UNIVERSITAT POLITÈCNICA DE VALÈNCIA

_ TELECOM ESCUELA TÉCNICA VLC SUPERIOR DE INGENIERÍA DE TELECOMUNICACIÓN

Time to Digital and Charge to Digital converters for SiPM front ends

Thesis by:
Alessandro Morini

Advisor:
Dr. Vicente Herrero Bosch

Trabajo Fin de Máster presentado en la Escuela Técnica Superior de Ingenieros de Telecomunicación de la Universitat Politècnica de València, para la obtención del Título de Máster en Ingeniería Telecomunicación

Curso 2018-19

Valencia, 07 de abril de 2020

Escuela Técnica Superior de Ingeniería de Telecomunicación
Universitat Politècnica de València
Edificio 4D. Camino de Vera, s/n, 46022 Valencia
Tel. +34 96 387 71 90, ext. 77190
www.etsit.upv.es

VLC/ CAMPUS
VALENCIA, INTERNATIONAL CAMPUS OF EXCELLENCE

EMAS
Gestión ambiental verificada
REG.NO. ES-CV-000030

Nonno

Questa TESI è dedicata a mio nonno che è venuto a mancare poco tempo fa. Ha sempre creduto in me e questo è il mio tributo a colui che è stato un esempio di vita, il mio esempio di vita.
GRAZIE.
Questo è per te.
CAPO VE'.

Ti voglio bene
Il tuo amato Alessandro

Acknowledgements

I would like to inscribe this chapter to all of the people support me throughout my master's degree, giving me the chance to achieve this result and growing as a person.

Firstly, I am grateful to my advisor in Valencia Prof. Vicente Herrero Bosch. The door to his office was always open whenever I ran into a trouble or had a question about my project work. He shared his knowledge and passion on electronics, following me over the master period step by step. Indeed, he gave me the material I needed to develop in the best way the final goal and helped me out in order to reach the requirements' machine. Moreover, he has always been patient with me answering my questions and sorting out my doubts. It has been a pleasure working with him and having a chance to work in this environment work. I will be open to have another opportunity to work with him in a future.

In second, thanks to the Prof. Andrea Giovanni Bonfanti who helped me out in the developing of my work in Valencia and support me in this journey. It would not have been possible without his aid.

Furthermore, I will always be thankful to the POLI for these unforgettable and, at the same time, tough years. Once you are a POLI student you will always be part of this "family".

Moreover, I am grateful to my girlfriend who has been by my side throughout this important path of my life. As it says in Spain "me ha aguantado" in the worst moment and she was there to support me and get out of any trouble. Thanks for what you have done.

I want to say thank to my family who always believe in me (even when I wanted to quite at the beginning of the bachelor) supporting me in any decision I took. Everyone has always been my side. In particular, I want to dedicate my success to my father Roberto, my mother Roberta and my brother Pietro who, even

though I was far away from home they have been always there for a tip or just to ask me how everything was going on. I love you all forever. Secondly, I feel thankful to my grandparents Arturo (finally one of your dream came true, one of your grandson is an engineer, unbelievable CAPO but it is true), Clara and Bruna (both of them they are as a mother for me and they risk to passed away over these years, but fortunately they are still here with me and they gave me the strength to achieve this important step on my life). Furthermore, I bless all of the friends I have met in these years who made this experience unforgettable. Particularly, my friends back home Filippo, Federico, Giulio, Davide, Francesco, Luca e Marco were there, and they will be always my points of reference. Then, all of my friends from the university in particular Adriano (you are the first one I met and the one will be always there even though you are following in your American dream), Matteo, Riccardo (you are the one will be always by my side in Milan, Valencia and wherever my captain), Edoardo e Lorenzo (my couple of Piazza Giolitti). I would like to say thank to more and more people, but they are too many. I will think about you though in this special day for me.

This is my personal win even if it is yours as well and I dedicate to you my success. I love you all. This is just the first step I have reached to, and I hope to keep going to this path. Go on and live on.

Thank you all,
Alessandro

Abstract

The nuclear medicine is largely used as diagnostic tool and capable to identify, through a radioactive compound, presence of tumoral cells, any endocrine, gastrointestinal or neurological disturbs and anomalies in the patient in exam. There a bunch of techniques for this type of medicine, but one of the most utilized is the Positron Emission Tomography (PET). PET is based on the coincident identification of gamma photons achieved from an annihilation process between an electron of a human cell and a positron released by a decade atom of the introduced compound. This procedure allows acquire 3-D distributions of the radioactive tracer.

Two tasks have been carried out in this master thesis: implementation of a single front-end channel (composed by an amplifier and a gated integrator) taking into account specification have been set in advance; a survey on a Time to Digital Converter (TDC) and Analog to Digital Converter (ADC).

The first one accomplishes firstly a preamplifier for the integrated SiPM using a 0.35 um technology. The output current will feed a TDC (boosted for fast signals) and an ADC (boosted for charge integration). During the second step a gated charge integrator has been carried out, which will be used for the analog chain needed for the ADC. It has been settled an integration start threshold and a configurable integrating window.

Regarding the second task, we focused on different configurations for TDC that could work with the given requirements. Furthermore, a Sample and Hold (S/H) and a Successive Approximation ADC (SAR) have been implemented. The SAR is composed by a quite fast comparator, a programmed logic in Verilog-A, necessary to study bit by bit, and a DAC in the end.

Overall, it is necessary a circuit study in advance in order to evaluate over the possible configurations (DC, noise, trans, AC, stability and so on). Afterwards, the picked configuration is built up on Cadence Virtuoso in order to see if there is correspondence between the theory and the reality.

Sommario

La medicina nucleare è utilizzata come strumento diagnostico in grado di individuare, attraverso l'impiego di una miscela radioattiva, la presenza di cellule cancerogene, disturbi dell'endocrino, gastrointestinali o neurologici e altre anomalie nel paziente in esame. Tra le tecniche utilizzate c' è la Tomografia ad Emissione di Positroni (PET). La PET è basata sul rilevamento di raggi gamma ottenuti durante un processo di annichilazione tra un elettrone di una cellula umana e un positrone rilasciato da un atomo decaduto della miscela introdotta. Questa procedura permette di acquisire distribuzioni 3-D del radio farmaco introdotto al paziente in esame.
Questo lavoro di tesi viene sviluppato in due parti distinte: realizzazione di un front-end a singolo canale (composto da un preamplificatore e un integratore) rispettando determinate specifiche in fase di lavoro e sondaggio su diverse topologie di Time to Digital Converter (TDC) e Analog to Digital Converter (ADC). Nella prima parte viene implementato un preamplificatore per il SiPM integrato usando una tecnologia 0.35 um. La corrente in uscita da tale amplificatore verrà poi fornita a un TDC (ottimizzato per segnali veloci) e a un ADC (ottimizzato per integrazione di carica). Il front-end prevede poi la realizzazione di un integratore gate con una determinata soglia di controllo e una finestra di integrazione configurabile.
Per quanto riguarda la seconda parte, ci si è prima concentrati su una possibile configurazione di TDC che potesse rispettare i requisiti di lavoro. Successivamente, è stato implementato un Track and Hold (T/H) e un ADC ad Approssimazioni Successive (SAR). Relativamente al SAR, è stato necessario realizzare un comparatore con una buona velocità, una logica ottenuta con Verilog-A capace di cambiarmi o meno il singolo bit ed infine un DAC.
Il lavoro complessivo prevede uno studio previo delle varie configurazioni possibili per i vari componenti delle due diverse catene. In un secondo momento, si incentrati sull'implementazione di tali configurazioni in Cadence Virtuoso per vedere quale raggiungesse le condizioni ottimali.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter gives an overview of the work done for the Master Thesis, focusing on the different aspects of the PET, in order to have a first touch with the machine. Furthermore, the goals we want to achieve are listed to give an overview of the work done throughout the thesis period.
In the end, the structure of the Thesis is presented explaining in detail chapter by chapter's content.

1.1 Motivation

The principal systems of the medical image such as the nuclear magnetic resonance (MRI), computerized tomography (CT) and echography are known to be able to provide information about the anatomic of the human body with a good spatial resolution. Nevertheless, in some occasions they are not capable to identify alterations in the metabolism and functions of the organism.
In order to recognize information on the metabolism, techniques in diagnostic nuclear medicine are widely used. One of the systems, which offers this advantage, is the Positron Emission Tomography, simply knowing as PET. It is characterized by the detection of a radioactive compound injected in the patient, summing up information related to substance's metabolism.
PET has a simple working principle based on the detection of gamma rays (γ) due to a positron of an atom of the aforementioned tracer that decays. In PETs two γ photons will be released in antiphase (180° from each other) and a couple of detectors catch them, and therefore resulting in a good quality detection process (they can take back any wrong event).
Nowadays, Silicon Photomultipliers (SiPM) represent a well consolidated technology for a large range of applications that need the detection of low light levels. Over the last years, thanks to the improvements in its performances (Photon-detection efficiency or PDE) and the reduction of its drawbacks (dark currents, cross talk and so on), the possible application spectrum has become wider and wider, considered as a really good option as a detector for PET [1].
Sometimes, front-electronics plays a fundamental role in meeting the relevant specifications of a detection system based on SiPMs. The front-end should operate with relative short dead time in order not to detect wrong events [2]. Wrong events are attributable to one or both annihilation photons that could scatter in the body or in the detector itself. A 511 keV photon emitted from positron-electron annihilation could undergo a Compton scattering process [3] [4] in the surrounding tissues and changes its direction. If the scattered photon is then detected by the PET detector and interpreted in a coincidence event, a false Line-of-Response (LOR) will be formed. This is called a scattered coincidence event. Scatter coincidence events give false position information and add noise to the

image, which in turns decrease the image contrast (degradation of the PET measurement).
Often, front-end electronics has a fundamental in meeting the specifications of a detection system based on SiPMs. From this point of view, the electronics front-end represents the bottleneck that limits the system performances. For example, the Single Photon Time Resolution (SPTR) which is possible to achieve with a SiPM detector of large area is strongly dependent on the contribution of the electronic noise. The electronic noise is responsible of causing statistic fluctuations of the instant when the output signal overcomes the chosen threshold (this is true in the case a voltage amplifier is implied to a SiPM). Thus, in order to exploit the best features of the detector, suitable solutions for the front-end electronics, well-tuned to the characteristics of the SiPM are necessary.
It is also important to highlight that, the front-end has to be able to preserve the intrinsic speed of the signal coming from the detector. Indeed, the relevant part of the charge released by the SiPM is the long tail of its voltage pulse with very slow time constant. The collection time of this fraction of the total charge is extremely long.
For these reasons, the most widespread methods, involved in several realizations of readout circuits for SiPMs, are based on interfacing the detector with a front-end at the same time capable of preserve the intrinsic speed of the signal generated by the detector, thanks to the well suited input impedance, and to reproduce at its output a replica of the signal, which can be applied to a fast disciriminator for the extraction of the time information or a slow discriminator in order to get a charge information, as depicted in Fig. 1.1.
There are two approaches for the read-out of the SiPM: voltage mode (Fig 1.2) and current mode (Fig. 1.3).
In case a voltage amplifier is used to readout the detector, R_{IN} cannot be large. In fact, the charge released by the detector is collected quicker if the input resistance of the front-end is reduced, because of the faster discharge of the detector's capacitance at high frequencies, due to a larger discharge current flowing into Rin. As consequence, if Rin increases, the rate of the event sustainable by the detection system is reduced because of pile-up effects. Furthermore, the timing performance is affected. To summarize, the considerations just provided recommend that a very low input resistance is preferable for the front-end electronics of a SiPM. For instance, in the case a SiPM detector is involved to read out a scintillator, low values of the input resistance are necessary to limit the variation of the voltage across the detector as the photons impinge on it. Large voltage variations on the SiPM will cause non-linearity in the energy measurements, due to the fact the micro cells subjected to avalanche breakdown at different times will have different gains (refer to [1] in order to have to the exhaustive explanation).
The amplifier should also have sufficient gain, to reproduce an output signal of adequate amplitude, in a way it can be processed by the next blocks, that are an integrator and a comparator for energy and time measurements, as depicted in Fig 1.2. These specifications, i.e. large GBWP (gain bandwidth product) and hence, sufficient gain and large bandwidth, are difficult to be reached without large power consumptions, making the voltage mode not effective in applications where very low levels of lights have to be detected, timing accuracy is relevant specification and the number of readout channels is large. On the other side, in the case the

dynamic range of the input signal is large, implying that the low voltage gain values are needed, and the specifications on the time accuracy are relaxed, the voltage mode method can be applied. As for the noise performance of the circuit, the total equivalent input voltage noise of the voltage amplifier (it is associated to a common source input transistor) is summed to the voltage across Rin, generating limitations in Signal to Noise Ratio (SNR) and in timing resolution.

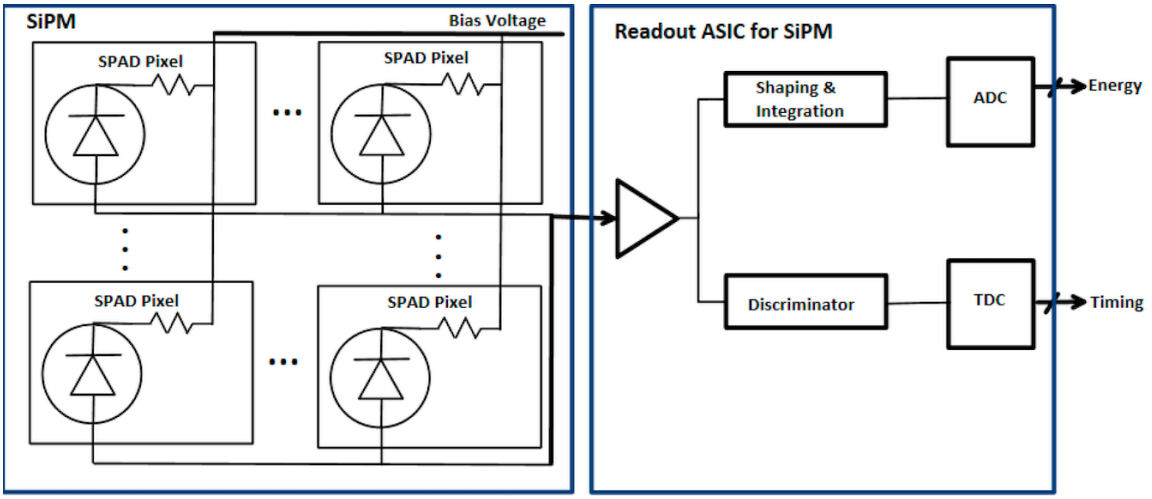

Fig. 1.1: SiPM and ASIC readout at the top

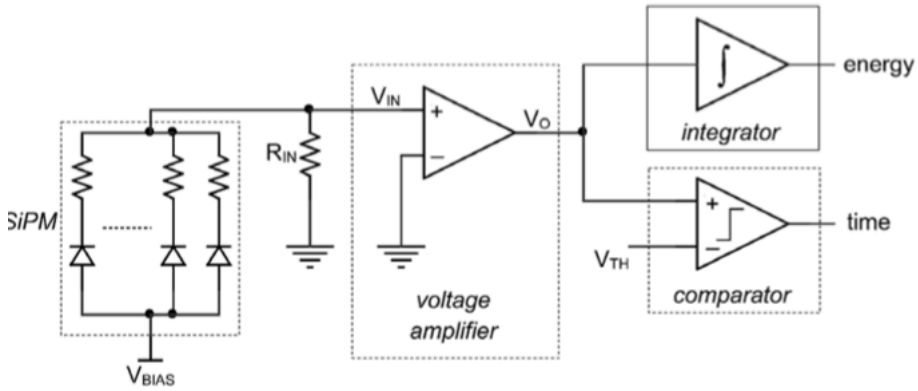

Fig. 1. 2: SiPM readout with voltage mode approach

In Fig 1.3 a current buffer with very low input impedance is coupled to the detector and exploits the advantage of small Rin. As it has been clarified in [1], a very effective readout can be modeled using a current mode amplifier that is able to read the current pulse generated by the detector at low impedance, rapidly discharging the capacitance of the detector, and can reproduce the current on high impedance, in a way that it will be processed for the extraction of the time information, thanks to a comparator connected to a Time to Digital Converter (TDC), and energy information. If it is followed by a slow path, the current buffer is connected to an integrator which charges the capacitance. The voltage at the end

of the integration will be saved on a capacitance (Sample and Hold or Track and Hold) and afterwards, the conversion will start (ADC).

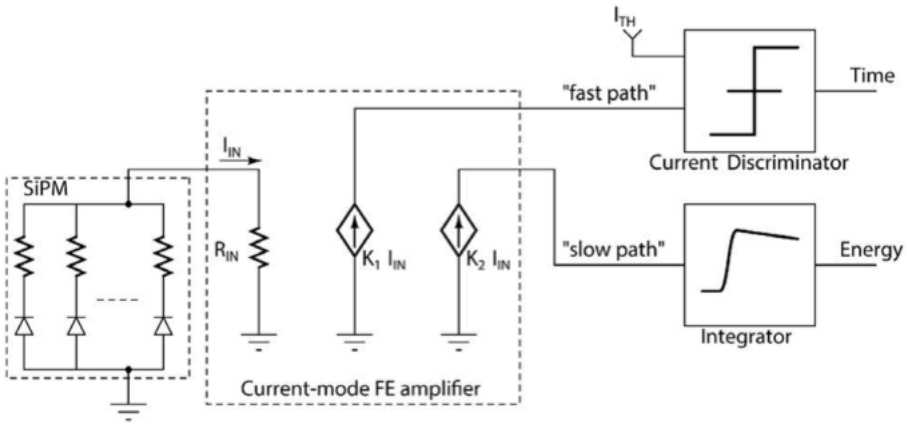

Fig. 1.3: SiPM readout with current mode analog chain

1.2 Goals/Content

This final thesis is a part of a bigger project in which a PET system will be developed. This work involves two parts: the first one focuses on designing the electronics for the front-end; the second one aims for a survey on an Analog Digital Converter converters (ADC), deeply focusing (design and test) on the electronics for acquisition and conversion of the data. The read-out chain is implemented for the whole SiPM (not for just a SPAD).
The thesis will focus on the detection system, and in particular on the front-end electronics:

1) Design and implementation of an input Preamplifier for the SiPM, placed right after the photodetector, using a 0.35 um technology from Austria Microsystem (AMS). The Preamplifier will give at the output a current that will enter into Time Branch (optimized for fast signals) and the Charge Branch (optimized for charge integration). The integrated front-end channel needs to have single photon electron (pe) resolution capability and high dynamic range (over 400 pe).

2) Design of a charge gated integrator with integration start threshold control and configurable integrating window [125 ns / 250 ns / 500 ns]. Discharge operation should be lower than 10% of the shortest integrating window.

Hereafter, we will give the requirements that have to be met in order to develop the analog chain:

Input Load Capacitance of the sensor > 500 pF

Bandwidth > 200 MHz

Silicon Photon Electron (SPE) = 0.4 pC (Pulse: Rise tau = 1 ns // Fall tau = 25 ns)

Dynamic Range (DR) = [1 photon electron (pe), > 400 pe]

SPE resolution < 5 % (Integration window = 250 ns)

In detail, the input load capacitance is the capacitance of the SPAD which is directly connected to the Preamplifier (it will give the dominant pole of the system). In the case study, the team involved in the PET's project took into account two different effective photosensors areas, the -3050CN and the -6050CN (3 x 3 mm^2 and 6 x 6 mm^2), that is the microcell of the SiPM (S13370 Vacuum Ultraviolet Multi-Pixel Photon Counter). Regarding the electrical characteristics, the first microcell has a terminal capacitance of 320 pF and the second one 1280 pF. The target, picking a SPAD capacitance equal to 500 pF, is to develop a model that could suit for the two cases. In this way, good performances will give for the -3050CN, since we are taking a margin in the capacitance of the microcell, while the performances for the -6050CN will be not enough satisfactory (they will give just a clue on how the total system work).
The bandwidth is the bandwidth of the total front-end (Preamplifier + gated integrator); the silicon photon electron, or simply SPE, corresponds to a charge equal to 0.4 pC and being the DR greater than 400 pe, the total charge on the SiPM will be:

$$0.4\ pC * 400\ pe = 176\ pC$$

Actually, we took a margin of the 10 %, considering 440 pe to have a better understanding of the problem. Moreover, the current generated by a SPE has an exponential trend with a peak of 20 uA (this value has been taken into account in laboratory test and given by the advisor).
The charge has to be read by the preamplifier because is proportional to the number of photons that the scintillator generates (Xe in our case) that is then proportional to the energy of the detected gamma ray. Thus, we read the energy (integrating the charge) because, in some occasions, it could happen that the gamma ray could suffer from scattering process (loss of energy interacting with some material), modifying its trajectory and introducing false information in the system. These rays lose energy in scatter events and, for this reason, can be detected measuring their energy.
The target for the integration is to fit the 400 pe in a time at most equal to 10 us (Tintegration has to be equal at 10 us at most, even if it is shorter around 1 us, it is better). Besides, the Treset (a Track and Hold will be involved in the digital chain) has to be around 200 ns (much smaller than Tintegration in order to avoid any dead time) and Thold (conversion time) 1 us at most for the ADC. In particular, Treset is the time in which the capacitance at the feedback of the integrator takes to charge at its initial voltage, once the integration period is over.
The SPE resolution is necessary to calibrate the system. In particular, it means that our system is able to distinguish a single photon and hence, we can calibrate the

functionality of the entire system counting individual photons. In this way, we can use a light source that generates few photons and measure the response of the system to SPE (SPE = n* LSB). There are other ways to calibrate, but the one we used (SPE resolution less than 5 %, integration window equal to 250 ns) is the best. To have a better understanding of the SPE resolution, we refer to Fig. 1.4. In fact, this procedure is for the calibration of the system and it allows to see the detected signal for one photon, two photons and so on. In this way, when I read a signal at the output of the integrator that has 10 LSBs, if the SPE = 2 * LSBs, the detected signal contains 5 photons.

Fig. 1.4: Spectrum gamma emitter (the graph has to be read from right to left)

The integrator stage is the best way to measure charges (number of photons) of a signal. There are other options to do it, but they often introduce errors or depends on the pulse shape (case FLEXTOT 3.2).
Regarding the survey we went through two different branches (they are both connected to the Preamplifier) in order to see a translation from the entering charge of the Preamplifier into timestamp and charge digital data:

- Review of multichannel ADC converter and proposal of an ADC structure for Charge Branch conversion.
  The ADC has to provide 9 effective bits output for each triggering channel. The conversion time must not exceed 1 us, taking into account frequency limitations in the 0.35 um technology kit being used in the test design.

Basically, what we want to realize is a model of the ADC in order to find the specifications for who is going to design it in reality. The goal of the system is to see how the system works using a SAR. Indeed, most of the chip in the market exploit a Wilkinson ADC [5].
Finally, it remains to give two important considerations about the working environment, Cadence. The HIT-kit is the kit we used in Cadence to denote every components of the blocks. The process is the CM35B4 from AMS 0.35 um CMOS process (that is the technology library).

1.3 Master Thesis structure

The second chapter includes some needed background in order to understand and get in touch with the most important points of the thesis. Indeed, we went through the working principle of a PET, with its components, and studied in detail how a photomultiplier is made and how it works. In particular, the sensor SiPM incorporated in the PET is presented, focusing on its features in order to be aware how it impacts on the read-out electronics.

Regarding the third chapter, it addresses the state-of-the-art of the chips available in the market in order to understand how the problem is faced and then solved. This work will be worth in order to conceive the motivations that are the roots of the project itself and the project challenges we face to. Furthermore, it is worth explaining how we have solved the problem itself respect to the anyone else and how we tried to improve the state-of-art. Specifically, three chips (TOFPET, FLEXTOT PETIROC) are compared between each other, centering more on the TOFPET that is the one the team, involved in the PET's project, used for the whole machine.

The fourth chapter focuses on studying the methodologies and then on the real implementations of the Preamplifier. Precisely, firstly we describe how we implement the input block (single photo electrons detected by the SiPM); in a second moment, we will study an input preamplifier studying its internal topology; then, after considerations about the DC simulation, we will define the DC Nulling block step by step, starting from the differential internal configurations up to the assembled Current Collector (CC). For each component, we will study its behavior in Cadence through simulations in order to link them with the computations made in the project phase.

In the fifth chapter, the main parts of the developed slow chain are described. In particular, the signals controlling the switches of the system are displayed (in

Cadence) to show the different working steps of the chain; then, different options for a charge gate integrator have been given in order to show advantages and disadvantages of the structures with respect to the other, and hence, explaining how we ended up picking the one having Monticelli' s output. Then, we developed the interface of the system. First, a Sample and Hold is studied in detail and a topology for the OTA (in this component) is presented. Secondly, the ADC SAR is explained in deep. Each part of this ADC has been designed, highlighting the functions.
At the end of each section, as we did in the fourth chapter for the Preamplifier, we showed the results we obtained in Cadence in order to compare the theory with the reality.

In summary, in the last chapter, the results in Cadence, once we merged the Preamplifier to the slow chain, are presented. Basing on the results, we will come up with the conclusions.

Chapter 2

Background

The positron emission tomography (PET) is a nuclear imaging technique which uses annihilation gamma photons from positron decay to generate three-dimensional functional images of the human body. Mainly, PET's applications are clinical research, clinical oncology, and brain function analyses.
PET stands out from any other body imaging techniques because it can provide information about the metabolism of the body. In fact, this machine utilizes the emission from radioactive blend (tracers) in order to localize tissues where a specific cancerogenic cell is allocated.

2.1 PET manners

2.1.1 Working principle at a glance

Preceding the scan conducted by a PET [6], an irradiated compound isotope (typically flu-orodeoxygluocose or knowing as FDG) is introduced into the human body. The aforementioned compound is incorporated into an active molecule; in a second moment, the compound is assimilated by tissues affected by a hypothetical disease and the patient is ready to be placed into the machine.
When a radioactive atom, of the tracer injected, declines, a positron is emitted from the nucleus and, after covering a distance between 0.1 and 1 mm in tissue, a participles' collision takes place. Basically, the same positron merges with an electron and a couple of gamma photons are released in opposite directions (180° apart from each other) as shown in Fig. 2.1. The coupled photons get to sensors, normally forming a ring of detectors, generating an emission of visible light that is perceived by photomultiplier tubes (PMTs) or silicon photomultiplier (SiPM) as it can be seen in Fig. 2.2.
The PET scanners detect the coupled emitted photons in order to determine the line of response (LOR) along that the annihilation occurs. Indeed, each detector has to establish the energy, position, and time of arrival (ToA) of the incoming gamma photons. All the data reaches a coincidence unit, in which at each detected couple of photons a specific electrical collision will be correspond to. On this purpose, firstly photons with the same energy will be selected while photons do not arrive coincidentally (in a time window roughly of few ns) will be ignored. At this point, the LORs are created based on the information from the photons' position.
After many thousands of LORs, a tomographic 3-D image of the patient is realized, knowing where the compound is concentrated and so the tumoral cells.
It is interesting to highlight that, in a PET system, the energy information is important as much as the time information. The energy information allows refusing any events that are not at 511 KeV (they are not gammas or they suffered from scattering that has modified their directions), while the time information

permits to select the gamma couple associated to an event (back to back) and, also, to measure the temporal distance between the gamma rays in order to estimate the origin zone of the event. However, this mechanism does not yield a perfect position in the space because we would need a temporal resolution equal to some picoseconds. On the other hand, it allows us to refuse many gamma rays that introduce error in the image reconstruction.

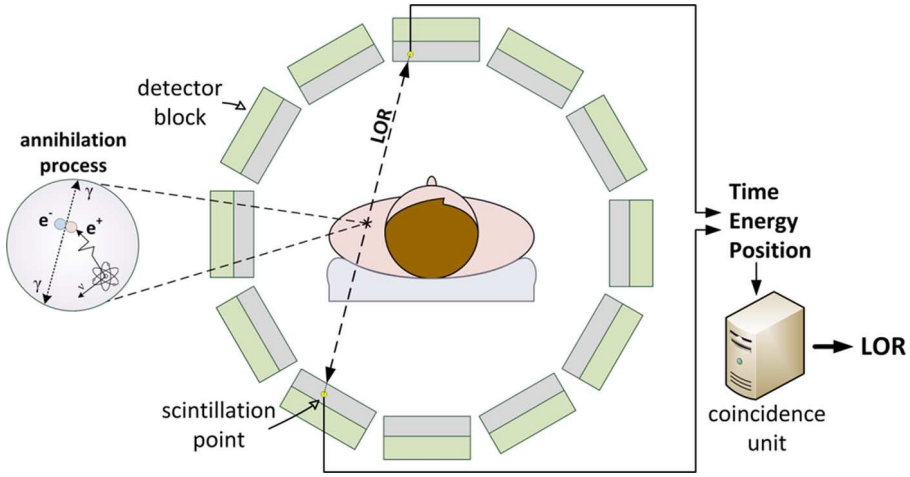

Fig. 2.1: PET working principle

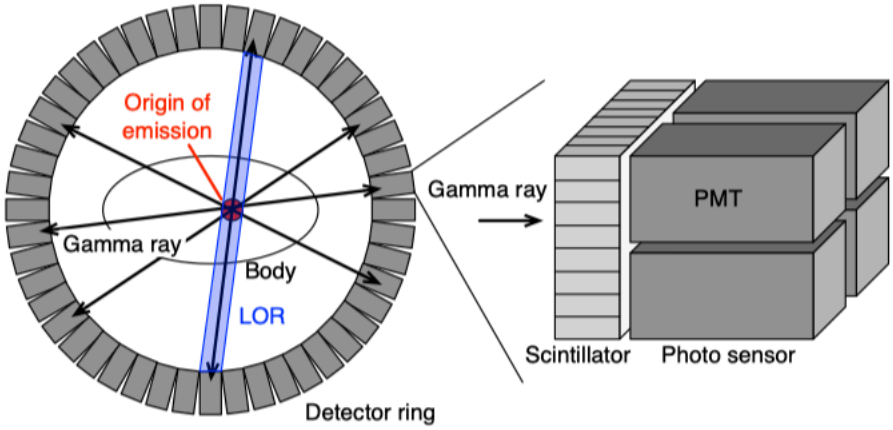

Fig. 2.2: PET scanner

2.1.2 Positron Emission

In proton- rich nuclei, a proton can decay to a neutron, a positron and a neutrino via the process:

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

which is identified by β^+(decay). β^+decay just occurs inside nuclei when the daughter nucleus has a greater binding energy than the mother one.
The emitted positron can lose its kinetic energy by Coulomb interactions with electrons and, after a while, it reaches its thermal energy. Then the positron annihilates with an electron. Then annihilation produces a pair of photons (gamma)

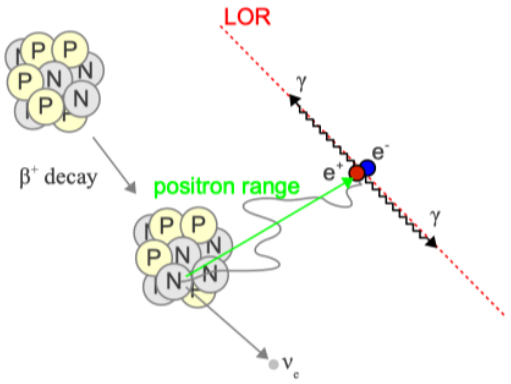

Fig. 2.3: Positron emission from a β^+ decay and its annihilation with an electron, and the definition of positron range and LOR

with a characteristic energy of 511 keV each, moving in opposite directions. This is the result of energy and momentum conservation. Both positron and electron have masses of 511 keV.
A PET scanner, at the same time (within a coincidence window of nanosecond range), detects the two gamma photons with two opposite detector sections, and forms a LOR which contains the information of the annihilation position. A set of LORs detected by the PET scanner is the used to reconstruct the positron emission occurring place.
The distance from the emission point of the positron to the annihilation is known as the positron range (see Fig. 2.3). It depends both on the energy of the emitted positron and the materials close by. The positron range produces an inherent error to the data collected by PET scanners and cannot be corrected. It is worth underlying that the total momentum of the annihilating positron and electron could not be zero, this variation generates an angular uncertainty [7] between the two gamma rays.

