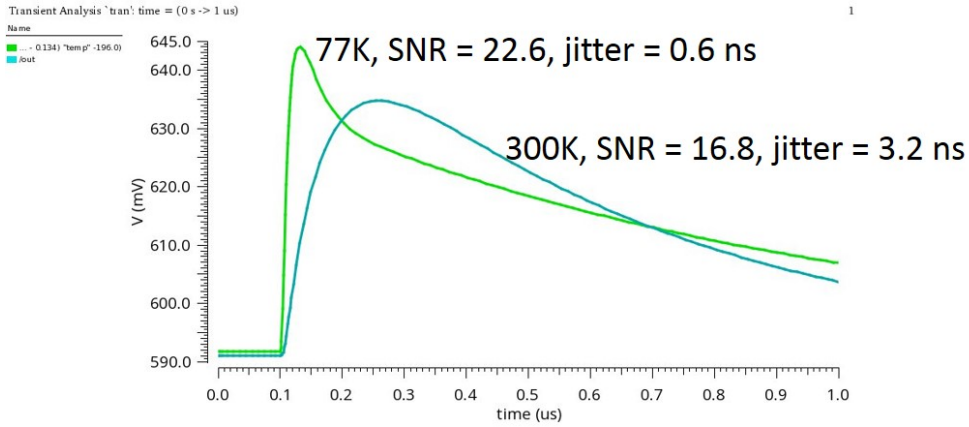


Figure 3.26: Simulation of the very front end (cathode-DC coupled) at both 300K and 77K

Figure 3.27 shows the TestChip2 layout. The digital blocks including signal synchronization module and LVDS transmitter are also marked in the TestChip2 layout.

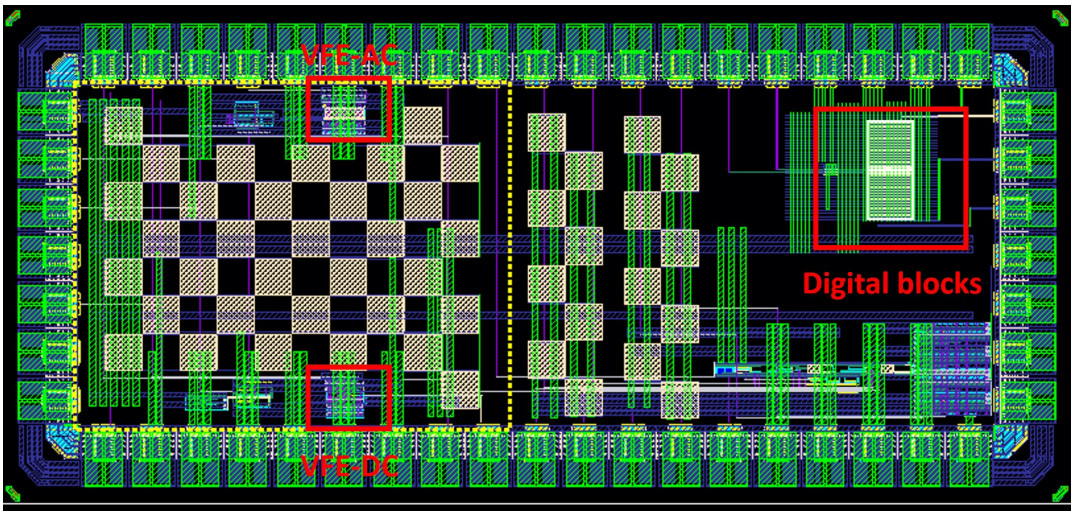


Figure 3.27: Layout view of TestChip2, 2.6 mm x 1.2 mm

3.4 Experimental results of test chips

3.4.1 Characterization environment set up

As discussed above, two test chips are taped out with UMC CMOS 110nm process to study the front-end circuits for large-area SiPMs readout and to operate at cryogenic temperature.

- The TestChip1 includes the first readout scheme: operating with the power supply of ± 1.25 V, 4 RCGs readouts 6 cm^2 SiPMs independently, which is composed by 6 of 1 cm^2 SiPMs and managed with 2s3p topology. Then a conventional CR-RC shaper with a shaping time of about 300 ns is used as TIA stage and sums 4 RCGs current signals to one voltage output.

- The TestChip2 includes two VFEs based on RCG current sensitive amplifier, and the output stages are implemented with DC and AC coupling schemes respectively. Besides, some digital blocks in order to verify the logic control and data transmission in cryogenic temperature (77K).

The test environment set-up is shown in figure 3.28. The two test chips are wire bonded to their test boards. The pulse generator is used to create test pulses and the oscilloscope probes the output signals. The power supply provides the 2.5 V supply for two test boards, and the on board commercial voltage regulators are implemented to generate the expected power supplies for test chips (TestChip1, ±1.25 V; TestChip2, 0 to 1.2 V). The applied Lecroy WaveRunner 104XI oscilloscope has the bandwidth of 1GHz and sample rate of 10Gs/s, meeting the requirements of

the test. The input and output signal are transmitted by Sub-Miniature version A (SMA) cable.

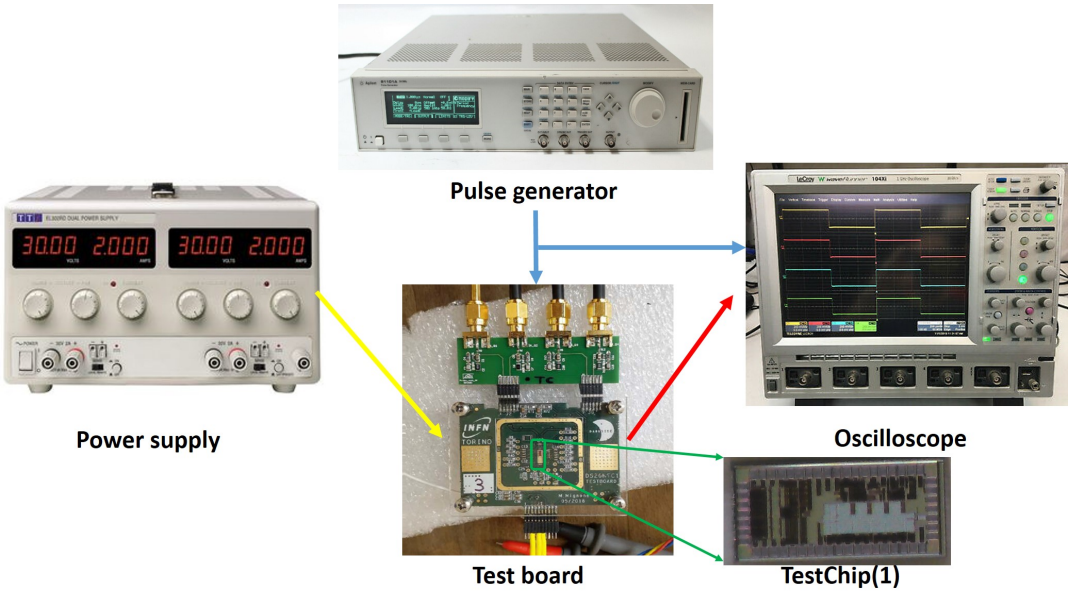


Figure 3.28: Test environment set-up

The cryogenic environment is realized by liquid nitrogen bath in a dewar (internal diameter of 15.5 cm and hight of 27.5 cm), shown in figure 3.29a and 3.29b.

Considering there is no configuration circuit on the test chips, the bias signals are set off-chip. Figure 3.30a shows the example of current type bias (PMOS) setting for the test chip, where the R_1 is a trimmable resistor off-chip and T_1 is used to monitor the voltage, and its value follows the simulation result. The voltage bias is implemented by voltage divider shown in 3.30b.

3.4.2 TestChip1 test

The input circuit of TestChip1 is designed on a sub-board which can be connected with the test board, shown in figure 3.31a to model the 6 cm^2 SiPM in 2s3p topology. The schematic of input circuit is shown in 3.31b, where the pulse generator creates a step voltage (V_{step}) and inject the amount of charge equal to $Q_{in} = V_{step} \cdot C_1$.

The passive resistor (R_2 and R_3) and capacitors (C_2, C_3, C_4) are used to model the SiPM sensor, the value and description of each part are shown in table 3.6.

An instability issue was observed at the output node in the first version of TestChip1 and has been subsequently verified by simulation. The reason for instability is linked to the design of the shaper stage (3.17b), where three closed loops

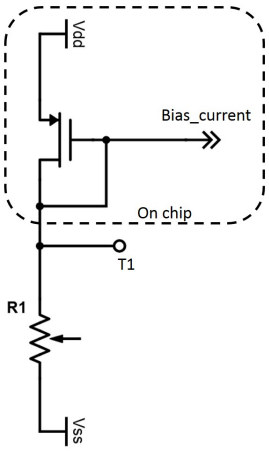
(a) Test environment in laboratory

(b) Liquid nitrogen bath to create cryogenic environment

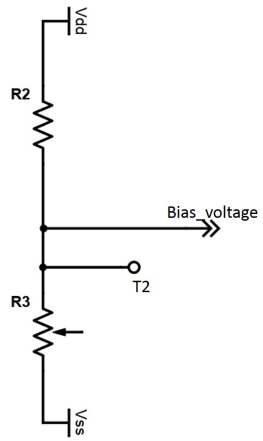
Figure 3.29: Study of the Test chips in cryogenic condition

Table 3.6: Parameters of TestChip1 input circuit

Parameter	Description	Value
V_{pulse}	pulse generator	variable step voltages
R_1	termination resistor	50 Ω
C_1	inject capacitor	18 pF
R_2	quenching resistors in SiPM	47 Ω
C_2	capacitance in SiPM	800 pF
R_3	series resistance	20 Ω
C_3	capacitance in SiPM	4.7 nF
C_4	capacitance in SiPM	4.7 nF

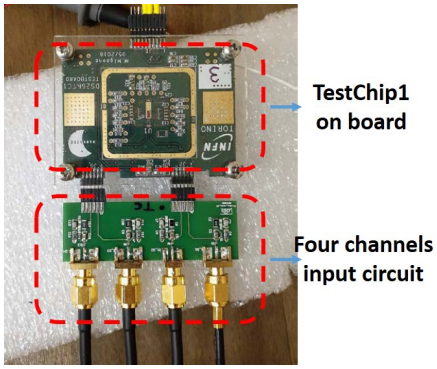


(a) Schematic example of current bias

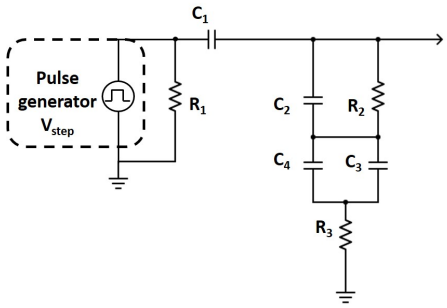


(b) Schematic example of voltage bias

Figure 3.30: Bias setting



(a) Test board is connected with 4-channel input circuit sub-board



(b) Schematic of input circuit

Figure 3.31: The input circuit of TestChip1 board

exist in this stage: the first and second shaper stage and the Baseline Holder. The shaper core amplifier is optimized for a 50 Ω load, while the first stage shaper has a load of 60 $k\Omega$.

A second version of the design was sent to foundry, and the revised schematic is shown in figure 3.32. In the new version, the Baseline Holder is removed because its quite small current (table 3.4) may cause some uncertain in 77 K and the shaper is replaced by a non-inverted voltage amplifier. The signal from RCGs are summed by a resistor R_1 to be converted to voltage and trimmed by a current source controlled

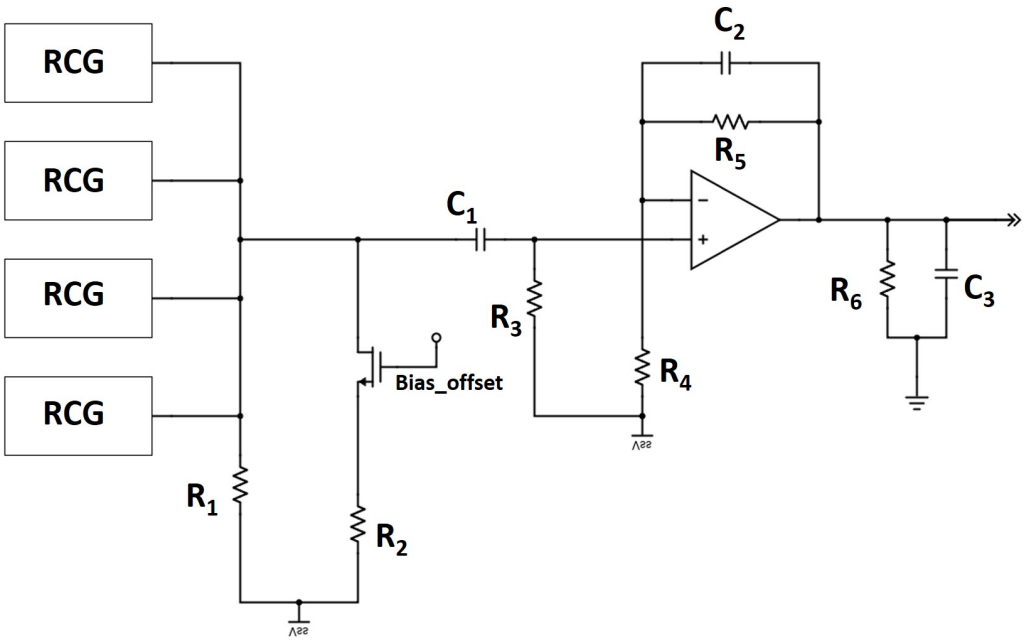
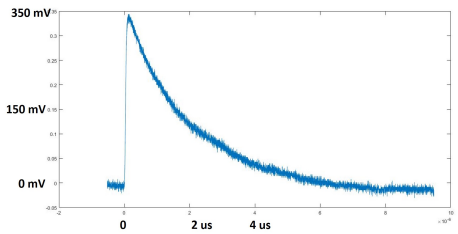


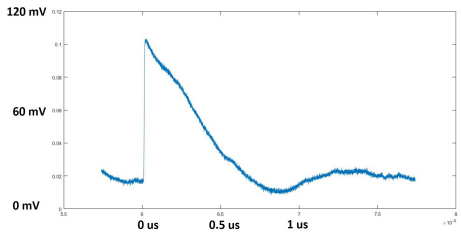
Figure 3.32: Architecture of new version of TestChip1

by $Bias\_offset$, then the signal is AC coupled (C_1 and R_3) to the next stage. The voltage amplifier of second stage has the same architecture of the original shaper-core. The ratio of R_5 and R_4 defines the voltage amplification value, the compensate capacitor C_2 is used to improve the signal to noise ratio of the second stage. R_6 and C_3 are used to model the load of oscilloscope.

Figure 3.33a and figure 3.33b show the waveforms sampled by oscilloscope both in 300 K and 77 K with a zero input capacitance load which prove that the summing stage is stable in both room and liquid nitrogen temperature.



(a) Waveform captured by oscilloscope at 300 K



(b) Waveform captured by oscilloscope at 77 K

Figure 3.33: Waveforms observed by oscilloscope

The input to output relation is characterized by the pulse generator and observed by the oscilloscope. The figure 3.34 shows the test result of four channels input to output response at 300 K, which means they have almost the same gain.

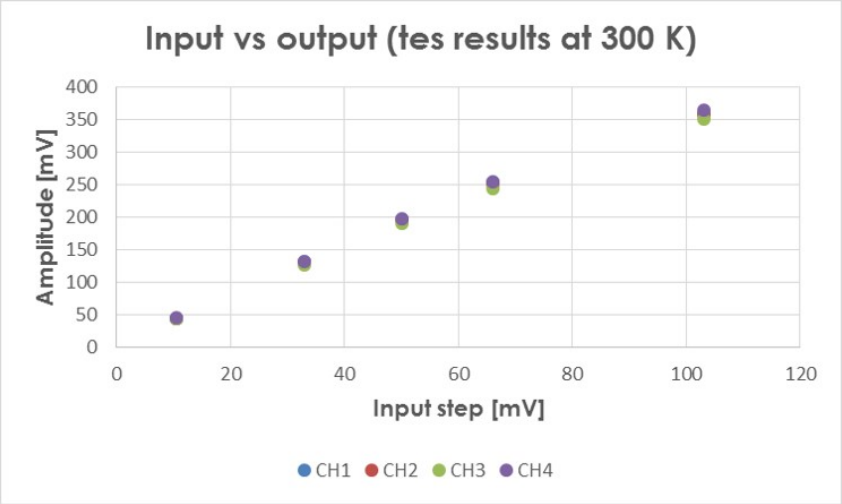


Figure 3.34: Test results of 4 channels input and output at 300 K

The comparison of post-layout simulation and the test of input to output relation is shown in figure 3.35, and two results have a good match.

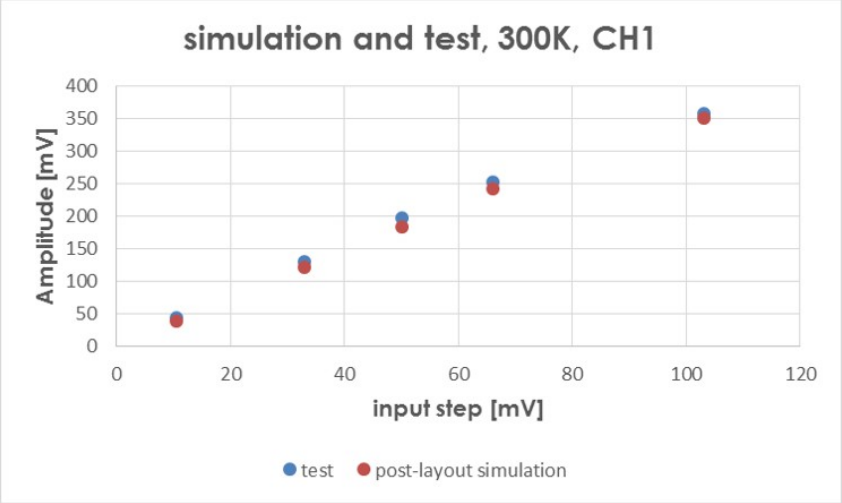


Figure 3.35: Simulation and test of input to output relation at 300 K (CH1)

Then, the table 3.7 summaries the characterization results in 300 K. The gain has a very good match between the test and simulation result. The output baseline has a little variation, resulting from the mismatch of the differential voltage

amplifier. The Equivalent Noise Charge (ENC) has increased by 36 %, the possible reason may be the interference from external sources. The larger rising time may be result from the parasitic parameters on the test board. To have a summary, TestChip1 can work as the expected.

Attributes	post layout simulation	test	mismatch
Gain	188 mV/pC	192 mV/pC	2 %
Baseline	4.7 mV	9.8 mV	5.1 mV
ENC	16.5fC	22.5fC	36%
Rising time	74 ns	87 ns	17.6 %
Power consumption	90.75 mW	96.25 mW	6.1 %

Table 3.7: Comparison results between post layout simulation and test in 300 K

However, the instability issue is still found in liquid nitrogen temperature with an large input capacitance load which is used to model the sensor capacitance, that indicates instability problem also exists in the input stages: the RCGs. However, the simulation results always show an enough phrase margin of the RCG feedback loop. Therefore, more studies should be carried out to reveal the root cause of the issue.

3.4.3 TestChip2 test

The output stage of the VFE in TestChip2 is implemented by a common source PMOS in series of a 10 $k\Omega$ resistor. Thus a single-end to the differential buffer on board is used to drive the output to be observed by the oscilloscope. Figure 3.36 shows the test setup and figure 3.37 shows the waveforms observed by the oscilloscope.

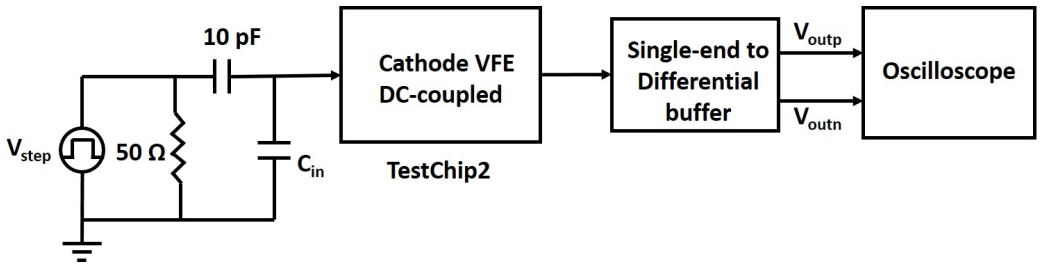


Figure 3.36: Setup of TestChip2 VFEs characterization (300K)

The very front-end circuits in Test-Chip2 contain four different working modules (DC-Cathode; DC-Anode; AC-Cathode; AC-Anode), the following experimental results are only from the DC-cathode circuit.

