Design and test of readout electronics for medical and astrophysics applications

ScuDo
Scuola di Dottorato ~ Doctoral School
WHAT YOU ARE, TAKES YOU FAR

Doctoral Dissertation
Doctoral Program in Electrical, Electronics and Communications Engineering
with curriculum in "Electronic Devices" in convention with the National Institute for Nuclear Physics (INFN) (30^{th}cycle)

Design and test of readout electronics for medical and astrophysics applications

By

Federico Fausti

******

Supervisor(s):
Giovanni Mazza,
Italian National Institute for Nuclear Physics, Turin section

Doctoral Examination Committee:
Doctor U. Weber, Referee, GSI Helmholtz Centre for Heavy Ion Research, Darmstad
Doctor S. Beolè, Referee, Turin University of Physics

Prof. M. Greco, Turin University of Physics
Prof. M. Martina, Polytechnic University of Turin
Prof. F. Bonani, Polytechnic University of Turin

Politecnico di Torino
2018

Declaration

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

Federico Fausti
2018

* This dissertation is presented in partial fulfillment of the requirements for **Ph.D. degree** in the Graduate School of Politecnico di Torino (ScuDo).

I would like to dedicate this thesis to some reference points, in my life:
to Giulia, my antiparticle, together we are brighting light
to my parents who always believed in me
and a special hug to anyone who gave me inspiration, a spark.

Acknowledgements

I would like to acknowledge the Turin TERA group, especially my two Roberto whose suggestion convinced me in this great Ph.D. way. Thanks to the group members: Flavio, Vincenzo, Simona, Anna , Omar, Leslie, Mohammad, Felix, Lorenzo and Andrea. Feeling like home is not mandatory for a workplace but it is necessary, for a successful collaboration. I'm glad and thankful to Giuseppe Pittà, Marco Lavagno, and the DE.TEC.TOR. Devices and technologies Torino, for boosting my career with the TERA09 project; working for a young team, for its R&D core, has been a terrific experience for the fresh-Master-graduated me, who really needed a concrete goal while looking for a high-tech profile target. Thanks to all the Ufficio_VLSI members, this place has been a solid but even sparkling spot, where I grew up shaping my technical knowledge, with the right motivation. Thanks to Jonhatan, Serena, Alessandra, Agostino and Fabio, the "senior colleagues"; thanks to Simona, Ramshan, Alejandro, Weishuai, Junying, Raffaele and Lorenzo, the new design force. Thanks to Gianni, my Ph.D. tutor, for the patient he had in the last four years and for being "the engineering point of view". Thanks for the kindness and the example given by Alexander and Pavel, my Russian Mini-EUSO colleagues, to Francesca for her strength and Mario for his generous professional sharing and for making this international collaboration possible. Another special thank goes to my martial art group and expecially to Master G. and Roberto with whom I spent so much time and shared much more than the sport passion.
Surfing this grateful wave, I'm glad and proud of being born in Brozzo, a very small village in Brescia province. Although the small boundary might appear too tight, I still need a reference frame that like me, helped other people to think broader. With Brozzo comes all the broad family including my parents, brother, sister, cousins aunts, uncles, grandparents and Giulia's family, that became part of my life.

I believe I took something good from each of you. I hope I left something in return.

Abstract

The applied particle physics has a strong R&D tradition aimed at rising the instrumentation performances to achieve relevant results for the scientific community. The know-how achieved in developing particle detectors can be applied to apparently divergent fields like hadrontherapy and cosmic ray detection. A proof of this fact is presented in this doctoral thesis, where the results coming from three different projects are discussed in likewise macro-chapters.

A brief introduction (Chapter 1) reports the basic features characterizing a typical particle detector system. This section is developed following the data transmission path: from the sensor, the data moves through the front-end electronics for being readout and collected, ready for the data manipulation. After this general section, the thesis describes the results achieved in two projects developed by the collaboration between the medical physics group of the University of Turin and the Turin section of the Italian Nuclear Institute for Nuclear Physics.

Chapter 2 focuses on the TERA09 project. TERA09 is a 64 channels customized chip that has been realized to equip the front-end readout electronics for the new generation of beam monitor chambers for particle therapy applications. In this field, the trend in the accelerators development is moving toward compact solutions providing high-intensity pulsed-beams. However, such a high intensity will saturate the present readout electronics. In order to overcome this critical issue, the TERA09 chip is able to cope with the expected maximum intensity while keeping high resolution by working on a wide conversion-linearity zone which extends from hundreds of pA to hundreds of μA. The chip gain spread is in the order of 1-3% (r.m.s.), with a 200 fC charge resolution. The thesis author took part in the chip design and fully characterized the device.

The same group is currently working on behalf of the MoVeIT collaboration for the development of a new silicon strip detector prototype for particle therapy

applications. Chapter 3 presents the technical aspects of this project, focusing on the author's contribution: the front-end electronics design. The sensor adopted for the MoVeIT project is based on 50 μm thin sensors with internal gain, aiming to detect the single beam particle thus counting their number up to 10^9 cm^2/s fluxes, with a pileup probability $< 1\%$. A similar approach would lead to a drastic step forward if compared to the classical and widely used monitoring system based on gas ionization chambers. For what concerns the front-end electronics, the group strategy has been to design two prototypes of custom front-end: one based on a transimpedance preamplifier with a resistive feedback and the other one based on a charge sensitive amplifier. The challenging tasks for the electronics are represented by the charge and dynamic range which are respectively the 3 - 150 fC and the hundreds of MHz instantaneous rate (100 MHz as the milestone, up to 250 MHz ideally).

Chapter 4 is a report on the trigger logic development for the Mini-EUSO detector. Mini-EUSO is a telescope designed by the JEM-EUSO Collaboration to map the Earth in the UV range from the vantage point of the International Space Station (ISS), in low Earth orbit. This approach will lay the groundwork for the detection of Extreme Energy Cosmic Rays (EECRs) from space. Due to its 2.5 μs time resolution, Mini-EUSO is capable of detecting a wide range of UV phenomena in the Earth's atmosphere. In order to maximize the scientific return of the mission, it is necessary to implement a multi-level trigger logic for data selection over different timescales. This logic is key to the success of the mission and thus must be thoroughly tested and carefully integrated into the data processing system prior to the launch. The author took part in the trigger integration in hardware, laboratory trigger tests and also developed the firmware of the trigger ancillary blocks.

Chapter 5 closes this doctoral thesis, with a dedicated summary part for each of the three macro-chapters.

Keywords: Front-end; Readout electronics; ASIC; Trigger logic; Data management; Particle therapy; Cosmic ray detections.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Particle detector systems: basic concepts

Particle detectors are the most accurate artificial senses which humans can take advantage of to study the deepest Nature details. The physical interaction between a particle and the interacting material reveals information appreciable only after a certain amount of manipulation steps depending on the researched data and the event phenomenology. Particle detectors are employed in many applications from astrophysics to microbiology, medical imaging, and quality assurance, to the fashionable high-energy physics and dark matter experiments. Cutting-edge research centers, as well as common life automotive or kitchen devices adopt these devices, that even though are target-customized, they are united by a common work-flow and basic functional blocks. Therefore, a typical particle detector is characterized by the fact that a sensor or one section of it detects a signal generated by interaction with the crossed sensitive volume due to the energy deposited, typically by ionizing the medium. This signal, generally represented by mobile charge carriers (e.g. electrons-positive ions in gas detectors or electrons-holes in solid-state detectors), is converted into an electrical pulse induced by the charge drifting to the detector collecting electrodes. At this point, an amplification stage is mandatory prior to further processing, storage, and analysis phases.
Figure 1.1 depicts the functional block representation of a typical detector system.

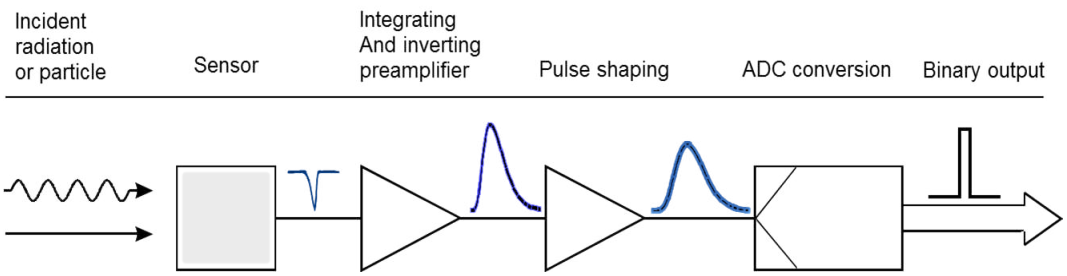

Fig. 1.1 Basic detector functions: radiation is absorbed in the sensor and converted into an electrical signal. This low-level signal is integrated in a preamplifier, fed to a pulse shaper, and then digitized for subsequent storage and analysis.

The very front-end electronics mainly amplify the detector signal following the pulse profile or shaping it (e.g. integrating a current signal), with a certain care in keeping the noise at a low value. Once the electrical pulse crosses a comparator threshold the information is digitized. Further processing steps elaborate the information in such a way it can be transmitted with high throughput and a reliability level which is intrinsic to the adopted standard format. The digital readout circuitry interacts with data links in order to transmit the information for the on-line data selection (i.e. trigger logic) or the direct on-disk archiving. The saved data are then deeply analyzed off-line through algorithms developed in a simulation where the detector geometry, behavior and the way the data is processed by the electronics have been modeled by means of powerful software. The detector mechanics (positioning, holding, the FoV), the cooling system and eventually the radiation hardening are important aspects that involve professional profiles from various fields.
In applications like the particle beam monitoring system or the quality assurance for radiation therapy, the detector is used to probe a certain physical quantity (e.g. the dose that is the energy per mass unit) which is a cumulative effect and is not related to the single particle. The electrical signal is integrated and through a certain technique, the detecting system reveals the needed information at the accuracy level required. As an example of this approach, the TERA chip (see Chapter 1) integrates the input signal and adopts the subtraction of a fixed charge quantum to the amplifier output. The modern detector system complexity allows detecting the single photon or the single charged particle managing the discrimination of such events at high-rates up to billions per square centimeter. The Ultra Fast Silicon Detectors (described in Chapter 2), for instance, are devices able to detect ions with tens of micrometers spatial resolution and tens of picoseconds timing detail. For the MoVeIT project,

such kind of detector has been chosen with the aim to discriminate the single ion crossing the silicon detector layer in a particle beam to $\sim 10^9$ protons $cm^{-2}s^{-1}$ flux. The sensor sensitive area segmentation is the standard approach adopted whenever high rates and/or high spatial resolution, are required. The single sensor segment is thus independent and it is associated to a dedicated front-end channel. The signal processing is typically parallelized and very often particle physics experiments take advantage of Field Programmable Gate Arrays (FPGA), dealing with high channel densities that mean high data throughput transmitted at high rates. FPGA are commercial chips with software programmable logic gates processing in parallel a large amount of rather easy tasks. Modern FPGAs are really System on Chip (SoC), embedding programmable logic (PL) and CPUs on the same silicon die. Although these high-performance commercial devices are getting better and better, the level of optimization of high-tech detector systems requires fully custom front-end electronics integrating a large number of channels and functional blocks. This integrating capability is guaranteed thanks to the deep sub-micron design technology nodes. It was in the field of high energy physics that in the 80s, for the first time a custom integrated electronics have been coupled with a silicon micro strip detector to measure short living particle decay tracks. Nowadays most of the integrated front-end electronics for particle detectors are developed as Application Specific Integrated Circuits (ASICs) fabricated in CMOS technologies. These ASICs generally embed a mixed-signal logic where from the analog very front-end, the signal is collected from the sensor and then digitized to be ready for an on-chip or off-chip manipulation.
A complex detector system e.g. a telescope is a device organized in multi functional sub-blocks including the sensor (typically photomultiplier tubes), the front-end (ASIC) and the readout (FPGA) electronics. Looking for rare events and detecting them over a wide spatial scanning campaigns (astronomical dimension, in the telescope case), it is mandatory to have a certain data selection, thus to manage the computing power and the storage volumes. A filtering algorithm, called trigger, is a fundamental piece in a detector work-flow dealing with a large amount of data. The trigger logic is based on physics models that have been coded and simulated with the goal to maximize the possibility to detect significant events, extracting them from the background signal. The trigger logic should recognize the event evolution features and distinguish among tricky or ambiguous cases induced by hardware complications of fake trigger sources (e.g. flying over a city in the nighttime, could be a source of fake triggers for a radiation sensitive telescope). The Mini-EUSO UV-telescope

trigger and its hardware integration is described in this thesis at Chapter 4.
Finally, it is possible to remark the here presented main points, highlighting the figures of merit for a general detection system functional blocks:

- Detectors: efficiency, speed, granularity, resolution;
- Trigger/DAQ: efficiency, data compression, data throughput, physics models;
- Off-line analysis: signal and background discrimination, physics models.

Further details and technicality concerning front-end and readout electronics systems will be given in the following chapters.

Chapter 2

Design and characterization of the TERA09 ASIC

2.1 Introduction

Particle therapy, also called hadron therapy, is an oncological technique based on accelerated ions hitting a target. The main goal of particle therapy is the tumor treatment while sparing the healthy tissues in the surrounding diffusion path. The idea is to deliver in a specific and confined volume a sufficiently high dose required for destroying the sick tissue, keeping safe the rest thus avoiding serious or even irreversible damage or complications [1] [2]. This approach takes advantage of the way ions interact with the medium. When ions cross the human body they are mostly not deflected, so that the tissue damage is reduced until these ions, stopping, release the main part of their energy as a peak. Taking advantage of this peak, called "Bragg peak", hadron therapy maximizes the dose delivered to the tumor sparing the surrounding healthy tissue. The result is an increase of complication-free tumor cure in comparison to conventional radiotherapy. Although the number of hadron therapy centers for proton and carbon ions treatment is constantly increasing, this clinical approach is not so widespread, principally because of the high costs for building and operating the facilities. For this reason, the trend in the new developments is towards the realization of smaller and compact particle accelerators. An accelerator with these characteristics provides pulsed beam with a much higher intensity in each pulse than conventional uniform beams to maintain the same released dose-per-treatment

to the patient. Particle therapy accelerators are equipped with monitor detectors as required by automated control systems and they are going to cope with this high intensity pulsed beam structure. As example of these new detectors, a multiple gap monitor chamber takes the advantage of using gaps of different width in the possibility to study and estimate the effects of the ion recombination generated into the gaseous detecting medium when a high ionization density occurs.

The Italian National Institute for Nuclear Physics (INFN), in collaboration with the University of Turin, developed a new front-end and readout electronics suitable for detectors that have to deal with the new technique features, keeping the compatibility with the last decades standard procedure which is still largely diffused. The core of this front-end electronics is a 64-channel Application Specific Integrated Circuit (ASIC) named TERA, consisting of a current-to-frequency converter followed by a counter. The TERA chips, tailored for clinical applications, have been used in several clinical devices both for quality control in radiotherapy [3] and for beam monitoring in particle therapy facilities [4] [5]. These chips are all based on the recycling integrator principle [6] to convert the input charge into pulse counts, or equivalently the input current into a pulse count frequency, each count corresponding to a fixed quantum of charge. This conversion method offers several advantages as the intrinsic lack of dead time and the very good linearity up to the maximum conversion frequency. Moreover, the charge collected at any input channel can be sampled asynchronously with the conversion operations by simply reading the corresponding counter.

The new version of the TERA chip, named TERA09, is an upgrade of its predecessor (TERA08) and has been specifically designed with the possibility to work with high-flux pulsed particle beams. Using a charge quantum of 200 fC, a linearity within $\pm 2\%$ for an input current range between 3 nA and 12 μA is obtained for individual channels, with a gain spread among the channels of about 3%. Moreover, the chip automatically calculates the partial and total sum of the counter values, which can be directly accessed in dedicated registers. In the design process, special effort has been devoted in maintaining as much as possible the backward compatibility such that the new chip can replace the older versions in any of the current devices with small impact on the acquisition system and on the power supply. The results of the tests, presented in [7], indicate the possibility to achieve an increase of about two orders of magnitude in dynamic range compared to the TERA08 without a significant loss in sensitivity and linearity, thus adding to this version the additional flexibility to

extend its use to high-flux particle beams applications. By connecting all the 64 ASIC channels to a common input, the current range can be increased 64 times preserving a linearity within $\pm 3\%$ in the range between and 20 μA and 750 μA. Furthermore, additional informations has been added at the ASIC characterization by means of radiation damage tests. As its predecessors, TERA09 is not going to be directly exposed to the particle beam flux; nevertheless, it has been important to estimate the expected Single Event Upset (SEU) rate in a clinical treatment room, mainly due to the interaction of backscattered neutrons in the extended digital logic area.

2.1.1 The author's contribution

The author's contribution on the work presented in Chapter 2 can be separated into the following three main parts:

ASIC design contribution: preliminary and feasibility study of the digital architecture; preamplifier optimization and design of the LVDS clock receiver.

Characterization of the TERA09 chip: development of the data acquisition software for the chip readout, laboratory characterization and radiation damage test through the single event upset detection.

PCB development: schematic design of the TERA09 test board (the one used during the characterization); technical advice on the design and test of the Front-End (the board embedded into the detector employed in clinical treatments).

2.2 Interaction of photons and ions with matter

Ions and photons interact differently with matter and this phenomenon shapes in a different way the depth-dose profiles in the two cases. Interacting with the medium, a high energy photon (X or γ)undergoes to strong energy release due to the atomic scattering and the ionization process. The stochastic interaction of photon with matter and the different amount of energy released within it may lead to different kind of events like inelastic scattering, Compton scattering, photoelectric effect and

pair production. Figure 2.1 depicts the depth-dose (dose is the energy per unit of mass) profile of photons and ions in water, which density is similar to human body tissues. It is possible to appreciate the contrast between the narrow energy release of ions and the broad effect of photons. The built-up at the photon entrance is then followed by an exponential dose decrease suggesting an undefined range. Charged particle radiation is directly ionizing and the released energy per unit length (stopping power), is higher for slower particles that suffer more for scattering:

$$\frac{dE}{dx} \sim \frac{1}{v^2} \tag{2.1}$$

Where v is the projectile velocity, dE/dx is the energy released per unit length.

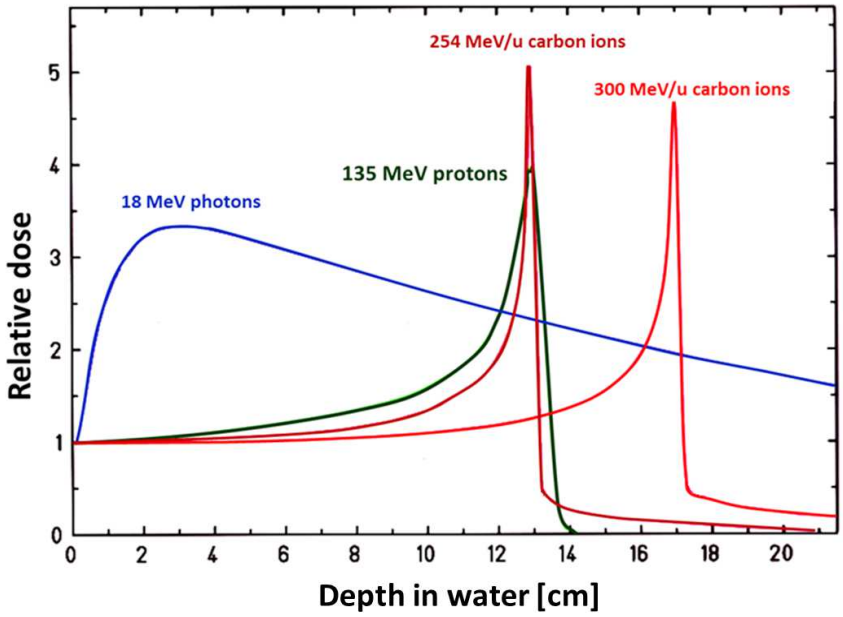

Fig. 2.1 Depth–dose distributions of protons, carbon ions and photons in water.

At energies > 1 MeV electrons have a relativistic energy; in this condition the dE/dx is no more dependent on the energy

$$\frac{dE}{dx} \sim \frac{1}{c^2} \tag{2.2}$$

with c the speed of light.

An electron with $v \approx c$ is responsible fo a few MeV/cm dose which is constant. The low mass of electrons implies a strong lateral scattering and makes the ions preferable for a confined energy deposition. As shown in Figure 2.1, protons and ions have a low dose profile at the entrance, which remains almost flat until the range end where at the Bragg peak, the particle extinguish their motion through a high energy release. Figures 2.2 qualitatively shows the different effect of photons and proton, in terms of target approach and boundary sparing.

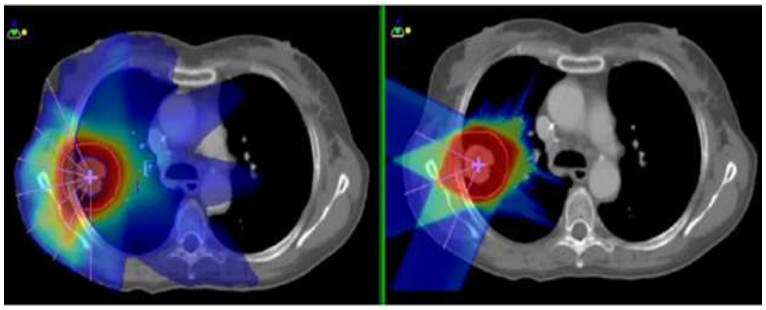

Fig. 2.2 Explicative computed tomography picture showing the selectivity differences between photons (left) and protons (right).

In particle therapy, the atomic nuclei are accelerated at the kinetic energy range $\approx$ 50 - 400 MeV/u and they mainly interact with the atomic electrons of the medium by Coulomb's force. The Bethe-Bloch equation describes the stopping power model for heavy charged particles:

$$-\frac{dE}{dx} = \frac{4\pi m_e N_A}{m_e c^2}\frac{z^2}{\beta^2}\frac{Z}{A}\rho ln\left(\frac{2m_e c^2\beta^2\gamma^2}{I}\right) - relativistic\_terms \qquad (2.3)$$

where $\beta = v/c$, v the particle speed, c the speed of light, ze is the particle charge, N_A is the Avogadro's number, m_e is the electron mass at rest. Z, A, ρ and I are the atomic number, the mass number, the density and the mean excitation energy of the medium.

When $v << c$, $\beta << 1$) therefore the formula can be reduced to

$$-\frac{dE}{dx} \sim \frac{Kn_0 z^2}{v^2}[ln(2m_e v^2/I)] \qquad (2.4)$$

where K is a constant, n_0 the electron density of the target material.
In this situation the terms $\frac{z^2}{v^2}$ explain that the stopping power increases if the projectile

velocity decreases with a logarithmic trend. The minimal $\frac{dE}{dx}$ occurs for $E = 3Mc^2$, with M the projectile mass. In a mono-energetic particle beam the particles undergo to different scattering trajectories thus resulting in different ranges. This phenomenon, called straggling, is quantifiable as 1% for the protons mean energy range ([6]) and it goes almost like the square-root of the particle mass, as depicted in Figure 2.3.

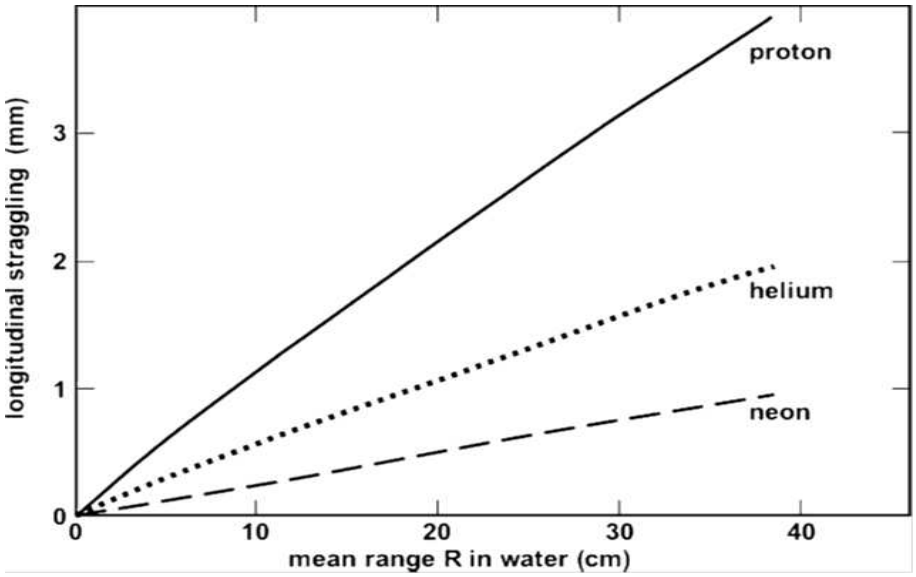

Fig. 2.3 Straggling effect highlighted comparing proton helium an neon beams in water.

The ion beam presents a lateral spreading due to multiple interaction with the lighter electrons. For a large diameter beam this effect involves only the beam external particles since the central part has a scattering-out rate which is compensated with an almost equivalent scattering-in; this fact is no longer true for narrow beams, as reported in Figure 2.4.

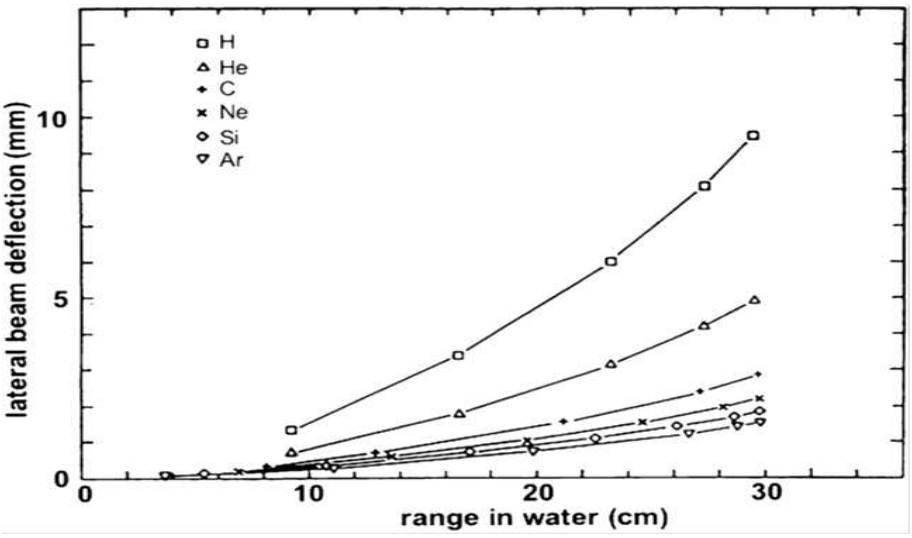

Fig. 2.4 The deflection of a beam with the increasing path length.

An infinite narrow proton beam deflects $\sim 5\%$ of its initial range whereas heavy charged particles suffer less for angular spread.

2.2.1 The Bragg peak profile

The ion energy loss measured with ionization chambers reveals a Gaussian shaped Bragg peak due to multiple scattering phenomena. With a minor interaction with the medium, particle with higher energy penetrate more until the resting point where the Bragg peak will be larger and this width is responsible for the gradient of the dose distribution distal fall-of. Treating a tumor 10 cm deep the FWHM will be 4-5 mm. A gradient at half of this value would be optimal but not easy to control, since the penetrated tissues have no homogeneous density.
Treating low depth targets it is usual to artificially enlarge the Bragg peak artificially using millimeters thick passive absorbers (as explained in the following section); with larger peaks the treatment time is reduced, especially in active scanning techniques where the treatment is divided in slices spaced by a peak maximum width.

2.3 Particle therapy: basic concepts

2.3.1 Technique overview

The main goal of a therapy in general is to treat the patient controlling as much as possible the interaction with the target and sparing the healthy tissues. In radiotherapy the reliability and the confined energy release of accelerated ions make them the most suitable choice, dealing with tumors while sparing the rest. In 2018 67 medical centers treat patients with protons and 11 uses carbon ions (numbers updated on April 2018 [8]). Nowadays the number of particle therapy centers is increasing and tens of centers are already under construction. The Particle Therapy Cooperative Group (PTCOG) publishes the number of patient treated worldwide with accelerated ions: the total number of hadron therapy patients was 108.238, in 2012 and rose up to 174.512 at the end of 2016. More details can be found in the hereafter reported table [3].

Table 2.1 comparison between current-mode amplifier and CSA

Type	**Number of treated patient**	**Period**
He	2054	1957-1992
Pions	1100	1974-1994
C-ions	21580	1994-2016
Other ions	433	1975-1992
Protons	149345	1954-2016
Grand Total	174512	1954-2016

Already in 1946 Robert R. Wilson, an Harvard physicist published a paper explaining some ion properties and possible application in medicine [4].
As already mentioned ions offer a better dose confinement than electrons and photons, having a narrow peak energy release and the peak position in the volume depends on the particle beam energy. Furthermore it is possible to optimize the dose conformation modulating the peak with a Spread-out Breagg Peak (SOBP). The SOBP technique (Figure 2.5) is the result of many Bragg curves grouped (partially overlapped) using passive absorbers or active systems. The active system techniques imply the particle deflection provided by magnets and the range variation tuning the energy [9].

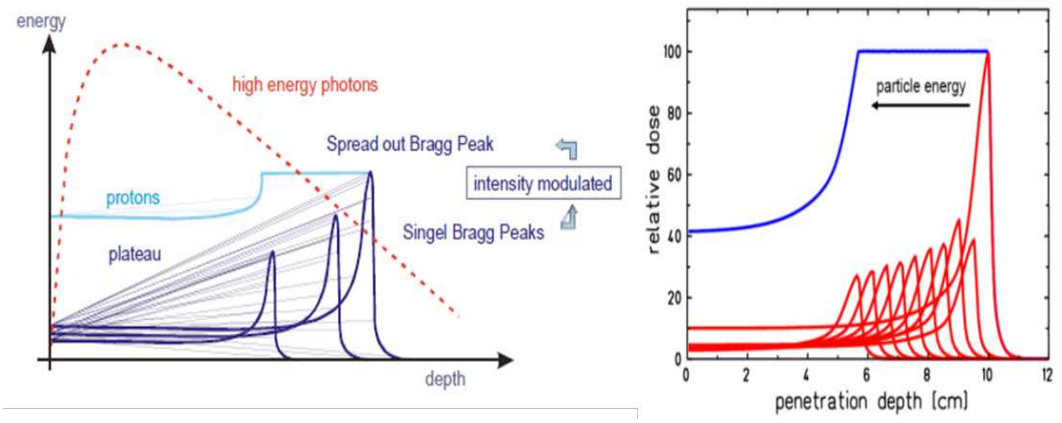

Fig. 2.5 Comparison between the photon dose distribution and a SOBP (left). SOBP principle: the distal peak has the highest amplitude (right). Picture taken from [10].

The fact that the penetration depth can be precisely selected, the path is finite and the maximum energy release is just before the ion rest are the great qualities of ions in particle therapy. Moreover, the shallow entrance with low $\frac{dE}{dx}$ is fundamental for the healthy tissue sparing, allowing the deep sited tumors.
A further improvement in the overall technique is given by heavier ions like the carbon ion. Carbon ions have $\sim 1/3$ of the lateral scattering effect of protons, providing a higher Relative Biological Effectiveness (RBE) that means a better damaging power for the sick tissues.
As already mentioned with SOBP the dose can be kept uniform using filters like rotating wedges, propellers or ridge filters; passive scatterers are used to spread the transverse dose whereas collimators regulate the contours of the field (Figure 2.6). Even though carbon ions offer advantages over protons in particle therapy, the former suffer for nuclear fragmentation that could be critical using passive systems for beam modulation. The nuclear fragments are responsible for a tail in the depth-dose profile that extends over the Bragg peak and the effect is emphasized increasing the penetration depth. This fact is due to their lower Z that leads to higher ranges (having almost the same velocity of the particle in the beam). Synchrotron accelerators allow to overcame these phenomena, varying the extraction energy in a dynamic beam delivery system.

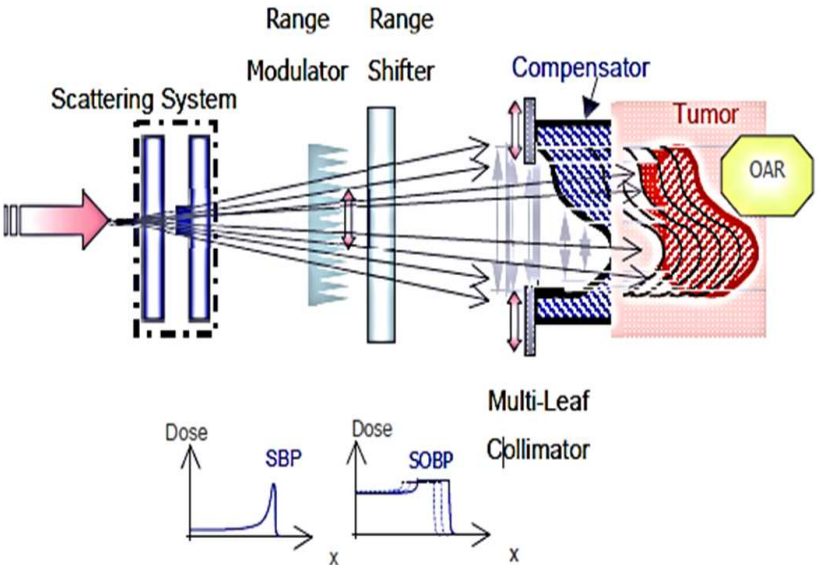

Fig. 2.6 Drawing of a passive modulating system. The transverse distribution is modeled with spreading scatterers. A Ridge filter followed by a range shifter provides the correct energy and depth for the SOBP components. The distal profile is then shaped with a properly modeled compensator. Picture taken from [10].

2.3.2 Ion accelerators in particle therapy

The new technological trend in charged particle radiotherapy is to realize compact, reliable accelerators with advanced beam performance. These characteristics imply for instance a significant money effort reduction, facilitating the access to hadron therapy to more patients. A more compact accelerators means a reduced radius of the accelerating structure therefore, the magnetic field $\bar{B}$ has to increase in order to keep particles in the correct orbit. The increase of $\bar{B}$ makes the focalization worse; to overcome this problem, radio frequency accelerating systems are used.
In a particle accelerator the Radio Frequency (RF) cavities are placed along the beam pipe. The RF cavity geometry is designed such as the electromagnetic waves created by a RF power generator resonate, pushing forward the incoming charged particles moving in the vacuum. The electromagnetic field drives the particles and it oscillates with a given frequency, needed to arrange particles arriving with a slightly inhomogeneous energy set: in-time particle are not accelerated whereas particle arriving earlier are decelerated as well as particle in late are accelerated. In this way the particle moves in bunches picking up energy from an increasing field provided

by the accelerator magnets.
In order to deliver a standard therapeutic dose of 2 $Gy/(min \cdot l)$, the typical beam intensities ranges between $1.8 \cdot 10^{11}$ and $3.6 \cdot 10^{11}$ particles per minute. However, the new projects of more compacted accelerators need higher magnetic fields to keep particles in the correct orbits and therefore higher intensities provided by radio frequency accelerating systems, to overcome the focalization problems. Thus, the intensity for a new generation accelerator for hadron therapy, is in the order of 10^{14} particles for minute. Such a high intensity is going to lead inefficiency problems in standard detector currently employed in the particle therapy facilities.

2.4 Gas filled detectors

A widely diffused type of charged particle detector is the one based on the ionization of a gas crossed by a charged projectile that releases enough energy to create ion-electron pairs along its path. Applying a certain voltage difference between two points at the gas volume boundary, those ion-electron pairs move as the electric field guide them at the electrodes collection, separating charges by their polarity. During the collection, the charge motion induces a certain signal that is the event evidence seen by the front-end electronics. The information about the intensity and the energy of a particle beam is therefore extrapolated from the resulting charge, current pulse or voltage difference at the detector electrodes. Considered that, in a gas detector the charge associated with an electron or a singe ion is by far too low to result in a detectable electric signal, this device needs a large number of charged pairs and has to collect them sharply, before recombination effects attenuate the signal. Engineering the detector, important features like the gas type, the geometry and the high-voltage range play a fundamental role in the device efficiency optimization, mitigating issues like the recombination phenomena.
The random thermal motion characterizes the dynamic of a resting gas, where $10^{-6} - 10^{-8} m$ is the the mean free motion range of the medium molecule. The same kind of motion affects also the free electrons or ions and this leads to a certain spreading of the charge that in a certain sense reject an ideal point-like collection resulting in a normal distribution with a standard deviation increasing with time.

2.4.1 Gas filled detector operation modes

A gas detector is a versatile device that generally can be adapted to different applications changing the voltage difference applied at its electrodes. Figure 2.7 depicts the behavior of a gas filled detector as a function of the biasing voltage. Since the main issue for high intensity pulsed-beam, is the recombination region we will focus on this part, toghether with the proper ion chamber region.

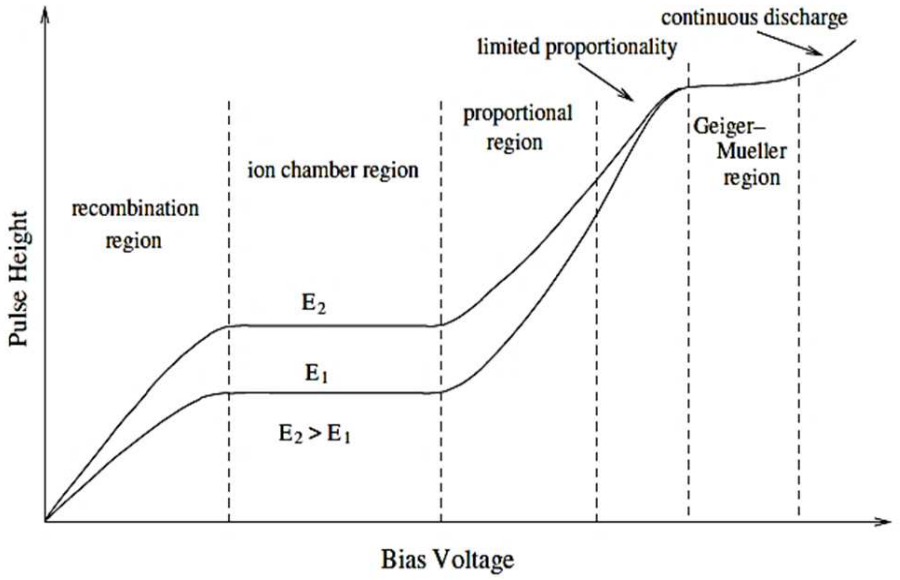

Fig. 2.7 Operation regions for a gas filled detectors. The pulse height depends on the bias voltage which define the detector mode of operation. E_1 and E_2 represents radiation with different energies. Picture taken from [11].

The Coulomb force drive the attraction of two opposite polarity particles and it is the basis of recombination. The electron-ion pairs recombine to form neutral molecules with a rate that is inversely proportional to the applied voltage. In order to precisely measure the particle beam energy it is crucial to reduce as much as possible the signal losses. Therefore an ionization chamber ideally works in a region where all the charges are detected and increasing the voltage does not increase the efficiency any longer. This working range is also called saturation region. In this situation the front-end electronics is fed with a signal which amplitude is proportional to the incident radiation. Increasing the bias voltage enhance the capability to collect charges nevertheless it has to be limited in order to avoid the detector discharge, that would transform a detector from inefficient to quantitatively blind.

2.5 Parallel plate ionization chambers

As depicted in Figure 2.8, the easiest way to develop a gas ionization chamber is adopting a parallel plates geometry: a gas volume is enclosed between two planar electrodes.

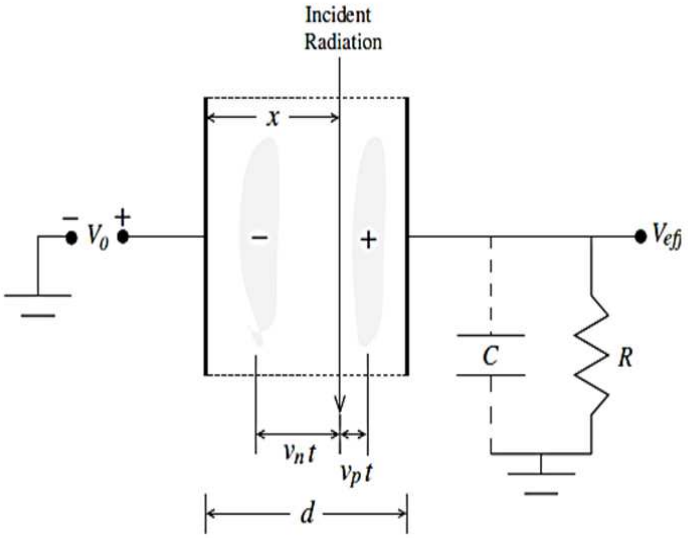

Fig. 2.8 Schematic representation of a parallel plate gas-filled ionization chamber. Considering a higher electron mobility they spread in a larger volume than positive ions; this is the reason why a larger negative zone is depicted.

The movement of the collected charges changes the electric potential. The result of this perturbation can be appreciated with an analytical modeling of the phenomenon:

$$V_{eff}(t) = V_0 - V_{np}(t) \tag{2.5}$$

where V_0 is the static bias and $V_{np}(t)$ is the potential difference at a certain time t, associated to positive and negative charges in the detector.
With N_0 ion pairs at t, the electrons with an average velocity v has a kinetic energy:

$$K_n(t) = N_0 eEv_n t = \frac{v_0}{d} N_0 e v_n t \tag{2.6}$$

The intensity of the electric field has been assumed uniform within the gas volume; $E = \frac{V_0}{d}$ with d the electrodes distance. The same results can be applied to the ions and replacing n (negative) with p (positive):

$$K_p(t) = N_0 e E v_p t = v_0 N e v t \tag{2.7}$$

The dualism with a parallel plates capacitor continues considering the detector capacitance C, and explaining the potential energy as

$$U_{ch} = \frac{1}{2} C V_{np}^2 \tag{2.8}$$

and the total energy delivered by the bias voltage V_0 is

$$U_{total} = \frac{1}{2} C {V_0}^2 \tag{2.9}$$

with a short circuit the just named U_{total} equals the sum of the potential and kinetic terms:

$$U_{tot} = U_{ch} + K_p + K_n \tag{2.10}$$

$$\frac{1}{2} C V_0^2 = \frac{1}{2} C {V_{np}}^2 + N_0 e E v_n t + N_0 e E v_p t \tag{2.11}$$

Describing the effective potential as $V_{eff} = V_0 - V_{np}$ it is possible to rearrange the formalism:

$$(V_0 - V_{np})(V_0 + V_{np}) = \frac{2 N_0 V_0 e}{C d} \cdot (v_p + v_n) t \rightarrow V_{eff} \approx \frac{N_0 e}{C d} \cdot (v_p + v_n) t \tag{2.12}$$

with

$$V_0 + V_{np} \approx 2 V_0 \tag{2.13}$$

With a mass $\approx 1/1000$ the one of the proton, the electron motion shapes the initial signal, as reported in Figure 2.9.

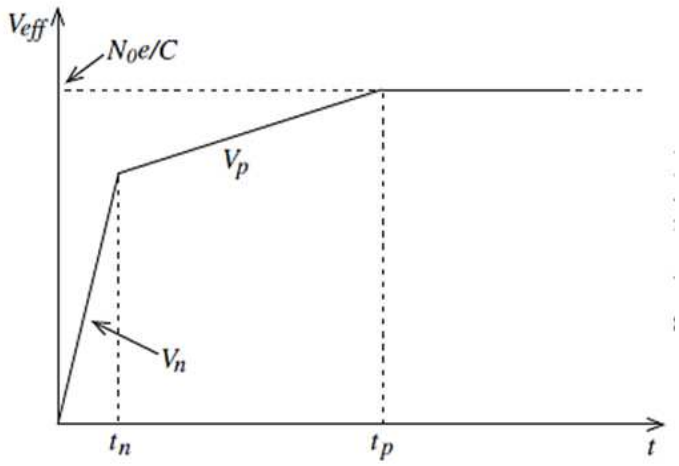

Fig. 2.9 Ideal representation of a pulse profile in a parallel plate ionization chamber. Electrons and ions shapes with different gradient, due to the mass difference and thus the different collection speeds. V_n and V_p are the signal amplitudes related to negative and positive charges respectively. Picture taken from [11].

The interval $t_n = \frac{x}{v_n}$ is the time needed by an electron to travel toward the anode form a distance x. With a lower rate, ions are collected to the cathode after a time $t_p = \frac{(d-x)}{v_p}$. The maximum voltage amplitude is reached after the last particle has been collected. The signal shape has three different slope:

$$\frac{N_0 e}{Cd} \cdot (v_p + v_n)t : 0 \geq t \leq t_n \tag{2.14}$$

$$\frac{N_0 e}{Cd} \cdot (v_p + x)t : t_n \geq t \leq t_p \tag{2.15}$$

$$\frac{N_0 e}{Cd} : t \geq t_p \tag{2.16}$$

In a real circuit, R-C effects smooth the signal profile as shown in Figure 2.10. Although sharp responses (i.e. small time constants) allows a high-rate discrimination, a too small R-C might lead to information losses and lack of linearity.

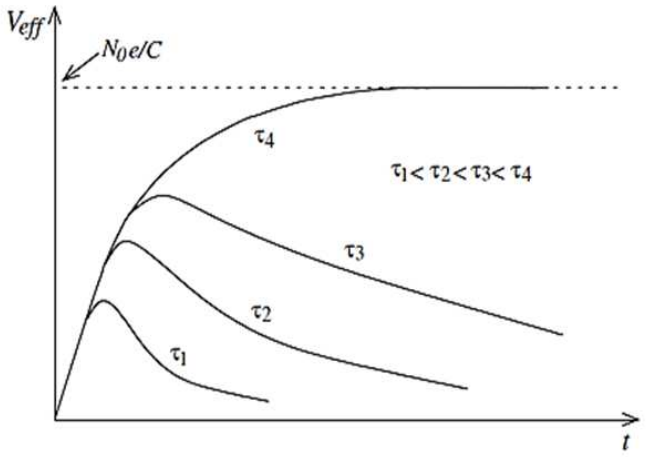

Fig. 2.10 Pulse profile representation. The smoother shapes (respect to Figure 2.9) are due to the difference between the detector time constant and the charge collection period. Picture taken from [11].

2.5.1 Advantages and drawbacks of ion chambers

Due to its versatility a ionization chambers can be used in different way therefore the positive and negative features of this device are somewhat related to the specific application. Nevertheless it is still possible to outline some general aspects to better understand where an ionization chamber could be placed among the applied physics detectors.

Using an ion chamber advantages in the saturation mode, the ionization current is directly proportional to the incident radiation energy and it is almost independent of the bias voltage; this increases its reliability under inevitable power supply fluctuation and drift. Avoiding the charge multiplication, a gas detector working as ionization chamber does not suffer for gas quality changes. Simple, widely diffused, ionization chambers have anyway technical limitations like the low current flowing through the ionization chamber that makes it a not suitable detector choice, for low radiation applications. Another drawback in using a radiation chamber is the dependence from temperature and pressure variation. The user should consider this aspect which is indeed relevant only for high resolution detection.

2.5.2 Collection efficiency in ionization chambers

In clinical application with charged ions, the dose released to the target is generally determined by measuring the amount of charge Q collected by ionization chambers placed just before the patient. As already mentioned, there is a certain difference between the measured charge and the one that would be collected with a saturated detector. The charge collection efficiency is defined as:

$$f(V) = \frac{Q}{Q_{sat}} \tag{2.17}$$

Rising the bias voltage there is a sharp increase in the charge collection efficiency until it reaches a plateau, once approached the saturation charge Q_{sat} (Figure 2.11). The bias voltage is limited by the fact that this applications requires to work in the saturation region, avoiding charge multiplication or the even worse breakdown situation.