2.1.3 Radiation Observation

One of the most important part of an imaging system is the radiation observation. This process is composed of a scintillation material (single-crystal scintillator thanks to his high density and atomic number that experiences better detection efficiency), that is comprised in each detector of the ring, and photon counting sensor. Indeed, silicon is characterized by a low stopping power, due to the its low density and atomic number, making it just capable of absorbing radiation up to few tens of KeV. A scintillator is a material with the capability to take in ionizing radiation, in our case gamma rays, and to convert fraction of the absorbed energy

into visible photons which is sensed by a photodetector and converted into an electrical signal.
At this point, it is important to mention the two main processes through which a radiation is absorbed: photoelectric absorption and Compton scattering.
Photoelectric absorption is identified by an incident photon which is completely absorbed in the material. The photon has enough energy to generate a photoelectron from an inner shell of an atom of the material:

$$E_{loss} = E_k = E_{ph} - E_{bind}$$

In the formula above, E_k (kinetic energy) is equal to the difference of the incoming energy and the binding energy of the extracted photo electron. The vacancy, due to the exchange of energy by means of radiative process (fluorescence) or non-radiative process (Auger effect), will be recovered by some electrons.
On the other hand, Compton scattering (called incoherent scattering as well) is an interaction process, characterized by the exchange of only a fraction of the energy from the incoming photon. In this mechanism, the gamma photon is dispersed by bound valence electrons. The scattered photon energy is the difference between the incident photon energy and the emitted electron one. This fact suggests that the probability of Compton interaction depends on the number of valence electrons of the medium. Another parameter to consider for this process is the momentum of the incoming photon. As a result, its energy transfer and interaction probability depend on the collision angle (θ), handled by Klein-Nishina [8] distribution and can be calculated as follow:

$$E_{loss} = E_{ph} * (1 - \frac{1}{1 + \frac{E_{ph}}{m_o c^2}(1 - \cos\theta)})$$

The interaction processes are shown in Fig. 2.4.

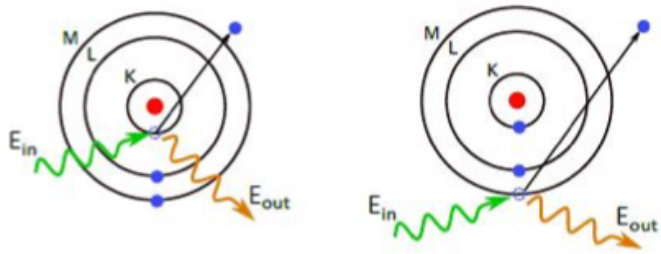

Fig. 2.4: Photoelectric process and Compton scattering

In the end, we want to illustrate the incidence of the process which is the function of the incoming radiation energy and the atomic number of the absorbing material (see Fig. 2.5). Having a look at the figure below, we can notice, in the energy range of gamma photons (511 KeV), the dominant process is the Compton scattering.

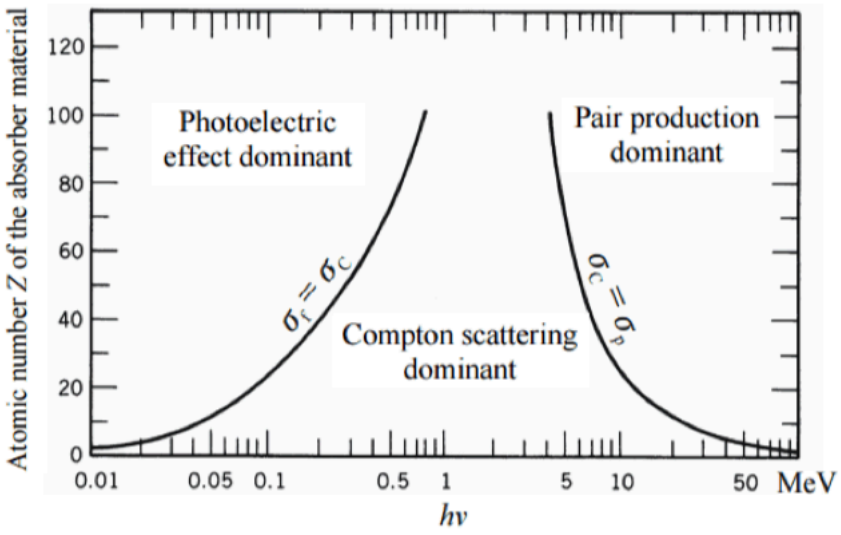

Fig. 2.5: Incidence of interaction process as function of the incoming radiation energy

2.1.4 Timing Performance

Another important key point of the PET is the timing resolution (difference between the times in Fig. 2.6). It is the statistical timing fluctuations or unpredictability due to the scintillator's and the photosensor's timing characteristics. The coincident detection of two detector is shown in Fig. 2.6. The output from each detector is distinguished by a threshold from detector noise or scattered gamma events which has low energy and emitted to a data acquisition system (DAQ).

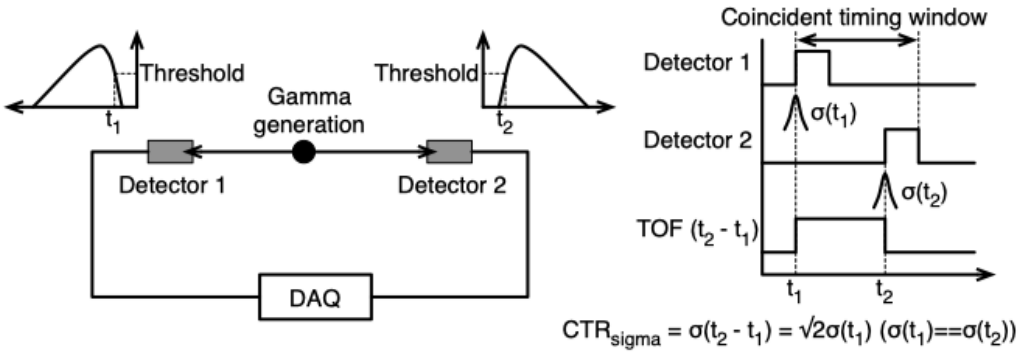

Fig. 2.6: Detecting coincident events in two detectors

Since the timing resolution is basically the variability in the signal arrival times (time-of-arrival or simply TOA) for different events, it requires to be properly measured in order to detect coincident events to distinguish true events from false ones. True coincidences happen when both photons from annihilation event are detected coincidently, and no other event is observed within the coincident timing window. There are various cases of false events, as shown in Fig. 2.7. A scattered

coincidence occurs as soon as a gamma photon is scattered before the detection. Since the direction of the gamma photon has changed during the Compton scattering process, the resulting coincidence will be registered to the wrong LOR. In addition to this, random and multiple coincidences generate false events. They are similar to each other, except that, in the case of multiple coincidences, three events from two annihilations are detected within the coincidence timing window. Scattered and random coincidences cannot be distinguished from true coincidences. However, they can overlap to true coincidences, adding statistical noise.

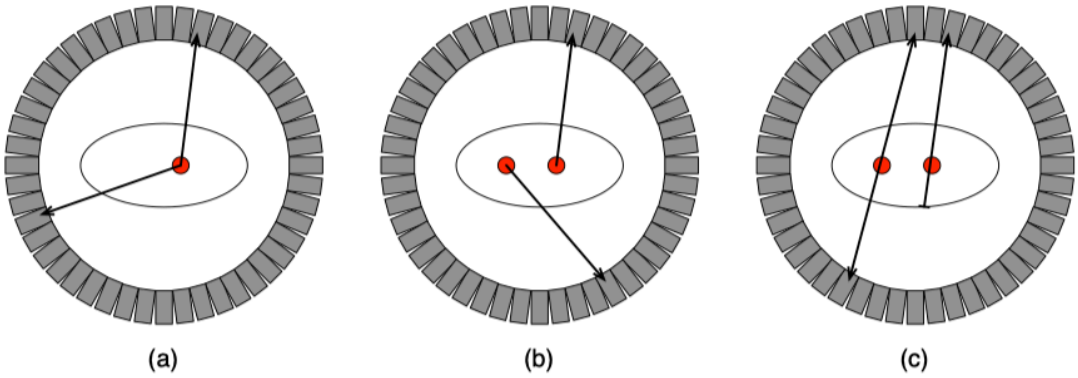

Fig. 2.7: In order by left to the right there are false events, random events and multiple events

The accuracy of the coincidence detection is defined as coincidence time resolution (CTR), as it can be seen in Fig. 2.6. Assuming that each detector has the same timing uncertainty to another, $\sigma_1(t)$, CTR is computed in the following way:

$$CTR_{sigma} = \sqrt{2}\,\sigma_1(t)$$

$$CTR_{fwhm} \approx 2.35\sqrt{2}\,\sigma_1(t)$$

Good timing resolution of a PET detector, besides reducing random coincidences, can be involved to evaluate the annihilation point between the two detectors by measuring the arrival time difference of two photons [9] [10] [11]. This PET is known as time-of-flight PET (TOF PET), which will be presented in detail in chapter 3.
The advantage of estimating the location of the annihilation point is an improvement in terms of signal-to-noise-ratio (SNR), as a result from the acquired image, due to a reduction on noise propagation during the image reconstruction process.

2.2 Photosensors

Photosensors are able to detect light coming from a scintillator. The scintillator itself emits light from gamma photons previously absorbed.
This section introduces two useful photosensor, SiPMs and PMTs, for PET application, explaining the reason why we focus on SiPM (due to its benefits respect to PMTs). In detail, the SiPM's structure and electronic model is explained. At the end, the Hamamastu S13370 (the SiPM used for the PET machine) is introduced.

2.2.1 SiPM versus PMT

Silicon Photomultipliers offer a highly alternative to PMTs because they exploit a solid-state (mainly silicon) technology. The advantages in using SiPMs instead of PMTs are listed:

- Higher internal gain, that helps out to reduce the impact of the electronic noise due to the analog front-end.
- Bias voltage range ranging between 25 V and 80 V (the voltage depends on the breakdown voltage of the photodetector) for SiPM, whereas for PMTs thousands of volts.
- Magnetic compatibility, which offers many possibilities in medical application [12] [13] in which PMTs could not be exploited.
- Fast and uniform response makes SiPMs well suited for highly demanding timing applications, for instance TOF-PET [14].
- Mechanical robustness and compact size allow to design much more compact modules (the costs are reduced in this way).
- High degree of fabrication uniformity offered by technological development of modern semiconductor facilities.

2.2.2 SiPM structure

A silicon photomultiplier integrates a parallel array of photon detectors, called microcells, as shown in Fig. 2.8. In the same figure, the physical implementation of the microcell is illustrated.

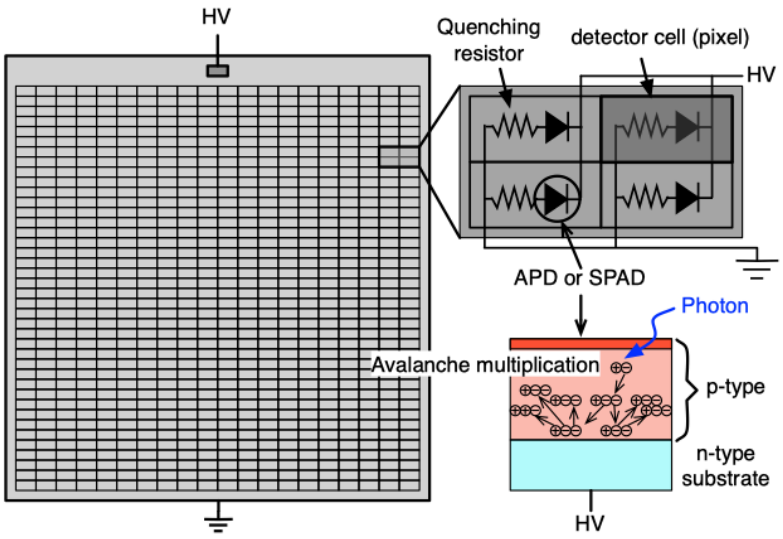

Fig. 2.8: SiPM' s structure (equivalent circuit)

Each microcell is made by a Geiger Mode Avalanche Photodiode (GM-APD, or in another way Single Photon Avalanche Diodes, SPADs) with an integrated quenching element. SPADs are devices that operate above the breakdown voltage: an electron-hole can be generated in the depletion region due to the absorption of a photon. These carries will be accelerated by an electric field, creating a self-sustaining current (impact ionization).
In order to stop the avalanche process, SiPMs feature a passive quenching resistor in series to the avalanche diode. This resistor is able to generate a voltage proportional to the avalanche current and to reduce the voltage across the junction below the breakdown, and hence, stopping the avalanche. Since no current flows anymore, the diode depletion capacitance is charged again to the bias voltage, and a new avalanche can be triggered. In Fig. 2.9 the equivalent circuit of the microcell and its operation cycle is displayed.

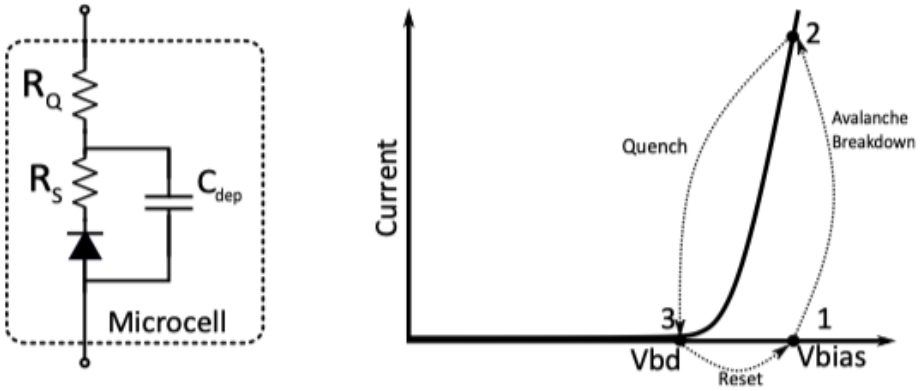

Fig. 2.9: Equivalent circuit for a SPAD and its operation cycle

At this point, it is important to underline that each microcell presents a bi-stable behavior. In fact, the avalanche is not proportional to the number of generating events, being the cell no more sensitive until the bias voltage has restored on the detector capacitance, and no further events will produce signals. The avalanche breakdown is the result of the electron-hole pair (EHP): the EHP can be originated by the absorption of a photon, thermal agitation and tunneling effect. These effects cannot be distinguished from each other, thus resulting in the same avalanche current. We need to point out that thermal agitation and tunneling effect are undesired effects and they are treated as noise sources (dark count events).
In Fig. 2.10 [15], a single photoelectron (SPE) spectrum has been obtained illuminating the device with a low light intensity source. The peaks corresponding to 1 to 9 pe can be clearly seen, showing the good SPE resolution of the SiPMs.

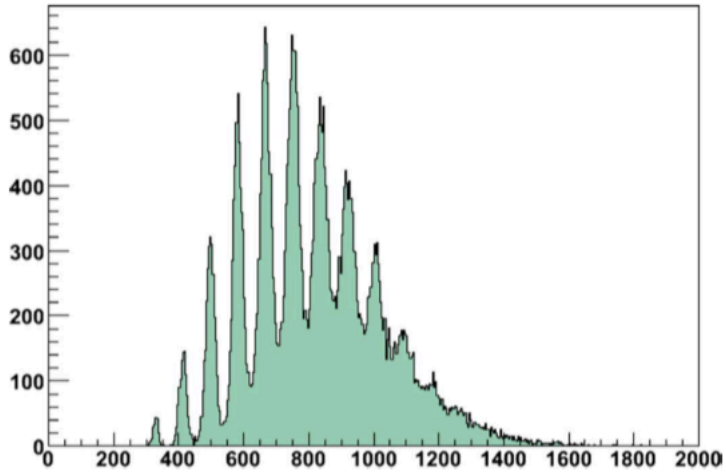

Fig. 2.10: Single photoelectron spectrum obtained illuminating the SiPM with brief, low-level light pulses

2.2.3 SiPM's electrical model

Referring to Fig.2.11, the electrical model [16] of a SPAD is shown in order to understand the electrical behavior of the detector.
In the figure, Rs is the detector series resistance, Cdep is the detector depletion capacitance (around few fF), Vbd is the breakdown voltage, Rq is the quenching resistance (KΩ), and Vbias is the voltage bias (higher than Vbd) that causes the high electric field that triggers the avalanche.

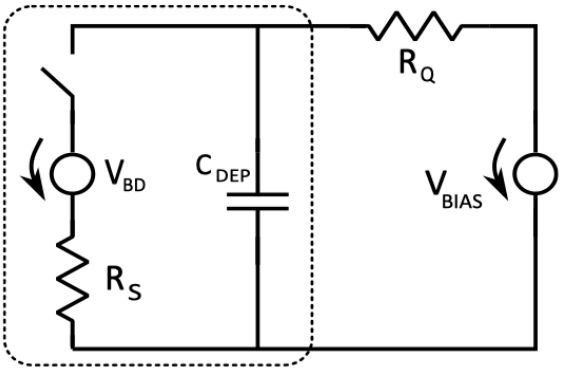

Fig. 2.11: Equivalent circuit of a GM-APD with Rq and external bias. The switch models the turn-on (photon absorption or dark event) and turn-off (quenching) probabilities.

When no avalanche is triggered, the switch is opened. In this condition, the voltage across the depletion capacitance of the detector is equal to Vbias, and no current flows in the circuit.
As soon as the avalanche is triggered (the switch closed) a current I_{INT} starts flowing in the sensor (it swiftly rises):

$$I_{INT} = \frac{V_{BIAS} - V_{BD}}{R_S}$$

In the formula above, $V_{BIAS} - V_{BD}$ is simply called overvoltage or Vov.
This current discharges Cdep at the SPAD cathode, in a way that the excess voltage decreases to zero with a time constant:

$$\tau_{fall} = C_{dep} * (R_S // R_Q)$$

The overall bias voltage never decreases below the breakdown voltage. For this reason, the avalanche is not quenched and current continues to flow through the device. The current final value is roughly:

$$I_{INT} = \frac{V_{BIAS} - V_{BD}}{R_Q}$$

If this value is high enough, the avalanche is self-sustaining. On the other hand, if the current is small, there will be a probability that, after a random time, all the carriers have left the space charge region, so avalanche multiplication can no longer take place and the avalanche will self-quench. The boundary between high and low value is about 100 uA.
When the electric field across the junction reduces, the avalanche will stop, and no more current will flow. Basically, this effect occurs when the switch opens, and the voltage on Cdep is reset to Vbias with the slower time constant:

$$\tau_{reset} = C_{dep} * R_Q$$

The reset time constant can be more than two order of magnitude greater than the quench one.
In the case Vd can be approximated with Vbias, the microcell will be ready to trigger another avalanche phenomenon. The current provided by the detector, according to the model, is shown in Fig. 2.12 [17].
However, the exposed model has limitations related to the missing of the loading effects of the surrounding microcells (indeed any effects introduced by the other microcells, composing the SiPM, have been neglected). For this reason, a more accurate model (not as simple as the one we proposed) is indicated in the paper [18].

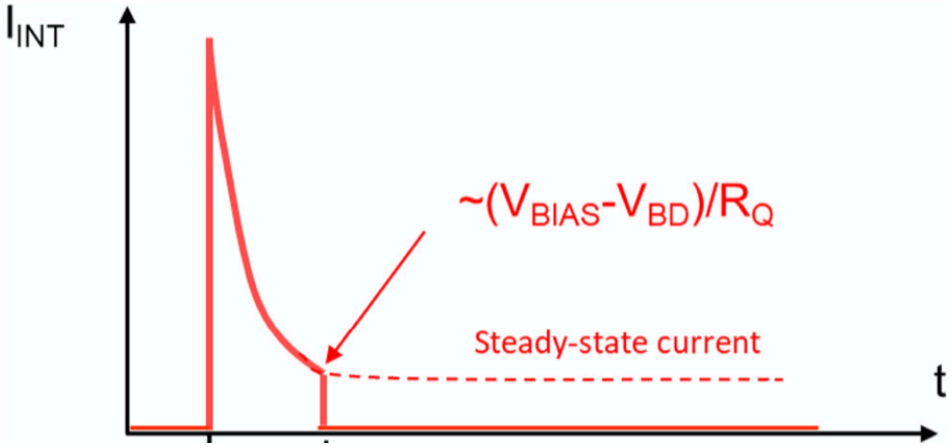

Fig. 2.12: Pulse waveform of SiPM, related to the electrical model

2.2.4 S13370 (VUV4 generation) Multi-Pixel Photon Counter

On the market, there is a variety of SiPM available from different manufactures, involving Hamamastu, SensL, Fondazione Bruno Kessler (FBK), and so on. Nevertheless, the detector that fits the specifications of project is the S13370 VUV4-MPPC (Vacuum Ultraviolet Multi-Pixel Photon Counter) made by Hamamastu. Therefore, in this section we will describe the main parameters and technological advantages offered by this SiPM.
VUV4-MPPC are detectors able to detect light down to 120 nm, covering scintillation wavelengths of liquid xenon (LXe) and argon (LAr) with cryogenically compatible, ultralow-RI packaging options. They feature [19]:

- high sensitivity for VUV
- stability for cryogenic temperature
- suitability for detection of LXe or LAr scintillation light
- can be operated at low voltage (< 60 V) in LXe
- single photon counting capability
- PDE (photon detection efficiency) close to 25% at 178 nm
- Gain larger than 2 x 10^6

Regarding the scintillators [20], they are used for dark matter search or neutrino less double-beta decay experiments. In particular, LXe is characterized by peak emission wavelength to 178 nm, a temperature of 165 K and they are directly detected by VUV photodetector. On the other hand, Lar has a peak emission wavelength of 128 nm, a temperature of 87 K and, as the LXe, they can be directly detected by VUV, or they can be caught indirectly (after WL-shifter) by UV/blue photodetector (around 420 nm).
A crucial point is the stability for cryogenic temperature. VUV-MPPC has a metal quenching resistor to maintain its pulse shape at low temperatures. The metal resistor has 1/5 the temperature coefficient of the poly-Si resistor, so its resistance has excellent stability against temperature changes. Moreover, VUV-MPPC with metal quenching is capable to maintain its pulse shape at both room and low temperatures, but MPPC with poly-Si resistor has longer pulse tails and recharge time at low temperatures.
Once we provided an overview about VUV-MPPC, it is the time to give some characteristics about the specific case study: S13370 (it is a basic VUV4 -MPPC). Two different S13370 are being taken into account for this project. The first one is the S13370-3050CN, characterized by an effective area of 3 x 3 mm^2, and the S13370-6050CN with an area of 6 x 6 mm^2. Both of them, have a pixel pitch of 50 um and a package made by ceramic. Besides the structure, it is important to underline the electrical and optical characteristics. In these particular cases, they both have gain equal to 2.55 x 10^6 (the gain of a SiPM is defined as G and it can be computed referring to this formula G = (Cd * Vov)/q), the same PDE that is 35 % and a breakdown voltage 53 $\pm$ −5 V. The first SiPM have a terminal capacitance of 320 pF, while the second one a capacitance of 1280 pF. As we said in the paragraph 1.2, the reason why we considered in the case study a capacitance of the single SPAD equal to 500 pF, is because we want to check the feasibility of the both microcells attached to the system we developed for the read-out. The system we have developed will work when the S13370-6050CN is used in the system, while the performances of the S13370-6050CN will not be satisfactory.

2.3 Photosensor Parameters

The understanding of some parameters [21] in photomultiplier, principally SiPM, for a PET is important. Indeed, there is a relation between the SiPM performance terms and the technological implementation of the parts that composed the PET machine. For this reason, the following parameters are presented in this section: breakdown voltage, PDE, dark count noise, correlated noise, SPTR and gain.

2.3.1 Breakdown voltage

The breakdown voltage, or simply called Vbd, is the minimum required voltage needed to the self-sustaining avalanche multiplication in Geiger-mode avalanche photodiode. Basically, it is the minimum voltage required in order to obtain a current pulse at the output of the detector. Examples of pulses are depicted in the following figure, 2.13 and 2.14:

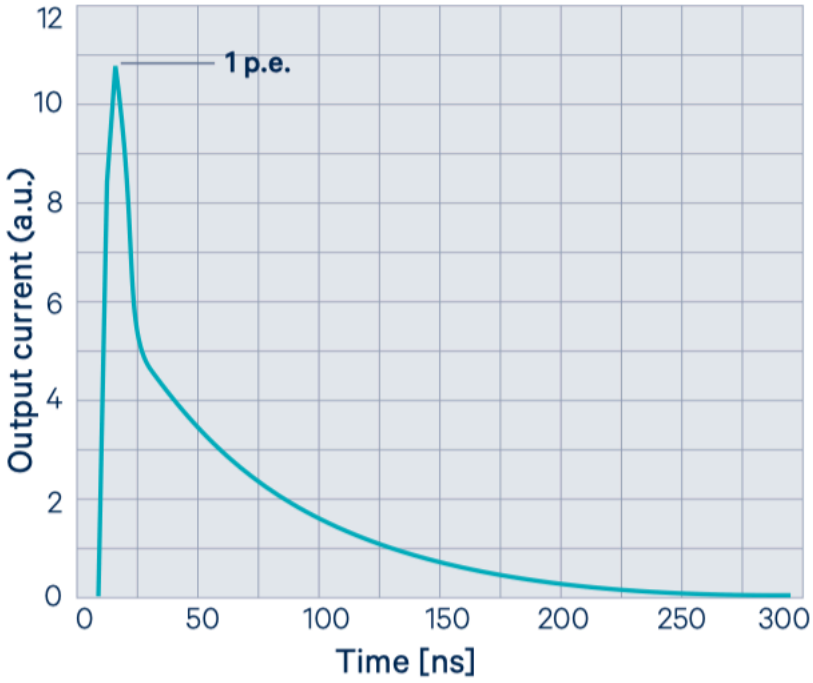

Fig. 2.13: Current pulse due to a photon absorption (single-cell photon). From the picture, the pulse has a fast-rising edge and slower recovery edge. The amplitude is 1 p.e. (photoelectron)

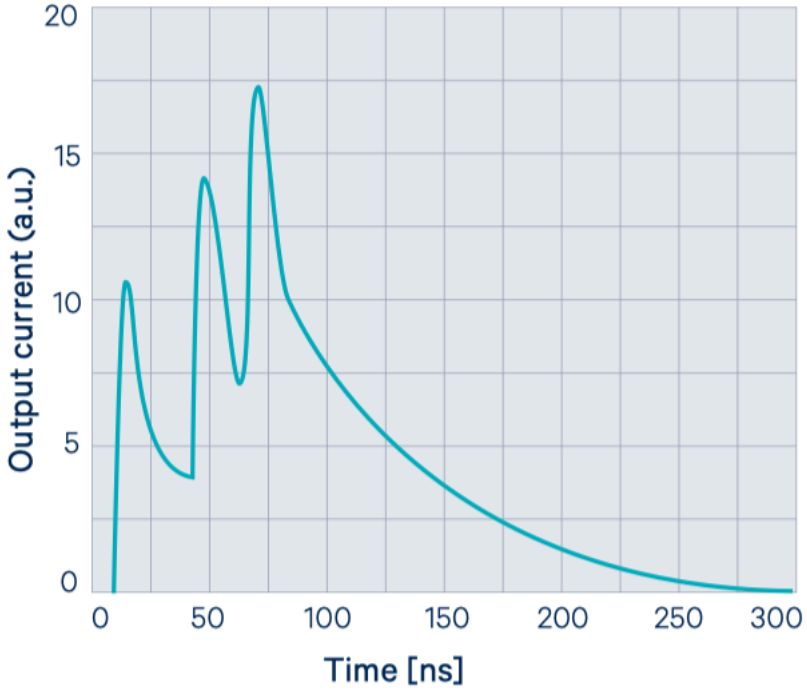

Fig. 2.14: Three micro-cells are impeded at different times. The output pulse is obtained summing up the three signals

It is important to underline, in the case of Vbias = Vbd the detection efficiency and the gain (as it is expressed in the formula in section 2.3.8) are zero. When Vbias > Vbd, output current pulses are actually noticed. The difference between of Vbias and Vbd is defined as overvoltage (Vov). The following formula is valid:

$$Vov = Vbias - Vbd$$

2.3.2 Photon-detection efficiency (PDE)

The Photon-detection efficiency, or simply PDE, is defined as the ratio between the average number of detected photons and the number of the incoming ones.
It gives the probability that visible light, produced by the scintillator (single photon), arriving on the SiPM surface is detected, yielding an output pulse.
PDE is computed combining QE, quantum efficiency, FF, fill factor, and triggering probability $P_{avalanche}$:

$$PDE\ (Vov\ ;\ \lambda) = QE(\lambda) \times\ FF\ \times\ \mathrm{Pt}\ (Vov\ ;\ \lambda)$$

QE gives information about the probability that a photon impinging on the SiPM is transmitted to the silicon, then absorbed in the surface and converted in an electron/hole pair.
P_t expresses the probability that the electron/hole pair successfully starts a self-perpetuating avalanche process and then an output current pulse.
FF describes the amount of the SiPM sensor surface area that reacts to light. It takes into account that each micro-cell in the SiPM has some dead area on its parameter.

2.3.3 Primary noise (Dark Count rate)

Current pulses are produced in SiPMs even if there is not a source of light. This effect is due to thermal agitation or tunneling when non-photon carries are generated. During the quiescent mode, if a carry originates inside the active region of the photodiode an avalanche is triggered (with a probability Pt) and the pulse can be seen. This is called dark event. The number of dark events per unit time is the dark count rate (DCR).
In a single GM-APD a dark event has the same shape of a photo-generated one. Indeed, in both cases, the output pulse has 1 p.e. amplitude, see Fig. 2.13.

2.3.4 Dead time

It is the time required by a SiPM in order to reestablish the output current to zero after a detection of a photon.

2.3.5 Correlated noise

In SiPM there are other two sources of noise, in detail, afterpulsing (AP) and optical crosstalk (OC). These events arise from an existing current pulse (photo-event or dark event) and therefore they are called correlated noises.

2.3.5.1 After pulsing

Afterpulsing is an undesired avalanche multiplication due to carriers trapped in silicon defects that are emitted during the recharge phase of the photodiode.
Basically, the net effect is the observation of another current pulse right after the original one, as shown in Fig. 2.15.
The probability, this event occurs, depends on aforementioned dead time, the Vov, linearly, and the cell size quadratically.

2.3.5.2 Optical crosstalk

Crosstalk features photons released in avalanche multiplication and then re-absorbed in the cells nearby or in the inactive region of the same cell, inducing current pulses.
We have to distinguish between direct-OC and delayed-OC though.
Direct-OC takes place when an ejected photon arrives at the active region of another cell, triggering an additional avalanche at the same instant of the first avalanche. The result is the double pulse, as depicted in Fig. 2.16.
Delayed-OC involves photons which are re-assimilated in the inactive regions of the SiPM. At this point, the electron (hole) has to propagate to the active region of a cell before triggering an avalanche. The delayed crosstalk occurs then right after the original one, as shown in Fig. 2.16 by the third peak.