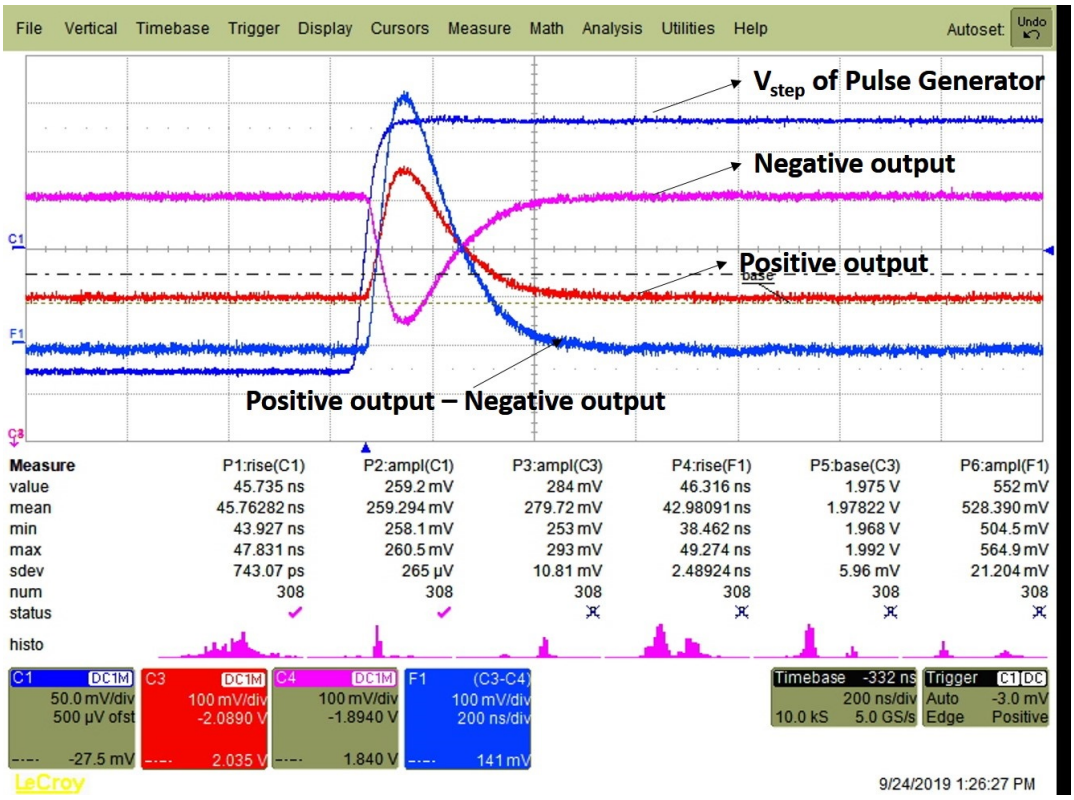
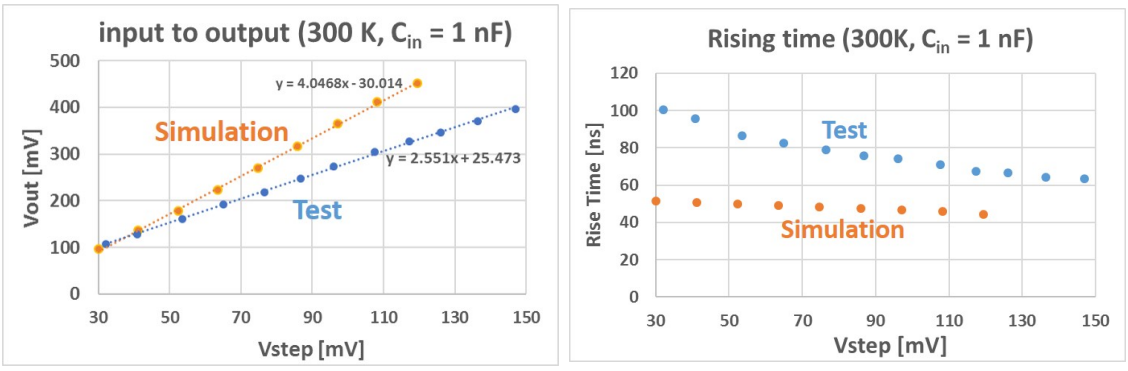


Figure 3.37: Waveform from VFE in Test-Chip2

VFE at 300 K

Figure 3.38 shows the input and output response of the very front-end circuit at room temperature.



(a) input and output in 300 K

(b) rising time in 300 K

Figure 3.38: Amplitude and rising time measurement

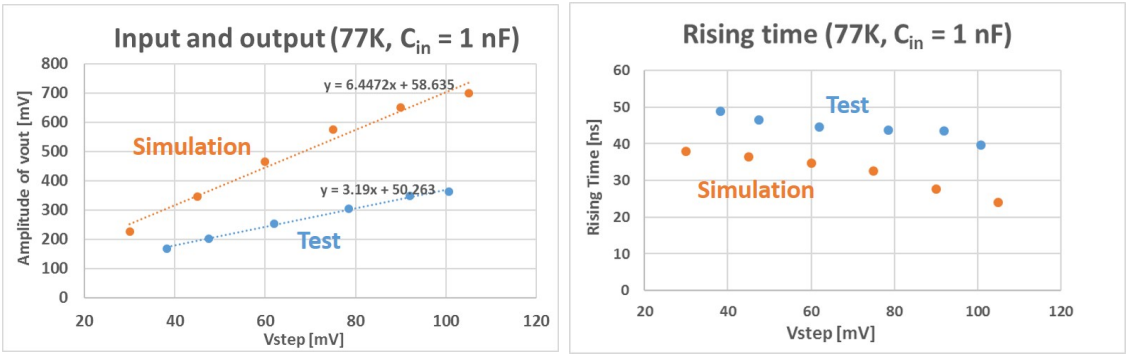
Table 3.8 summarizes some basic parameters. Compared with the simulation results, the gain of the circuit is reduced by about 40 %, and the value of the r.m.s. noise is basically the same, so it can be inferred that the signal-to-noise ratio of the circuit is reduced by about 40 %. Considering the similar trend of the rising time, the possible reason of the mismatch may be the variation of the input resistance of RCG, where the increasing of input impedance can reduce the bandwidth (enlarge the rising time, reduce the gain) and increase the noise. The possible sources of this variation could be the underestimate of the poly-silicon resistance at gate of common source transistor (the large size transistor in RCG) and also some parasitic resistance on board. Furthermore, the process variation of the constant resistor value, the load resistor of the output stage, may also contribute the gain mismatch.

Attributes	post layout simulation	test	mismatch
Gain [mV/pC]	405	255	-40 %
rmsNoise [mV]	1.55 mV	1.47 mV	-5.2 %
ENC [fC]	3.83	5.76	50 %
Rising time	47.6 ns	75.9 ns	59.5 %

Table 3.8: 300K, Test-Chip2 VFE, parameters comparison between the post layout simulation and test (DC-cathode VFE, input capacitance 1 nF, the rising time is measured with input step voltage of 85 mV)

VFE in 77 K

Figure 3.39 shows the input-output response of the chip and rise time at liquid nitrogen temperature.



(a) input and output in 77 K

(b) rising time in 77 K

Figure 3.39: Amplitude and rising time measurement in 77K

Table 3.9 summarizes the measurement results of the basic parameters.

It can be observed that, similar with the room temperature condition, the cryogenic test results shows a reduce of the gain and improvement of the rising time. Thanks to the reduction of also the noise, the signal to noise ratio is similar with the simulation.

Attributes	post layout simulation	test	mismatch
Gain[mV/pC]	645	319	-50.5 %
rmsNoise [mV]	2.0	1.1	-45 %
ENC [fC]	3.1	3.4	9.7 %
Rising time	32.5 ns	43.8 ns	34.8 %

Table 3.9: 77K, TestChip2 VFE, parameters comparison between the post layout simulation and test(DC-cathode VFE, input capacitance 1 nF, the rising time is measured with input step voltage of 85 mV)

Comparing the measurement results of TestChip2 VFE from room to liquid nitrogen temperature: The gain improves about 25%, the ENC decreases about 40%, and the rising time decreases about 40%. Therefore, it can be summarized that the very front end circuit has a much better performance working at cryogenic temperature than room temperature.

3.5 Summary

In this chapter, two test chips are designed and characterized to study two schemes used for large area SiPM readout and working at cryogenic temperature. This chapter presents also the SiPM electrical model and a review of the effects of cryogenic temperature on MOSFETs.

TestChip1 integrates a single-channel front-end structure that reads out 24 cm^2 SiPM, following the structure of original scheme of Darkside-20k. It contains four RCG input stages and each to read out the SiPM of 6 cm^2. A summing amplifier combine the current signals of the four branches and bring the one channel output that can drive the transmission cable.

The first version of TestChip1 shows an instability caused by the lack of phase margin of the second stage (CR-RC shaper). The second version of TestChip1 resolve this problem by changing the shaper to a voltage amplifier, and receives the expected experimental results as expected.

However, this version of the chip still can not work stably in liquid nitrogen with a large input capacitance. It can be inferred that the reason of this instability comes from the RCG structure since the oscillation is related to the input capacitance load. Considering that the corresponding results cannot be obtained from the simulation

and the simulation at the liquid nitrogen temperature is not necessarily accurate. Therefore, further study is required to reveal the specific reason.

TestChip2 integrates the very front end circuit (RCG + TIA) and the basic digital signal modules to verify the basic blocks of the second scheme, expecting to read out large area SiPM by the multi-channel ASIC, with the digitization on-chip and working at cryogenic temperature. The test results show that the front-end circuit can work properly at both room temperature and liquid nitrogen temperature. The signal to noise ratio and the bandwidth of the test structure are sufficiently improved from the room temperature to liquid nitrogen temperature: the ENC has a reduction of about 40% and the rising time has a reduction of about 40%.

The digital signal modules include signal synchronization module and LVDS transmitter. Experimental results show that they can work properly at room temperature and at liquid nitrogen temperature (The LVDS transmitter has a good performance working at the frequency up to 320MHz).

The IP blocks developed, produced and tested with these two prototypes, are used in the design described hereinafter. Based on the verification of TestChip2, chapter 5 covers the design of first prototype of mixed-signal ASIC for large area SiPM readout working at cryogenic temperature.

Chapter 4

Design of multi-channel front-end ASICs for capacitive sensors in 110 nm CMOS

This chapter describes the design of a versatile mixed-signal front-end ASIC for the readout of highly capacitive semiconductor and gaseous detectors. Designed in an area of 5 x 5 mm^2, this chip with 64 parallel channels features a full chain readout providing amplification, signal conditioning and discrimination, and a data payload containing the channel ID, the time stamp and charge information for each event. The programmable gain and input impedance of the front-end amplifier allows to match the requirements of different detectors.

The chip has been fabricated in standard 110 nm CMOS technology and operates with a core 1.2V power supply and 2.5V for the I/O. Table 4.1 shows the key features of the ASIC.

Parameters	Values
Number of channels	64
Events rate	$>$ 100 kHz per channel
INL	$<$ 1 %
Dynamic Range	up to 400 fC
Sensor capacitance	tens to hundreds of pF
Power Consumption	$<$10 mW/ch
Technology	UMC 110 nm CMOS

Table 4.1: Design parameters of the chip

This chapter is organised as follows: Section 2 gives an overview of the chip structure. Section 3 describes the architecture of the pre-amplifier, including mainly

the transfer function and noise analysis. Section 4 makes the description of the shapers and discriminators. Section 5 introduces the back-end part, including the TDC, S&H and ADC working principles. Section 6 reports the test results.

4.1 Overview of the ASIC architecture

The development of this chip was done in parallel with that of the TIGER ASIC developed for the readout of a Cylindrical Triple-GEM detector, in the framework of the BESIII Inner Tracker upgrade program[22][9]. The re-use of key IPs between the two ASICs, such as the Time-to-Digital Converters, the DACs and of most of the control logic shortened the design time, while the sharing of the same dedicated fabrication reticle allowed for a significant cost reduction. For a detailed review of the TIGER ASIC and associated on-detector electronics the reader is referred to [85].

Figure 4.1 shows the block diagram of one channel. The signal from the detector is firstly read out by a Regulated Common Gate (RCG) pre-amplifier, which works as a current conveyor and provides programmable gain and input impedance. The current output signal is then split into two branches: the timing branch consists of a fast shaping TIA (Transimpedance Amplifier) with a peaking time of about 60 ns used for accurate timing measurements, while the energy branch has a slower shaper with a peaking time of about 170 ns to minimise the equivalent noise charge (ENC).

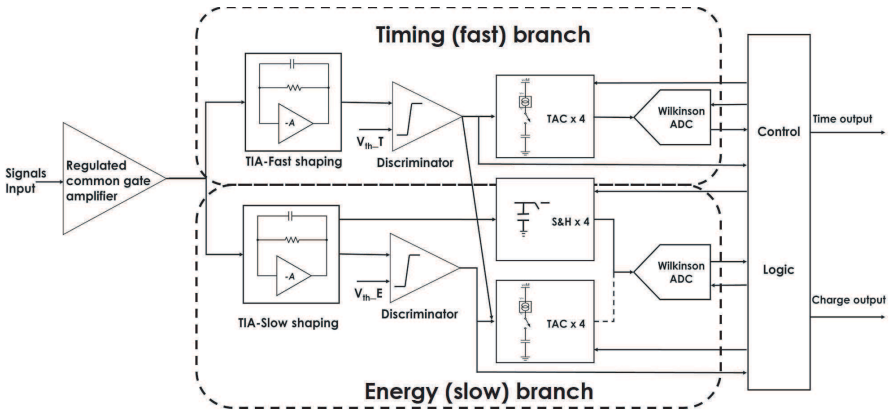


Figure 4.1: Architecture of one channel

The output signal of the timing branch amplifier is fed to a fast leading-edge threshold voltage-mode discriminator, generating a trigger signal which is used for the time-to-digital conversion of the crossing time. The low-power Time-to-Digital Converter (TDC), based on time interpolation, uses up to four Time to Analogue Converters (TACs) and one Wilkinson Analogue-to-Digital Converter (ADC).

The circuit allows for two different methods to be used for the charge measurement: Sampling and Hold (S&H) of the voltage signal at the output of the energy branch, or a time-based readout of the Time over Threshold (ToT). The S&H circuit samples and holds the peak voltage from the slow shaper output, which is then digitised by a Wikinson ADC. Alternatively, the ToT method is implemented using two TACs to record the time stamps of the rising edge and falling edge. This method, despite its intrinsic non-linearity when using CR-RC filters, is a versatile solution for the energy measurement in case the input charge exceeds the dynamic range of the S&H circuit. The trigger signal for the S&H circuit and the rising edge time-stamp for the ToT can be generated both by the leading-edge crossing of the fast or slow shapers. Similarly, the falling edge for the ToT measurement can be selected either using the fast or the slow signal branch.

Control logic in each channel handles the operation of the back-end digitisation circuitry. This digital core operates at 200 MHz and manages the TACs, TDC/ADCs and data/control interface with the chip global back-end.

4.2 Versatile Front-End Amplifier Design

Figure 4.2 depicts a simplified transistor-level schematic of the front-end amplifier. The design parameters of the pre-amplifier MOSFETs are listed in Table 4.2.

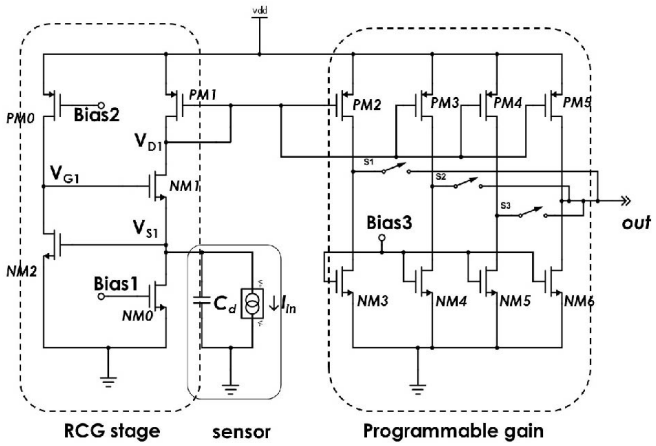


Figure 4.2: Transistor-level schematics of the pre-amplifier

The input stage is based on a common gate topology with g_m-boosting and works as a current conveyor. This regulated common gate amplifier topology allows for the realisation of a controllable very low input impedance front-end [86].

A programmable gain stage, shown in figure 4.2, is implemented with a configurable parallel connection of the output PMOS of the current mirror. The series

NMOS	Width[μm]	Length[μm]	PMOS	Width[μm]	Length[μm]
NM0	15	5	PM0	180	0.8
NM1	50	0.5	PM1	10	0.5
NM2	8000	0.6	PM2	16	0.5
NM3	16	2	PM3	8	0.5
NM4	8	2	PM4	4	0.5
NM5	4	2	PM5	4	0.5
NM6	4	2			

Table 4.2: Dimensions of front-end transistors

switches s_1, s_2 and s_3 allow for 8 programmable gain settings. The programmable gain stage (figure 4.2) is replicated for the fast and slow branches, and the current-mode output signal is fed to each one of the shapers. The control voltage $Bias3$ (configured by a 6-bit DAC) can be adjusted and allows for a fine setting of the output DC current of the gain stage, effectively controlling the amount of DC current sank from the shaper stage.

Each channel features a 6-bit DAC and a 5-bit DAC to set the currents in common gate ($Bias1$) and g_m-boosted stage ($Bias2$) respectively. This allows for a configuration range of the bias current in the order of 2.5 μA to 10 μA in the common gate stage, and 0.1 mA to 3.3 mA in the g_m-boosted stage. In the RCG circuit, the input transistor NM1 is in a common gate configuration, and NM2/PM0 implement the common source amplifier used to decrease the impedance seen at the source of NM1. Figure 4.3 shows a CAD layout detail of the front-end of a single channel, which highlights the large silicon area of the g_m-boosting transistor NM2.

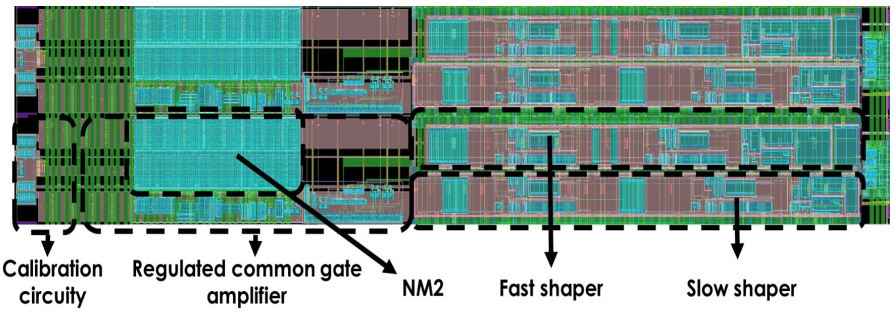


Figure 4.3: Layout of front-end

4.3 Shaper stages and discriminators

The simplified schematics of the two shapers are illustrated in figures 4.4 and 4.5.