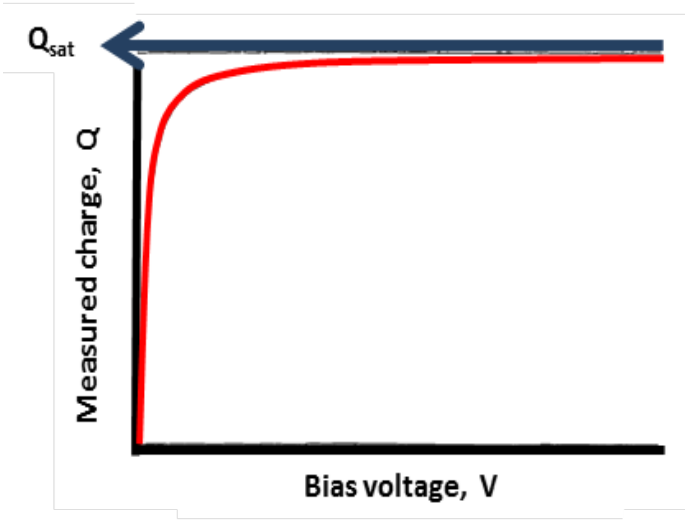

Fig. 2.11 Typical plot of measured charge Q as a function of the chamber polarizing potential V. The measured charge increases almost linearly for low voltages and approaches asymptotically the saturation charge Q_{sat} at higher voltages.

Knowing the collection efficiency for a certain bias voltage and electrodes separation, the $Q(V)$ measure allows to extrapolate Q_{sat}. Two main recombination processes are considered: the initial recombination and the general recombination. Initial recombination concerns the events occurring inside the ionization track and

it is not dose-dependent and $1/Q$ varies linearly with $1/V$. General recombination happens with ions created in different track therefore, it is dose-dependent. For general recombination, close to the saturation region ($f \geq 0.7$) l/Q varies linearly with l/V^2 with continuous beams and linearly with $1/V$ in pulsed beams. For ionization detector filled with electronegative gases, only the general recombination is significant if the device is working in the saturation region [12].

2.5.3 Slow rate issues in ionization chambers

In active scanning techniques for particle therapy, where the beam energy has to be selected and changed on-line from the accelerator control system, synchrotron accelerators are generally used. One of the main characteristics of a synchrotron is that the radio-frequency accelerates the particles in bunches. The accelerating machine has to be designed with thigh constraints in order to satisfy the beam quality features required in medical applications. An important parameter is the spill homogeneity during the beam extraction period. In order to perform on-line beam monitoring during an active scanning procedure, it must be possible to switch the beam on and off according to the required dose. Moreover, a slow beam extraction (through slow synchrotron extraction, RF-KO extraction or betatron core driven extraction techniques) is needed for meeting quality assurance standards [13]. Slowing down the beam extraction carries technical issues mainly due to the fact that the final dose during a patient treatment has to be kept constant (and the treatment time can not increase). Therefore, smaller pulses (10^8 p/cm^2 s^{-1} instead of 10^9-10^{10} p/cm^2 s^{-1}) with higher repetition rate are provided. smaller pulses means less statistics, especially for the segmented detectors adopted for the beam centroid position detection (that has to be measured typically with a space resolution within 0.1 mm). The most accurate beam monitoring systems (e.g. the CNAO one) have a charge resolution in the order of 200 fC, value that does not allow to measure the charge fluctuation between two different pulses [9]. A possibility to overcome the small signal intrinsic statistics would be introducing an additional gain layer to the IC, e.g. in Gas Electron Multiplier detectors (GEM).
During the data acquisition with ICs, it is important to integrate and measure the entire collected signal, in order to avoid the introduction of underestimation effects. In a typical 3 mm gas thickness, the total drift time for electrons is $\sim$0.3 μs and the one for ions is $\sim$100 μs. The classic accelerators adopted nowadays in clinical

centers have a spot duration typically larger than 200 μs, therefore, the monitor chambers large collection time is not influencing the quality of the measure. Another situation shows up with the new accelerator versions where the monitoring system has to deal with a 1-10 μs pulsed beam structure and the related charge recombination issues. More details are hereafter reported.

2.5.4 Recombination rate

The signal loss due to ion pair recombination depends on the concentration of different polarity charges in a given point and on the interaction time. The recombination rate volume-normalized can be expressed as:

$$\frac{d}{dV}\left(\frac{dN}{dt}\right) = \alpha C^{+}C^{-} \tag{2.18}$$

where α is the recombination coefficient, and C^{+} and C^{-} are the positive and negative ion concentrations. The same equation can be written highlighting the charge loss $q = eN$ as a function of the charge densities $\rho^{+} = eC^{+}$ and $\rho^{-} = eC^{-}$ (e is the electron charge).

$$\frac{d}{dV}\left(\frac{dq}{dt}\right) = \frac{d\rho}{dt} = \frac{\alpha}{e}\rho^{+}\rho^{-} \tag{2.19}$$

The bias voltage V, the electrodes distance d and the positive and negative ions mobility in the detector gas, k^{+} and k^{-}, are directly responsible to ion concentration and interaction time. Once the electrodes are polarized, negative ions drift toward the anode whereas positive ions are collected at the cathode, both having a drift velocity proportional to the electric field E as:

$$v^{+} = k^{+} \cdot E v^{-} = k^{-} \cdot E. \tag{2.20}$$

In an electronegative gas (e.g. air), free electrons easily approach gas molecule creating negative ions; in this situation the two ion polarities have an almost equivalent mobility. Higher mobility ions tend to escape more to recombination. If the gas is not electron affine, the free electrons recombination is not observed and k^{-} is $\sim 3 \cdot k^{+}$.

2.5.5 Collection efficiency in pulsed radiation beams

For high-intensity pulsed beams, the particle pulse duration is shorter than the mean transit time for an ion in the cavity, the ionization has to be studied for the time interval concerning the single pulse. Therefore the charge density per pulse ρp is just instantaneous [14].
The drift of the ions creates in the inter-pulse period a positive ion region near the cathode N,a negative ion region near the anode P and a central recombination region. The recombination region width can be modeled as:

$$w(t) = d - t(k^+ + k^-)\frac{V}{d} \tag{2.21}$$

mettere la figura 4.4

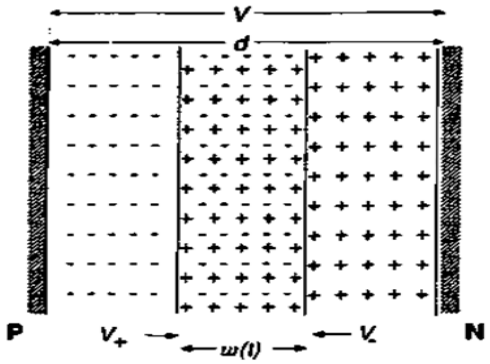

Fig. 2.12 Schematic diagram of a parallel-plate chamber in which a pulsed radiation beam has produced a uniform distribution of positive and negative charges. Positive charges drift toward the negative electrode N with velocity V^+ , while negative charges drift toward the positive electrode P with velocity V^-. Charge recombination is possible only in the overlap region having a width w chat decreases with time.

A the beginning $w = d$ ($t = 0$, $w(0)$) and then this space reduces to zero after a recombination period T. For $w(T) = 0$

$$T = \frac{d^2}{(k^+k^-)V} \tag{2.22}$$

If it is assumed that the number of free ions is due to a pure ionization result, positive and negative ions will be present in equal number and this balance is kept also

during the recombination too because this phenomena interest the two ion polarities (since positive ions recombine with negative ones). Therefore $\rho = \rho^+ = \rho^-$ and

$$\frac{d\rho}{dt} = \frac{\alpha}{e}\rho^2 \tag{2.23}$$

integrating:

$$\rho(t) = \frac{\rho_p}{1 + (\frac{\alpha}{e})\rho_p t} \tag{2.24}$$

with ρ_p, the instantaneous charge density produced during each pulse.

The recombination fraction f_r exists until free ions are present and can be expressed in the following way:

$$f_r = \frac{A}{\rho_p A d} \int 0^T (\frac{\alpha}{e})\rho^2(t)w(t)dt \tag{2.25}$$

where A is the electrode area .

$$f_r = [1 - 1/uln(1+u)] \tag{2.26}$$

where

$$u = (\frac{\alpha}{e})\frac{\rho_c d^2}{k^+ k^- V} \tag{2.27}$$

The collection efficiency $f = 1 - f_r$ is thus

$$f = \frac{1}{u} ln(1+u) \tag{2.28}$$

by approximation for $V \to \infty$, i.e., for $u \to 0$, expanding $ln(1+u)$ for, $u \to 0$

$$f = \frac{u - \frac{u^2}{2} + \frac{u^3}{3}}{u} \approx 1 - \frac{u}{2} \approx \frac{1}{1 + \frac{u}{2}} \tag{2.29}$$

For a pulsed radiation beam crossing a gas ionization chamber which is working close to the saturation, the charge collection efficiency is:

$$f^p g(V) = \frac{Q(V)}{Q_{sat}} = \frac{1}{1+\frac{\lambda^p g}{V}} \tag{2.30}$$

where $\Lambda_g^p = \alpha\rho_p d^2/2(k^+ + k^-)eV$. Equation (1.35) may also be written as

$$\frac{1}{Q} = \frac{1}{Q_{sat}} + \frac{\lambda^p g}{V} \tag{2.31}$$

with λ_g^p defined as $\lambda_g^p = \Lambda_g^p / Qsat$. If Q_H and Q_L are charges measured at V_H and V_L , respectively, then $f_g^p(V_H)$ can be written as:

$$f_g^p(V_H) = \frac{Q_H}{Q_{sat}} = \frac{Q_H/Q_L - \left(\frac{V_H}{V_L}\right)}{1-\left(\frac{V_H}{V_L}\right)} \tag{2.32}$$

2.5.6 The multiple gap ionization chamber

The pulsed beams in particle therapy require an optimized version of the detectors allowing a high precision even in the operation region where current detectors undergo to high recombination rates.
This thesis chapter is focused on the development of a new ASIC designed to be able to extend its dynamic range from a classical configuration to a high intensity pulsed beam. In the latter case mitigation approaches must be considered already from the detector that has to deal with particle recombination. One quite straightforward solution involves a ionization chamber multiple gaps. An example of double gap chamber is represented in Figure 2.13 where the picture shows a detector realized from a collaboration between INFN and the Turin University.

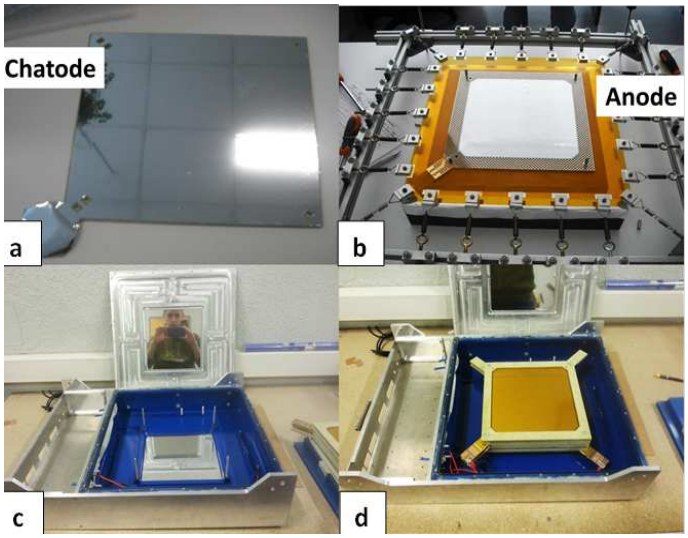

Fig. 2.13 Double gap chamber realized by INFN and University of Turin. a) cathode, made of aluminum deposited over a thin layer of mylar; b) anode, made in aluminum deposited over a layer of kapton; c) empty chamber box; d) the assembled double gap ionization chamber.

The two gaps have been designed with different electrodes distance in order to have different charge recombination effects. Since the chambers have different collection efficiencies, a procedure based on the data collected in a pair of chambers was developed which allows to correct for the recombination and to determine the total ionization charge.
When pulsed beams with intensities of 10^{12} ions/s are considered, and the gaps of the chamber are in the order of few millimeters, the detector will work in the recombination region (Figure 2.7); since in order to prevent the discharges, the bias voltage is limited and the released charges cannot be totally collected. Therefore, working in this unsaturated zone, the recombination of the charges becomes non-negligible and it is necessary to correct for it.
In order to simplify the notation, from this point on, Q stands for the charge released into the chamber gas (previously named Q_{sat}) and Q' is used for the charge collected by the electrodes. The collection efficiency is expressed as:

$$f = \frac{Q'}{Q} \tag{2.33}$$

Considering now two gaps with different thickness and different applied voltages, both filled with nitrogen. Assuming that the electrodes are thin and the two gaps are

close enough such that the beam crossing the perturbation of the beam is negligible. Therefore:

$$f_1 = \frac{Q'_1}{Q_1}; f_2 = \frac{Q'_2}{Q_2} \tag{2.34}$$

where Q_1 and Q_2 are the charge created by ionization respectively in the first and second gap. The collected charges Q' are measured with the electronic read-out of each chamber.
The ratio between the collected charges is equal to:

$$\frac{Q'_1}{Q'_2} = \frac{f_1 Q_1}{f_2 Q_2} \tag{2.35}$$

Since the released charges Q_1 and Q_2 are proportional to the distance d between the electrodes,

$$\frac{f_1 Q_1}{f_2 Q_2} = \frac{f_1 d_1}{f_2 d_2} = \frac{d_1 \frac{1}{u_1} ln(1+u_2)}{d_2 \frac{1}{u_2} ln(1+u_2)} \tag{2.36}$$

$$u_1 = \frac{\alpha/e}{k_- + k_+} \cdot \frac{\rho_1 {d_1}^2}{V_1} = \rho_1 u_1 \frac{{d_1}^2}{V_1} \tag{2.37}$$

$$u_2 = \frac{\alpha/e}{k_- + k_+} \cdot \frac{\rho_1 {d_2}^2}{V_2} = \rho_2 u_2 \frac{{d_2}^2}{V_2} \tag{2.38}$$

Where μ_1 and μ_2 are constant values.

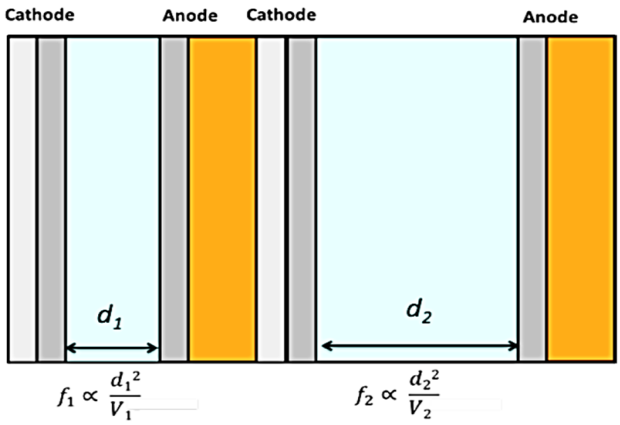

Fig. 2.14 Double gap chamber: conceptual representation.

If the same beam passes through both chambers, $u_1\rho_1 = u_2\rho_2$ and u_2 can be expressed in terms of u_1 as:

$$u_2 = u_1 \frac{V_1}{V_2} \frac{{d_2}^2}{{d_1}^2} \tag{2.39}$$

$$\frac{{Q_1}'}{{Q_2}'} = \frac{f_1 Q_1}{f_2 Q_2} = \frac{f_1 d_1}{f_2 d_2} = \frac{d_1}{d_2} \cdot \frac{ln(1+u_1)u_1}{ln\left(1+u_1\frac{V_1 {d_2}^2}{V_2 {d_1}^2}\right)u_1} \cdot \frac{V2{d_2}^2}{V1{d_1}^2} \tag{2.40}$$

$$\frac{V_1}{V_2} = a; \frac{d_1}{d_2} = b \tag{2.41}$$

therefore

$$\frac{{Q'}_1}{{Q'}_2} = \frac{a}{b} \frac{ln(1+u_1}{ln\left(1+u_1\frac{a}{b^2}\right)} \tag{2.42}$$

$$\left(1+u_1\frac{a}{b^2}\right)^{\frac{{Q'}_1}{{Q'}_2}} = (1+u_1)^{a/b} \tag{2.43}$$

$$u_1 = \left(1+u_1\frac{a}{b^2}\right)^{\frac{{Q'}_1{}^b}{{Q'}_2{}^a}} - 1 \tag{2.44}$$

This is equation in the form $u = f(u)$ can be solved for u using an iterative method based on linear approximations.

Let u_0 be an estimate of u_1 and let $u_1 = u_0 + h$. The number h is the true root of the equation u_1 , and $h = u_1 - u_0$ quantifies the discrepancy from the truth. A linear approximation can be adopted (h small):

$$0 = f(u_1) = f(u_0 + h) \approx f(u_0) + hf'(u_0) \tag{2.45}$$

thus, unless $f'(u_0)$ is close to 0

$$h \approx -\frac{f(u_0)}{f'(u_0)} \tag{2.46}$$

where u_n is the current estimate and u_{n+1} the following one with f' the first derivative of f. Once obtained u, it is possible to calculate the collection efficiency for the first gap:

$$f_0 = \frac{1}{u_1} ln(1u_1) \tag{2.47}$$

and the same procedure can be adopted to calculate f_2.

$$u_1 = u_2 \frac{V_2 {d_1}^2}{V_1 {d_2}^2} \tag{2.48}$$

$$u_2 = \left(1 + u_2 \frac{b^2}{a}\right)^{\frac{{Q_2}'a}{{Q_1}'b}} - 1 \tag{2.49}$$

$$f_2 = \frac{1}{u_2} ln(1 + u_2) \tag{2.50}$$

Once the collection efficiency f is determined for both the chambers it is possible to determine the charge Q released into the chamber.
The just described method cannot be used on-line for reasons of computational speed; therefore, with the aim to estimate the value of f_1 and f_2 , we plan to use the following graphical method. Following the development of the above equation, the ratio f_2/f_1 can be expressed as:

$$\frac{f_2 Q_2}{f_1 Q_1} = \frac{f_2 d_2}{f_1 d_1} \rightarrow \frac{Q_2}{Q_1} = \frac{d_2}{d_1} \tag{2.51}$$

and

$$\frac{f_2}{f_1} = \frac{Q_2' Q_1}{Q_1' Q_2} = \frac{Q_2' d_1}{Q_1' d_2} \tag{2.52}$$

knowing the distances d between the two gaps, it is possible to measure the collected charges Q_1' and Q_2'. Then, with these values, changing the initial ion density n_0 (i.e. changing the beam intensity), it is possible to tabulate analytically a series of values of f_2/f_1 . The plot of f_2/f_1 vs n_0 is a monotonic function and for each value of f_2/f_1, f_1 and f_2 are uniquely determined, for a specific n_0 , as represented in Figure 2.15.

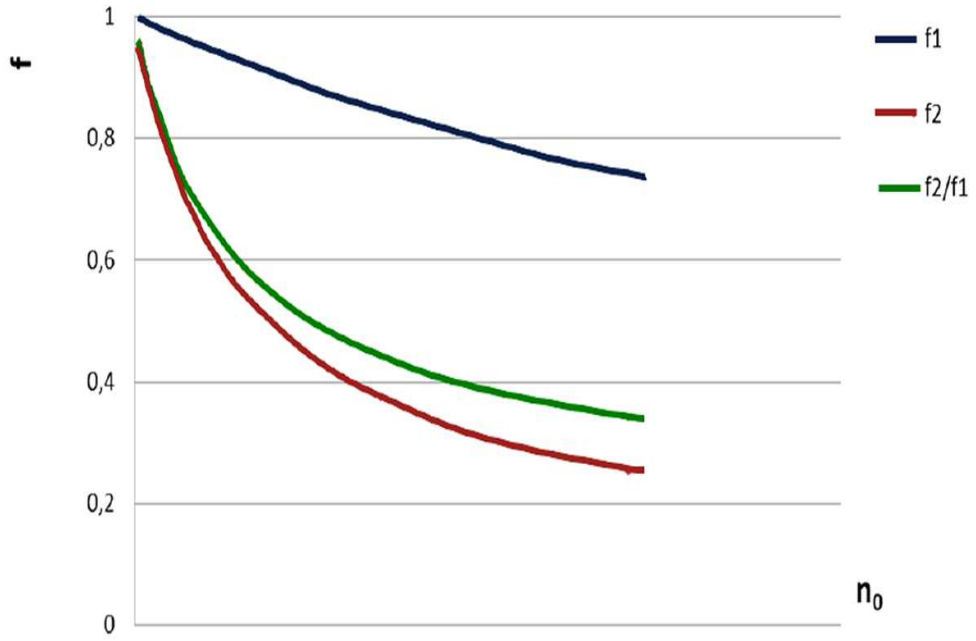

Fig. 2.15 Example of the graph used for the collection efficiency estimation.

2.6 The TERA09 ASIC

2.6.1 Motivations

The TERA09 ASIC is a current-to-frequency converter device, designed as an improved version of its predecessor, TERA08. This previous version features 64

channels operating in parallel, each accepting input currents of both polarities and implementing a 32 bits counter with up/down counting capability [15]. TERA08 operates with a clock frequency of 100 MHz and a maximum conversion frequency of 20 MHz. With a charge quantum of 200 fC the maximum current that a channel can convert without saturation is about 4 μA. Even if in conventional particle therapy facilities the adopted currents are in the order of hundreds of nano amperes, this limit can be an issue for the pulsed beam structure provided by new systems and by the whole planned next generation of accelerators where, with the aim of reducing the complexity and increasing the performance of the machines, new accelerating technologies are exploited [16] [17]. Short beam pulses of 1-10 μs duration with a repetition rate of 1 kHz or less will replace the almost constant beam flux used in the present clinical facilities, leading to an effective beam duty cycle two to three orders of magnitude smaller. Therefore, to achieve a similar dose, the beam flux in each pulse has to increase accordingly. The new design project started after a preliminary test performed increasing the current range with TERA08. A discrete upper-board adapter applied over the ASIC test board, allowed to evenly split the input current of a detector element into several readout channels and then adding up the counts of these channels to reconstruct the input current. It was shown [18] that this method is suitable to increase the maximum input current up to 64 times, preserving the good linearity achieved with the individual channels and with a limited increase in the standard deviation of the measurement. Nevertheless, there are drawbacks with this method; one is the lower number of detector elements which can be read out with a chip, the second is the necessity of reading out the values of a large number of counters, up to 64, and perform their sum. Both affect the versatility of the chip and strongly limit the range of application of TERA08. In addition, a 64 increase in the dynamic range could not be sufficient for the target application. This is the reason why TERA09 has been designed.

2.6.2 TERA09 circuit architecture

Figure 2.16 shows the architecture of the TERA09 chip. The ASIC contains 64 identical channels equipped with a current-to-frequency converter, described in detail later, followed by a 32-bit counter. Currents of both polarities can be converted, leading to increments or decrements of the counters depending on the current polarity. The readout of the counters can be done independently from the operations of the

converters. By asserting a common external latch signal the content of all the counters are loaded simultaneously in 32-bit registers. This operation does not stop the activity of the counter, thus there is no dead time due to the readout. An integrated system of adders triggered by the latch signal provides the sum of groups of 4, 16 and 64 channels.

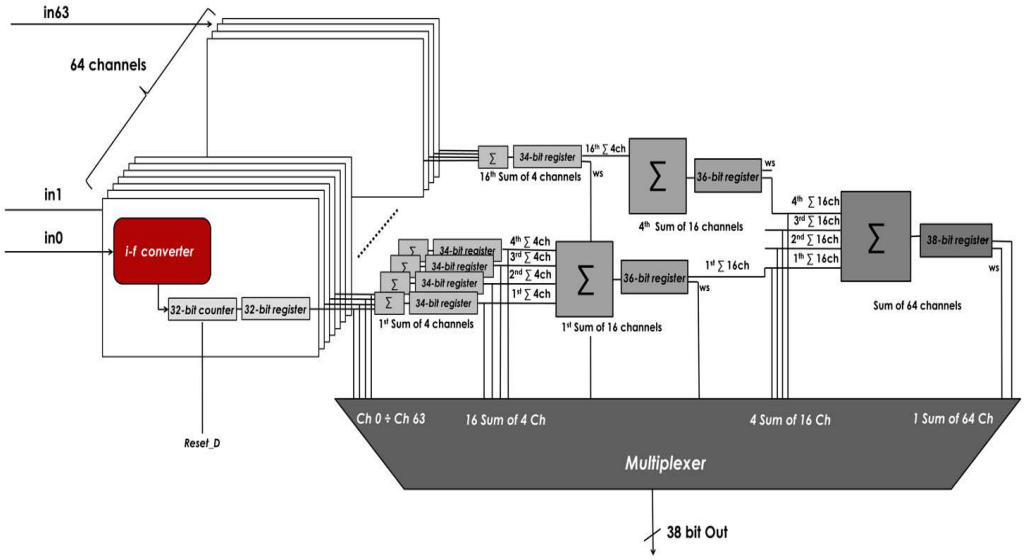

Fig. 2.16 TERA09: functional blocks representation.

These values are stored in additional 34-, 36-, and 38-bit wide registers. Any of these registers can be addressed via seven digital Channel Select lines and read out on a 38-bit output bus through a multiplexer. It is possible to directly read the sum of the counters of 4,16 or 64 channels if, in order to increase the dynamic range, the input current is split among the channels, as explained in the previous section.
Particular care has been taken to prevent the overflows of the registers to corrupt the corresponding sums. Indeed, the converter was designed to operate at the maximum conversion frequency of 80 MHz; as an example, in this conversion frequency limit condition the counter 32-bit capacity will be exceeded, and the counter will reset to the starting value, approximately every 50 s. Such a condition is easily identified and corrected for if each individual channel is acquired separately, but it can be more difficult to identify and correct when only the sum of a large number of channels is acquired. To identify in advance possible flips of individual counters, an output warning signal was implemented which is set whenever any of the 64 counters

exceeds half of its capacity (see Figure 2.16). When such condition occurs, the asynchronous digital $reset_D$ can be used to zero all the counters soon after the latch signal.
The converter of TERA09 is based on the charge recycling technique and is implemented as shown in Figure 2.17.

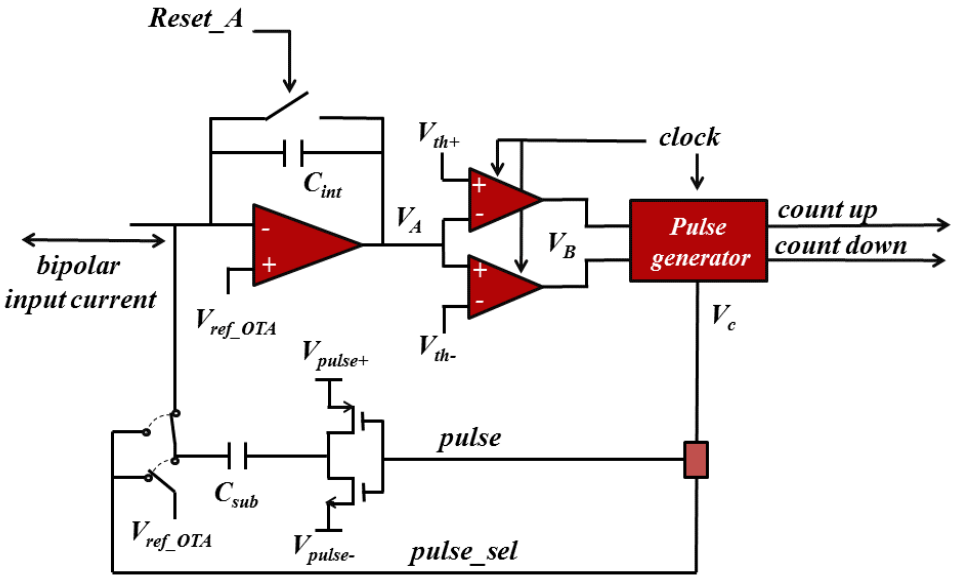

Fig. 2.17 Block diagram of the TERA09 electrical current-to-pulse converter.

The input current is integrated over a 1.2 pF capacitor C_{int} via a folded-cascode operational transconductance amplifier (OTA), described in the following subsection. The output voltage V_A increases when the current exits from the chip (negative current) and vice versa. This voltage is compared with two fixed thresholds, V_{th_+} and V_th_- , by two synchronous comparators. Whenever the comparator input voltage crosses the threshold, the corresponding comparator sets a logic level 1 at one of the input V_B of the pulse generator (PG). When one of its inputs goes high, the PG generates a pulse V_C with a duration of 2 clock cycles that sends a current pulse with polarity opposite to the input current to discharge of the capacitor C_{int}. The pulse adds or subtracts a fixed amount of charge Q_c, depending on the outputs of the comparators; this results in a change of voltage across V_A given by Q_C/C_{int} . In parallel, the PG sends an increment or a decrement signal to the counter. The waveforms of these signals are shown in Figure 2.18 for the case of a steady negative current.

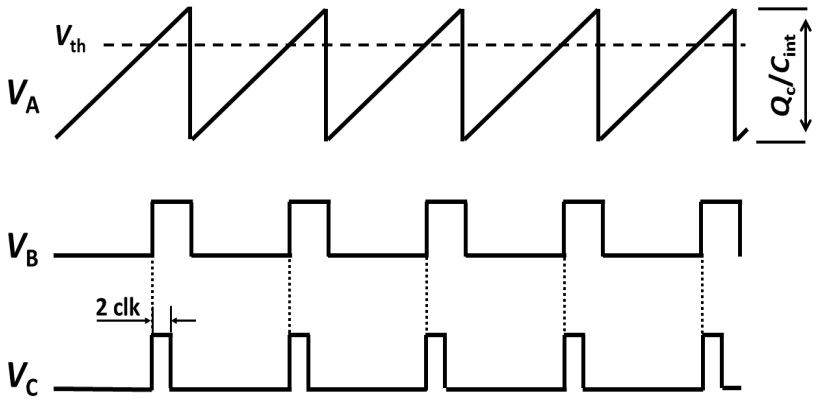

Fig. 2.18 Voltage waveforms at the output of the charge integrator (V_A), the comparator (V_B) and the pulse generator (V_C) for a constant negative input current.

The circuit for the discharge of the integrating capacitor is shown in Figure 2.19.

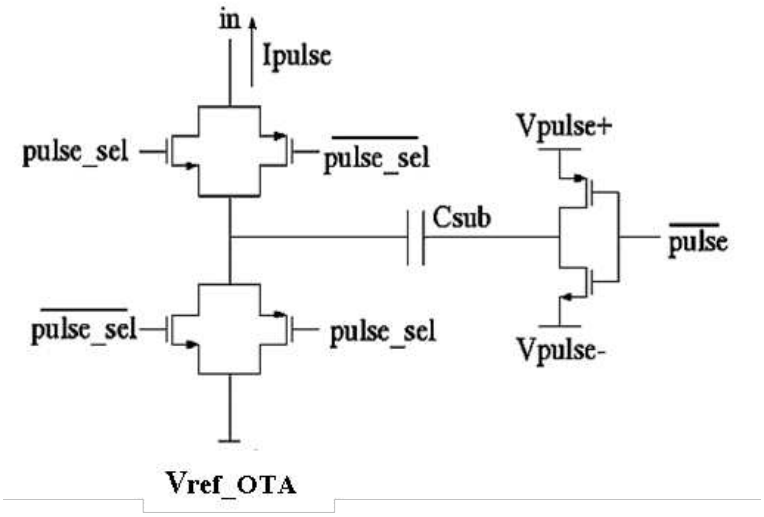

Fig. 2.19 Diagram of the circuit for the discharge of the integrating capacitor.

The circuit for the discharge of the integrating capacitor is shown in Figure 2.19. The pulse V_c is split in two pulse signals, pulse and $pulse_{sel}$, which are delayed by one clock cycle. The first signal is used to send a voltage pulse to a 200 fF capacitor C_{sub}, whose amplitude is given by the difference $\Delta V = (V_{pulse+} - V_{pulse-})$ which can be selected in the range between 0.25 and 3.3 V. At the edges of the pulse, the

capacitor generates two δ-like current pulses of equal absolute value and opposite polarity, each corresponding to a charge given by

$$Q_c = C_{sub} \cdot \Delta V \tag{2.53}$$

Depending on the value set for ΔV , the charge quantum can therefore be selected in the range between 50 fC and 660 fC. One of the two current pulses is used to add either a positive or a negative charge quantum to C_{int} while the other is discharged toward the reference voltage of the OTA (V_{refOTA}). The selection is achieved by acting on the two CMOS switches via the $pulse_{sel}$ signal. As an example, the waveforms shown in Figure 2.55 lead to a positive charge quantum being added to C_{int}. It should be noted that the voltage across the switches is always equal to the OTA reference, thus limiting the leak of charge through the switch. Another factor which could limit the resolution of the circuit is the charge injected by the parasitic capacitors of the switches. This charge is minimized by the use of CMOS switches and by choosing the minimum size for the transistors.

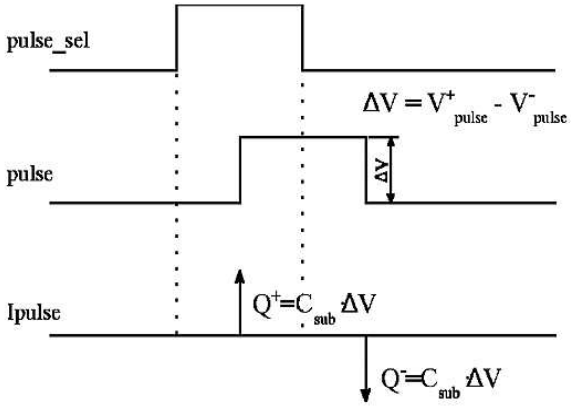

Fig. 2.20 Example of voltage waveforms of the two pulses used by the circuit for the discharge of the integrating capacitor.

Finally, the integration capacitors of all channels can also be fully discharged via the $Reset_A$ common digital input. The total charge collected at the input of the channel in a given time interval is given by the number of pulses generated by the *PG* in the same time interval, measured as the increment or the decrement of the

counter, multiplied by the value of the charge quantum Q_c. For an average input current I_{in} , the average output frequency f of the converter is given by

$$f = \frac{I_{in}}{Q_c} \tag{2.54}$$

All the operations described above are synchronized to an external master clock and supervised via a Moore-style finite state machine. Since four clock cycles are required to perform the C_{int} discharge sequence, the maximum conversion frequency is $1/4$ of the master clock frequency. The chip has been designed to operate at a maximum clock frequency of 320 MHz; in this condition, the maximum output frequency of the converter is thus 80 MHz, an increase of a factor of four compared to the predecessor.
Forty chips TERA09 were produced in the CMOS 350 nm technology by AMS taking advantage of a Multi Project Wafer organized by Europractice [19]. The chips, whose size is $4.68 \cdot 5.8\ mm^2$ (layout shown in Figure 2.21), were encapsulated in an MQFP 160 package prior to delivery. In the following, the results of the tests performed on the chips are reported.

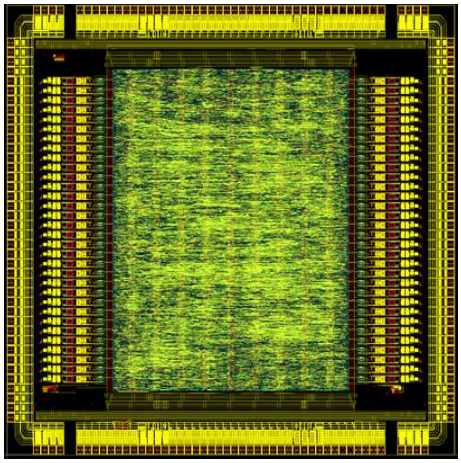

Fig. 2.21 Layout top view of the TERA09 ASIC.

2.6.3 TERA09 OTA: design informations

As mentioned TERA09 works with bipolar input currents. For this reason its amplifier has been designed to integrate the input signal, providing as output a voltage ramp which has a positive angular coefficient if the input current is negative and vice

versa. The adopted amplifier type is the OTA.
Considered stand-alone (i.e. without the capacitive feedback adopted to work as an integrator), a conventional Operational Transconductance Amplifier (OTA) is classified as a class A amplifier that produces an output current from an input voltage difference, through a transconductance gain parameter:

$$G_m = \frac{i_0}{V_i} \tag{2.55}$$

the maximum output current is equal to the applied bias current. Ideally the input and output resistances are infinite therefore no current is crossing the resistances depicted in Figure 2.22. The open circuit gain is $A = G_m R_o$.

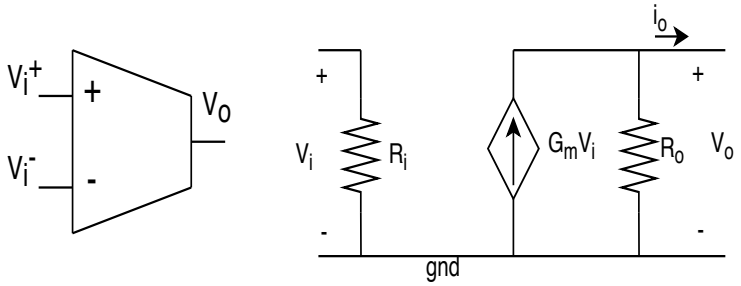

Fig. 2.22 OTA conventional symbol (left) and equivalent circuit representatin (right).

A single stage OTA presents a cascoded differential input pair and also the load is configured with a cascode. The DC gain results from the parallel of the input and output cascode equivalent resistances, multiplied by the input transistor g_m. $GBW = g_{mPM1}/C_l$ is the gain-bandwidth product. C_l is the load capacitance.
The core amplifier is based on the folded cascode topology. As shown in 2.23, the folded cascode architecture is obtained when the common gate transistor and the common source one are complementary types. The advantage in using a folded cascode configuration is the increased output swing and a Common Mode Rejection Ratio, whenever a differential architecture is adopted.

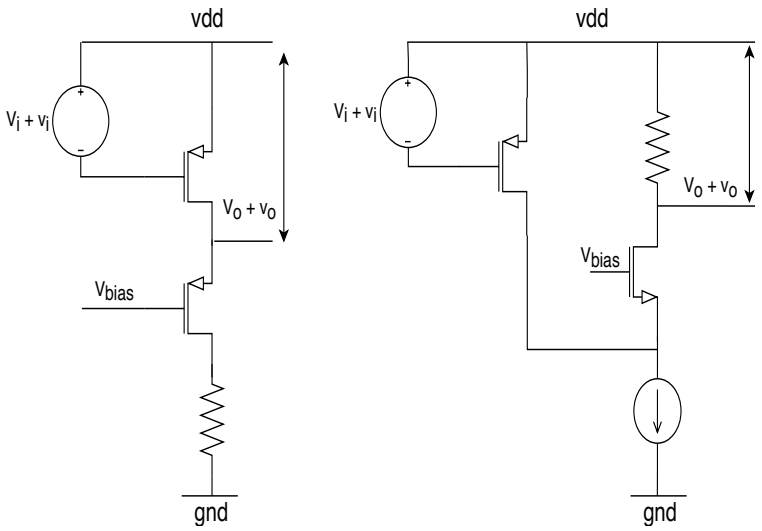

Fig. 2.23 Schematic representation of the conventional cascode scheme (left) and the folded cascode one (right). In both the cases the current in the MOSFET drain of the transistor closer to vdd is directly transferred to the second transistor.

In Figure 2.24 the schematic representation of the TERA09 folded cascode OTA is depicted.

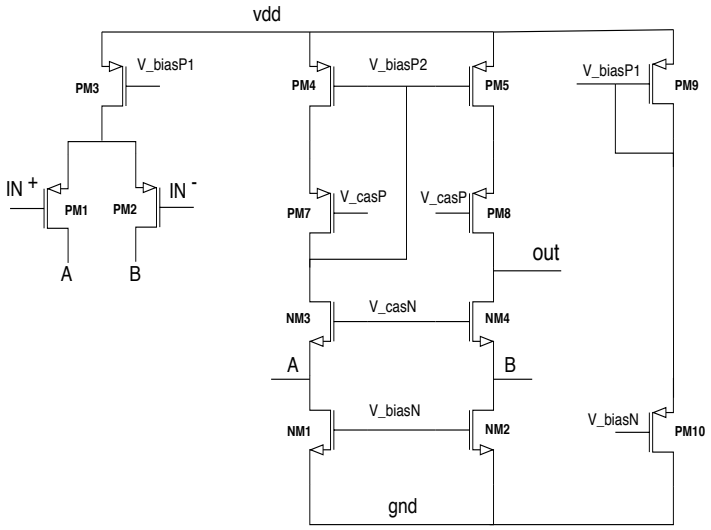

Fig. 2.24 Schematic representation of the TERA09 Folded Cascode amplifier.

The current crossing $NM1$ and $NM2$ in an equilibrium state, is twice the one in the differential input pair and the resulting excess moving in the diode connected branch is downwards mirrored by $PM5$ sunk by $NM2$. Reducing the gate voltage of

$PM2$ will reduce the current in $PM1$, decreasing the current through $NM4$ increases the one in $PM5 - PM7$; I($PM7$) - I($NM4$) is sourced to the output load. If the voltage on $PM1$ gate is lower than the one in $PM2$, the situation is going to be specular. In this architecture, the bandwidth has a second pole limitation depending on the ratio between the PM4 transistor transconductance and the equivalent parasitic capacitance experienced from the PM4 gate.

Is hereafter reported a table showing the TERA09 OTA simulation results, considering the temperature operation range and model corners. Gain, Gain · Bandwidth product and margin phase values have been obtained for 0, 27 and 85 Celsius degrees and changing among the typical models (tm), worst case for speed (ws) and worst case for power (wp).

The folded cascode OTA is feedback in the TERA09 channel, with a 1.2 pF capacitance (C_f) which is twice the value implemented in its predecessor, the TERA08 chip. This increased capacitance improves the front-end stability at high rates, up to 250 MHz. The gain reduction descending from an higher C_f is compensated with a proper transistor sizing thus to increase the input transistor transconductance and sustain an increased bias current overall.

Table 2.2 Process corner analysis results. Settings: V_{ref} = 1.65 V; I_{bias} = 800 μA; C_l = 1 pF. The data have been obtained simulating the schematic view.

Model library	**T [° C]**	**vdd [V]**	**G [dB]**	**GBW [MHz]**	**Phase margin [○]**
tm	0	2.97	40.1	636.7	49.2
	27	2.97	39.3	569.8	48.2
	85	2.97	38.3	534.0	47.6
	0	3.3	40.6	780.7	47
	27	3.3	39.8	658.0	48.2
	85	3.3	38.5	507.2	49.5
	0	3.63	41.0	727.4	51.6
	27	3.63	40.1	736.3	47.8
	85	3.63	38.8	552.9	49.4
ws	0	2.97	41.7	488.2	47.7
	27	2.97	40.9	428.4	47.9
	85	2.97	39.6	323.5	48.4
	0	3.3	42.2	571.4	44.7
	27	3.3	41.4	508.2	45.6
	85	3.3	40	365.3	49.7
	0	3.63	42.5	655.5	46.5
	27	3.63	41.6	593.8	49.2
	85	3.63	40.3	428.4	48.3
wp	0	2.97	39.1	995.9	48.3
	27	2.97	38.3	885.1	48.1
	85	2.97	37.2	669..4	49.5
	0	3.3	39.6	1127.6	47.6
	27	3.3	38.7	996.4	47.3
	85	3.3	37.4	773.1	48.3
	0	3.63	40.1	1294.8	46
	27	3.63	39.1	1061.1	47.8
	85	3.63	37.7	842.9	48.1

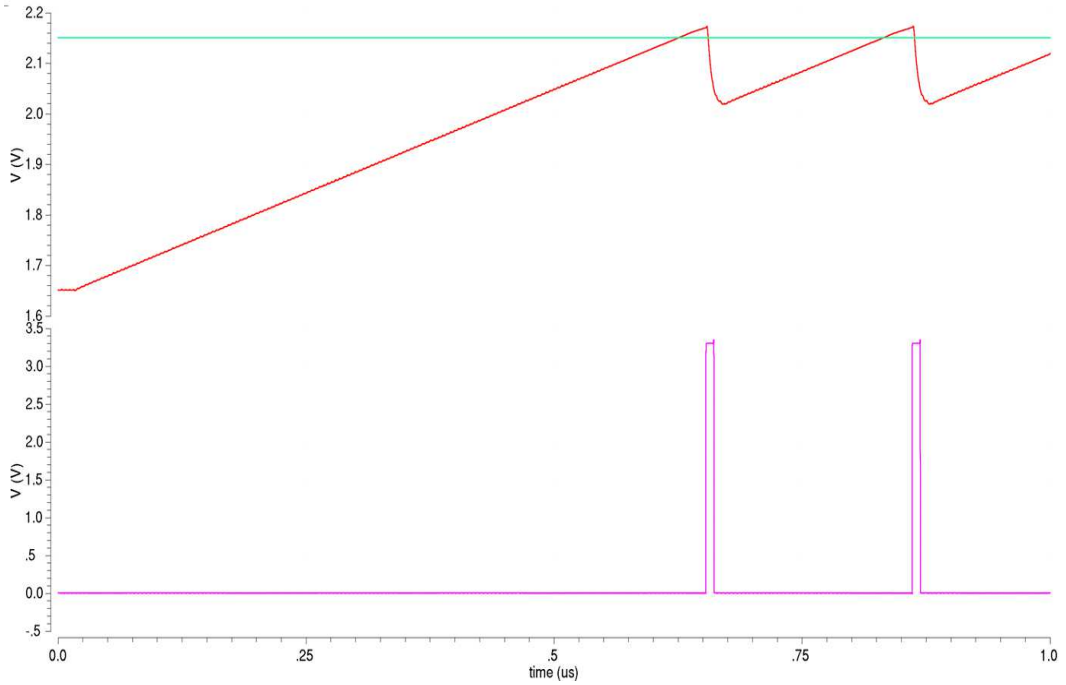

Fig. 2.25 Schematic simulation result showing the recycling integrator architecture in action. The TERA09 channel is here fed with a 1μA input current and the system clock is set at 250 MHz. The red curve is the OTA output voltage ramp which is brought under threshold (in green), subtracting a 200 fC charge quantum. The digitized output is reported in the lower part of the picture, in violet.

Figures 2.25, 2.26 shows the result of schematic simulations for the TERA09 channel with a 1 μA and 10 μA input current respectively. The system clock was set at 250 MHz therefore the recycling integrator with its pulse subtraction FSM works up to 62.5 MHz ($\frac{clockperiod}{4}$).

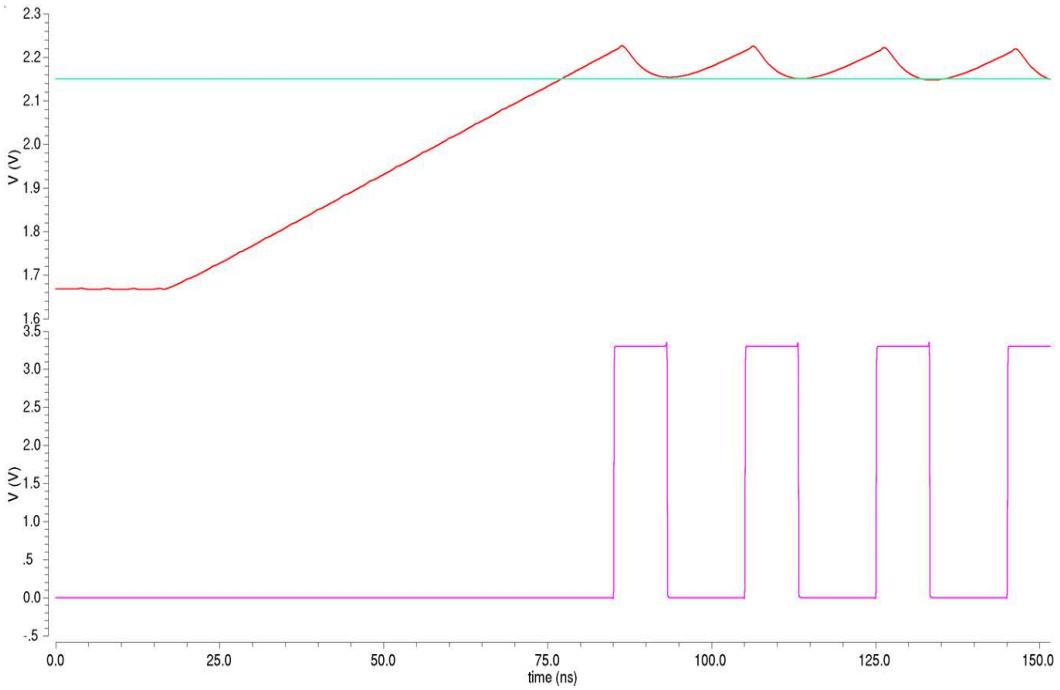

Fig. 2.26 Schematic simulation result showing the recycling integrator architecture in action. The TERA09 channel is here feeded with a 10μA input current and the system clock is set at 250 MHz. The red curve is the OTA output voltage ramp which is brought under threshold (in green), subtracting a 200 fC quantum charge. The digitized output is reported in the lower part of the picture, in violet.