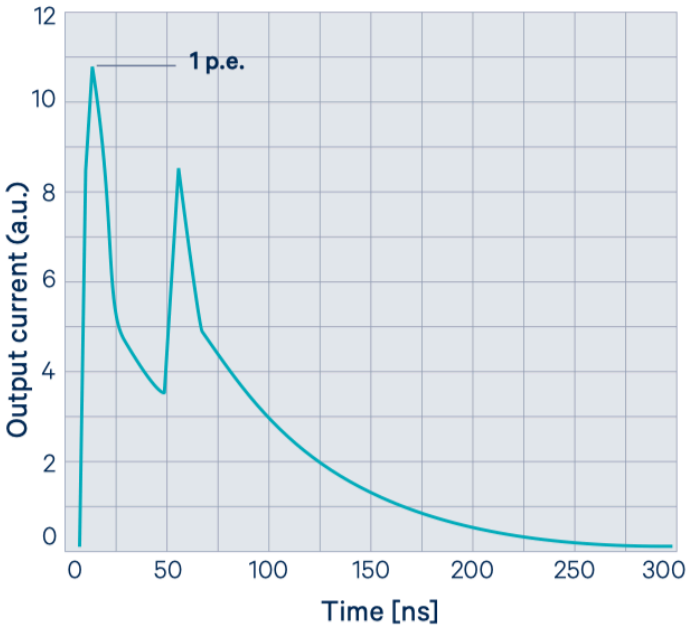

Fig. 2.15: The same micro-cell fires twice in the recharge phase due to a carrier emitted from a silicon trap. The amplitude is less than 1 p.e. because Cd has not restored to Vbias when AP happens

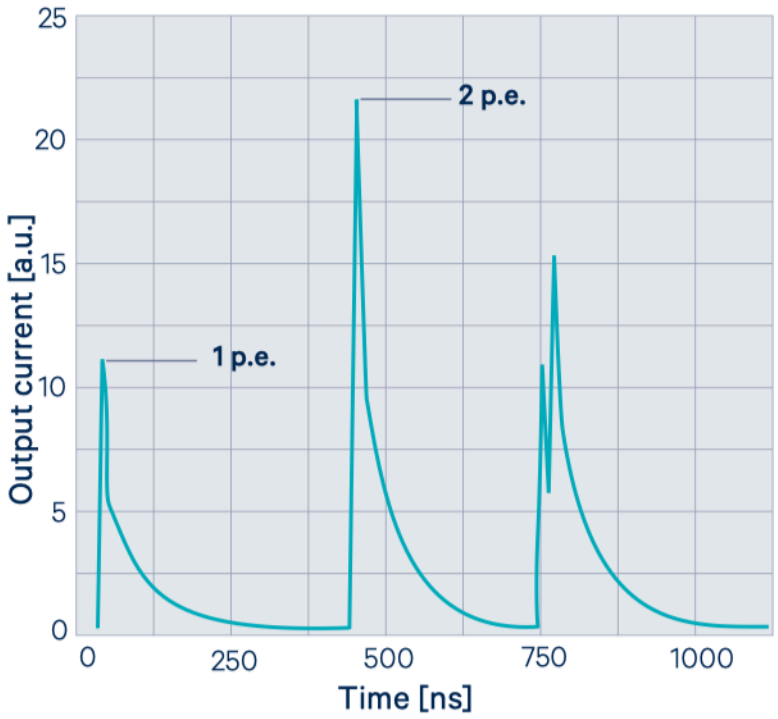

Fig. 2.16: In this graph we have: a single cell signal (1 p.e.), a direct crosstalk signal (2 p.e.; two cells firing at the same time), and a delayed crosstalk (another cell fires after the first one, typically few ns)

2.3.6 Single-photon Time resolution (SPTR)

SPTR represents the precision of arrival time evaluated for a single detectable photon, colliding randomly on the detector surface.
The timing performance depends on the SiPM parameters and on the read-out chain as well. Indeed, it is the statistical jitter of the delay between the time reference of the light pulses, most of the time a triggered signal of the main source, and the measured arrival time of the SiPM signal. There could be a variance due to sources of uncertainty in timing measurement introduced by the detector and the read-out.

2.3.7 Energy Resolution

The energy resolution is determined as the ratio of FWHM of the full energy peak and the energy value of the full energy peak maximum. A SiPM needs to have a minimum number of photo-detecting cells, pixels, in order to compute the gamma photons' energy. Nevertheless, there is a trade-off between FF and energy

resolution. In fact, SiPM should have an adequate number of pixels to accomplish linearity for photon counting without saturation effect. On the other side, it needs to have a quite high FF though and, increasing with pixel size, these parameters are therefore in contrast between each other.

2.3.8 Gain

The gain of a SiPM sensor is defined as the number of carriers contained in the single cell current pulse [22] since this is the number of carries generated during the avalanche in response to an absorbed gamma photon.
The gain of a microcell (and therefore the sensor) is defined as the ratio of the charge (the SPE charge given in the system specifications corresponds to the one in nominal bias conditions of the Hamamastu SiPM) from an activated microcell to the charge of an electron. The gain can be calculated from the overvoltage Vov, the microcell capacitance Cd (the gain is given by the discharge of the capacitor from Vbias = (Vbd + Vov) to Vbd), and the elementary charge (q) of an electron:

$$G\,(Vov) = (Cd \times Vov)/q$$

Chapter 3

State of art

In this section, we provide the different systems in the literature, in particular focusing on the following chips TOFPET2, FLEXTOT PETIROC on the market. We will go through the problem in order to give the motivations of the work, to explain the challenges we had to face in the project phase and how we solved them respect to the others, and finally explaining how we tried to improve the state of art.

3.1 TOFPET2

The TOFPET2 (Time of Flight PET2) ASIC, designed by PETsys Electronics, Oeiras, Portugal [23] [24], is a new 64 channel chip for the readout and digitalization of signal from fast photon detectors in applications where a high data rate and fast timing is required. The scheme of one channel is shown in Fig. 3.1. The selected preamplifier provides a signal to two post-amplifiers and a pulse-charge integrator. The post-amplifiers, that are specialized for time resolution and pulse triggering, have an adjustable transimpedance gain and supply voltage signals to two discriminators. In reality, there is just a preamplifier with a current mirror at the output that is connected to the discriminators and the charge integrator (the Fig. 3.1 is to clarify the fact there are two separated channels, one for timing (TDC) and one for energy (ADC)).
Considering the discriminators, one is used for timing measurement with a programmable threshold down to a few photoelectrons, whereas the second one has higher threshold and is used for chip rejection (dark counts) of low amplitude signals (the threshold's signal is above the level of the dark count of the SiPM; the ADC is activated only when the signal is greater than threshold itself), avoiding dead time. All thresholds are separately configurable for each channel.

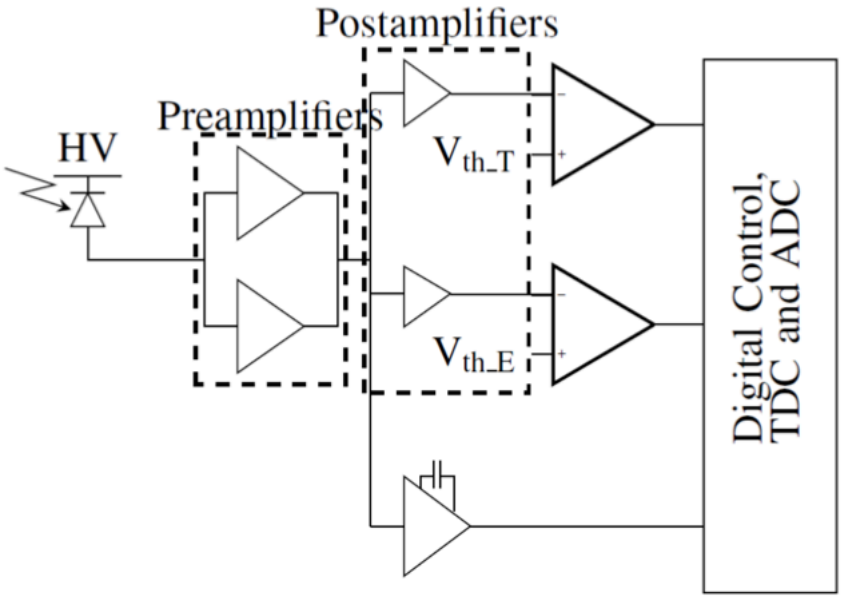

Fig. 3.1: Top level scheme of the single channel of a TOFPET2

The signals from the discriminators are fed to a digital logic block (the logic works at 200 MHz clock frequency) that controls two TDCs and a charge integrator. The first TDC measures the phase of the rising edge of the low threshold discriminator output with respect to the reference clock. The second TDC measures the falling edge of the high threshold discriminator output. The mixed mode TDC block is based on four-fold-time-to-analogue converter (TAC), for signal de-randomization and a 10-bit Wilkinson analogue-to-digital converter (ADC). The TDC has 30 ps time binning.
The channel digital logic also handles the charge integration time interval (it is possible to be configured in different ways). The output of the charge integrator is digitized by a Wilkinson ADC. One of the two ADCs per channel is configured to measure either the charge amplitude (default mode) or the fine time of the falling edge of the high threshold discriminator (in this way, it is possible to reject any event characterized by low charge and not to saturate). In the first case, the second time measurement has the resolution of the coarse time counter (200 MHz). The difference between the time measurements can be used to calculate the time-over-threshold (ToT) that depends on the amplitude of the pulse. The digital data corresponding to each input pulse (event) is then stored in local buffers before being transferred to the FIFO.
Once we give an overall of the chip it is the time to give an overview about its main features:

- Designed in standard CMOS 110 nm technology.
- Signal amplification and discrimination for each of 64 independent channels.
- t1, t2 are configurable separately.

- Rejects dark counts without triggering, allowing to handle large dark counts rates.
- Configurable charge integration time up to one microsecond.
- Quad-buffer TDCs and charge integrators for each channel. The first branch is used for timing measurement. The second can be either used for time-over-threshold (ToT) or charge measurements with a Wilkinson ADC
- Dynamic range: 1500 pC.
- TDC time binning: 30 ps.
- Gain adjustments per channel in the charge branch: 1, 0.5, 0.25.
- On-chip charge calibration pulse generator with-6-bit programmable amplitude.
- Main clock frequency: 160-200 MHz.
- Max output data rate per ASIC: 3.2 Gb/s.
- Max event rate per channel: 600 kevents, 80 bits per event.
- Power dissipation per channel: 8.2 mW, for the recommended settings.

3.1.1 Frontend Amplifiers and Discriminators

The preamplifier are two currents that can be selected, in the case of positive and negative polarity. The amplifier is based on a modified of the common gate transimpedance amplifier (TIA). The conveyors give a low input impedance for the detector and a high impedance current output. The circuit for the positive signal version is depicted in Fig. 3.2 (better DR and preservation of timing performance). Regarding the negative signal, the topology used is the same of Fig. 3.2, with the transistors (NMOS or PMOS) swapped to the complementary type. A preamplifier provides a low frequency amplification of 25 dB and a passband of 330 MHz and a has a power consumption of 2.5 mW.
Once the processing has taken place, the signal is processed by two amplifiers and an integrator. The preamplifier is connected to the TIAs (PMOS current mirrors) in AC. The discriminators are fed by the TIAs. The first discriminator uses a low threshold at the level of 1 pe and 20 pe, while the second one has a threshold up to 1000 pe.

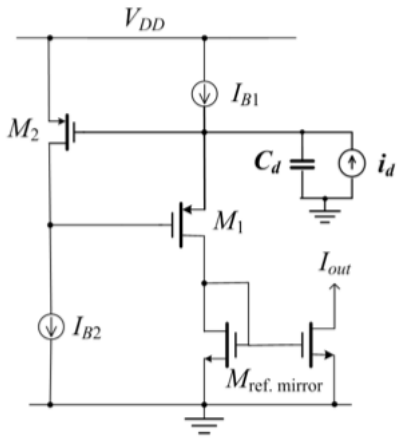

Fig. 3.2: Preamp. Internal structure

3.1.2 Charge Integrator

The purpose of the charge integrator is to process the signal coming out from the preamplifier within a time window that is generated by the digital logic (this is based on the output of the discriminators). In the scheme in Fig. 3.3, there are four flipped capacitors sharing a single readout differential amplifier perform the signal integration, allowing de-randomization of the signals. The analogue signals are fed to a 10-bit a Wilkinson ADC for the conversion.

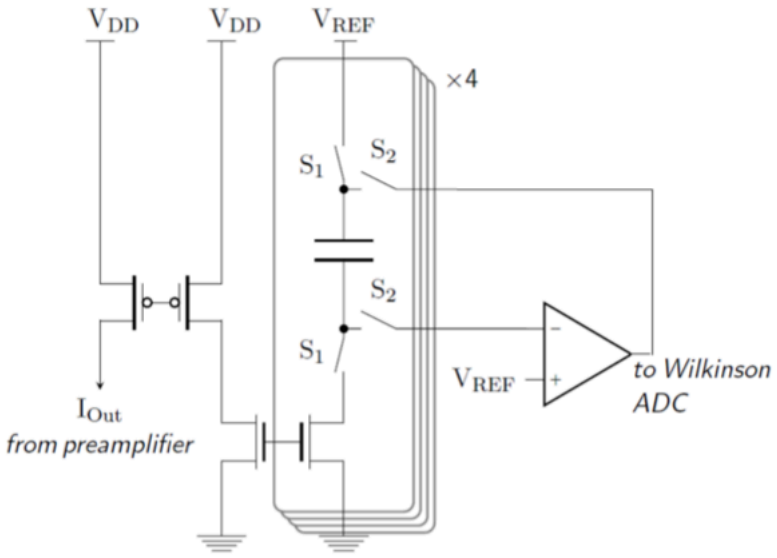

Fig. 3.3: Integrator structure and ADC

3.2 FLEXTOT

The FlexToT (Flexible Time Over Threshold) ASIC has been designed by the Universitat de Barcelona and CIEMAT, then fabricated by Austria Micro System (AMS) [25] in HBT technology (Heterojunction Bipolar Transistor, HBT), using SiGe Bi-CMOS 0.35 um technology, with the aim of supplying a fast, low-power front-end readout for silicon photomultiplier array scintillator-based PET detectors. It has 16 independent channels with DC-coupled inputs, permitting individual adjustments of the threshold and bias voltage for each SPAD, making it possible to correct for any gain variations due to temperature changes and/or pixel non-uniformity. Basically, the ASIC is able to send an output signal for each channel with duration proportional to the amplitude of the current at its input (as it can be seen for the peaks in Fig. 3.4).
At the input stage three signal paths are created [26]. The basic signal processing chain following this stage is shown in Fig. 3.4. The different signal paths are simple mirror outputs with different gains from the input stage.

The high gain output is implemented for the timing information (first incoming photons will give the time information), with a very low threshold setup to detect as soon as possible. The fast comparator is needed to output the width of the input signal for the selected threshold. The timing outputs of the channels are the inputs of an OR that will send a common time signal. It is necessary to make an effort in order to avoid the malfunction of the circuit when this branch is under saturation. As a result, there is an increase in dynamic range, keeping good timing measurement.
A second output with smaller gain will be integrated with linear discharge and then connected to a discriminator. The width is proportional to the incoming signal (energy output). Some digitally configurable parameters allow to adjust to different signal ranges with better linearity for different input timing constants or ranges.
In the end, the third output with new pile-up detection circuitry is also generated. The output of the third signal path is differentiated using a capacitor and then compared with a threshold to generate a digital output. This output is then connected to two cascaded flip-flops that will give a digital output high when input peaks are detected.
The configuration of this chip is carried out by a standard JTAG (Joint Test Action Group). Following this standard, it can be possible to access to the common registers (for the 16 channels of the ASIC) and to the independent register of each channel. Moreover, the threshold of the discriminators of each branch is accessible (*Energy* for LG (low gain), I*th* for HG (high gain), *PileUp* for LGb). The discharge constant of the preamplifier of the LG branch and the DC value at the input of each channel (Offset) are accessible as well.
Regarding the principal characteristic [27] of the ASIC FlexToT, we can present the following ones:

- The input amplifier should have at 250 MHz of BW in order to work well with scintillators.
- Proportionate a wide dynamic range in order to guarantee the operation with different detectors (0.1 mA – 18 mA).
- Low power consumption per channel (11 mW)
- Guarantee the direct connection in DC with the SiPMs and the option to variate the DC voltage at the input of the channel (in the SiPMs anode). This option is offered to balance the difference in gain due to the different voltages of the pixels of the SiPMs when they are fired by photons.

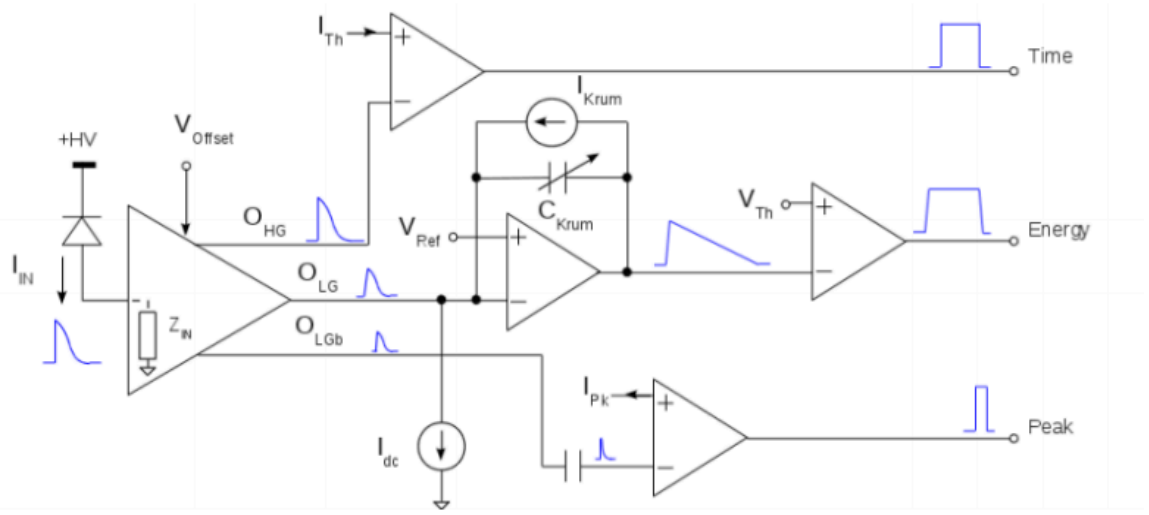

Fig. 3.4: Block Diagram of the FlexToT electronics

There are two versions of the ASIC FlexToT (v1 and v2) with the same architecture of the one in Fig. 3.4. In the second version, an input for analog calibration has been added through 16 bits of the configuration register. By means of this control, the commutation of this input to each of the 16 channels of the ASIC has been enabled, allowing the automatic calibration of the ASIC itself. The dynamic range of the ASIC is increased in order to operate with currents below 1 mA (monolithic blocks can be used). Also, the homogeneity between the channels of the ASIC has been improved.

3.3 PETIROC

Petiroc is a 16-channel front-end ASIC (see Fig. 3.5) designed in AMS 0.35 um SiGe technology [28] by Weeroc. The die size has an area equal to 3.5 x 5.7 mm^2 (it is a naked die since we want to get good result in terms of front-end's bandwidth). A drawback is due to the use of package, since this implies high parasitic inductances and hence, worse performances of the ASIC.
Each channel is composed by a very-low power input DAC to adjust the high voltage of the SiPM up to 800 mV. A wider range is not possible to be adjusted, without unbiasing the circuitry, due to the fact the electronics front-end is DC coupled to the SiPM.
It is also present a fast trigger line, composed of an RF preamplifier followed by a fast discriminator followed by a digital latch. A charge measurement is composed of a shaped preamplifier followed by an analogue memory.

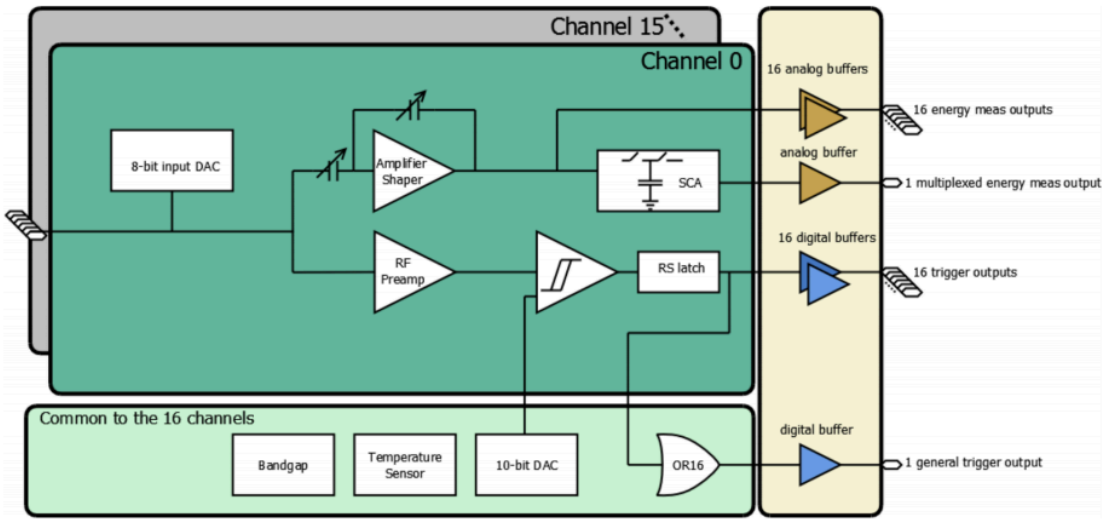

Fig. 3.5: Petiroc block diagram

A 10-bit DAC ensure the trigger level adjustment from 0.5 to few tenth of photoelectrons. The DAC's input has been designed to be very low power because that part of the ASIC cannot be power pulsed. Indeed, the stability of the SiPM high voltage is an important parameter for the stability of a system and the large capacitance of the SiPM as well. A power consumption of 1 uW/DAC is ensured.
A bandgap provides a stable reference for different biasing and reference voltage throughout the ASIC. A temperature sensor allows an analogue measurement of the ASIC temperature that will reflect the temperature of the system. This effect is due to the fact the Petiroc is placed close to the SiPM.
Petiroc provides 16 analogue charge measurement outputs for prototyping and also bench measurement as a multiplexed serial analogue output to be used in a multichannel multi-ASIC system. The 16 trigger outputs are then supplied to the user and can be disabled singularly. A trigger OR is given as well in order to perform an ASIC level trigger. It is important to underline that the large panel of outputs permits a versatile use of Petiroc, either for few-channel prototyping or small system design.
Hereafter an overview of the features [29] of the PETIROC:

- Number of channels: 16.
- Sensitivity: trigger on first photoelectron.
- Timing resolution: 20 ps FWHM in analogue mod (2 pe injected) – 50 ps FWHM with internal TDC.
- Dynamic range: 2000 photoelectrons.
- Power consumption: power supply 3.3; analogue core 56 mW (excluding analogue outing buffer), 3.5 mW/ch.
- Inputs: 16 voltage inputs with DC adjustments for SiPM HV tuning.

- Outputs: Digital output (energy on 10-bit, time on 10 bit – 40 ps bin); 16 trigger outputs; 1 multiplexed charge output, 1 multiplexed hit register; 2 ASIC trigger outputs (Trigger OR on 16 channels, 2 levels).
- Internal Programmable Features: 16 HV adjustment for SiPM (16 x 8 b), trigger threshold adjustment (10 b), charge measurement tuning, 16 trigger masks, internal temperature sensor, trigger latch.

3.4 Final Considerations

After an initial evaluation of the chips in the market, the project group decided to use the TOFPET2 ASIC since it is the best one fitting the requirements of the entire system. Indeed, the thesis objective is not to realize and hence, customize an ASIC. The objective is to explore the different alternatives in order to see which one between them is the best one (given the requirements of work). We do not know how the parts of the commercial ASICs are made (type of the integrators, preamplifiers, etc.) and for this reason we set out to realize a model in order to compare the different features and detect where the limits are, etc. In another words, the objective of the thesis is "prospective", an investigation of the model options we have in order to carry out systems as the ASICs (that we took as guidelines) are.
Specifically, in the project thesis we developed a preamplifier with a current mirror at the output that is connected to a discriminator (timing measurements), that is the first block of the fast chain (TDC), and to a charge integrator, which establishes the beginning of the slow chain (ADC). We did not take into account the dark noise discriminator (this part was not included in the project thesis).
Moreover, the sizing of the mirror transistors is different from a chain to the other, depending on the characteristics of the chain itself.
Focusing on the slow chain, after the charge integrator we inserted a Track and Hold (we did not consider the energy discriminator right after the integrator). This block was necessary to track the voltage on the feedback capacitance at to reduce as much as possible the charge injection effect.
Originally the TOFPET2 uses a Wilkinson ADC, but in our case, we decided to exploit the better performance of a SAR ADC. Indeed, first of all, the conversion time of a Wilkinson ADC is much slower than a SAR. Furthermore, in the case we considered more bits (more than 9) the resolution would be as not accurate as the resolution of a SAR.

Chapter 4

Input stage

Regarding this chapter, the methodology used for the input stage measurements is explained in deep, justifying why and how these simulations are carried out in order to achieve the expected objectives. The input stage is characterized by three blocks with three different purposes.
The first stage (Summing parcel) has been implemented to mimic the current at the input of the Preamplifier. We considered 440 photoelectrons DR at the input of the Preamplifier. For this reason, we came up with SOMMA_Stimuli_tot. This is the final block including the 440 photoelectrons (causing 440 avalanche events) and from which a current will be generated and processed.
In SOMMA_Stimuli_tot the SPAD's output is represented by an exponential voltage. At the beginning, the signal is at 0V and then, after reaching a peak of 20 uV, decreases exponentially. We used voltage sources, instead current sources, because in the Cadence functional library only voltage adders were available. Therefore, after summing the voltages, a current source controlled in voltage (transconductance amplifier) is placed in order to realize the connection in current. Indeed, the output of a SiPM is a current entering in the input low impedance of the Preamplifier.
The second one is the core of the input stage, the Preamplifier. This is necessary in order to transfer the current we have at the output of the first block (SOMMA_Stimuli_tot) to the input of the integrator.
The third and the last one is a block to null the DC offset at the output of the Preamplifier.
Once presenting all the components in theory, we will show the results of the Cadence's simulations of each block.
Finally, the last section presents the results of the simulation performed on the entire preamplifier.

4.1 Summing Parcel

The Summing block or SOMMA_Stimuli_tot is needed at the input of the Preamplifier in order to emulate the DR of the system.
This block is composed by four sum blocks (each of them labeled as SOMMA), each of them containing 11x10 sub-blocks (in total there will be 440 avalanche events due to the 440 p.e. DR, causing a voltage in the SPAD and hence, a current flowing through it). In deep, each sum block will have a structure equal to the one in Fig. 4.1, where each Stimuli block is summed up to another through an adder, having in total 10 Stimuli block (SOMMA1). As we can see in Fig. 4.2, each Stimuli block is made by 11 blocks represented by an exponential voltage generator, having a peak equal to 20 uV (it starts to 0 V, after a time it goes up to 20 uV abruptly and

finally decreases with an exponential trend). Every single block of Stimuli corresponds to the SPAD's output (that is actually a voltage) composing the SiPM. Each Stimuli block is identified by a delay (it is a variable) and each SPAD's output of a Stimuli block appears at 20 ns after the previous one, covering a range of 200 ns in total.

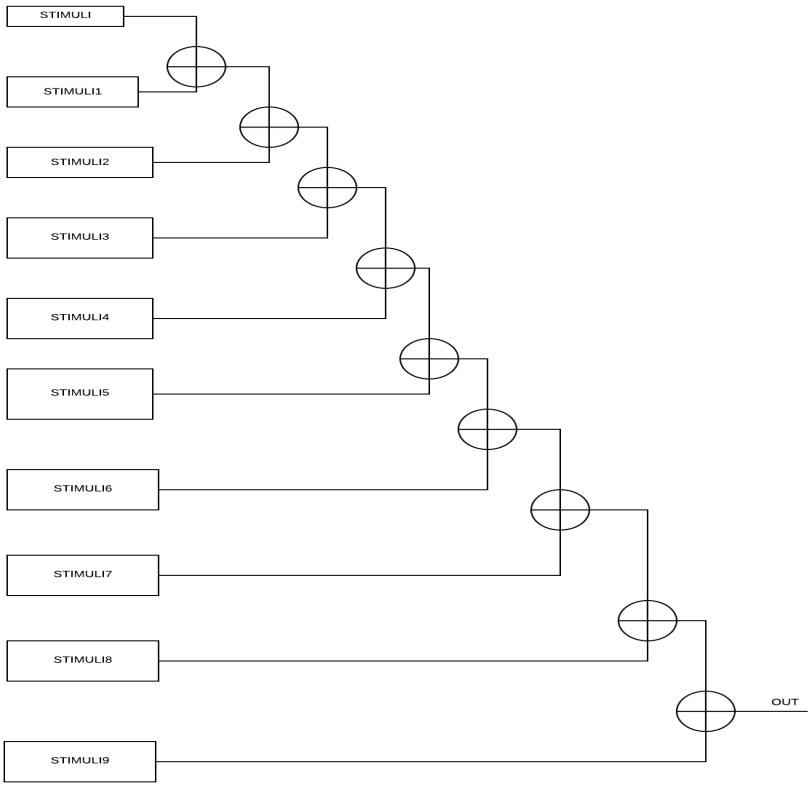

Fig. 4.1: SOMMA1 block's structure

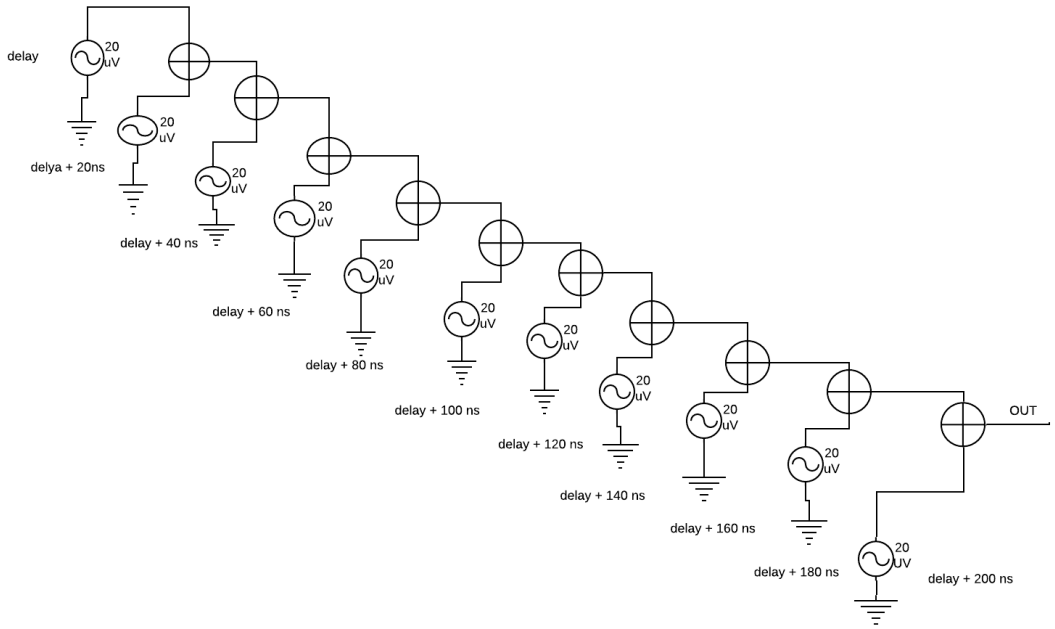

Fig. 4.2: Single Stimuli block

In Fig. 4.3 the four SOMMA blocks are summed to each other and connected at the end of the chain to a Voltage Controlled Current Source (vccs) component that is basically a conductance equal to 1 MS (Mega Siemens). In this way, we transform the total voltage into current which will be available at the input of the Preamplifier. Indeed, the output of a SiPM is a current.