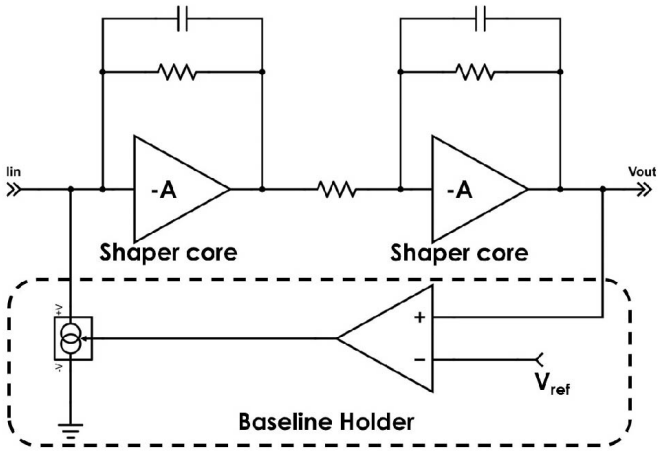


Figure 4.4: Schematic of Fast Shaper

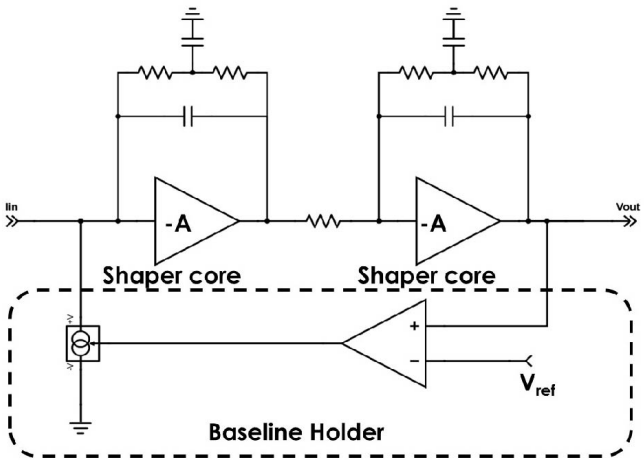


Figure 4.5: Schematic of Slow Shaper

The feedback resistors and capacitors are carefully designed in order to minimise the spread in key parameters, like peaking time and gain, due to statistical device mismatch and process variations. In the timing branch a conventional CR-RC shaper is employed. In the energy branch the feedback uses a pair of complex conjugate poles, whose impulse response has a better approximation to Gaussian

shape, thereby resulting in a lower noise (by increasing the peaking time) for a comparable rate capability.

The shaper cores share the same structure, using a single-ended input stage and a class-AB output [33][34]. The transistor level schematics are shown in figure 4.6.

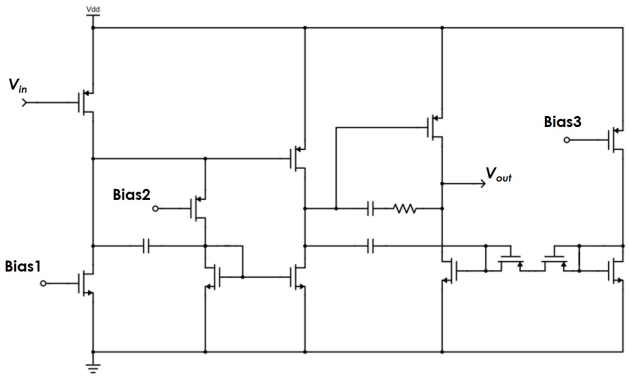


Figure 4.6: Transistor level design of shaper core

Both shapers employ a Baseline Holder (BLH) structure, whose working principle is mainly based on a very low frequency feedback [32], to set a defined output baseline. Figure 4.7 shows the transistor level design.

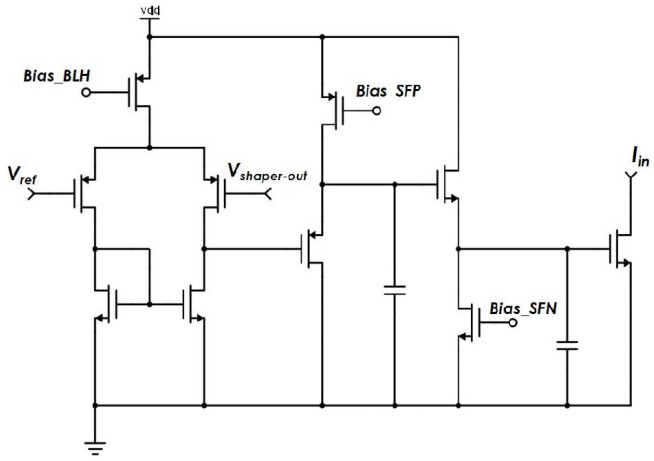


Figure 4.7: Transistor level design of baseline holder

The voltage outputs of both the fast and slow shapers are fed to input of the fast discriminators that generate CMOS-level trigger signals for the channel control

logic. The two discriminators share the same structure and the transistor-level schematic is shown in figure 4.8. The bias current of the differential input amplifier is controlled by V_{b1} and that of the output stage is set by V_{b2}. Both voltage bias are set by a 6-bit DAC at the periphery of the chip, and their value is thereby common to all 64 channels. The global setting V_{hyst} is configured by a 3-bit DAC for an adjustable hysteresis amplitude.

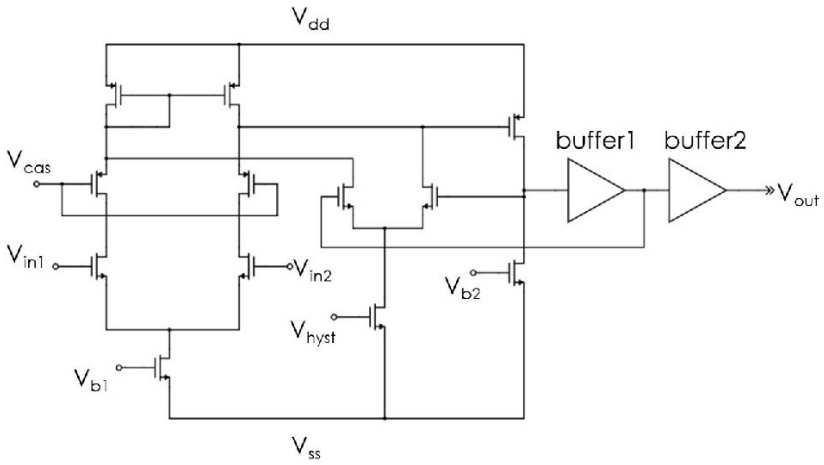


Figure 4.8: Schematic of the discriminator

4.4 Time and Amplitude Digitisation Circuits

The time measurement is performed by 2 low-power TDCs based on analogue interpolation [65]. A set of 4 Time-to-Amplitude Converters (TACs) per TDC are used, allowing for the de-randomisation of the coming events. For each event, the TAC generates and stores a voltage signal that is proportional to the time difference between the trigger and a known leading edge of the system clock. This voltage information is subsequently transferred into a second capacitor C_{TDC} and processed by the Wilkinson ADC, while the buffer is reset to an idle state. Any trigger occurring during the conversion time of the TDC will be processed by the next buffer in the queue, following a round-robin scheme for assignment. Any event occurring while all 4 buffers are occupied will be discarded.

The block diagram of the multi-buffered TDC is illustrated in figure 4.9 :

In the event of a trigger, the switch S_1 closes and the current source I_{TAC} (25 μA) discharges the C_{TAC} (0.5 pF) until the next clock cycle rising edge. A synchronous finite-state machine (FSM) closes the switch S_2, transferring the voltage stored on C_{TAC} into C_{TDC} (2 pF). In order to cope with the RC constant created by the R_{on} of the CMOS switch, this operation takes 20 clock cycles. The C_{TDC} capacitor

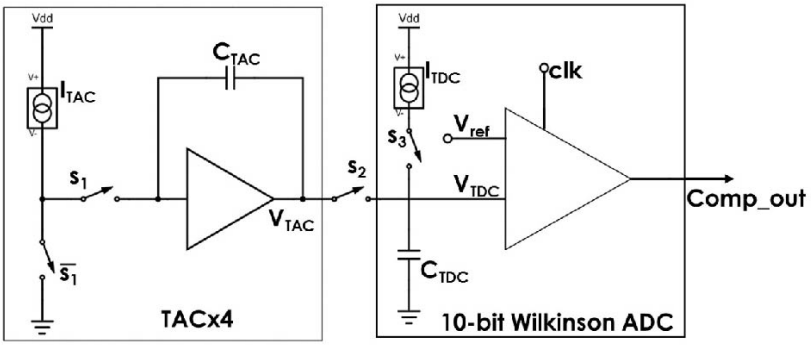


Figure 4.9: Multi-buffer TDC block diagram

is thereafter recharged with a smaller current I_{TDC} (0.78 μA) until V_{TDC} reaches the steady-state voltage V_{ref}, which is the working principle of the 10-bit Wilkinson ADC. Thereby, the time is interpolated by a factor of 128, considering the following design parameters:

$$32 \times I_{TDC} = I_{TAC}$$

$$C_{TDC} = 4 \times C_{TAC}$$

A system clock of 200 MHz provides a TDC time binning of 40 ps (LSB = 5 ns / 128). The fine counter (T-fine) information is convoluted with a 16-bit time stamp (T-coarse) provided by a global binary counter, which state is distributed to the channel, running at the chip clock frequency of 200 MHz: T-coarse and T-fine together provide the time stamp information. When the conversion is completed, the voltages on C_{TAC} and C_{TDC} are reset to the reference value (V_{ref}) by the control logic.

The conversion time defines the event rate that the TDC can handle, and is therefore a function of both operation clock and the interpolation factor. The design specification of a TDC capable of providing a time binning of 40 ps is driven by the fact that we expect, based on simulation results, the intrinsic time resolution of the front-end to be better than 500 ps and 300 ps r.m.s. for an input charge of 50 fC and considering, respectively 100 pF and 10 pF input capacitance. In these conditions, we expect the quantisation error of the TDC, which adds quadratically to the full channel intrinsic time resolution, to have a negligible contribution.

Two different charge measurement modes are implemented in the chip: ToT (Time-over-Threshold) and S&H (Sample and Hold) mode. In ToT mode both the rising and falling edges of the discriminator are digitised by the TDCs and

the charge information can be extracted from the pulse duration. The ToT measurement can be performed on the output of either the Timing branch or Energy branch.

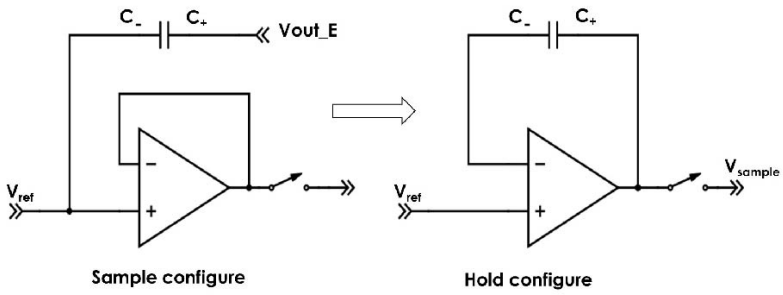


Figure 4.10: The principle of S&H mode

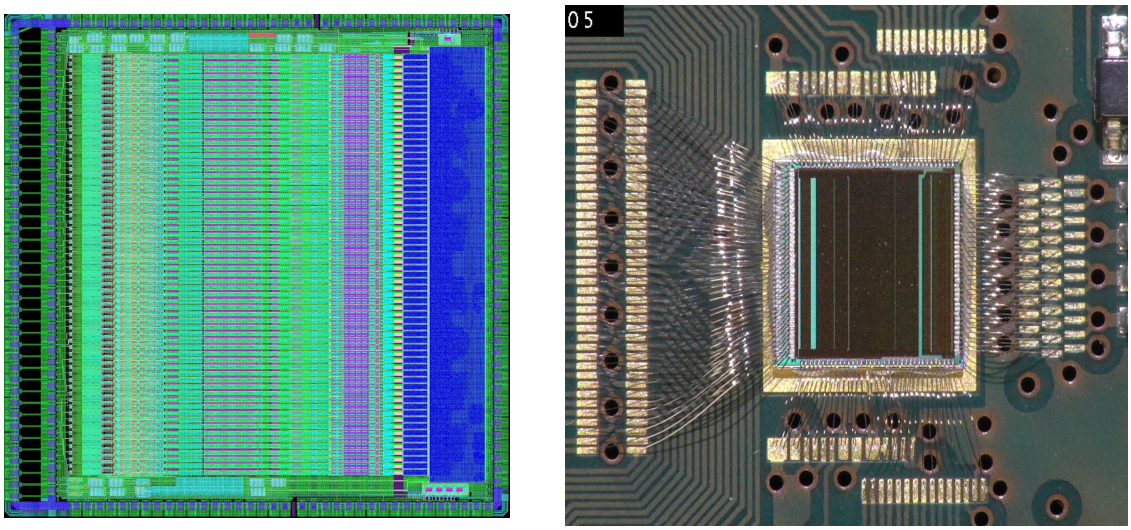
Figure 4.10 shows the basic process of S&H mode for charge measurement. The S&H circuit records and holds the peak voltage of the signal from the slow shaper on a capacitor. The configurable sampling time window is managed by the channel control logic, with the start provided by the discriminator of the fast branch (due to the smaller time walk). The voltage stored on the capacitor is then digitised by the Wilkinson ADC of the energy branch which is shared with the TACs, providing a linear measurement of the input charge. Similarly to the method adopted by the TDC, each branch employs four S&H buffers allowing for events de-randomisation.

4.5 Characterisation results

Figure 4.11 shows the CAD ASIC layout and the silicon chip wire-bonded to a test board. It is configured, controlled and readout using a commercial FPGA board with standard LVDS links.

A test pulse can be applied using on-chip calibration circuit or injected with an external pulse generator and a C-R circuit. The internal test pulse circuitry is implemented in the chip periphery and uses either a trigger signal generated by the global control logic or a digital test pulse fed directly from an external trigger generator. The circuit generates a voltage step function which amplitude is configurable using a 6-bit DAC. The voltage pulse is propagated to the channel under test, and a current-mode signal is generated locally by each channel enabled for calibration.

Figure 4.12 shows the measured peak amplitude at the output of the fast shaper for several gain settings ranging from 10 mV/fC ("set1") down to the minimum 1.2 mV/fC ("set8"). A test point at fast shaper output is built on chip and an on-board buffer is applied to bring the signal outside. The amplitude of the fast shaper output



(a) Layout of the chip

(b) Silicon chip on test board

Figure 4.11: Layout and silicon chip on test board

is characterized by the oscilloscope and the gain is calculated by the amplitude of output amplitude to input charge ratio.

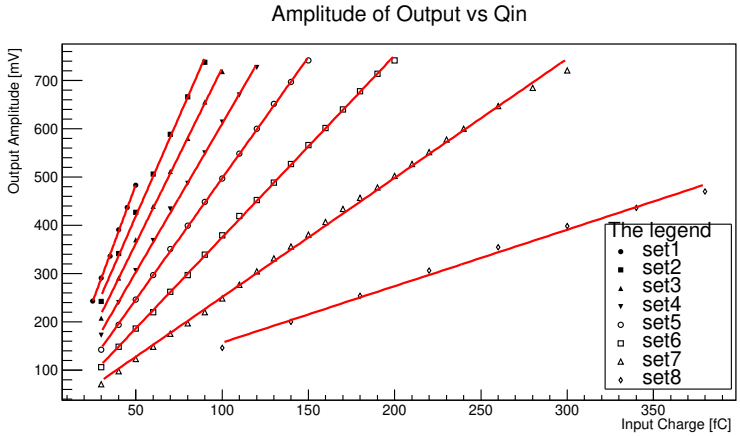


Figure 4.12: Adjustable Gains Characterisation

Table 4.3 shows a summary of the gain measurement characterisation results. The mismatch of experimental vs. simulated results at the minimum gain settings, which correspond to the worst Integral Non-Linearity (INL) results, is caused by a dynamic modulation of the V_{sg} of the PMOS devices on the amplifier output stage, which drives these MOSFETs into linear region. Although in this first prototype, this non-linearity could be corrected offline after calibration, a design fix will be

required in the final version of the chip. A cascoded topology would increase the output resistance of the current mirror, enhancing the linearity of the circuit.

Gain Set	Gain Test [mV/fC]	Gain Simulated [mV/fC]	Gain Mismatch	INL Test	Max Q_{in} [fC]
1	9.67	9.89	2.22 %	0.64 %	50
2	7.95	8.71	8.73 %	0.69 %	90
3	7.13	7.52	5.19 %	0.66 %	100
4	6.08	6.31	3.65 %	0.71 %	120
5	5.01	5.10	1.76 %	0.69 %	150
6	3.77	3.86	2.33 %	1.3 %	200
7	2.50	2.60	3.85 %	2.15 %	280
8	1.17	1.33	12.03 %	2.78 %	380

Table 4.3: Gains test of timing branch

The noise measurement is performed by scanning the amplitude of a fixed charge test pulse with a variable discriminator threshold level V_{th}. The curve consisting of triggered counts can thereafter be analysed by fitting with an S-curve, which the slope is a direct measurement of the noise. Figure 4.13 shows the characterisation result of the noise as a function of the input capacitance in the timing branch, compared with the post-layout simulation. In these tests, the capacitive load at the input was forced with an external capacitor on the test board.

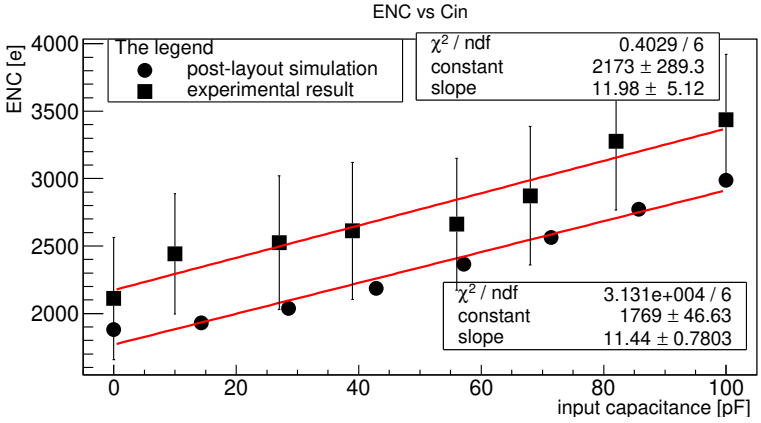


Figure 4.13: Timing branch noise versus input capacitance

The measured noise is higher than expected by a factor of 20%. This excess might be due to interference noise from the test environment and power supply.

In order to study the channel intrinsic time resolution, a sequence of test pulses synchronised to the FPGA system clock are used for the injection of a calibrated charge to the front-end. The σ of the Gaussian-Fit of the measured time distribution is a direct measurement of the jitter. Figure 4.14 plots the measurement and simulation results of the timing jitter as a function of the input capacitance, in the condition of an injected charge of 14 fC and gain setting of 10 mV/fC. The test was repeated by scanning the phase of the trigger signal in respect to the clock in steps of 135 ps (38 points on a 5 ns period).

The simulation values are obtained using the following function:

$$\sigma_t = \frac{rmsNoise}{Slope} \tag{4.1}$$

where $Slope$ is the slew-rate of the leading edge of the timing shaper output at a fixed threshold, and $rmsNoise$ is the total output r.m.s. noise voltage, both obtained with the simulation of a post-layout netlist. The systematic mismatch of simulation versus experimental results is not fully understood and further investigation is needed. The excess of noise in test data is independent of the input capacitance, but we were not able to replicate the same conditions simulating the post-layout simulations. Thereby, this systematic offset of 440 e^- r.m.s. could be related to digital interference noise at the level of the discriminator circuit.

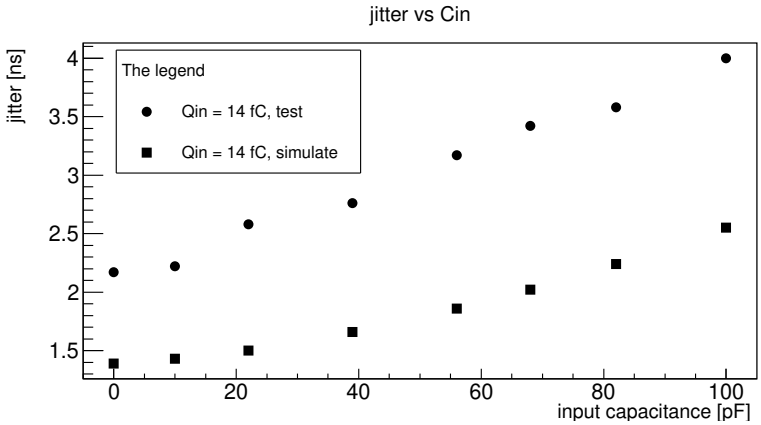
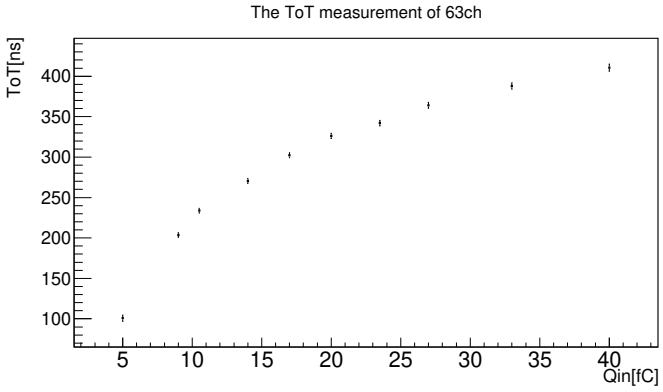
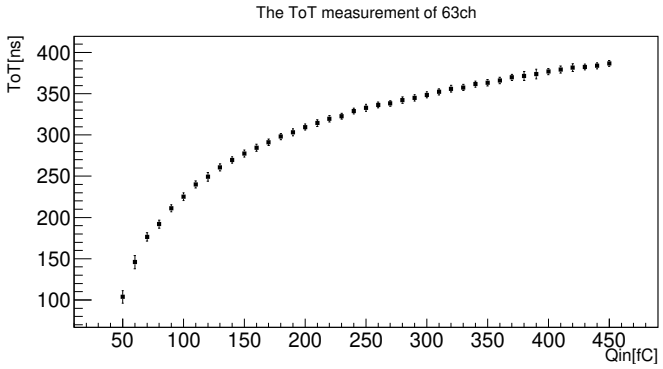


Figure 4.14: Timing resolution test in T-branch

The charge measurement, as aforementioned, can be performed using ToT (Time-over-Threshold) or S&H (Sample and Hold) circuit. The electrical characterisation is performed injecting a test pulse with different charges. Figures 4.15 and 4.16 show the results of a charge measurement with ToT and S&H modes, and both in the gain "set1" and "set8", respectively. The digitised output of the S&H is converted to the analogue peak voltage according to the calibration.