2.7 The ASIC characterization

2.7.1 Experimental setup

The experimental setup used to characterize the ASICs was based on a National Instrument PXI chassis with a 7813R FPGA board, interfaced to the host PC using the LabVIEW FPGA software toolkit [20]. In order to manage a multi-ASICs test routine, a dedicated PCB has been realized implementing a socket that allows to plug-in the chip under study and plug-it-off after the procedure (more details are reported in the following subsection).

For some of the tests, a voltage generator was used to inject a precise steady current, either into a single channel or evenly distributed in parallel to all the 64 channels. In the first case, the source was connected to the input via a 10 MΩ series resistor. The configuration for the parallel connection is represented in Figure 2.27.

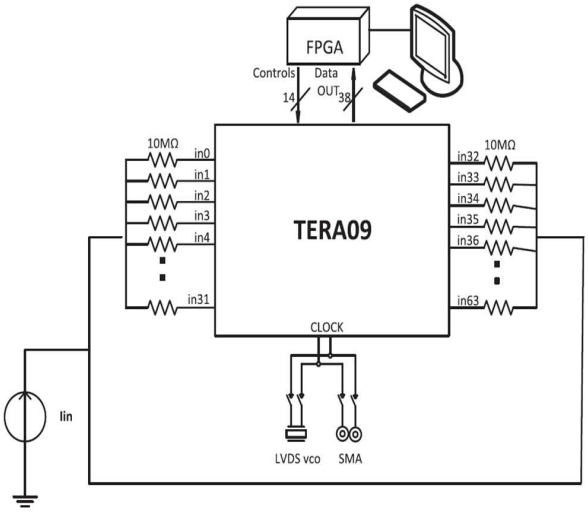

Fig. 2.27 Schematic representation of the acquisition setup where 64 channels are connected in parallel to the current source.

The connection of groups of channels to a common input of the TERA08 ASIC has been investigated in [18]. From those test results has been observed that this kind of signal splitting through a parallel connection of the channel is only possible through resistors of 10 MΩ or more, high enough to limit to a negligible value the offset currents which originate from the small voltage offsets of the input stages of the channels. In this case, the input impedance seen by the generator amounts to 10/64 MΩ. A simple fan-in board with SMD resistors was prepared for this purpose (see Figure 2.28).

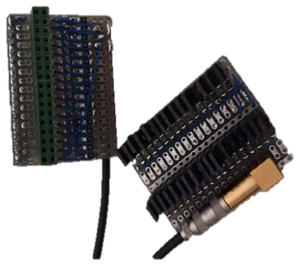

Fig. 2.28 Fan-in boards with the surface mounted resistors for the TERA09 channel parallel connection. The connector adapted to the PCB inputs is visible in green, on the left side of the photo; on the right, it is possible to appreciate the SMD resistors and the jumpers with black hats which are used to physically close (or not) the electric paths.

The Keithley 2400 voltage generator was used, which provides a precise voltage source in the range between 1 mV and 211 V but is current-limited at high voltages. For higher currents, we used a Bertan 323 model HV power supply, which can provide a voltage source up to 3 kV.

2.7.2 The TERA09 test board

As previously mentioned, a custom test board has been developed for testing the TERA09 ASICs. The PCB is a six-layer system with the following organization:

two layers dedicated to grounding

one layer is the 3.3 V CMOS power plan

one layer is used for the OTA voltage reference

the two remaining layers are used for the signals routing with special care for the clock lines.

The PCB mounts a 160 pins MQFP socket (Figure 2.29)that houses the chip in a removable way. Just two PCB tes boards have been realized to test the 40 chips production.

Fig. 2.29 The 160 pins MQFP socket open (left) and close (right). A mechanical pressure of the top cover, guarantee the electrical contact between the ASIC pins and the socket ones.

A further crucial aspect for the PCB design has been the choice to implement the possibility to provide the clock signal either on board through a voltage controlled oscillator (VCO) or from external sources, via SMA connectors. The clock signal

format is LVDS and the operating modality (external or internal clock source) can be selected by soldering two zero-ohm resistors as bridge connection.

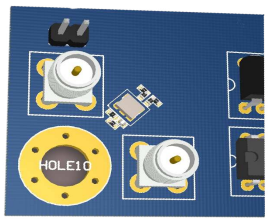

Fig. 2.30 3D model detail for the VCO-SMA clock source splitting. The SMAs are represented with the withe round components and the oscillator with LVDS output is placed between them.

The PCB embeds the circuits for adapting the 3.3 V CMOS digital levels to the 5V TTL levels of the DAQ board and all the voltages needed to operate the chip. Finally, two connectors allowed to inject a current separately in each input channel of the chip (two connectors are needed because TERA09 has 32 channels on one edge and the other 32 channels in the opposite edge).
Figure 2.31 shows two views of the test board design and Figure 2.32 is a rendering of the test board.

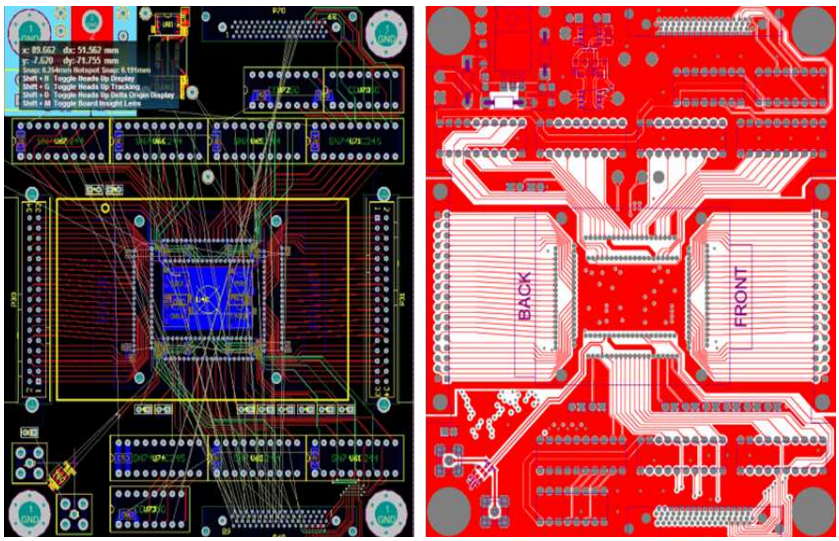

Fig. 2.31 Test board design views. Floor plan view (left). Gerber file of the top layer (right).

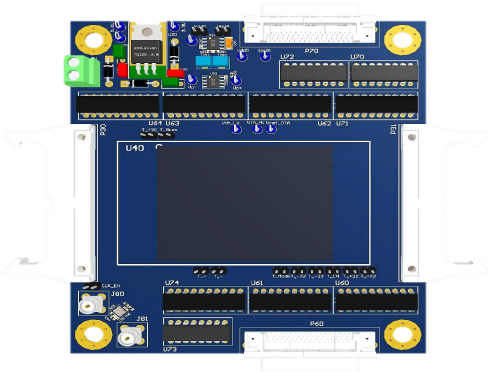

Fig. 2.32 3D model or rendering of the TERA09 test board. The chip inputs are routed through flat cables on the lateral edges (represented in white). The upper part has the voltage references and the voltage regulator for the 5 V to 3.3 V levels; in the left-high corner, the connector for the external voltage supply is represented in green. In the left-lower corner of the picture, there is the VCO-SMA clock source splitting path.

The TERA09 chip has a 32-bit register for each of the 64 channels, 16 registers 34-bits deep for the sum of four channels, 4 registers 36-bits deep for the sum of sixteen channels and one 38-bits register. Therefore eighty-five register are readout with a given frequency, by the DAQ software. The TERA09 output is the result of a register selection through a control multiplexer, operated with the following addressing logic:

8 bits covers 85 register possibilities ($2^8 = 128$); b_0 is the LSB, b_7 is the MSB.

b_7 is the chip select: it is used to enable or disable the chip data exchange, though for a multiple ASICs system (e.g. the TERA09 front-end board described in the following sections).

b_6 separates the selection from channel level to channel-sum level. When $b_6 = 0$, bits from b_0 to b_5 select the 64 channels in the ordinary way (00000 = channel 0, 00001 = channel 1 and so on). With $b_6 = 1$, the situation is slightly more complex.

With $b_6 = 1$, the bits b_0 and b_1 are responsible for the sum level selection: if $b_0 = 0$ then the bits form b_2 to b_5 are used to chose the first level sums i.e. the sum of four channels (for example, excluding the b_7, 10000x0 = first sum of four registers from channel zero to channel three. The x character means that

it does not care its value. 10001x0 = second sum of the first level thus the sum of the channel from four to seven included).

If $b_0 = 1$ then, if $b_1 = 0$, the selection is confined to the second level sums, the sixteen channels sum, selected through the b_5 and b_4 (100xx01 selects the first sum of sixteen channels, from channel zero to channel fifteen included).

If $b_0 = 1$ and $b_1 = 1$ the selection target is unequivocally the register containing the sum of all the sixty-four channels.

Resuming: b_7 = chip select b_6 = single channel or register sum selection; the string $1b_5b_4b_3b_2x0$ is used to select the sums of four registers; the string $1b_5b_4xx01$ is used to select the sums of sixteen registers; the string $1b_5b_4xx01$ is used to select the sums of sixty-four registers.

The bit configuration for the TERA09 register addressing is resumed in Figure 2.33

Bit configuration					selection
$b_7 = 0$					the chip is not selected
$b_7 = 1$					chip selected
	$b7 = 1; b_6 = 0$				channel 0-63 selection
		10 000000			channel 0
		10 000001			channel 1
		…			…
		10 111111			channel 64
	$b7 = 1; b_6 = 1$				
		$b_0 = 0$			
			11 00000 x0		ch0-ch3 sum
			11 00001 x0		ch4-ch7 sum
			…		…
			11 11111 x0		ch12-ch15 sum
		$b_0 = 1$			
			$b_1 = 0$		
				11 0000 01	ch0-ch15 sum
				…	…
				11 1111 01	ch47-ch63 sum
			$b_1 = 0$		ch0-ch63 sum

Fig. 2.33 TERA09 register addressing bits.

Figure 2.34 highlights the test board connections while the data flow and the interconnections between the ASIC and the FPGA is depicted in Figure 2.35. The FPGA sends the controls through the high-density connector $VHDCI_1$ (Figure 2.36):

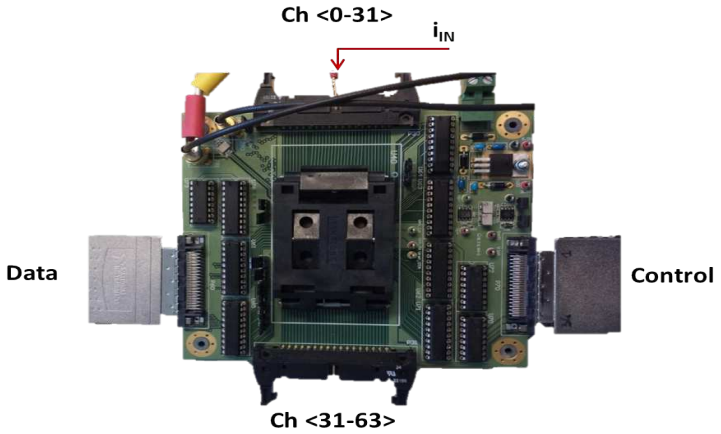

Fig. 2.34 TERA09 test board with the connector detail: two flat connectors for the 64 channel inputs (top and bottom), two VHDCI connectors for data and control management.

$bias\_mode_0$ and $bias\_mode_1$ controls the bias current of the OTA, in order to decrease the power consumption when the chip is locked to a lower frequency.

the *load* signal stores at the same time the counters content to the registers, and it occurs with a frequency selected in the DAQ (usually 1 - 10 kHz).

rst_synch is the synchronous reset of the digital and analog information.

rst_D is the digital reset used to empty the digital registers.

rst_A is the analog reset used to force the OTA feedback capacitance discharge. This signal is generally used at the beginning of the data acquisition in order to start without charge accumulated in the feedback capacitance.

$b_7 - b_0$ are used to select which register content is going to be reported as output.

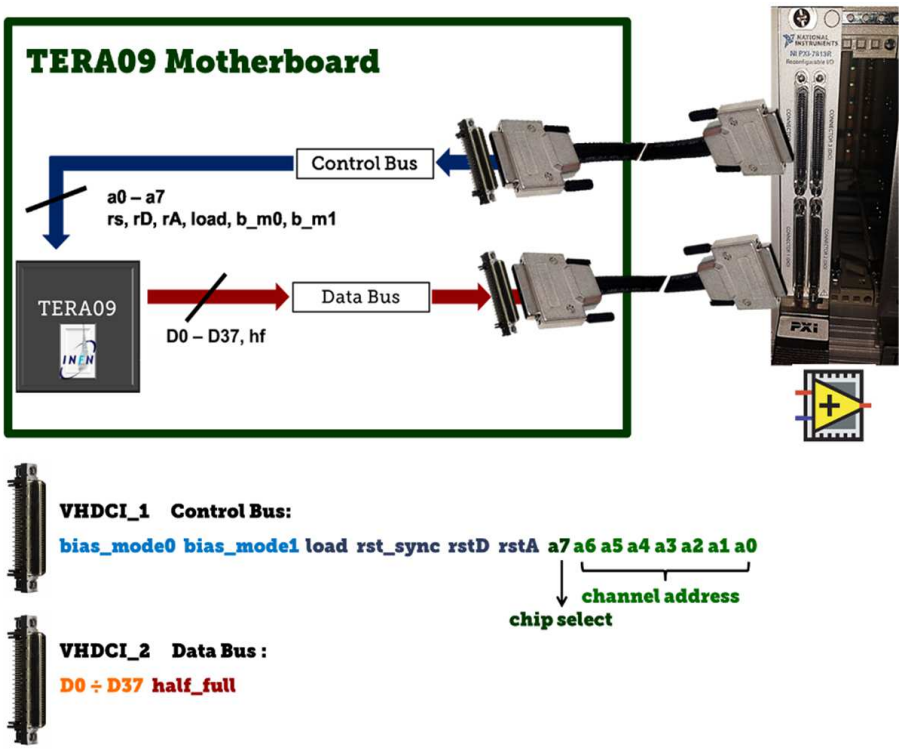

Fig. 2.35 Schematic representation of the ASIC-FPGA data exchange. The connection between the TERA09 test board (named Motherboard, in the picture) and the NI chassis are realized using VHDCI connectors.

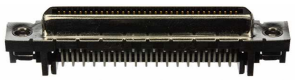

Fig. 2.36 High density, 68 poles VHDCI connector.

The $VHDCI_2$ is used for the ASIC output data: $D0 - D37$ are the data 38 bits. The bit $half\_full$ rises to 1 whenever the sum of the 64 channels register reaches half of this counting width (the MSB flips from 0 to 1); in this way, the user has time to deal with incoming register overflow avoiding data loss.

2.7.3 Characterization results

TERA09 has been designed to achieve a good linearity over a large input current range, therefore the tests were mainly addressed to verify this feature. At first, this section describes the performance of the individual channels in terms of gain

uniformity, background current and the linearity range. Afterward, the linearity over the extended input range is obtained with a parallel connection of the inputs.

The basic functionality of the chips was checked by injecting a steady current in individual channels selected randomly and by measuring the corresponding number of counts. This check was repeated several times increasing the clock frequency. It was not possible to reach the maximum value of 320 MHz because the chip becomes unstable at frequencies above 280 MHz. The problem was investigated in detail and was found to be related to the cross-talk between digital signals of the chips and the clock signal on the test board. The results reported in this paper were obtained with a 250 MHz clock, corresponding to a maximum counter increment frequency of 62.5 MHz, and should be considered conservative in terms of the measured dynamic range. All the chips were found to be working and providing a similar number of counts. One of them was selected randomly for the additional tests reported in the following.

As a preliminary step, for each value of the charge quantum Q_C used in the tests, a calibration procedure was adopted. The external reference voltages, common to all the channels, were set acting on trimmers provided on the test board. The charge quantum of each channel was then derived, applying equation (2), by injecting channel by channel a steady current of 1 μA and measuring the corresponding counter frequency. The value, averaged over all the channels, was found to deviate by a few percents from the expected value, well within the tolerance of the capacitance values of Csub which, for the used technology, is specified to be $\pm 10\%$. The small deviation of the chip average charge quantum was corrected by adjusting the external reference voltages.

Figure 2.37 shows the deviation of the charge quantum from the average as a function of the input channel number for a Q_C average value of 200 fC. This figure indicates that the gain uniformity across the channels of the chip is around 3%. A structure with a certain symmetry can be noticed, where the trend observed in channels 0 to 31 appears to be similar to that observed in channels 32 to 63. This similarity may originate from the geometric arrangement of the channels in the chip, where the two groups of channels are positioned symmetrically along two opposite edges. Similar results are obtained at different values of Q_C, as expected considering

that the relative deviation of the charge quantum is mainly proportional to the relative deviation of the capacitance of C_{sub}.

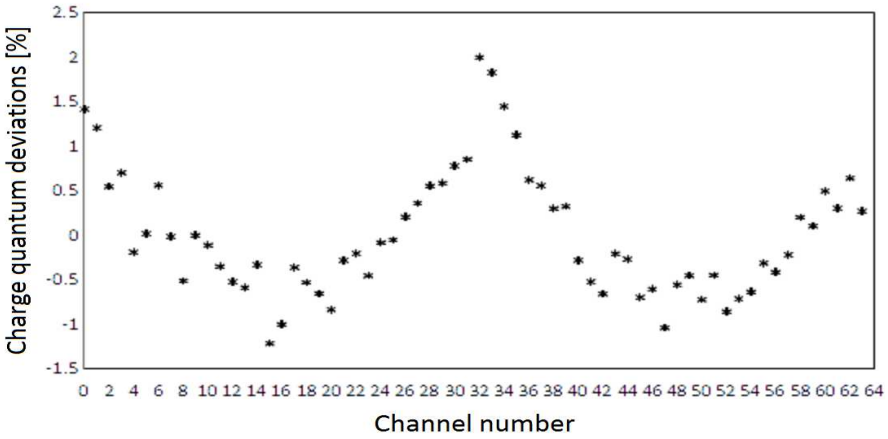

Fig. 2.37 Charge quantum relative deviations as a function of the input channel, obtained from 1μA input current measurements. Channels are numbered from 0 to 63

The background current was measured by acquiring data with the inputs of the chip unconnected. The measurement was repeated for charge quantum values ranging from 50 fC to 600 fC in 50 fC steps. The results, shown in Figure 2.38, are expressed in terms of counts per second, where the following rule was adopted: the number of counts is assigned a positive sign when the counter is incrementing (i.e. for negative input currents) and a negative sign when the counter is decrementing (i.e. for positive input currents).

As expected, the result scales approximately with the charge quantum value. For a value of 200 fC, a typical choice for beam monitoring in particle therapy [21], an average background of about -10 counts per second is observed, corresponding to a positive pedestal current of 2 pA, almost two orders of magnitude smaller than the minimum currents measured in clinical applications. For the analysis presented in the following, the charge quantum was fixed to 200 fC, and the corresponding values of the pedestal of each input channel were measured and subtracted from the data in all the measurements.

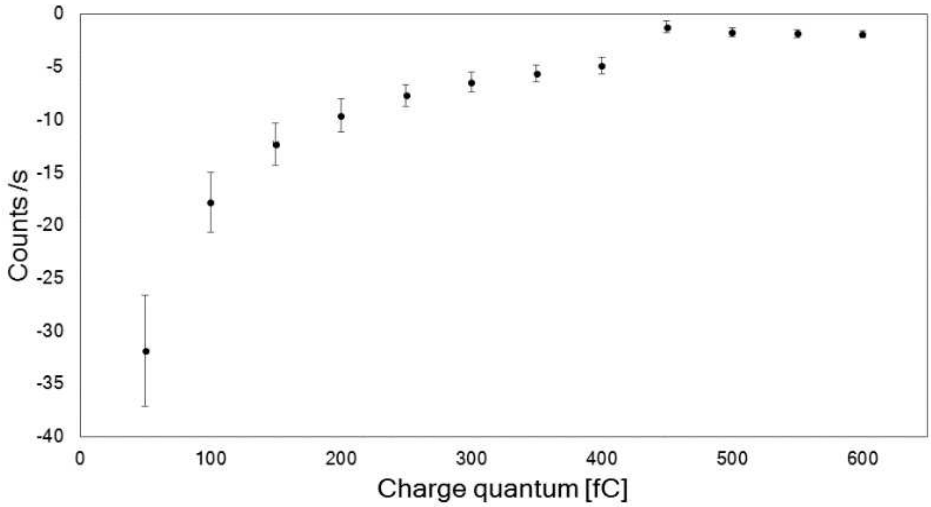

Fig. 2.38 Average background frequency as a function of the charge quantum Qc. The vertical bar indicates the standard deviation of the 64 channels of the chip.

Figure 2.39 shows the value of the counter frequency as a function of the value of the input current for a typical channel of the chip, where positive and negative currents are shown separately.

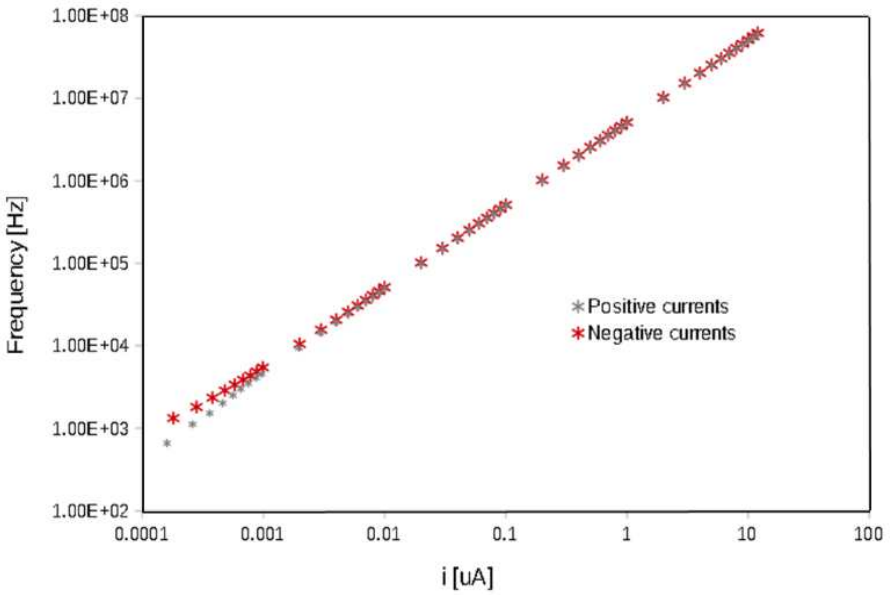

Fig. 2.39 Count frequency as a function of the absolute value of the input current for a single channel for a 200 fC charge quantum. Positive and negative currents are shown separately.

The measurement was performed for an input current ranging from $\pm$80 pA to $\pm$12 μA and the results were fitted to a line separately for positive and negative currents. The relative deviations from linearity as a function of the input current are reported in Figure 2.40.

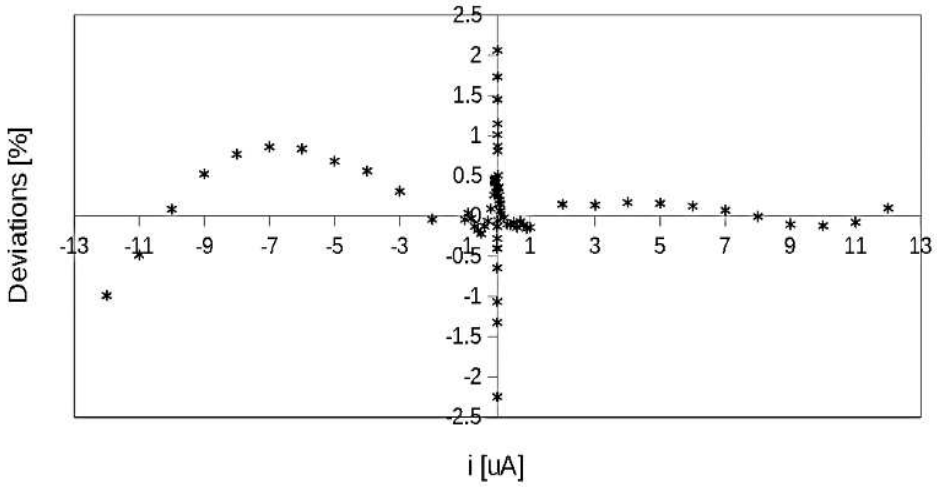

Fig. 2.40 Relative deviation from linearity as a function of the input current for the measurements reported in Figure 9.

A good linearity is observed with deviations well within $\pm$1% for most of the range except for low absolute values of the input current, where it is found that the deviation exceeds 2% below 3 nA. For comparison, a maximum deviation from linearity of 1.5% was reported in the range 500 pA to 3 μA for the chip TERA08 [22]. As anticipated, TERA09 extends considerably the input current range compared to its predecessor. The worse linearity, compared to TERA08, at very small input currents was unexpected and may originate from leakage currents in the test board used for the measurements. The compact front-end printed circuit board for TERA09, currently under development, will be designed to overcome this problem. As explained in the previous sections, by reducing the number of detector elements that can be served by a chip, the dynamic range can be further extended by splitting evenly the current through several channels of the chip and by reading out the sum of the corresponding counters. For example, reducing to 16 detector elements, the dynamic range is fourfold increased. All the results which follow were aimed at probing the performance with such arrangement and were obtained connecting all the 64 channels of the chip to the same current source.

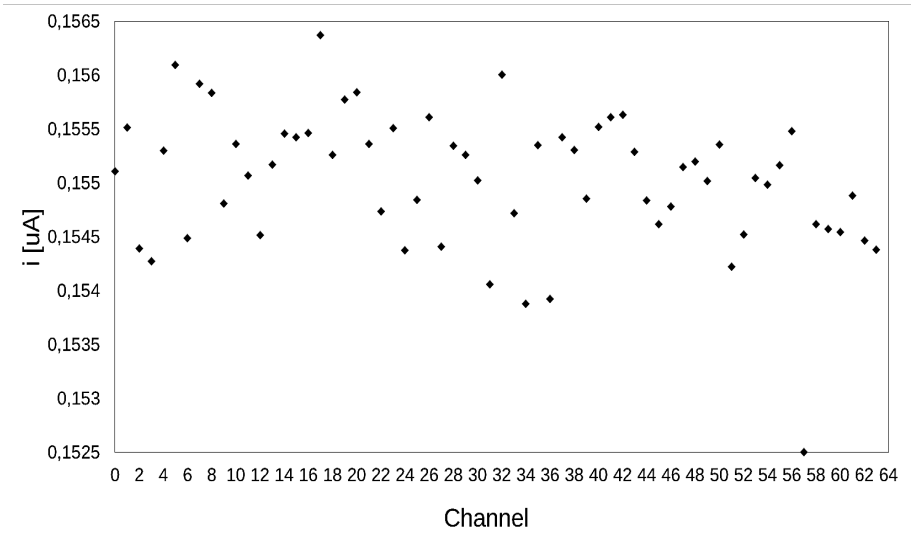

Fig. 2.41 Distribution of a 10 μA input current among the 64 channels of the chip. Channels are numbered from 0 to 63.

To convert the frequency of counts of each channel into a current, the equation $f = i_{[}in]/Q_C$ was used correcting for the channel-by-channel variations of the gain shown in Figure 2.37. The results indicate that the current is uniform across the channels with a maximum deviation of $\pm 1\%$, compatible with the tolerance of the 10 MΩ resistors connected to each input. The linearity obtained with this arrangement was tested, using a charge quantum of 200 fC, in a range of currents between $\pm 10\mu$A and ± 750 μA, yielding to the results reported in Figure 2.42.

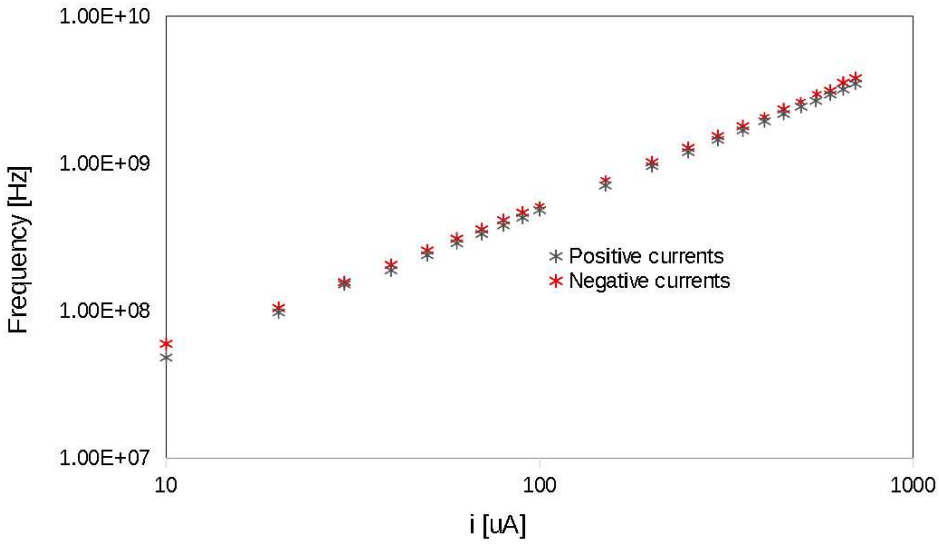

Fig. 2.42 Absolute value of the count frequency as a function of the absolute value of the input current for the sum of 64 channels, as described in the text. Positive and negative currents are shown separately.

The corresponding deviation from linearity, obtained with the same method as for the single input channel, is shown in Figure 2.43.

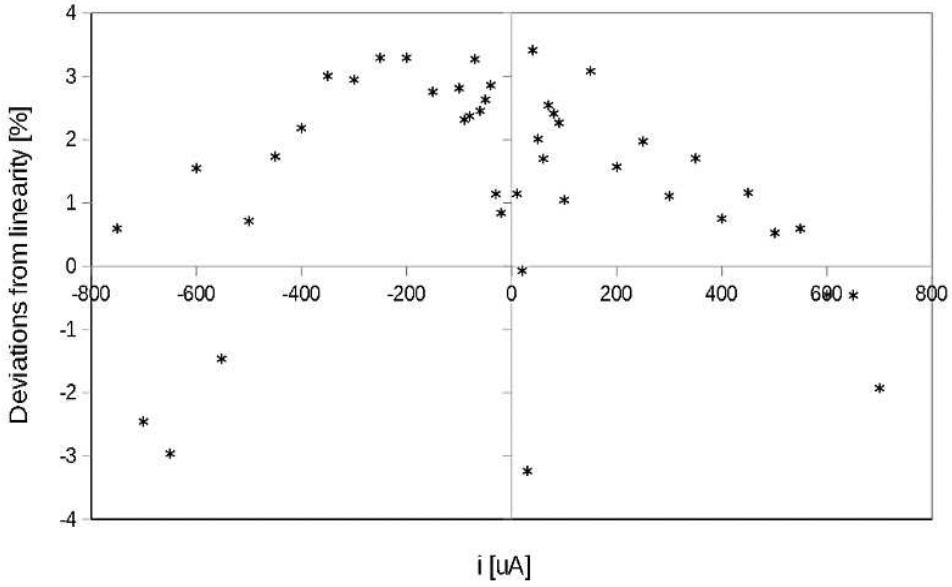

Fig. 2.43 Relative deviation from linearity as a function of the input current for the sum of 64 channels reported in Fig. 12.

It is found that a linearity better than $\pm 3\%$ can be achieved in a range of currents between 20 μA and 750 μA for both current polarities.

2.8 Radiation damage test

2.8.1 Introduction

The TERA09 chip is designed to be located aside of the monitoring chambers, far from the therapeutic beam, and no protection from data corruption from single events was implemented in its design. However, considering the relatively large area of the chip covered by data registers and the secondary neutrons field produced during the irradiation, the potential exposure to data corruption by Single Event Effect phenomena need to be addressed. With this aim TERA09 has been directly exposed to various ion fluxes with different LET, to study the upset rate as a function of the energy deposited by single events. From the analysis of the data, it is possible to predict the single event effect cross-section in a clinical environment and estimate the readout failure probability in a real application scenario.
The effects of the total ionizing dose on the performance of the TERA chips were studied and reported in previous publications [23]. In this thesis section the investigation of radiation-induced Single Event Upsets (SEU) on TERA09 is reported. This is particularly relevant considering that the chips were not designed for operating in a harsh radiation environment and no protection from data upsets originating from single events was implemented in the design.
The sensitivity of a device to single events is generally determined by measuring the rate of occurrence of a given effect as a function of the energy deposited, by irradiating the device with ions of different Linear Energy Transfer (LET). The SIRAD (Silicon RAdiation Damage) facility [24], located at the 15 MV Tandem of the Legnaro National Laboratory (LNL) of the INFN, offers the possibility to work with different ion sources, select the beam fluence and the beam incident angle on the Device Under Test (DUT), thus allowing to collect statistics of single event occurrence as a function of the energy deposited by the ions. We focused on the digital part of TERA09, described in the following section, where the incident ion deposited energy may induce bit flip phenomena (i.e. SEU) in a large set of data

registers, leading to data corruption. Analyzing the data from this SEU tests it is possible to determine the TERA09 upset probability in a typical clinical environment.

2.8.2 Basis on Single Event Effects in CMOS

In CMOS technology the reliability of a system is an aspect getting worsening (or at least getting more challenging), with the design detail downscaling. Intuitively, the smaller is the device and more sensitive to parameters variation or system perturbation it will be the device itself. In CMOS technology it is common to refer to Single Event Effects distinguishing among heavy and soft damages. Examples of heavy or permanent damages are the Single Event Burnout that is a destructive effect or the Single Event Latch-up, a short-circuit that can lead to burnout if not mitigated in time. In the soft event category, the Single Event Transient and the Single Event Upset are the most common; the former results in a charge transient caused by a single proton or heavy ion passing through a sensitive node in the circuit whereas the latter results in a bit-flip, thus a logic state change due to energy deposition in a digital cell. Several studies proved that SEU and SEL effects are physically separated in terms of silicon region where they occur. Figure 2.44 shows the fact that SEUs are confined in the first micrometers thickness under the device surface whereas SELs occur deeper in the silicon bulk [26].

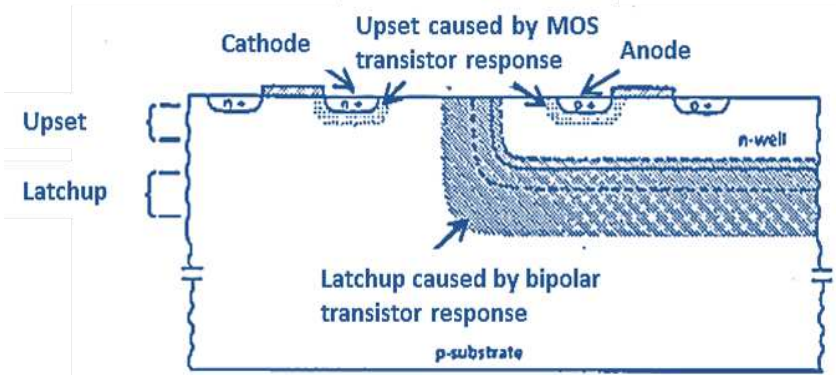

Fig. 2.44 Spatial distribution of different SEE: SEU are confined in the first micrometers thickness, under the device surface whereas the SEL are deeper in the silicon bulk.

As mentioned before, a Single Event Latch-up occurs in the deep volume of the silicon bulk where, in a CMOS process, the combination of n-well, p-well, and substrate forms a parasitic n-p-n-p structure named thyristor. During a latch-up, both

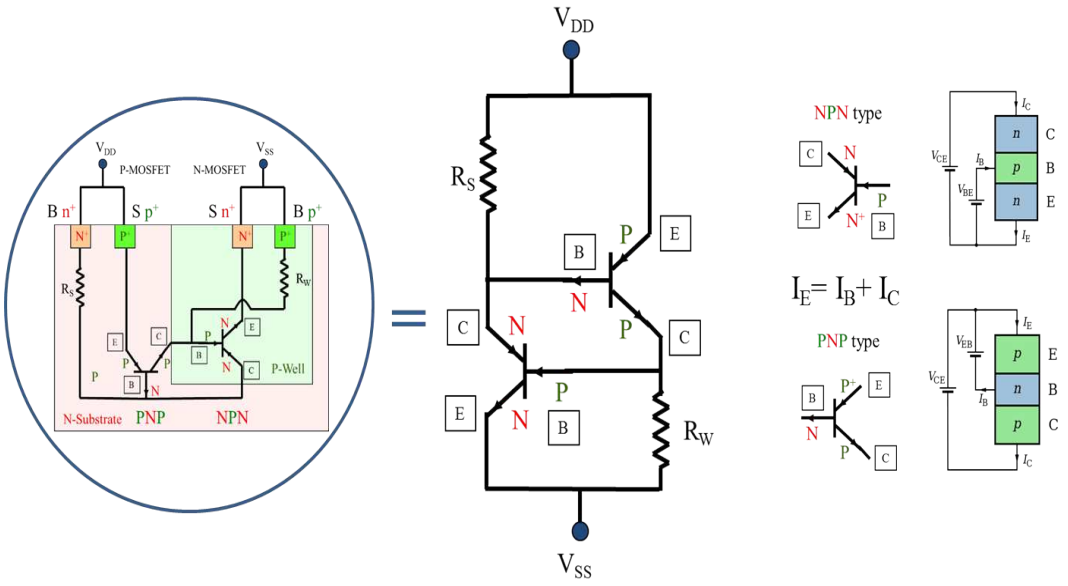

Fig. 2.46 Schematic representation of the thyristor-like path, responsible of SEL. From left to right: the CMOS n-p-n-p structure, the electric circuit and the BJT connections.

MOSFETs are conducting resulting in a power-ground short circuit. The permanent and destructive event is avoided turning off the power supply. Figure 2.45represent this kind of path which details are shown in Figure 2.46.

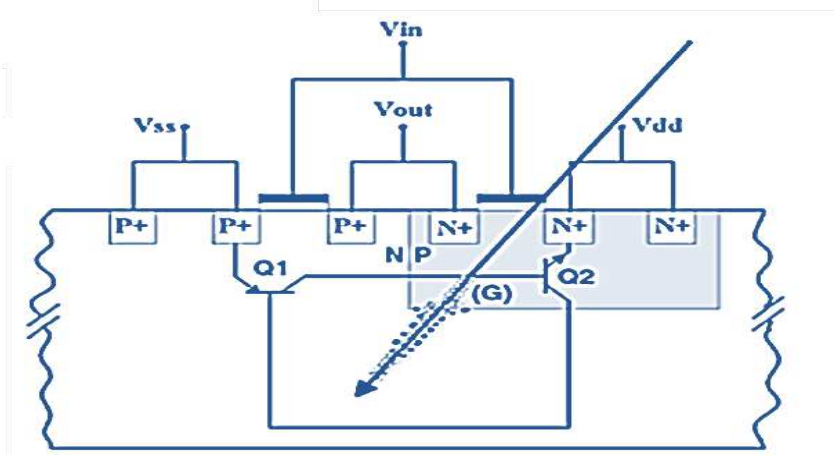

Fig. 2.45 The Single Event Lachup occurs in the deep volume of the silicon bulk. In a CMOS process, the combination of n-well, p-well and, substrate forms a parasitic n-p-n-p structure named thyristor. During a latch-up, both the MOSFETs are conducting resulting in a vdd-gnd short circuit. The permanent and destructive event can be avoided turning off the power supply.

Considering the TERA09 chip and its application, the study has been focused on the SEU phenomena. High LET particles increase the SEU probability in a CMOS circuit and this is related to the released energy in the crossed medium and the related ionization density. Hitting the silicon bulk, these ions create ion-holes pairs and their collection at the source/drain diffusion regions may result in a p-n junction current pulse, driving a voltage change in the corresponding circuit node [27]. The event occurs whenever in a sensitive node the charge injected by the current pulse exceeds a certain critical value Q_{crit}. It is possible to model the SEE with the following equation:

$$V \geq \frac{Q_{crit}}{C} = \frac{1}{C}\int_0^{t_{sw}} i_{ds}dt \tag{2.56}$$

where C is the capacitance of the discharging path and t_{sw} is the time delay between the particle strike and the logic state change (voltage exceeding a certain threshold value). i_{ds} is the drain-source current flowing into the transistor of the SEU interested node. Figure 2.47 represents an example of a logic state switching occurring in a CMOS inverter.

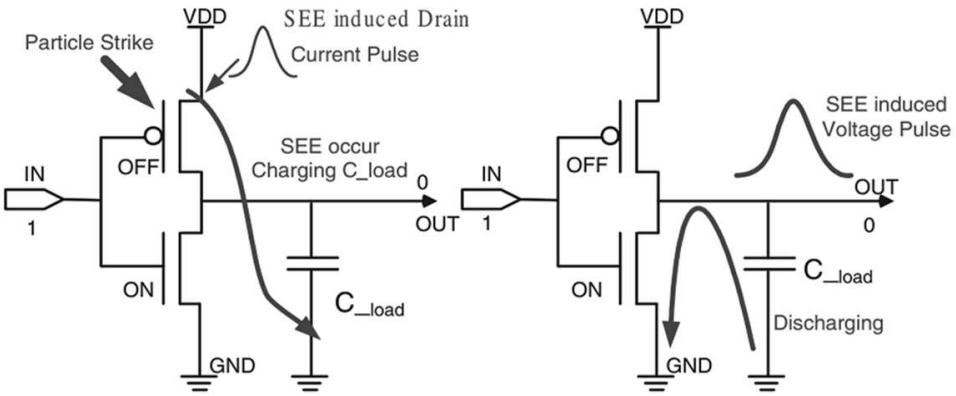

Fig. 2.47 Schematic representation of an inverter bit-flip due to a SEE.

2.8.3 SEU test setup

As just mentioned, the SEU is a not a permanent damage of the chip circuitry and it results as a bit-flip originated by a high energy deposition by a single track in a small sensitive volume located into the digital circuitry. For single event effect studies, the

particle fluence rate must be carefully selected to be low enough to distinguish the effects caused by the impacts of single ions and high enough to observe a significant number of single effects in the measurement time. Typical ion fluxes are in a range from 10^3 to 10^5 ions/cm^2s [24]. The experimental setup used to test the TERA09 for SEU consisted in a PCB test board with a socket to carry and properly bias the ASIC, interconnecting it to the data acquisition system (DAQ), based on a 7813R FPGA board, interfaced to the host PC using the LabVIEW FPGA software toolkit. Figure 2.48 shows the cable connections through the vacuum irradiation chamber. An external clock source provided the 250 MHz LVDS clock whereas the PCB 5 V bias was supplied by a voltage generator that was current-limited to prevent latch-up triggered burnout. A program was prepared which, after issuing an initial latch signal to load the values of the counters in the chip registers, reads out at the desired frequency the chip registers detecting and displaying any bit flip. The contents of the registers were also saved in a file for off-line analysis.

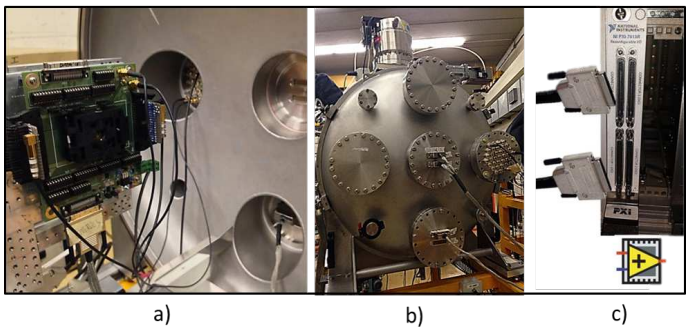

Fig. 2.48 a): The DUT thus the TERA09 ASIC carried by the PCB socket, fixed inside the vacuum chamber on the metal plate holder.
b): Vacuum chamber external flange connections.
c): VHDCI connectors carrying the ASIC data and controls to the National Instrument FPGA.

Figure 2.49 shows the GUI of the DAQ created for the TERA09 SEU test.

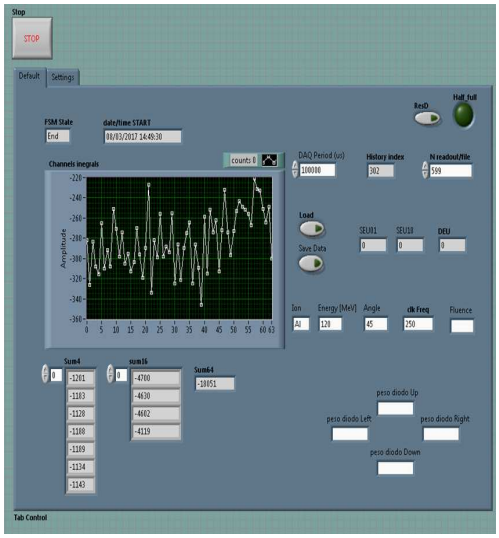

Fig. 2.49 SEU DAQ graphical user interface. The left panel reports the number of counts for the 64 channels: after a first starting phase of current injection, the load signal (counters to registers latch) is turned off and the SEU test starts.

A Keithley 2400 voltage generator was used to provide a steady current to the 64 inputs of the chip in order to let the counters increment rapidly after the power-up of the chip. This was necessary considering that upsets leading to a 0-1 bit and 1-0 transitions could occur with different probabilities and we wanted to measure the upset rate in a condition where zeros and 1 are uniformly distributed in the register cells. Figure 2.50 shows the test board with the current splitting among the ASIC channels.

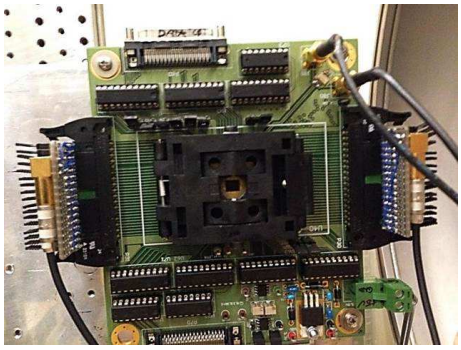

Fig. 2.50 SEU test setup: TERA09 test board with the current injection. The input current is equally split among all the 64 channels, thanks to custom upper boards with SMD resistors. An input current is needed to set the register values at random values before freezing them and then starting with the SEU test.

The PCB board with TERA09 has mounted on a metal plate holder inside a vacuum chamber placed on the beam line (2.51).

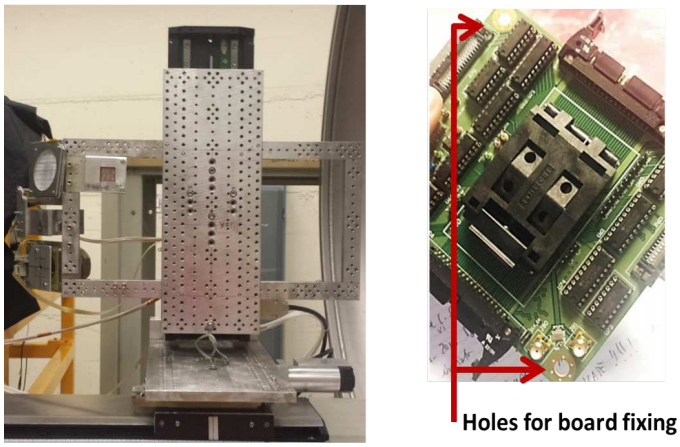

Fig. 2.51 Steel plate with 3-4 mm of diameter holes for the DUT holding (left). TERA09 test board with the holes in the corners for the mechanical fixing over the metal holder (right).