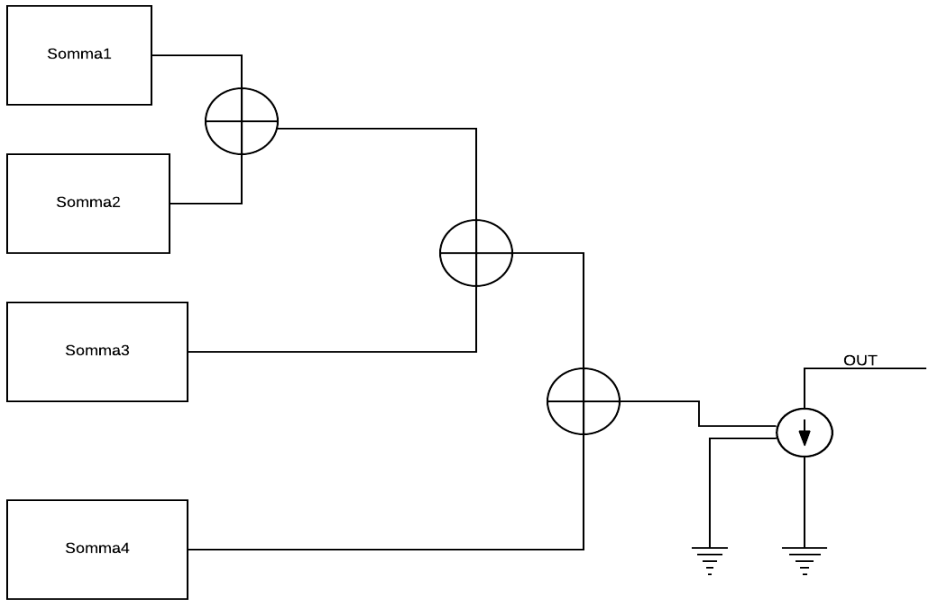

Fig. 4.3: Total Sum block, SOMMA_Stimuli_tot

4.2 Preamplifier

From the analog requirements, the BW of the total front-end (Preamplifier + OTA inside the charge integrator) has to be equal or greater than 200 MHz. For this reason, in our specific case, we want to reach at least a 100 MHz of Preamplifier's BW that could not be reached by a simple transimpedance amplifier with resistive feedback and an integrator. Moreover, if we had used a resistive feedback, we would have had problems in terms of stability, since we would have coupled the capacitance of the Silicon Photomultiplier (CSiPM), at the output of the resistive feedback, with the capacitance of the integrator (Ch) in the loop. A wider BW, using the topology for the preamplifier we came up with, could not be reached due to the high equivalent capacitance of the SiPM (500 pF), that introduces the dominant pole in the system, and the limit on the power dissipation. At most, the power dissipation of the system can be 5 mW (from specification).
A very effective read-out approach, based on a current mode preamplifier, is capable of reading the current pulse coming from the SiPM at very low impedance, discharging quickly the large equivalent capacitance of the detector (CSiPM), and reproduces this current (with a scaling factor) on a high impedance node, so it can be further processed for the extraction of the time and energy information. Typically, large bandwidths are easier to be achieved with current mode amplifiers, because of the absence of high impedance nodes (at the input node), therefore the output signal can chase the very fast leading edge of the current pulse generated by the SiPM, resulting in good performance in terms of time resolution.
The Fig. 4.4 shows the internal Preamplifier's structure. As we can see, it is composed by different stages: an internal loop (M2-M3-M15-M7); at the top we have an enhanced cascode mirror (M10-M11 and M1-M4) needed to the transfer the current, resulting from single photons, to the input of the integrator block, connected to Vout; a DC offset correction; and finally, there is another enhanced mirror at the bottom of the Preamplifier input (the current in DC is mirrored from the input to the output). The mirror M10-M11 and M1-M4 is not a mirror 1-1, but the gain, that is actually factor scaling, is equal to 0.04 ($G < 1$).
In deep, we will go through each stage in order to understand better the functionality of the amplifier. Besides, we will clarify why there is a DC offset block inside the Premamp.

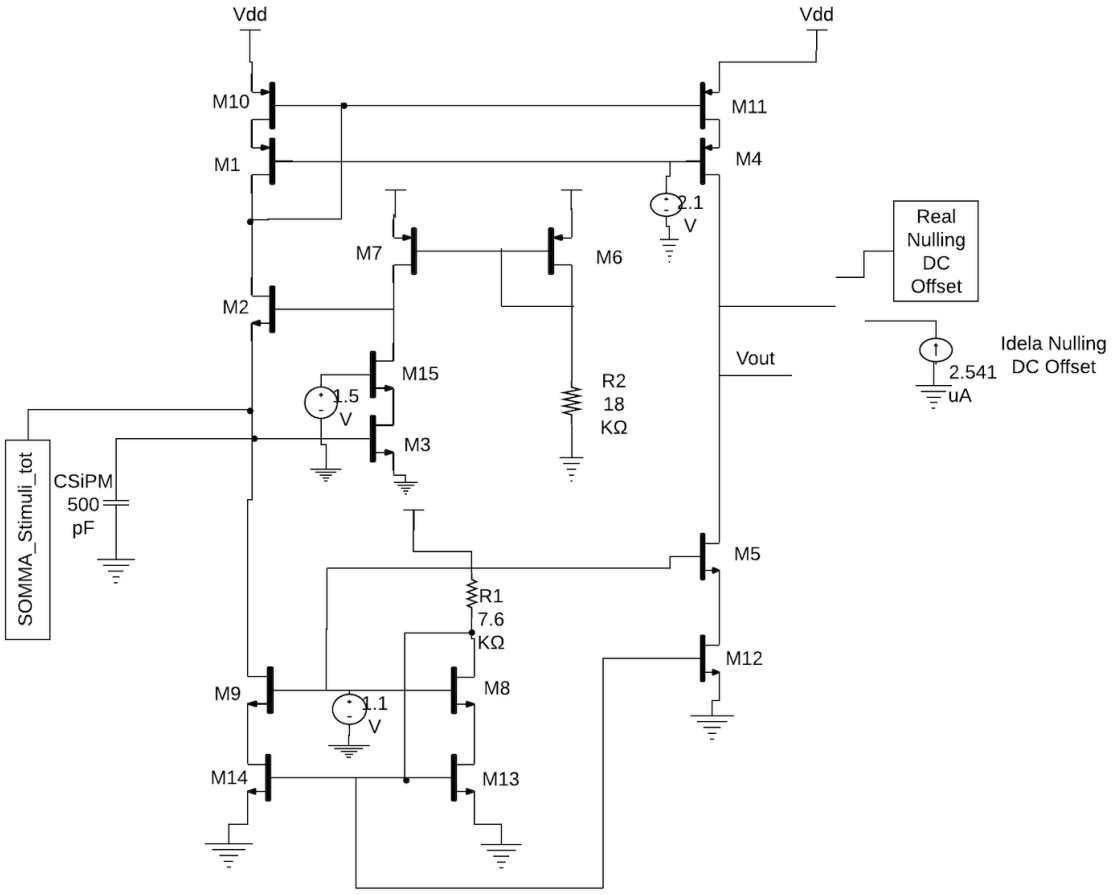

Fig. 4.4: Internal Preamplifier's structure

4.2.1 Internal Loop

The internal loop is composed by M2-M3-M15-M7, as shown in Fig. 4.4. The reason why the feedback has been employed is to decrease the input resistance of the current mode amplifier while saving power.
First of all, we will focus on the analysis of the stability of the stage. The DC loop gain Gloop (0), is:

$$Gloop\ (0) = -gm3 * r7$$

Fixing the value of Gloop (0) to 100 and, setting the value of r7 equal to 20 KΩ, we compute the value of gm3, resulting to be 5 mA/V. The choice of picking r7 = 20 KΩ has been made for stability necessity. Since at the drain of M7 there is an equivalent parasitic capacitance (small), in order to have the pole introduced by this capacitance at high frequency we needed r7 around some KΩ.
The current flowing into M3 and M7 will be the same:

$$I_3 = \frac{gm_3 * V_{ov3}}{2} = 500\ uA = I_7$$

The reason why we came up with this value for the current of M3 is because we wanted to reduce the noise effect introduced by it.
Being the BW of the preamplifier 100 MHz and considering the relationship between BW and GBWP as $BW \approx GBWP * \frac{\pi}{2}$, the GBWP results to be 65 MHz). Therefore, we now are able to compute gm2 from the GBWP's formula:

$$GBWP = Gloop(0) * f1 = \frac{gm2}{2 * \pi * CSiPM} * gm3 * r7 = 65\ MHz$$

$$gm2 = \frac{2 * \pi * CSiPM * GBWP}{Gloop(0)} = 2.04\ \frac{mA}{V}$$

We decide to fix gm2 = 2.09 $\frac{\text{mA}}{\text{V}}$ and from this value and the knowing relationship for a transistor in saturation region (gm = (I*2)/Vov):

$$I_2 = \frac{gm_2 * V_{ov2}}{2} = 314\ uA$$

Knowing gm2, it is possible to compute the frequency of the dominant pole introduced by CSiPM:

$$f1 = \frac{gm2}{2 * \pi * CSiPM} = 667\ KHz$$

At this point, in order to have the second pole larger than the GBWP, we consider the shortest L (length in um) for M3 and M15 so that we have a small contribution in the drain capacitance. The second pole is due to the capacitance C_2 at the gate of M2 and the drain capacitances of M7 and M15: $C_2 = \left(C'_{db} * (3 * L_{min}) * (W15 + W7)\right) + (\frac{2}{3} * C'_{ox} * \left((\text{W2} * \text{L2})\right) = 33.2\ fF$.
In the formula, we included the contribution of the area of the drain regions, $C'_{db} = 0.1\ \frac{fF}{um^2}$, $3 * L_{min} = 0.35\ um$, as spacing of the drain contact areas and $C'_{ox} = 5\ \frac{fF}{um^2}$, as oxide capacitance.
The frequency of the second pole is computed hereafter:

$$f2 = \frac{1}{2 * \pi * C_2 * r7} = 240\ MHz$$

4.2.2 Enhanced Cascode Mirror

As it can be seen in Tab. 4.1, there is a reduction factor of 0.04 from the current flowing into M10 to the one in M11 (gain of the stage). As well known, de-amplifying is deleterious at the input of a front-end stage since we worsen the SNR. At the same time, this is the only solution due to a problem related to the area occupied (equivalent SiPM capacitance). Being the gain of the SiPM big, the circuit cannot deal with the amount of charge it has. In detail, considering $Q =$

$C * V$ where Q is equal to 176 pC since the charge of the SPE is 0.4 pC and the DR is 440 pe and V is the maximum excursion at the output of the integrator, we obtain:

$$C = \frac{Q}{V} = \frac{176\ pC}{2\ V} = 88pF$$

The capacitance we would need at the output would be too big. Indeed, the max capacitance we could have at the output is 15 pF. Since it has been involved the channel Metal1/Metal1, that has a parasitic capacitance of 2 fF/um², if the capacitance at the output was 15 pF, it would mean an area of 7500 um² would be occupied. If it was true for each channel, the system would be enormous. Hence, in order to have some margin we ended up picking a capacitance equal to 8 pF. Doing so:

$$Q_{11} = C_{out} * V_{out} = (8\ pF) * (2\ V) = 16\ C$$

$$\frac{Q_{11}}{Q_{max}} = \frac{Q_{11}}{Q_{10}} = \left(\frac{16}{176}\right) C = 0.0909$$

We decided to take a margin from the result above ($\frac{Q_{11}}{Q_{max}} = 0.0909$), considering 0.04 if I had more charges at the input of the Preamplifier.
At this point, being the charge proportional to the current flowing into a transistor ($Q = \int I * dt$), the same expression for the charge can be applied to the current.
Once the analysis in terms of stability and the aspect relative to the gain OF THE preamplifier are carried out, we can compute the sizing of the remaining transistors of the amplifier and the relative characteristics, coming up with Tab. 4.1.

	I(uA)	Vov(V)	gm(uA/V)	W/L	W(um)	L(um)	rd (Kohm)
M1, M10	314	0.2	3140	314	220	0.7	44.6
M2	314	0.3	2090	69.78	24.4	0.35	22.3
M3	500	0.2	5000	250	625	2.5	100
M4, M11	12.6	0.2	126	12.56	8.79	0.7	1101
M5, M12	12.6	0.3	84	2.79	0.98	0.35	557
M6	8.33	1.5	111	1.5	0.74	0.5	120
M7	500	1.5	667	8.89	4.44	0.5	20
M8, M13	314	0.3	2090	69.78	24.4	0.35	22.3
M9, M14	314	0.3	2090	69.78	24.4	0.35	22.3

M15	500	0.3	3330	111.11	38.9	0.35	14

Tab. 4.1: Preamplifier sizing

4.2.3 Noise Evaluation

Secondly, we go through the analysis of the Preamplifier's noise showing the contribute of each component at its output (then this noise will be transferred to the ADC) in terms of power spectral density. At the end, the total power spectral density (taking into account the single contributions) will be computed.

$$Svout_{M7}^2 = 4KT\lambda \cdot gm_7 \cdot (\frac{gm_2}{gm_3})^2 \cdot \left(\frac{gm_{11}}{gm_{10}}\right)^2 = 1.4 \cdot 10^{-27}\,\frac{A}{Hz}$$

$$Svout_{M6}^2 = 4KT\lambda \cdot gm_6 \cdot (\frac{gm_2}{gm_3})^2 \cdot \left(\frac{gm_7}{gm_6}\right)^2 \cdot \left(\frac{gm_{11}}{gm_{10}}\right)^2 = 0.85 \cdot 10^{-26}\,\frac{A}{Hz}$$

$$Svout_{M3}^2 = 4KT\lambda \cdot gm_3 \cdot (\frac{gm_2}{gm_3})^2 \cdot \left(\frac{gm_{11}}{gm_{10}}\right)^2 = 1.05 \cdot 10^{-26}\,\frac{A}{Hz}$$

$$Svout_R^2 = 4KTR \cdot gm_7^2 \cdot (\frac{gm_2}{gm_3})^2 \cdot \left(\frac{gm_7}{gm_6}\right)^2 \cdot \left(\frac{gm_{11}}{gm_{10}}\right)^2 = 2 \cdot 10^{-26}\,\frac{A}{Hz}$$

$$Svout_{M11}^2 = 4KT\lambda \cdot gm_{11} = 1.12 \cdot 10^{-24}\,\frac{A}{Hz}$$

$$Svout_{M12}^2 = 4KT\lambda \cdot gm_{12} = 7.5 \cdot 10^{-24}\,\frac{A}{Hz}$$

$$Svout_{M14}^2 = 4KT\lambda \cdot gm_{14} \cdot \left(\frac{gm_{11}}{gm_{10}}\right)^2 = 2.5 \cdot 10^{-26}\,\frac{A}{Hz}$$

$$Svout_{M13}^2 = 4KT\lambda \cdot gm_{13} \cdot \left(\frac{gm_{11}}{gm_{10}}\right)^2 = 2.5 \cdot 10^{-26}\,\frac{A}{Hz}$$

$$Svout_{M10}^2 = 4KT\lambda \cdot gm_{10} \cdot \left(\frac{gm_{11}}{gm_{10}}\right)^2 = 3.76 \cdot 10^{-26}\,\frac{A}{Hz}$$

$$Svout_{M2}^2 = 4KT\lambda \cdot gm_2 \cdot \left(\frac{gm_{11}}{gm_{10}}\right)^2 = 2.5 \cdot 10^{-26}\,\frac{A}{Hz}$$

Summing up, the most significant contributions to the total power spectral density (at Vout) are due to M_{12} and M_{11} because the mirror de-amplifies (factor 0.04). The total power spectral density results to be:

$$Svout_{Vout}^2 \cong \sqrt{Svout_{M11}^2 + Svout_{M12}^2} = \left(1.4 \cdot \frac{pA}{\sqrt{Hz}}\right)^2$$

4.3 Nulling DC offset

This block has been attached to the Preamplifier's output once we run the DC spectre simulation. In fact, we realised a mismatch between the currents flowing into the PMOS at the top and the NMOS at the bottom of the output stage with a voltage generator of 2V (this is the maximum voltage in order to guarantee 3.3 V at the input of the ADC) at the output of the amplifier (attaching a big impedance this effect will not occur). PMOS generates 9.317 uA while NMOS emits 11.858 uA and therefore a current equal to 2.541 uA is necessary to fix this offset.
To do that, we firstly connected an Ideal Nulling DC Offset that is simply an ideal current generator that is characterized by a current equal to 2.541 uA (difference between PMOS and NMOS current, as said before). This is just a way to simulate the circuit.
A real option is also provided, as depicted in Fig. 4.4 (Real Nulling DC Offset block).
At the top level, the Real Nulling DC Offset is identified by the Current Collector (CC) [30] topology (see Fig. 4.5) and the ladder R-2R structure (the whole structure is displayed in Fig. 4.6 and the explanation of its working principle is shown in Fig. 4.7) connected to each other (they work simultaneously). The two blocks are connected between each other through two voltage suppliers: DUMP and OUT
The current collector (CC) structure is shown in Fig. 4.5. It is necessary to keep DUMP and OUT voltages as close as possible, while taking out current from Mout transistor (this is the output of Real Nulling DC Offset) to inject it to Preamplifier's output. In this way, we are able to compensate the current flowing in the output branch of such block (it is the target of the NULLING OFFSET block). The internal common mode voltage of the fully differential operational transconductance amplifier (OTA) sets the value of the DC voltages of OUT and DUMP. At the same time, the OTA supervises the voltages keeping the difference between them as low as possible thanks to the high gain of the feedback mechanism.

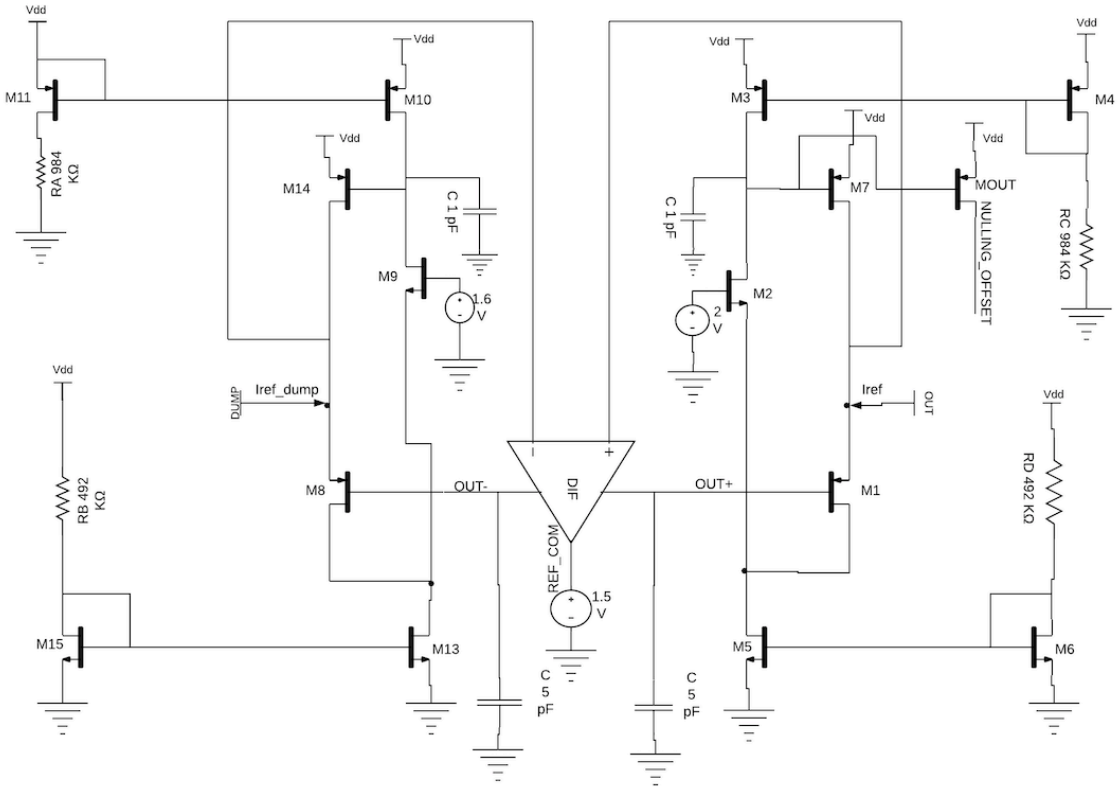

Fig. 4.5: Fully Differential Current Collector

Before analyzing the sizing of the transistors composing the CC, it is important to understand how the ladder R-2R works (working principle) and how it is connected to the CC. Hence, we firstly study the working principle of the DAC R-2R in deep (Fig. 4.6) referring to a Bult and Geelen [31] current splitter.

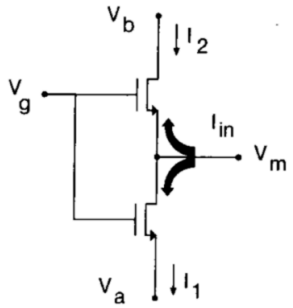

Fig. 4.6: Current division's principle

Both transistors have the same gate voltage (V_G = Vdd = 3.3 V). Voltages Va and Vb could have any value as long as the transistors are on. A current flowing in or out of the circuit will be split in two parts. The current will be equally split up between the two transistors only if Va = Vb.

The current division principle claims that the fraction of this input current flowing to one side is:

- constant and independent of Iin
- independent of Va and Vb
- independent of whether one or both devices operate in strong or in weak inversion

This makes the current division technique independent of process variations.

Once the principle of the current division has been investigated and understood, it is the time to present the structure we implemented in our project. The goal of the ladder R-2R [30] is to provide Iref and Iref_dump, which allow us to correct the offset (in current) at the output of the Preamplifier.
The implementation of the MOS current divider (ladder R-2R) is depicted in Fig. 4.7. It exploits 6 bits digitally programmable downscaled. The transistors of each couple, that represent the n-th bit, are connected to voltage generators (3.3 V or 0 V) that turn the transistors either ON or OFF. Specifically, when VControl is 3.3 V (NVControl is 0 V) the n-th bit is 1, while on the other way around (VControl is 0 V and NVControl is 3.3 V) the n-th bit is 0. The final sequence of bits is 111001.

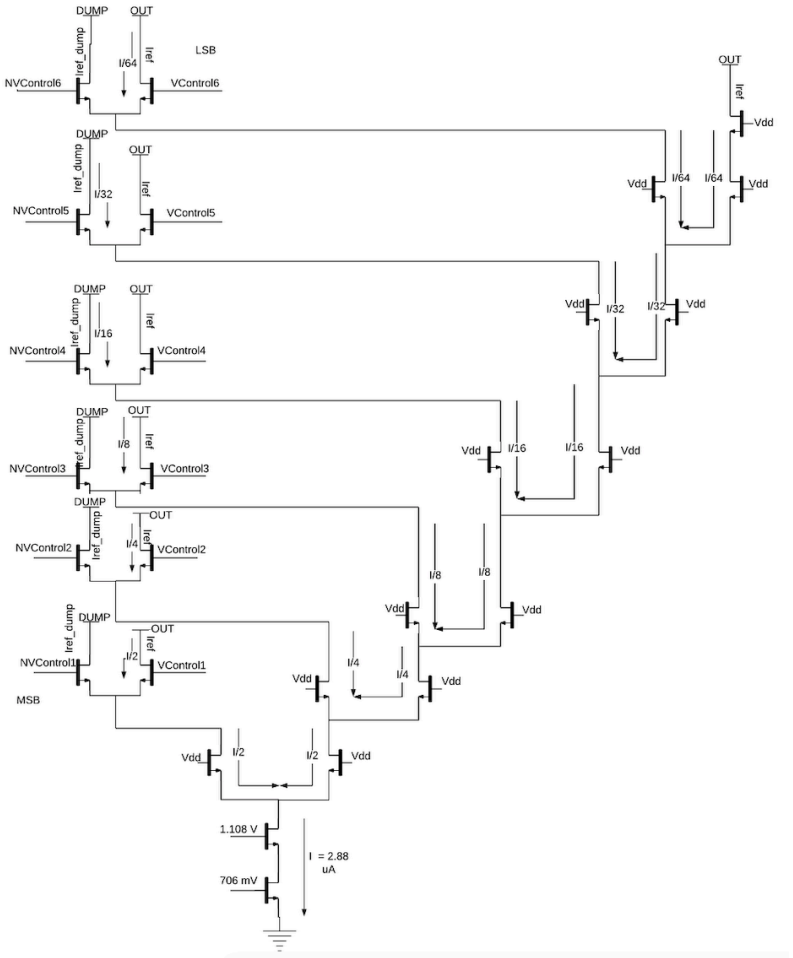

Fig. 4.7: DAC R-2R structure

Since the current of the tail generator is equal to 2.88 uA, we simulated the circuit (CC and R-2R) in order to have the best approximation of Iout (2.541 uA) (take into account the sizing of Tab 4.3). The current in Mout is actually 2.547 uA, that is a little bit greater than the target.
Once the current divider R-2R has been introduced, it remains to study the sizing of the CC of Fig. 4.5.
First of all, we set Iref equal to 2.541 uA and the current of M1 (I1) equal to this current as well. Therefore, the current into M7 will be 5.08 uA (Iref + I1). Then, for sake of simplicity, we set the current of M2 (I2) to 2.541 uA. Through the mirror stages, we set the currents into M3 and M5. Regarding the current of M3 (I3), it is the same of M2 and hence, having a Vov4 equal to 0.2 V and $(W/L)_4 = (W/L)_3$, we have to set Rc equal to 984 KΩ. On the other hand, being I2 and I1 equal to 2.541 uA, the current into M5 (I5) has to be 5.08 uA. In order to get this result, we set $(W/L)_6 = (W/L)_5$ and, being Vov6 0.2 V, Rd will be equal to 492 KΩ.
The same process has been applied to size the left part of the circuit. In this case, the only difference concerns Iref_dump, we considered it equal to 0 A.
In Tab. 4.2, there are the parameters of the transistors deployed in the CC.

	I(uA)	Vov(V)	gm(uA/V)	W/L	W(um)	L(um)	rd (Mohm)
M1	2.54	0.2	25.4	2.54	2.54	1	7.87
M2	2.54	0.3	16.9	0.565	0.565	1	7.87
M3, M4	2.54	0.2	25.4	2.54	2.54	1	7.87
M5, M6	5.08	0.2	50.8	2.54	2.54	1	3.94
M7	5.08	0.2	50.8	5.08	5.08	1	3.93
Mout	2.54	0.2	25.4	2.54	2.54	1	7.87
M8	2.54	0.2	25.4	2.54	2.54	1	7.87
M9	2.54	0.2	25.4	1.27	1.27	1	7.87
M10, M11	2.54	0.2	25.4	2.54	2.54	1	7.87
M13, M15	5.08	0.2	50.8	2.54	2.54	1	3.94
M14	2.54	0.2	25.4	2.54	2.54	1	7.87

Tab. 4.2: Fully Differential Current's sizing

At this point, it is the time to go to a deeper level, taking into account the differential amplifier topology of the CC (see Fig 4.5).
First of all, we start studying the CMFB topology (Common mode feedback) [32], shown in Fig. 4.8, where r7//r8 (these transistors work in ohmic region) adjusts the bias current of M5 and M6. The output CM level settles r7//r8 such that I5 and I6 exactly balance I9 and I10. Being I9 = I10 = I, so Vb – Vgs5 = 2* I * r7//r8 and therefore r7//r8 = (Vb – Vgs5) / (2* I).

$$\frac{1}{un * Cox * \frac{W}{L}_{7,8} * (V_{out2} + V_{out1} - 2V_{TH})} = \frac{V_b - V_{GS5}}{2 * I}$$

$$V_{out2} + V_{out1} = \frac{2 * I}{un * Cox * \frac{W}{L}_{7,8}} \frac{1}{V_b - V_{GS5}} + 2V_{TH}$$

The CMFB suffers from several disadvantages. First of all, the output CM level is a function of device parameters. Secondly, the voltage drop across r7//r8 limits the output voltage swings. Thirdly, in order to reduce this drop, M7 and M8 are wide, introducing substantial capacitance at the output.

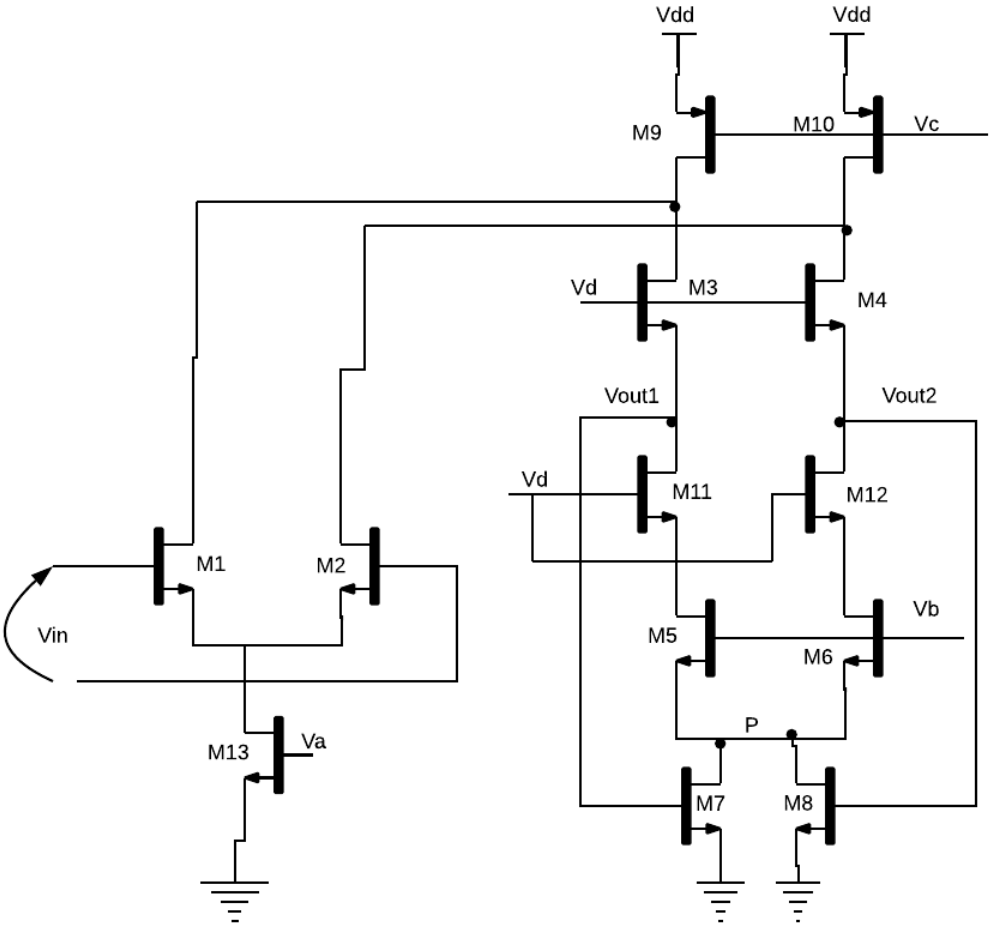

Fig. 4.8: CMFB using triode devices

In order not to have the drawbacks of the CMFB, we came up with the circuit in Fig. 4.9 [32]. In fact, the output swings are not limited any more, thanks to the feedback from the tail current generator, through the input differential pair, and

the output (M1-M6-M8-M9). In addition, the output levels are relatively independent from the device parameters and the sensitivity respect to Vb (control pin of the common mode voltage) has been lowered.
The idea is to define Vb by a current mirror arrangement such that Id9 "tracks" I15 and Vref (REF_COM). We set $(W/L)_{15} = (W/L)_9$ and $(W/L)_{16} = (W/L)_7 + (W/L)_8$. In this way, Id9 is equal to I15 only if Vout, cm = REF_COM. Besides this consideration, it is important to underline that the transistors 16, 7 and 8 work in linear zone.
Moreover, Vds15 and Vds9 would be different due to the channel modulation effect. To fix this problem, M17 and M18 are in the circuit to reproduce at the drain of M15 a voltage equal to the source of M1 and M2.