(a) gain 10 mV/fC



(b) gain 1.5 mV/fC

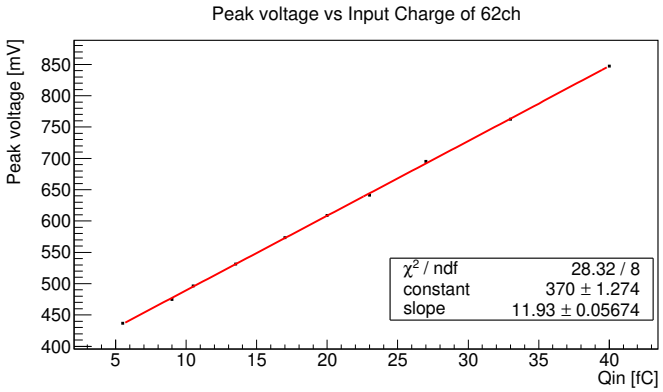
Figure 4.15: Input charge measurement in ToT mode

Experimental data obtained in S&H method show a INL better than 1% in both the minimum and maximum gain settings. Nevertheless, the non-linearity in minimum gain conditions worsens above 300 fC, for reasons that were already discussed earlier in this Section. The residuals of the linear fit of the characterisation data is shown in figure 4.17. For the ToT mode, the results are shown in figure 4.15. The non-linear ToT versus Qin behaviour of the front end requires a 3rd order polynomial fit or an offline Look-Up Table for the charge reconstruction. Despite the advantage in terms of higher dynamic range of the ToT method, a linear fit makes the S&H mode more advantageous, since it does not require an offline Look-Up Table.

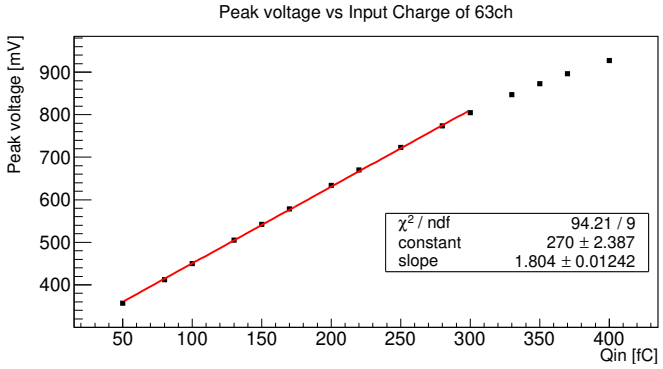
Table 4.4 provides a brief summary of the test results.

4.6 Conclusions and outlook

This chapter depicts the circuit-level design and electrical test results of a versatile 64-channel mixed-signal ASIC developed for the readout of high-capacitance



(a) gain 12 mV/fC



(b) gain 1.8 mV/fC

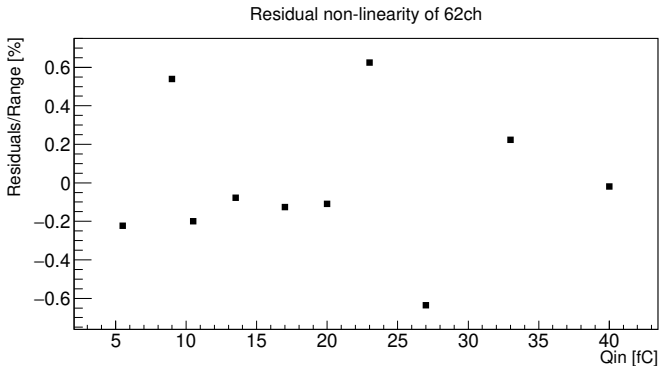
Figure 4.16: Input charge measurement in S&H mode

Attributes	Test Results
Power Consumption	9 mW/ch
INL	< 1 % (up to 300 fC)
Dynamic Range	up to 400 fC
Gain	1.8 to 12 mV/fC (E-branch)
ENC @ 100 pF	$3500e^-$
Jitter @ $Q_{in} = 14$ fC, $C_{in} = 100$pF	4 ns

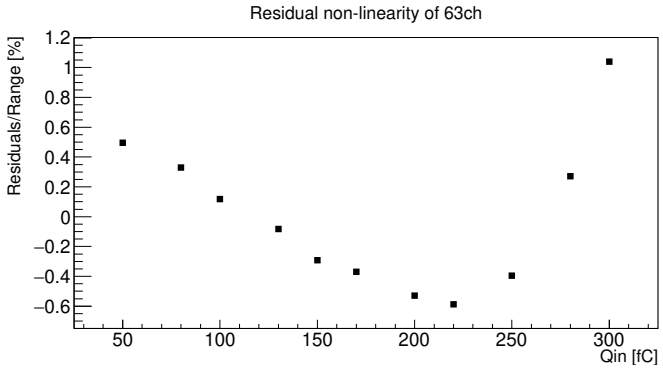
Table 4.4: Summary of electrical test results

sensors, providing the time stamp and charge measurement of each event. This chip was produced in a 110 nm CMOS technology engineering run, sharing the reticle with the TIGER ASIC[85], which was developed for the readout of the CGEM Inner Tracker detector for the BESIII Upgrade.

The intrinsic time resolution is better than 4 ns r.m.s. for an input charge of 14 fC with a 100 pF detector capacitance. The charge measurement is performed



(a) gain 12 mV/fC



(b) gain 1.8 mV/fC

Figure 4.17: Residuals in S&H linear fit

by two alternative modes : ToT and *S&H* mode. The dynamic range up to 400 fC allows for the use of this ASIC in a wide number of semiconductor and gaseous detectors, while the low-impedance front-end maximises the PSRR and reduces the susceptibility to external interference noise. A possible application of the ASIC for the readout of a Cherenkov detector using a SiPM detector plane is currently under study.

Chapter 5

A mixed-signal ASIC for SiPM readout

5.1 Overview the full chip

In this chapter, ALCOR (A Low Power Chip for Optical sensors Readout), the prototype of a mixed-signal ASIC for SiPM detectors, is described. Based on a time-based readout architecture, this chip inherits the studies mentioned above and is expected to operate both at room and cryogenic temperature (down to the boiling temperature of Liquid Nitrogen 77 K).

This chip core consists of 32 parallel pixels arranged in a matrix of 4 × 8, as shown in figure 5.1.

The pixel architecture is shown in figure 5.2, each pixel has a dual-polarity very front end based on RCG structure to readout either the anode or cathode signal. The amplified signal is fed into two independent branches both followed by a leading-edge discriminator. A set of four TDCs, based on analogue interpolation, are employed in each pixel to increase the count rate ability and de-randomize the the incoming photons (Possion distribution).

The data payload generated for each event is collected by an End of Column circuit, and the full digitized data is finally transmitted off-chip using 4 LVDS transmitters.

5.2 Very front end

5.2.1 Very front end design

The very front end (VFE) shown in figure 5.3 is implemented by the complementary RCGs both for anode or cathode signal readout. Then, the signal selected either from the anode or cathode RCG is divided into two independent branches,

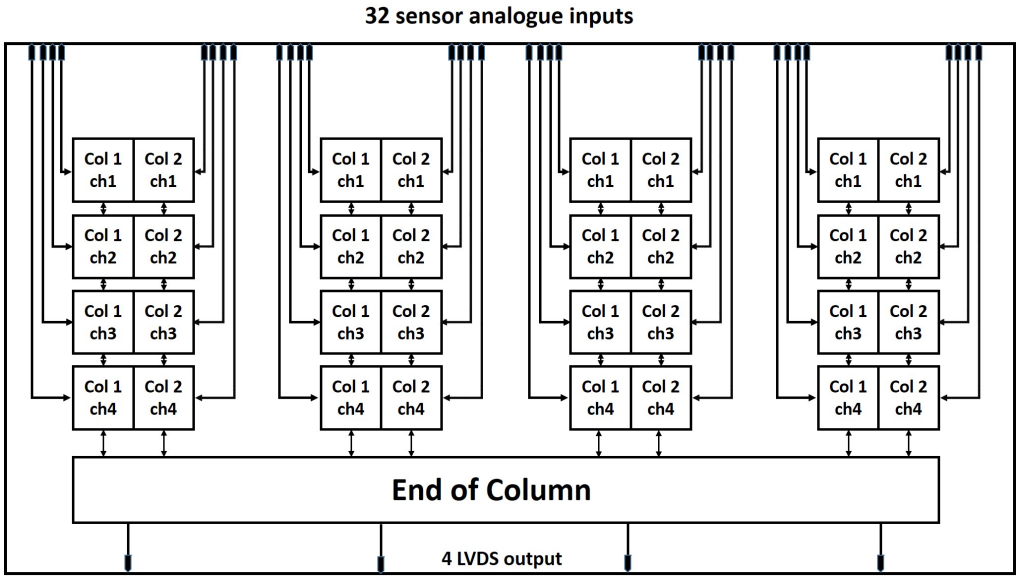


Figure 5.1: Architecture of full chip

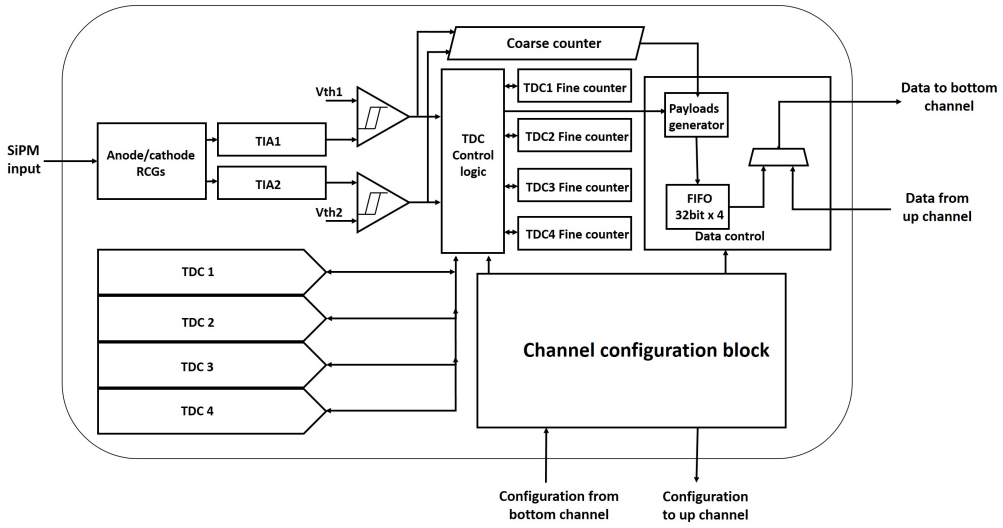


Figure 5.2: Architecture of each readout pixel

and a common source amplifier is adopted to transfer the current to voltage signal (TIA).

Figure 5.4 shows the transistor level design of the two RCGs, and table 5.1 details the dimensions of transistors. Bias circuits are implemented in the analogue bias section of the End of Column block (figure 5.1). The schematics of the bias cells for the VFE are shown in figure 5.5, which details also the configuration bits for the setting of the common gate bias (Bias_cg1 and Bias_cg2) and gm-boost

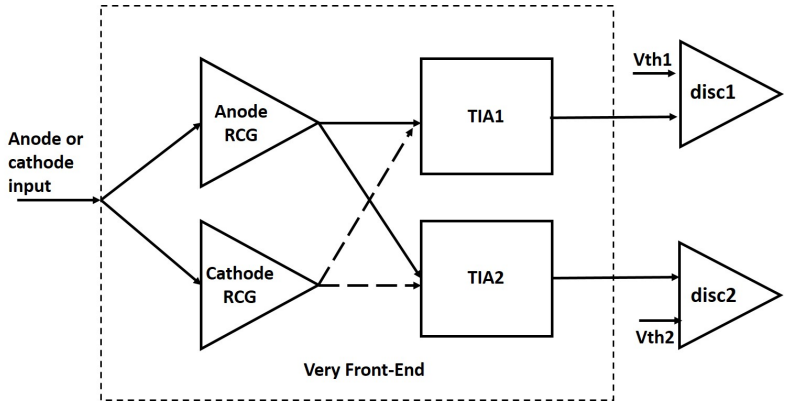
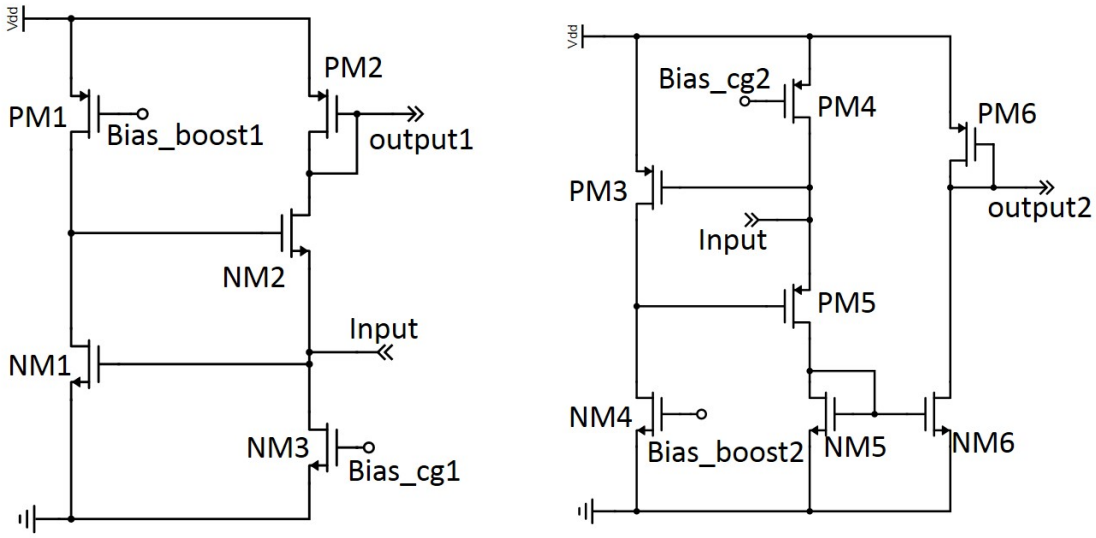


Figure 5.3: Architecture of very front end



(a) Transistor level design of anode RCG

(b) Transistor level design of cathode RCG

Figure 5.4: Schematic of complementary VFE

bias (Bias_boost1 and Bias_boost2) currents.

When the anode VFE is enabled, the bias currents of cathode VFE will be switched off to save power and make it possible to sum the two inputs in the same node, considering that the parasitic capacitance (in the order of 10 pF) is much less than the sensor capacitance (in the order of nF). The nominal values for the common gate current and boost stage currents is 50 uA and 3 mA, respectively. The bias 5-bit DACs allow us to trim those currents in the range 25 uA - 100 uA and 1.5 mA - 5 mA.

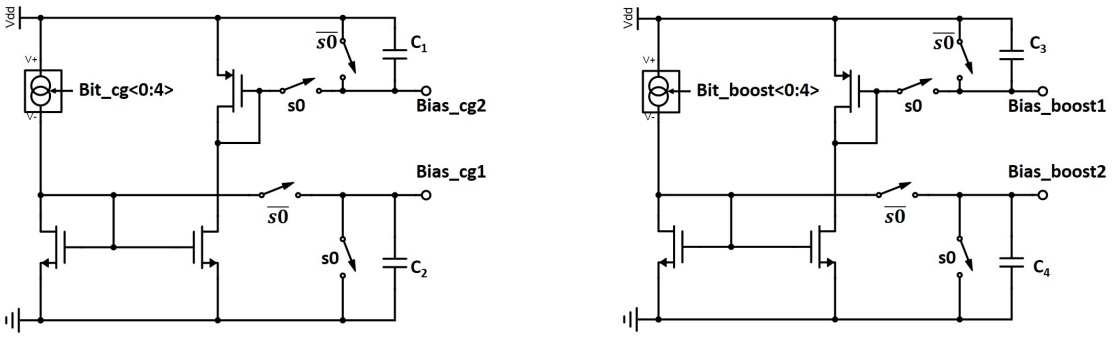
One bit $s0$ is used to select anode or cathode readout mode, which is also adopted to select the signal of anode VFE or cathode VFE to be fed to the output

NMOS	Width[μm]	Length[μm]	PMOS	Width[μm]	Length[μm]
NM1	3200	0.5	PM1	200	0.4
NM2	500	0.15	PM2	18	0.4
NM3	100	2	PM3	4000	0.25
NM4	8	2	PM4	40	1
NM5	12	1	PM5	800	0.12
NM6	12	1	PM6	18	0.4

Table 5.1: Dimensions of ALCOR very front-end transistors

stages (TIAs).

Decoupling capacitors (C_1 to C_4) are implemented by NMOSCAPs to filter the noise from bias circuits to VFEs. It should be mentioned that both the common gate and gm-boost current are increased in respect to the design values of the VFEs prototypes in TestChip2, discussed in Chapter 3. The current sources, in particular the gm-boost current source (PM1 and NM4 in figure 5.4), have a current of several mA, thus the PM1 and NM4 are quite sensitive to the coupled noise from bias circuits. Besides, the switch $s0$ and $\bar{s0}$.are implemented by NMOS and PMOS, which have a series resistance in hundreds of Ohms and can couple dominant thermal noise to the VFEs. Consequently, a large area NCAPs (nF order) are used for decoupling gm-boost bias (C_3 and C_4).



(a) Bias circuit of common gate current

(b) Bias circuit of gm-boost current

Figure 5.5: ALCOR RCGs Bias circuits integrated in the End of Column part of the chip

The two independent output stages share the same transimpedance configuration, and the transistor level design is seen in figure 5.6. Each branch has 2-bit ($s1$ and $s2$) to configure the gain independently. Thus four gain settings are available (LSB, 4LSB, 7 LSB and 10 LSB). The DC-coupling scheme is selected between the RCGs and output stages.