The holder allows to retract the DUT during the setup of the accelerator and to align it in front of the beam for the measurement (Figure 2.52).

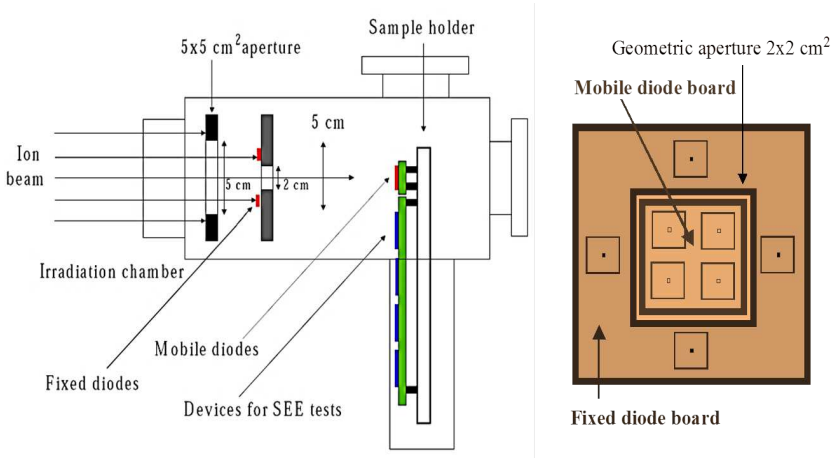

Fig. 2.52 Schematic draw of the irradiation chamber with the dosimetry system and the device under test holder (left). Diodes geometry and placement scheme (right).

A set of silicon diodes in a fixed position in front of the final beam collimator is used to monitor the beam fluence during the irradiation. Before starting the measurements, the DUT is kept in a retracted position and the beam is centered and focused with the aid of a scintillator imaged by a CCD camera. Then, the diodes are

cross-calibrated with a set of silicon diodes which are inserted in the same position covered by the DUT during the tests and then retracted. Figure 2.53 shows the quartz-diode calibration system.

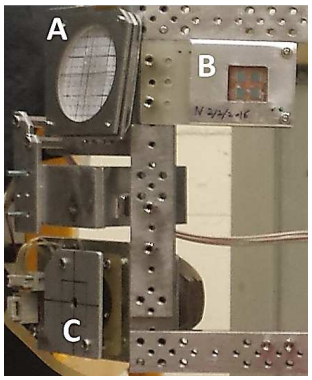

Fig. 2.53 Vacuum chamber inner photo: A) quartz; B) fixed diodes; C) mobile diodes.

The TERA09 chip is packaged in an MQFP 160 pins ceramic structure and hold to the board by means of the plastic socket. In order to expose the 4.68·5.8 mm^2 silicon area of the chip, the ceramic cover of the packaging was removed and a hole was drilled in the socket, as shown in Figure Figure 2.54.

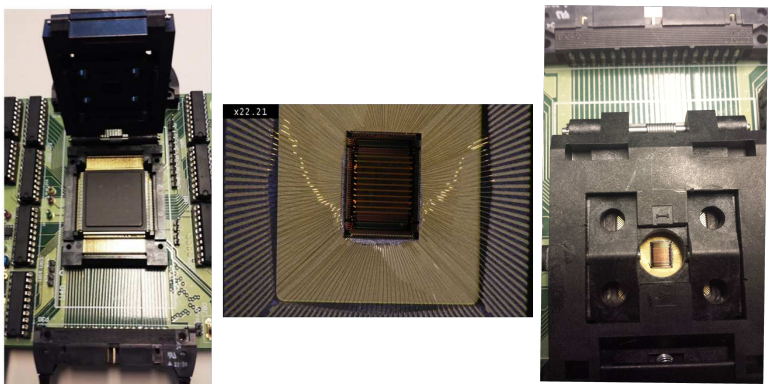

Fig. 2.54 The TERA09 test board with the socket hosting the ASIC open (left). Microscope picture of TERA09, after the ceramic lid removal (center). Hole in the cover lid of the ASIC carrier (right)

The clock cables have SMA connectors that can be directly connected to the external flange of the SIRAD vacuum chamber. All the data and controls lines are twisted pairs (with ground lines) and packaged with custom connectors. A remote

Table 2.3 List of used ions with the respective energy, angle, cross section and deposited energy.

Ion	**Energy [MeV]**	**Angle [°]**	**Edep [MeV]**	σ_{SEU}
F^{19}	122	0	0.94	$2.76E^{-12}$
F^{19}	122	20	1.00	$1.34E^{-11}$
Si^{28}	157	0	2.08	$2.554E^{-09}$
Si^{28}	157	20	2.21	$5.83E^{-09}$
Cl^{35}	171	0	3.07	$1.88E^{-09}$
Cl^{35}	171	15	3.18	$7.10E^{-09}$
Cl^{35}	197	0	2.87	$1.53E^{-08}$
Cl^{35}	197	20	3.06	$1.21E^{-08}$
Br^{79}	241	0	10.12	$3.6E^{-06}$

control desktop, placed outside of the accelerator room, allowed the operator to control the data acquisition and monitor the data on-line during the measurement.

2.8.4 Results and data analysis

As the probability per unit fluence and per bit cell to have a bit-flip in the cell is expressed as the SEU cross section, defined as $\sigma_{SEU} = N_{errors}/\phi N_{bit}$ the probability . The results of the DUT ion irradiation are represented in Figure 2.55 as the SEU cross section (log10 of the cross section) as a function of the deposited ionization energy (the LET-deposited energy conversion has been calculated, as explained in [28]). The experimental data have been fitted with the Weibull function, as suggested in [26]. This function suitably describes a phenomenon characterized by a threshold activation that saturates its trend at a certain value. In the SEU case, the threshold value is the minimum energy required to trigger the event E_0, whereas σ_0 is the saturation level representing the situation in which the upset probability in the sensitive area is 1.

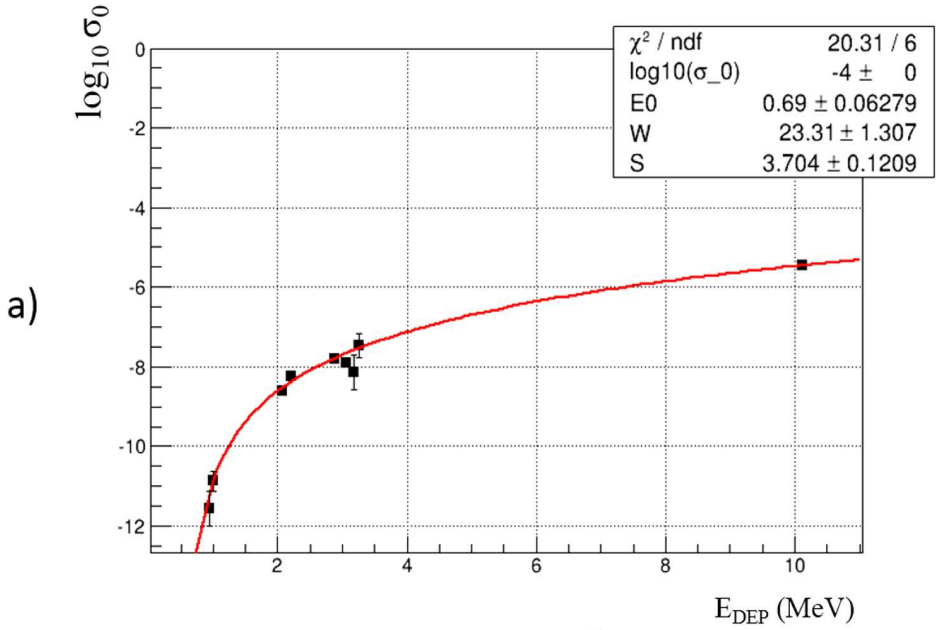

Fig. 2.55 a) SEU cross-section (y-axis) as a function of the deposited energy (x-axis). The fit has been performed using the Least Square Method. b) Weibull function: E_{DEP} is the energy deposited in the CMOS silicon; E_0 is the threshold energy to trigger a SEU event; σ_0 is the cross-section saturation that represent the situation when all the sensitive volumes are under SEU; s and W are fit shape-parameters.

This representation underlines the typical threshold behavior of the SEU effect. It has been important to choose the right ions sets, in order to achieve a reference in the plateau region of the fit and for the minimum energy triggering SEUs.
For ions with a LET larger than Chlorine, the measurement was affected by latch-up events in the silicon bulk. In these cases, the current limitation of the voltage supply avoided destructive consequences due to short-circuiting.

Using a Bromine ions beam, the data acquisition was interrupted by latch-up just after few seconds, thus acquiring very short runs. On the contrary, with ions lighter than Fluorine, no SEU was observed. As previously mentioned, changing the incident angle between beam and DTU allows to slightly increase the deposited energy and adds a second energy-deposited point, for the same ion. Regarding the DAQ, by not asserting the load signal from the acquisition procedure, the expected

content of the registers should not change from one reading step to the following one (the DAQ time was set to 100 ms). At this point, an iterative control over all the 2774 memory bits is performed and every time a bit-flip occurs, a SEU counter is updated.

2.8.5 Expected SEU rate in a clinical room

The predecessors of the TERA09 chip, named TERA06 and TERA08 [29] are currently adopted in hadron therapy facilities worldwide, especially at the National Center for Oncological Hadron therapy (CNAO [2]) in Pavia [9] (Italy), where our group has a historical role of research and technological collaboration since the foundation.
The goal of the ion test is to estimate the SEU rate for TERA09 in a typical application; as a reference, the CNAO treatment room has been considered. In the CNAO facility, a 25 m diameter synchrotron accelerates protons in the energy range 60 MeV-250 MeV and carbon ions at 120 MeV/u-400 MeV/u. In the monitor chamber setup, the chips are not directly exposed to the particle beam thus it is possible to assume that the damaging events are mainly due to the backward produced secondary neutrons. FLUKA Monte Carlo simulations, provided by [25], report that the main contribution that could be relevant for the radiation damage on the readout electronics is related to the secondary neutron backward emitted by the interaction between the 400 MeV/u carbon ions and the target materials (see Table II); in the same document the authors claim that by using carbon ions of 400 MeV/u, the fluence rate of secondary neutrons at the nozzle, where two monitor chambers are installed, is approximately 3.4 10^{10} n/cm^2 per year. From the same publication, the here reported Figure 2.57 shows that the average energy of the backward produced neutrons spectrum is 20 MeV. Figure 2.56 from [26], shows the energy deposition probabilities in a Sensitive Volume of 1x1x1 $1\mu m^3$ (from the adopted method, is the elementary volume that can occur into SEU) for four proton energies. Since neutrons and protons of the same energy produce almost identical upset rates if the incident energy exceeds 20 MeV" [26].

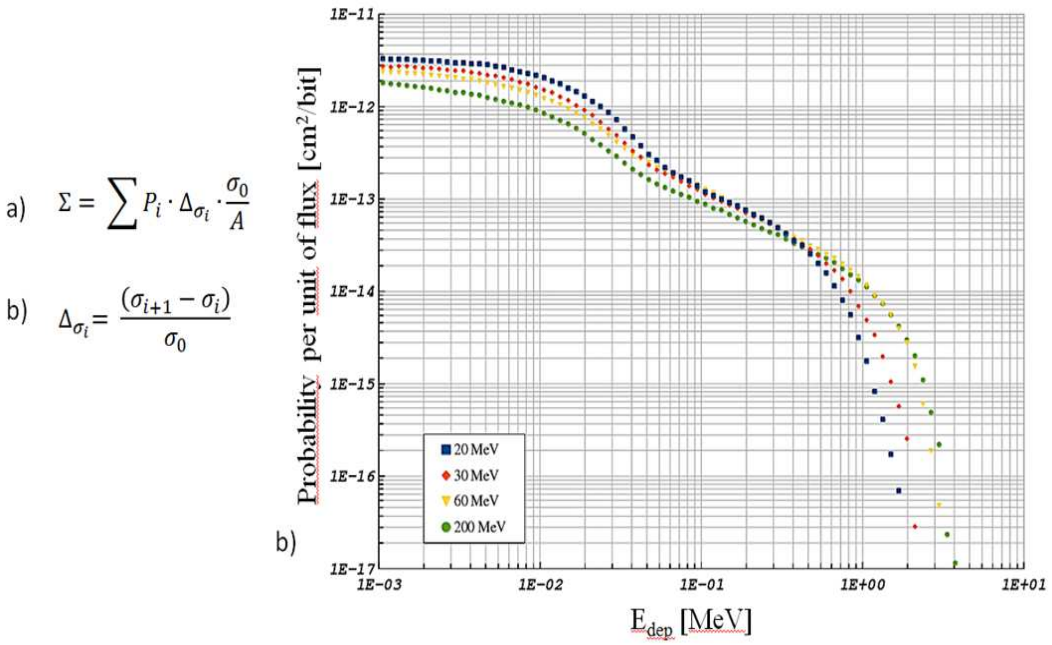

Fig. 2.56 a): SEU cross section Σ in the proton environment,where A is the cross-sectional area of the SV adopted.
b): The increased cross section in the energy interval between two energy bin values i.
c): Energy deposition probabilities for protons of different energies. The curve shows the probability to have within the SV an ionizing deposition greater or equal to the indicated E_{DEP}. The curve selected for the data analysis is the one for 20 MeV protons (due to the average value in Figure 2.57), selecting the plot area from the threshold energy E_0 obtained with experimental data fit (0.69 MeV).

.

Using the equations reported in Figure 2.56 a) and b) and the data from Figure 2.56 c) is it possible to adjust the results related to protons to our experimental data with ions thus to calculate the SEU probability in a neutron environment. From the data analysis, the expected rate of upset for TERA09 is about 70 SEU/year. Albeit this value has to be taken as an order of magnitude due to the method inaccuracy, such upset rate is easily manageable thanks to a redundant detector present on the beam line, like in every monitoring system. The counts number is considered as the true value only if the discrepancy in the measures provided by the main detector and the redundant one is within a certain range, considering the standard detectors variability. Moreover, the probability that both, the main monitor chamber and the redundant one suffer for SEU simultaneously, in the same bit, is close to zero.

Table 2.4 Number of secondary neutrons and protons produced by carbon ion and proton beams on ICRU tissue.

Target	Primary particles	n/primary	p/primary
ICRU tissue	$400 MeV/u$ *carbon ions*	2.66	1.50
ICRU tissue	120*MeV protons*	0.11	0.10

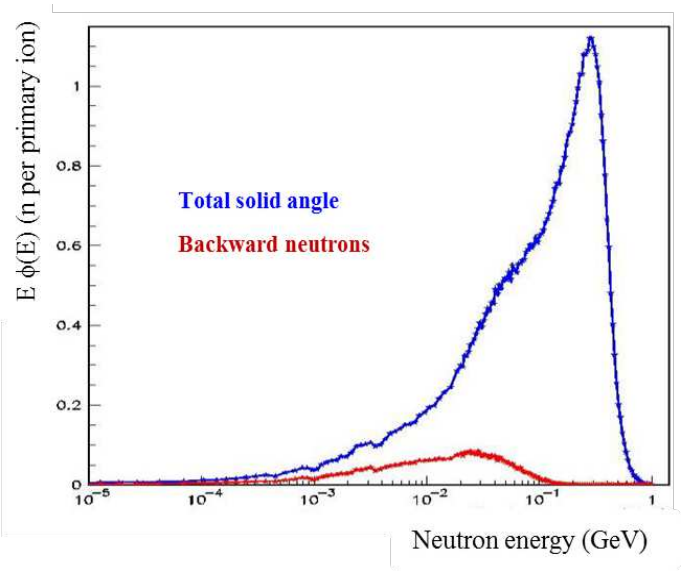

Fig. 2.57 Spectrum of secondary neutrons produced in the backward direction by 400 MeV/u carbon ions, hitting a phantom made of ICRU tissue; a comparison with the total energy spectrum of secondary neutrons is made.

2.9 TERA09 Front-End board

TERA09 ASIC results from a collaboration between the INFN and the Turin University and has been the goal of the technology transfer with an industrial partner named DETECTOR. DETECTOR Devices and Technologies Torino [31] is a company born ten years ago as a Turin University spin-off; in this last decade, DETECTOR dealt mainly with the development of ionization chambers for quality assurance and beam monitor tasks for the main hadron therapy centers in Europe. This company sponsored the thesis author's fellowship.

After the design and characterization phases, the TERA09s have to be embedded in dedicated electronics for the integration with the detector. At this point factors like encumbrance and heating become more important, requiring the design of a

more compact PCB. One common application is the detection of the particle beam profile using strip chambers where a large number of channels has to be readout. For this reason, a new board has been developed, named Front-End board. This PCB, depicted in Figure 2.58, hosts two ASICs grouping the digital signals on a 100 poles connector.

Fig. 2.58 The TERA09 front-end board hosting two ASICs is going to be the new monitor chamber readout electronics, adopted by DETECTOR Devices and Technology Torino.

The register addressing code of TERA09 provides for the chip selection through the eighth bit. An SMD switch controlling the board selection can be used for the readout of multiple Front-End boards working in parallel.
The new TERA09 Front-End board has been characterized at DETECTOR laboratories in summer 2017 (Figure 2.59), paying particular attention to the design modifications from the previous test board. The new PCB improves the protection of the more sensitive paths with a better isolation of the clock signal which is currently running on a dedicated layer and packaged between two ground layers. Furthermore, the delicate signal management (e.g. the digital reset and latch) increased the routing complexity, considering that the new PCB is more compact than the previous one and it hosts two ASICs instead of one.

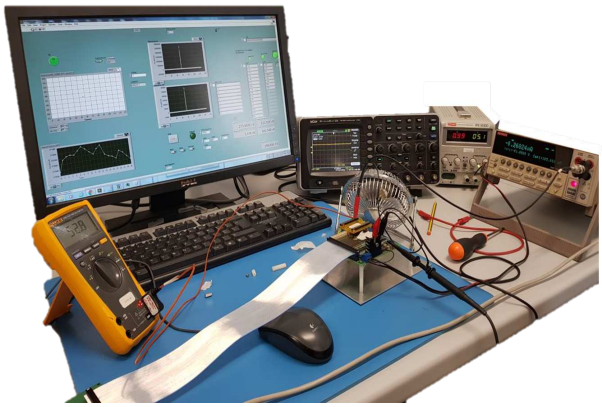

Fig. 2.59 TERA09 Front-End board characterization setup at DETECTOR laboratories.

Both channel counting noise and the linearity with the new board are similar to those previously presented. Concerning the power dissipation, a metal heat sink cools the ASIC ceramic packaging from 80 to 50 Celsius degrees (Figure 2.60).

The TERA09 Front-End board are currently in a test phase at GSI in Darmstadt, Germany [32]. From these tests it resulted that a dedicated ring biased at the preamplifier reference voltage and shielding the input channels, would be needed to improve the channel-by-channel counting homogeneity.

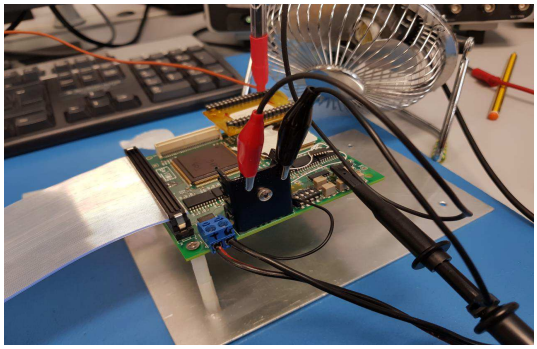

Fig. 2.60 TERA09 Front-End board characterization: power dissipation check. The crocodiles probes detect the heat sink temperature.

2.10 Summary

A new design of the 64 channels TERA front-end readout ASIC, named TERA09, has been developed and characterized. Each channel features a current-to-frequency converter followed by a 32-bit counter. Compared to the previous versions, the chip was designed to extend the current range for applications of beam monitoring with clinical pulsed particle beams. This was achieved by improving the maximum frequency of the current-to-frequency converter and by providing the possibility to access the sum of the counters of groups of channels directly on the chip.
Using a charge sensitivity of 200 fC, the gain uniformity across the channels of the chip was found to be within 3%, with a background current of approximately 2 pA. It was shown that the chip can be adapted to read out currents from few nA to several hundreds of μA, with deviations from linearity at the percent level. This input range corresponds to a dynamic range of almost 6 orders of magnitude, an increase of two orders of magnitude compared to its predecessors.

Concerning the radiation damage, the TERA09 ASIC does not have embedded protection techniques since is not meant to be placed directly on the beam during its activity. However, an irradiation test has been carried out in order to estimate the SEU probability. The test has been performed at the SIRAD Tandem accelerator at LNL in Padova, using a set of heavy ions with different energies and target incidence angles. Due to its field of application, there has been an interest in predicting the expected upset rate in a typical treatment room of CNAO. From the literature data regarding the secondary neutron fluence at the CNAO nozzle and and using the model developed in [26], we extrapolated a number (about 70 SEU/year) which has to be considered correct within an order of magnitude. However, this rate is then easily controllable through redundancy, with a second independent monitor chamber already present at CNAO as in every standard monitoring systems. The probability to have a simultaneous bit-flip in the same bit of both the monitor chambers, in a certain detector readout cycle, is therefore absolutely negligible. Terminated the ASIC characterization phase, a new Front-end board carrying two TERA09s has been developed and tested. Currently, this new PCB is at the GSI laboratories for further investigations, before being embedded into a gas detector to be tested on the field with a particle beam.

References

[1] Jatinder *et al.*, Intensity-Modulated Radiation Therapy. The state of the Art. American Association of Physicists in Medicine, Medical Physics Publishing, ISBN 1-930524-16-1, 2003.

[2] U. Amaldi, B. Larsson, Hadrontherapy in Oncology, Excerpta Medica, Elsevier, SBN 0-444-81918-5, 1994.

[3] http://www.oncolink.org/.

[4] L. R.R. Wilson, Radiological use of fast protons. Radiology 47, 487–491 (1946).

[5] U. Linz (ed.), Ion Beam Therapy, Biological and Medical Physics, Biomedical Engineering, DOI 10.1007/978-3-642-21414-1 4, Springer-Verlag Berlin Heidelberg 2012, Pag46.

[6] M. Hollmark, J. Uhrdin, D. Belkic, et al., Influence of multiple scattering and energy loss straggling on the absorbed dose distributions of therapeutic light ion beams: I. Analytical pencil beam model. Phys. Med. Biol. 49, 3247–3265 (2004).

[7] F. Fausti *et al.*, Design and characterization of a 64 channels ASIC front-end electronics for high-flux particle beam detectors, NIM A (2017)

[8] https://www.ptcog.ch/index.php/facilities-in-operation

[9] S. Giordanengo, Design, implementation and test of the hardware and software for the fast control of the dose delivery system of Centro Nazionale di Adroterapia Oncologica (CNAO).

[10] M. Regier et al., Medical accelerator for hadrontherapy with protons and carbon ions

[11] G. F. Knoll, Radiation detection and measurements, third edition, 2000.

[12] J. W. Boag and T. Wilson, The saturation curve at high ionization intensity.

[13] A.Molodozhentsev et al., "Design of dedicated proton synchrotron for the Prague Radiation Oncology Centre", in Proc. of the EPAC'98 Conference, Stockholm, Sweden.

[14] J. W. Boag, Radiation Dosimetry, edited by F. H. Attix and W. C. Roesh, (Acadernic Press, New York, 1966).

[15] A. La Rosa, et al., Nucl. Instr. Meth. Phys. Res. A 583 (2007) 461.

[16] U. Amaldi, et al., Nucl. Instr. Meth. Phys. Res. A 620 (2010) 563.

[17] F. Romano, et al., Nucl. Instr. Meth. Phys. Res. A 829 (2016) 153.

[18] R. Cirio, et al., Nucl. Instr. Meth. Phys. Res. A 798 (2015) 107.

[19] http://www.europractice.stfc.ac.uk/welcome.html.

[20] http://www.ni.com/labview/.

[21] S. Giordanengo, et al., The CNAO dose delivery system for modulated scanning ion beam radiotherapy, Med. Phys. 41 (2015) 263.

[22] A. La Rosa, et al., Nucl. Instr. Meth. Phys. Res. A 565 (2006) 833.

[23] R. Cirio *et al.*, Radiation damage studies of a recycling integrator VLSI chip for dosimetry and control of therapeutical beams, NIM A 482 (2002), 752 – 760

[24] D. Bisello, et al., SIRAD: an irradiation facility at the LNL Tandem accelerator for radiation damage studies on semiconductor detectors and electronic devices and systems, NIM A 462 (2001), 426-434.

[25] "Radiation Effects on the On-line Monitoring System of a Hadrontherapy Center", A. Ansarinejad, A. Ferrari ,Iranian Journal of Medical Physics Vol. 11, No. 4, (2014), 294-300

[26] M. Huhtinen, F. Faccio, Computational method to estimate Single Event Upset rates in an accelerator environment, NIM A 450 (2000), 155-172

[27] M. Omana, G. Papasso, D. Rossi, and C. Metra, “A Model for Transient Fault Propagation in Combinatorial Logic,” in Proc. 9th IEEE On-Line Testing Symp., 2003, pp. 111–115.

[28] David G. Mavis,Paul H. Eaton, SEU and SET modeling and mitigation in deep submicron technologies (2007), IEEE International Reliability Physics Symposium Proceedings. 45th Annual

[29] G. Mazza *et al.*, A 64-Channel Wide Dynamic Range Charge Measurement ASIC for Strip and Pixel Ionization Detector, Nuclear Science, IEEE Transactions 52, (2005) 847 – 853

[30] https://fondazionecnao.it/it/

[31] http://detector-group.com/it/

[32] https://www.gsi.de/en/start/news.htm

Chapter 3

Design of a single ion discriminator for hadrontherapy

3.1 Introduction

In the field of particle therapy the typical detectors used as beam monitors are gas ionization chambers (as explained in Chapter2). With the aim to measure the beam flux and the beam position, a ionization chamber collects the free charges generated by ionization effects in a gas volume, confined between a pair of metal electrodes. Recently in the Turin section of the National Institute of Nuclear Physics (INFN), the medical physics group is working on a new kind of silicon detectors to be used as beam monitor with the purpose to count the single beam particle. This task is a MoVeIT (Modeling and Verification for Ion beam Treatment planning) work package goal [1]. The MoVeIT project is developed by a interdisciplinary collaboration involving various INFN groups and the three Italian hadron therapy facilities (CNAO [2], LNS [4], TIFPA [3]), with the main purpose of rising the technical level in innovative treatment planning systems (TPS) and new verification devices. The novelty in these TPSs will be the implementation of biological models, the involvement of target fragmentation, radio biological effectiveness (RBE) and treatment of intra-tumor heterogeneity, such as hypoxia. As previously said, one goal of the MoVeIT collaboration is the development of a detector able to discriminate the single ion in a particle beam; although the milestone is primarily targeted to radio biology beams from low fluxes down to 1 Hz, up to 10^8 Hz, the MoVeIT community

is also interested in hadron therapy beams. The development is motivated by the well-known gas ionization chambers lack of efficiency due to charge recombination phenomena. The ambitious idea is to replace gas detectors with solid state ones, starting with a small size prototype. In this sense, even the high rate limit could be pushed further but at the particle therapy rate (up to $10^9 cm^{-2}\ s^{-1}$), state of the art silicon detectors are needed. Our group is collaborating with the Ultra Fast Silicon Detectors group (both groups belong to INFN and the University of Turin), with the aim to take the advantage of the know-how rose in the recent years R&D efforts in the High Energy Physics community (the CERN RD50 collaboration) and devoted to the development of detectors for extremely precise time measurements. The detector for MoVeIT will be equipped with Low Gain Avalanche Diodes (LGAD) called UFSD (the same acronym is used for the group), in a strip segmentation geometry. UFSDs are n-on-p silicon sensors featuring an internal moderate gain due to a thin p+ and low resistivity diffusion layer. The layer is located close to the bottom side of the n++ electrode of a heavily doped junction. Basically, the particle crossing the sensor releases hole-electron pairs from primary ionization. When the charges, moving thanks an electric field action, cross the gain layer, charge multiplication occurs followed by its collection at the electrodes [5]. The gain is limited to a value 10-20 in order to reduce noise perturbations and electric field confinement complexity in segmented detector configurations [5].

The main advantage of a UFSD is to provide an enhanced signal in thin detectors with similar noise level of a thicker silicon sensor of the same geometry. This allows producing detectors as thin as 50 μm providing signals of very short time duration and excellent time resolution. Moreover, UFSDs allow increasing the signal-to-noise (S/N) ratio by increasing the voltage bias thus better separating the noise from the signal [7].

The MoVeIT concept is based on the fact that, combining the counting number with the charge measured by a classic gas ionization chamber, it is possible to find the deposited energy by a single ion. In order to satisfy the 1 mm spatial resolution requirement for clinical practices, the number of particles should be measured within the 1% accuracy. The UFSD thickness allows a nanosecond level charge collection ($1 \sim$ ns in 50 μm) [5] and tens of picoseconds time resolution (35 ps time resolution for MIP particles has been observed in beam tests).

3.1.1 The author's contribution

The author's contribution on the work presented in Chapter 3 mainly consists in the microelectronics design of ABACUS, a multi-channel single ion discriminator for the MoVeIT silicon detector prototype. To fully understand the readout electronics requirements, a continuous interaction with the group colleagues it has been necessary, in order to fully understand the detector performances and studying the the detector response model, taking part of the sensor characterization during tests with ion beam. The ASIC design has been submitted on January 22^{nd} 2018 and the foundry tape-out is foreseen for the beginning of May. The author continued the technical collaboration required in defining the test board details.

3.2 The MoVeIT single ion counter: general aspects

The MoVeIT counting detector is a prototype idea, proof of concept for what could be a revolution in the particle therapy field: replacing the parallel plane gas ionization chambers with silicon detectors. Being such a challenging goal, it is mandatory to start from an in-scale device with a covered surface that is two orders of magnitude smaller than a standard monitor chamber. The single ion discriminator is going to be a 3×3 cm^2 silicon detector segmented in strips. These strips are UFSD sensors custom developed for this project. At the moment the group tested hundreds of strips mainly grouped in two different classes, due to their geometry: a set of 20 strips 15 mm long, 80 μm wide, with a 146 μm pitch and a second type of sensors grouped in 30 strips, 30 mm long, 150 μm wide and with a 216 μm pitch (further details related to these sensors will be provided in the following sections). Although at the moment it is not already clear which geometry is the best one, the short sensors showed a higher gain uniformity along their length. Concerning the geometry setup, both the sensor sets are $\sim$5 mm wide, therefore, for the 30 mm long strips a 6 sets row covers the target area whereas for the 15 mm long strips, a double row setup is needed. Each silicon strips will be coupled through wire bonding to a front-end channel, part of a custom-designed ASIC. The detector prototype is going to have an FPGA as readout system which will perform a first data manipulation for pile-up correction.
This thesis chapter is focused mainly in the description of the single ion discriminator chip, designed by the author on behalf of the MoVeIT collaboration. This chip main requirements are a wide input charge range (3-150 fC), and a discrimination rate of

at least 100 MHz (the radio biology target), over the entire input signal range. In order to deal with a particle therapy rate, the entire detector would manage a 10^9 Hz rate, therefore, the design phase was looking for a higher counting rate. Nevertheless, a proper readout system and an off-line analysis will improve the counting efficiency of this prototype.

3.3 Silicon detector basis

The functionality of world wide diffused devices such silicon detectors originates from the p-n junction physics. Considering an n-on-p silicon sensor, the sensitive volume is the depleted region at the interface between the two differently doped zones by means of external biasing. A charged particle crossing the sensitive volumes ionizes the material along its track thus creating electron-hole pairs (e-h) which are collected drifting respectively toward the n++ contact (electrons) and p++ contact (holes). As reference number, a Minimum Ionizing Particle (MIP) passing through a silicon sensor, creates 73 e-h pairs per micron thickness. The Figure 3.1 qualitatively represents the operating principle of a silicon sensor. These free e-h pairs, moving under the influence of the weighting field that represents the capacitive coupling between the charge and the readout pad, induce a current signal at the electrodes that ends when the last charge carrier reaches its electrode. The integral of this induced current is roughly 1.3 fC every 100 micron.

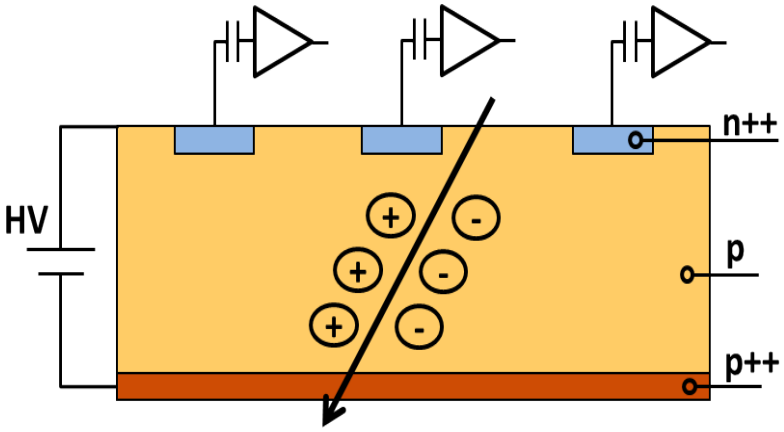

Fig. 3.1 Basic operational principles of a silicon detector: an external bias voltage polarizes inversely the p-n junction, creating a large depleted volume. When an incident charged particle crosses the sensor, it creates electron-hole pairs whose drift generates an induced current in the electronics.

In a typical condition of a 300-micron thick sensor at 600 volts, the average electron drift time is 3 ns and the average holes drift time is 5.5 ns. An important features is that the maximum signal amplitude is independent from the sensor thickness. This comes from the Ramo – Shockley's theorem; the concept can be imagined considering the fact that the electric field is inversely proportional to the sensor thickness (parallel plane geometry) whereas the number of generated charge carriers is directly proportional to it, resulting in no dependence from this parameter. The desired signal in order to have a precise and fast detection has a large amplitude and sharp rising profile. A silicon sensor featuring an internal charge gain is a good candidate to fulfill both requirements.

3.4 UFSD as LGAD evolution

As previously mentioned, UFSD are Low Gain Avalanche Detector that mainly differentiates from the well-known Avalanche Photo Diode for the gain factor: 10-20 for UFSD, hundreds for APD. LGAD sensors maintain the principle of the avalanche process activated by the movement of a charged particle in a $\sim 10^5$V cm^{-1} electric field but they have a modified doping profile with a micrometers thick p+ layer doped with Boron or Gallium, placed on the bottom of the n++ electrode to limit

the multiplication volume. A simplified drawing showing a traditional n-in-p silicon detector and an LGAD are shown in Figure 3.2.

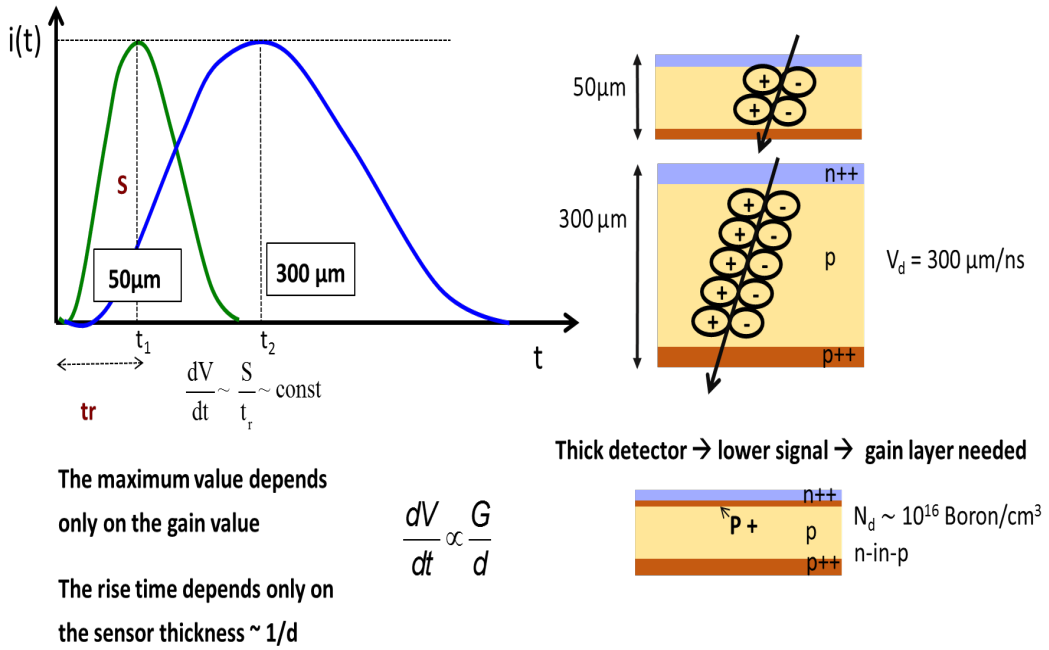

Fig. 3.2 Qualitative drawing representing the comparison between a classical 300 μm thick silicon sensor an a 50μm thick UFSD.

Thanks to the heavily doped thin layer underneath the n++ electrode (for the n-on-p type), the LGAD doping profile has a sharp transition responsible of two different electric field zones. A charge collected to the electrode crosses a drift volume with typically ∼30kV cm−1 before the thin multiplication zone with ∼300 kV cm−1. This features requires a specific effort in shaping the implants to allow high bias-voltage operation without breakdown. The multiplication mechanism is activated by the charge carriers that cross the additional gain layer therefore i.e. electrons in n-in-p (p++ gain layer under the n++ electrode) and holes in p-in-n (n++ gain layer under the p++ electrode). Thanks to their higher mobility, the elections require a lower electric field to activate the multiplication; for this reason the n-in-p type is preferable, aiming to control the multiplication process at the point that it is possible to limit the activation process to the electron, thus avoiding to operate the device in avalanche mode. With this achievement, the gain is less sensitive to the electric field variations and the excess noise factor (fluctuation in charge) is reduced too. An important feature is the internal gain amplifies the noise more than the signal

nevertheless, since the electric noise of the detector is larger, the sensor SNR is eventually improved.
UFSDs are silicon sensors based on thinning the LGADs and keeping the gain at its minimum value needed to detect the single particle and allowing accurate time measurement. High SNR and fast charge collection are obtained with low gain and thin sensitive volumes. In fact, the high gain not only implies an intrinsic increased sensor noise but it complicates the electrode segmentation for the high electric field that even increases after irradiation, leading to power consumption issues and limits. As already mentioned, the signal amplitude depends only on the sensor gain and its rising edge steepness is related to the drift time of the electron crossing the sensor thickness [5]. The weighting field uniformity is a priority and leads to the choice of the simplest geometry: the parallel plane one. The capacitance of such type of sensor increases with an inverse proportionality respect to the two electrodes distance. Therefore thin detectors are needed in order to maximize the data readout speed. From simulation and experimental results it has been found that the balance between the two effects is reached with a $\sim 50 \mu$m thin UFSD.

3.4.1 UFSD issues and mitigation solutions

UFSDs are devices developed aiming to achieve the best time resolution σ_t. Although the MoVeIT single ion discriminator prototype does not perform time measurements, the intrinsic fast charge collection (nanosecond level) plus the UFSDs achievable granularity make them excellent candidate for high rate signal discrimination.
The phenomena responsible to the time accuracy degradation and the relative mitigation techniques are hereafter reported. The time resolution parameter depends on various contributions:

$$\sigma_t{}^2 = \sigma_{TimeWalk}{}^2 + \sigma_{LandauNoise}{}^2 + \sigma_{Distortion}{}^2 + \sigma_{Jitter}{}^2 + \sigma_{TDC}{}^2 \tag{3.1}$$

where the equation terms are hereafter explained.
The Landau statistic regulates the physics of the energy deposition for a charged particle crossing a silicon volume. These fluctuations characterize the signal amplitude and even for a fixed amplitude, the charge deposition is non uniform (current fluctuations). What happens is that a ionizing particle has a certain probability to interact with a silicon electron that is not the valence one. If extracted, an inner

electron causes secondary ionization since it is more energetic. As a result of these rare events the deposited energy spectrum has a Most Probable Value (MPV) which is approximately 0.7 times the mean value.

For a MIP the mean energy loss per unit length is 3.88 MeV cm^{-1} or 116 keV for 300 μm of silicon with a MPV of 81 keV. Since the minimum energy required to create an e-h pair in silicon is 3.6 eV, the mean number of charge carriers created is 108 per micrometer and 737μm as MPV. This results in a mean signal from a 300 μm thick detector of 32000 electrons, i.e. a most probable value of 22500 electrons; which is equal to 3.6 fC. In a 50 μm UFSD with a gain factor $\sim$15, the charge collected at the electrodes is 8 fC MPV, as shown in Fig 3.3, where typical UFSD signal shapes are reproduced with the Weightfield2 and TCAD Synopsys softwares [6], as explained in the following sections.

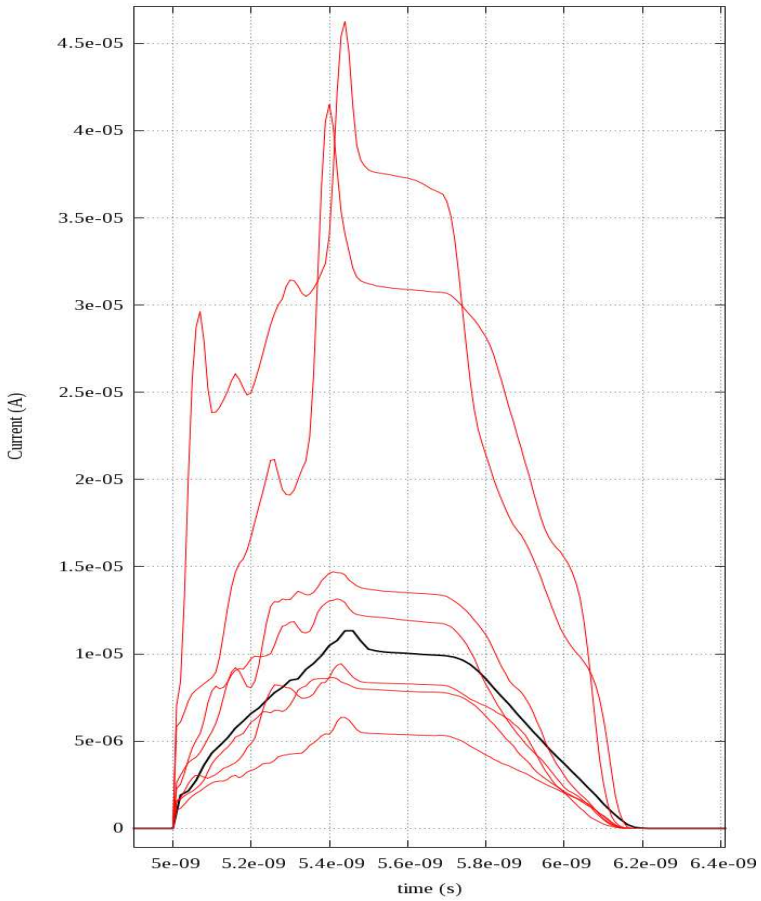

Fig. 3.3 Sensor signal for different charge values extracted from Landau distribution. The reference sensor is 50/*mu*m thick and has a nominal gain factor of 15. The black line represents a MIP signal resulting in a peak current of $\sim$ 10 μA for a 8 fC charge

As a result of Landau fluctuations, the Time Walk effect is responsible for discriminator triggering delays due to the fact that large amplitude signals are steeper thus they cross the threshold in advance with respect to low amplitude ones. Intuitively, for a linear signal there is a proportion that ties the couples "time needed to reach the threshold, rising time" and "threshold level, signal amplitude". This concept is represented in Fig 3.4.

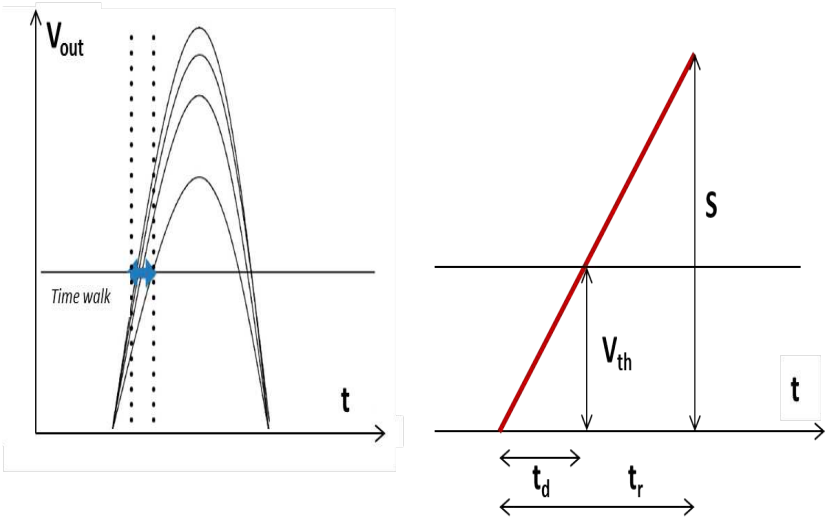

Fig. 3.4 Time Walk effect: geometrical explanation.

Let's assume for simplicity a linear signal, with amplitude S and rise time t_r. This signal crosses the threshold V_{th} with a delay t_d. Using the geometrical relationship $t_d/t_{rise} = V_{th}/S$, the moment when the particle crosses the threshold can be written as:

$$\sigma_{TimeWalk} = [t_d]_{RMS} = [\frac{V_{th}}{(S/t_{rise})}]_{RMS} \propto [\frac{N}{(dV/dt)}]_{RMS} \tag{3.2}$$

where we used $S/t_{rise} = dV/dt$ and N is the rms noise.

Time walk compensations could be based on higher slew rates and appropriate discriminator techniques like the constant fraction discrimination (the latter is in general associated to an increased complexity that reduces the bandwidth).

As explained before, once the charge is released inside the silicon volume, its movement induces a current signal on the electrodes that is modeled with the Ramo-Shockley's equation:

$$i(t) = -q\bar{v} \cdot \bar{E} \tag{3.3}$$

where q is the charge of the particle, $\bar{v}$ is the drift velocity and $\bar{E_w}$ the weighting field. Looking at the equation, a sensor designer sees two goals: keep the drift velocity constant throughout the sensitive volume and keep the weighting field homogeneous along the electrode pitch. The target is to saturate the drift velocity (this is achievable with a $\sim$30kV cm^{-1} electric field, at room temperature) and implants need to have a width very similar to the pitch but larger than the sensor thickness. The Jitter term is used to describe the effect of the sensor or electronics noise that overlaps the rising edge of the signal, resulting in early or late threshold crossing (discriminator firing). As expressed by the following equation, a steeper signal i.e. a high slew rate signal should in principle minimize the jitter effect, since a theoretically vertical signal (delta-like) does not suffer by noise overlapping; unfortunately high slew rate means wide bandwidth that carries a higher noise. $dV/dt = S/t_{rise}$ and therefore:

$$\sigma_{Jitter} = N/(dV/dt) \approx t_{rise}/(S/N) \tag{3.4}$$

3.5 UFSD design and simulation

The development of the MoVeIT detector prototype started simulating the charge collection in silicon for different shapes and dimensions of terminals and silicon layers. The Weightfield2 [6] and TCAD Synopsys Sentaurus [9] softwares are used to get an estimation of the sensor output signal features, which are critical for the front-end trade-off requests. Weightfield2 (WF2) is a tool developed by the UFSD group in Turin to properly simulate the behavior of a silicon sensor with gain, oriented to timing performance studies. Through a graphical user interface, as shown in Figure 3.5, the user can select among different types of particle (MIP, α-particle, laser, x-ray), changing the incident angle and hit position on the sensor; moreover it is possible to vary the sensor geometry, doping concentrations, the value of an external B-field, ambient temperature and thermal diffusion. This software allows to estimate an approximated oscilloscope and front-end electronics response. WF2 embeds a library with the energy release information, as result of GEANT4 simulation with

default parameters [8]. With this pre-calculated data, WF2 simulates the e-h pairs generated by ionization along the particle track. It is up to the user to decide the adopted accuracy in time steps or in percentage of pairs simulated for a given event. For each time step, the induced current is derived using Ramo-Shockley's theorem by summing over the single charge carrier contribution with the related charge, position and weighting field.