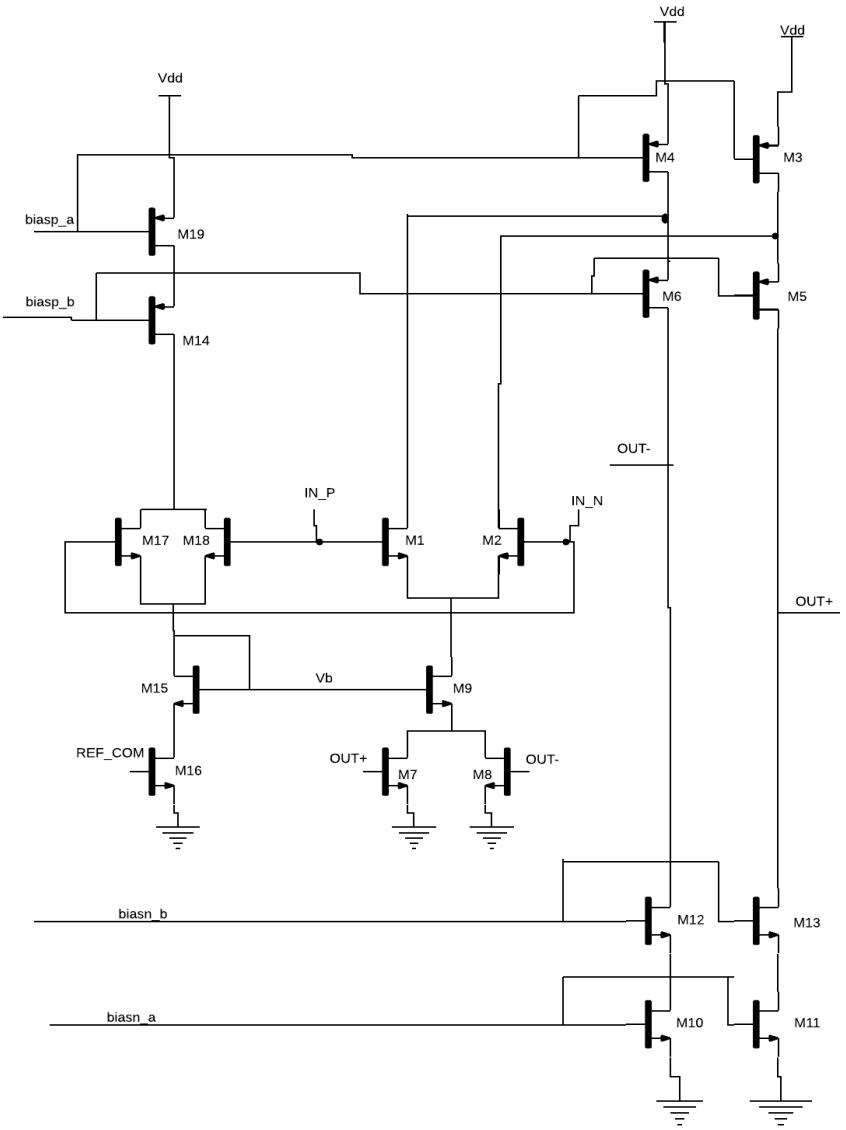

Fig. 4.9: Internal Differential amplifier's topology

Once we have an overview of the system, we can focus on the computation of the Gcm (Common mode gain) of the circuit [32]. To do that, we set IN_P - IN_N to 0 and we simplify the circuit as shown in Fig. 4.10. As said before, gm7 and gm8 have to be calculated in triode region (gm = un* Cox * (W/L) * Vds).

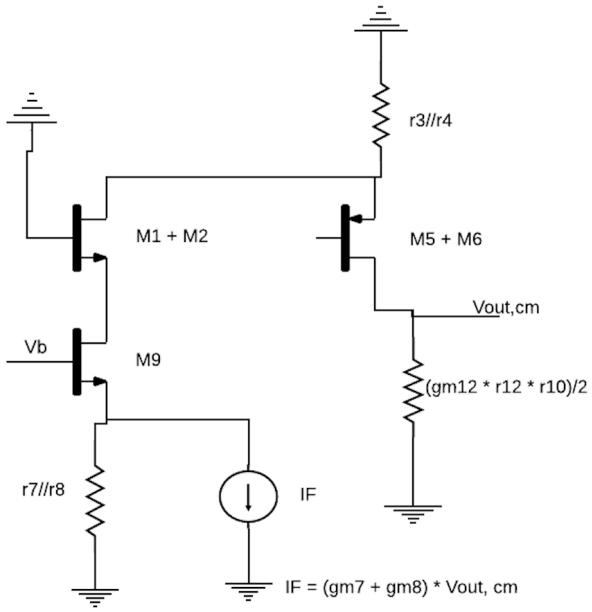

Fig. 4.10: Circuit to study the loop of the differential amplifier

In a well-designed circuit, the gain of the loop (common mode loop of the circuit in Fig. 4.9) has to be high. We therefore deduce that the closed-loop gain (refer to the differential amplifier closed in buffer configuration) is approximately equal to $1/\beta$ (we suppose the forward gain high), where β is the feedback factor.

$$\beta = \frac{V_2}{V_1} = -(gm7 + gm8) * (r7//r8) =$$

$$= -2 * un * Cox * \frac{W}{L}_{7,8} * \left(\frac{V_{DS7,8}}{2 * un * Cox * \frac{W}{L}_{7,8} * (VGS_{7,8} - V_{TH7,8})}\right)$$

$$= \frac{V_{DS7,8}}{VGS_{7,8} - V_{TH7,8}}$$

where $VGS_{7,8} - V_{TH7,8}$ is overdrive voltage of M7 and M8. Thus,

$$\frac{dVout,cm}{dVb} = \frac{VGS_{7,8} - V_{TH7,8}}{V_{DS7,8}}$$

Since $VGS_{7,8}$ (that is the output CM level) is typically around VDD/2, the equation suggests that $V_{DS7,8}$ has to be maximized. This final relation we have obtained is

important since, being Vb the control pin of the common mode voltage, it is possible to adapt the voltage of M9 through the differential couple.
In table 4.3 the characteristics of the transistors been involved in the differential amplifier are shown.

	I(uA)	Vov(V)	gm(uA/V)	W/L	W(um)	L(um)	rd (Mohm)
M1, M2	10.2	0.2	102	2.8	14	5	9.8
M3, M4	20.4	0.19	214.7	12.4	62	5	4.9
M5, M6	10.2	0.2	102	6	30	5	9.8
M7, M8	10.2	0.2	102	0.2	2	10	19.6
M10, M11	10.2	0.17	120	4	20	5	9.8
M12, M13	10.2	0.185	110.3	3.4	17	5	9.8
M15	20.4	0.2	204	5.6	28	5	4.9
M16	20.4	0.2	204	0.4	4	10	9.8
M17, M18	10.2	0.2	102	2.8	14	5	9.8
M9	20.4	0.2	204	5.6	28	5	4.9
M19	10.2	0.19	107.4	12.4	62	5	9.8
M14	10.2	0.2	102	12.4	62	5	9.8

Tab. 4.3: Differential Sizing when it is closed in a buffer configuration

4.4 Differential amplifier's Simulations

First of all, the differential amplifier of Fig. 4.9 has been simulated in a buffer configuration (see Fig. 4.11) in order to see its performances separately from the rest of the NULLING OFFSET block. Once check out its stability and the DC simulation, we insert it in a structure, building the Current Collector.

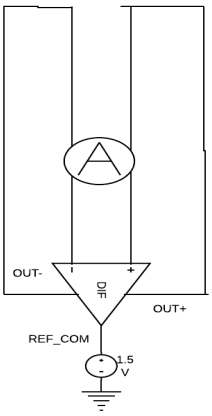

Fig. 4.11: Differential structure in buffer configuration

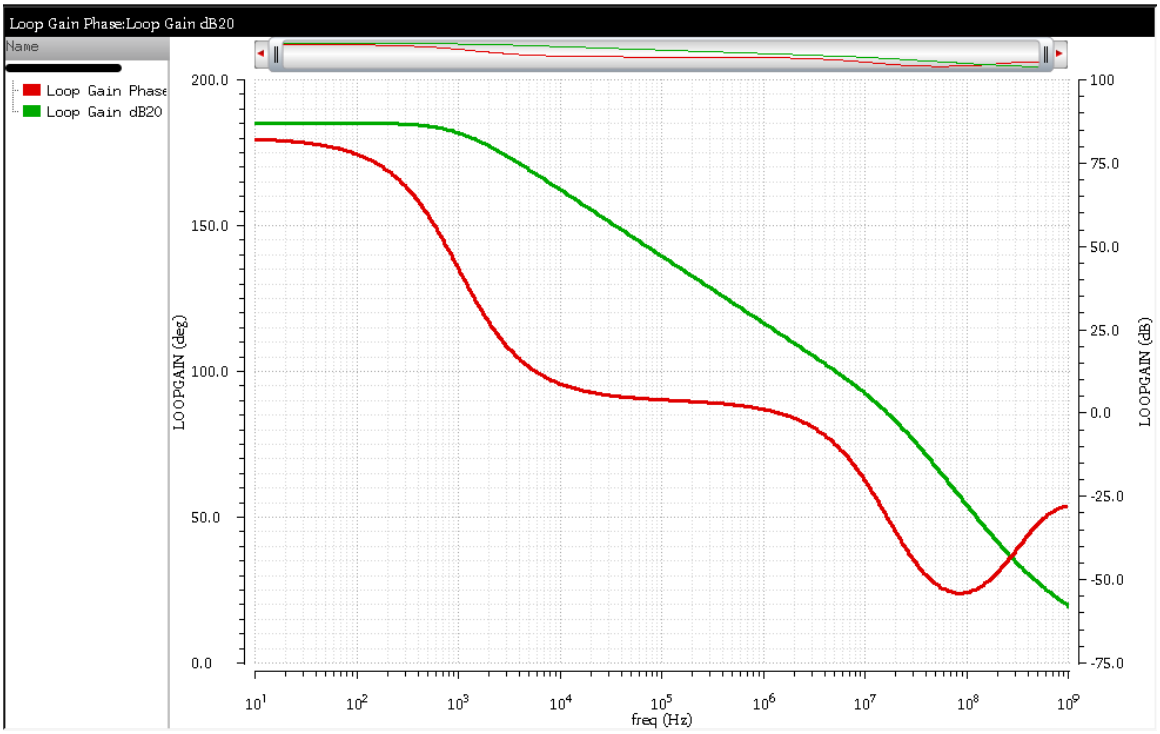

Fig. 4.12: Gain and phase of the loop of the differential amplifier closed in buffer configuration

The result of the simulation in Fig. 4.12 effectively says that the amplifier is stable with a phase margin of 51.82°. The measurement we want to realize is on the total circuit, CC + Ladder R-2R and not on the differential amplifier itself. For this reason, the phase margin is not low even though the simulation is on a typical corner frequency (we just want to have a stable differential amplifier, so this is not important how much is stable in the case study).
On the other side, the gain bandwidth product (GBWP) is equal to 15.45 MHz. From the result, the bandwidth of the amplifier is not so wide, but this is not a big

deal since we just want to exploit the component to cancel out the DC offset in the Preamplifier.
At this point, we built the fully differential amplifier current collector and we simulated it together with the ladder network. At the begin, we did not insert any capacitances and we saw that the circuit (CC) was not stable. For this reason, we firstly added two capacitances (5 pF) at the output (OUT- and OUT+) of the differential amplifier. Simulating, we realized that the circuit was not stable and, for this reason, two capacitances (1 pF) had been inserted at the other two critical points of the two loops (Iref and Iref_dump) to get stability.
Finally, we achieved stability, reaching 68.72° of phase margin and a GBWP equal to 3.28 MHz, as it is shown in Fig. 4.13 (the bandwidth is quite narrow, but as we said before, this is not important since we used this structure just to eliminate the offset of the Preamplifier output's branch).

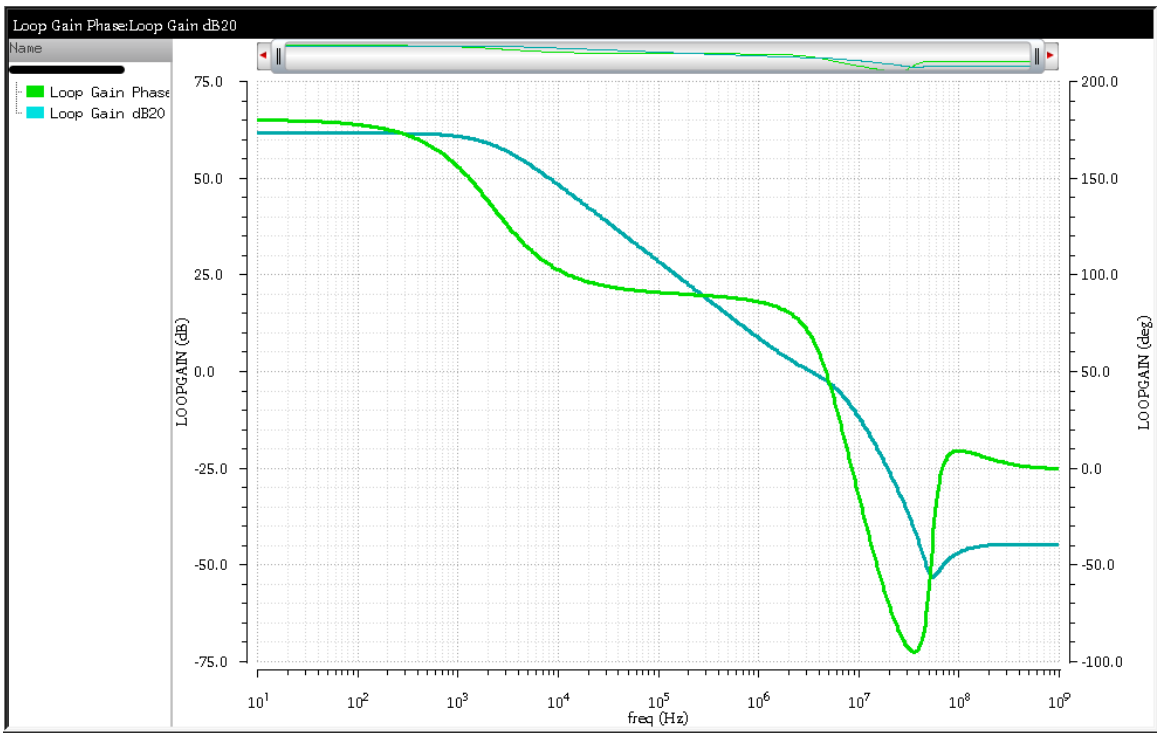

Fig. 4.13: Loop gain and phase of the CC

4.5 Preamplifier's Simulations

The Fig. 4.14 shows how we studied the performances of the Preamplifier (INGRESSO), feeding the input with an exponential current generator. This emulates a current coming out from a SiPM. The signal generator is characterized by a current1 of 0 A (starting value of the exponential signal) and a current2 of 20 uA (peak of the exponential signal), two delay times of 100 ps and a damping factor1 (τ_{rise}) of 500 ps and the second one (τ_{fall}) of 50 ns, and at the output a voltage generator vdc of 2 V (DC voltage generator). The damping factors are actually $\tau_{rise} and\ \tau_{fall}$ characterizing the exponential waveform.

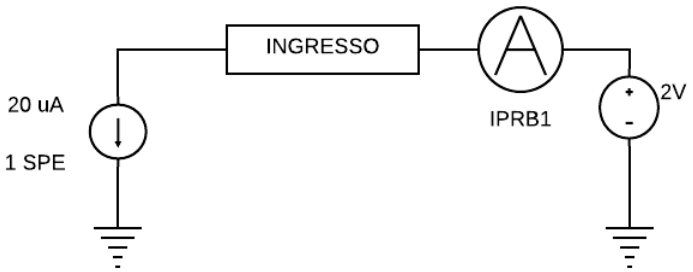

Fig. 4.14: Preamplifier studying simulation

First of all, we started analyzing the performance stability of the circuit. Fig. 4.15 shows the loop gain (M2-M3-M15-M7) magnitude and the phase. As it is clear from the figure, the circuit is stable, having a phase margin of 54° and a GBWP almost equal to 62 MHz.

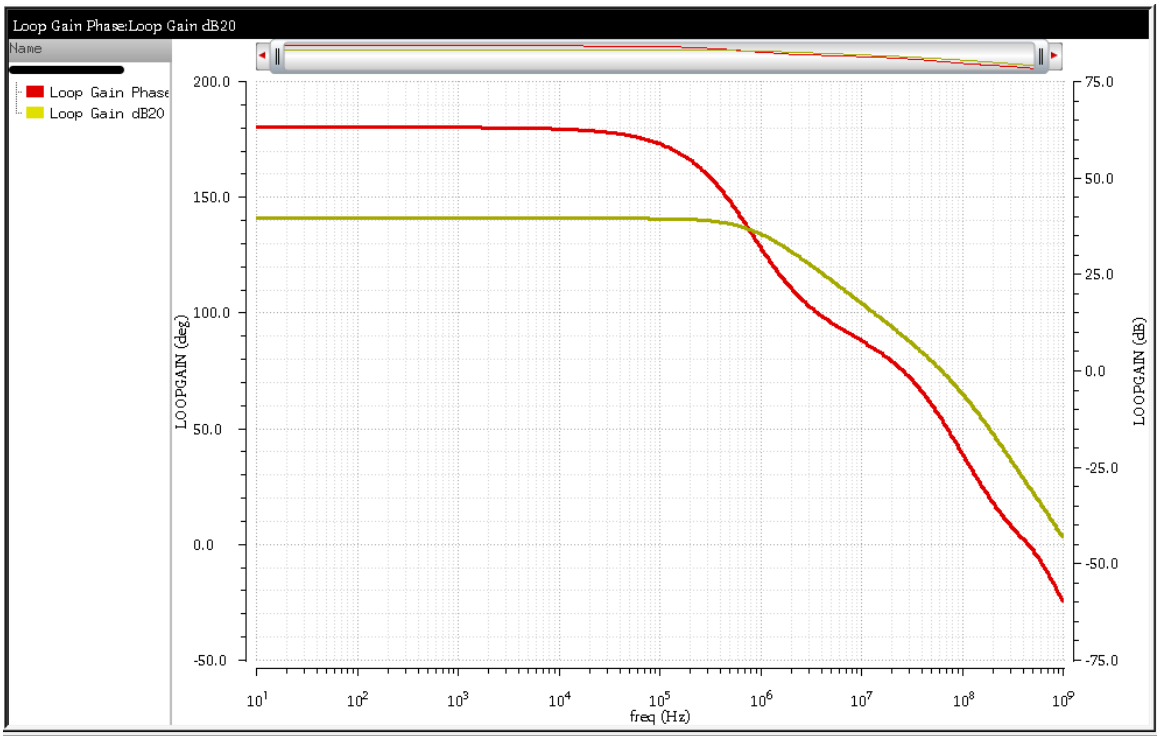

Fig. 4.15: Loop gain and phase trends

Before going through the other simulations than involve the preamplifier, it is important to get some information about the power dissipated (P) by the circuit. Indeed, P results to be 4.6 mW. The total power consumption is bigger than the one computed in theory (3 mW). Anyway, our target for this stage was around 5 mW of power dissipated at most, so we reached the goal.

Moving forward, to better estimate of the bandwidth of the Preamplifier, we resort to an AC simulation (see Fig. 4.16).
As we can see from the figure 4.16, there is an overshoot equal to 3.8 mA for an input current of 1 A (linear analysis) in correspondence of the GBWP. At -3 dB, the bandwidth of the system is equal to 103.9 MHz, slightly larger than our target of 100 MHz.

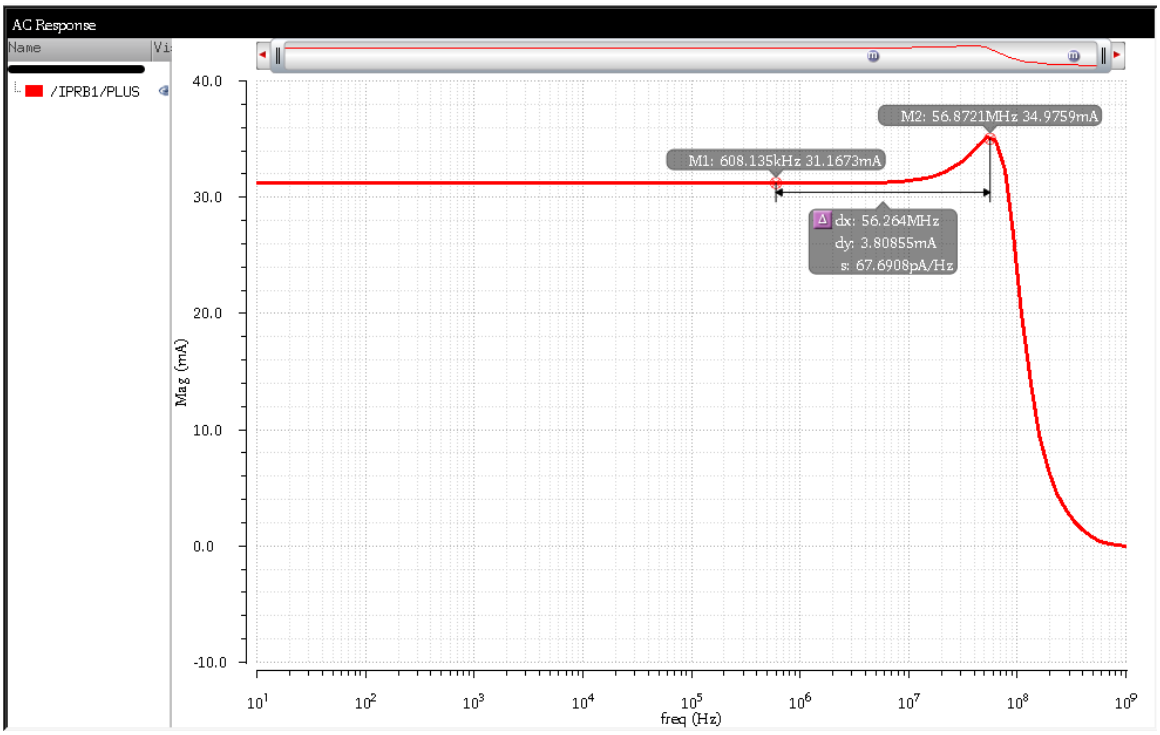

Fig. 4.16: AC response

In Fig. 4.17, the total output current power spectral density noise (A^2 / Hz) is shown. Taking into account the root of the power spectral density, there is a peak at 87 MHz that is equal to 16.6 pA/$\sqrt{Hz}$ (overshoot effect). The integral output noise in the amplifier's BW is equal to 37.8 nA. This value reported at the input of the amplifier (37.8 nA/0.04 = 945 nA) and compared to the peak current generated by a SPE (20 uA) results to be smaller. Anyway, the noise is not a requirement of the thesis.

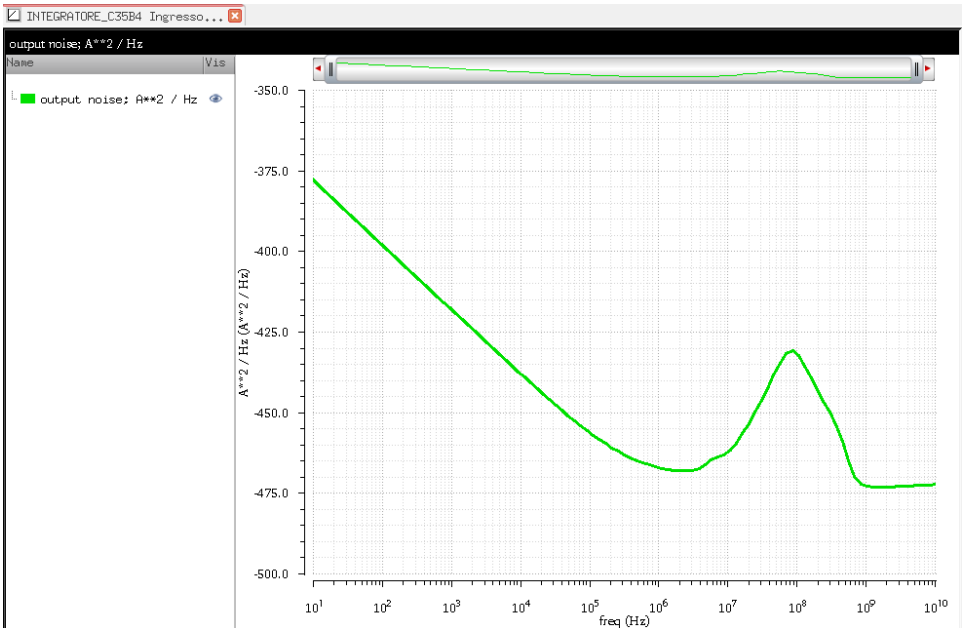

Fig. 4.17: Noise behavior at the output

Hereafter, we explain how to eliminate the overshoot effect. This is not part of thesis objective. It is just worth to have a better understanding on how the bandwidth and the noise are linked to each other.
To do so, we introduced a capacitance (1 pF) in correspondence of enhanced cascode mirror of M10, which is able to filter the aforementioned effect. Indeed, comparing Fig. 4.17 and Fig. 4.18, the peak at the 87 MHz has been alleviated in Fig. 4.18. Introducing the filter capacitance, the peak reaches 4.4 (pA/$\sqrt{Hz}$).

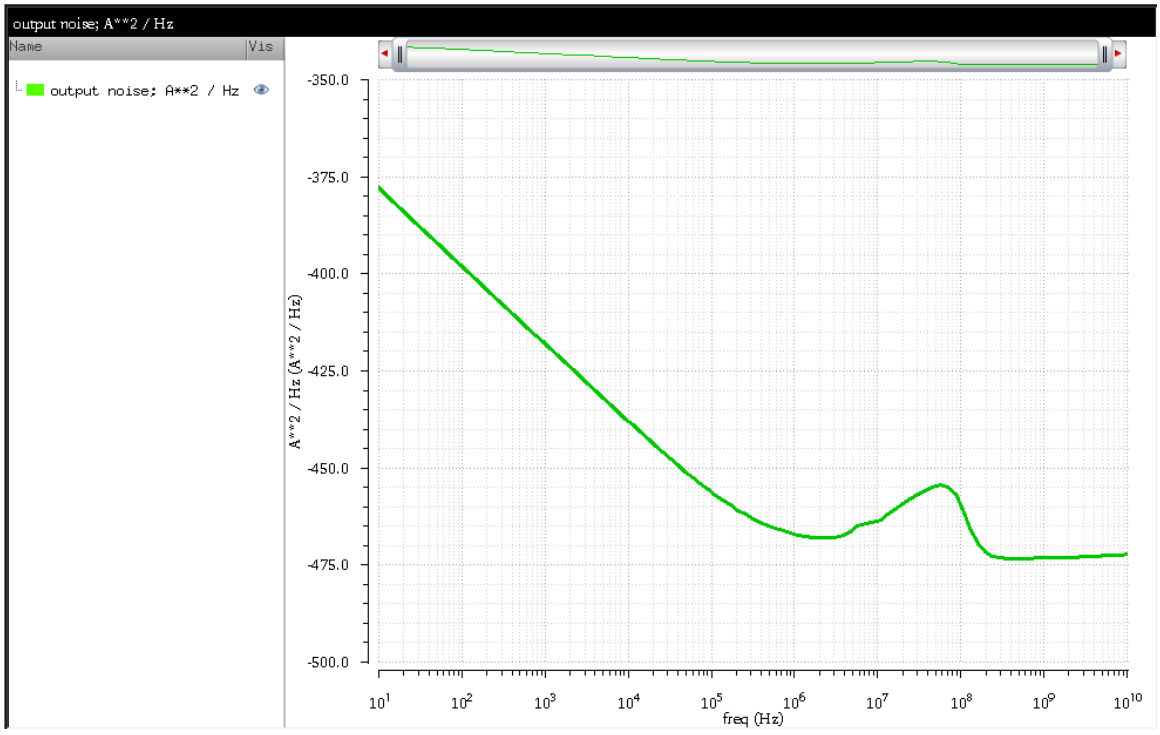

Fig. 4.18: Output current noise expressed as power noise spectral density (red line) and root of the power spectral density (yellow line)

Taking into account the AC simulation, introducing the capacitance at the gate of M10, we would be able to eliminate the overshoot effect, as we can see in Fig. 4.19. At the same time there is a disadvantage due to the fact that the bandwidth is now 19.29 MHz at -3 dB.
The filter capacitance is just able to cancel the peak and flatten the curve, worsening the BW. This capacitor reduces the BW since

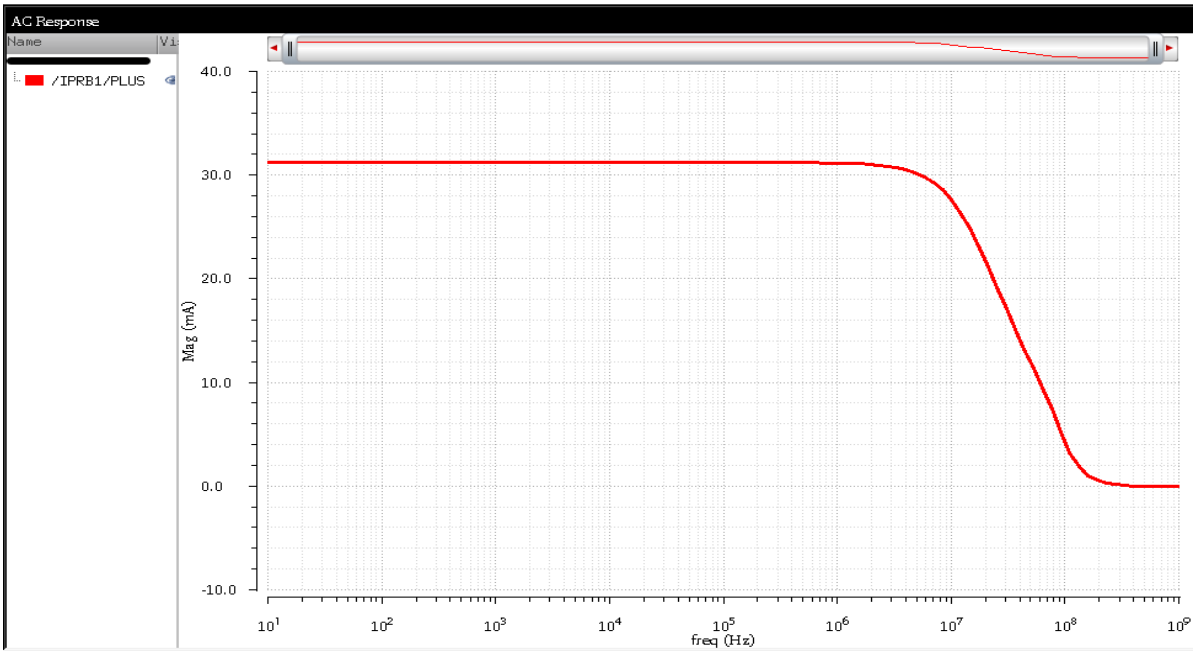

Fig. 4.19: AC response with

Summing up, there is a clear trade-off between noise and bandwidth due to the capacitance used for filtering. Being our goal to reach a BW of 100 MHz, we do not insert any capacitance to filter the overshoot effect. Obviously, the noise performance is worse than the case with the filter capacitance, but there were not limitations on noise in the specifications of the thesis. Thus, it is not worth to insert the filter capacitance.

Chapter 5

Slow Chain

In the fifth chapter, we focus on the slow chain (see Fig. 5.1) presenting the switches with the relative signals connected to them, the internal OTA structure of the gated integrator; successively we concentrated on the implementation of a Track and Hold (T/H) and the ADC SAR. Considering the SAR, we developed a comparator needed to link the voltage saved on the T/H and the signal coming out from the DAC, and ultimately a programmed logic in order to change or not the single bit from the MSB to the LSB.
At the end of each section, the Cadence simulations for each block are presented in order to determine the real performances and to compare them with the theoretical computations/evaluations.