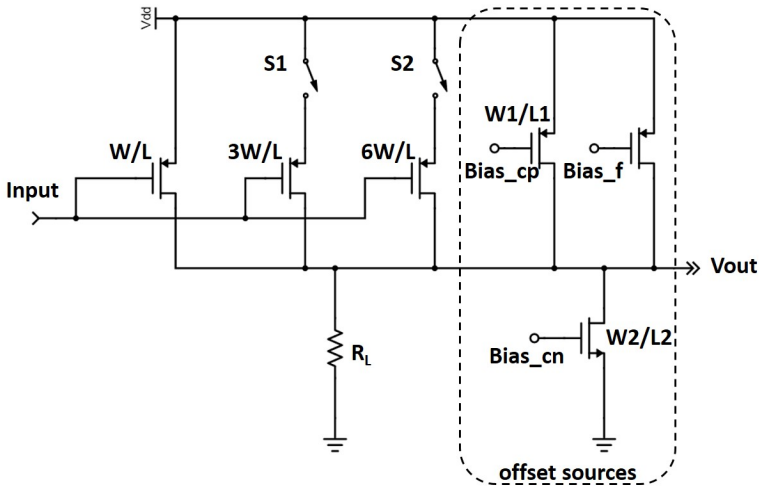
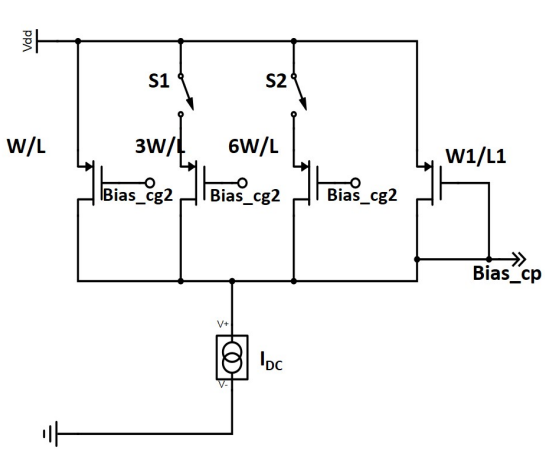
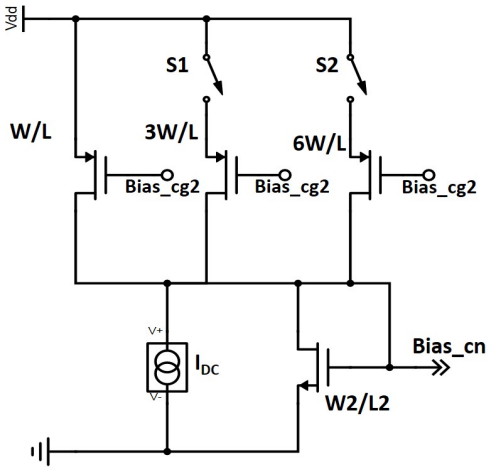


Figure 5.6: Programmable gain in output stage of ALCOR VFE

When $s1$ or $s2$ switches on or off, the DC current which flows through the constant resistor ($R_L = 3.5$ kΩ) will also be modified to change the baseline of the output. Therefore, several offset sources are used to keep the baseline a constant value. In total, three current sources are injected to the output node: the coarse PMOS current (Bias_cp), coarse NMOS current (Bias_cn) and fine PMOS current (Bias_f).



(a) Bias of coarse offset PMOS current



(b) Bias of coarse offset NMOS current

Figure 5.7: Bias circuits of coarse offset to keep baseline constant

The dedicated bias circuits of coarse current sources (shown in figure 5.7) are used to create specific offset currents cope with the gain and common gate current

setting. In order to have a constant baseline, the current I_{DC} which flows through the load resistor (R_L) should be a constant. The output stage is DC-coupled with the common gate stage of RCG, thus the current of output stage will follow the variation of the common gate current. Besides, the configuration of the output stage gain will also modify the current flowing to the constant resistor. Therefore, the bias setting in figures 5.7a and 5.7b duplicate the current value configured by the two settings, and the compensate currents are generated to keep a constant current in the resistor (R_L in figure 5.6). Assuming the expected baseline as 500 mV, the dedicated current value is obtained:

$$I_{DC} = 500mV/3.5k\Omega = 143\mu\text{A} \tag{5.1}$$

This is the current value of I_{DC} in figures 5.7a and 5.7b. The PMOS and NMOS coarse current sources (configured by) are set to be off in nominal configuration (I_{cg} = 50 μA and both $s1$ and $s2$ close). When the current of output transistors is smaller than the nominal value, the PMOS coarse current (configured by Bias_cp) will inject compensate current to R_L; In the opposite case, the NMOS coarse current (configured by Bias_cn) will reduce the current flowing through R_L.

The third offset current (configured by Bias_f) with 3-bit and LSB = 6 uA, which is set to cope with the variation in the process (especially the absolute value of R_L). The fine offset current is configured by pixel logic and has the maximum capability to adjust the baseline in 200 mV ($8 \times 6\mu\text{A} \times 3.5\text{k}\Omega$).

All the current bias signals are generated from the conventional current reference generation circuits shown in figure 5.8, which is located in both pixel bias cells and EoC part.

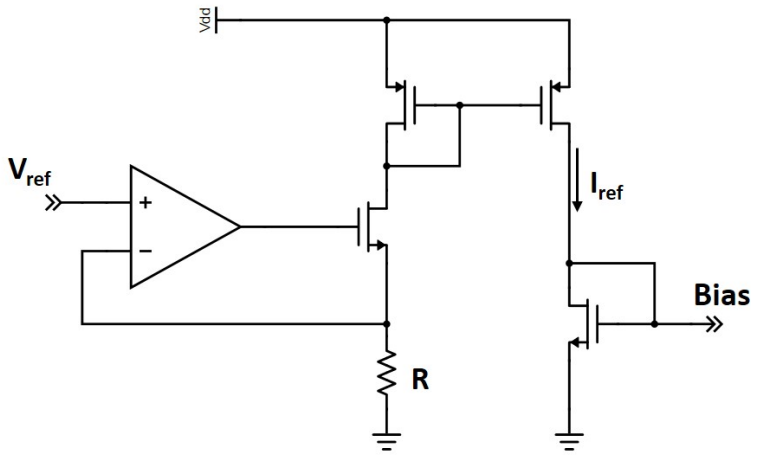


Figure 5.8: Reference current generation circuit and bias distribution

The reference voltage is provided off-chip by a commercial regulator, and the relation can be easily obtained:

$$I_{ref} = \frac{Vref}{R} \tag{5.2}$$

The "golden reference" current is transferred to the voltage by the diode-connected transistor and distributed in voltage mode to the local analogue circuitry.

Table 5.2 summaries the configure information in the VFE, where the End of Column (EoC) is the chip level configuration.

gurat

Name	description	# bit	configure level	nominal
Bias_cg	common gate current	5	EoC	50 μA
Bias_boost	g_m-boost current	5	EoC	3 mA
Bias_f	fine offset to adjust baseline	3 x 2	Pixel	6 μA
s0	"1": anode VFE; "0", cathode	1	EoC	1
s1	LSB bit of gain setting	1	Pixel	0
s2	MSB bit of gain setting	1	Pixel	0

Table 5.2: Summary of configuration in ALCOR VFE

The very front end has two independent output stages, also their gains and the threshold voltages of the corresponding discriminators. The very front end circuit is able to work at dual thresholds setting, as shown in figure 5.9, the four TDCs are controlled to achieve the expected conversions.

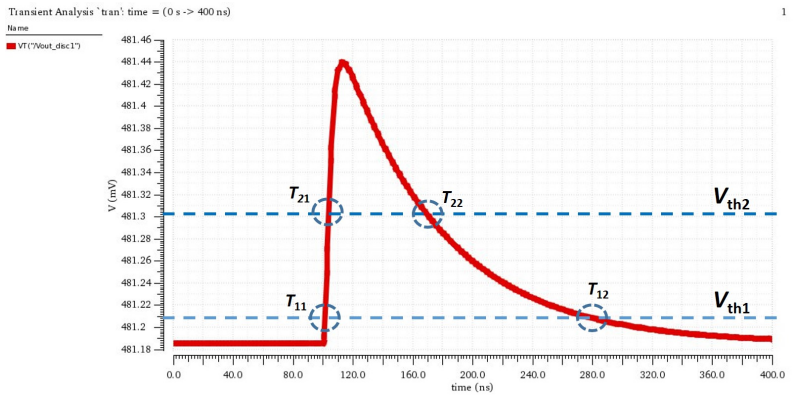
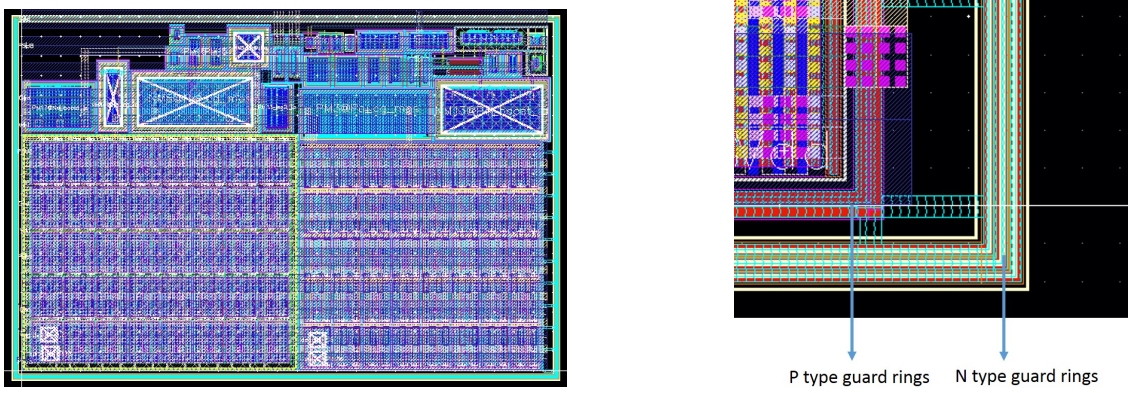


Figure 5.9: Working principle of dual threshold voltages

A lower threshold voltage (V_{th1}, usually less than 1 pe amplitude) can be set to obtain a more accurate leading edge timing measurement (T_{11}), because of a bigger signal slope. The energy information extracted from Time over Threshold (ToT) measurement can be used to calibrate the time walk. The ToT resolution is dominant by the falling edge measurement which has a smaller slope (like T_{12}).

The falling edge time (T_{22}) can be measured with a higher threshold voltage (V_{th2}), where the falling slope is better than the previous one. Furthermore, the slew rate of the signal can be obtained by the two leading edge time measurement (T_{11} and T_{21}) of dual thresholds. which can be used to characterize the bandwidth of the very front end circuit and the sensor capacitance.

Figure 5.10a shows the layout view of the very front end circuit, with a final size of $120\times90\mu m^2$.



(a) Layout view of very front end circuit, $120\times90\mu m^2$

(b) Multiple guard rings to suppress the cross talk between different blocks

Figure 5.10: Layout design of very front end

Considering the cross talk between different blocks through the substrate, especially from digital block to analogue block, the following techniques are adopted to suppress this problem:

1. The conventional NMOS transistors share the same substrate, thus easily to be influenced by the interference from the substrate. In this work, NMOS transistors with independent substrate are used, thanks to the UMC CMOS technology. A local N well is doped firstly in the PSUB, then the P well in created inside the N well that is used for the substrate of NMOS.

2. The digital and analogue parts are separated to keep a long distance in the layout.

3. More guarding rings are added around the key transistors. Multiple guard rings are added also around each modules to suppress the substrate interference from the others (shown in figure 5.10b).

5.2.2 Very front end simulation

Input impedance and bandwidth

The input impedance of very front end is simulated by applying a ac stimulate, as shown in figure 5.11. Both anode and cathode readout RCG has a low input impedance, 12Ω and 18Ω respectively. As discussed before, low input impedance makes the circuit read out the sensor signal in a short time and reduce the voltage variation of the input node. Besides, the lower impedance is observed in the anode RCG, where the NMOS common source transistor has a larger transconductance than the PMOS with the same power consumption.

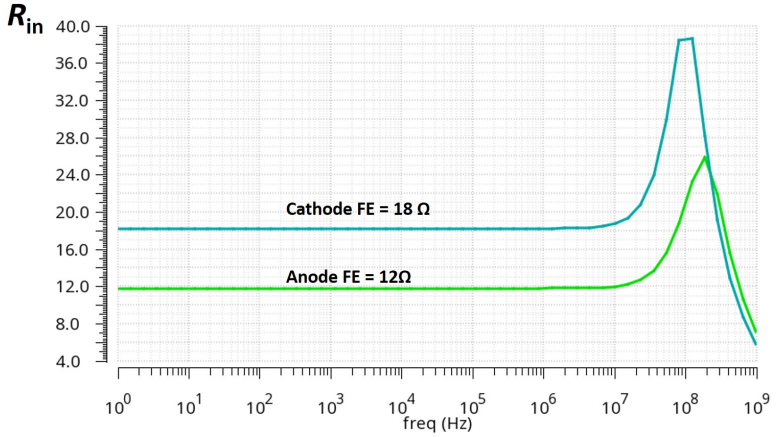


Figure 5.11: ALCOR input impedance simulation ($C_{\text{in}} = 0$)

Analogue bandwidth is obtain by measuring the rising time of output, when a $\delta(t)$ like signal is fed to the input. The bandwidth is calculated by equation 5.3.

$$F_{\text{3dB}} \approx \frac{0.338}{t_{\text{r}}} \tag{5.3}$$

Where t_{r} is the rising time (10%-90%) of the output. As shown in figure 5.12, with the detector capacitance (C_{in}) of 0 (5 nF), the rising time is 1.46 ns (5.85 ns), and the bandwidth is 226 MHz (58MHz).

Linear dynamic range

The very front end circuit consists of the RCGs and TIAs parts, where the TIA is implemented by a basic common source amplifier with a constant resistor load. The limitation of the linear response range is from the TIA part, when the voltage variation drives the output transistor out of saturated region.

It should be mention that, the linearity itself doesn't affect the performance of the very front end, because ALCOR is a time information measurement based

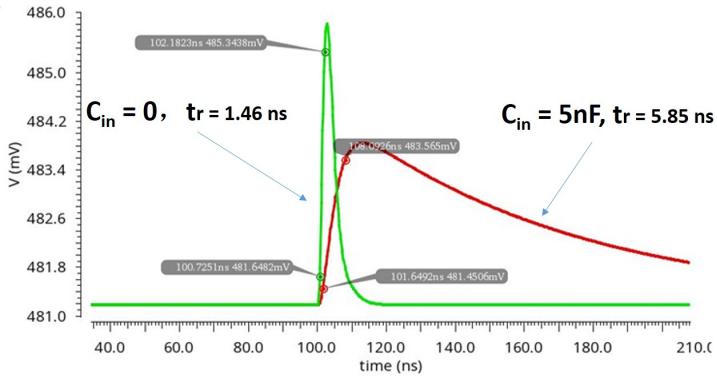


Figure 5.12: Bandwidth simulation of very front end circuit

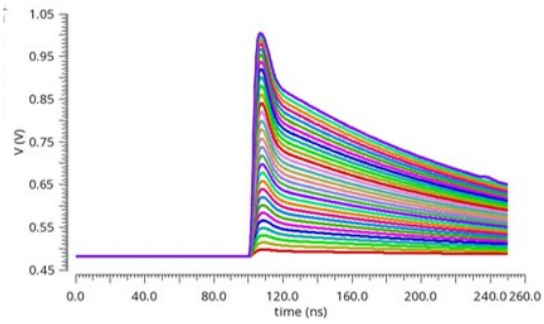
chip. Besides, the basic common source amplifier with constant resistor load has the intrinsic non-linearity, because the transconductance of common source transistor is not a constant value during the amplification.

While a much worse linearity can be observed if some transistor is driven out of the saturated region, which can make a mismatch between the transient circuit performance and the ordinary simulation and analysis (based on small signal). Considering the time stamp is measured at the trigger point of a given threshold, the linear range has to fulfil the setting range of threshold voltage.

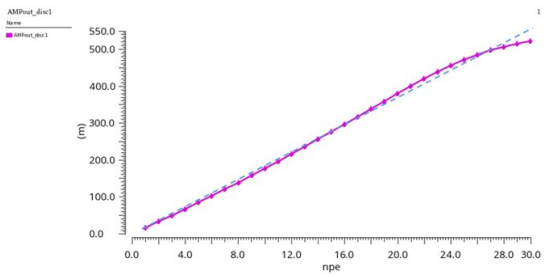
Figure 5.13 shows the simulation results of the maximum gain setting (18 mV/pe). The output of very front end shows a worse non-linearity when the input signal is larger than about 27 photons (INL $>$ 3%). The peak voltage of the output is about 980 mV at the 27 photons input, which indicates the linearity is limited by the voltage swing bedroom of the output transistor (VDD = 1.2 V). The maximum gain setting is usually used to have the more accurate leading edge time measurement (a lower threshold), and linear dynamic range up to 27 photon is enough to choose a optimal threshold voltage.

Figure 5.14 gives the simulation results with the minimum gain setting (2.7 mV/pe). The linear range is up to 61 photons (INL $>$ 3%), which is also enough for a higher threshold voltage setting. The peak voltage of output is only about 610 mV at the input of 61 photons. Because the larger signal gives the output of RCG a larger voltage swing, this is, the gate voltage of common source transistor. In the same time, the drain node of the common source transistor has a opposite variation. The two variations drive the transistor out of saturated region at a smaller peak voltage.

The different gain settings make the dual threshold working scheme more flexible.

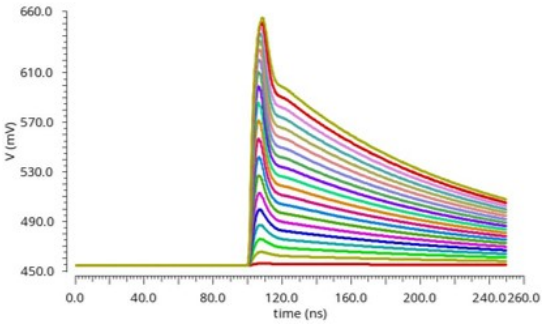


(a) Output waveforms and the input number of photons

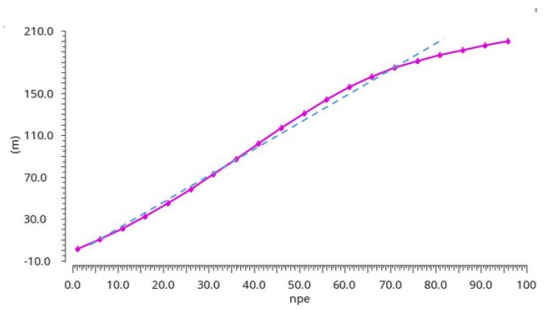


(b) Output amplitude and the input number of photons

Figure 5.13: The linear dynamic range simulation with maximum gain setting (300K)



(a) Output waveforms and the input number of photons



(b) Output amplitude and the input number of photons

Figure 5.14: The linear dynamic range simulation with minimum gain setting (300K)

Output baseline simulation

The output baseline of very front end is important for the working condition itself and the discriminator, and the threshold voltage of the discriminator is programmable from 550 mV to 875 mV.

The baseline of very front end is related to, as discussed above, the common gate current, gain settings, and the process variations. Figure 5.15 shows the waveforms at all possible common gate current (5-bit, 32 conditions) and the gain (2-bit, 4 conditions) settings ($32 \times 4 = 128$ conditions). The simulation results show a baseline variation of less than 30 mV, which proves that the baseline compensate circuit of output stage works well.

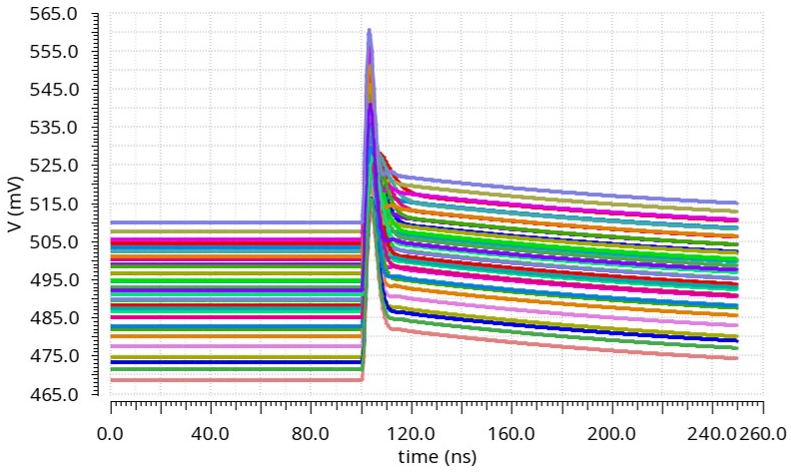


Figure 5.15: Output waveforms of the different common gate current and gain settings

The corner simulation is performed to evaluate the affect from the process parameters variation. The simulation results are shown in figure 5.16a and summarized in table 5.3. The corner "fnsp-f" means: NMOS (fast), PMOS (slow), Resistor, (fast). Others can be seen by analogy. Considering a 3-bit fine current is used to cope with this variation and the maximum voltage trim ability is of about 200 mV, the baseline change caused by the process parameters variation can be compensated by the configuration (pixel level).