WF2 implements different models for the impact ionization rate. The van Overstraeten [10] and Massey [11] models, are based on the Chynoweth law whereas other two, the Bologna [12] and the Okuto-Crowell [13] models propose their own law for the ionization coefficients. Except for Massey, Synopsis Sentaurus [9] implements these models too.

The high accuracy of the WF2 modelization has been confirmed through crosscheck for MIP and alpha particles using TCAD Sentaurus results and measured signals.

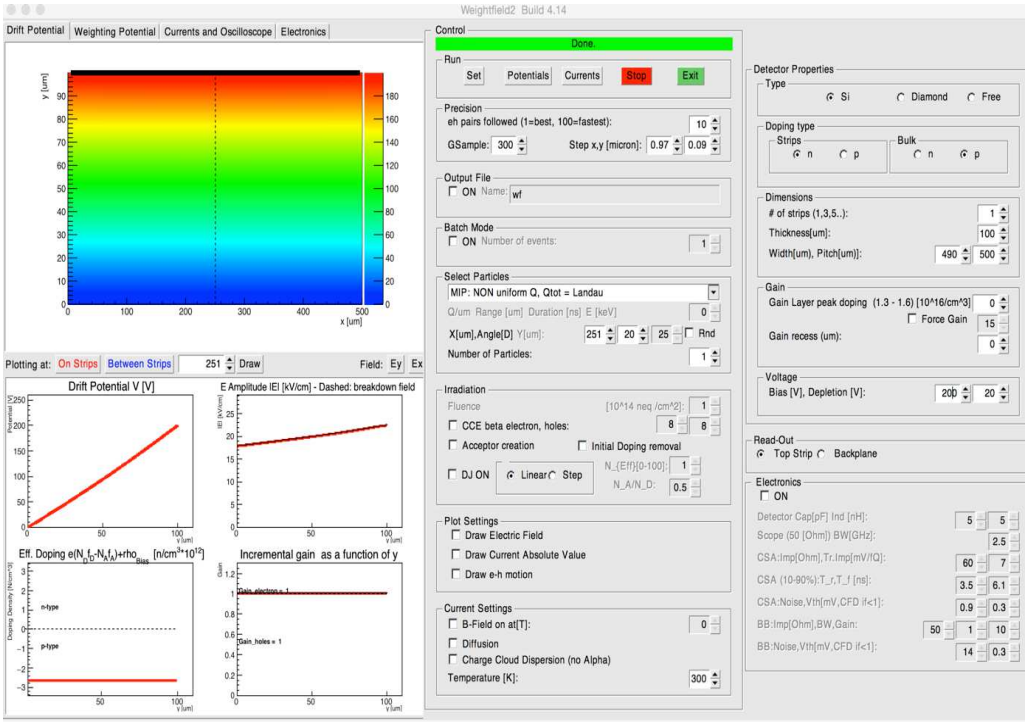

Fig. 3.5 Snapshot from the Weightfield2 graphical user interface. Figure taken with permission from [7].

For a 50 μm thick UFSD with a gain factor of 15, a MIP particle generates a 1.2 ns signal with a charge of approximately 8 fC MPV [17]. Figure 3.3 shows

the sensor signal modelization for different charge values, considering the Landau distribution: the black line refers to the MIP effect.

The MIP signal can be approximated by a trapezoid with a rise time and a fall time of 450 ps, a minor base b1 of 300 ps and the area that corresponds to the input charge Q_{in}. The trapezoid major base b2 is 1.2 ns and the height can be calculated with the following equation:

$$h = \frac{2Q_{in}}{b1 + b2} \approx 1.33 \cdot 10^9 [s^{-1}] \cdot Q_{in} \tag{3.5}$$

This simplification has been used during the front-end design since it allows an easier parametrization of the signal charge and repetition rate, suitable for CAD transient sweep analysis. The spread of the amplitude probability can be fitted with a Landau distribution, as reported in Fig 3.6.

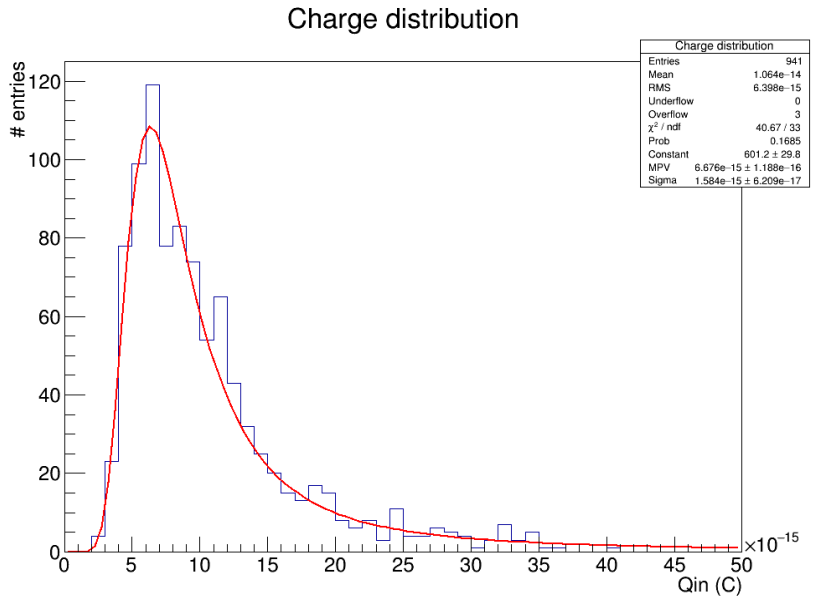

Fig. 3.6 Simulated input charge fitted with Landau distribution

Taking into account a typical proton beam for therapy applications, the 60 MeV - 250 MeV energy range corresponds to 2-6 MIP particles and a range of charge collected by the sensor electrodes from 3-4 fC to 150 fC. The same simulation procedure for the Landau distributions has been adopted in these cases, for both protons and Carbon ions with therapeutic energies (the CNAO therapy center synchrotron

accelerator works with these two ions). Figure 3.7 shows the Landau distribution for different proton energies, considering a sensor gain of 10.

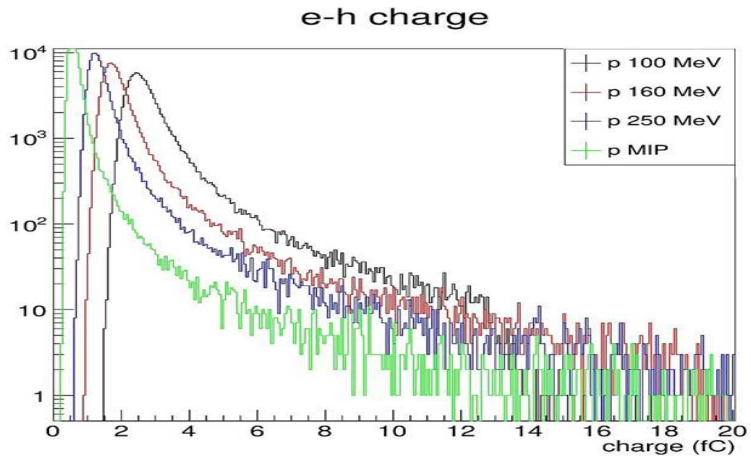

Fig. 3.7 Charge signal distributions, in a 50 μm UFSD with gain 10. The different profiles represent 4 different proton projectiles energies: MIP equivalent, 100 MeV, 160 MeV and 250 MeV.

Figure3.8 refers to Carbon ions (C^{6+}) and gain less sensors, due to the fact that having an higher Z, the C^{6+} stopping power is higher thus the greater released energy does not require to be amplified.

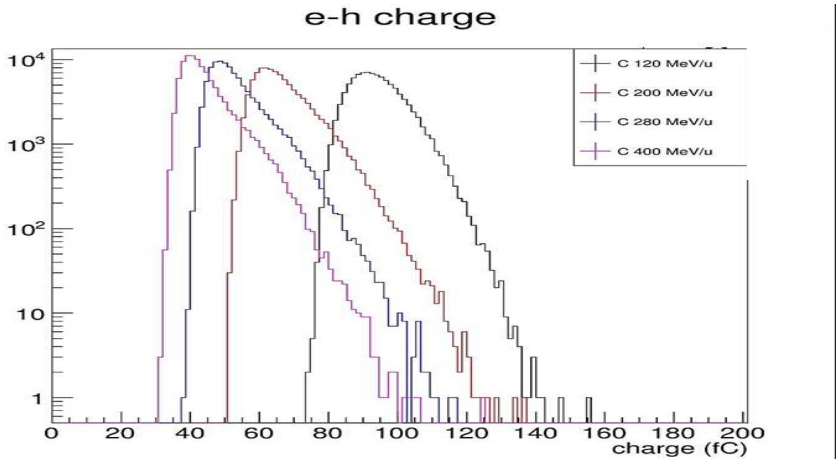

Fig. 3.8 Charge signal distributions, in a 50 μm UFSD with no gain. The different profiles represent 4 different Carbon ion energies: 120 MeV/u, 200 MeV/u, 280 MeV/u and 400 MeV/u.

A further investigation step has been performed with beam test campaigns at the National Center for Oncological Hadron therapy (CNAO) at Pavia, Italy, using UFSD sensors (as reported in the following section). In the meanwhile, the data coming from Weightfield2 in text format as charge vs time or pulses vs time have been used as input stimulus during the front-end electronics design.

3.6 Radiation damage and hardening

UFSD sensors are developed in a group mainly devoted to High Energy Physics (HEP) and this means that these devices must be radiation tolerant up to high fluences. One important application for UFSD is the equipment of the CERN CMS inner barrel for timing, targeted at the High Luminosity upgrade of the Large Hadron Collider (HL-LHC) [14]; for HL-LHC, UFSDs have to guarantee a proper functionality up to $5 \cdot 10^{15}$ neutron-equivalent cm$^{-}2$. The CERN RD50 collaboration [15] is in charge to investigate on radiation damage effects for UFSDs. From several studies it came out that, UFSDs suffer for classical silicon effects like decrease of charge collection efficiency, increase in the leakage current and changing in the doping concentration. Although this list of results is not surprising, the fact that UFSDs have a inner and thin gain layer make them more sensitive to leakage current and to the modification of doping concentrations.
In silicon the ionizing radiation is responsible of surface currents that are not multiplied by the gain and saturate at the level of Mrad. The non-ionizing loss is responsible for heavier damage at the bulk level, introducing charge on neutral defects in the silicon band gap. Dopant-lattice defects interaction change the junction equilibrium.

For a certain volume V of silicon sensor exposed to a radiation fluence ϕ, the leakage current increases as:

$$\delta i = \alpha V \phi \tag{3.6}$$

with α = 2.5 10^{-17} A cm^{-1} for protons and α = 4.0 10^{-17} A cm^{-1} for neutrons, respectively.

Since this effect depends exponentially on the temperature (as described in the following equation), it can be mitigated by cooling the system.

$$i(t) = i_0 e^{-t/\tau_{eff}} \tag{3.7}$$

Working with thin sensors reduces the trapping probability during the charge drift because this phenomena is proportional to the charge free path; for a fluence of ϕ = 10^{15} neq/cm^2, the average free path is about 50μm. Another advantage of thin sensors over the thick ones is the reduced bias voltage and leakage current.

In UFSDs, the main noise source is the shot noise which reveals whenever charge carriers cross a potential. This effect is accentuated with the increasing of the leakage current. The shot noise current density is given by the following equation:

$$i^2_{Shot} = 2q \cdot i_{Det} = 2q \cdot [i_{Det} + (i_{Bulk} + i_{Det})] G^2 G^x \tag{3.8}$$

where q is the electron charge and G^x is the excess noise factor expressed as a power of the gain value.

The shot noise effect is important for irradiated detectors where it can be equal or even greater than the electronic noise and the SNR is reduced even if the sensor has an inner gain. In order to mitigate these effects thus hardening the UFSDs, alternative sensor wafers have been produced where the Boron dopant has been replaced with Gallium; the theory and the simulations suggest that this approach should reduce the formation of the acceptor-interstitial interactions. More in detail, gallium shows less sensitivity to bias voltage changes and this is related to its more extended, with a lower gradient doping profile just after the implantation. The smoother gain curve of a gallium-doped silicon sensor, reaches the break-down condition at voltage values that are typically higher than the boron-doped sensors. A second even more sophisticated solution is based on Carbon spray: Carbon enriched wafers exploit the higher mobility of these atoms that fills the interstitials instead of Boron atoms.
Thanks to Weightfield2, it is possible to simulate the effect of charge trapping and initial acceptor removal in a silicon sensor. The Figure 3.9 shows a simulated current signal for a 50 μm thick UFSD exposed to different fluences.

From process simulations, gallium revealed to have a wider and less pronounced doping profile immediately after implantation with respect to boron, leading to less steep gain curves and a lower sensitivity to bias voltage changes, reaching break-down at higher voltages than boron.

In the framework of the MoVeIT application, the ions crossing the sensor have an higher stopping power due to the energy range which is in the order of tens or hundreds of MeV and not GeV, as in HEP. For this reason, the hitting particles are considered as 2-6 MIPs. Therefore the particle therapy scenario is rather different and need to be explored properly. Some results related to this aspect are reported in the following sections.

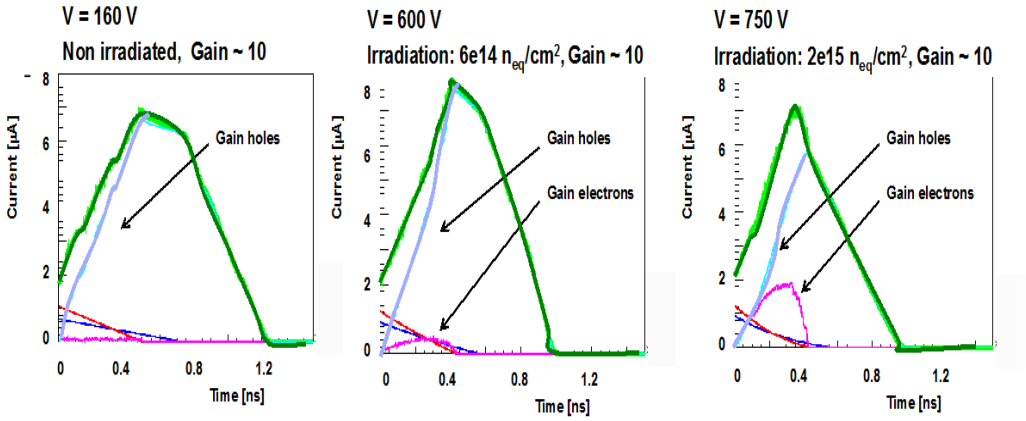

Fig. 3.9 50 μm UFSD current signals. Left: un-irradiated sensor with gain $\sim$ 10; center: UFSD current signal after a fluence ϕ = 6e14 neq / cm^2; right: UFSD current signal after a fluence $\phi = 2 \cdot 10^{15}$ n eq /cm^2 . Is it possible to notice than the decreasing of the current induced by trapping is higher for longer drift time, while the changes of the location where multiplication happens, from the gain layer to the bulk, affects the shape of the induced current signal since the contribution from gain electrons starts to be relevant. Figure taken with permission from [7].

3.7 UFSD productions

In 2018, there are three centers in the world that are producing 50 micron thick UFSD sensors: the Centro Nacional de Micrelectrónica (CNM) Barcelona, the Italian Fondazione Bruno Kessler (FBK), in Trento and Hammamatsu Photonics. CNM has been the first group proposing and developing the LGAD technology in

2016 [16]; they proposed a variety of designs, differing mainly in the substrate manufacturing technique like the float zone (FZ), silicon-on-insulator (SoI) and epitaxial (epi) with high and medium resistivity. On the other hand, the UFSD group in Turin is collaborating and is weekly in contact with FBK where they are currently working on a third jointed sensor production run (planned for summer 2018), after the first with 300μm thick LGAD and a second production with 50 μm thinned wafers with various dopant configurations as reported in Table 1.2.

3.8 UFSD coupled amplifiers

For UFSDs applications the most challenging aspect concerning the detector electronics is the very front-end design. More specifically, the design of the amplifier and how the other blocks interact with it, starting form the detector. Figure 3.10 depicts the general detector-preamplifier system.

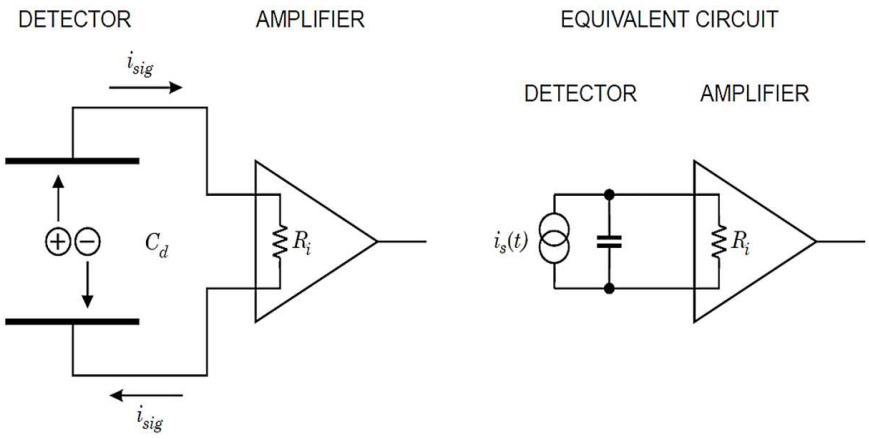

Fig. 3.10 The detector and amplifier (left) and the equivalent detector circuit (right).

An amplifier is a device which applies a multiplication factor to a signal i.e. it is a device with a certain gain. The most basic and diffused concept of amplifier is represented by a transistor. It is well known that bipolar or MOSFET transistors behavior is non perfectly linear and even their gain is not constant. A typical method used to overcome these issue is to feedback the amplifiers with passive linear components such resistors or capacitances; a related improvements in using these kind of feedback is the fact that the gain is independent of the signal amplitude. Another

relevant aspect is the way the amplifier is connected with the circuit. The amplifier input impedance is designed following the input signal generator features (the sensor, in this case) whereas the output impedance choice is based on the load, independently form the input impedance. It is possible to distinguish the amplifier types among four basis configurations: voltage amplifiers, current amplifiers, transconductance amplifiers and trans-impedance amplifiers (more details are reported in the following table).

Table 3.1 Amplifier types.

Type	**Input**	**Output**	**Gain**	Z_{in}	Z_{in}
Voltage	V_i	V_o	$A_V = V_o/V_i$	high	low
Current	i_i	i_o	$A_i = i_o/i_i$	low	high
Transconductance	V_i	i_o	$A_g = i_o/V_i$	high	high
Transimpedance	i_i	V_o	$A_z = V_0/i_i$	low	low

The UFSD sensors employed in the MoVeIT project provide a current signal for the front-end and the microelectronics team decided to adopt the trans-impedance architecture for the amplifier design. Considered the project R&D character, the design of two different configurations has been developed and two different ASIC prototypes are going to be produced. One prototype is based on a resistive feedback, named for simplicity TIA and a second configuration, with a capacitive feedback, is a Charge Sensitive Amplifier (CSA).

It is possible to point out some features and the critical aspects emerging form coupling TIA and CSA amplifiers with UFSDs.
A typical TIA amplifier is characterized in having a wide bandwidth that allows to tightly follow the time structure of the current sourced by the sensor (low shaping effect) and convert this values in a high voltage slew rate (dV/dt) at its output. This quality maximize the use of thin sensors where the current pulse derivative is large. Although an high slew rate minimizes the jitter contribution, the required high bandwidth introduces noise effects and is fed by high bias current (power budget and heat dissipation aspects). Optimizing the TIA performance suggest to keep the amplifier input time constant ($C_d \cdot R_i$) in the order of magnitude of the current rise-time ($\sim$ 1-2 nanoseconds, for UFSDs).
For a CSA, the approach is complementary. The bandwidth is shorter and the output voltage depends on the amplifier response to the input stimuli and it does not follow

the input signal shape. The output signal amplitude is proportional to the input charge. The CSA jitter minimization is due to the noise filtering related to the signal integration. The slew rate term is here smoother but, since $dV/dt \sim dQ(t)/dt = i(t)$, the maximum slew rate corresponds to the signal current maximum. The minor steepness (compared to the current-mode) does not affect the fast charge collection that depends on the sensor thickness. The two fronts of the CSA output signal have different dependences: • The rise time: t rise $\sim (C_d + C_l)/g_m$ where g_m is the input stage transconductance and C_l is the capacitance value of the circuit loading the preamplifier output. • The fall (discharge) time: t fall $\sim R_f\, C_f$, controlled by the feedback components. The fall time, t_f , should be longer than the rise time, $t_f >> t_r$, otherwise the charge will discharge as predicted by the ballistic deficit effect:

$$V_{out}^{prek} = \frac{Q_f}{C_f}\left[\frac{t_f}{t_r}\right]^{t_r/(t_f+t_r)} \tag{3.9}$$

The detector capacitance strongly influence the CSA performances, in term of noise, signal rise time and signal amplitude. This latter aspect means that the fraction of signal charge Q_s stored on the feedback capacitor C_f (the one that is amplified), is not the total amount of the charge collected by the detector, as happens for current-mode amplifiers. The relationship that explain this effect, is hereafter reported.

$$\frac{Q_f}{Q_s} = \frac{Q_f}{Q_Det + Q_f} = \frac{(1+A_0)C_f}{C_Det + (1+A_0)} \tag{3.10}$$

where A_0 is the input transistor open loop gain.

A comparison between a high-bandwidth TIA amplifier and a CSA is shown in following Table.

Table 3.2 comparison high-bandwidth TIA and CSA

Type	**SR**	**Rise time**	**Fall time**	**Ballistic deficit**	**Available charge**
high-BW TIA	d_i/d_t	$R_{in}C_d$	$R_{in}C_d$	1	Q_{tot}
CSA	$i(t)$	$(C_l + C_d)/g_m$	R_fC_f	$(Q_f/C_f)\cdot[t_f/t_r]^{t_r/t_f+t_r}$	$\frac{(1+A_0)C_f}{C_d+(1+A_0)C_f}$

The non uniform e-h pairs creation is differently managed by TIA and CSA. While TIAs suffer to Landau fluctuations at the pulse beginning, CSAs have a cumulative effect for the total charge collected.

Since this thesis work focuses on the CSA-based ASIC prototype, few more details regarding this topology structure are hereafter presented.
A charge-sensitive amplifier is an active integrator obtained starting from an inverting voltage amplifier configuration. The input impedance is ideally infinite (no charge flowing through the input MOSFET gate) and a feedback capacitance shorts the amplifier output with its input terminal. A CSA representation is reported in Fig 3.11.

The voltage across C_f is:

$$V_f = (A+1) \cdot V_i \tag{3.11}$$

thus, assuming $Z_i \approx \infty \rightarrow Q_i = Q_f$, the charge deposited on C_f is:

$$Q_f = C_f \cdot V_f = C_f \cdot (A+1) \cdot V_i \tag{3.12}$$

Whereas the effective input capacitance, also called dynamic capacitance, is:

$$C_i = Q_i / V_i = C_f \cdot (A+1) \tag{3.13}$$

The gain A_q:

$$A_q = \frac{dV_0}{dQ_i} = \frac{AV_i}{C_i V_i} = \frac{A}{C_i} = \frac{A}{A+1} \cdot \frac{1}{C_f} \tag{3.14}$$

if A » 1:

$$\approx \frac{1}{C_f} \tag{3.15}$$

The gain of an ideal CSA is regulated selecting a well-controlled component, the feedback capacitance.

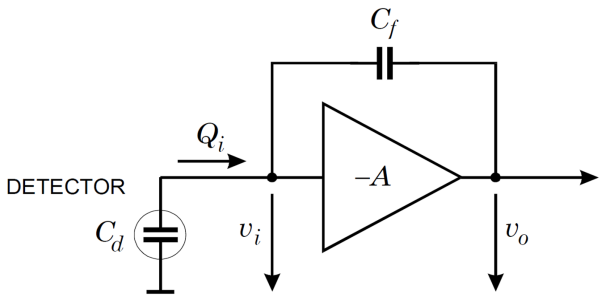

Fig. 3.11 Schematic representation of the charge sensitive amplifier topology.

3.9 Electronic noise: general considerations

In a detector, both the sensor and the electronics introduce noise perturbation effects like the resistor thermal noise, leakage current from the sensor and the amplifier flicker and white noise components. It is possible to represent the effect of each noise source through a model describing the power associated to this component, in function of the frequency (the spectral power density). The spectral power density of an amplifier is strongly related to its architecture whereas passive components like capacitors and resistors influence the circuit with noise effects that do not depend on the circuit configuration; even the sensor spectral power density is circuit-independent.

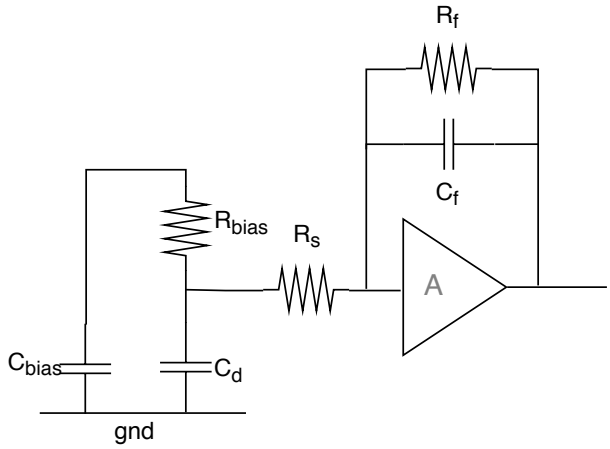

Fig. 3.12 Resistive and capacitive components in a generic sensor-amplifier connection.

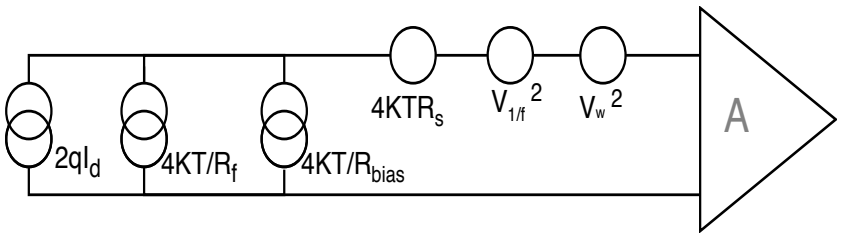

Fig. 3.13 Sensor-amplifier noise model: the spectral power density model representation.

The biasing and feedback resistor thermal noise components, as well as the detector leakage current are conventionally modeled with current sources parallel connected with the amplifier inputs. With the same approach, other noise sources like the series resistors thermal effects, the noise frequency-independent (white noise) and the flicker noise are represented with voltage sources connected in series with the amplifier. The amplifier transfer function properly calculated for the noise modelization, is convoluted with the power density components. At this point, an indicator named Equivalent Noise Charge (ENC) is used to quantify the single noise source impact. ENC is the number of electrons required as input signal to produce an output perturbation equivalent to the one produced by the noise component under study. Since the series, parallel and flicker components do not depend one from the other, it is possible to calculate the total effect from a quadratic sum of the terms as

$$ENC_{tot} = \sqrt{{ENC_s}^2 + {ENC_p}^2 + {ENC_f}^2} \tag{3.16}$$

$ENC_s \propto \sqrt{1/t_p} \cdot (C_d + C_{in})$; $ENC_p \propto \sqrt{t_p}$ and $ENC_f \propto (C_d + C_{in})$; C_{in} is the equivalent capacitance in parallel with the detector, t_p is the preamplifier peaking time.

The flicker noise component is related to the technology features and is not relevant for fast architectures, while it is generally better to adopt small capacitance detectors. The shot noise source results from the leakage current and its spectral power density is

$${i_n}^2 = 2qI_{bl}G^{2+x} \tag{3.17}$$

with I_{bl} the bulk leakage current, G^{2+x} the contribution of the sensor gain and q the electron charge. The ENC_p term calculated for a CSA amplifier is

$$ENC_p = \sqrt{\frac{I_{bl}\tau_f}{2q}} G^{1+x/2} \tag{3.18}$$

with $\tau_f = R_f \cdot C_f$, the feedback constant.

3.10 Beam test with UFSD pads

With the aim to evaluate the counting and timing properties of UFSD sensors in a real application, four beam tests have been performed at CNAO particle therapy center in Pavia, Italy.

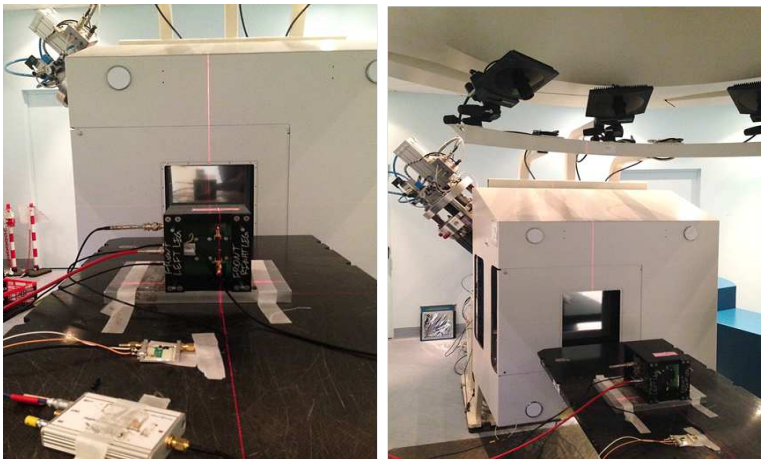

Fig. 3.14 Test setup at CNAO: the squared black box fixed over the treatment bed contains two PCB mounting a 50 μ m thin UFSD pad each. the white device with a central metalized window is the beam monitor chamber, placed at the end of the beam extraction line.

The results presented in [18] and here reported, have been obtained via off-line analysis of the collected waveforms.
The standard testing condition presented a proton beam in the 62 to 227 MeV energy range. As an example, the third test is reported. In this test the data acquisition took place during 32 runs of 2 10^{10} protons each, with a beam FWHM of 1 cm and a flux range from 20% to 100% of the maximum value.
Two sensor pads of 50 μm active thickness (1.2 x 1.2 mm^2) have been mounted with a telescope setup that fixed them at 1 cm distance one form the other; by means of a metallic box (telescope), the two sensors were aligned to the beam (Figure 3.14).

The pads have been realized by the National Center for Microelectronics (CNM) in Barcelona and by Hamamatsu (Japan). The sensors outputs were fed to broadband amplifiers (CIVIDEC 40 dB [19]), readout through an oscilloscope (Teledyne Lecroy WaveRunner 640Zi, 40 GS/s sampling rate [20]), and acquired in parallel through a digitizer (CAEN DT5742, 5 GS/s sampling rate [21]), providing snapshot of 200 ns duration. The setup included also a PTW PinPoint ionization chamber (T31015 [22]) aligned to the beam after the sensors, used to provide a reference rate, HV and LV power supplies. Two computers, one in the treatment room and one in the control room have been used to acquire the measurements and to remotely control all the instrumentations. Figure 3.15 shows the pad sensors used in the test.

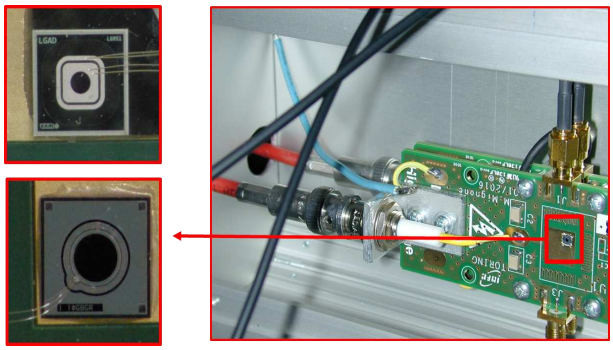

Fig. 3.15 Picture of the sensors mounted in a metallic box and aligned to the beam. Left-down corner: Hamamatsu sensor (1 mm Ø x 50 μm thickness). Left-up corner: CNM sensor (1,2 x 1,2 x 50 μm thickness).

Using the digitizer it is possible to record data as time window (Figure 3.16); this allowed to study the shape and duration of the signal produced by proton tracks. The measured signal duration was less than 2 ns, which limits the pile-up effect for incoming beam with a Poisson distribution of particles up to 10^8 p/s on the single channel. Under these conditions, the choice of the best discrimination threshold is clearly fundamental to deal with pile-up issues, as represented in Figure 3.17.

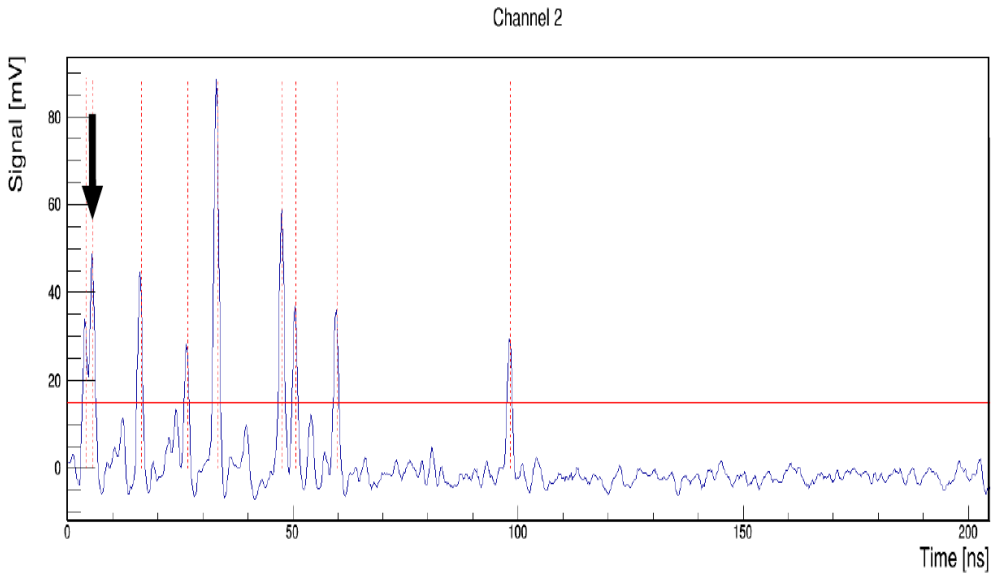

Fig. 3.16 A 200 ns snapshot acquired with the channel 2 of the digitizer. It is possible to estimate the nanosecond structure of the CNAO proton beam, looking at the tens of millivolts peaks, where each peak corresponds to a single ion signal. The black arrow highlights a pile-up event.

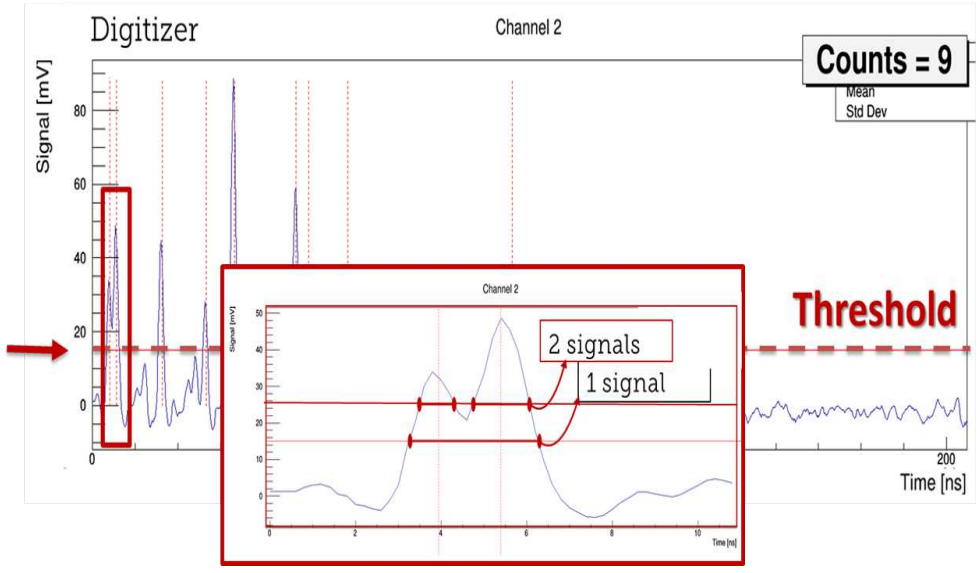

Fig. 3.17 Example of a 200 ns time window collected with the digitizer. The arrow points to a peak with pileup effect and the rectangular box zooms-in the area to underline the importance of the threshold selection in pileup situations.

In order to optimize the signal discrimination, a threshold scan technique has been used on data collected with various sensors. Furthermore dynamic algorithms

have been adopted for peak counting with specific effort dedicated to overlapped signals. The signal amplitude distribution reported on the right side of Figure 3.18, has been obtained deriving the peak-rate distribution reported on the left side of Figure 3.18 as a function of the discrimination threshold.

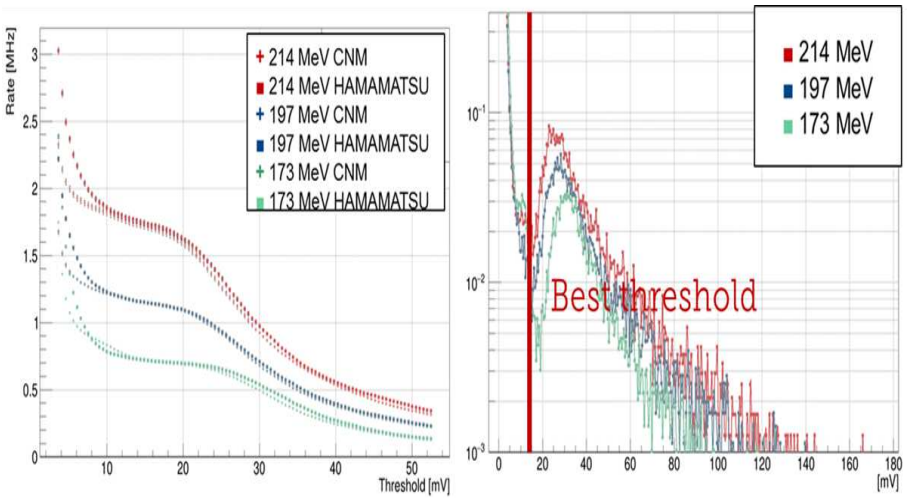

Fig. 3.18 On the left: rate versus threshold for three different energies, for both sensors (C = CNM at 250 V, H = Hamamatsu at 190 V). For low values of the threshold, a high contribution of the noise is clearly evident, while for high values of the threshold there is a significant loss of the signal. On the right: amplitude distribution of the signals for three different energies for the CNM sensor, in which the vertical scale is given in arbitrary units

As previously mentioned, the Landau fluctuations regulates the statistics of the the charge released into the silicon; at low beam energy the stopping power is higher, as confirmed by the MPV in 3.18. The signal-to-noise ratio can be deduced by the separation of this two therms, highlighted in 3.18 and this separation can be even increased using higher bias voltages. A PinPoint dosimeter [22] positioned behind the detectors has been used as charge reference for the efficiency estimation reported in 3.19.

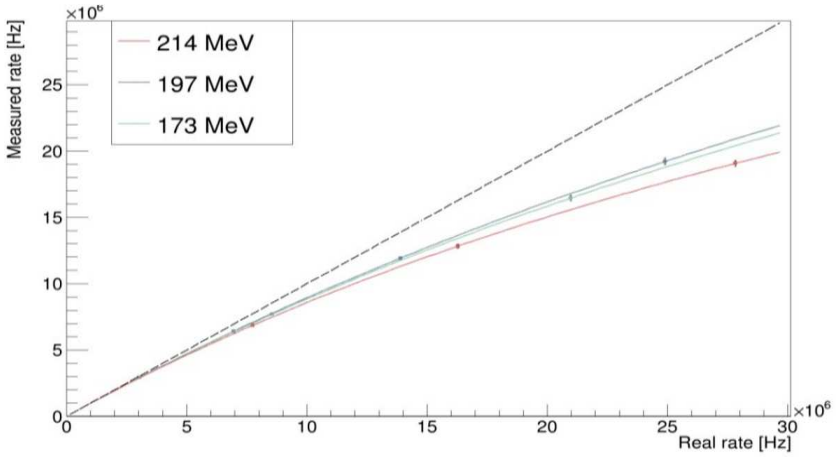

Fig. 3.19 Measured ion rate vs the expected one. The lack of efficiency is related to the fact that the particle beam distribution has been considered as Poissonian while, for the CNAO synchrotron, it is RF bunched with an istantaneous peak intensity higher than expected.

From this plot, a linearity deviation is marked starting from 5 MHz. The data were fitted with pile-up models assuming a Poisson distribution for the arrival times and a fixed acquisition dead time [29]. The models were found to describe the results with an acquisition dead time of about 10 ns, i.e. much larger than the 2 ns pulse duration observed in Figure 3.16. The reason of this discrepancy originates from the highly non-uniform time structure of the CNAO beam, as showed in figure 3.20, where a longer time window was acquired on the oscilloscope.

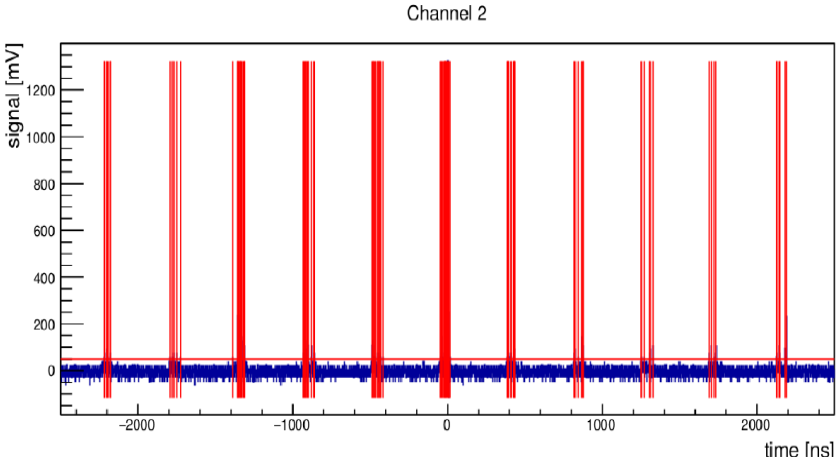

Fig. 3.20 The CNAO particle beam structure. The protons are collected in bunches distributed with a radio frequency given by the synchrotron. The instantaneous intensity reaches $10^{10} p$ cm^{-2} s^{-1}.

A bunched structure with a frequency of few MHz was observed, probably originating from the radio-frequency of the CNAO synchrotron acceleration system. Since the instantaneous intensity reached $10^{10} pcm^{-2}s^{-1}$, this caused a larger pile-up probability, if compared with the one related to a uniform time distributed beam. Acquiring the data in parallel from two sensors, it has been possible to appreciate the signal correlation; the result of this last data analysis is reported in Figure 3.21.

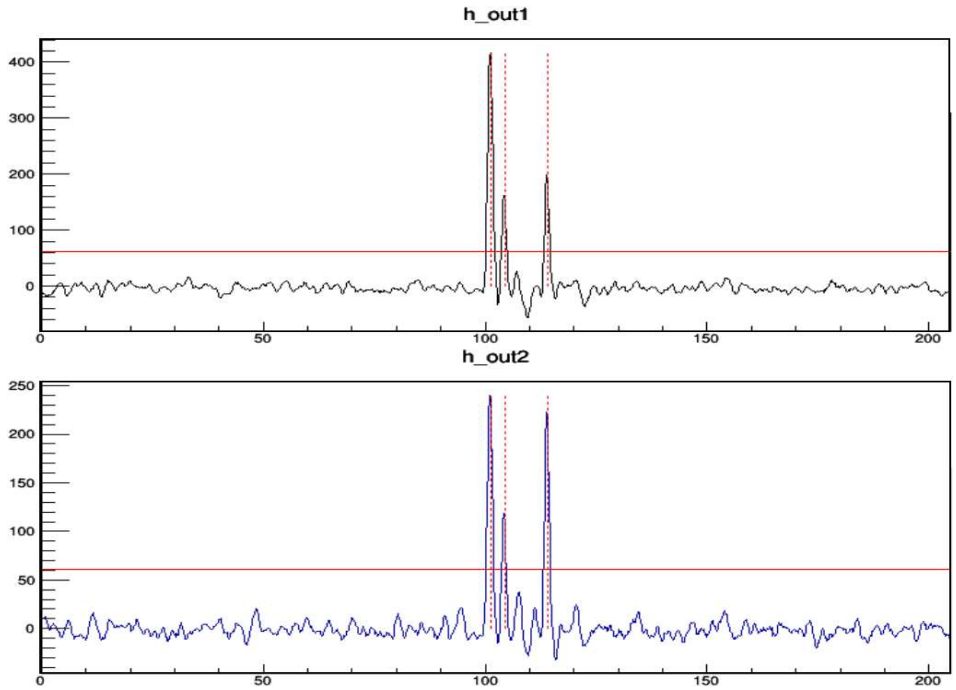

Fig. 3.21 Digitizer shapshot for two different UFSD pads detecting the same signal, thanks to a telescope setup. The simultaneous acquisition by two digitizer channels allows to appreciate the signal coincidences.

In order to reduce the time walk effect related to the fact that large amplitude signals are detected earlier than small signals (amplitude dependent slew rate), the Constant fraction discrimination technique has been used calculating the time difference distribution between two corresponding signals. This method leads to a 50 ps time difference resolution that means $\sim$35 ps for a single pad ($\frac{50}{\sqrt{2}}$).
Another critical aspect related to these sensors is the radiation resistance performances. For this reason, the distribution of the signal amplitude for the same sensor was compared before and after 32 runs, corresponding to about 10^{12} $p\ cm^{-2}$. As reported in Figure 3.22, the output pulse shows a gain loss of $\sim$ 20% between the

two distributions. Similar effects responsible for a loss of gain after irradiation of the sensors were observed and reported in [23].

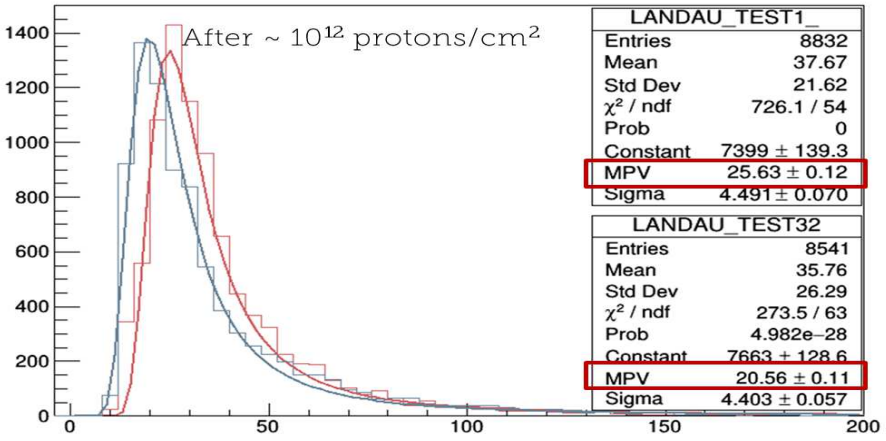

Fig. 3.22 The MPV value of the amplitude distribution passed from 25 10^{-12} Vs before irradiation to 20 10^{-12} Vs after 32 runs of proton irradiation

As explained in the previous dedicated section, in order to address the radiation damage problems that came out form the tests at CNAO, new sensors have been produced in a run called UFSD2, which presents different dopant species (B, Ga, and C spray) and concentrations [24]; Table 3.3 shows a summary of the 18 new produced wafers with the dopant type and concentrations.
Further test campaign are required to estimate the radiation hardness for the new UFSD productions.
The high resolution obtained with particle counting up to 10^9 p/cm^2s fluxes, confirm the effectiveness of a multi-channel ASIC-based systems that embedes amplification and discrimination stages, in order to cover several square centimeters of detector area with high spatial and time resolution.