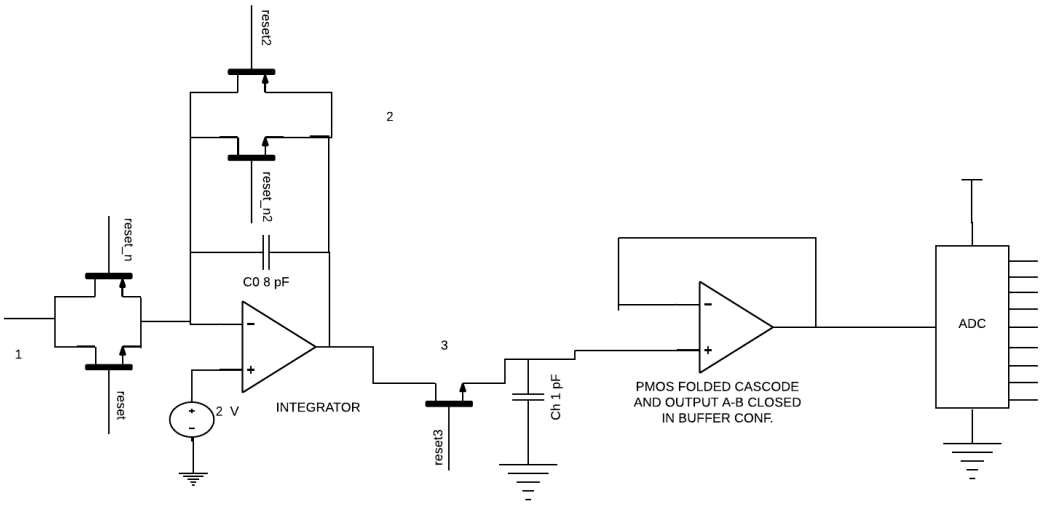

Fig. 5.1: Slow chain

5.1 Switches

In the slow chain there are three switches which are controlled by signals produced by step generators, characterized by their period, delay time, fall and rise times and pulse width. In order to understand how the switches work, we refer to Fig. 5.2. As soon as reset triggers from 0V to 3.3 V and reset_n triggers from 3.3 V to 0 V, the first switch closes. If there is a current, being the result from SPEs, it is transferred to the input of the integrator through the amplifier. Regardless, this is not our case because the current will just arrive at the input of the integrator after 10.41 us.
In the meantime, the second switch (switch in the feedback of the integrator) remains opened (indeed its control generator has a delay) and the capacitance Co

is at 0V. Regarding the third switch, it is closed from the beginning of the simulation to 20 us and then, it opens. At 210 ns, the second switch closes, charging the capacitance (C0) at 2 V. This switch keeps closing for 200 ns and then it opens, leaving the capacitance charged at 2 V. The capacitance of the integrator will be at 2 V till 10.41 us and, right after this time, the signal coming from the amplifier, arrives at the input of the integrator. The integration process starts and the capacitance discharges to a voltage equal to 300 mV (in the case at the input there are 440 pe). The integration period will last for 8.68 us. Being the third switch still closed at the end of the integration (8.68 us + 10.41 us = 19.09 us), the voltage in C0 will be saved on Ch and transferred through the OTA at the input of the ADC.
When the feedback switch closes again at 20.21 us, C0 will be charged again to 2 V, while Ch will be at 300 mV (voltage at the end of the integration in C0) since at 20 us switch3 opened.
Referring to Fig. 5.2, out is the singal at the output of the integrator (from 10.41 us the discharge of C0 can be appreciated), I30/OUT is the total current at the output of the Preamplifier and then, the other signals are the ones that control the switches, as already explained.

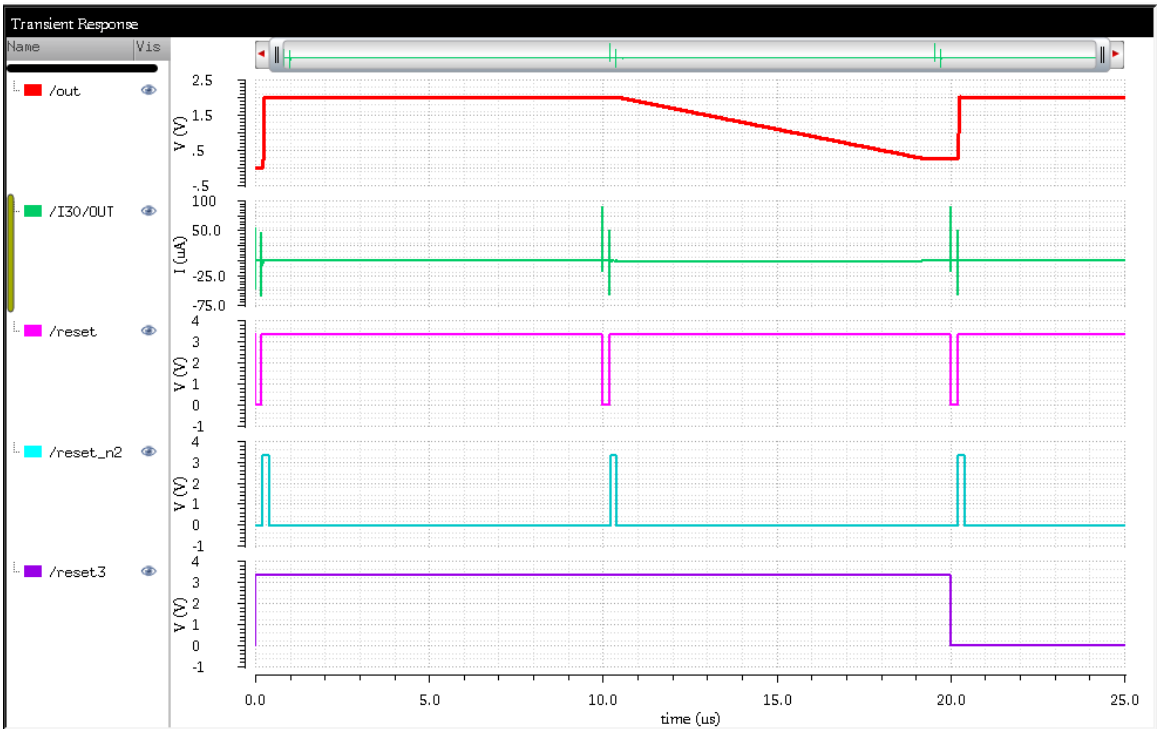

Fig. 5.2: Signals controlling the switches of the system and the signals at the output of the integration (out) and in the sample capacitance (out_sample)

5.2 Gated Integrator

This is one the most important part of the entire chain and the one we focused on more. The internal configuration of the gated charge integrator is basically composed by three parts: a bias stage (as the name says, it is necessary to bias the entire circuit), PMOS folded cascode input stage and a Monticelli's topology at the output. A folded-cascode input stage is used to provide a high open-loop gain. The common-gate level shifter circuit is inserted in the output stage of the folded-cascode differential stage.
First of all, let's focus on a new the output stage (see Fig. 5.3). This structure [33] is necessary to understand the working principle of the Monticelli output stage.

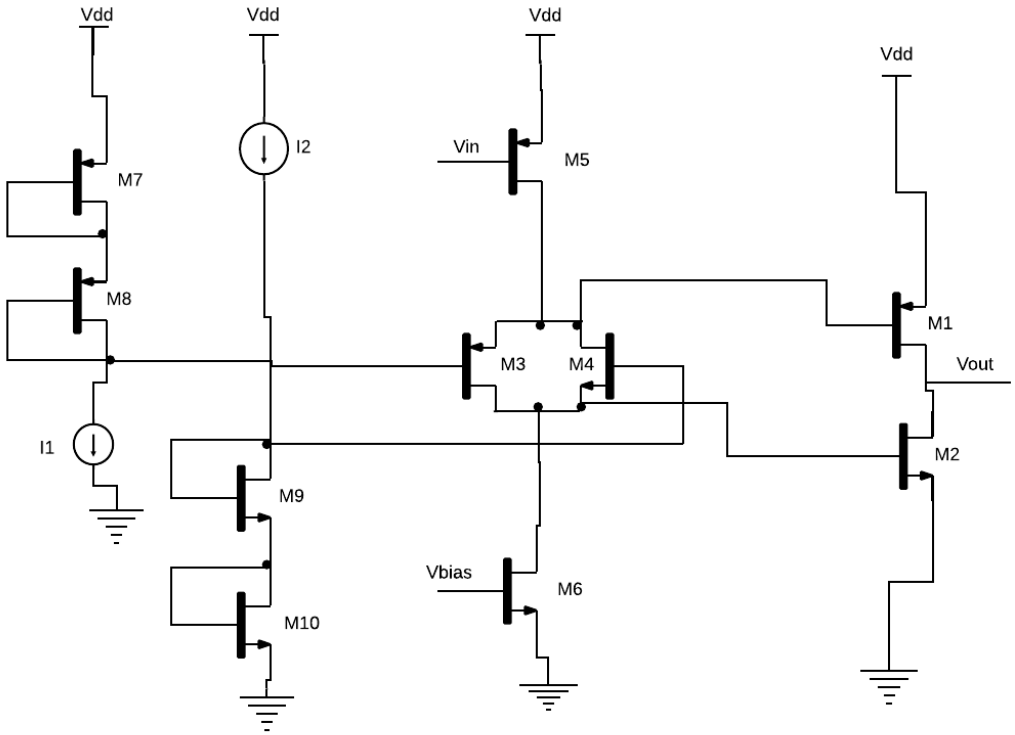

Fig. 5.3: Rail-to-rail output stage with common-gate level shifter

The circuit consists of the push-pull output stage made by M1 and M2. In this circuit, transistors (M5, M6) form a common source configuration where the input signal alters the relative conduction levels of the common-gate devices (M3, M4). The class A-B bias circuit sets up the two loops (M1, M3 and M2, M4), that fixes the voltage drop between the gates of the output devices.
The quiescent conditions of the output stage are established as follow. The complementary currents I1 and I2 (I1 = I2 = I) flow into complementary stacks of diode-connected transistors M7, M8 and M9, M10, whose drain potentials are used to bias the gates of the common-gate transistors M4 and M3. In steady state, the current through M6 is equally divided between M3 and M4.
At this point, let's assume that:

$$(W/L)_4 = (W/L)_9 = (W/L)_{10}$$

$$(W/L)_3 = (W/L)_7 = (W/L)_8$$

and since M3, M8 and M4, M9 carry the same drain currents, they will also have the same gate-to-source voltages, and hence we have Vs4 (=Vd3) = Vg10 = Vs9 and Vs3 (=Vd4) = Vg7 = Vs8. As a result, M2, M10 and M1, M7 will have the same gate-to-source voltages as well, being the output current in steady state equal to:

$$Iout = I * \frac{\left(\frac{W}{L}\right)_1}{\left(\frac{W}{L}\right)_7} = I * \frac{(W/L)_2}{(W/L)_{10}}$$

When the output stage is driven to sink a large load current, Vin goes low (common source) and pulls Vs3 and Vs4 up to high level close to Vdd. Under this condition, M4 is turned off and M3 carries the full current coming from M5. The source of M3 moves up to its maximum point, therefore cutting back on the conduction of M1. The drain of the common-gate device M3 rises as well, dragging the gate of M2 at high level, becoming highly conductive.
On the other hand, under the conditions of strong sourcing, Vin goes high (common source), and Vs3 and Vs4 are pulled down. In this case, M3 is shut off and M4 carries the full current. The source of M4 pulls the gate of M1 at low level, making it highly conductive.
Once the working principle of the Monticelli is clear, we can study the total circuit of the gated integrator. Hence, moving to the circuit in Fig. 5.4 [33], we will explain the DC behavior, coming up with the values in Tab. 5.1. We decided to set the GBWP equal to 200 MHz. This decision is due to the specifications of the project.
Before computing the bias of the differential pair transistors (M11 and M12) through the GBWP equation, it is necessary to compute the impedance we see at the output of the first stage (Z0) and the gain of the second stage (G2 = (gm1 + gm2) * R2).
Regarding the output of the first stage (source of M3), the equivalent output impedance is not just the parallel between the resistance at the source of the Monticelli stage's transistor (1/gm) and the equivalent impedance of the tail generator (Z_{s1}), that is composed by the cascode configuration including M5A and M7A, connected to the Monticelli. We also have to take into account the effect of the Gloop of the Monticelli that decreases the impedance (when the loop does not act). For this reason, we came up with:

$$Z_0 = \left(\frac{1}{gm_3} // Z_{s1}\right) * \frac{1}{1 - A * \beta}$$

In the formula above $A * \beta$ is the Gloop of the Monticelli structure and it is given by $A * \beta = \frac{gm_3 \cdot Z_{s1}}{1+gm_3 \cdot Z_{s1}} \cdot \frac{gm_4 Z_{s2}}{1+gm_4 Z_{s2}}$. Specifically, Zs1 and Zs2 are the equivalent impedances of the tail generators connected to the source of M3 and to the source of M4 respectively.
After some computation it is easy to verify that:

$$Z_0 = \frac{1}{gm_3} * \frac{K}{1-K^2} = \frac{1}{gm_3} * 9.74 \ = 7.21\ k\Omega$$

Where K is equal to $\frac{gm_3 \cdot Z_{s1}}{1+gm_3 . Z_{s1}}$. In our case, $K = \frac{gm_3 \cdot (r_{5A} \cdot gm_{5A} \cdot r_{7A})}{1+gm_3 \cdot (r_{5A} \cdot gm_{5A} \cdot r_{7A})} \approx 0.95$.

On the other hand, the gain of the second stage is:

$$G2 = (gm1 + gm2) * (r1//r2) = (gm1 + gm2) * R2 = 82.4 = 38.32\ dB$$

At this time, we are able to compute the minimum gm11 and then I11 that respect this condition on the GBWP, multiplying the differential gain and the dominant pole (introduced by the Miller capacitance of 1 pF) of the OTA:

$$GBWP = Gdiff * f1 = \frac{gm11 * R1 * (gm1 + gm2) * R2}{2 * \pi * Cc * (gm1 + gm2) * Z0 * R2} = \frac{gm11}{2 * \pi * Cc}$$

$$gm11 \geq \ GBWP * \ 2 * \pi * Cc \geq 1.26\ mA/V$$

From the computation above, we ended up picking gm11 equal to 1.3 mA/V. Therefore, the new GBWP will be 207 MHz. Fixing gm11, we can compute the gain of the first stage, being $G1 = Zo * gm11 = 9.38$.
In a second moment we calculated the poles and the zeros of the system in order to check the stability of the entire system.
The dominant pole of the system is introduced by the Miller capacitances (it can also be seen in the formula to get the GBWP at the denominator):

$$f1 = \frac{1}{2 * \pi * Cc * (gm1 + gm2) * Z0 * R2} = 268\ KHz$$

Once the effect of the Miller capacitance has been taken into account, it is the time to calculate the other poles in the circuit. The second pole is due to the load capacitance (1 pF) in parallel, since Cc is already closed, with the contributions of the gate/source capacitance of M1 (Cgs1), and of the drain/source capacitance of M5A (Cds5A) and of the drain/source capacitor of M4 (Cds4). Obviously, the load capacitance is in parallel as well with the contributions given by the transistors M2 (Cgs2 gate/source capacitor) and, regarding the drain/source, by M6A (Cds6A) and M3 (Cds3). We will refer to Cgs as the sum of Cgs1 and Cgs2, while the total drain/source capacitance is labeled as Cds. Hence, $C_{gs=} \frac{2}{3} * C'_{ox} * \big((W1 * L1) + (W2 * L2)\big) = 1.48\ pF$ and $C_{ds} = C'_{db} * (3 * L_{min}) * (W5A + W4 + W6A + W3) = 70.6\ fF$.
Being Cgs and Cds in parallel, they can be summed up. The sum of these two capacitances is indicated as C1 and it is equal to $C_1 = C_{gs} + C_{ds} = 1.55\ pF$. At this point, we can compute the sum of C1 and Cl, hence the frequency of the second pole. Being $C_2 = C_1 + \ C_L = 2.55\ pF$, the frequency of the second pole (f2) is:

$$f2 = \frac{(gm1 + gm2)}{2 * \pi * C_2} = 1.87\ GHz$$

In the formulas, we included the contribution of the area of the drain regions, $C'_{db} = 0.1\ \frac{fF}{um^2}$, $3 * L_{min} = 1.05\ um$, as spacing of the drain contact areas and $C'_{ox} = 5\frac{fF}{um^2}$, as oxide capacitance.
Moving forward, we show how we get the expression of Rn (nulling resistor), estimating the zero, introduced by the Miller capacitance. The zero will be fixed at the same frequency of the second pole in a way to cancel out its effect.

$$fz = f2 = \frac{1}{2 * \pi * Cc * \left(Rn - \left(\frac{1}{gm1 + gm2}\right)\right)}$$

$$Rn = \frac{1}{(2 * \pi * Cc * f2) + (\frac{1}{gm15 + gm16})} = 118\ \Omega$$

Once the value of the nulling resistor has been computed, the frequency of the third pole of the system can be calculated as follow:

$$f3 = \frac{1}{2 * \pi * R_n * (C_c // C_L // C_1)} = 1.41 * 10^{10} Hz$$

Finally, it remains to mention the effect of the zero due to the Miller capacitances:

$$fz = \frac{1}{2 * \pi * Cc * \left(Rn - \left(\frac{1}{gm1 + gm2}\right)\right)} = 1.87\ GHz$$

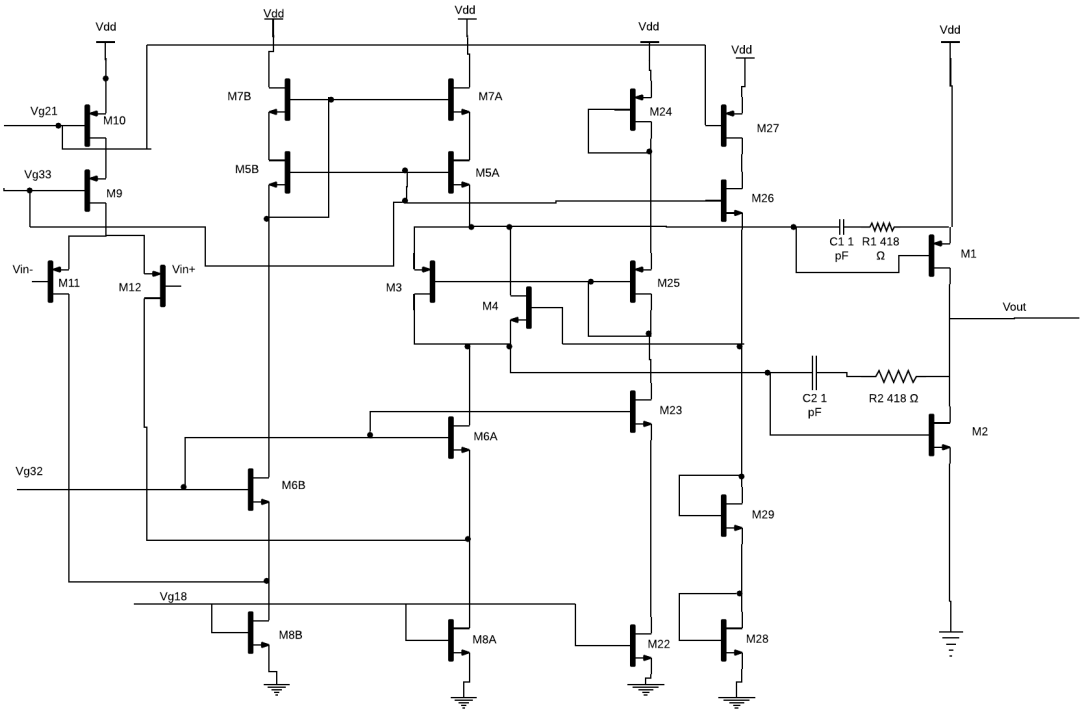

Fig. 5.4: Final OTA's structure (being part of the integrator) with Monticelli block composed by M3 and M4

The opamp of Fig. 5.4, with an output stage that uses common-gate level shifter to bias the push-pull output devices, is very compact and power efficient (the static consumption is concentrated the output stage). If we applied a symmetric small signal at the input of the block, we would have the same behavior for the transistors at the top and the ones at the bottom of the system.
Once the stability analysis is over, we can compute the sizing of the remaining transistors of Fig. 5.4 and the relative parameters, coming up with Tab 5.1:

	I(uA)	Vov(V)	gm(mA/V)	W/L	W(um)	L(um)	rd (Kohm)
M11, M12	130	0.2	1.3	130	130	1	154
M9, M10	260	0.2	2.6	260	260	1	76.9
M6A, M6B	270	0.2	2.7	135	135	1	74.1
M8A, M8B	400	0.2	4	200	200	1	50
M7A, M7B	270	0.2	2.7	270	270	1	74.1
M5A, M5B	270	0.2	2.7	270	270	1	74.1
M3	135	0.2	1.35	135	135	1	148
M4	135	0.2	1.35	67.5	67.5	1	148
M24, M25	135	0.2	1.35	135	135	1	148

M28, M29	135	0.2	1.35	67.5	67.5	1	148
M1	1500	0.2	15	1500	750	0.5	6.67
M2	1500	0.2	15	750	375	0.35	4.67

Tab. 5.1: T/H's PMOS folded cascode with A-B output sizing

Once the differential pair transistors, the Monticelli's structure and the output transistors have been sized, it is the time to size the bias transistors (Fig. 5.5).

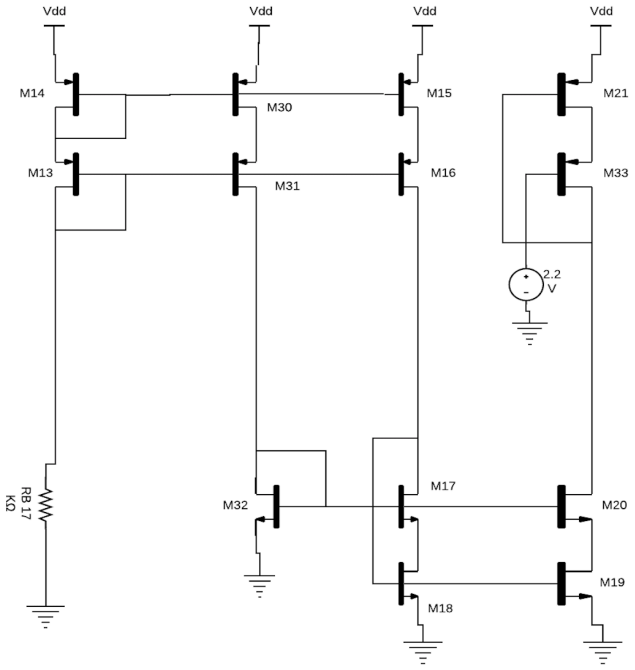

Fig. 5.5: Bias transistors

In Tab 5.2 there are the parameters of the transistors involved for biasing the differential stage and the output stage.

	I(uA)	Vov(V)	gm(mA/V)	W/L	W(um)	L(um)	rd (Kohm)
M13, M14	100	0.2	1	100	100	1	200
M30, M31	100	0.2	1	100	100	1	200
M15, M16	400	0.2	4	400	400	1	50
M32	100	0.5	0.4	8	8	1	200
M17, M18	400	0.2	4	200	200	1	50
M19, M20	400	0.2	4	200	200	1	50
M21, M33	400	0.2	4	400	400	1	50

Tab. 5.2: Sizing of the bias transistor

The power consumption of the system (P) is obtained taking into account the DC currents flowing in the transistors of the circuits in Fig. 5.7 and Fig. 5.8. Thus, the total power dissipated by the circuit results to be:

$$P = Vdd * (I_{14} + I_{30} + I_{15} + I_{20} + I_{10} + I_{7A} + I_{7B} + I_{24} + I_{27} + I_1) = 10.7\ mW$$

The target in terms of power consumption for this stage was to dissipate at most 12/13 mW.

5.2.1 Noise Evaluation

The goal of this section is to carry out the computation of the input referred current noise of the OTA stage and then $vout_{rms}$ (since the current will be integrated by the integrator). Being a two-port circuit, the input-referred noise sources will not depend on the load impedance. It is therefore convenient to short the output node to ground, thus looking at the output short circuit current noise. In order to calculate the input referred current noise [34] we consider the circuit in Fig.5.4 when the components are noiseless, and the input referred current noise source is driving the differential stage. Note that for this current to flow through the input port we should take into account the capacitive impedance across the gate/source terminals. Being the i_n a noise signal from the input-referred current source that flows through C_{gs} (gate/source capacitance). This signal drives a current (i_{n*} gm_{11})/s* C_{gs11} through M11. Then, by symmetry, the same current flow through M12 and the corresponding C_{gs}. Since i_n flows into C_{gs} thus making the component (i_{n*} gm_{11})/s*C_{gs11} drained by M12. Therefore, the result from the differential current flowing into the input pair is an output current of (2*i_{n*} gm_{11})/s* C_{gs11}. In terms of noise power spectral density:

$$S_{out}^{\infty} = 4 \cdot Si_{eq}^2 \cdot (\frac{\omega_T}{\omega})^2 \cdot Z_0^2 \cdot (gm_1 + gm_2)^2$$

Being $\omega_T = gm_{11}/C_{gs11}(C_{gs11} = \frac{2}{3} \cdot C'_{ox} \cdot (W_{11} \cdot L_{11}) = 0.4\ pF$ and $\omega = (2 \cdot \pi \cdot GBWP)/2$.

In order to compute the input referred current noise source we will firstly go through the computation of the input-referred voltage noise source. Once the input-referred voltage noise source is known, we will link it to the input referred current noise source (our purpose).

To compute the input-referred voltage noise source, the output noise in the circuit model, where a voltage generator equal to Sv_{eq}^2 is connected at the input differential pair and the components of the circuit (Fig. 5.4 and Fig. 5.5) are considered noiseless, is compared to the output noise in the real circuit (Fig. 5.4 and Fig. 5.5) where a short is made between the input terminals.

When Sv_{eq}^2 is connected to the input differential pair (no other noise contributions), it can be seen that the short circuit current noise power is linked to Sv_{eq}^2 by:

$$S_{out}^{\infty} = Sv_{eq}^2 \cdot (gm_{11})^2 \cdot Z_0^2 \cdot (gm_1 + gm_2)^2$$

Let us now derive the contribution (for each component) to the output voltage power spectral density in the real circuit (Fig. 5.5) and comparing it to the output current noise in the circuit driven by the input referred noise, it turns out:

$$Sv_{eq\ M11,M12}^2 = \frac{4KT\lambda \cdot (gm_{11} + gm_{12})}{{gm_{11}}^2} = 1.64 \cdot 10^{-17} \frac{V^2}{Hz}$$

$$Sv_{eq\ M7A,M7B}^2 = \frac{4KT\lambda \cdot (gm_{7A} + gm_{7B})}{{gm_{11}}^2} = 3.42 \cdot 10^{-17} \frac{V^2}{Hz}$$

$$Sv_{eq\ M8A,M8B}^2 = \frac{4KT\lambda \cdot (gm_{8A} + gm_{8B})}{{gm_{11}}^2} = 5.06 \cdot 10^{-17} \frac{V^2}{Hz}$$

We also have to take into account the contributions due to the transistors that bias the circuit (Fig. 5.5).

$$Sv_{eq\ M18}^2 = \frac{4KT\lambda \cdot gm_{18} \cdot (\frac{gm_{8B}}{gm_{18}})^2}{{gm_{11}}^2} = 2.53 \cdot 10^{-17} \frac{V^2}{Hz}$$

$$Sv_{eq\ M18}^2 = \frac{4KT\lambda \cdot gm_{18} \cdot (\frac{gm_{8A}}{gm_{18}})^2}{{gm_{11}}^2} = 2.53 \cdot 10^{-17} \frac{V^2}{Hz}$$

$$Sv_{eq\ M18}^2 = \frac{4KT\lambda \cdot gm_{18} \cdot (\frac{gm_{8B}}{gm_{18}})^2}{{gm_{11}}^2} = 2.53 \cdot 10^{-17} \frac{V^2}{Hz}$$

$$Sv_{eq\ M18}^2 = \frac{4KT\lambda \cdot gm_{18} \cdot (\frac{gm_{22}}{gm_{18}})^2 \cdot (\frac{gm_3}{gm_{25}})^2}{{gm_{11}}^2} = 2.93 \cdot 10^{-18} \frac{V^2}{Hz}$$

It remains to compute the noise (power spectral density) associated to M13, M14, M15, M16, M30, M31, M32 and Rb. We will refer to this contribution to the total output power spectral density as $Sv_{eq\ BiasInput}^2$.

$$Sv_{eq\ BiasInput}^2 = \frac{4KT\lambda \cdot (gm_{32} + gm_{30} + gm_{31} + gm_{14} \cdot \left(\frac{gm_{30}}{gm_{14}}\right)^2 + gm_{13} \cdot \left(\frac{gm_{31}}{gm_{13}}\right)^2 + \frac{1}{R_b} \cdot \left(\frac{gm_{31}}{gm_{13}}\right)^2) \cdot (\left(\frac{gm_{17}}{gm_{32}}\right)^2 \cdot \left(\frac{gm_{8A}}{gm_{18}}\right)^2)}{{gm_{11}}^2} = 2.79 \cdot 10^{-15} \frac{V^2}{Hz}$$

$$Sv_{eq\ BiasInput}^2 = \frac{4KT\lambda\cdot(gm_{32}+gm_{30}+gm_{31}+gm_{14}\cdot\left(\frac{gm_{30}}{gm_{14}}\right)^2+gm_{13}\cdot\left(\frac{gm_{31}}{gm_{13}}\right)^2+\frac{1}{R_b}\cdot\left(\frac{gm_{31}}{gm_{13}}\right)^2)\cdot((\frac{gm_{17}}{gm_{32}})^2\cdot\left(\frac{gm_{8B}}{gm_{18}}\right)^2)}{gm_{11}{}^2} = 2.79\cdot 10^{-15}\frac{V^2}{Hz}$$

$$Sv_{eq\ BiasInput}^2 = \frac{4KT\lambda\cdot(gm_{32}+gm_{30}+gm_{31}+gm_{14}\cdot\left(\frac{gm_{30}}{gm_{14}}\right)^2+gm_{13}\cdot\left(\frac{gm_{31}}{gm_{13}}\right)^2+\frac{1}{R_b}\cdot\left(\frac{gm_{31}}{gm_{13}}\right)^2)\cdot((\frac{gm_{17}}{gm_{32}})^2\cdot(\frac{gm_{22}}{gm_{18}})^2\cdot\left(\frac{gm_3}{gm_{25}}\right)^2)}{gm_{11}{}^2} = 3.22\cdot 10^{-16}\frac{V^2}{Hz}$$

$$Sv_{eq\ M15,M16}^2 = \frac{4KT\lambda\cdot(gm_{15}+gm_{16})\cdot(\frac{gm_{8A}+gm_{8B}+gm_{22}}{gm_{18}})^2}{gm_{11}{}^2} = 2.2\cdot 10^{-16}\frac{V^2}{Hz}$$

Summing up all the contributions:

$$Sv_{eq}^2 = 6.3\cdot 10^{-15}\frac{V^2}{Hz} = (79\ \frac{nV}{\sqrt{Hz}})^2$$

Knowing Sv_{eq}^2, it is possible to compute Si_{eq}^2, that is actually our target, from the following expression:

$$Si_{eq}^2 = Sv_{eq}^2\cdot\frac{gm_{11}^2}{4}\cdot\left(\frac{\omega}{\omega_T}\right)^2 = 9.6\cdot 10^{-23}\frac{A^2}{Hz} = (9.8\ \frac{pA}{\sqrt{Hz}})^2$$

The input referred current noise will be then integrated by the integrator [35]:

$$vout_{rms}^2 = Si_{eq}^2\cdot\frac{T_s}{C_0^2} = 5\ nV^2$$

$$vout_{rms} = \sqrt{vout_{rms}^2} = 70.7\ uV$$

This value will be transferred to the input of the ADC through the OTA closed in buffer configuration.
In this circuit we have also to take into account the noise $\frac{K\cdot T}{C}$ due to the presence of the switch in the feedback configuration. Indeed, due to the thermal noise originating from the "on" resistance of the switches, a $\frac{K\cdot T}{C}$ noise in injected in the switched capacitor. The $\frac{K\cdot T}{C}$ contribution is:

$$\frac{K\cdot T}{C_0} = (22.3\ uV)^2$$

In the case the capacitor was 200 pF, the $\frac{K \cdot T}{C}$ introduced would be:

$$\frac{K \cdot T}{C} = (4.47\ uV)^2$$

That is smaller than the case of the feedback capacitance equal to 8 pF. Nevertheless, picking the feedback capacitance equal to 8 pF, we have advantages in terms of the cost being the area much smaller (8 pF) compared to the one with 200 pF, as already explained. There is a trade-off between cost (area) and noise.

5.2.2 OTA and Monticelli's topology in Cadence

The Fig. 5.6 displays how we analyzed the performances of the Integrator (the OTA with the Monticelli topology is closed in a buffer configuration), feeding the input with a pulse voltage generator. The generator is characterized by a step signal equal to 2 V (the signal goes up and down with a step shape (width equal to 100 ns), starting from 0 V to 2 V, over a period of 200 ns; delay time of 2 ns and rise and fall times equal to 100 ps), a DC value of 2 V, an AC magnitude of 1 V.
Moreover, a load capacitance (8 pF) is connected at the output of the structure.