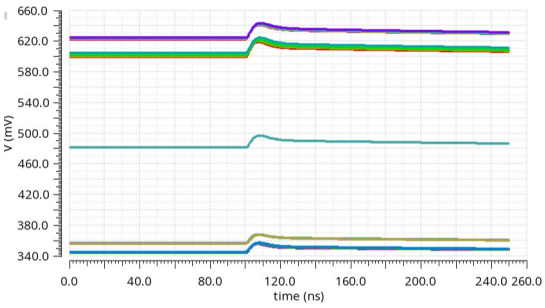
Table 5.3: Corner simulation of the baseline

Corner label	Num	Corner type	V_{base}
Nominal	1	nominal (27°C)	480 mV
C00-C03	4	fnsp-f (-40, -20, 0, 27°C)	356 mV
C10-C13	4	fnsp-f (-40, -20, 0, 27°C)	623 mV
C20-C23	4	fnsp-f (-40, -20, 0, 27°C)	345 mV
C30-C33	4	fnsp-f (-40, -20, 0, 27°C)	600 mV

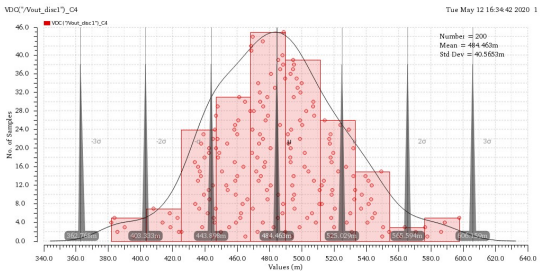
Figure 5.16b shows the Monte Carlo simulation results, the standard deviation of the baseline voltage is about 40 mV, which can be easily compensated by the configuration.

Time jitter and SNR simulation

Different SiPM products have different parameters, which can also make the influence on the performance of the front end circuit. In general, the most important



(a) Output waveforms of corner simulation

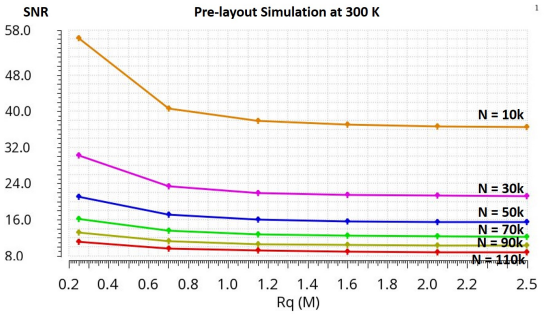


(b) Monte Carlo simulation of baseline voltage, σ =40 mV (200 runs)

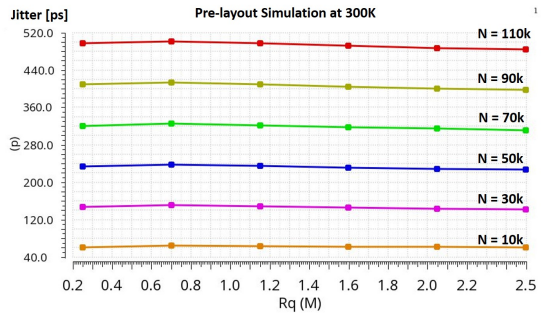
Figure 5.16: Corner and Monte Carlo simulation of the very front end output baseline

parameters of SiPM (for electronics performance) are the detector capacitance and quenching resistance, and the detector capacitance mainly depends on the number of sub-cells. Considering the SiPM parameter listed in table 3.1 (Section 3): for the application in Darkside-20k experiment, the sensor area for one channel readout is expected about 1 cm^2, corresponding to the sub-cells number of 110k, detector capacitance of about 5 nF; for the application in ToF-PET, a much smaller size usually 9 mm^2 is expected, corresponding to the sub-cells number of 10k, detector capacitance of about 450 pF. The quenching resistance of the SiPM listed in table 3.1 is quite high to suppress the after pulse noise which is critical for the rare event condition. In the other application, like ToF-PET, the quenching resistance can be down to hundreds of kΩ.

Figures 5.17 and 5.18 give the simulation results at 300K and 77K respectively: the SNR and jitter of very front end circuit with different number of cells (N) and quenching resistance values.

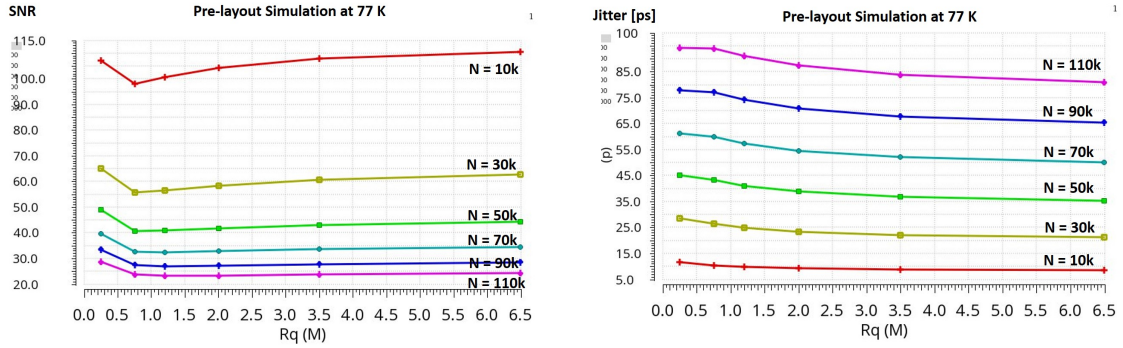


(a) Signal to Noise Ratio simulation



(b) Jitter simulation

Figure 5.17: The SNR and jitter of very front end with different number of sub-cells (N) and quenching resistance at 300K, single photon signal and anode signal readout



(a) Signal to Noise Ratio simulation

(b) Jitter simulation

Figure 5.18: The SNR and jitter of very front end with different number of sub-cells (N) and quenching resistance at 77K, single photon signal and anode signal readout

From the simulation results, some conclusions are summarized below:

1. The detector capacitance (number of sub-cells) makes the greatest effect on the electronics performance. The very fornt end can achieve better than 1 ns (100 ps) timing resolution for a 5 nF (450 pF) detector capacitance. For all conditions, the SNR is larger than 8, giving the capability for single photon event detection.

2. The quenching resistance is a not key parameter for the SNR and jitter of ALCOR very front end.

3. The ACLOR very front end is expected to have much better performance at 77 K than 300K.

Expected working at room and cryogenic temperature, either for anode or cathode SiPM signal readout, in total four different working situations have been studied, and the post-layout simulation for single-photon specifications are listed in table 5.4. Each mode should have a single photon SNR larger than 8, which should provide the conditions for a single-photon detection capability. The fast peaking time (O(10 ns)) minimizes the intrinsic jitter and allows for high event rate ability. The cathode readout chain has a bit worse jitter performance (slower peaking time) but a bit better SNR because the cathode RCG uses the PMOS as the common source transistor which has a lower transconductance and also lower noise. The simulation results, using the standard CMOS PDK provided by the foundry, anticipate much better performance in 77 K than 300 K.

Working mode	jitter	SNR	Amplitude	Peaking time	Power consumption
anode 300 K	500 ps	9.7	17.5 mV	10.4 ns	3.7 mW
cathode 300 K	718 ps	10.4	12 mV	14.9 ns	4.4 mW
anode 77 K	113 ps	21.6	29.2 mV	4.5 ns	3.4 mW
cathode 77 K	171 ps	18.4	20.1 mV	6.1 ns	4.1 mW

Table 5.4: List of simulation results of the VFE for different working modes, default SiPM parameters (1 cm^2), single photon signal

5.3 Discriminator and TDC

The transistor level schematic of ALCOR leading edge discriminator is shown in 5.19. The threshold voltage can be set from 550 mV to 878 mV, with a step of 2 mV. The discriminator is connected with the output of very front end part and is used to generate the CMOS trigger signal for analogue to digital conversion.

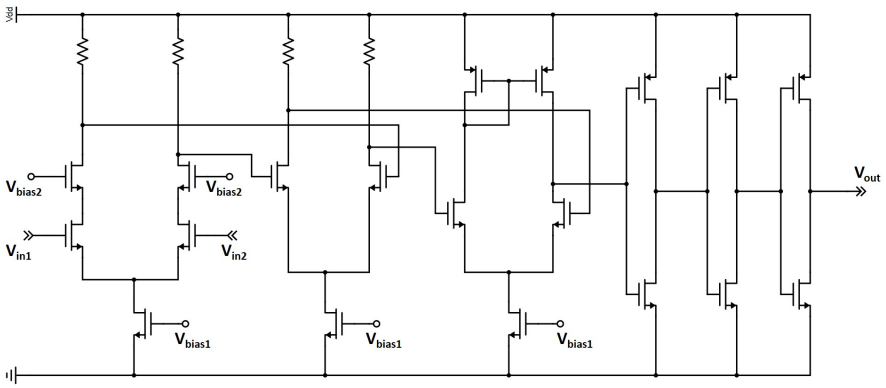


Figure 5.19: Discriminator schematic in ALCOR

The working principle of the TDC block is shown in figure 5.20, which is still implemented with the analogue interpolation method. Before each measurement, the voltage of $C1$ and $C2$ will be reset to a reference voltage V_0. The switch $s1$ is close, and the coarse current I_{coarse} starts to charge the capacitor $C1$, at the time the discriminator is triggered. $s1$ switches off and the charge stops at the next clock rising edge. The voltage value V_A is recorded by $C1$, the clock rising edge is the coarse time and recorded by a 15-bit counter. During the conversion phase, switch $s2$ is closed and $C2$ is charged with a smaller current I_{fine}. The control state machine stops the conversion when the latched comparator changes state, and the fine time stamp is recorded with a 9-bit counter.

The two capacitors C1 and C2 have the same value, and the interpretation factor is defined by the coarse and fine current ratio.

With the relation:

$$I_{coarse} = 64 I_{fine} \tag{5.4}$$

or,

$$I_{coarse} = 128 I_{fine} \tag{5.5}$$

the bin of fine time resolution is 50 ps or 25 ps.

The TDC has a time bin equals to 50 ps and a dead time of 150 ns when the master clock runs at the maximum frequency of 320 MHz. This fine time information is convoluted with a a coarse time information provided by the system clock, which is latched with a 15-bit gray-encoded counter.

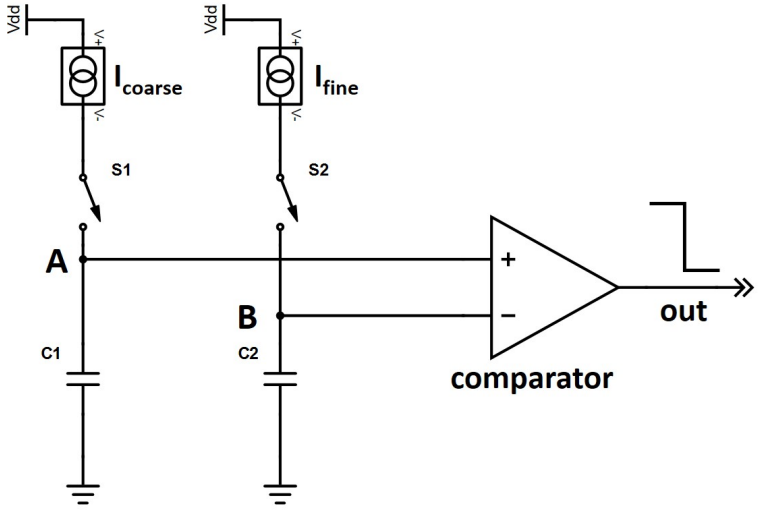


Figure 5.20: Working principle of ALCOR TDC

Figure 5.21 describes the simulation of TDC in ALCOR, to see the behaviour of analogue waveform, discriminator output and two voltage ramps.

Compared with the TDC used in the ASIC described in Chapter 4, The ALCOR TDC has two charge processes instead of firstly discharge and then recharge. Four independent TDCs are integrated into one channel instead of 4 buffered time to analogue converter (TAC) capacitors, increasing considerably the event rate capability (also for derandomization).

Another modification is the structure of the current ramp, the TDC discussed in Chapter 4 uses a double-cascode current source. While in TDC of ALCOR, the basic current mirror is implemented. For the analogue interpolation TDC, two fundamental aspects are uniformity of the clock and quality of current ramp for charge and discharge to achieve good linearity and resolution. We neglect the degradation of the linearity caused by the clock jitter, since it is mainly depending

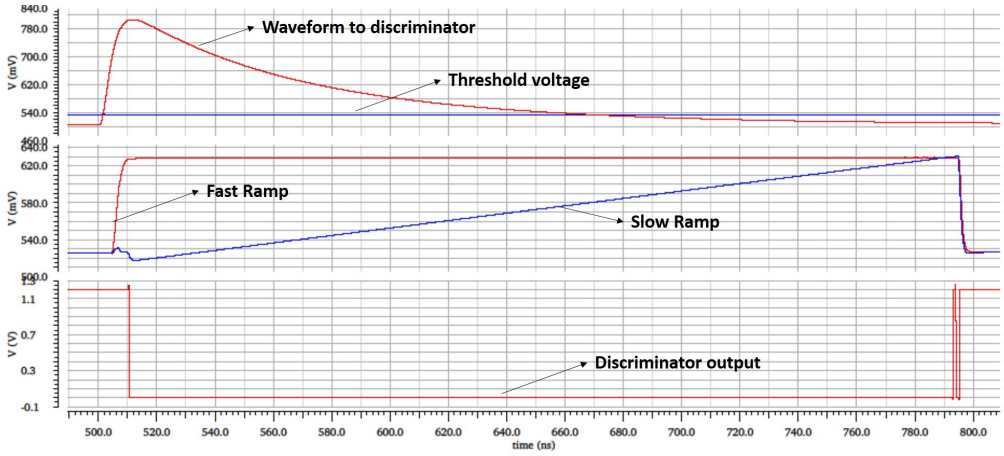


Figure 5.21: ALCOR TDC simulation

on the design of the LVDS receivers. The latter is a matter of how ideal a current ramp is generated.

An ideal current source features an infinite output resistance. In CMOS technology, the current mirror is used to copy a current from a reference which is based on the equation 5.6 (take the NMOS for example), this equation is already discussed before. λ is a process related constant to define the channel-length effect. Figure 5.22a shows the basic structure of current mirror, where I_2 makes the copy from reference current I_1.

$$I_{DS} = \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{th})^2 (1 + \lambda V_{DS}) \tag{5.6}$$

And the output resistance is defined as:

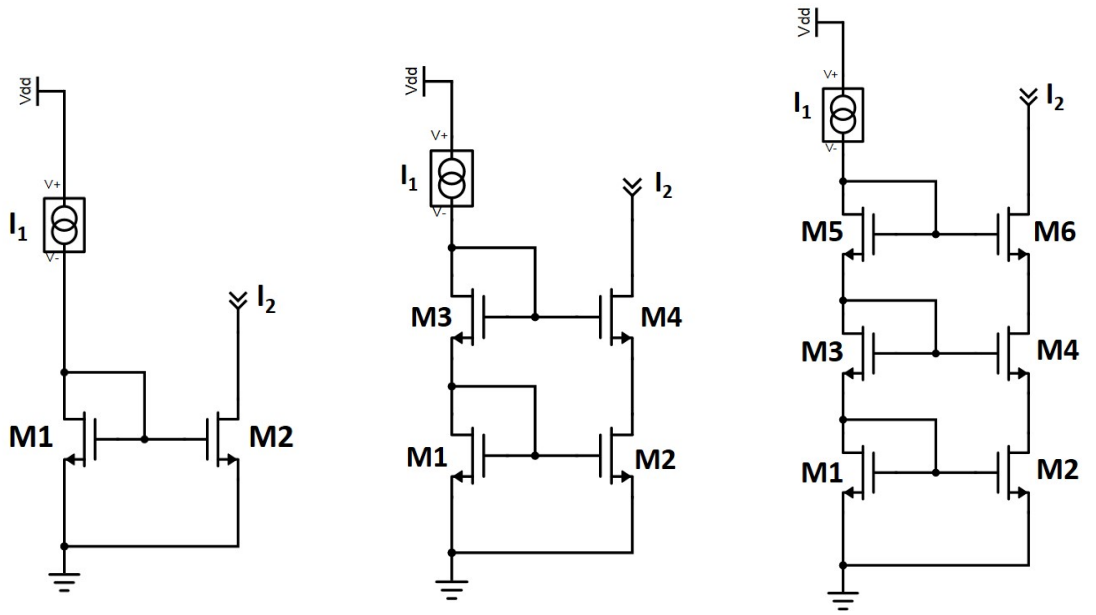
$$r_o = \frac{\partial V_{DS}}{\partial I_{DS}} \tag{5.7}$$

Considering the gate voltages are the same since they are connected directly, the equation 5.7 means the higher the output resistance r_o of a current source, the smaller is the modulation of the drain current I_{DS} due to the variation on the drain-to-source voltage V_{DS}.

From equation 5.6, the ration between I_2 and I_1 is expressed:

$$\frac{I_2}{I_1} = \frac{\mu_n C_{ox} \frac{W2}{L2} (V_{GS} - V_{TH})^2 (1 + \lambda V_{DS2})}{\mu_n C_{ox} \frac{W1}{L1} (V_{GS} - V_{TH})^2 (1 + \lambda V_{DS1})} = \frac{(W2/L2)(1 + \lambda V_{DS2})}{(W1/L1)(1 + \lambda V_{DS1})} \tag{5.8}$$

Furthermore, it can be acknowledged that the accuracy of the current mirror is strongly related to the drain to source voltage and the parameter λ. Consequently, there are two basic approaches to minimize the mismatch of current mirroring.



(a) Basic current mirror (b) Cascode current mirror

(c) Double-cascode CM

Figure 5.22: Different structures of current mirror

The first can be achieved by improving the output resistance of the current source with cascode topology. As shown in figure 5.22b, the output resistance of I_2 (source of $M4$) becomes:

$$r_{o,cascode} = r_{o2} \cdot g_{m4} \cdot r_{o4} \tag{5.9}$$

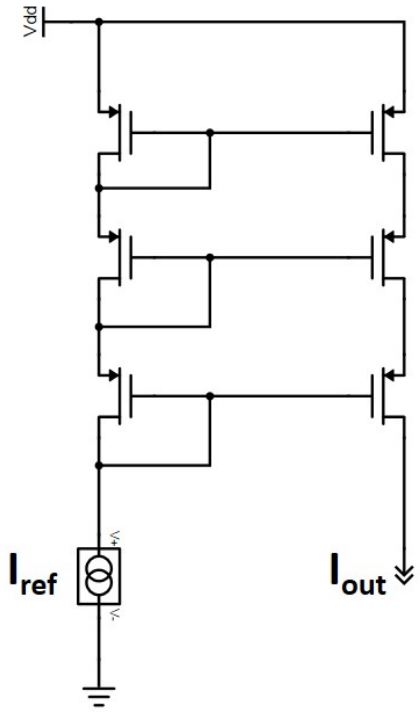
Where the output resistance of current source I_2 in cascode structure improved a factor of $gm4 \cdot r_{o4}$ than basic one. Further, double cascode topology (figure 5.22c) has the output resistance equals to:

$$r_{o,double-cascode} = r_{o2} \cdot g_{m4} \cdot r_{o4} \cdot g_{m6} \cdot r_{o6} \tag{5.10}$$

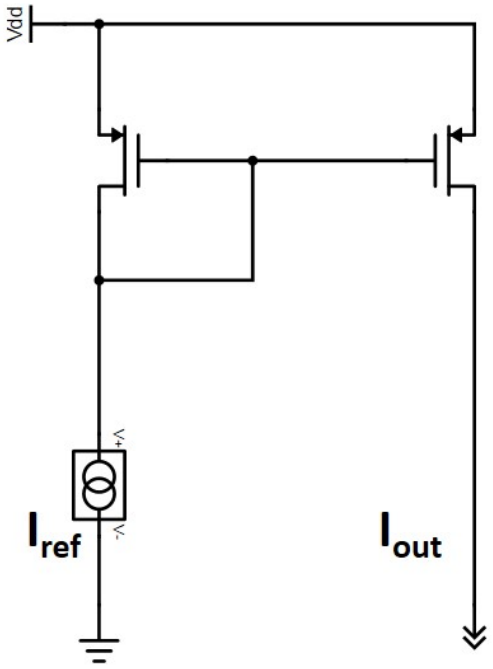
The cascode topology improves the output resistance with the price of less headroom. A second approach is to use large length transistor to reduce the channel length modulation, which is much more severe in short length device. This method saves the headroom but has to cost more layout area.

The current source in TDC of Chapter 4 adopts the double cascode current mirror as shown in figure 5.23a. While the ALCOR TDC current source is implemented with basic structure shown in figure 5.23b, considering the improved

threshold voltage at 77 K reduces the bed room of the cascode structure.