3.11 UFSD for MoVeIT

On the base of the beam test results obtained with silicon pads, the group decided the features for the strip-segmented sensors, to be developed in the next UFSD production. These strips have been produced by the Trento (Italy) FBK (Fondazione

Table 3.3 Summary of the UFSD2 sensors production. From the left to the right: the wafer number from 1 to 18, the dopant element, the gain dose, the presence of carbon spray and the diffusion type.

Wafer number	**Dopant**	**Gain dose**	**Carbon**	**Diffusion**
1	Boron	2.45	no	low
2	Boron	2.50	no	high
3	Boron	2.50	no	high
4	Boron	2.50	yes	high
5	Boron	2.50	yes	high
6	Boron	2.55	yes	high
7	Boron	2.55	yes	high
8	Boron	2.55	no	high
9	Boron	2.55	no	high
10	Boron	2.60	no	high
11	Gallium	2.50	no	low
12	Gallium	2.50	no	low
13	Gallium	2.60	no	low
14	Gallium	2.60	no	low
15	Gallium	2.60	yes	low
16	Gallium	2.60	yes	low
17	Gallium	2.27	no	low
18	Gallium	2.27	no	low

Bruno Kessler) research center. Two geometries of 50 μm thick strips, are currently under test:

30 strips 30 mm · 0.08 mm, 146 μm pitch

20 strips 15 mm · 0.15 mm, 216 μm pitch

Figure 3.23 shows the gdsII file with the UFSD2 wafer design whereas, Figure 3.24 is a wafer picture.

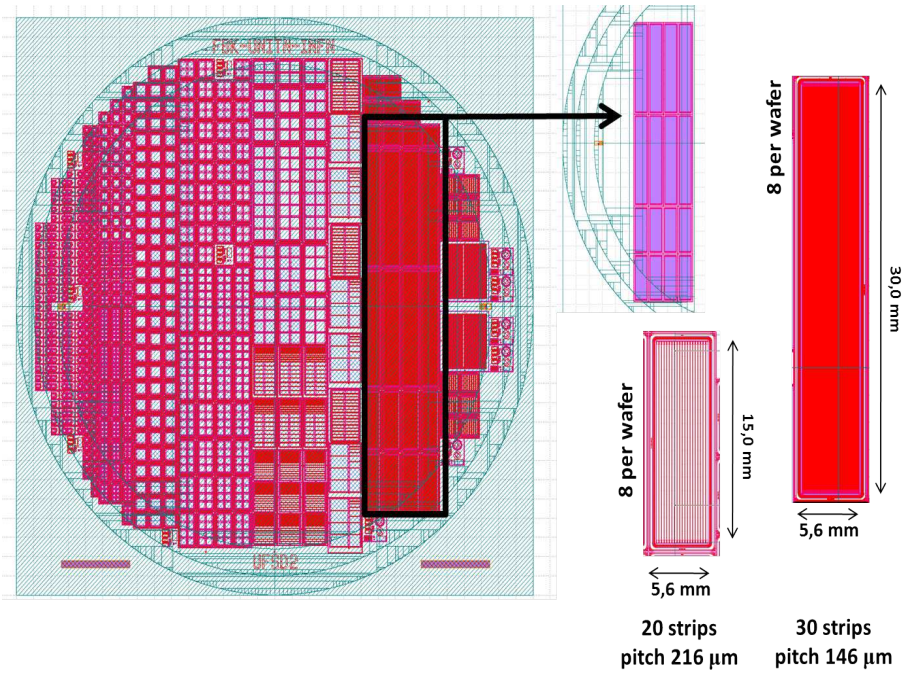

Fig. 3.23 UFSD strips for MoVeIT. On the left, the gerber file of the UFSD2 production wafer. On the right, the zoom-in on the MoVeIT sensors: set of 20 short strips and set of 30 long strips.

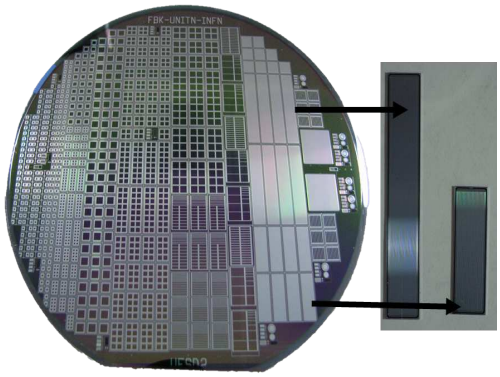

Fig. 3.24 Picture of the MoVeIT strips. On the left, the wafer from the UFSD2 production. On the right, the short and long type strips.

An extensive sensor test campaign is still ongoing, in order to characterize the strips with current vs voltage and capacitance vs voltage curves. With hundreds of sensors available and with the aim to speed up the work, the testing procedure has been automated by mean of a needle-based probe-card, with a software driven switch for the channels control. The test setup is shown in Figure 3.25.

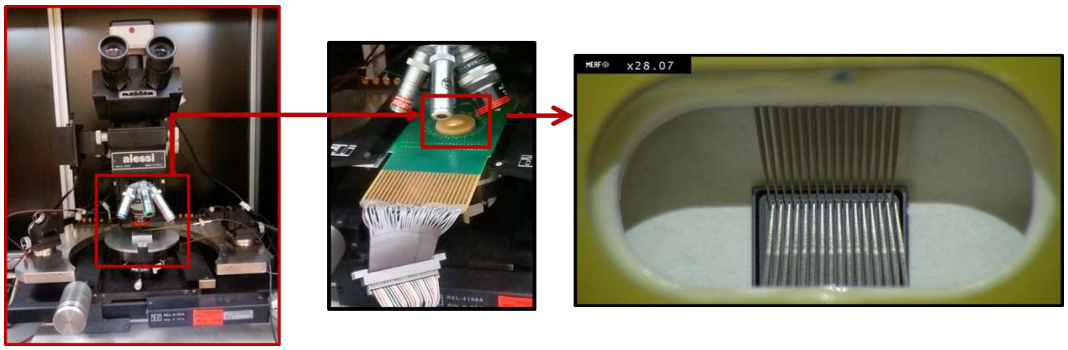

Fig. 3.25 Test setup photo. From left to right: the microscope station, the probe card with the connector for the needles control, the needles zoom-in.

An interesting test performed in laboratory consisted in shooting with a laser source from the sensor edge, thus to cross a couple of strips; analyzing the data of the collected charge, it has been possible to obtain informations about the sensor gain profile. Although the details of this procedure lie outside the interest of this thesis, the qualitative aspects of these results are useful from a front-emd design point of view. Figure 3.26 represents a 2-D chromatic spatial distribution of the gain for two strips and a bias voltage of 230 V. Figure 3.27 represents the gain profile as

signal amplitude in millivolts, vs the material depth in μm; in the same picture, for better understanding is reported and overlapped to the previus result the same kind of profile for gain-less sensors.

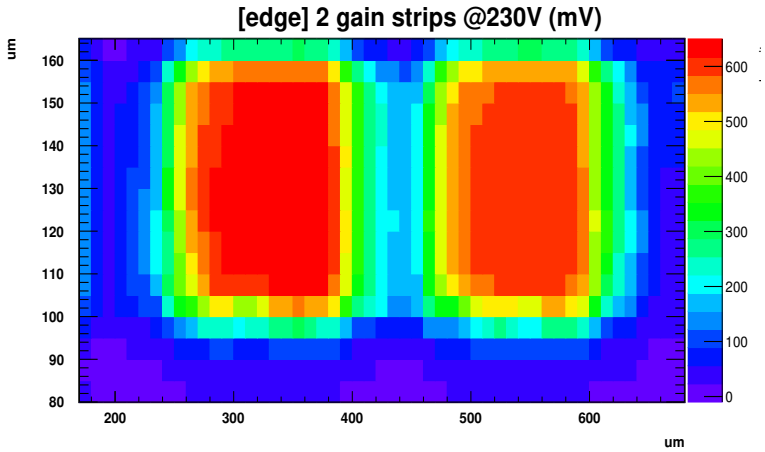

Fig. 3.26 Chromatic representation of the 2-D spatial distribution of the gain for a couple of strips biased at 230 V. The strips collected the charge as result of a laser shooting through the sensor edge.

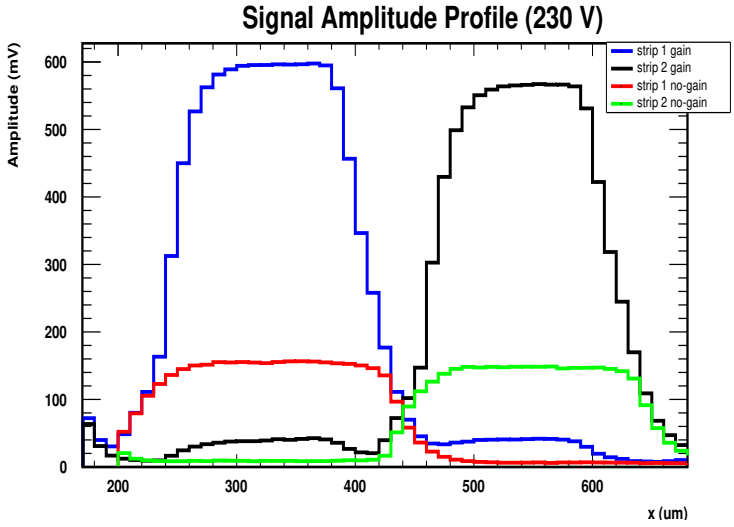

Fig. 3.27 Signal amplitude vs depth for a couple of strips biased at 230 V. The strips collected the charge as result of a laser shooting through the sensor edge.

The behavior of the strips gain in function of the bias voltage, is shown in Figures 3.28.

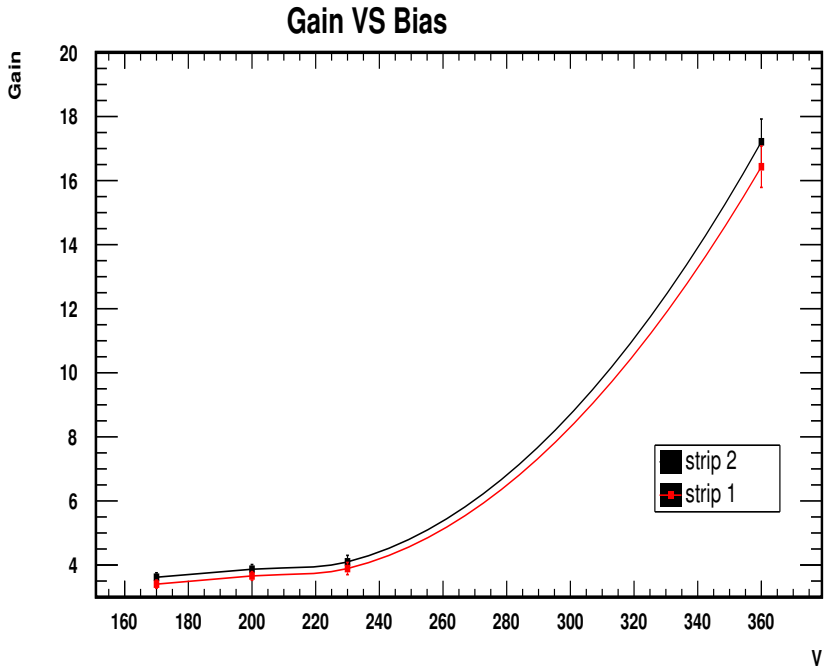

Fig. 3.28 Strip gain in function of the voltage, for a couple strips belonging to the same sensor.

All the the above reported results are related to short strips form the wafer 8, in the UFSD2 production.

3.12 The ABACUS design

3.12.1 Introduction

UFSD sensors have been chosen for the MoVeIT project thanks to their fast charge collection with a reasonable signal amplitude. As previously described, the MoVeIT front-end electronics has to deal with a nanosecond level rise time with an additional challenge represented by the 10^8Hz signal rate (100 MHz as milestone, 250 MHz the ideal R&D target). In this condition it is no more obvious that the amplifier output returns to the baseline before a new pulse can be processed; this undesired aspect could lead to signal pile-up. The quite extreme condition of a therapeutic particle beam (fluxes up to 10^9 protons $cm^{-2}s^{-1}$), considered the requirement of a pile-up kept $\leq 1\%$, push the detector designers to develop front-end electronics with

faster return to baseline supported by a high sensor granularity (to reduce the single channel event rate).
In order to equip the MoVeIT single ion discriminator prototype with a custom electronics, our group decided to explore the design of two different architectures in parallel: one based on a resistive feedback and the second one, based on a capacitive feedback. The two alternative solutions will be characterized both standalone and coupled with the UFSDs. The prototype resulting as more efficient, will be upgraded in a second and complete chip embedding the digital logic of counters and registers together with the analog part (as the TERA ASIC family does [[31]]).

The resistive feedback design is based on a differential trans-impedance amplifier (TIA) adopting a differential cascode structure to achieve high trans-impedance gain, high bandwidth, and low input referred noise, with good Power Supply Rejection Ratio (PSRR) and Common Mode Rejection Ratio (CMRR). Albeit less sensitive to low input charges with respect to the integrating architecture, the TIA approach can ideally easily deal with the time limits given by the detector. The TIA channel presents a preamplifier with a relatively low open loop gain (around 10), due to the low load resistor required to have high speed. Two stages have been connected in cascade in order to increase the open loop gain and thus decrease the effect of the input capacitance. Moreover a quite high closed loop gain is needed to keep low the noise contribution of the following stages.
This chapter focuses on the capacitive feedback prototype, where the charge is integrated following the Charge Sensitive Amplifier (CSA) concept. This specific CSA has a NMOS input telescopic cascode common source preamplifier, equipped two independent branches with degeneration resistors for noise reduction. The signal is discriminated by a two stages comparator and fed to a differential CML driver. Considering the 10^8 input signal rate, the preamplifier output must manage high slew rate for both signal rising and falling edges, in order to speed up the detection and be ready to receive a close following pulse. With the aim to manage the signal overlapping and pile-up phenomena while keeping the baseline stable, the circuit has a balanced switch-reset for the feedback capacitance, activated as discrimination feedback.
For both TIA and CSA based solutions, a 10 mA order bias current is required to feed the input transistors and achieve high cut-off frequency and low noise. Such a high current translates in high power consumption (which is not critical in this application) and high input parasitic capacitance. The cascode structure has been widely used to

eliminate the effect of Miller capacitance and thus enhance the bandwidth.
The design of two ASIC has been submitted on January 22nd, 2018, as an Europractice Multi Project Wafer ([32] MPW) which tape-out, is planned for the Beginning of May 2018.
The intent of this thesis section is to describe the technical features and the adopted design choices for the CSA based device, called ABACUS. ABACUS stands for Asynchronous Based Analog Counter for Ultra fast Silicon strips.

3.12.2 ASIC overview

ABACUS is a fast single ion discriminator integrating 24 independent channels, designed in a standard 110 nm CMOS technology, biased at 1.2 V. This number of channel allows to study an fairly complex system, coupling an entire set of MoVeIT UFSD strips (20 strips). The technology node choice derives from the design group experience gained in the last years design submissions. Moreover, we estimated that the 110 nm technology node is suitable in terms of performances, transconductance parameter and noise levels, considering the coupling with UFSD type sensors. A further support has been the know-how that the Turin microelectronics and the UFSD groups gained with the TOFFEE ASIC [33]. TOFFEE is a 8 channel preamplifier-discriminator-driver ASIC designed for timing measure with UFSD pads in high energy physics experiments (CERN CMS-TOTEM Precision Proton Spectrometer). Although the TOFFEE amplifier features a t_r $\sim$1.5 ns, the R-C signal shaping is in the order of 10-15 ns. Moreover the charge input range properly managed by TOFFEE is 3-20 fC. For there reasons this chip could not be adopted for the MoVeIT task; nevertheless, studying the behavior and testing the design choices with a physical chip helped to check the reliability of some technical decisions and allowed to understand the limits in order to overcome or mitigate them.
In a silicon area of $2 \cdot 5$ mm^2, ABACUS integrates 24 channels and each channel is equipped with the following functional blocks:
CSA preamplifier
OTA buffer
A two stages discriminator
Single ended to differential converter
Current Mode Logic driver
6-bit Digital to Analog Converter

Pulser
Recovery circuit
Signal reshaping inverters.

In Figure 3.29 the ABACUS channel is represented in its functional blocks view. The main concept that characterizes the channel logic is the discrimination-triggered reset of the feedback capacitance. The high input pulse rate is managed forcing to the baseline level the falling edge of the CSA out, every time the the discriminator threshold is crossed. The 500 fF feedback capacitance (C_f) fixes the amplifier gain while influencing the feedback stability. This value for C_f is responsible to a large R-C therefore, the reset operation is intended as the fastest way to keep ready the preamplifier in following the next pulse that can be 4-5 ns close in time. This choice unfortunately introduces a strong perturbation of the system that has to deal with high frequency reset signal (peaked signals) and an immediate configuration swap for the shorted feedback. The reset action tends to generate an undershoot of the signal respect to the baseline and the designer has to control even signal overshoot for the NMOS switch-reset charge injection. It is moreover true that resetting the CSA output signal, the device loses the correspondence between the input signal amplitude and the digitized signal one. Even if the purpose of this prototype is only the signal discrimination for peak counting, pile-up correction algorithms will most probably be used during the data analysis. Currently it is expected to clearly distinguish from the digitized signal duration if overlapped signals have been detected or not but most probably it would not be possible to extrapolate the number of overlapped pulses: a typical CML outut has a time duration of 1.2-1.6 ns whereas in case of overlapping, this is going to be longer but with the reset action, the signal duration vs number of peaks proportionally is lost.

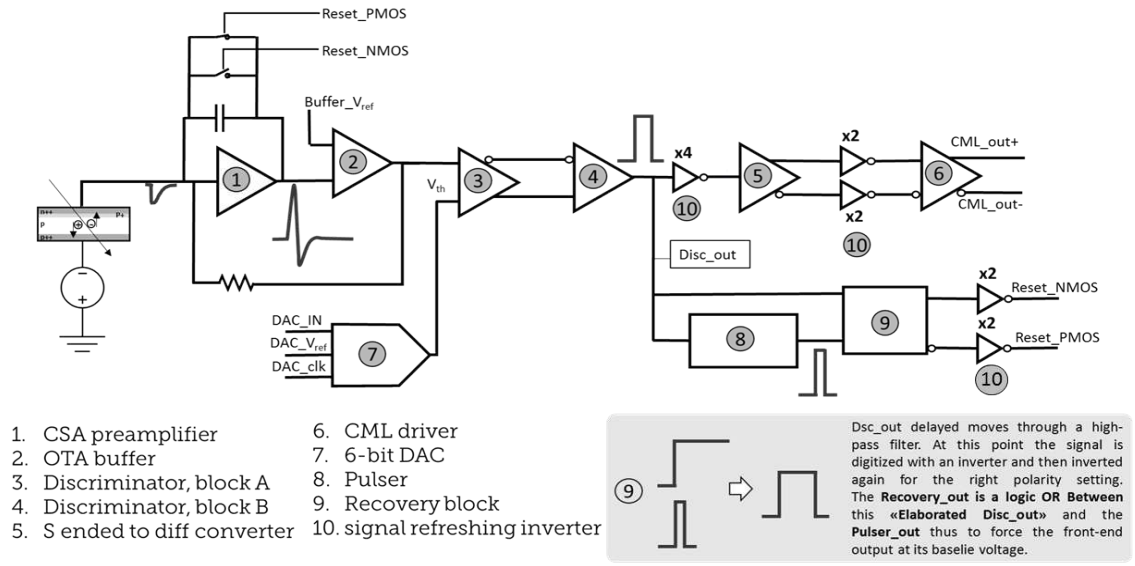

Fig. 3.29 Functional block representation of the ABACUS channel.

3.12.3 Preamplifier

Like most of the radiation and particle sensors, UFSDs return single polarity signals and this has been the starting point for a single-ended front-end design. Due to the fact that ABACUS has to deal with high repetition rates and small charges, the intrinsic amplifier noise plays a relevant role in the discrimination efficiency; this was an additional reason in choosing single-ended stages since they should offer a better noise figure than differential ones [34]. Once fixed the architecture type, the input transistor modeling is crucial, especially in a charge sensitive amplifier. The challenging aspect is due to the fact that bigger transistors implies large transconductance and lower spectral density for the 1/f noise. At the same time a large MOSFET contributes more on the input capacitance, worsen the thermal noise component. For a 110 nm technology and when the thermal noise is reduced by a long shaping time (like the CSA case), PMOS transistors would be preferable because of their intrinsically lower $1/f$ noise. Although this is true, the fact that in modern MOSFETs electrons have 3-4 times the holes mobility makes NMOS preferably preferable as preamplifier input transistors whereas the PMOS are generally used as active loads. For a fixed power budget NMOSs offer an higher g_m because of its higher mobility and the fact that the inversion coefficient is inversely proportional to it. Although in weak inversion there is no distinction anymore between PMOS and NMOS, since

the g_m depends only on the bias current, an NMOS works easier in weak inversion whereas a PMOS works easier in strong inversion. It is thus possible to say that a NMOS is preferable whenever a preamplifier requires a high input transconductance. Another important advantage offered by the modern deep sub-micron technologies is the fact that both PMOS and NMOS transistors can be isolated from substrate pick-up noise due to the sharp digital switching: the insulated n-well approach is used for the PMOS shielding whereas deep n-wells are currently available and they are used to guarantee a shield effect on NMOSs.
In Figure 3.30 the schematic representation of the ABACUS preamplifier is shown. At functional level the block is a charge sensitive amplifier which consists of a telescopic cascode with two independent branches, where the NM1 NMOS input transistor is a common source with split current bias.

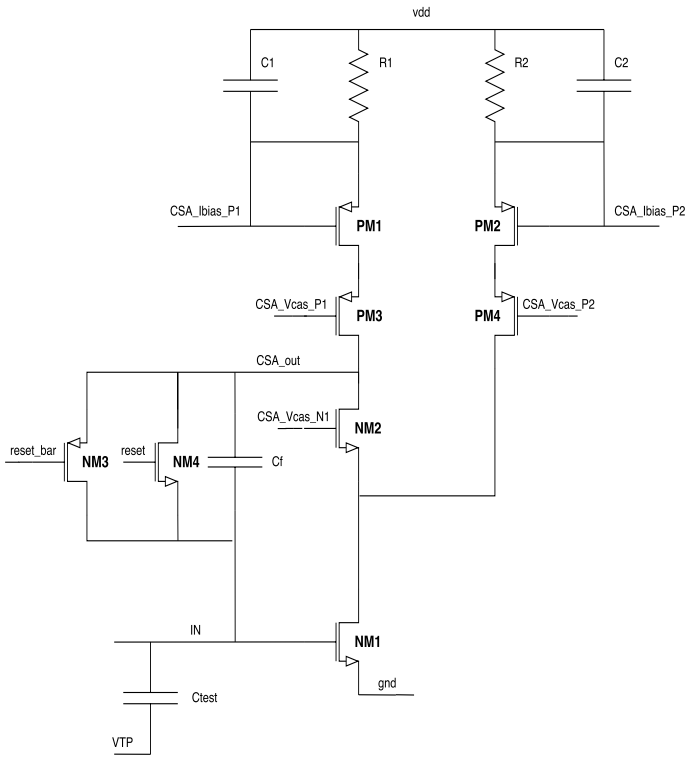

Fig. 3.30 Schematic representation of the ABACUS pramplifier, a telescopic cascode amplifier with split bias current and NMOS input transistor, working as CSA.

The cascode transistors NM2 and PM3 are adopted to increase the output resistance of the cascoded NM1 and PM1 respectively (the current source). The open loop gain is represented by the product between the transconductance g_m and the

drain-source resistance of the cascoding device. A second independent branch is used to increase the open-loop gain. The left branch biases the cascodes PM3 and NM2 ($CSAI_{biasP1} \sim 200\mu$A); increasing the value of this current increases the output voltage slew rate but at the same time decreases the swing. The the most part of the CSA bias current current feeds NM1 transistor flowing through the right branch ($CSAI_{biasP2} \sim 10mA$) and is provided by an additional and completely independent current source (PM2 and PM4 transistors). This $\sim 10mA$ current is required to have a large input transconductance and speed-up the front-end response.
The drain voltage of the input transistor is rather stable due to the cascode configuration. This allows to have a $CSAV_{casP2}$ lower than $CSAV_{casP1}$ thus to free more the PM2 dynamic, increasing its output impedance. Moreover since ABACUS is a prototype, it is useful to have the possibility to tune both branches independently, for instance to mitigate the temperature variation effect and control the baseline (described in the ABACUS layout section).
In the upper part of Figure 3.30 it is possible to notice that source degeneration resistors have been used for noise reduction, performing an effect on the node between the CSA branches:

$$g_{m3}r_{ds3} \cdot g_{m4}r_{ds4} \cdot R_1, \tag{3.19}$$

$$g_{m2}r_{ds2} \cdot (r_{ds1}//g_{m_5}r_{ds5} \cdot g_{m6}r_{ds6} \cdot R_2) \tag{3.20}$$

The open-loop gain of the ABACUS CSA is obtained by the equation:

$$A_0 = -g_{mNM1} \cdot B \approx 15. \tag{3.21}$$

Where

$$B = (C \cdot D)//E \tag{3.22}$$

and $C = g_{mNM2} \cdot R_{NM2}$;
$D = R_{NM1}//[g_{mPM4} \cdot R_{PM4} \cdot g_{mPM2} \cdot R_{PM2} \cdot R2]$;
$E = g_{mPM3} \cdot R_{PM3} \cdot g_{mPM1} \cdot R_{PM1} \cdot R1$
for notation simplicity the drain-source resistances are here named R.

The value of the current in transistor NM1, its transconductance g_{mPM1} and dimensions have been chosen to cope with the short collection time of the sensor and to manage the trade-off between input charge range and rate. If it is true that large gain would reduce time walk effects and would facilitate the signal discrimination for small charges, on the other hand high rate instability issues can arise with large signals. From Figure 3.31 it is possible to see that the -3dB gain is 23.6 dB at 60 MHz leading to a gain-bandwidth product of 900 MHz.

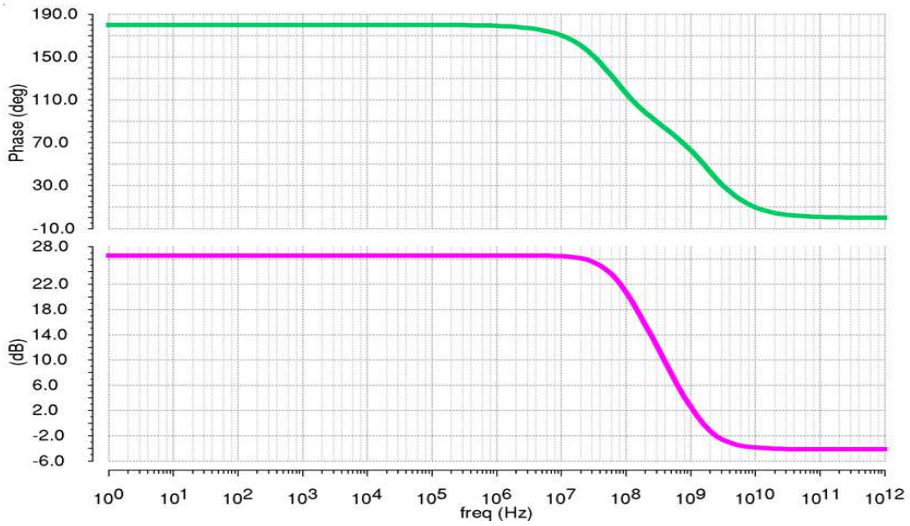

Fig. 3.31 Bode diagrams for the CSA representing the phase vs frequency (up) and the gain module in dB vs frequency (down).

In the ABACUS channel, by setting a high threshold and thus blinding the discriminator it is possible to avoid the reset of the signal capacitance. In this way the amplifier gain linearity can be studied by means of a charge sweep analysis, as reported in Figure 3.32.

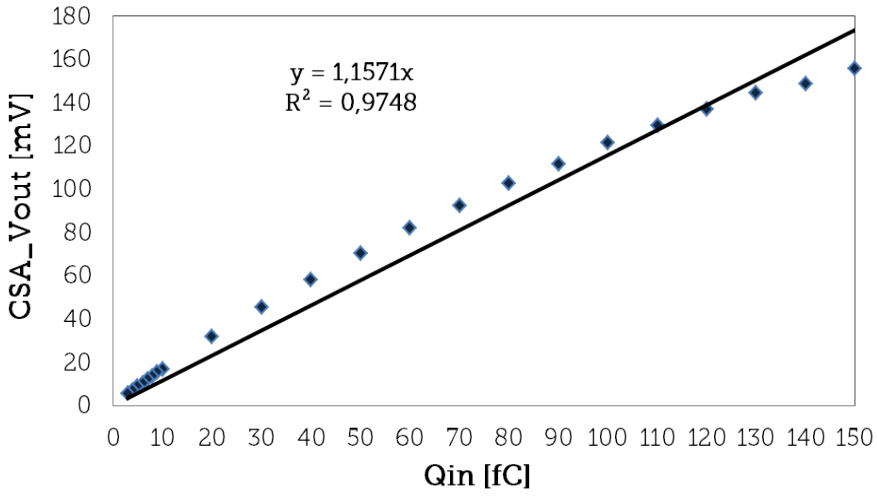

Fig. 3.32 Gain linearity performance of the CSA, in a 3 fC - 150 fC input charge range (schematic simulation).

0

More information regarding the CSA impulse response have been achieved with the same high threshold technique. Results are reported in Figures 3.33, 3.34 where the output voltage of the OTA Buffer (component described in the following subsection) and the CSA with a single MIP equivalent input pulse (8 fC), and with different input small charges (3 fC, 5 fC, 7 fC, 9 fC) are shown, respectively.

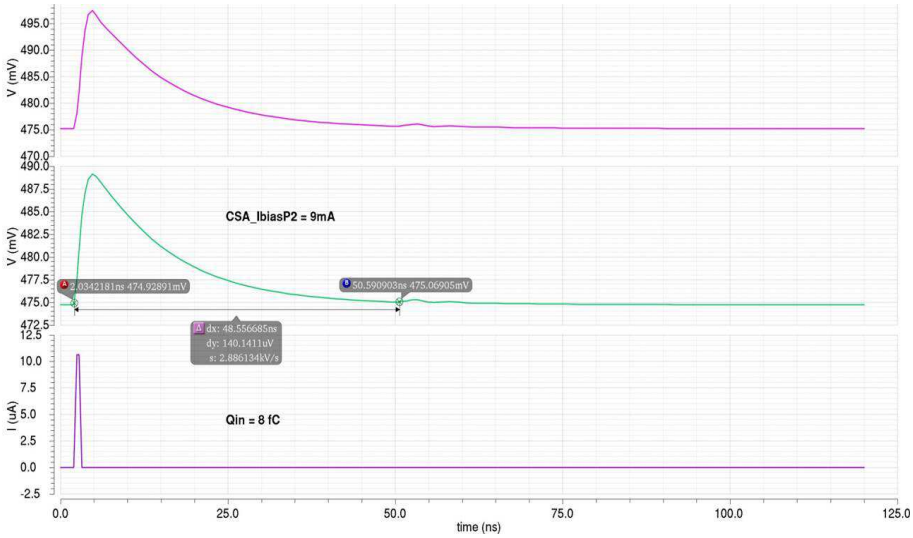

Fig. 3.33 RC shaping of the CSA for a MIP equivalent signal (8 fC in a 50 μm UFSD). From the top to the bottom: OTA Buffer output voltage, CSA output voltage, input signal (layout simulation).

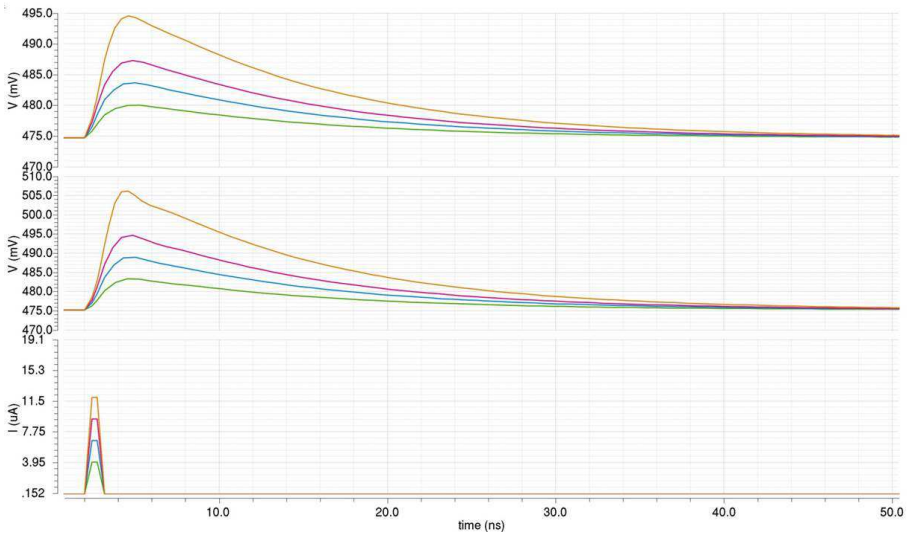

Fig. 3.34 RC shaping of the CSA for 4 different signals: 3 fC, 5 fC, 7 fC, 9 fC. From the top to the bottom: OTA Buffer output voltage, CSA output voltage, input signal (layout simulation).

The effect of a proper setting of the discriminator threshold is visible on Figure 3.35, where the CSA output voltage has a sharp cut induced by the discrimination activated reset. The small ripple with the undershoot is the effect of the switch-reset NMOS charge-injection, slightly emerging even after its mitigation with a properly sized compensation-PMOS.

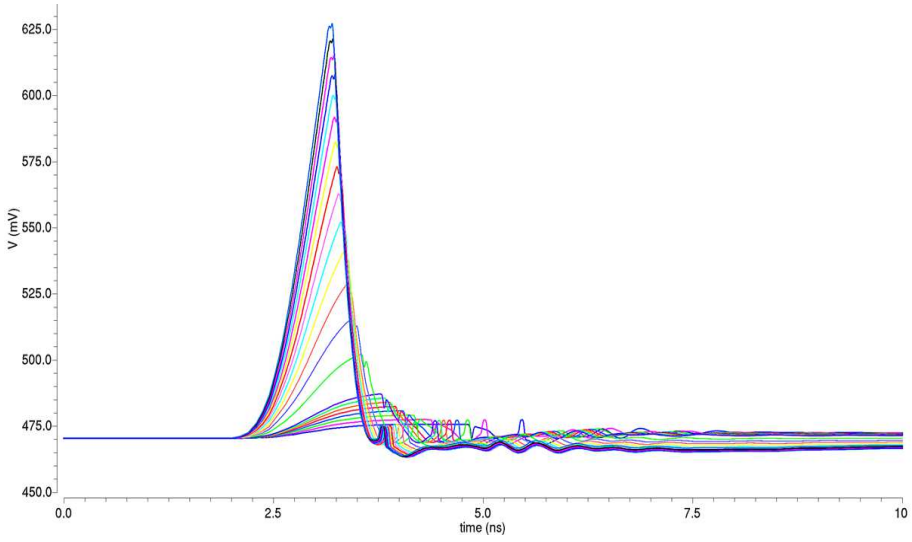

Fig. 3.35 Front-end shaping for a 3 fC - 150 fC input charge range. The sharp falling edge is the result of the discrimination triggered reset (schematic simulation).

CSA transient noise analysis Transient noise analysis have been performed to estimate the degradation effect in terms of the front-end output signal shape. This procedure consists in repeating the injection of a single pulse input signal to which the CAD applies a certain noise frequency component selected from a user specified interval. The effects of a transient noise analysis on the front-end output signals, performed in the 1 mHz-10 GHz frequency range for the ABACUS charge boundary values (3fC-150fC) have been reported in Figures 3.36 and 3.37. Due to the presence of noise the repeated input signal does not result in preamplifier output signals perfectly overlapped and the rms noise represents this effect. This information is obtained assuming a normal distribution of the events and calculating the standard deviation:

$$V_{std} = \sqrt{\frac{1}{N-1} \cdot \sum_{k=1}^{N} (V_k - V_{av})^2} \tag{3.23}$$

where V_k is the amplitude of the k^{th} pulse and V_{av} is the mean value.

In particle detector the terms Equivalent Noise Charge (ENC) is referred to the amplifier input and is calculated dividing the rms noise by the gain. Considering again the ABACUS case, a schematic test-bench setup with a 5 pF detector capacitance results in $\sim$370 mV rms noise; with a preamplifier gain of 1.15 mV/fC (from Figure 3.32), $\frac{0.37mV}{1.15mV/fC} = 0.323fC$ corresponding to an $ENC = 2020electrons$. From a deeper analysis it is possible to point out that the contribution of the input transistor amounts to more than 50% of the total CSA noise.

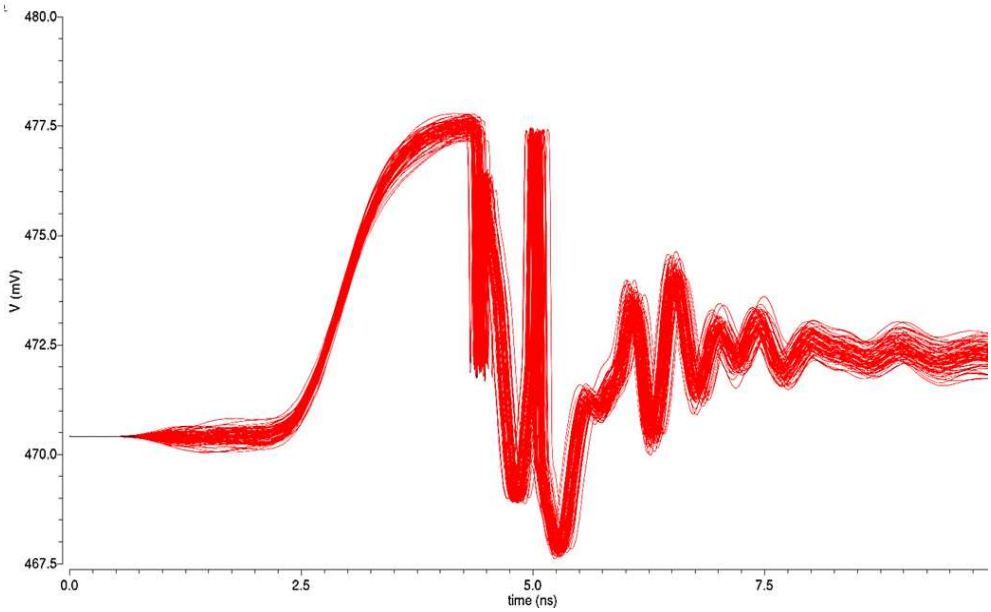

Fig. 3.36 Result from a 100 Hz - 100 GHz transient noise analysis. The noise effect has been observed on the CSA output, with a 3 fC input signal.

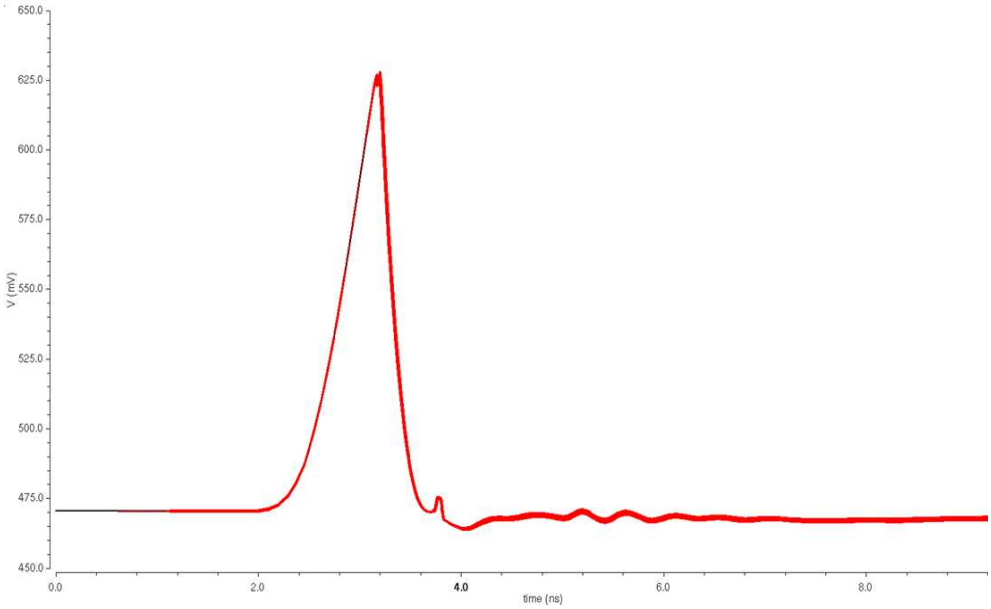

Fig. 3.37 Result from a 100 Hz - 100 GHz transient noise analysis. The noise effect has been observed on the CSA output, with a 150 fC input signal.

Defining Jitter the ratio between the rms noise and the slew rate (intended as the signal slope in its $10\% - 90\%$ interval), it is possible to appreciate the results

in Figure 3.38, where the Jitter vs Q_{in} is shown for both schematic and layout models of the preamplifier. Figure 3.39 offers a more complete understanding of the phenomena: while the rms noise is almost constant increasing the input charge, the slew rate increases enhancing the time resolution (i.e. the Jitter).

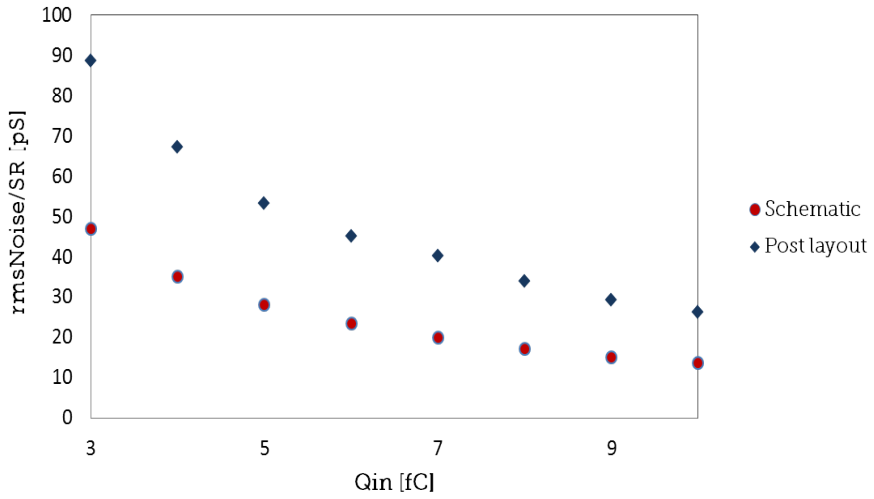

Fig. 3.38 CSA jitter vs input charge: comparison between schematic and post-layout simulation behavior.

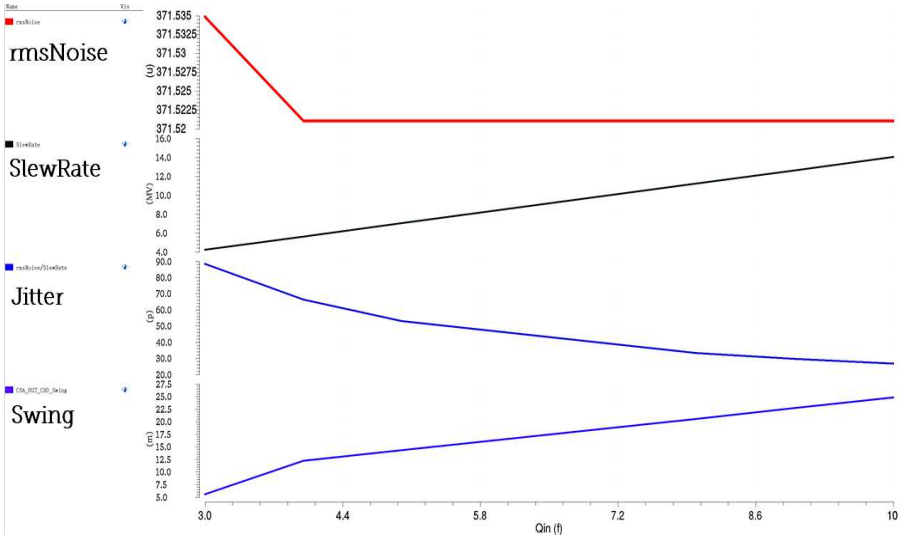

Fig. 3.39 CSA details for the low-charge range (layout simulation).

Another information regarding the Jitter is its inverse proportionality with the detector capacitance, as confirmed by the simulation result in Figure 3.40. In a

parallel planes structures like silicon detector the capacitance goes inversely with the thickness but it is not possible to use thicker detectors because it would increase the collection time. Another approach colud be to reduce the area of the sensor; this choice carries the drawback to increase the sensitive area with a granularity that is limited by mechanical and electrical constraints. High granularity detectors have an implicit level of complexity descending to the higher number of front-end channels, the power consumption and dissipation issues, dead areas etc..

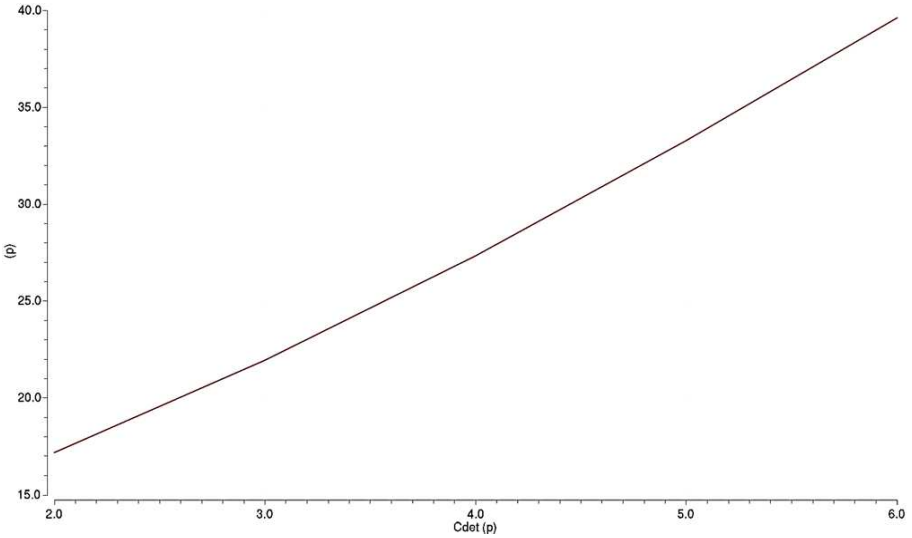

Fig. 3.40 CSA jitter vs the detector capacitance (layout simulation).

3.12.4 Buffer

As represented in Figure 3.29, the preamplifier feedback has the integration capacitance (C_f) connected between the CSA input and its output and then there is a a resistive connection (R_f) between the OTA Buffer and the CSA input. The R_f resistor is used to discharge C_f after a signal is detected and it provides a DC value at the preamplifier input thus keeping the proper NMOS input transistor operating point. The OTA Buffer placed between the amplifier and the discrimination stage is therefore used as impedance adapter and to protect the CSA input by a direct resistive feedback connection. Its architecture is based on a source follower with an active feedback, intended to present a low output impedance; this is suitable to deal with high frequency signal outputs.