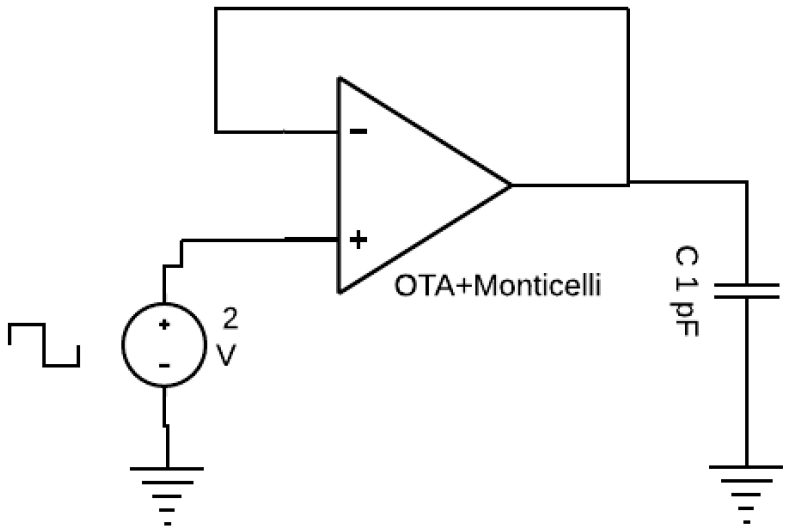

Fig. 5.6: OTA + Monticelli structure's TB (test bench)

First of all, we evaluated the phase and the gain of the OTA (see Fig. 5.7), being part of the integrator, are shown. In detail, the phase margin of the system is equal to 49.7° and the GBWP of the system is 69.79 MHz. Our GBWP target was 200 MHz. We could not achieve the frequency we estimated in theory. In fact, if we wanted to reach the 200 MHz (GBWP) we would need to increase the overall current, worsening the power consumption of the circuit. The total power dissipated by the system is equal to 10.27 mW, almost equal to the value we got in theory. The requirement on the power dissipation has been satisfied.
Summarizing, there is a trade-off between the power consumption and the BW.

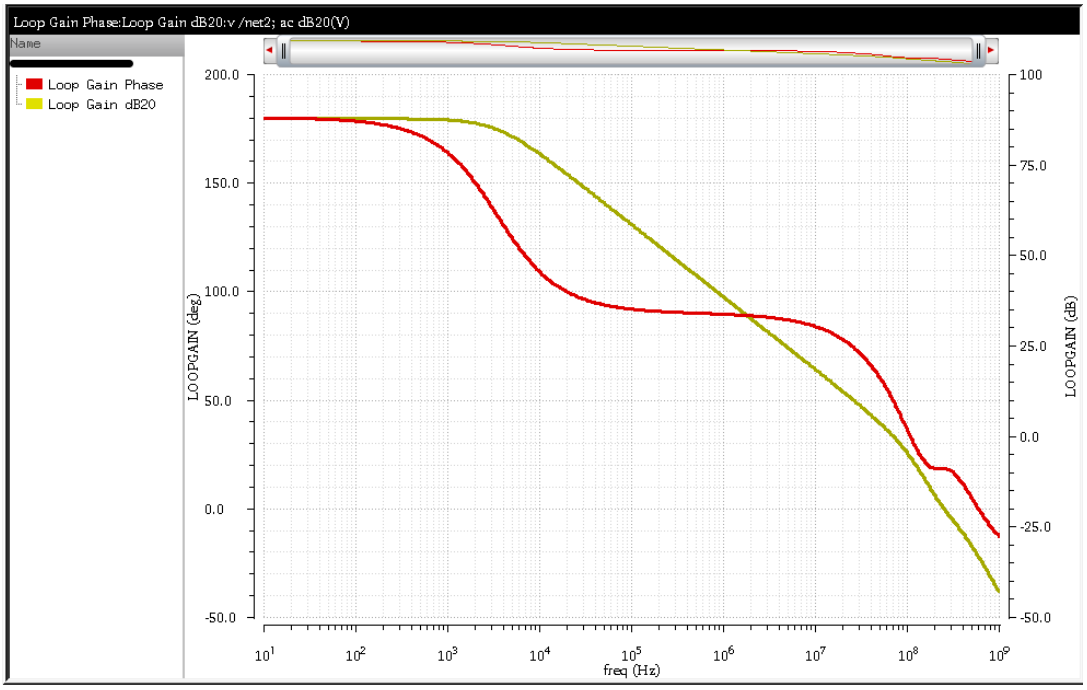

Fig. 5.7: Phase and gain of the loop of the OTA composing the integrator

Moreover, we performed the AC simulation of the OTA for an input voltage of 1 V (see Fig. 5.8), highlighting the net2 that is actually the voltage at the output of the stage. From the graph, we computed the bandwidth of the system that is equal to 111 MHz in correspondence of -3 dB.

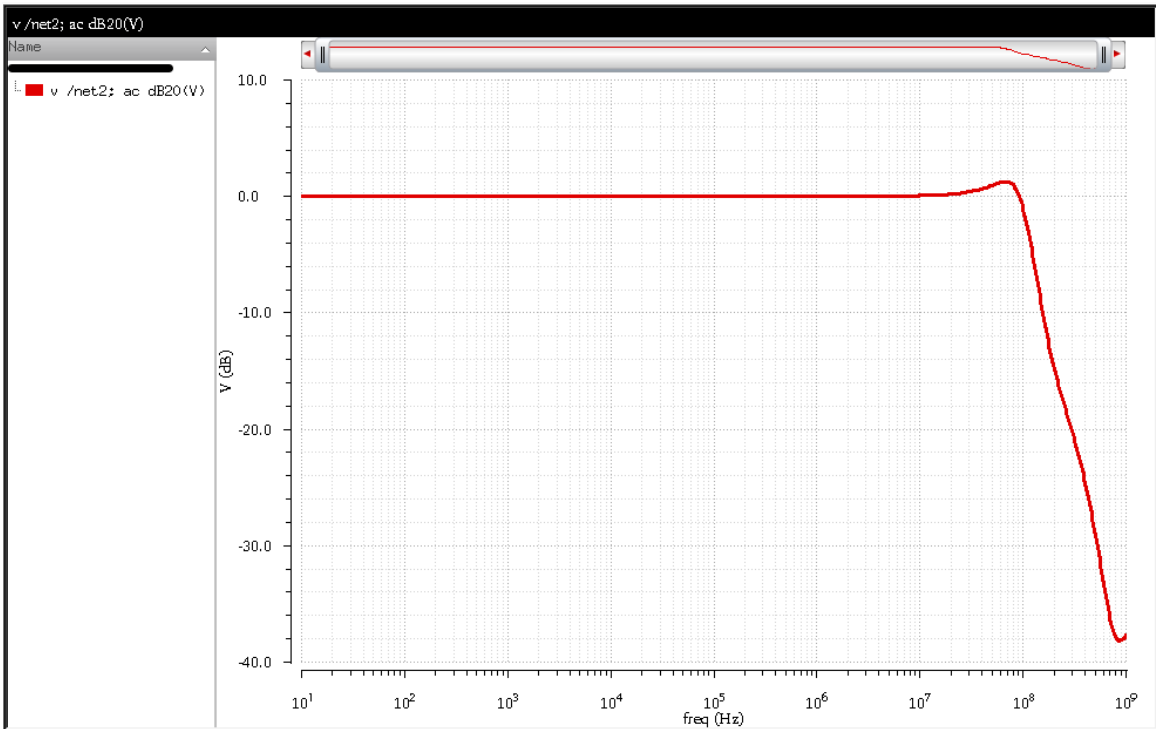

Fig. 5.8: AC response of the OTA with Monticelli

The transient behavior of the OTA, in response of a step singal at its input, is shown in Fig. 5.9. The limit on the SR in the output capacitance is too big even though the output current is 1 mA. This effect is due to a problem in the dynamic of the output supply. The signal changes a bit; hence we should have a greater swing. The rise and fall time are too big, in fact they should be around 5/6 ns and they are actually 18 ns and 14 ns. In order to correct this effect, we should decrease the Miller capacitor and increase the tail current connected to the Monticelli. There is a trade-off between the transient response and the stability of the system.

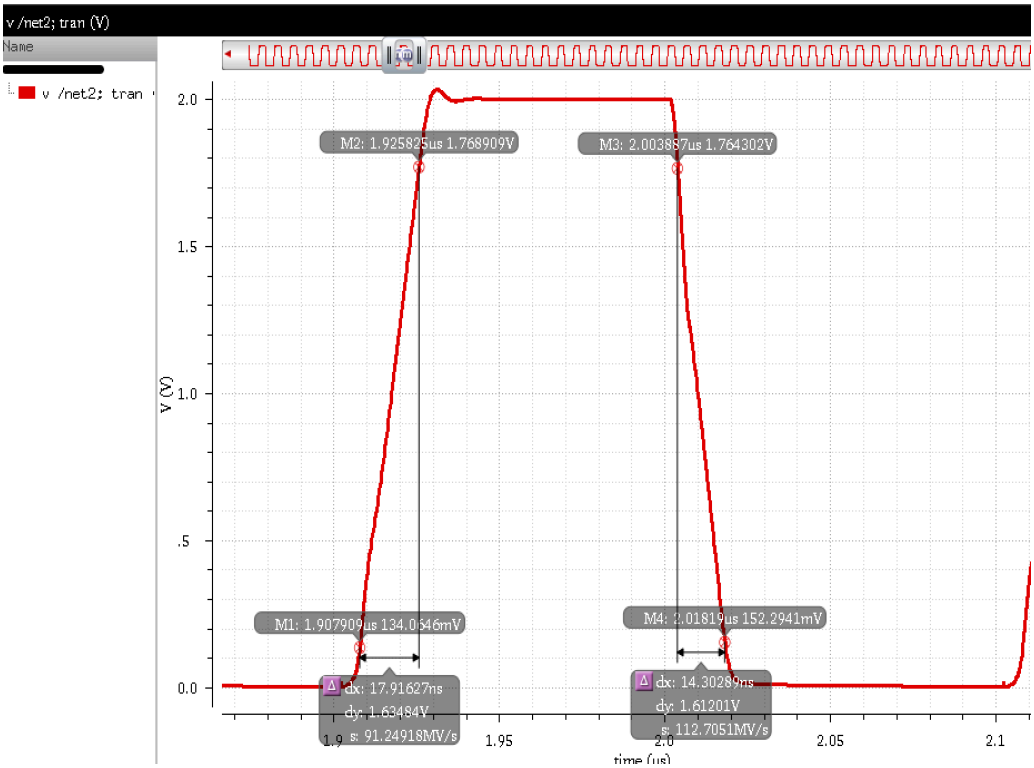

Fig. 5.9: Transient response

In Fig.5.10, the total output voltage power spectral density (V^2/Hz) is shown. From the graph we computed the $Vout_{rms}$ using the calculator in Cadence. The $Vout_{rms}$ results to be 166.7 uV. This value is much smaller than the case we would not consider the integration of the noise (refer to Fig. 5.8, integrator close in buffer). The theoretical case, taking into account the value we have for the input referred voltage at the input of the OTA, is 1.4 mV.

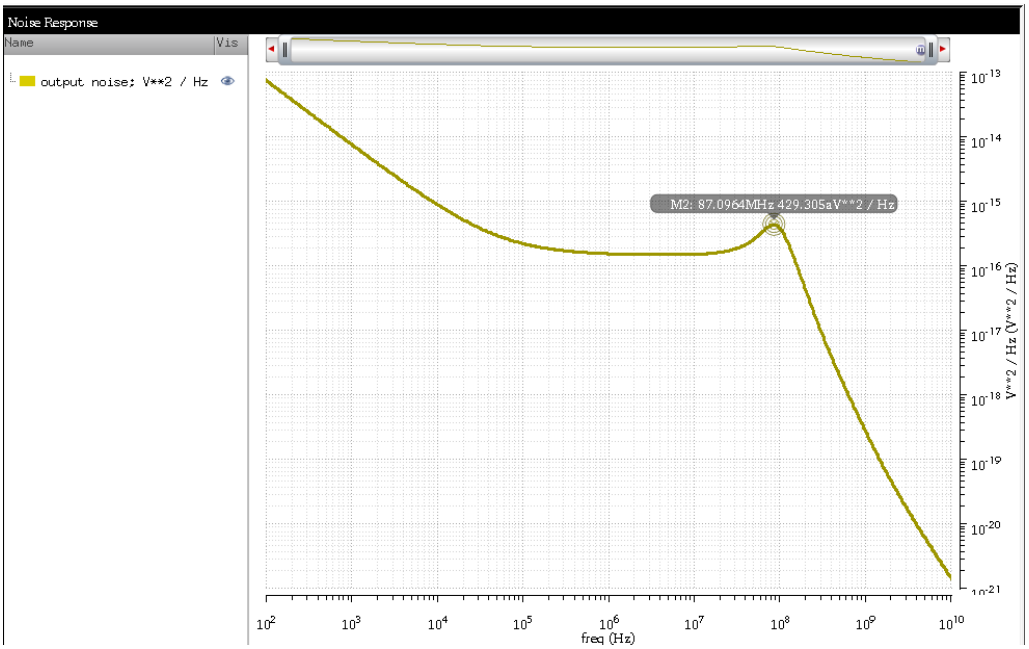

Fig. 5.10: Power spectral density expressed in V^2/Hz at the output of the OTA

5.3 Track and Hold

The Track and Hold (T/H) is made by three parts, as shown in Fig. 5.11: an NMOS working as a switch and controlled by the signal reset3, a capacitance (1 pF) where the data (that is actually the integrated input signal) will be saved on, and an OTA (in particular a PMOS folded cascode has been implemented and an output A-B has been attached to the first stage) closing in a buffer configuration. Using this configuration for the Track and Hold, we limit the effect of the charge injection (less charges coming from the DAC, since the DAC code changes continuously due to the logic). Moreover, we do not need a wide bandwidth since the block has just to isolate the integrator (the signal does not move at any level after the integration) from the rest of the circuit.
Finally, it remains to explain the reason why we came up with a Ch (Hold capacitance) equal to 1 pF. Indeed, since we did not know the charge we would have had (inside the chip) we put a capacitance of a 1 pF. Typically, this is the capacitance we expect in a wire at this node (depends on the chip).
We will present how the OTA works in DC and its stability in terms of phase and GBWP.

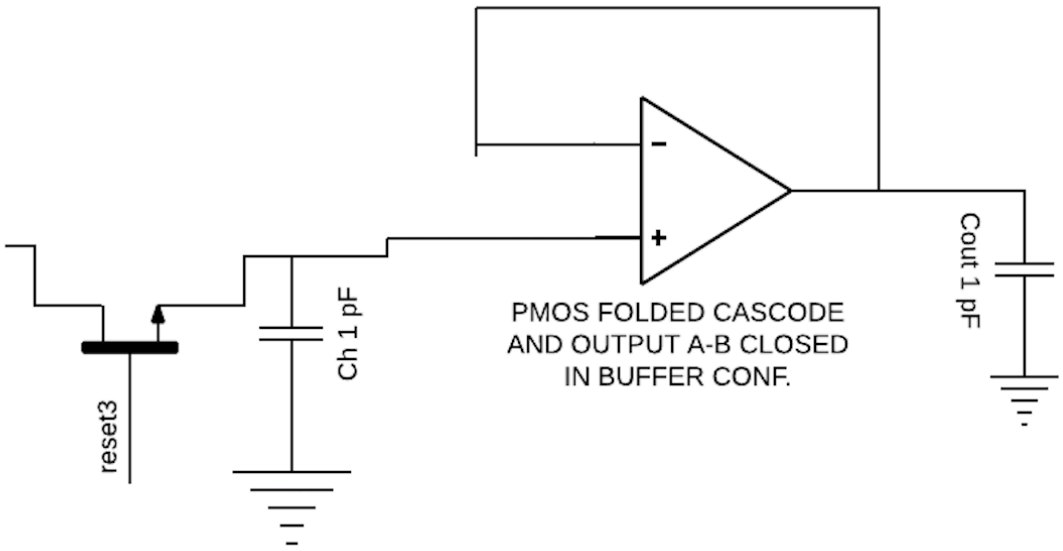

Fig. 5.11: Track and Hold at the top level

5.3.1 OTA's configuration

In Fig. 5.12 the internal configuration of the OTA composing the sample and hold appears. The OTA is composed by a first stage, made by a PMOS folded cascode, and a second one with an A-B output. We considered a GBWP equal to 40 MHz, a Miller capacitance (Cc) of 1 pF and we connected at the output an external load capacitance equal to 1 pF. The reason why we set the GBWP equal to 40 MHz is because we estimated the maximum bandwidth (it could be even smaller) at the output of the integrator to be equal to 40 MHz after the integration (indeed after an integration the BW of a system is reduced). Regarding the choice of the Miller capacitance, if it was bigger than 1 pF, we would limit the slew rate (SR).
Being the expression of the GBWP:

$$GBWP = Gdiff * f1 = \frac{gm1}{2 * \pi * Cc}$$

$$gm1 \geq 2 * \pi * Cc * GBPW \geq 251\ uA/V$$

We decided to set gm1 equal to 260 uA/V and in this way, we computed again the GBWP that is equal to 41.4 MHz. At this point, we studied the stability, calculating the poles and eventual zeros in the system. We will refer to the R1 and R2, being the former the resistance at the output of the first stage, whereas the latter the resistance at the output.

$$R1 = r9 * gm9 * r11//(r7 * gm7 * (r3//r1)) = 34.7\ M\Omega$$

$$R2 = r15//r16 = 8\ K\Omega$$

$$f1 \cong \frac{1}{2 * \pi * Cc * R1 * R2 * (gm15 + gm16)} = 34.9\ Hz$$

Before computing the other poles of the system, let focus on the evaluation of the parasitic capacitances. The capacitance at the gate of M15 (and thus M12) is given by the following expression $Cg15 = \left(\frac{2}{3} * C'_{ox} * \left((\mathrm{W15} * \mathrm{L15}) + (\mathrm{W12} * \mathrm{L12})\right)\right) = 963\ fF$. On the other side, at the drain of M9 and M7 there are two contributions $Cd = \left(C'_{db} * (3 * L_{min}) * (\mathrm{W7} + \mathrm{W9})\right) = 2.58 fF$. Since they are at the same node, they are in parallel and we can sum up the two capacitances $Cgd = Cg15 + Cd = 164.6\ fF$. Finally, the contribution of the drain regions of M15 and M16 have to be taken into account: $Cout = \left(C'_{db} * (3 * L_{min}) * (\mathrm{W15} + \mathrm{W16})\right) = 79.3\ fF$. Being Cload = 1 pF, at the output node there will be a total capacitance equal to $Cext = Cload + Cout + Cgd = 1.19\ pF$.
Once Cc closed (at high frequency), its effect has already gone, the effect due to the capacitances just computed is taken into account:

$$f2 = \frac{(gm15 + gm16)}{2 * \pi * Cext} = 1.3\ GHz$$

At this moment, we show how we get the expression of Rn (nulling resistor), estimating the zero, introduced by the Miller capacitance, far away from the GBWP (indeed we consider the zero at same frequency of the second pole):

$$fz = f2 = \frac{1}{2 * \pi * Cc * \left(Rn - \left(\frac{1}{gm15 + gm16}\right)\right)}$$

$$Rn = \frac{1}{(2 * \pi * Cc * (2 * GBWP)) + (\frac{1}{gm15 + gm16})} = 183\ \Omega$$

Knowing the value of Rn, the frequency of the third pole can be computed:

$$f3 = \frac{1}{2 * \pi * Rn * (Cc // Cext // Cgd)} = 2.58\ GHz$$

Once computing the poles of the circuit and the nulling resistor, it is the time to calculate the zero. This is due to the Miller capacitance and it is equivalent to:

$$fz = \frac{1}{2 * \pi * Cc * \left(Rn - \left(\frac{1}{gm15 + gm16}\right)\right)} = 2.58\ GHz$$

When the stability analysis of the circuit is ended, it remains to size the other transistors of the OTA. Finally, in Tab. 5.3 there are the principal parameters of the transistors composing the OTA of the T/H.

	I(uA)	Vov(V)	gm(uA/V)	W/L	W(um)	L(um)	rd (Mohm)
M0	52	0.2	520	52	86.2	1	0.385
M1, M2	26	0.2	260	26	43.1	1	0.77
M3, M4	52	0.2	520	26	20.6	1	0.385
M5	52	0.2	520	52	86.2	1	0.385
M6, M7	26	0.2	260	13	10.3	1	0.77
M8, M9	26	0.2	260	11.6	43.1	1	0.77
M10, M11	26	0.3	260	26	19.2	1	0.77
M12, M13	8.67	0.2	86.7	4.3	2.06	1	2.31
M14	8.67	0.2	86.7	4.3	2.06	1	2.31
M15	1000	0.3	6670	444.4	134	0.8	0.016
M16	1000	0.2	10	500	103	0.8	0.016

Tab. 5.3: T/H's PMOS folded cascode with A-B output sizing

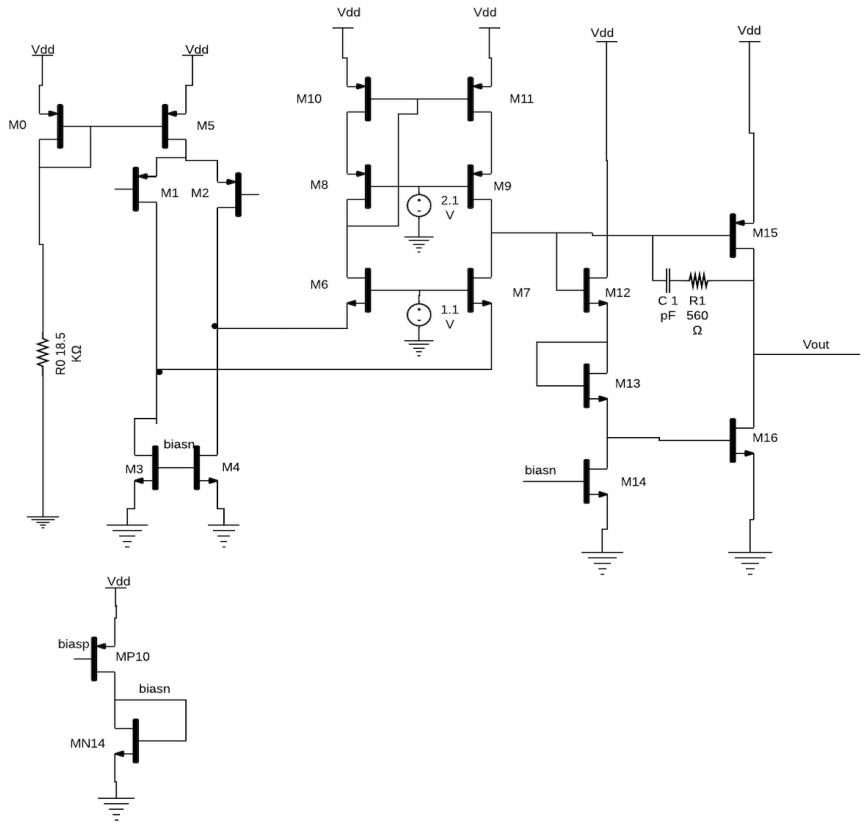

Fig. 5.12: OTA' structure

5.3.2 Noise Evaluation

The goal of this section is to calculate the input referred voltage noise source of the OTA stage and then $vout_{rms}$. As we did before with the OTA of the integrator, in order to compute the input-referred voltage noise source, the output noise in the circuit model, where a voltage generator equal to Sv_{eq}^2 is attached at the input differential pair and the components of the circuit are considered noiseless, is compared to the output noise in the real circuit (Fig. 5.12) where a short is made between the input terminals.
In the case of Sv_{eq}^2 at the input differential pair (no other noise contributions), it can be seen that the short circuit current noise power is linked to Sv_{eq}^2 by:

$$Svout_{tot}^2 = Sv_{eq}^2 \cdot (gm_1)^2 \cdot R_1^2 \cdot (gm_{15} + gm_{16})^2$$

And hence:

$$Sv_{eq}^2 = \frac{Svout_{tot}^2}{(gm_1)^2 \cdot R_1^2 \cdot (gm_{15} + gm_{16})^2}$$

Let us now derive the contribution to $Svout_{tot}^2$ (for each component) in the real circuit (Fig. 5.12) and comparing it to the output current noise in the circuit driven by the input referred noise source, it turns out:

$$Sv_{eq,M1,M2}^2 = \frac{4KT\lambda}{gm_1} = 2.05 \cdot 10^{-17} \frac{V^2}{Hz}$$

$$Sv_{eq,M3,M4}^2 = \frac{4KT\lambda \cdot (gm_3 + gm_4)}{gm_1^2} = 1.23 \cdot 10^{-16} \frac{V^2}{Hz}$$

$$Sv_{eq,M10}^2 = \frac{4KT\lambda \cdot gm_{10} \cdot \left(\frac{gm_{11}}{gm_{10}}\right)^2}{gm_1^2} = 2.05 \cdot 10^{-17} \frac{V^2}{Hz}$$

$$Sv_{eq,M11}^2 = \frac{4KT\lambda \cdot gm_{11}}{gm_1^2} = 2.05 \cdot 10^{-17} \frac{V^2}{Hz}$$

$$Sv_{eq}^2 = 1.85 \cdot 10^{-16} \frac{V^2}{Hz} = (13.6 \frac{nV}{\sqrt{Hz}})^2$$

At this point, knowing the noise equivalent BW of the OTA ($BW \approx GBWP \cdot \frac{\pi}{2} = 62.8\ MHz$) and being the OTA closed in buffer configuration (the transfer function of the input-referred noise is 1), it is possible to compute the value of $vout_{rms}$:

$$vout_{rms} = \sqrt{Sv_{eq}^2 \cdot BW \cdot 1} = 108\ uV_{rms}$$

The analysis of the system has been carried out, even if this is not the relevant noise in the circuit. Indeed, the $\frac{K \cdot T}{C}$ is the noise we have to take into account. It results to be:

$$\frac{K \cdot T}{C} = (63\ uV)^2$$

This contribution is not significant compared with the signal at the input of the T/H that is around 300 mV. The signal is not worsened by the noise of the T/H.

5.3.3 Track and hold simulation

In Fig. 5.13 the performances in terms of loop gain (blue line) and phase margin (green line) when at the input is applied a vdc generator (DC generator that emulates the voltage in Ch once the integration is over) are shown. In particular, the gain of the OTA is 68.74 dB, the GBWP is equal to 113.5 MHz and the phase margin of 85.33°.

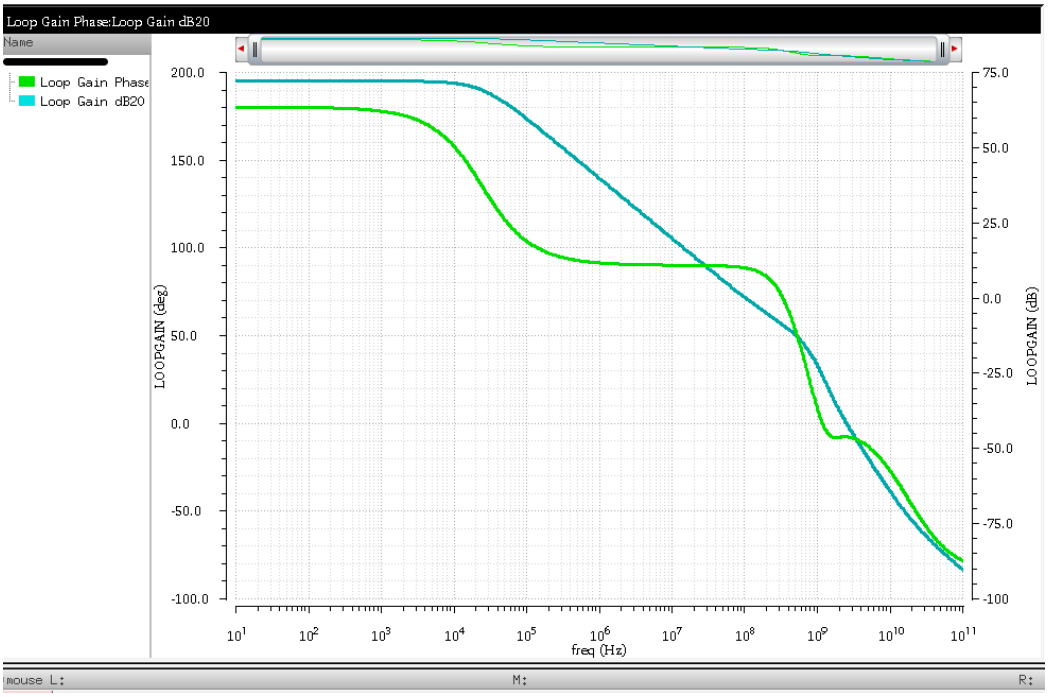

Fig. 5.13: Phase (red) and gain (yellow) of the OTA's loop

Moreover, in order to check the feasibility and the comportment of the T/H we reproduced the transient simulation, having as result the plot in Fig 5.14. To do that, we applied at the input (Vin) of the T/H a vsin generator (sinusoidal generator) to emulate the integration process. The sinusoidal singal is characterized by an amplitude of 300 mV and a frequency of 100 KHz (period (T_s)

= 10 us). When the signal (red line) controlling the switch triggers from low to high level, the output signal (green line) will sampled. Hence, this value will be saved in Ch (the voltage will follow Vin) and, through the OTA in buffer mode, will be transferred to the input of the ADC's comparator.

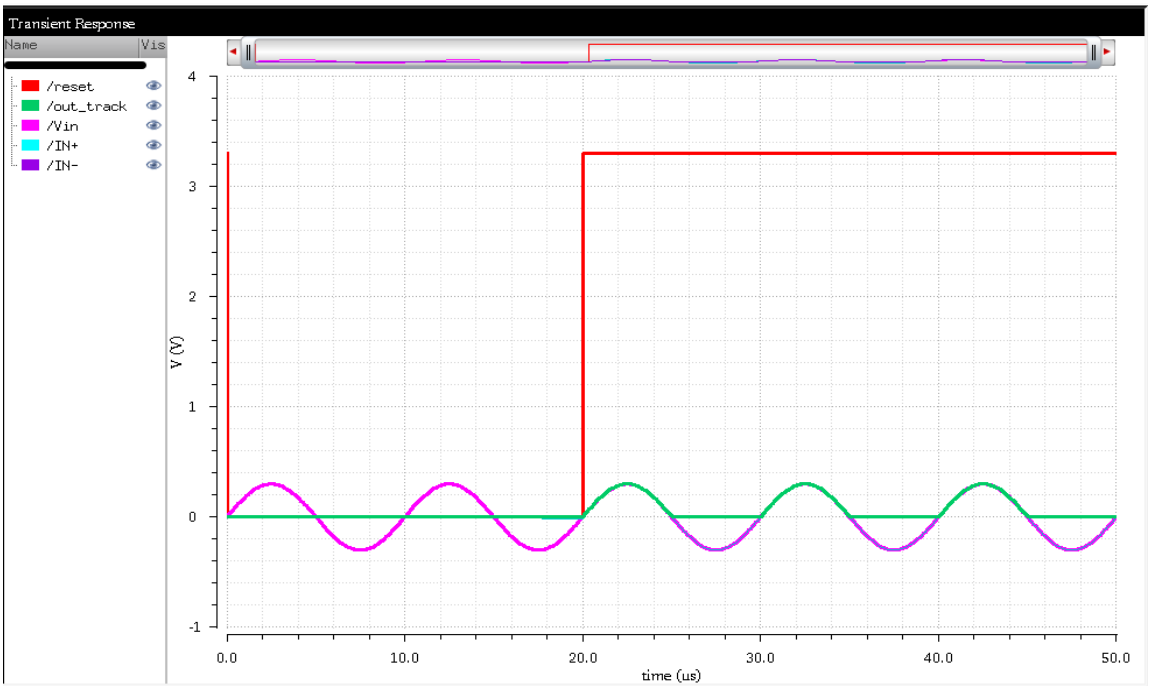

Fig. 5.14: Transient simulation of the T/H

Finally, the total output voltage power spectral density (V^2/Hz) noise is displayed in Fig. 5.15. As we can appreciate from the graph, there is a decreasing trend and, after a stabilization of the power spectral density, the trend keeps decreasing. The integral voltage noise referred to the T/H is 75.95 uV. Comparing the result with the one computed theoretically, it can be seen they are similar.