(a) Double cascade current source in TDC of the work in chapter 4



(b) Basic current source in TDC of ALCOR

Figure 5.23: Current ramp source in TDC of the ASIC in chapter 4 and ALCOR

Figure 5.24 shows simulation results of four TDCs at 320 MHz clock frequency and interpolation factor of 128. The corresponding fine time counter is recorded, when 350 test pulses with a time phase step of 15 ps are sent to each TDC. No fine counter value is lost when the scan range (350×15 ps $= 5.25$ ns) is larger than one clock cycle. Which proves the basic current source in ALCOR TDC has a good enough quality.

5.4 Calibration circuit and debug buffer

The test pulse generation circuit in located in each pixel, and the schematic is shown in figure 5.25a. When Cal_disable is "0", the calibration circuit is disabled. When Cal_trigger = "1" (Cal_ trigger_ inv = "0"), $I_{\text{calvfe}} = I_{\text{calin}}$, as shown in 5.25b, the injected charge to VFE is $I_{\text{calin}} \cdot t_0$.

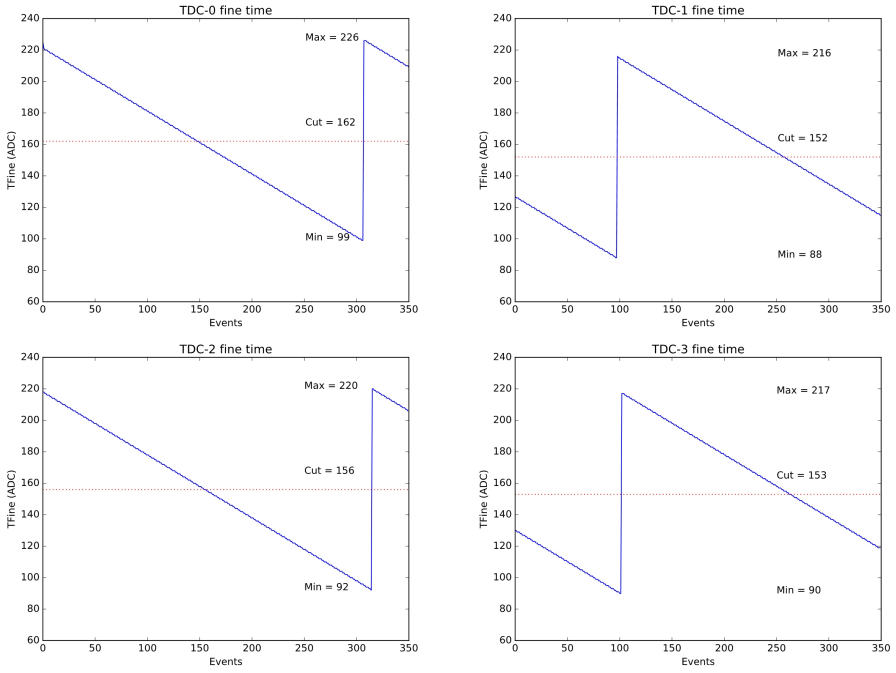
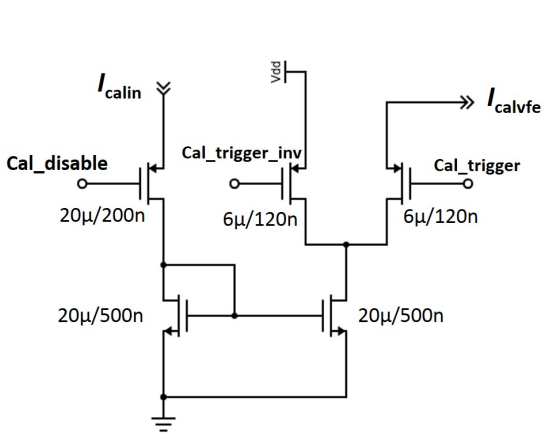
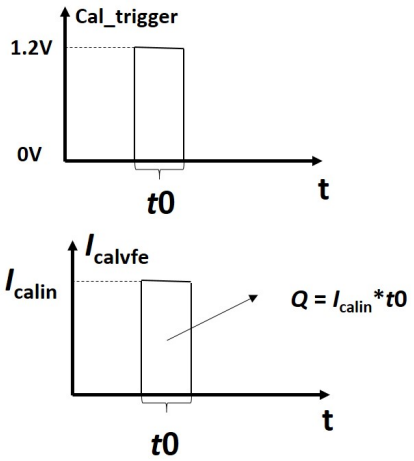


Figure 5.24: Test pulse scan of the four TDCs, the TDCs are configured with time bin of 25 ps



(a) Test pulse generation circuit located in pixel

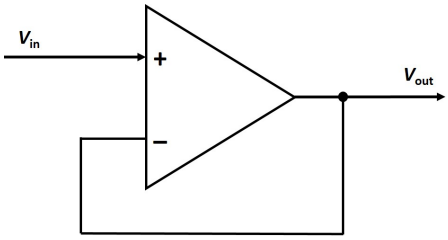


(b) Calibration current generation (Cal_disable = 0)

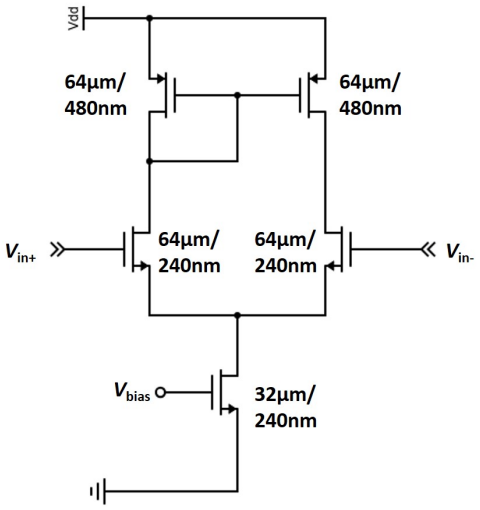
Figure 5.25: The calibration circuit

In order to check the working condition of the very front end, two debug buffers are used to bring the outputs of VFEs in pixel col1/ch1 and col8/ch1 (figure 5.1). The buffer structure and schematic are shown in figure 5.26a and figure 5.26b,

where about 400 μA current is consumed by each of buffer.



(a) Debug buffer



(b) Schematic of debug buffer amplifier

Figure 5.26: Buffer used to debug the output of ALCOR VFE

Figure 5.27 shows the simulation waveforms before and after the debug buffer. The ALCOR VFE has a high bandwidth, thus the output waveform of the buffer shows a larger peaking time. The mismatch of the buffer amplifier contributes a little variation of the baseline. But the two variations don't hinder to have a basic observation of the ALCOR VFE working condition.

5.5 Full pixel operation and data transmission

Five working modes of the pixel are configured by the pixel control logic, distributing the working of four TDCs.

1. OFF
2. LET: the rising edge of TRG1
3. TOT: the rising and falling edges of TRG1
4. TOT2: the rising edge of TRG1, the falling edge of TRG2
5. SR: rising edges of TRG1 and TRG2

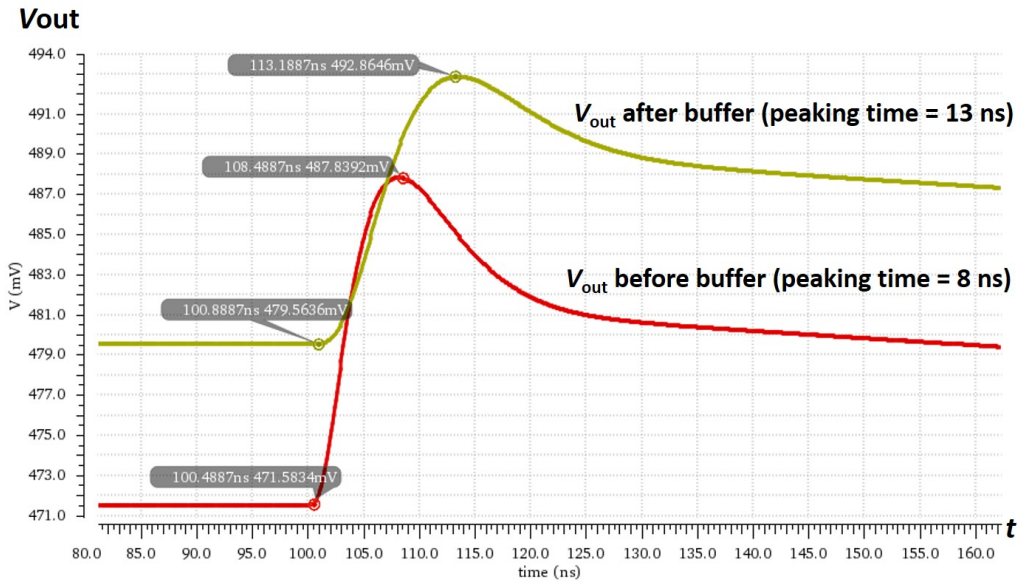


Figure 5.27: The input and output simulation waveforms of the debug buffer, with 50Ω and 18pF output load to model the oscilloscope

Where the TRG1 and TRG2 are trigger signal of two discriminators respectively.

The "OFF" mode will turn off the four TDCs, no conversion is carried out in this mode. The last four working modes have been discussed in the section of very front end: the dual threshold voltages measurement. "LET" mode is used for photon count, achieving high timing resolution and high count rate by using the four TDCs in turn. The event rate up to 5 MHz can be achieved; "ToT" and "ToT2" mode is used for energy measurement, and "ToT2" mode can have a better resolution by dual voltage setting. "SR" mode is used to characterize the VFE bandwidth and sensor capacitance.

Figure 5.28 shows the last four different working modes and the TDCs distribution.

At the end of the conversion, the data control logic generates a 32-bit payload containing the time-stamp, the channel ID and the specific TDC address, shown in figure 5.29.

In addition, if all four TDCs are occupied when a new event fired, another register is used to record the Status Word, consisting of the channel ID, TDC number and the number of cases to be discarded, and the number of single event upset (SEU) errors. Payloads are firstly queued and stored in a FIFO register in the channel. They are then transmitted to the periphery of the chip, where the End of Column (EoC) collects data from channels and acts as the interface between the channels and outside world.

Figure 5.30 describes the schematic diagram of data transmission in the ALCOR

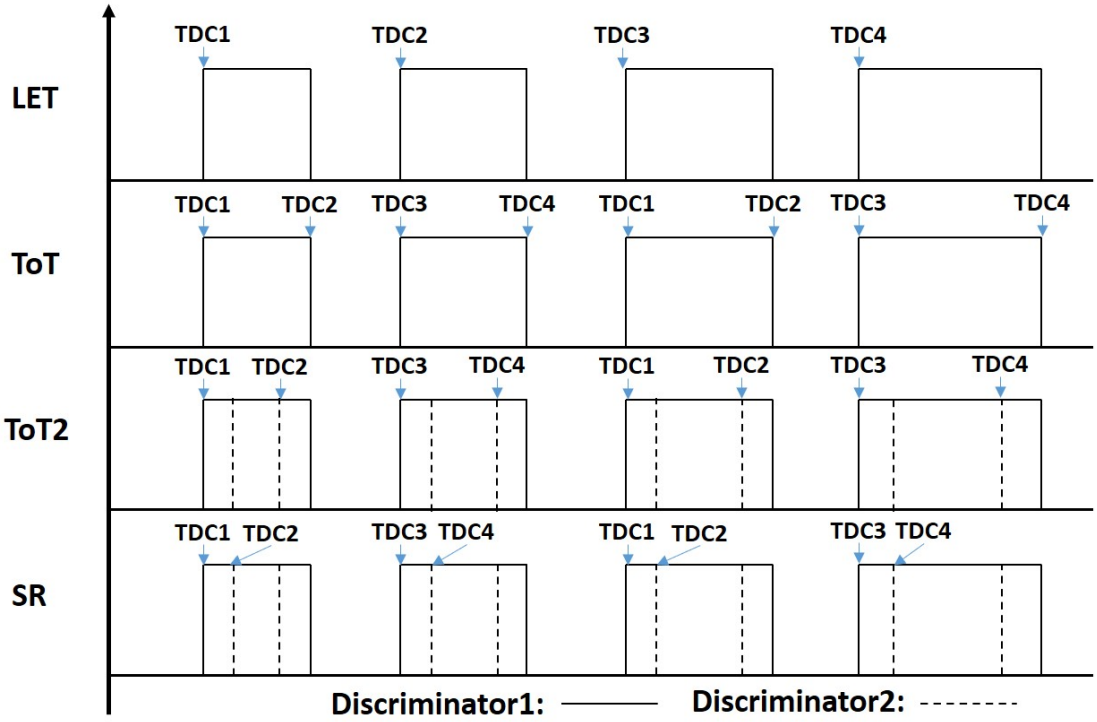


Figure 5.28: TDCs distribution with different working modes

Event word

Column ID	Channel ID	TDC ID	Coarse counter	Fine counter
3 bits	3 bits	2 bits	15 bits	9 bits

31 0

Status word

Column ID	Pixel ID	Lost event counter	Lost e.w. counter TDC1	Lost e.w. counter TDC2	Lost e.w. counter TDC3	Lost e.w. counter TDC4
3 bits	3 bits	6 bits	4 bits	4 bits	4 bits	4 bits

31 0

Figure 5.29: Format of event payload and status word

chip. Each column has a dedicated finite state machine (FSM), performing a scan of that column to read pixels with valid storage data. That is, only the channel

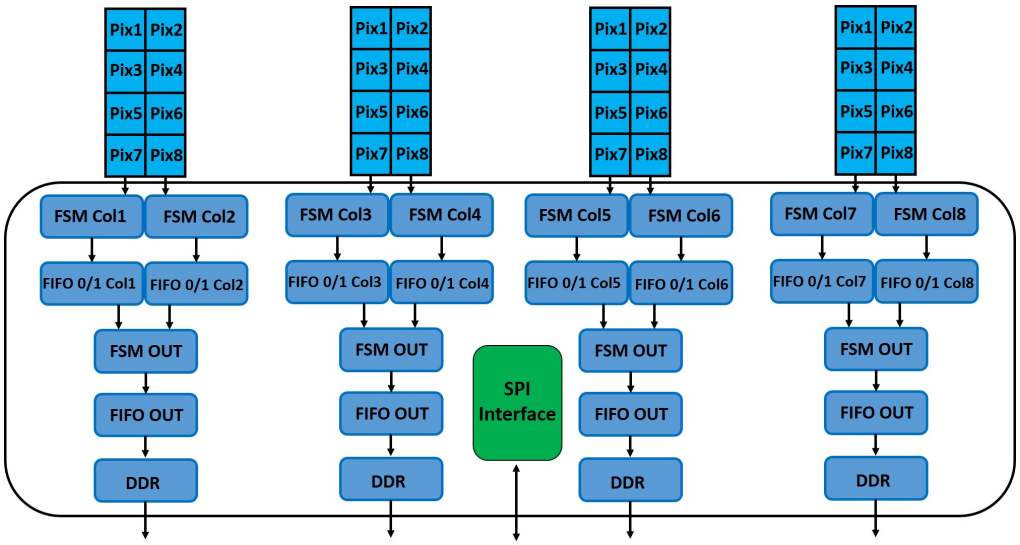


Figure 5.30: Data transmission in ALCOR

that stores the event data has the ownership of the bus. In each read cycle, each channel can only extract a set of data to prevent buffer overflow in other channels. During the readout cycle, the newly entered data can be processed and stored so that it can be read in the next cycle. In the above readout logic, it is not possible to sort the cases with time order. To avoid the ambiguity of the time sequence, an additional FSM sorts the two columns of data according to the coarse time information and outputs them to the next stage.

Configuration signals are provided to EoC using an SPI interface, while four LVDS drivers are used to transmit data off-chip (each two columns, 8 pixels share one LVDS transmitter). Each output link works in double data rate (DDR) mode with a maximum transmission speed of 640 Mbit/s.

The layout views of the full chip are shown in figure 5.31.

ALCOR chip integrates 32 readout channels, with pixel dimension of about 500 $\mu m\times$ 500 μm and full chip of 4.95 mm $\times$ 3.78 mm. The analogue and digital part is located in different regions; the adjacent readout pixels are mirror assembled to keep the distance between the analogue and digital part; In the full chip level, the analogue inputs are set on top while the digital outputs are in the bottom. All these strategies can reduce the cross talk from digital to analogue blocks.

The readout channels are squared and organized in a matrix, in this prototype, 32 channels are mounted per chip, and it is possible to study the silicon results easily with standard wire bonding pads on the PCB board. In the meanwhile, ALCOR is the proof of principle for 3D readout. Once verified by its silicon results, this ASIC can be easily expanded to a large scale integration array to readout large area of SiPMs with bump bonding pads.

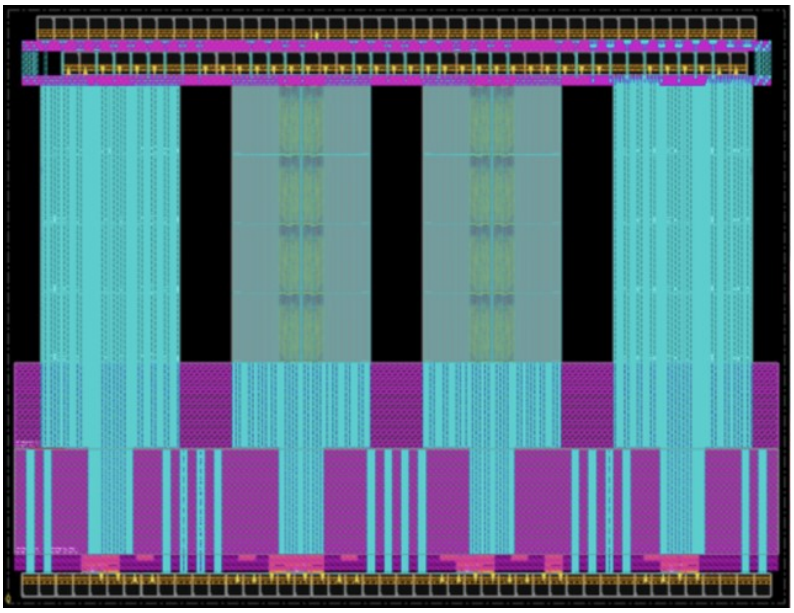


Figure 5.31: Layout view of ALCOR

5.6 Summary

ALCOR is a mixed-signal ASIC developed to readout SiPMs at both room and low temperature. The chip is designed in a standard 110 nm CMOS technology. Based on timing information measurement, both time of arrival and Time-over-Threshold (ToT) measurement are supported for timing and energy characterization. With high bandwidth very front end and 4 TDCs in one channel, an event rate of up to 5 MHz per channel can be accommodated. The TDC's time bin is about 50 ps (25 ps), and the power consumption of each TDC is only about 0.2 mW. This ASIC is a prototype for the study of the large area of SiPMs readout in cryogenic temperature and can be expanded to large scale readout channels in the next step study. The chip was sent for fabrication in November of 2019, and table 5.5 summaries the basic parameters and specifications.

Specifications	Value
Max sensor capacitance	5 nF
Event rate	up to 5 MHz/ch
Clock frequency	40 - 320 MHz
Time resolution	50 ps (25ps)
Channels	32
Power consumption	< 10 mW/ch
Layout area	4.95 mm $\times$ 3.78 mm

Table 5.5: List of ALCOR parameters

Chapter 6

Conclusions and future perspectives

In this work, the development of SiPM readout ASICs are carried out, expecting to read out large area SiPM in both cryogenic and room temperature.

The working principle and properties of SiPM are firstly introduced. Then the applications of SiPM are discussed, mainly in dark matter experiments and medical imaging. All of these explain the motivation to develop the readout ASIC which can work at cryogenic temperature and perform the digitization on-chip. The review of SiPM readout ASICs applied in different fields gives a basic overview of the related techniques, and a working scheme of this work is described.

The two test chips are designed and tested in liquid nitrogen temperature, aimed to study the properties and features of CMOS circuits working at cryogenic condition. TestChip1 integrates the front-end circuit used to readout out 24cm^2 SiPM and bring a single analogue output, following the original scheme proposed by the Darkside-20k project. TestChip1 works as expected at room temperature, and shows a good performance for the large area SiPM readout. While it is stable at liquid nitrogen temperature with a large detector capacitance load, more studies need to be carried out to find the reason.

In this work, a new scheme is proposed to for the front-end electronics. This scheme is expected to readout the SiPM by multi-channel ASIC, and the digitization is on-chip, which has several advantages: The sensor size readout by each channel is smaller, result in the better electronics performance; Full digital output data avoids the signal quality loss during the long distance transmission; Digital data is easy to be multiplexing and serialization, the number of transmission cables can be reduced; The warm electronics system can be simplified considerably; The readout ASIC can be also used in the LAr ToF-PET application.