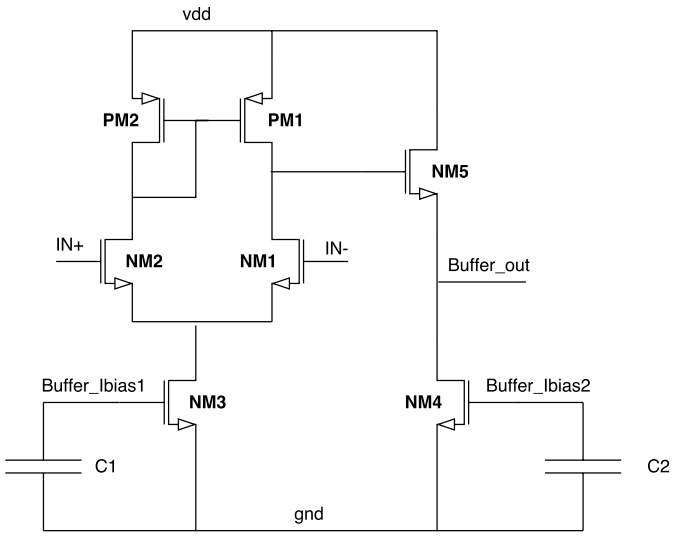

Fig. 3.41 Buffer pdf

3.12.5 Discriminator

The MoVeIT detector for single ion discriminator is based on a binary system concept: a comparator fires if the front-end output goes above a preset threshold, generating a digital pulse; no further information are recorded. A common practice adopted to achieve high speed discriminator is to design cascade architectures such that the gain is improved, being the product of the block gains (Figure 3.42).

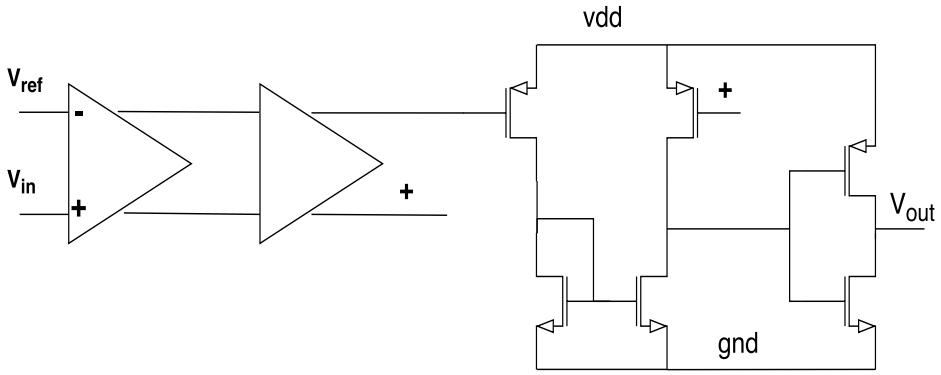

Fig. 3.42 To maximize the speed, a discriminator must be implemented cascading low-gain cells.

The ABACUS comparator implements as first stage a differential amplifier where the two input NMOS (NM1, NM2) are cascoded to mitigate the Miller capacitance

effect on their gate-drain connection (3.43). This stage increases the signal amplitude to the minimum value required to make a decision and isolates the input from the the second stage, avoiding to introduce the kickback noise from the discrimination switch. The load of this block is passive (high resistance polysilicon film) to minimize the parasitic capacitance that would be present with a PMOS load. The presence of two PMOSs diode connected and placed between the differential pair arms is intended to limit the voltage swing to boost the baseline recovery after detection [34]. Considering the wide amplitude range that ABACUS manages, this aspect has a critical impact over the system-level performance (as is explained later).

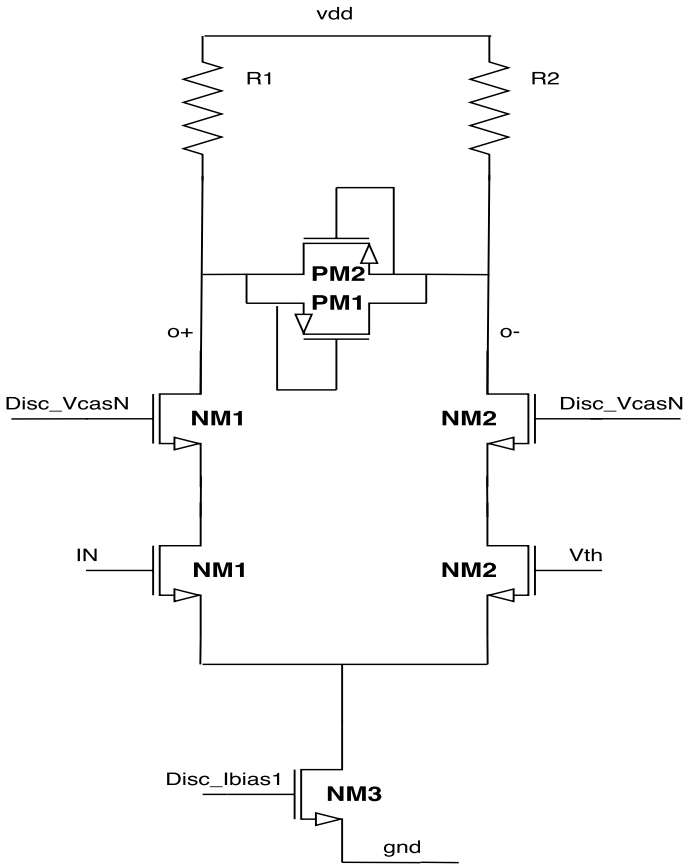

Fig. 3.43 Schematic representation of the comparator first block.

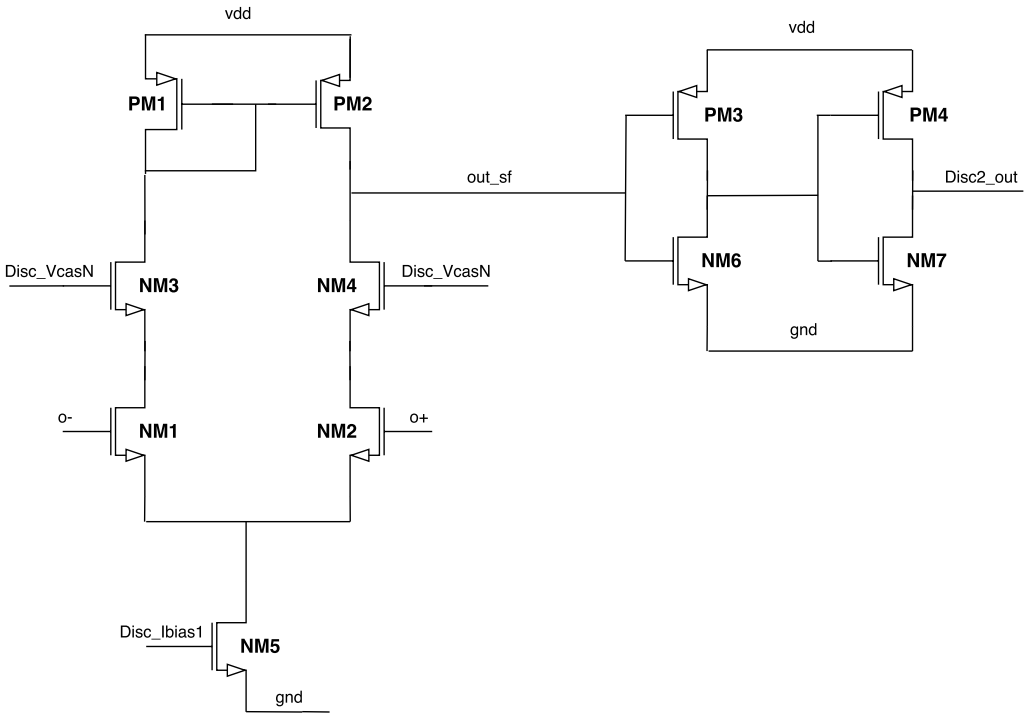

Fig. 3.44 Schematic representation of the comparator second block.

The second discriminating stage, Figure 3.44 provides a further signal amplification and performs the differential to single ended conversion. A couple of inverter buffers the output, increasing the impedance thus to properly drive the following block in the chain.

An important drawback of this configuration is the fact that it suffers from time walk effect. Figure 3.45 shows that even after the application specific optimization, the ABACUS comparator experiences simulation results with time walk induced discrimination delays up to $\sim$1.7 ns between the highest and the lowest charge that the ASIC should manage.

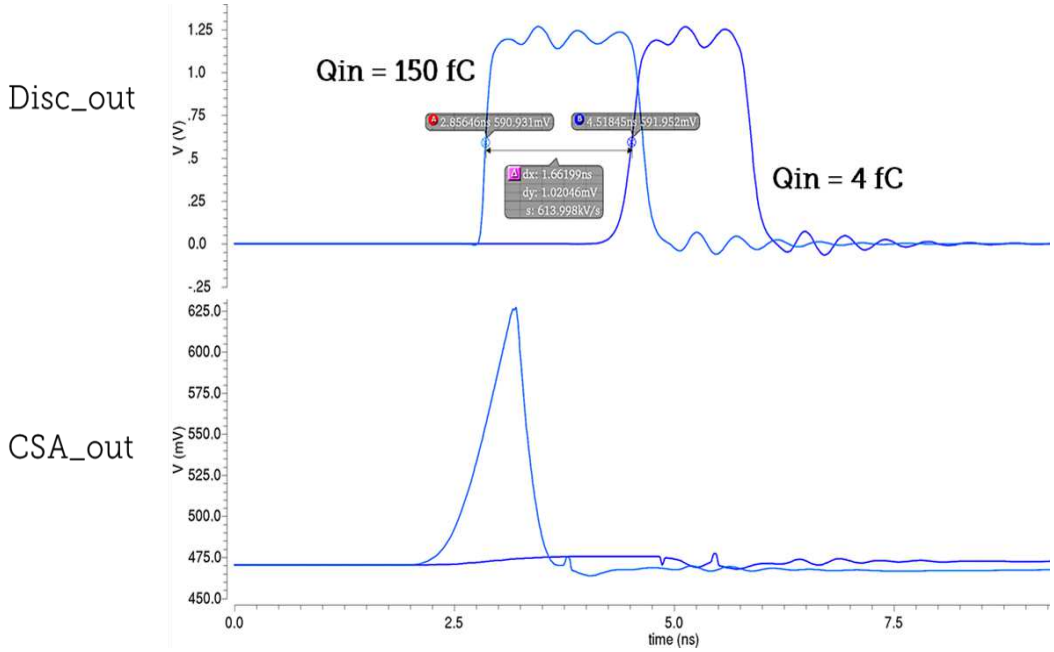

Fig. 3.45 timewalk

In the design of a multi-channel ASIC, a key point is the study of the threshold spread. ABACUS has a single ended front-end that is therefore more sensitive to mismatch effects than a fully-differential device and this aspect can be quantitatively estimated performing Monte Carlo analysis. Figure 3.46 represents the histogram result of a 200 runs Monte Carlo analysis over the complete ABACUS channel. This simulation shows that ABACUS has a threshold dispersion with $3\sigma \sim 20$ mV and for this reason ABACUS has channel-level embedded DACs for the threshold fine tuning.

For the sake of completeness, the Monte Carlo study highlighted as 95% the preamplifier influence the overall result and specifically, the preamplifier input transistor actively contributes as 25%.

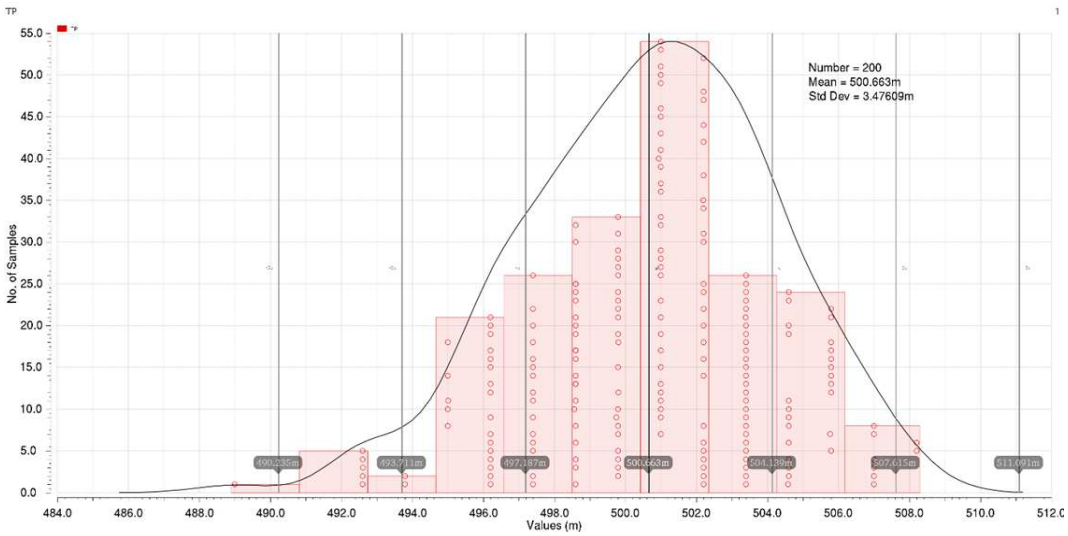

Fig. 3.46 Monte Carlo analysis for the discriminator threshold dispersion.

3.12.6 Local threshold DAC control

In a binary multi-channel system, the comparator threshold homogeneity among all the ASIC channels is a crucial point. ABACUS deals with charges extending from few femto-coulombs up to 150 femto coulombs. The UFSD sensor gain is limited (factor 10) to avoid the avalanche phenomena and control the signal collection while the front-end preamplifier open-loop gain is limited too (26 dB) to cope with the discriminator dynamic at high signal rates. In the low charge limit condition, the ABACUS comparator has to discriminate signal that are 5 mV over the baseline (as shows in Figure 3.34) and a sub millivolt threshold control is therefore mandatory. A common technique adopted to minimize the mismatch is to increase the dimensions of the discriminator input transistors but this approach is limited by the increasing of the parasitic capacitance. A more effective compensation method is to have a channel level threshold fine tuning employing a Digital to Analog Converter (DAC) through which an analog voltage can be programmed sending a digital code from outside. A chip level global threshold is therefore channel-by-channel properly tuned. The basic idea is depicted in Figure 3.47.

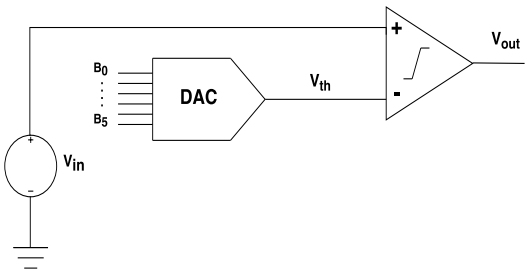

Fig. 3.47 Discriminator offset compensation with a fine tuning performed with an Digital to Analog Converter (DAC)

In a DAC, increasing the number of bits allows to achieve a higher resolution i.e. a lower value of the LSB:

$$V_{LSB} = \frac{V_{fs}}{s^n} \tag{3.24}$$

where n is the number of bits.

The DAC bit number is mainly limited by the available silicon area. The typical way to implement a DAC for threshold fine tuning is through an array of binary weighted current mirrors. In ABACUS the MonteCarlo theshold dispersion rms is $\sim$ 20 mV and a factor 2 is considered cautioning the on-silicon effect. The ABACUS DACs requirement is thus 6 bits over a full range scale of 40mV, centered on the nominal value ($\sim$475 mV).

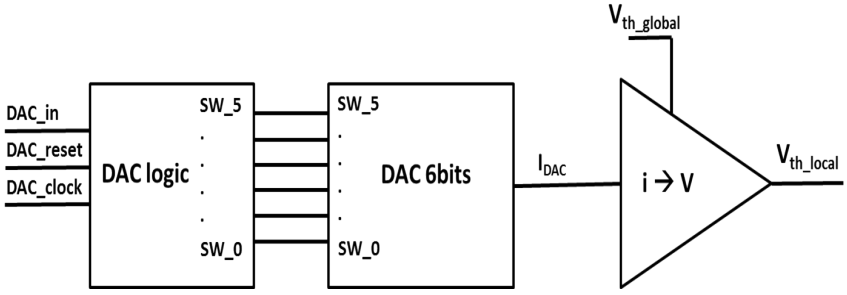

Fig. 3.48 ABACUS functional blocks dedicated to the digital-to-analog conversion. The V_{th_global} is fine tuned adding or subtracting a certain value selected by the user through a binary configuration.

Figure 3.48 represents the three components of the calibration DAC: the control logic containing six flip-flops implemented with digital standard cells (Figure 3.51),

the 64 unity current cells driven by the control logic (the signals are inverted to take into account that the current sources are made of PMOS transistors). Figure 3.49 represent the DAC bit control and Figure 3.50 is the inner bit logic. The third block is the linear current-to-voltage converter. The current in each DAC bit cell can be steered to a current-to-voltage converter, shown in in Figure 3.53. The generated voltage difference is fed to the comparator stage.

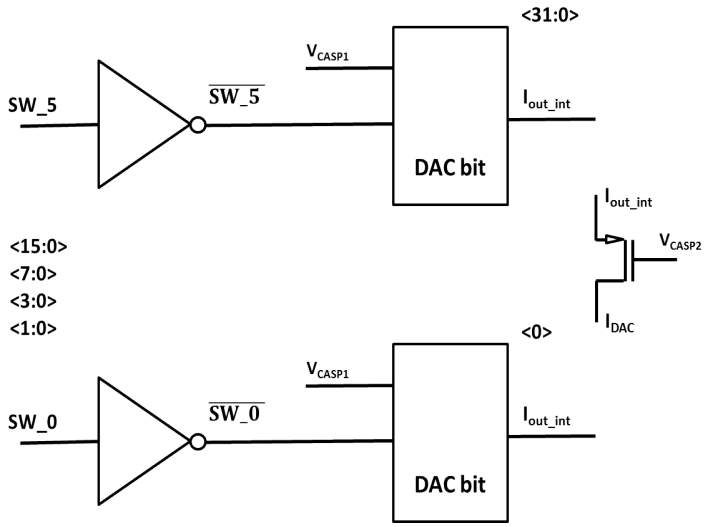

Fig. 3.49 Block representation of the DAC bit control. The inverters are used to keep an intuitive behavior with the stream of configuration bits: increasing the binary number means adding a positive value to the global threshold.

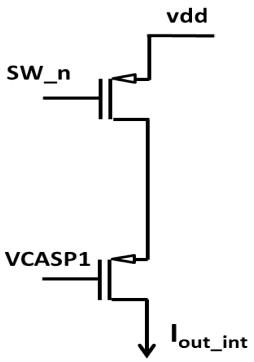

Fig. 3.50 Transistor detail of the unity current cell logic. The cascoded PMOS is controlled with the switch signal SW_n.

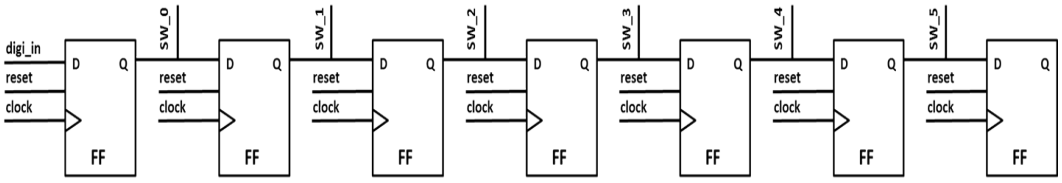

Fig. 3.51 DAC logic: block representation of the standard cell flip-flop chain and the extraction of the switch activation signals to control the bits.

In order to linearly convert i_{DAC} into the voltage threshold V_{th} sent to the comparator, a trans-impedance amplifier is used. The logic of this amplifier is represented in Figure 3.52 whereas Figure 3.53 shows the transistor level schematic.

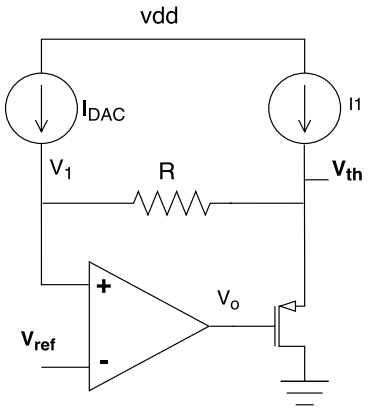

Fig. 3.52 Working principle of the adopted architecture for linear current to voltage conversion. The triangular block is the high gain transconductance amplifier.

The block is based on a single stage differential amplifier connected to the gate of a PMOS transistor and a feedback resistor placed between the amplifier output and the input node, suitable to perform a linear current to voltage conversion. This solution guarantee a low impedance connection with the DAC aiming to keep the node voltage constant even if the current sourced by the DAC changes (DAC code modification). The DAC and the amplifier circuit are therefore decoupled.

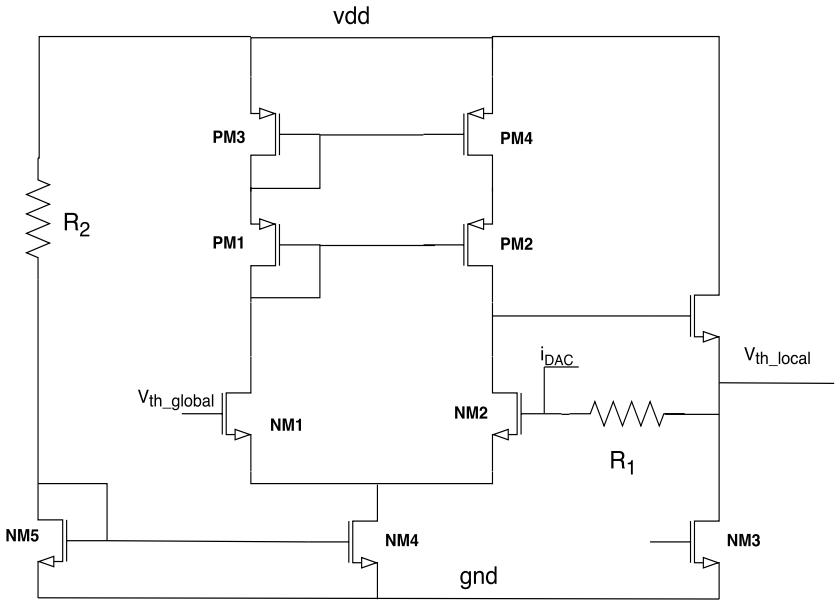

Fig. 3.53 Schematic representation of the trans-impedance amplifier employed for $i_{DAC} \rightarrow V_{th}$ conversion.

Since all the current sourced by the DAC flows through the resistance, the voltage V_{th} changes linearly, as required. In fact, considering a gain A for the differential amplifier, follows that

$$V_O = (V_1 - V_{ref}) \cdot A \tag{3.25}$$

Since all the I_{DAC} current flows through the resistance R:

$$V_1 - V_{th} = R \cdot I_{DAC} \tag{3.26}$$

If the PMOS transistor works in the saturation region, the current I_{SD} is given by:

$$I_{SD} = I_1 + I_{DAC} \approx \frac{1}{2} \cdot \mu_p \cdot C_{ox} \cdot \frac{W}{L} (V_{GS} - V_{th_P})^2 \tag{3.27}$$

And from straight quite basic calculations, assuming that the differential amplifier gain A is high enough to adopt the limit a $\rightarrow \infty$:

$$V_{th} \approx (V_{ref} + R \cdot I_{DAC}) \tag{3.28}$$

Figures 3.54, 3.55 prove the quality of the conversion linearity and Figures 3.56, 3.57 show the DAC voltage steps obtained with the 64 combinations of the 6 bits. It is possible to observe that a $\sim$ 0.5 mV step accuracy has been achieved over a more than 30 mV range.

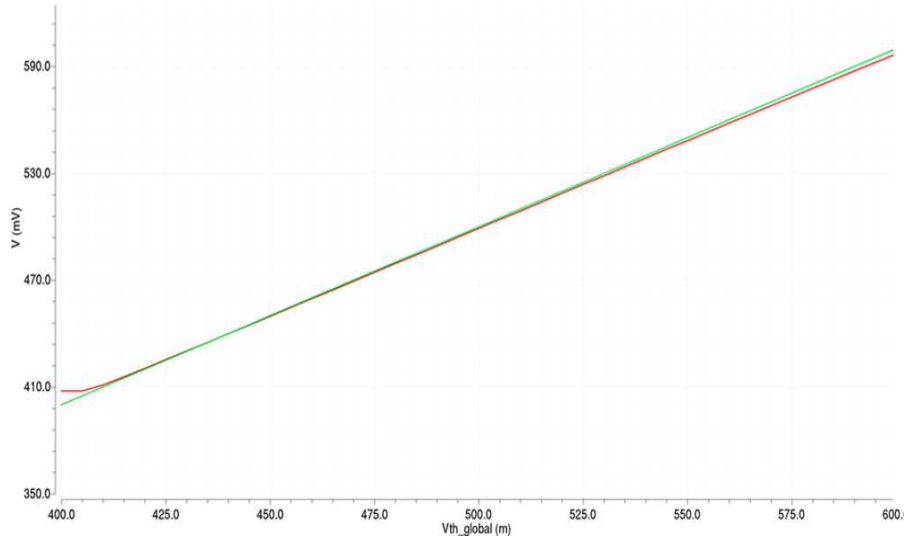

Fig. 3.54 DAC linearity: the conversion linearity is maintained over a$\sim$40 mV range. The plot shows the DAC output voltage in function of the global threshold voltage V_{th_global}. The latter is the top level threshold shared among all the chip channels and is then fine tuned with DACs, at channel level.

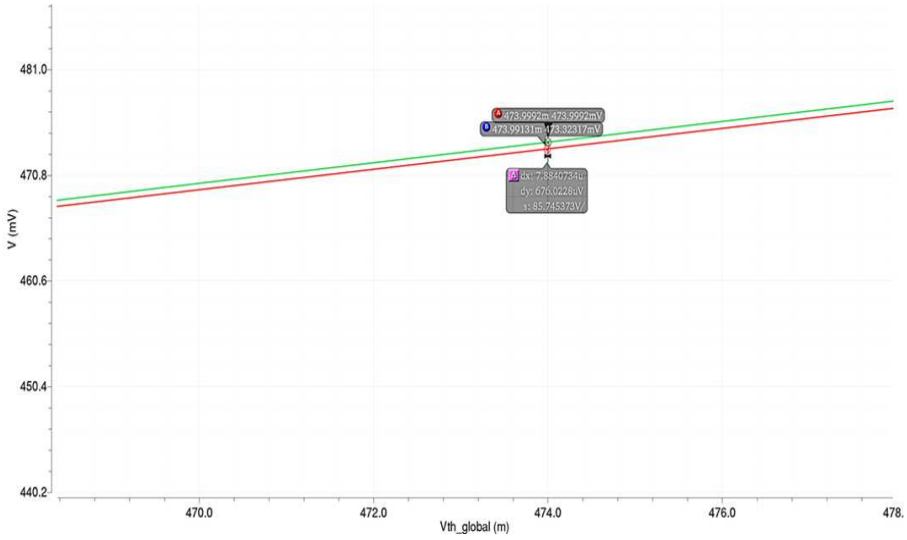

Fig. 3.55 DAC linearity: offset at 475 mV, a typical global threshold voltage.

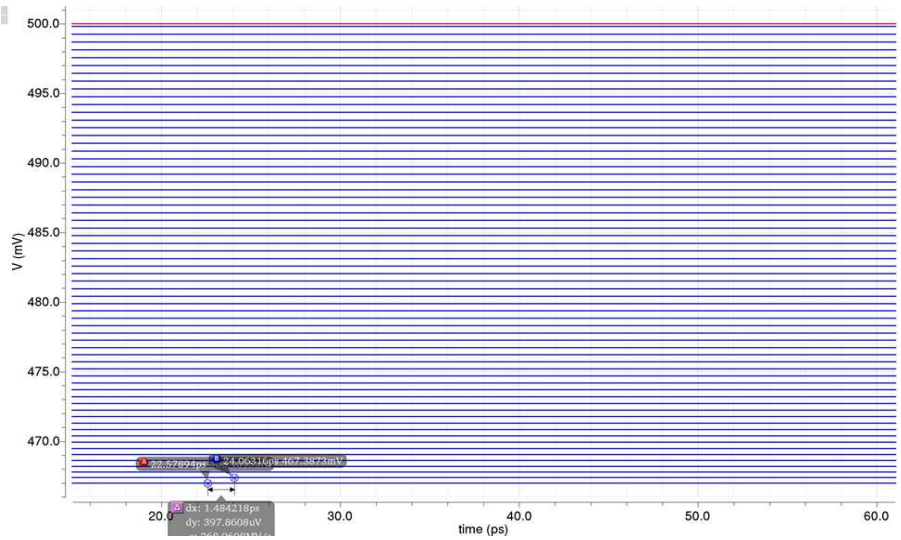

Fig. 3.56 The DAC voltage steps obtained with the 64 combinations of the 6 bits

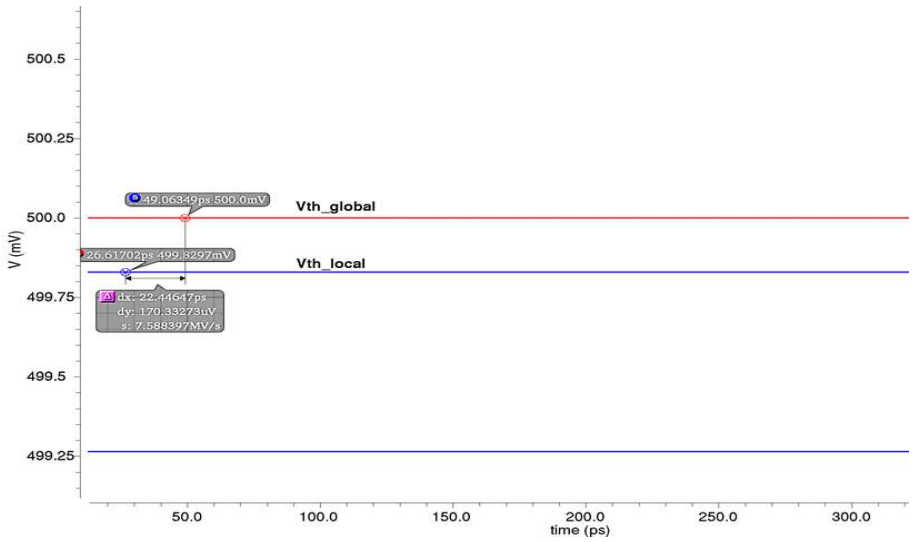

Fig. 3.57 DACsteps3

Thanks to its 6-bits DAC, each ABACUS channel works with his own threshold, with the aim to keep all the 24 channels equalized and ready to switch triggering at the same point, if all the front-ends would be ideally stimulated with an identical input. DAC-based compensation methods require also a local digital memory to store the chosen bit pattern for the converter; this is performed with a shift-register which input and output connections are linked to external pads for configuration via

FPGA. The configuration shift-register has a self-reset that is performed with a R-C low-pass filter at the ASIC power-on.

3.12.7 Pulse generator

ABACUS yearns for being a single ion discriminator which mode of operation is based on a discrimination activated reset of the amplifier feedback capacitance. When the OTA Buffer output voltage crosses the discriminator threshold the discriminator fires and the resulting digitized signal takes two paths: on one way this signal moves to the CML driver while in parallel, the discriminator output drives the CSA feedback reset. The discriminator output suffers on time walk effects and has a duration which depends on the input signal amplitude therefore it is not suitable, to be directly used for the C_f reset, especially in a 10^8 Hz application field. This task is conducted by a block called Pulser (shown in Figure 3.58).

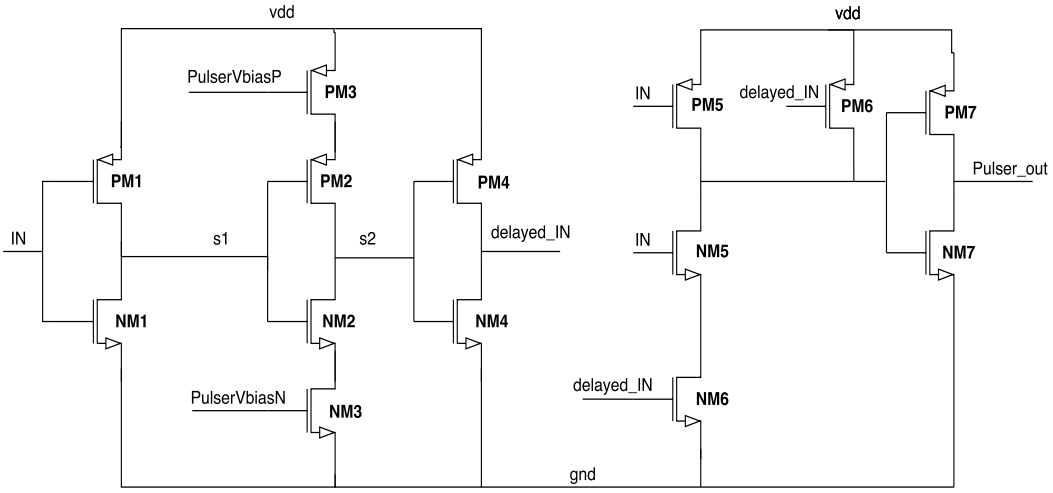

Fig. 3.58 Pulser: transistor level schematic. The output signal results form a bolean AND between the input and a delayed and inverter copy of the input. The two signals overlapping duration can be regulated with the PulserVbiasP and PulserVbiasN voltages.

Basically, referring to Figure 3.58, the Pulser takes the discriminator output that is inverted one first time (net S1), then is inverted again with a starved inverter to introduce a delay (net S2); finally a third inverter digitizes again the signal (delayed_IN). The inverted and delayed signal goes through a boolean AND gate with the original discriminator output (Pulser_out). Figure 3.59 shows the resulting pulse form the two signal overlapping (logic AND) whereas Figure 3.60 reports all

the pulser elaboration steps. It is possible to change the pulse duration regulating the R-C shaping with the two signals PulserVbiasP and PulserVbiasN that control the second inverter delay. Figure 3.61 shows a simulation of the pulse width selection.

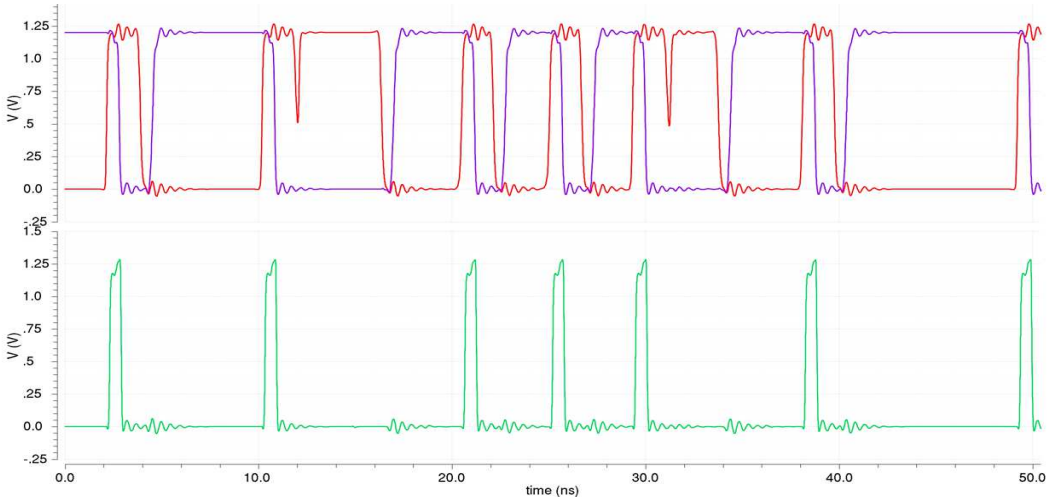

Fig. 3.59 Pulser working principle: the logic AND between the input and a delayed and inverter copy of the input (up) results in the pulse signal (bottom).

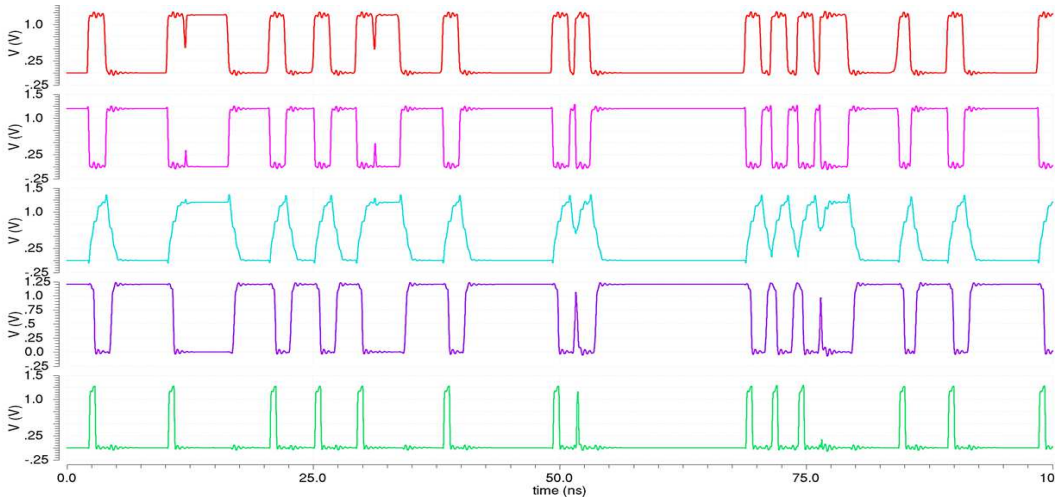

Fig. 3.60 Pulser, signal elaboration. Form the top to the bottom: the discriminator output (IN), the inverted discriminator output (S1), S1 inverted and delayed through a controlled RC shaping (S2), S2 inverted and digitized and in the picture bottom, the resulting pulse.

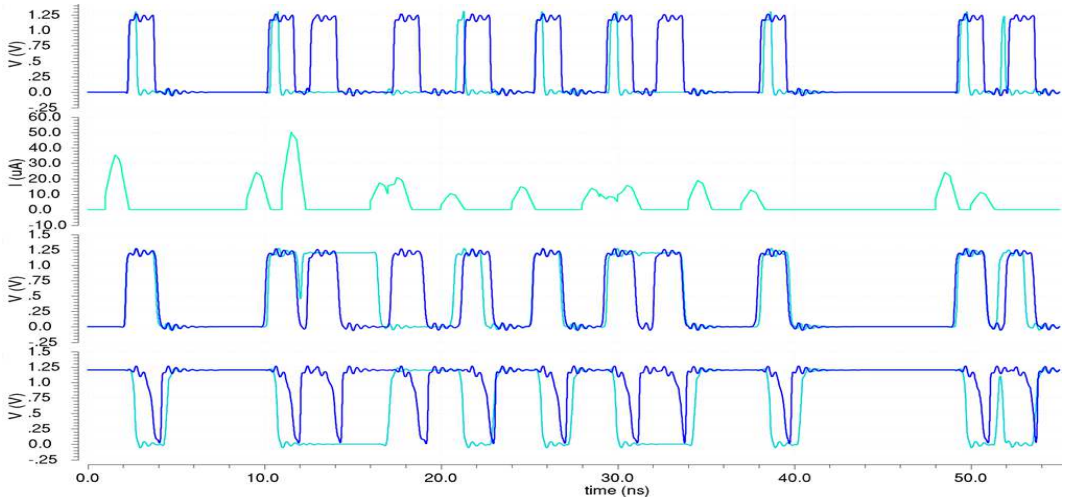

Fig. 3.61 Pulser feature: pulse widht control. Light blue: PulserVbiasP = 700 mV, PulserVbiasN = 500 mV; blue: PulserVbiasP = 785 mV, PulserVbiasN = 700 mV.

3.12.8 Recovery system

The proposed solution has the drawback that whenever the pulse duration is not enough to bring the front-end signal under threshold, the comparator remains fixed in a stack configuration and the system is completely blind. To avoid this effect, a Recovery circuit has been designed and implemented (shown in Figure 3.62.

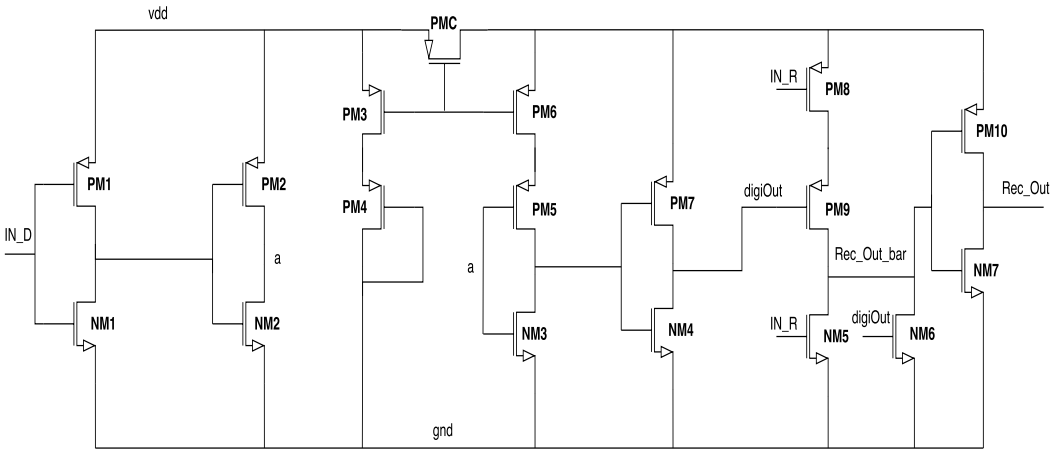

Fig. 3.62 Recovery: transistor level schematic

The working principle of this Recovery block is based on a boolean OR between the Pulser outputs and those discriminator outputs signals after an high-pass filter. The Recovery functional principle is represented in Figure 3.63.

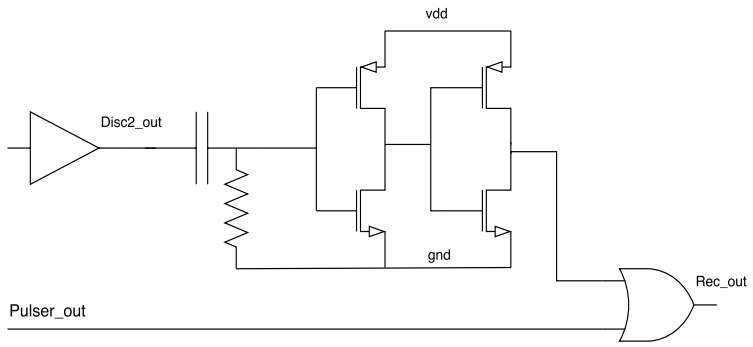

Fig. 3.63 The Recovery block working principle: the Recovery is different from the Pulser output only when the discriminator remains over threshold for more than a R-C fixed at 3.5 ns; whenever this happens, the boolean OR between the delayed discriminator output and the pulser signal, creates an extended reset signal.

In this way, the front-end output is forced down only in the case in that a discriminator signal remains over threshold for a time that can be critical. Studying the discriminator output shape as resulting from different input signal amplitudes in both schematic and layout configurations, it is possible to say that the signal width (i.e. the time duration of the signal) remains in the 1.5 ns - 2.1 ns range. Knowing that, the high-pass filter time constant has been fixed at 3.5 ns: a discriminator output signal that remains over threshold for this time (or more) is considered as a stuck one. After this time the Recovery actively participate to the reset action adding to the Pulser output a delayed discriminator output. In the general and more common situation, the Recovery output is equivalent to the Pulser one, with a unavoidable small delay induced by the CMOS inverter chain. The Recovery block has a differential output since it has to control the NMOS switch-reset transistor and the PMOS adopted for the collection of the charge injection. The PMOS control is operated with a small delay with respect to the NMOS, in order to properly interact after the former action. The Recovery signal elaboration is represented in Figure 3.64.

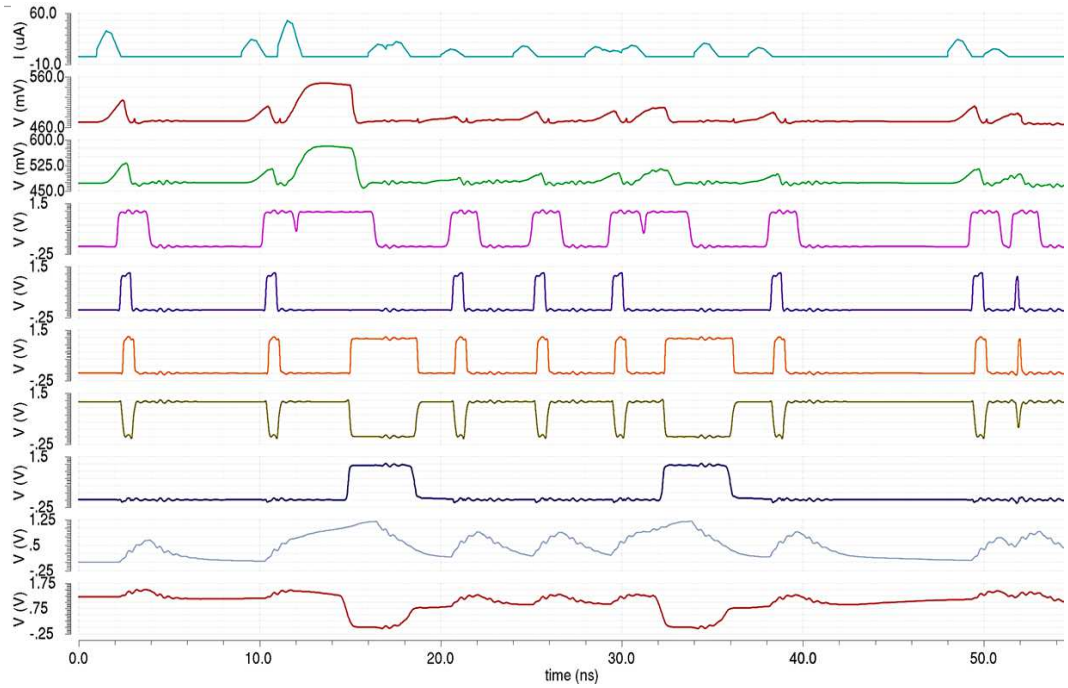

Fig. 3.64 Recovery, signal elaboration. From top to the bottom: the front-end input signal followed by the CSA output and the OTA Buffer output, the digitized discriminator output and the Pulser output. In the picture center, in orange, the Recovery output and its complement, followed by the signal added as active action from the Recovery (blue) and the RC with its digitized and inverted version, in the picture bottom.

3.12.9 The CML driver

The ABACUS ASIC is a prototype device which output is a differential signal that is going to be readout by an FPGA. The choice of the FPGA-type is related to the number of I/Os with the capability to perform a synchronous sampling with a frequency of at least 1 GHz. Concerning the chip embedded logic, the ABACUS discriminator output passes through a block that performs the single-ended to differential conversion (Figure 3.65) which in turn drives a Current Mode Logic driver which schematic is shown in Figure 3.66 (the CML bias OTA schematic is reported in Figure 3.67).

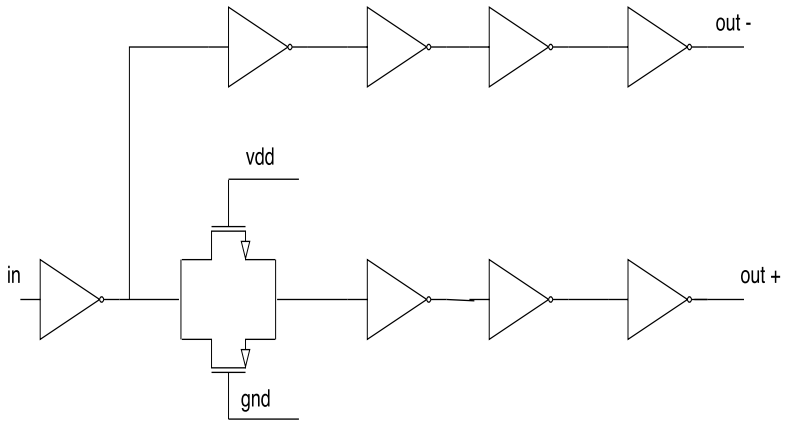

Fig. 3.65 Single-ended to differential converter: transistor level schematic.