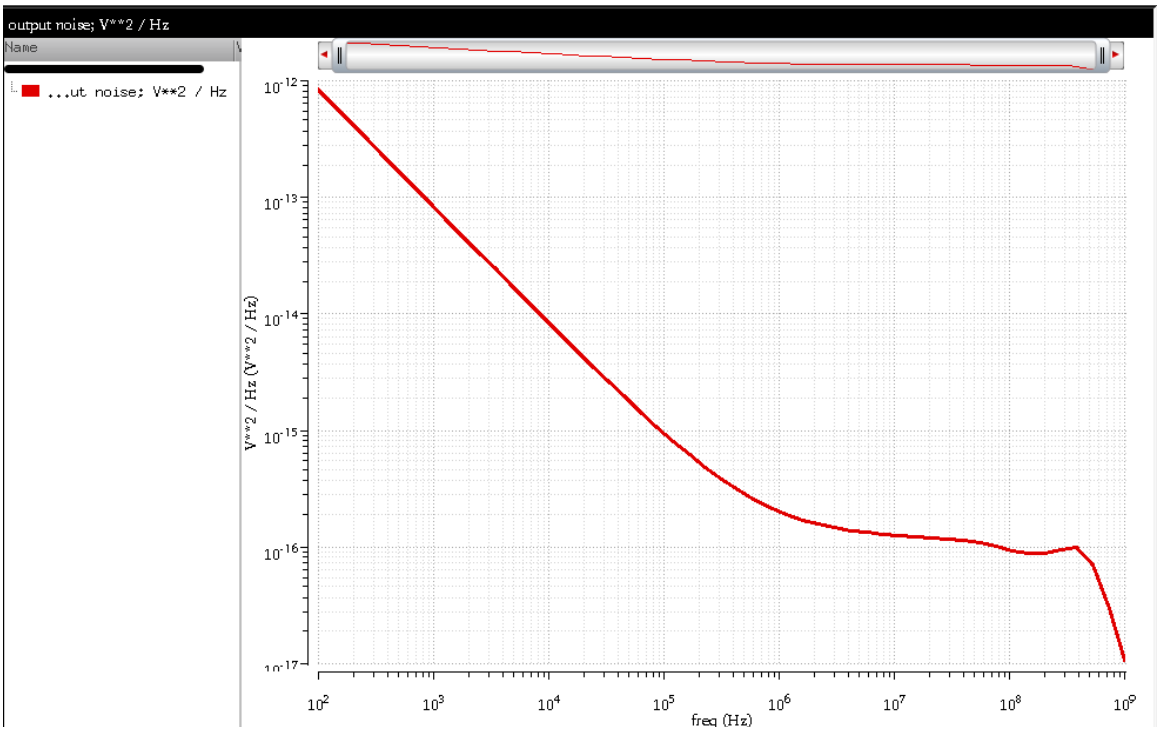

Fig. 5.15: Power spectral density expressed in V2/Hz at the output of the OTA

5.4 ADC

In this sub-section, we analyze the components of the ADC in Fig. 5.16. In our case, the ADC will be identified by 9 bits. Being the voltage range of the ADC from 3.3 V to 0 V, the ADC's resolution (LSB) will be: $LSB = \frac{3.3\ V}{2^9} = 3.9\ mV$.
We focus on the DAC, explaining the tool we exploited to originate this block. Then, we go through the internal structure of the comparator, giving its sizing and its behavior in transient. In the end, the Logic circuit of the ADC is presented, considering the first 3 iterations of the whole code (in total there are 9 iterations having the ADC a resolution of 9 bits). The other remaining iterations (6) are implemented in the same way as the first three and hence, for this reason, they are not displayed.
The ADC (SAR) has been developed just at high level in order to see its real functionality.

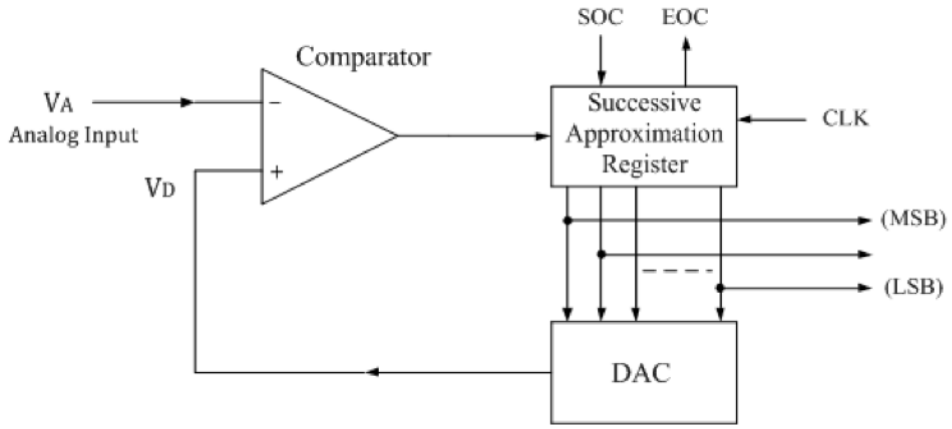

Fig. 5.16: Internal ADC SAR's structure

5.4.1 DAC

Regarding the DAC, the first approach we wanted to apply was the R-2R structure, that was used for nulling the offset as well (our target is an ADC of 9 bits, so in the case we had wanted an higher resolution (≥11 bits), we would have considered a capacitive DAC (and we would have used a Wilkinson ADC as the state-of-art's case). Indeed, taking into account 10 bits we would have had problems for the matching of the transistors

Then, we decided to take advantage from a utility of Cadence. In fact, the tool provides "*model writers*" views which are models already existing for different components. Considering the specifications, it is necessary to set the principal characteristics of the model, as shown in Fig. 5.17. In particular, we fixed the number of bits of the DAC (equal to the ones of the ADC) equal to 9, the voltage range is from 3.3 to 0 V, the output rise and fall times are 1 ns (it has to be fast), the conversion time for a single bit is 100 ns (having a maximum ADC conversion time and thus 9 bits to analyze) and, lastly, a digital input threshold at the middle of the dynamic DAC.

The generated DAC will have two inputs: the nine bits (represented by a latch which include them) and a clock. The clock of the DAC will be the same used for the Logic (see section 5.4.3) in order to be synchronous. Indeed, in the case the two clocks were different, there would not be a correspondence between the datum coming out the logic and the voltage at the output of the DAC.

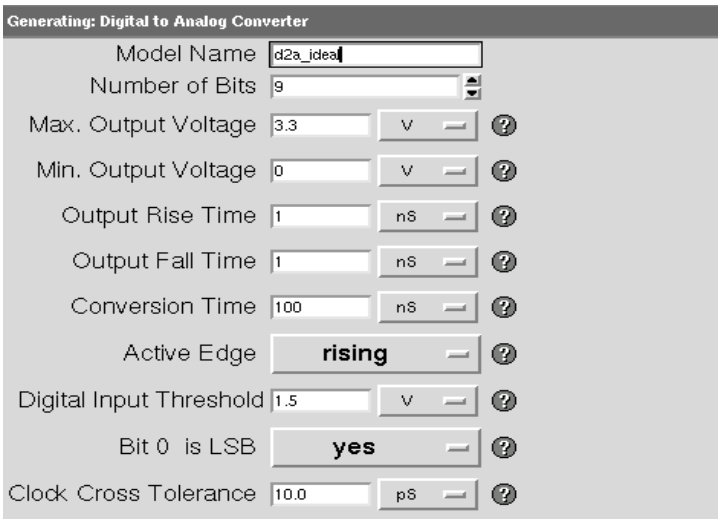

Fig. 5.17: DAC model writer

5.4.2 Comparator

The comparator is a key element for the working circuit of the ADC. Indeed, it is necessary in order to correlate the integrated signal (saved in the Ch of the T/H) at the input of the slow chain and the one coming from the logic (bit sequence) of the SAR, then transformed by the DAC in voltage.
We realized the comparator in three different stages, as shown in Fig.5.18: in the first one we used a comparator based on it; in the second one we took the differential output of the first stage and we turned into a single ended one through a current mirror; in the third and last one we used an output buffer in order to square, being inverter, and buffer digitally.
The first stage consists of a source-coupled pair [36], a tail current, and cross-coupled load. The load devices are connected such that in differential mode, the outer transistors act as positive resistors, while the cross-coupled devices act as negative resistors. The negative resistance cancels the positive, hence presenting a high differential out impedance. An advantage of the cross-coupling is that the NMOS load provides local common-modes feedback with no extra devices; the common-mode impedance is low (1/2gm), thus, the common-mode voltage is stabilized at one Vgs above ground.
The target of this structure is to be quite fast (necessary for the ADC) in response to a signal at the input.

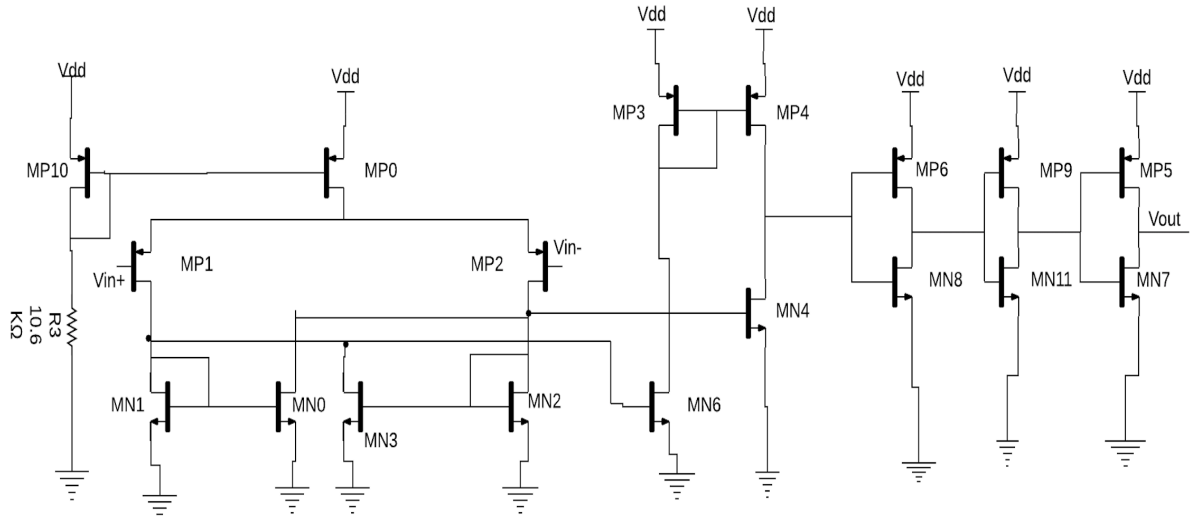

Fig. 5.18: Different stages of the ADC's comparator

In the following table, we showed the different sizing for the transistors of the comparator when in DC we apply a signal (Vin+) of 500 mV at MP1 and a voltage signal of 500mV (Vin-) at MP2:

	Vov(V)	W/L	W(um)	L(um)
MP0	0.7	27	13.5	0.5
MP1	0.65	17.14	6	0.35
MP2	0.65	17.14	6	0.35
MN1	0.35	7.14	2.5	0.35
MN2	0.35	7.14	2.5	0.35
MN0	0.35	5.7	2	0.35
MN3	0.35	5.7	2	0.35
MN4	0.35	24.3	8.5	0.35
MN6	0.35	17.14	6	0.35
MP3	1.2	5.7	2	0.35
MP4	1.2	5.7	2	0.35
MP6	2.3	17.14	6	0.35
MP9	OFF	34.3	12	0.35
MP5	2.5	91.4	32	0.35
MN8	OFF	8.6	3	0.35
MN11	2.6	17.14	6	0.35
MN7	OFF	45.7	16	0.35
MP10	0.7	28.57	10	0.35

Tab. 5.4: Sizing Comparator

In Fig. 5.19 the transient simulation of the comparator is displayed when at MP1 is applied a step signal (Vin+) from 502 mV to 0 V (the difference between Vin+ and Vin- corresponds to LSB/2), that has a period equal to 10 us and the rise fall time are 100 ps). We could not apply the proper step signal at the input (from 2 mV to 0 V equal to LSB/2) because the software did not recognize and hence displayed the DC value of Vin+ due to the offset we have (even if we set, at DC, Vin+ to 500 mV, it started from 0 V in the transient simulation). Nevertheless, we showed the comparator commutes.
It has been displayed the transient behavior of the comparator when it switches in the worst possible condition.
In the transient performance out_single (light blue) is the single ended output, while Vout (purple) is the output of the comparator after the buffer series. As we can see from the figure, the comparator acts after a time equal to s (f = 1/T ≈ GHz) more or less. Therefore, we demonstrated our target for the comparator, being fast.

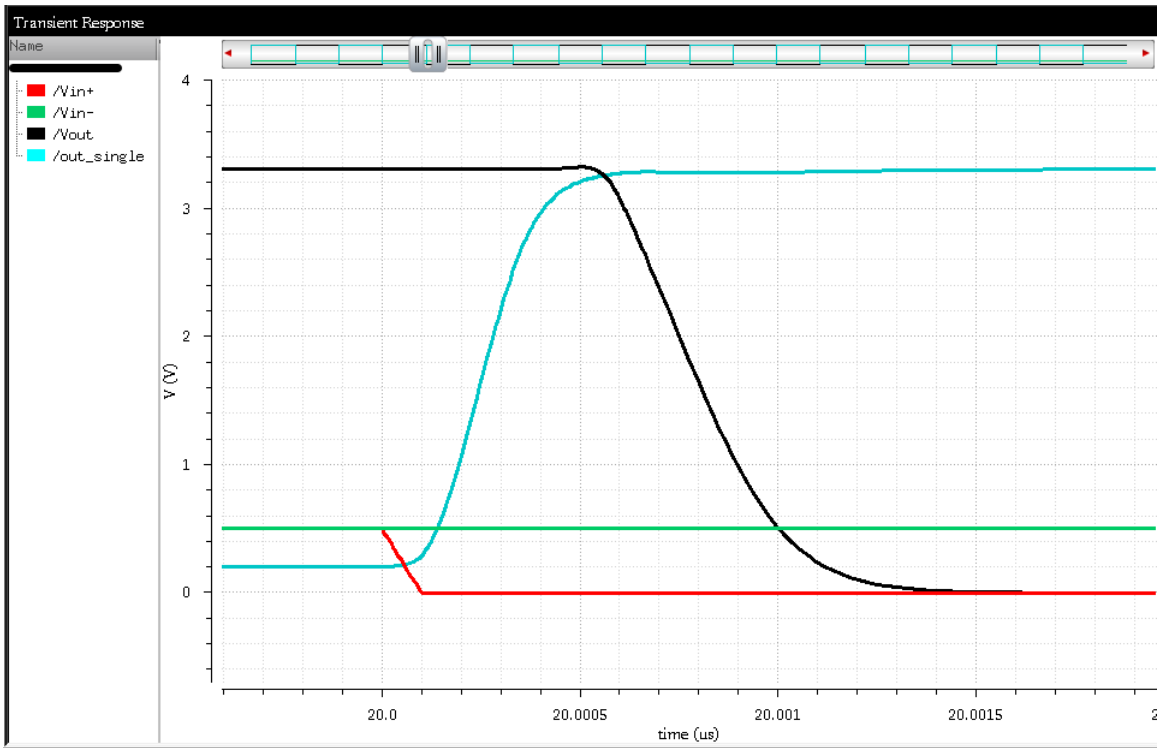

Fig. 5.19: Comparator behavior in transient mode

5.4.3 Logic

The Logic of the ADC is a model and it has been programmed in Verilog-A (not at transistor level) since the goal of this block was to analyze how the rest of the circuit behaves. In particular, we want to evaluate the performances of the integrator and the Track and Hold.

This logic assumes a critical role in the whole circuit, due to the fact it is fundamental for the best approximation of the integrated signal. Indeed, once the integrated signal in the sample capacitance (coming from the SiPM) has been compared with the signal coming out from the DAC, the result is sent to the logic of the SAR which will decide if the n-th bit keeps being 1 for the remaining cycles or set to 0. This logic has four inputs: a clock (clk_logic) that is actually the same of the DAC, a reset (RST) signal (it is asynchronous respect to the clock), the output of the comparator (comp_out) as said before, and the command signal of the switch connected between the integrator and the T/H(reset_3) which is necessary since in this way the logic knows when a voltage is caught by Ch. There is an output connected directly to the DAC and at the same time it is the one which provide the final result.
The logic implementation is presented for the first 5 bits in Fig. 5.20 and Fig. 5.21. Referring to these figures, we will go through the logic implementation, studying the first three bits of the sequence.
Firstly, RESET is analyzed, and, in the case, it is equal to 1, the bits of state are set to 0 and the same works for count, otherwise we follow the else path assigning next_state to state and next_count to count. In a second moment, we start studying the different cases. We begin examining the configuration equal to VFSR/2 (half of the total dynamic) that is 100000000. As can be seen in the first iteration, 100000000 is assigned to next_state. Then, next_state, passing through the DAC is converted into its analog value and compared to the value in Ch. If COMP_c is equal to 1 the n-8 th bit will remain 1 till the conversion is over and next_count will be set to 110000000, if not, next_count is set to 010000000. In the following iteration (7'd3) the operation made before is repeated. Hence, COMP_c is taken into account one more time. If it is equal to 1, next_count will contain 111000000 or 011000000 depending on the result of the previous iteration. On the other hand, if COMP_c is equal to 0, next_state will be 110000000 or 010000000, meaning the n-7 th bit will remain 0 till the end of the conversation. This process will last until when the n-1 th bit is studied. Finally, as a result at the output of the logic we will get the set of bits approximating in the best way the integrated signal.

```
always @(posedge CLK or negedge RST) //RST asincrono con el CLK
begin
        if (RST == 1'b0)
                begin
                state <= 7'b0000000;
                count <= 9'b00000_0000;
                end
        else
                if (START == 1'b1)
                        begin
                        state <= next_state;
                        count <= next_count;
                        end
end


always @(COMP_c, START, count, state)
begin

        case (state)

        7'd0: begin
                next_state = 7'd1;
                next_count = 9'b100000000;
                end

        7'd1: begin
                next_state = 7'd2;
                if (COMP_c == 1'b1)
                        begin
                        next_count = count | 9'b010000000;
                        end
                else
                        begin
                        next_count = (count & 9'b011111111) | 9'b010000000;
                        end
                end
```

Fig. 5.20: ADC logic, first two cycles (next_count = 100000000 in the first cycle is FSR/2)

```
        7'd2: begin
                next_state = 7'd3;
                if (COMP_c == 1'b1)
                        begin
                        next_count = count | 9'b001000000;
                        end
                else
                        begin
                        next_count = (count & 9'b101111111) | 9'b001000000;
                        end
                end

        7'd3: begin
                next_state = 7'd4;
                if (COMP_c == 1'b1)
                        begin
                        next_count = count | 9'b000100000;
                        end
                else
                        begin
                        next_count = count & 9'b110111111| 9'b000100000;
                        end
                end

        7'd4: begin
                next_state = 7'd5;
                if (COMP_c == 1'b1)
                        begin
                        next_count = count | 9'b000010000;
                        end
                else
                        begin
                        next_count = count & 9'b111011111| 9'b000010000;
                        end
                end
```

Fig. 5.21: ADC logic, cycles 3,4 and 5

Chapter 6

Results and Conclusions

This thesis work was aiming at developing a test design of an integrated front end for large arrays of SiPM. In a particular, we focused on a single channel of the readout system.
First of all, a single front-end channel has been implemented in 0.35 um technology to check the feasibility of the design. The SiPM read out has been modeled through a preamplifier which was supposed to have 200 MHz of bandwidth (BW) and 5 mW of power consumption from specifications.
The preamplifier is connected to a slow chain (ADC) and to a fast chain (Time to Digital Converter or TDC).
The need of a wide bandwidth is due to the fast chain because, having an acquisition time as low as possible (1/BW), the risk of losing a photon is reduced and the system is also more precise. Nevertheless, we implemented a topology which has a GWBP (= BW) equal to 62 MHz and in AC 104 MHz of bandwidth when we applied at the input a current of 1 A. This system could not satisfy the requirement of the bandwidth. Either way, this was not the goal of the project (we did not have to look for the best amplifier option in terms of bandwidth). Indeed, a preamplifier was only necessary to implement the slow chain.
Moreover, we got another problem over the implementation of the preamplifier. In fact, being the Dynamic Range (DR) 400 pe (we took a margin considering 440 pe) and the charge of the Single Photo Electron (SPE) 0.4 pC, the total charge at the output of the SiPM resulted to be 176 pC, and being the maximum voltage excursion at the output of the preamplifier 2 V, the capacitance, we would have need, had been 88 pF (huge for the channel Metal1-Metal1 taken into account). It would have meant an enormous cost in terms of area occupied. For this reason, we came up with a capacitance of 8 pF. In doing so, we introduced a de amplification meaning worse performances in terms of SNR.
Secondly, a charge integrator has been studied. An OTA (PMOS folded cascode) and a Monticelli structure have been used for this purpose. From the specifications, the integrator supposed to have an integration time of 10 us at most, a power consumption of 12 mW roughly and a fast SR. We could satisfy the requirements of the integration time since it resulted to be 8.8 us and power consumption which was 10.7 mW. Regarding the SR, the rise time and the fall time when a transient signal was applied at its input resulted to be slower (17 ns and 15 ns) than what we wanted (5 ns/6 ns). We could have decreased the Miller capacitance and increased the tail current of the Monticelli, worsening power consumption and stability as well. There is a tradeoff between SR and stability-power consumption.
At this point, once the analog part of the system has been taken into account, we focused on the T/H and the ADC.

The T/H is composed by an OTA (PMOS folded cascode) and an A-B output. Being the time to set the value on the hold capacitance equal to 200 ns at most, we picked a GBWP 40 MHz.
Regarding the ADC SAR (from requirements the conversion time has to be 1 us), a fast comparator has been implemented, while the DAC has been taken from the library of Cadence, exploiting the option model writer. Finally, the logic has been programmed in Verilog-A.
The final result in terms of linearity of the chain for different pe detected can be seen in Fig. 6.1:

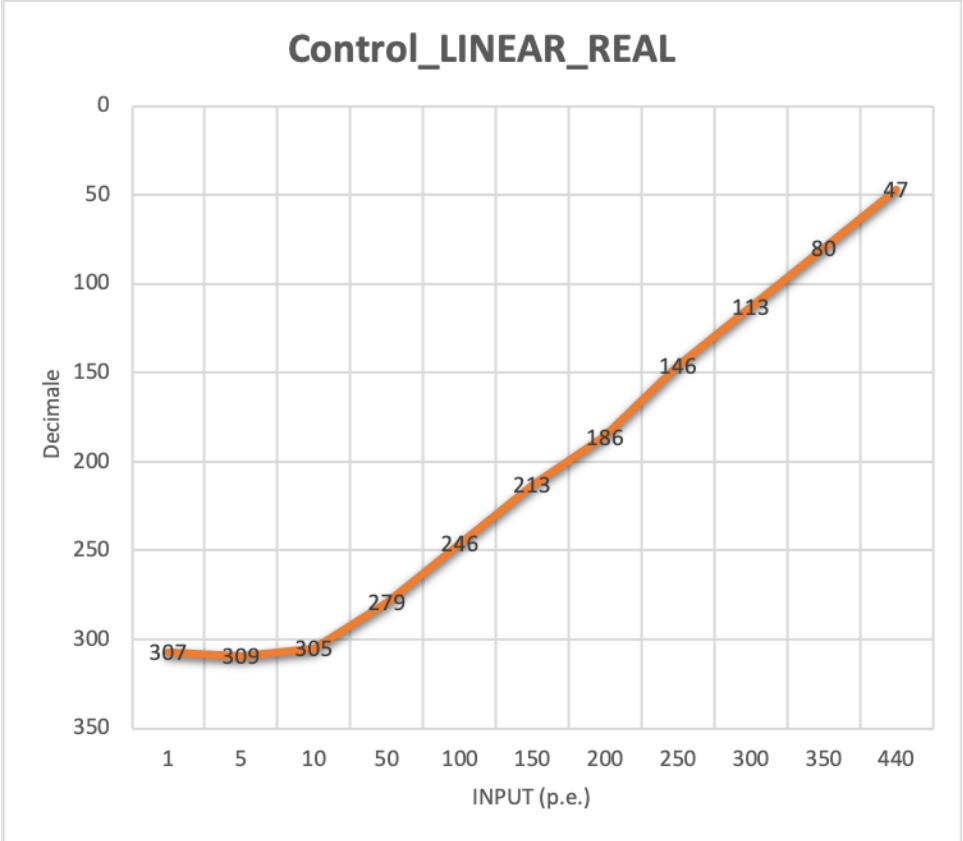

Fig. 6.1: Linearity check

As we can see from the figure, apart the initial part of the curve, the slope is linear from 10 pe.

Bibliography

[1] P. P. Calò, S. Petrignani, C. Marzocca and F. Ciciriello, "SiPM readout electronics," *Nuclear Inst. and Methods in Physics Research, A.*

[2] S. R. Meikle and R. D. Badawi, "Quantitative Techniques in PET".

[3] B. Guerin and G. E. Fakhri, "Novel Scatter Compensation of List-mode PET Data Using Spatial and Energy Dependent Corrections," *Transactions on medical imaging,* vol. 30, no. 3, 2011.

[4] M. Conti, I. Hong and C. Michel, "'Reconstruction of Scattered and Unscattered PET Coincidences Using TOF and Energy Information," *Physics in medicine and biology,* vol. 57, pp. 307-317, 2012.

[5] B. J. B. Neena Nambiar, M. N. Ericson and M. H. Kao, "A Current Mode Multi-Channel Wilkinson ADC," in *INTERNATIONAL CONFERENCE ON SIGNALS AND ELECTRONIC SYSTEMS*, Krakow, 2008.

[6] L. H. C. Braga, S. Member, L. Gasparini, L. Grant, R. K. Henderson, N. Massari, M. Perenzoni, D. Stoppa and R. Walker, "A Fully Digital 8 16 SiPM Array for PET Applications With Per-Pixel TDCs and Real-Time Energy Output," *JOURNAL OF SOLID-STATE CIRCUITS,* vol. 49, no. 1, 2014.

[7] D. W. Rickey, R. Gordon and W. Huda, "On Lifting the Inherent Limitations of Positron Emission omography by Using Magnetic Fields (MagPET)," *Automedica,* vol. 14, pp. 355-369, 1992.

[8] G. F. Knoll, "Radiation detection and measurement," *John Wiley & Sons,* 2010.

[9] W. W. Moses, "Time of Flight in PET Revisited," *Trans. Nucl. Sci.,* vol. 50, no. 3, p. 1325–1330, 2003.

[10] M. Conti, "Effect of Randoms on Singal-to-Noise Ratio in TOF PET,," *Trans. Nucl. Sci.,* vol. 53, no. 3, p. 1188–1193, 2006.

[11] M. Conti, "Focus on time-of-flight PET: the benefits of improved time resolution," *Eur. J. Nucl. Med. Mol. I,* vol. 38, no. 6, p. 1147–1157, 2011.

[12] P. M. Düppenbecker, B. Weissler, P. Gebhardt, D. Schug, J. Wehner, P. K. Marsden and V. Schulz, "Development of an mri compatible digital sipm based pet detector stack for simultaneous preclinical pet/mri," *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC),* no. pages 3481–3483, 2012.

[13] C. Fiorini, A. D. Butt, M. Occhipinti, R. Quaglia, P. Trigilio, K. Nagy, T. Bükki, M. Czeller, Z. Nyitrai and C. Piemonte, "Development of a sipm-based anger camera for insert, a new multi-modality spect/mri system for preclinical and clinical imaging," *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC),,* pp. 1-4, 2014.

[14] C. Piemonte, A. Ferri, A. Gola, N. Serra, A. Tarolli and N. Zorzi, "Performance of a novel, small-cell, high-fill-factor sipm for tof-pet," *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC),* pp. 1-3, 2013.

[15] G. Llosá, R. Battiston, N. Belcari, M. Boscardin, G. Collazuol, F. Corsi, G.-F. D. Betta, A. D. Guerra, N. Dinu, G. Levi, S. Marcatili and S. Moehrs, "Novel Silicon

Photomultipliers for PET Applications," *TRANSACTIONS ON NUCLEAR SCIENCE,* vol. 55, no. 3, 2008.

[16] A. Gallivanoni, I. Rech and M. Ghioni, "Progress in Quenching Circuits for Single Photon Avalanche Diodes," *Nuclear Science,* vol. 57, 2010.

[17] C. Piemonte and A. Gola, "Overview on the main parameters and technology of modern Silicon Photomultipliers," *Nuclear Inst. and Methods in Physics Research, A,* 2018.

[18] H. T. v. Dam, S. Seifert, R. Vinke, P. Dendooven, H. Löhner, F. J. Beekman and D. R. Schaart, "A Comprehensive Model of the Response of Silicon Photomultipliers," *Transactions on Nuclear Science,* vol. 57, 2010.

[19] F. Arneodo, M. Benabderrahmane, G. Bruno, V. Conicella, A. D. Giovanni, O. Fawwaz, M. Messina, A. Candela and G. Franchi, "Cryogenic readout for multiple VUV4 Multi-Pixel Photon Counters in liquid xenon," 2017.

[20] HAMAMATSU, "VUV-MPPC(VUV4)," *KSX-0046 C,* 2016.

[21] First Sensor, "Introduction to silicon photomultipliers (SiPMs)," 2015.

[22] P. Buzhan, B. Dolgoshein, A. Ilyin, Kantserov, V. V. Kaplin, A. Karakash, A. Pleshko, L. Filatov, S. Klemin and F. Kayumov, "AN ADVANCED STUDY OF SILICON PHOTOMULTIPLIER," *ICFA Instrumentation Bulletin.*

[23] PETsys Electronics, "TOFPET2: a high-performance ASIC for time and amplitude measurements of SiPM signals in time-of-flight applications".

[24] A. D. Francesco, R. Bugalho, L. Oliveira, A. Rivetti, M. Roloa, J. C. Silvaa and J. Varelaa, "TOFPET 2: A high-performance circuit for PET time-of-flight," *Nuclear Instruments and Methods in Physics Research,* 2016.

[25] J. Castilla, J.-M. Cela, A. Comerma, J.-M. Fernández-Varea, L. Freixas, L. Garrido, D. Gascón, J. Marín, G. Martínez, J.-M. Pérez, P. Rato-Mendes and I. Sarasola, "Evaluation of the FlexToT ASIC on the Readout of SiPM Matrices and Scintillators for PET," 2013.

[26] A. Comerma, D. Gascon, L. Freixas, L. Garrido, R. Graciani, J. Marın, G. Martınez, J. M. Perez, P. R. Mendes, J. Castilla, J. M. Cela, J. M. Fernandez-Varea and I. Sarasola, "FlexToT - Current Mode ASIC for Readout of Common Cathode SiPM arrays," 2013.

[27] I. S. Martín, "Detectores monolíticos y sensores compatibles con altos campos magnéticos para tomografía por emisión de positrones," pp. 75-76, 2015.

[28] J. Fleury, S. Callier, C. d. L. Taille, N. Seguin, D. Thienpont, F. Dulucq, S.Ahmad and G. Martin, "Petiroc, a New Front-End ASIC for Time of Flight Application," 2013.

[29] Weeroc, "Read-Out Chips Catalog," *Electronic Instrumentation.*

[30] V. Herrero-Bosch, C. W. Lerche, M. Spaggiari, R. Aliaga-Varea, N. Ferrando-Jodar and Ricardo Colom-Palero, "AMIC: An Expandable Front-End for Gamma-ray Detectors with Light Distribution Analysis Capabilities," 2011.

[31] K. Bult and G. J. G. M. Geelen, "An Inherently Linear and Compact MOST-Only Current Division Technique," *JOURNAL OF SOLID-STATE CIRCUITS,,* vol. 27, 1992.

[32] B. Razavi, Design of Analog CMOS Integrated Circuits, Singapore: McGraw-Hill, 2001.

[33] R. Gregorian, Introduction to CMOS OP-AMPS and COMPARATORS, USA: John Wiley & Sons., 1999.

[34] A. L. Lacaita, "Single Ended Differential Stage," no. 10-12, 2018.

[35] Cova and Rech, "Slide_Filters_LPF2," pp. 13-22, 2017.

[36] K. Bult and A. Buchwald, "An Embedded 240-mW 10-b 50-MS/s CMOS ADC in 1-mm(^2)," *JOURNAL OF SOLID-STATE CIRCUITS,* vol. 32, no. 12, 1997.