TestChip2 consists of the very front end part, logic control and data transmission module, in order to verify and study the basic blocks for the mix-signal ASIC design. The very front end circuit is composed by RCG input stage and the TIA

stage implemented by a basic common source amplifier. The experimental results shows that the very front end can work properly at both room and liquid nitrogen temperature. Furthermore, the signal to noise ratio is improved by 40%, the rising time is reduced by about 40% from 300 K to 77 K. The digital signal modules, the signal synchronization module and the LVDS transmitter is verified in both 300 K and 77 K, at a maximum clock frequency of 320 MHz. The study of TestChip2 shows that both the key blocks of analogue and digital part can work properly in cryogenic temperature, and the very front end can have a significant improvement of the performance.

ALCOR, a 32-channel mixed-signal ASIC is developed as the first prototype of SiPM readout at cryogenic temperature. In each channel, anode and cathode signal readout RCGs are adopted as the input stage, providing low input impedance and high bandwidth readout. The signal from RCG is then sent two TIAs with programmable gains and dual threshold voltage setting can be applied for timing and energy (ToT) measurement. Four TDCs based on analogue interpretation principle are integrated in one channel, achieving a high time resolution (Time bin of 50 ps or 25 ps), high count rate (up to 5 MHz) and low power consumption (about 0.2 mW for each TDC). The digital data is sent off-chip by four LVDS transmitters, each output link works in double data rate (DDR) mode with a maximum transmission speed of 640 Mbit/s.

The readout channels of ALCOR is organized in an array of 4×8, which can be used as the the proof of principle for 3D readout. The next step of this work can be the development of 3D readout ASIC which can expand to a large scale readout array.

Bibliography

[1] J Aalbers et al. "DARWIN: towards the ultimate dark matter detector". In: *Journal of Cosmology and Astroparticle Physics* 2016.11 (2016), p. 017.

[2] CE Aalseth, F Acerbi, et al. "DarkSide 20K: A Yellow book technical proporsal/pre-technical design report". In: *INFN*. 2016.

[3] Craig E Aalseth et al. "DarkSide-20k: A 20 tonne two-phase LAr TPC for direct dark matter detection at LNGS". In: *The European Physical Journal Plus* 133.3 (2018), p. 131.

[4] K Abe et al. "The T2K experiment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 659.1 (2011), pp. 106–135.

[5] Fabio Acerbi et al. "Cryogenic characterization of FBK HD near-UV sensitive SiPMs". In: *IEEE Transactions on Electron Devices* 64.2 (2017), pp. 521–526.

[6] Fabio Acerbi and Stefan Gundacker. "Understanding and simulating SiPMs". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 926 (2019), pp. 16–35.

[7] C Adloff et al. "Construction and commissioning of the CALICE analog hadron calorimeter prototype". In: *Journal of Instrumentation* 5.05 (2010), P05004.

[8] Salleh Ahmad et al. "Triroc: A multi-channel SiPM read-out ASIC for PET/PET-ToF application". In: *IEEE Transactions on Nuclear Science* 62.3 (2015), pp. 664–668.

[9] Antonio Amoroso et al. "A cylindrical GEM detector with analog readout for the BESIII experiment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 824 (2016), pp. 515–517.

[10] MG Bagliesi et al. "A custom front-end ASIC for the readout and timing of 64 SiPM photosensors". In: *Nuclear Physics B-Proceedings Supplements* 215.1 (2011), pp. 344–348.

[11] Arnout Beckers et al. “Characterization and modeling of 28-nm FDSOI CMOS technology down to cryogenic temperatures”. In: *Solid-State Electronics* 159 (2019), pp. 106–115.

[12] Maria Giuseppina Bisogni et al. “INSIDE in-beam positron emission tomography system for particle range monitoring in hadrontherapy”. In: *Journal of medical imaging* 4.1 (2016), p. 011005.

[13] S Blin et al. “Dedicated very front-end electronics for an ILC prototype hadronic calorimeter with SiPM readout”. In: *Report-No: LC-DET-2006-007* 12 (2006).

[14] Nilay Bostan. “Performance of Photon Detectors in ProtoDUNE”. In: *Bulletin of the American Physical Society* (2020).

[15] P Buzhan et al. “The advanced study of silicon photomultiplier”. In: *Advanced Technology And Particle Physics*. World Scientific, 2002, pp. 717–728.

[16] Augusto Carimatto et al. “11.4 A 67,392-SPAD PVTB-compensated multichannel digital SiPM with 432 column-parallel 48ps 17b TDCs for endoscopic time-of-flight PET”. In: *2015 IEEE International Solid-State Circuits Conference-(ISSCC) Digest of Technical Papers*. IEEE. 2015, pp. 1–3.

[17] F Carnesecchi. “Light detection in DarkSide-20k”. In: *Journal of Instrumentation* 15.03 (2020), p. C03038.

[18] Hucheng Chen et al. “Readout electronics for the MicroBooNE LAr TPC, with CMOS front end at 89K”. In: *Journal of Instrumentation* 7.12 (2012), p. C12004.

[19] Tianbing Chen et al. “CMOS reliability issues for emerging cryogenic Lunar electronics applications”. In: *Solid-state electronics* 50.6 (2006), pp. 959–963.

[20] W Cheng et al. “A mixed-signal large dynamic range front-end ASIC for high capacitance detectors”. In: *Journal of Instrumentation* 14.08 (2019), P08013.

[21] William F Clark et al. “Low temperature CMOS-a brief review”. In: *IEEE Transactions on Components Hybrids and Manufacturing Technology* 15.3 (1992), pp. 397–404.

[22] BESIII Collaboration et al. *BESIII Cylindrical GEM Inner Tracker*. Tech. rep. Technical report, July, 2014.

[23] LHCb Collaboration et al. *LHCb tracker upgrade technical design report*. Tech. rep. Technical report, February, 2014.

[24] Gianmaria Collazuol et al. "Single photon timing resolution and detection efficiency of the IRST silicon photo-multipliers". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 581.1-2 (2007), pp. 461–464.

[25] Albert Comerma Montells. *PACIFIC: The readout ASIC for the SciFi Tracker of the LHCb detector*. Tech. rep. Technical report, September, 2018.

[26] F Corsi et al. "Modelling a silicon photomultiplier (SiPM) as a signal source for optimum front-end design". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 572.1 (2007), pp. 416–418.

[27] José Mazorra de Cos et al. "PACIFIC: SiPM readout ASIC for LHCb upgrade". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* (2017).

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

[29] Marco D'Incecco et al. "Development of a novel single-channel, 24 cm 2, SiPM-based, cryogenic photodetector". In: *IEEE Transactions on Nuclear Science* 65.1 (2017), pp. 591–596.

[30] M Dabrowski et al. "Analog Front-ends for Cold Experiments". In: ().

[31] Nguyen Cong Dao et al. "An enhanced MOSFET threshold voltage model for the 6–300 K temperature range". In: *Microelectronics Reliability* 69 (2017), pp. 36–39.

[32] G De Geronimo, P O'Connor, and J Grosholz. "A CMOS baseline holder (BLH) for readout ASICs". In: *1999 IEEE Nuclear Science Symposium. Conference Record. 1999 Nuclear Science Symposium and Medical Imaging Conference (Cat. No. 99CH37019)*. Vol. 1. IEEE. 1999, pp. 370–374.

[33] Gianluigi De Geronimo and Shaorui Li. "Shaper design in CMOS for high dynamic range". In: *IEEE Transactions on Nuclear Science* 58.5 (2011), pp. 2382–2390.

[34] Gianluigi De Geronimo et al. "VMM1—An ASIC for micropattern detectors". In: *IEEE Transactions on Nuclear Science* 60.3 (2013), pp. 2314–2321.

[35] A Di Francesco et al. "TOFPET2: a high-performance ASIC for time and amplitude measurements of SiPM signals in time-of-flight applications". In: *Journal of Instrumentation* 11.03 (2016), p. C03042.

[36] Jonathan Echevers, nEXO Collaboration, et al. "Large Area SiPM Readout and Signal Processing for nEXO". In: *APS* 2019 (2019), J10–007.

[37] A Falcone et al. “Cryogenic SiPM arrays for the DUNE photon detection system”. In: *arXiv preprint arXiv:2001.09051* (2020).

[38] Thomas Frach et al. “The digital silicon photomultiplier—Principle of operation and intrinsic detector performance”. In: *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*. IEEE. 2009, pp. 1959–1965.

[39] Fritz H Gaensslen et al. “Very small MOSFET’s for low-temperature operation”. In: *IEEE Transactions on Electron Devices* 24.3 (1977), pp. 218–229.

[40] Rajesh Ganai et al. “A Proof-of-Principle for Time of Flight-Positron Emission Tomography Imaging”. In: *XXII DAE High Energy Physics Symposium*. Springer. 2018, pp. 125–128.

[41] Alberto Gola et al. “NUV-sensitive silicon photomultiplier technologies developed at Fondazione Bruno Kessler”. In: *Sensors* 19.2 (2019), p. 308.

[42] Stefan Gries. *Quantitative Corpus Linguistics with R: A Practical Introduction*. Routledge, 2009.

[43] Stefan Gundacker et al. “Time of flight positron emission tomography towards 100ps resolution with L (Y) SO: an experimental and theoretical analysis”. In: *Journal of Instrumentation* 8.07 (2013), P07014.

[44] S Gundacker et al. “On the comparison of analog and digital SiPM readout in terms of expected timing performance”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 787 (2015), pp. 6–11.

[45] G Hall. “Recent progress in front end ASICs for high-energy physics”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 541.1-2 (2005), pp. 248–258.

[46] Hamamatsu. *MPPC Technical Note*. URL: https://www.hamamatsu.com/resources/pdf/ssd/mppc_kapd9005e.pdf. (accessed: 05.12.2019).

[47] Hamamatsu. *MPPCs (SiPMs)/MPPC arrays*. URL: https://www.hamamatsu.com/eu/en/product/optical-sensors/mppc/mppc_mppc-array/index.html. (accessed: 05.12.2019).

[48] Zhong He. “Review of the Shockley–Ramo theorem and its application in semiconductor gamma-ray detectors”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 463.1-2 (2001), pp. 250–267.

[49] James R Hoff et al. “Cryogenic Lifetime Studies of 130 nm and 65 nm CMOS Technologies for High-Energy Physics Experiments”. In: *IEEE Transactions on Nuclear Science* 62.FERMILAB-PUB-15-078-PPD (2015).

[50] David FC Hsu et al. "Studies of a next-generation silicon-photomultiplier–based time-of-flight PET/CT system". In: *Journal of Nuclear Medicine* 58.9 (2017), pp. 1511–1518.

[51] Chenming Hu et al. "Hot-electron-induced MOSFET degradation-model, monitor, and improvement". In: *IEEE Journal of Solid-State Circuits* 20.1 (1985), pp. 295–305.

[52] Brian F Hutton et al. "Development of clinical simultaneous SPECT/MRI". In: *The British journal of radiology* 91.1081 (2018), p. 20160690.

[53] Fan Ji, Mikko Juntunen, and Iiro Hietanen. "Evaluation of electrical crosstalk in high-density photodiode arrays for X-ray imaging applications". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 610.1 (2009), pp. 28–30.

[54] RK Kirschman. "Cold electronics: an overview". In: *Cryogenics* 25.3 (1985), pp. 115–122.

[55] Glenn F Knoll. *Radiation detection and measurement*. John Wiley & Sons, 2010.

[56] Antje-Christin Knopf and Antony Lomax. "In vivo proton range verification: a review". In: *Physics in Medicine & Biology* 58.15 (2013), R131.

[57] Helmut Kopka and Patrick W. Daly. *Guide to LaTeX*. 4th. Boston: Addison-Wesley, 2004.

[58] C de La Taille et al. "Performance of 2nd generation CALICE/EUDET ASICs". In: *Journal of Physics: Conference Series*. Vol. 293. IOP Publishing. 2011, p. 012016.

[59] Shaorui Li et al. "LAr TPC electronics CMOS lifetime at 300 K and 77 K and reliability under thermal cycling". In: *IEEE Transactions on Nuclear Science* 60.6 (2013), pp. 4737–4743.

[60] F Liu et al. "Cold Electronics System Development for ProtoDUNE-SP and SBND LAr TPC". In: *2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. IEEE. 2017, pp. 1–4.

[61] Gerhard Lutz et al. *Semiconductor radiation detectors*. Vol. 40. Springer, 1999.

[62] Michela Marafini et al. "The INSIDE project: innovative solutions for in-beam dosimetry in hadrontherapy". In: (2015).

[63] D Marano et al. "Silicon photomultipliers electrical model extensive analytical analysis". In: *IEEE Transactions on Nuclear Science* 61.1 (2013), pp. 23–34.

[64] Marco Marcante. "Development and characterization of Silicon Photomultipliers for cryogenic applications". PhD thesis. University of Trento, 2019.

[65] S Martoiu et al. "A pixel front-end ASIC in 0.13 μm CMOS for the NA62 experiment with on pixel 100 ps time-to-digital converter". In: *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*. IEEE. 2009, pp. 55–60.

[66] Manuel de Medeiros Silva and Luis B Oliveira. "Regulated common-gate transimpedance amplifier designed to operate with a silicon photo-multiplier at the input". In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 61.3 (2013), pp. 725–735.

[67] Konstantinos Misiakos and Dimitris Tsamakis. "Accurate measurements of the silicon intrinsic carrier density from 78 to 340 K". In: *Journal of applied physics* 74.5 (1993), pp. 3293–3297.

[68] Frank Mittelbach et al. *The LaTeX Companion*. 2nd. Addison Wesley, 2004.

[69] Miki Nishimura et al. "Pixelated positron timing counter with SiPM-readout scintillator for MEG II experiment". In: *PoS* (2015), p. 011.

[70] Adam Nepomuk Otte et al. "Characterization of three high efficiency and blue sensitive silicon photomultipliers". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 846 (2017), pp. 106–125.

[71] Roberto Pagano et al. "Dark current in silicon photomultiplier pixels: data and model". In: *IEEE transactions on electron devices* 59.9 (2012), pp. 2410–2416.

[72] Slawomir Piatek. *Silicon Photomultiplier Operation, Performance and Possible Applications*. URL: https://www.hamamatsu.com/sp/hc/osh/sipm_webinar_1.10.pdf. (accessed: 04.17.2020).

[73] Claudio Piemonte et al. "Characterization of the first prototypes of silicon photomultiplier fabricated at ITC-irst". In: *IEEE Transactions on Nuclear Science* 54.1 (2007), pp. 236–244.

[74] RG Pires et al. "Carrier freezeout in silicon". In: *Cryogenics* 30.12 (1990), pp. 1064–1068.

[75] N Poonthottathil et al. "Reliability studies of electronic components for the operation at cryogenic temperature". In: *arXiv preprint arXiv:1911.10285* (2019).

[76] Jan M Rabaey, Anantha P Chandrakasan, and Borivoje Nikolić. *Digital integrated circuits: a design perspective*. Vol. 7. Pearson education Upper Saddle River, NJ, 2003.

[77] Marco Ramilli. "Characterization of SiPM: temperature dependencies". In: *2008 IEEE Nuclear Science Symposium Conference Record*. IEEE. 2008, pp. 2467–2470.

[78] Alejandro Ramirez, Xinran Li, and Andrew Renshaw. "3D -Three Dimensional Positron Identification with Liquid Argon Total-Body TOF-PET". In: *Bulletin of the American Physical Society* 63 (2018).

[79] Behzad Razavi. *Design of analog CMOS integrated circuits*. Tata McGraw-Hill Education, 2002.

[80] Janet S Reddin et al. "Performance evaluation of the SiPM-based Siemens Biograph Vision PET/CT system". In: *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*. IEEE. 2018, pp. 1–5.

[81] Dieter Renker. "Geiger-mode avalanche photodiodes, history, properties and problems". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 567.1 (2006), pp. 48–56.

[82] V Revéret et al. "CESAR: Cryogenic electronics for space applications". In: *Journal of Low Temperature Physics* 176.3-4 (2014), pp. 446–452.

[83] A Rivetti et al. "TIGER: A front-end ASIC for timing and energy measurements with radiation detectors". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* (2019).

[84] Angelo Rivetti. *CMOS: front-end electronics for radiation sensors*. CRC press, 2018.

[85] M Da Rocha Rolo et al. "A custom readout electronics for the BESIII CGEM detector". In: *Journal of Instrumentation* 12.07 (2017), p. C07017.

[86] MD Rolo et al. "A low-noise CMOS front-end for TOF-PET". In: *Journal of Instrumentation* 6.09 (2011), P09003.

[87] MD Rolo et al. "Tofpet asic for pet applications". In: *Journal of Instrumentation* 8.02 (2013), p. C02050.

[88] C Romo-Luque, PETALO Collaboration, et al. "PETALO: Time-of-Flight PET with liquid xenon". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* (2019), p. 162397.

[89] Victoria Rosén et al. "An Open Infrastructure for Advanced Treebanking". In: *META-RESEARCH Workshop on Advanced Treebanking at LREC2012*. Ed. by Jan Hajič et al. Istanbul, Turkey, 2012, pp. 22–29.

[90] Daniela Schulz-Ertner and Hirohiko Tsujii. "Particle radiation therapy using proton and heavier ion beams". In: *Journal of clinical oncology* 25.8 (2007), pp. 953–964.

[91] SensL. *Silicon Photomultipliers (SiPM) products*. URL: https://www.onsemi.com/products/sensors/silicon-photomultipliers-sipm. (accessed: 05.12.2019).

[92] SensL. *Work shop slides in LPC Clermont-Ferrand, 12-Mar-2014*. URL: https://indico.cern.ch/event/306859/contributions/705883/attachments/584419/804472/Schaart_fast_timing_workshop_Clermont_140312.pdf. (accessed: 10.12.2019).

[93] Wei Shen. "Development of high performance readout ASICs for silicon photomultipliers (SiPMs)". PhD thesis. Ruperto Carola University of Heidelberg, Germany, 2012.

[94] Frank Simon, CALICE Collaboration, et al. "Studies of scintillator tiles with sipm readout for imaging calorimeters". In: *Journal of Physics: Conference Series*. Vol. 293. IOP Publishing. 2011, p. 012074.

[95] Alessandro Spinelli and Andrea L Lacaita. "Physics and numerical simulation of single photon avalanche diodes". In: *IEEE Transactions on Electron Devices* 44.11 (1997), pp. 1931–1943.

[96] Simon Tam, Ping-Keung Ko, and Chenming Hu. "Lucky-electron model of channel hot-electron injection in MOSFET's". In: *IEEE Transactions on Electron Devices* 31.9 (1984), pp. 1116–1125.

[97] PAOLO TRIGILIO. "Development of an ASIC for SiPM readout in SPECT applications". In: *PhD thesis* (2016).

[98] Antonin Vacheret et al. "Characterization and simulation of the response of Multi-Pixel Photon Counters to low light levels". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 656.1 (2011), pp. 69–83.

[99] Marc Van Dongen. *LaTeX and Friends*. Springer, 2012. URL: http://csweb.ucc.ie/~dongen/LAF/LAF.html.

[100] Stefaan Vandenberghe et al. "Recent developments in time-of-flight PET". In: *EJNMMI physics* 3.1 (2016), p. 3.

[101] LMK Vandersypen et al. "Interfacing spin qubits in quantum dots and donors—hot, dense, and coherent". In: *npj Quantum Information* 3.1 (2017), p. 34.

[102] JJ Vaquero et al. "MRI compatibility of position-sensitive photomultiplier depth-of-interaction PET detectors modules for in-line multimodality preclinical studies". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 702 (2013), pp. 83–87.

[103] Neil HE Weste and David Harris. *CMOS VLSI design: a circuits and systems perspective*. Pearson Education India, 2015.

[104] Habib Zaidi and Alberto Del Guerra. "An outlook on future design of hybrid PET/MRI systems". In: *Medical physics* 38.10 (2011), pp. 5667–5689.

[105] Milan Zvolsky, EndoTOFPET-US Collaboration, et al. "EndoTOFPET-US–A Miniaturised Calorimeter for Endoscopic Time-of-Flight Positron Emission Tomography". In: *Journal of Physics: Conference Series*. Vol. 587. IOP Publishing. 2015, p. 012068.

This Ph.D. thesis has been typeset by means of the TEX-system facilities. The typesetting engine was pdfLATEX. The document class was `toptesi`, by Claudio Beccari, with option `tipotesi=scudo`. This class is available in every up-to-date and complete TEX-system installation.