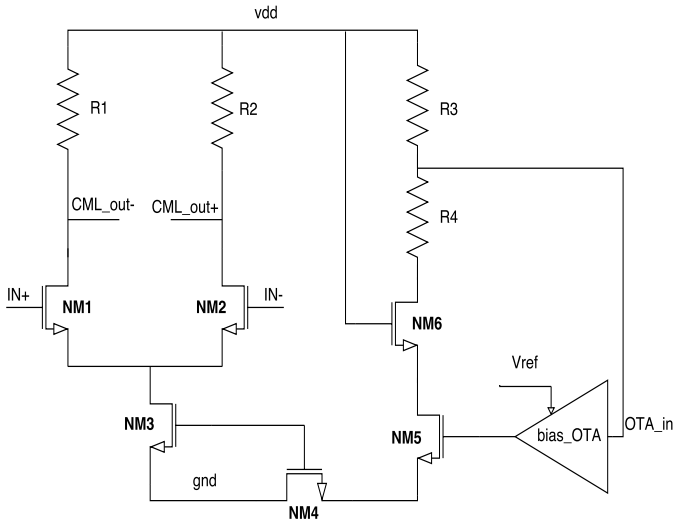

Fig. 3.66 CML driver: transistor level schematic

This logic standard has been chosen after a comparison with LVDS (Low Voltage Differential Signaling). The fact that CML requires a 1.2 V power supply domain (CMOS compatible) instead of the 2.5 V for the LVDS, added to the fact that the CML voltage swing can be controlled by a current source (thus the power too) with a simpler output stage, favors this driver type. Another option would be the scalable Low Voltage Signaling format which turned out to be not compatible compatible with the FPGAs already available. A comparison among the three data transmission formats is reported in Appendix1.

A CML driver operates in the MOS active region to maximize the speed (i.e. triode region avoided). The reduced voltage swing (compared to CMOS static circuits), the

fact that the current switching takes place at the input differential pair transistors and considering also that a CML driver is terminated at its input, make this device more suitable for high speed data transmission.

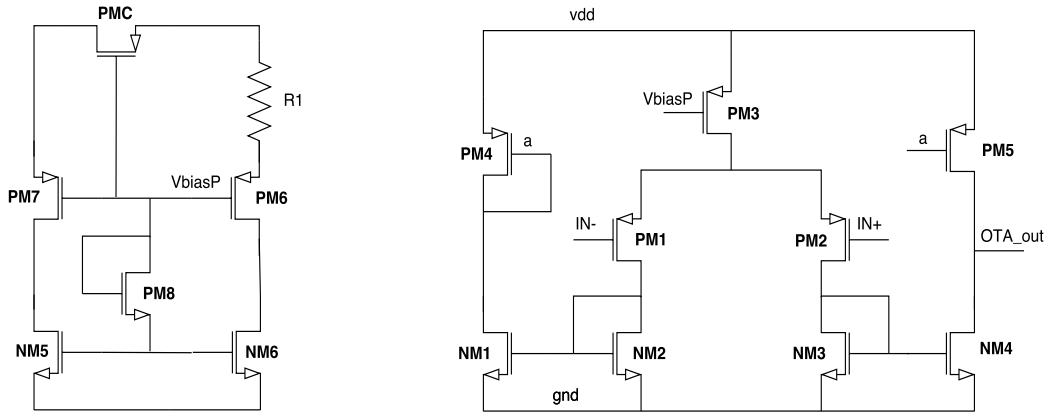

Fig. 3.67 CML embedded buffer: transistor level schematic

3.13 ABACUS layout

3.13.1 Introduction and general aspects

The level of complexity typical of a VLSI device, is such that the approach on microelectronics design is distinctly split in schematic and layout phases. The layout design phase is based on the conversion of a symbols system into a geometric shapes one, introducing metals, silicon wells, polysilicon contacts etc.. During the layout the designer deals with physical layers placing objects in two-dimensional surfaces that are interconnected through VIA, considering the coupling capacitance of overlapped layers as well as the R-Cs that each line carries in the layout view.
The layout standard flow requires a series of verification steps like the Design Rule Check (DRC). The DRC mainly controls the minimum width of metal lines with a given current, the minimum distance between two different layers, the minimum number of vias required for connection, the maximum distance between the channel and its closest bulk connection. To ensure the correct function of the fabricated chip and enhance the yield, some margin on those limits are usually requested.
Consider now the parasitic effects. For metal transmission lines, the smaller width they have, the larger parasitic resistance they will introduce. Considering the current

density limitation for a specific metal layer, large biasing currents would require wide metal paths however, this would introduce strong capacitive coupling between overlapped layers. At the same time, long metal connection on MOSFET gates violate antenna rules. As previously mentioned different layers are interconnected through VIAs in large number, avoiding highly resistive paths. Concerning the transistors an important effort has been taken during the ABACUS design in protecting the substrates from pick-up noise generated by the digital signals fast clipping (a dedicated subsection focuses on this aspect).
Once all the above implicit and the explicit rules are satisfied a further control is provided by the Layout Versus Schematic (LVS) check: the software verifies the correspondence between the schematic and the layout in terms of components, nets and interconnection equivalence. Only after the design is LVS clean, the flows allows to proceed with the extraction of the parasitics (PEX).

3.13.2 Layout view

The ABACUS ASIC has been designed with a 8-metals technology and a minimum MOSFET gate length of 110 nm (it is a 130 nm shrink). The design area is $2 \cdot 5\ mm^2$ and the number of pads is 140. The number of blocks that a channel can integrate can be limited by both area and pad number, since proper bias and control signals are usually required.
During the layout design, particular attention has been paid on the block interconnections in order to reduce RC delays. The electro migration rules have been checked with particular care on the high current nodes. The channel level topology takes into account first of all the signal propagation by placing the front-end as close as possible to the input pad and the driver close to the output pads; the comparator logic is close to the front-end to optimize the signal integrity but it is biased with a different power domain to protect the analog side from the digital switching. The pulser and recovery block are still quite close to the amplifier due to the feedback reset connected to the integrating capacitance: long paths would pick-up noise and provide delayed reset which fast reaction is essential with high signal repetition rates. The central part of the channel hosts the blocks fed by signals that are already digitized and/or that do not interact directly with the front-end (DAC and its control logic, the single-ended to differential converter, the shift register for DAC programming).

Figure 3.68 shows the ABACUS channel layout view and a front-end zoom-in is depicted in 3.69.

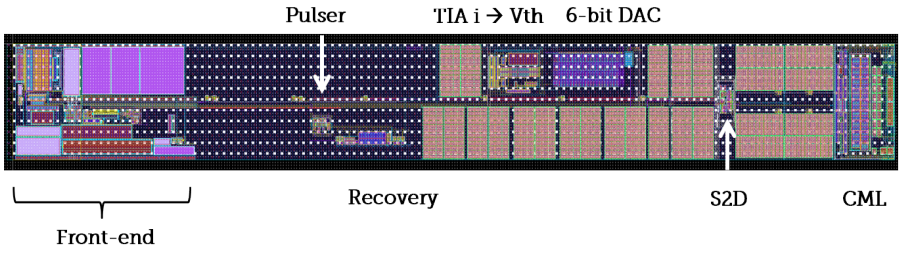

Fig. 3.68 ABACUS channel in its layout view.

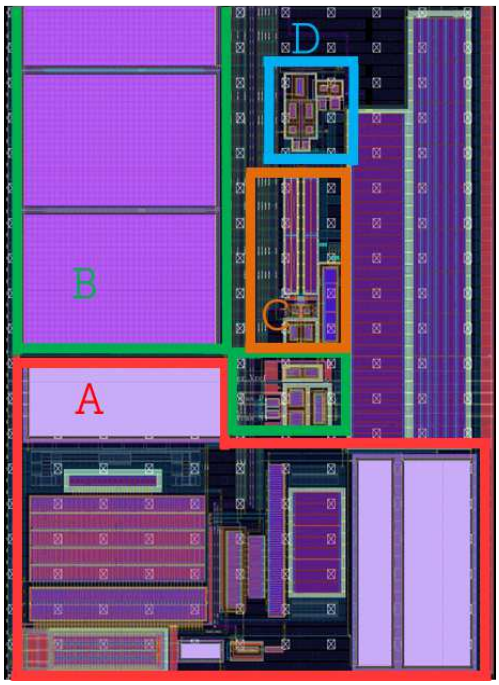

Fig. 3.69 Front-end layout with block identification.

The ASIC top level is arranged in four sectors by six channels each (Figure 3.70). This choice descends from various reasons, starting from the fact that this chip is a prototype and and thus it must be suitable for detailed testing; therefore, the possibility to tune as much signal as possible (e.g. the CSA and discriminator cascode bias currents and voltages) has to be used. Due to the limited number of pads it is not possible to route out many signals at channel level but it is at least possible to group channels to make intermediate signal externally available. Another reason

supporting the sectors division is the implementation of sector level bias cells instead of biasing all the channels with pads mirrored on the chip edges: this choice has been preferred in mitigating the bias gradients responsible for channel-by-channel behavior inhomogeneity.

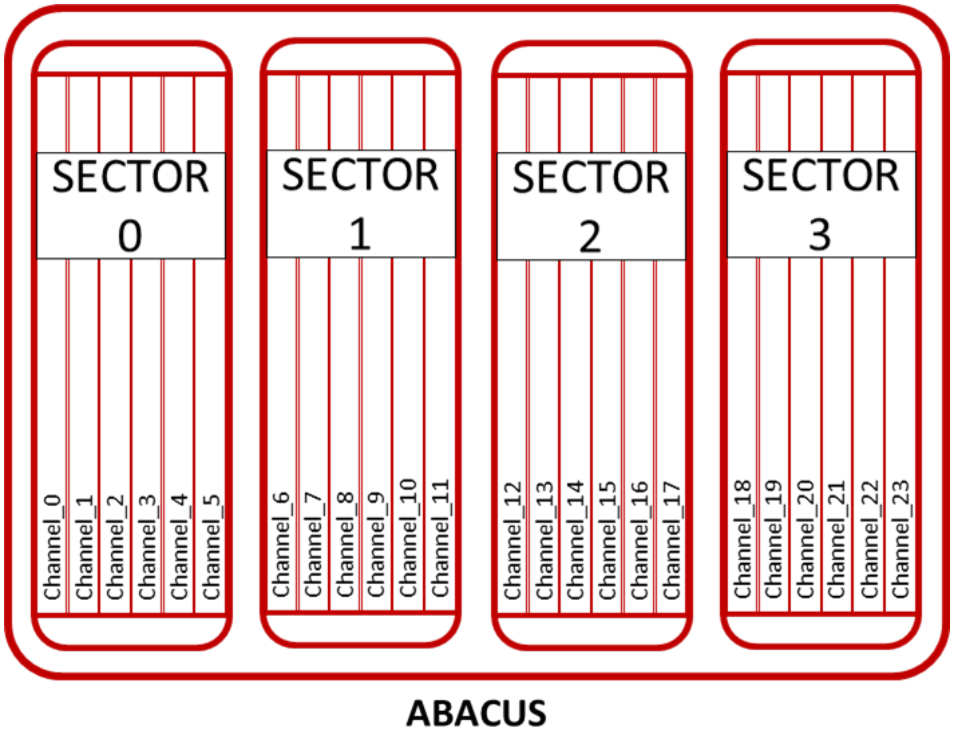

Fig. 3.70 ABACUS: sector division.

Having a sector modularity has been a great advantage in the heavy simulation phase that flanked the designed till the end. It has been possible to simulate increasing the level of complexity, from the single channel, moving to the six channel packed in a sector, a sector with only one channel (mini-sector) and the chip made by mini-sectors. Only at the end it has been possible to simulate the entire chip moreover, this is even more complex and slow for the layout view simulation with the parasitics. The channel to channel interconnections as well the sector to sector ones have been simulated only with an higher level block that routes those lines. Figures 3.71, 3.72, reports a top level view of the ABACUS layout. In the bottom of the pictures it is possible to see the inputs packaged between grounds, to shield and better isolate the channel signal. In the upper part there are the two polarity driver outputs. On the shorter edges the bias currents, voltages and the set and control signal are placed, as

well as the contacts for the power domains. The available area and the number of the channels are responsible of the right and left separation with two sectors each side. The ASIC center is filled with filtering capacitors and the higher metals are here used for the power distribution net.

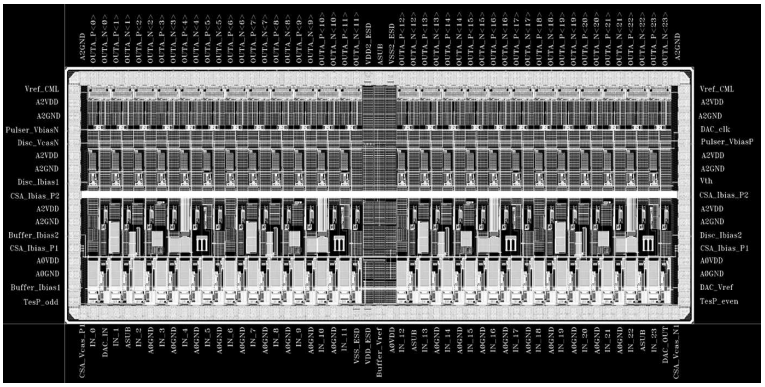

Fig. 3.71 Black and white layout view of the ABACUS top.

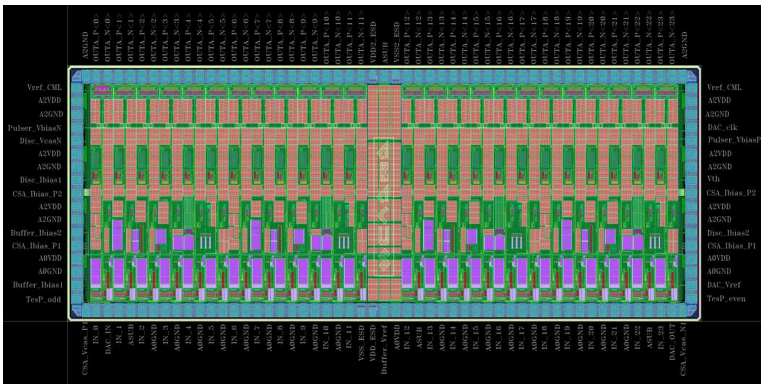

Fig. 3.72 Colored layout view of the ABACUS top.

Parasitic induced delays

The simulation of the net-list back-annotated with parasitic components extracted by the PEX tool reveals variations in the circuit characteristics like the bandwidth, signal amplitude and noise performance (if large resistors appear in the signal path). The component mismatch issue and the fact that small transistors suffer more for flicker noise would push to use bigger devices but the area constraint is against this. Moreover, big components complicates the routing and spread blocks require long path connections that degrade the signal. Figure 3.73 is useful to appreciate the

propagation delay for a digital signal, into the ABACUS channel; Figure 3.74 is a detail of the previous one.

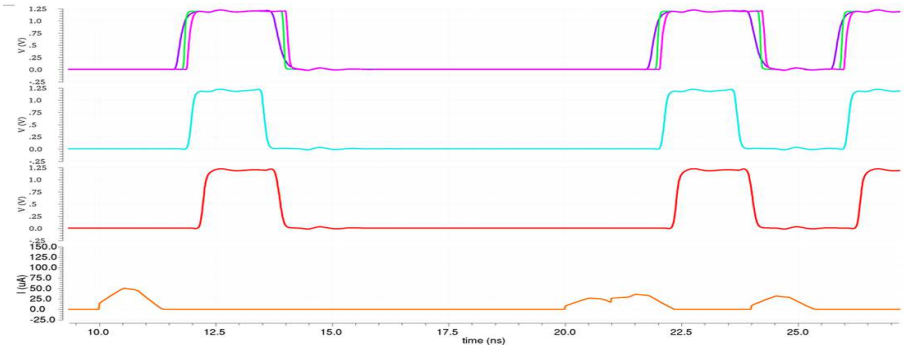

Fig. 3.73 Signal propagation delay among the discriminator (top), pulser and recovery output signals. The input signal is the bottom one. In the upper part, there is the superposition of three discriminator signals as the channel experiences having two CMOS inverter for signal reshaping. The waveforms are taken from a post-layout simulation.

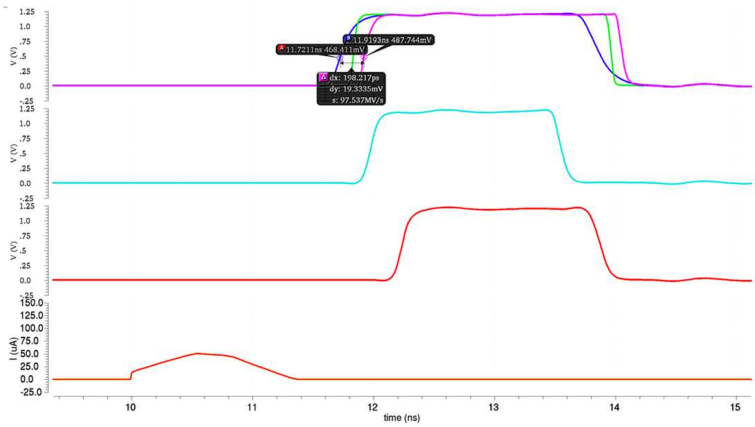

Fig. 3.74 Propagation delay among the discriminator, pulser and recovery output signals (post layout simulation): focus on the 2 inverter chain for the discriminator output signal reshaping.

3.13.3 Power domain splitting

The ABACUS ASIC has separated power domains for the analog and the digital sectors, as showed in Figures 3.75 3.76 3.77 3.78. This choice has been adopted to separate as much as possible the front-end (intended as the preamplifier plus OTA buffer), which is the most sensible part, form the rest of the chain especially from the discriminator and the driver that, during switching, can inject large spikes into

the power lines (up to $\sim$ 40mV peak-to-peak). Unfortunately this technique was not sufficient to avoid instability, emphasized in a multiple channel firing situation. Details related to this issue are reported in the following subsection.
Going back to the first aspect, the front-end-dedicated power domain is also suitable for keeping the current signal in a loop that starts from the detector and possibly finishes as close as possible, in order to minimize the noise pickup.

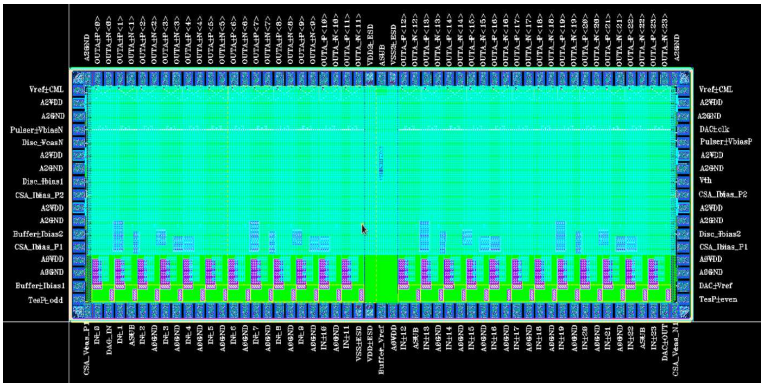

Fig. 3.75 ABACUS layout tip view: digital vdd highlighted.

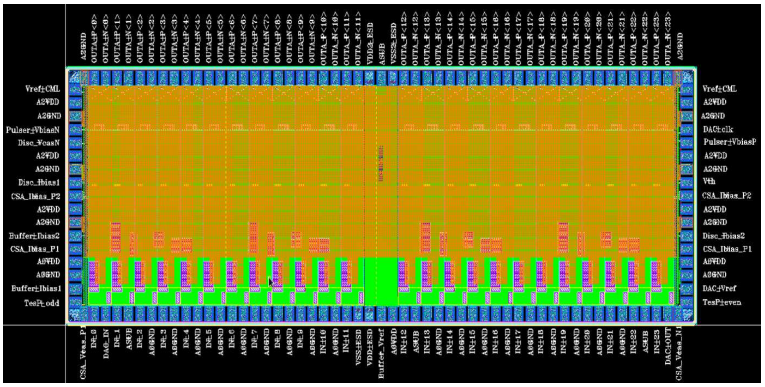

Fig. 3.76 ABACUS layout tip view: digital gnd highlighted.

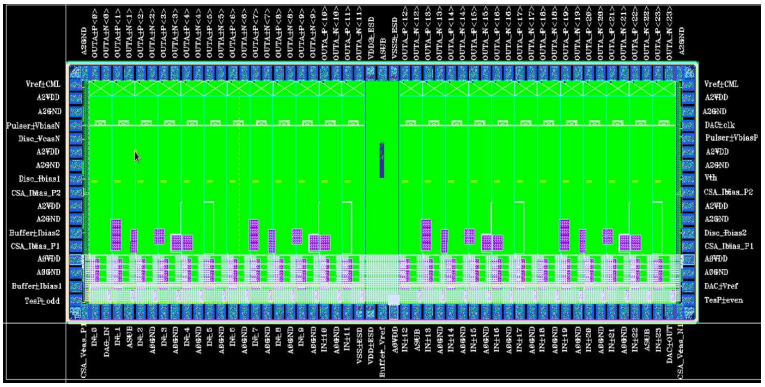

Fig. 3.77 ABACUS layout tip view: analog vdd highlighted.

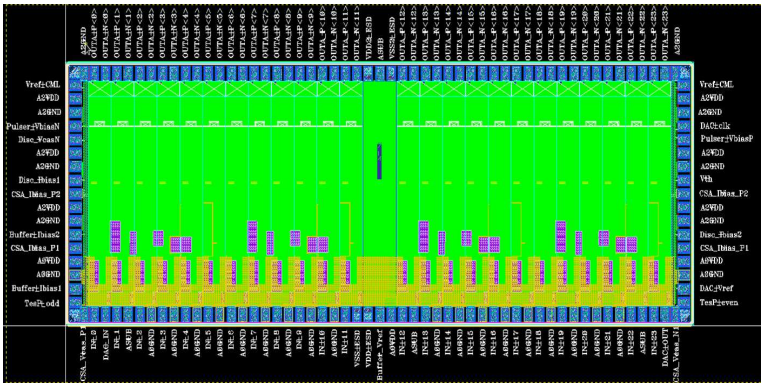

Fig. 3.78 ABACUS layout tip view: analog gnd highlighted.

3.13.4 Substrate noise coupling mitigation

In the field of integrated circuit design, high speed and high power digital blocks feature current and voltage glitches in their transients, that can arise as parasitic currents through the silicon substrate. The presence of a common substrate causes an electrical coupling of devices in various parts of the circuit meaning degradation in performance and reduction of the system performances, especially for analog circuits. More challenging fields in this sense are those integrating mixed signal architectures and/or radio-frequency (RF) components, commonly added since the continuous shrinking of the technology size pushed for higher level of integration.
A crucial point in the ABACUS layout design has been the stability control of the CSA baseline, whenever the CML driver switches. In addition to have separate power and ground two more techniques allowed to overcome the fact that the

CMOS substrate is eventually shared among the whole ASIC. First of all, the CML high-dynamic transistors (e.g. the input ones) have been designed as triple-well (Figure 3.79) thus to disentangle the typical inner ground-substrate or vdd-substrate connection, respectively for NMOS and PMOS. In this way it is possible to change the substrate connection of certain CML transistors forcing them to follow a longer path to reach the ground. It is important to remark that a direct connection substrate-ground-source for a NMOS has a different impact than a substrate-ground connection at symbol level, that means for instance a longer path for the signals.

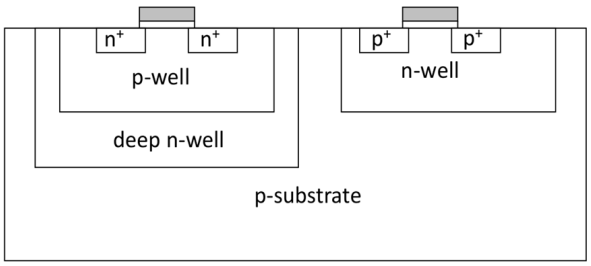

Fig. 3.79 Drawing of a CMOS triple well solution for transistor shielding.

A qualitative but highly indicative evidence of the mitigation results can be understood looking at Figures 3.80 3.81. These two pictures report one over the two polarity for the CML signal (which is differential), for ten channels that are stimulated with different kind of signal inputs (different amplitudes, rates and starting time). The idea is to highlight the possible crosstalk and interference effects as result of high frequency parasitic signal, coupled from the drivers to the CSA. Figure 3.80 shows a situation in witch the CML has ground and bulk shorted at transistor level whereas in Figure 3.81, the CML RF transistors have been modified as triple well and the substrate is connected to the channel one, the ground is connected to the channel digital ground.

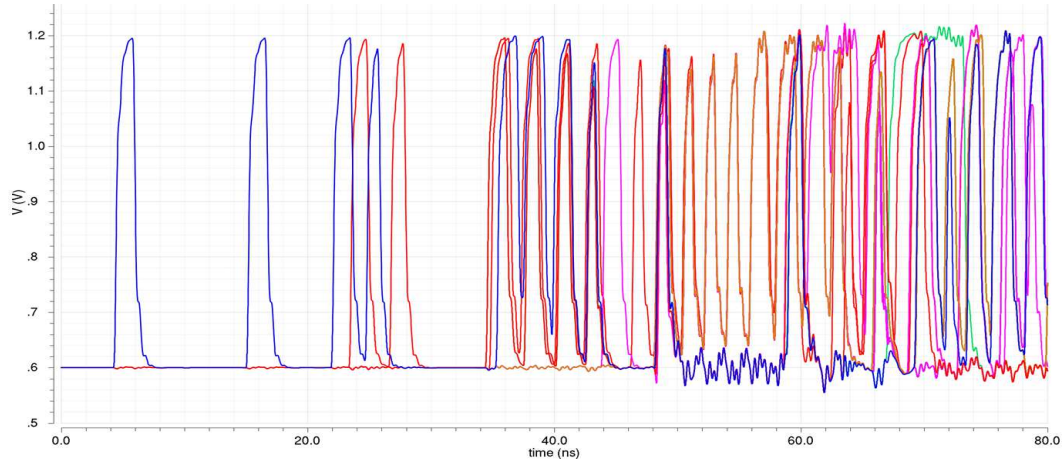

Fig. 3.80 CML positive output signals. The substrate connection propagates spikes on the power lines.

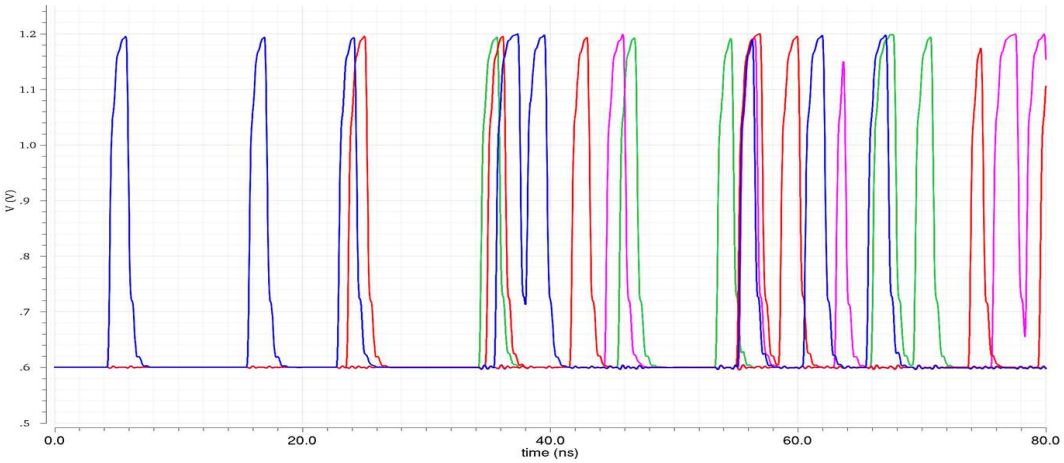

Fig. 3.81 CML positive output signals. Result for the substrate noise coupling mitigation technique: the CML output is connected to the digital domain gnd, thanks to triple-well transistors.

Another technique that has been adopted to mitigate the interference noise propagation through the substrate is based on a design layer named P-Well Blockage (PWBL). Using this approach it is possible to keep the substrate continuity in a critical region enclosing the critical blocks with a trench that excludes p++ doping elements in the substrate thus to increase its local resistivity. This (PWBL) structure is placed into the layout by the designer and physically realized before the transistor, during the semiconductor device fabrication process. Unfortunately the effect of this technique could not be simulated with the CAD, since the designers did not

have the proper tools for doing it (neither the substrate models nor the proper simulator licenses). Nevertheless this approach is rather common and has been already successfully adopted by the group in other projects.
Figures 3.82 3.83 are snapshot captured form the ABACUS channel layout and show respectively the PWBL for the S2D block and the latter shows the PWBL for the CML driver.

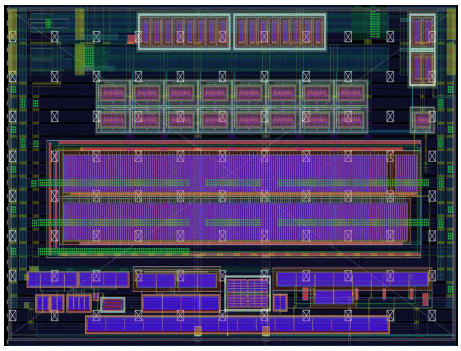

Fig. 3.82 P-Well Blockage adopted to mitigate the substrate noise coupling of the S2D. The PWBL is the semi-transparent trench boundary enclosing the block in the centre.

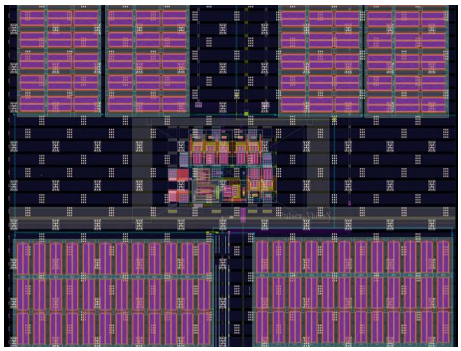

Fig. 3.83 P-Well Blockage adopted to mitigate the substrate noise coupling of the CML driver. The PWBL is visible at the very picture boundary.

3.13.5 Discrimination capability

Among the channel single functional block design and characterization, the figure of merit for ABACUS remains the ability to discriminate signals resulting from 60 MeV - 250 MeV protons hitting a UFSD strip. This subsection collects a series of CAD simulation results for both schematic and layout test benches, where GEANT4 time distributed and Weightifeild2 shaped signals are set as ABACUS channel inputs.

During the efficiency tests a standard digital counter in Verilog has been used to count the 0→1 transition number for the discriminator output. The number of the input pulses is known from the file that has been artificially created, therefore an efficiency estimation is obtained dividing the digital discriminator pulses by the analog input pulses coming from the sensor (as explained in previous sections, the sensor is modeled with a capacitor connected in parallel with a current generator linked to a text file with the simulated signals).
Firstly schematic simulation results for MIPs at the limits of the considered frequency range are reported: MIP, 50 MHz in Figure 3.84 and MIP 250 MHz in Figure 3.85.

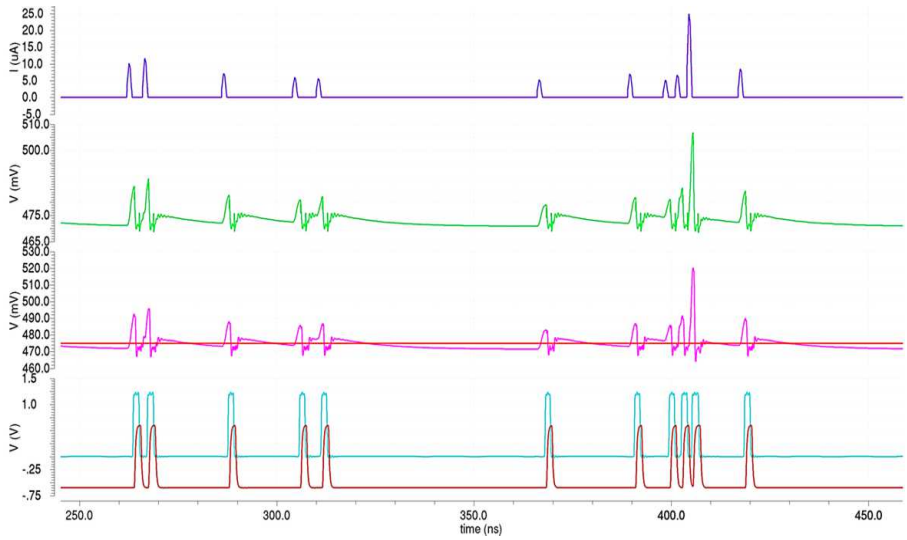

Fig. 3.84 Schematic simulation waveforms for MIPs at 50 MHz. From the top: the input signal, the CSA output voltage, the OTA buffer output voltage with the discriminator threshold superimposed and the CML+ and CML- on the bottom.

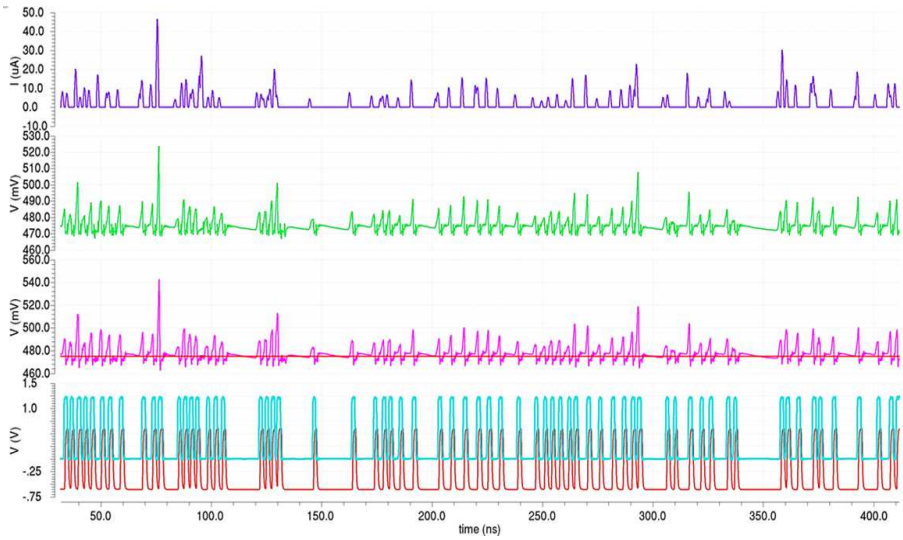

Fig. 3.85 Schematic simulation waveforms for MIPs at 250 MHz. From the top: the input signal, the CSA output voltage, the OTA buffer output voltage with the discriminator threshold superimposed and the CML+ and CML- on the bottom.

Fixing the rate at the higher value, the effect of the two energy boundaries are reported in Figure 3.86, the 60 MeV and Figure 3.87 the 250 MeV and 250 MHz, keeping in mind that lower energy ions have an higher stopping power and are thus responsible for larger amount of energy released into the crossed medium.

The time of arrivals of the events on a radiation detector usually follows a Poisson distribution, described by the following equation:

$$P(n) = \frac{\mu^n \cdot e^{-\mu}}{n!} \tag{3.29}$$

where P(n) is the probability to observe n events in a process having μ as mean value of events.
Marked in Figure 3.86 with time difference indicators, the fact that the statistics regulating the input pulses temporal distribution, is responsible for overlapped signals or signals that are 1-2 ns close already as sourced from the detector. These cases are going to result in unavoidable pile-up events.

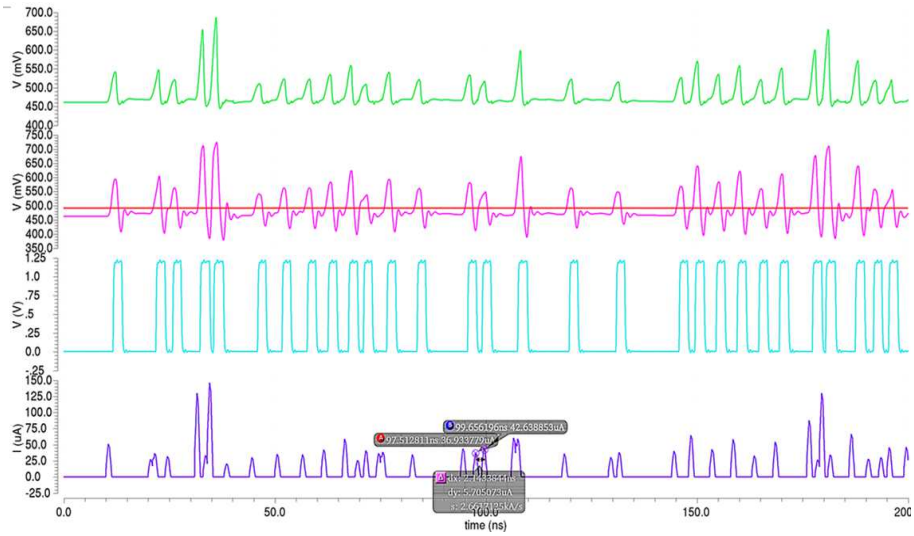

Fig. 3.86 Schematic simulation waveforms for 60 MeV protons at 250 MHz. From the top: the CSA output voltage, the OTA buffer output voltage with the discriminator threshold superimposed, the discriminator output the input signal.

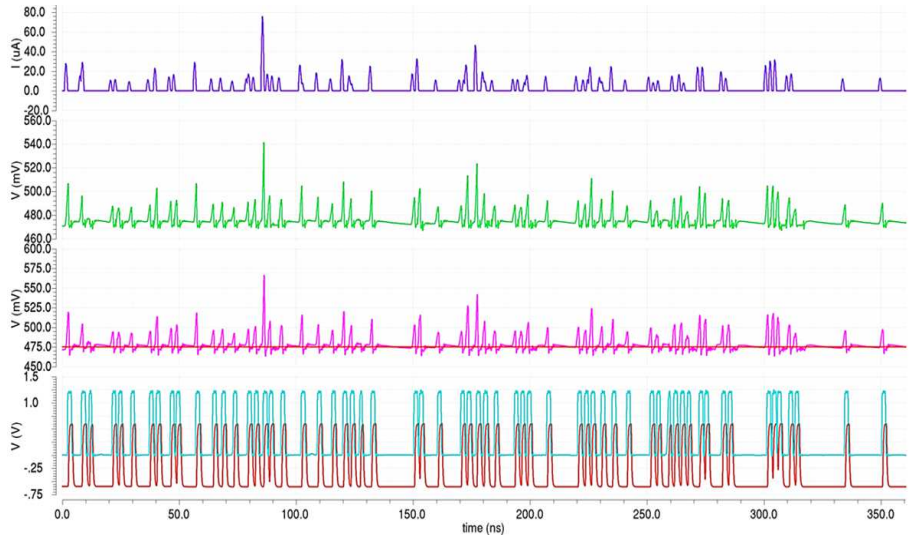

Fig. 3.87 Schematic simulation waveforms for 250 MeV protons at 250 MHz. From the top: the CSA output voltage, the OTA buffer output voltage with the discriminator threshold superimposed, the discriminator output the input signal.

Looking at Figure 3.88, it is possible to understand a certain amount of effects that come out or are just emphasized with the layout version of the channel. It is possible to observe walk effects for different amplitude pulses and the delay between the time when the OTA buffer signal crosses the threshold and the discriminator starts

to rise; an increasing delay propagates until the recovery signal is finally used to reset the feedback capacitance. Another interesting situation is highlighted by a red circle in the right lower corner that indicates the action of the recovery circuit, as a logic OR between the pulser signal and a delayed discriminator output but only if the latter remains up at least for 5 ns (otherwise the recovery is equivalent to the pulser output as explained in the dedicated subsection). Since it is a feedback command that brings down the discriminator output at the low logic level, it is possible to see only the shorten discriminator output signal, as modified by this action.

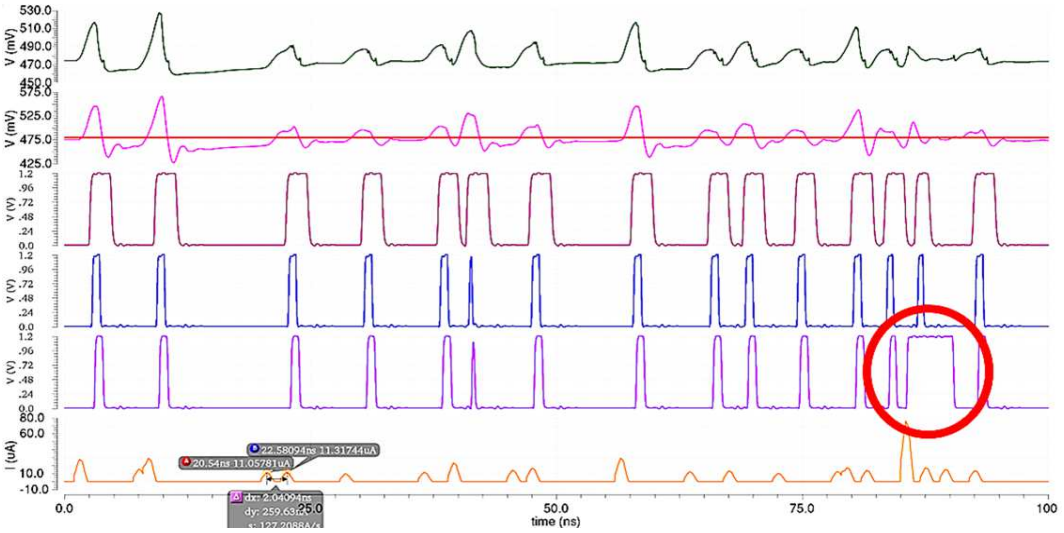

Fig. 3.88 Post-layout simulation waveforms for 60 MeV protons at 250 MHz. From the top: the CSA output voltage, the OTA buffer output voltage with the discriminator threshold superimposed, the discriminator output, the pulser output, the recovery block output and the input signal on the bottom.

In the following couple of pictures (3.89 3.90) a more complex test bench has been set to maximize the simulation run contents: an ABACUS sector (six channels) as been arranged with 4 different input sources and two channels unconnected. Figure 3.89 reports a 100 ns transient simulation with the waveforms of the analog input signals and the digitized signal as discriminator outputs, with four different configurations; from top to bottom: 250 MeV_200 MHz, 60 MeV_200 MHz, 60 MeV_100 MHz, 250 MeV_100 MHz. The same configuration setting but for the layout channel is reported in Figure 3.90.

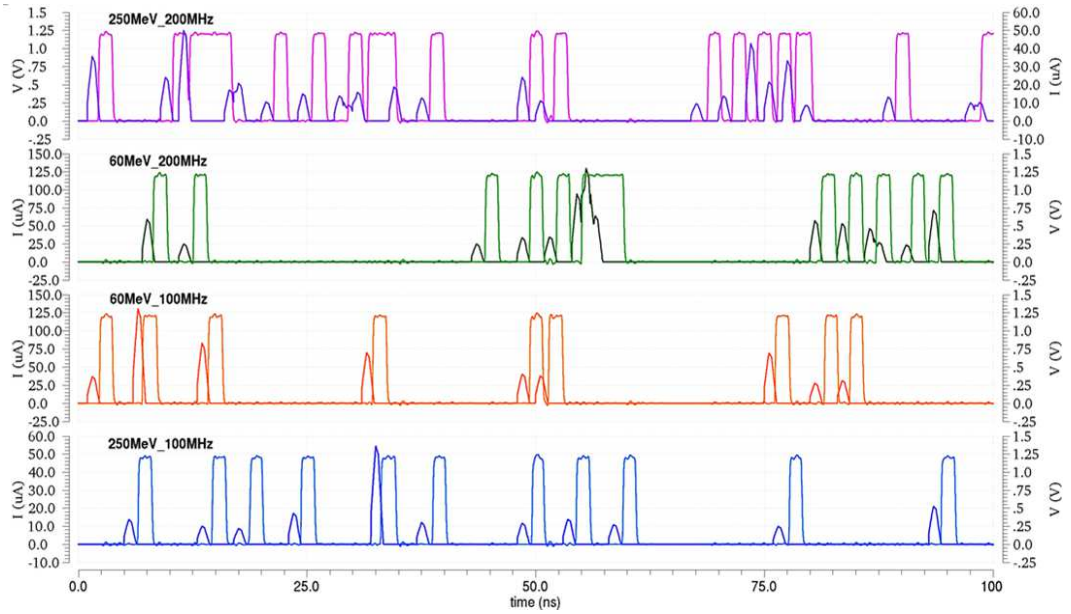

Fig. 3.89 Counting efficiency, schematic simulation. The waveforms represent the front-end input signal and the discriminated one, for four different configuration. From top to bottom: 250 MeV_200 MHz, 60 MeV_200 MHz, 60 MeV_100 MHz, 250 MeV_100 MHz.

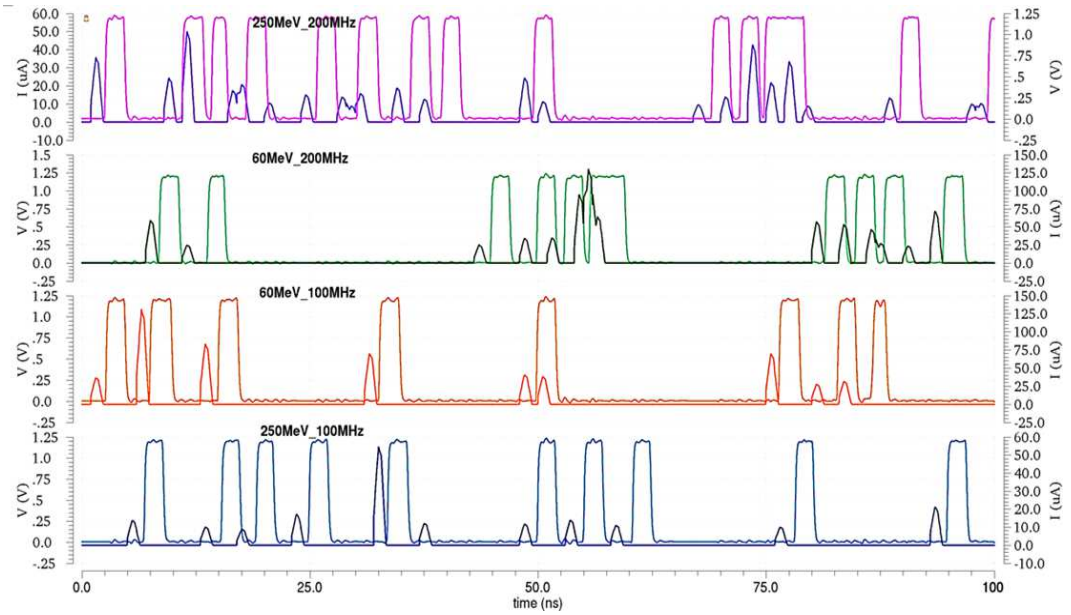

Fig. 3.90 Counting efficiency, layout simulation. The waveforms represent the front-end input signal and the discriminated one, for four different configuration. From top to bottom: 250 MeV_200 MHz, 60 MeV_200 MHz, 60 MeV_100 MHz, 250 MeV_100 MHz.

Table 3.4 Summary of the channel level power consumption, separating the analog and the digital domain

Domain	**mW**	**blocks**
AVDD, AGND	12.5	CSA, OTA buffer, DAC logic
DVDD, DGND	32	all the rest
Total	44.5	channel

Table 3.5 Summary of the ASIC top level power consumption, separating the analog and the digital domain

Domain	**mW**	**blocks**
AVDD, AGND	304	CSA, OTA buffer, DAC logic
DVDD, DGND	764	all the rest
Total	1070	channel

Power consumption and temperature sweep analysis In all the simulation results presented in the previous sections the temperature parameter was set at 55 Celsius degrees. This is a typical value expected for a custom chip of this dimension. Even if the power consumption was not a MoVeIT requirement, it influences the ASIC temperature. Considering the CSA nanosecond rising time and the high frequency input signals, it is easy to say that the circuit would be a considerable sink. Details concerning the power consumption are reported in tables 3.4, 3.5.

It is well known that the expected $\sim$1 W chip power consumption and without a proper mechanical heat sink or an active cooling system, the ASIC will most probably overheat up to 80-90 Celsius degrees. By studying the temperature influence in the ABACUS channel working condition it is possible to resume claiming that the most important thing that influence a correct and efficient behavior is the preamplifier (and thus the OTA buffer) baseline value. Specifically, the OTA buffer baseline should remain in the 465-480 mV range, with an optimum in the intermediate value of 474 mV. Typically happens that the baseline moves inversely with the temperature.

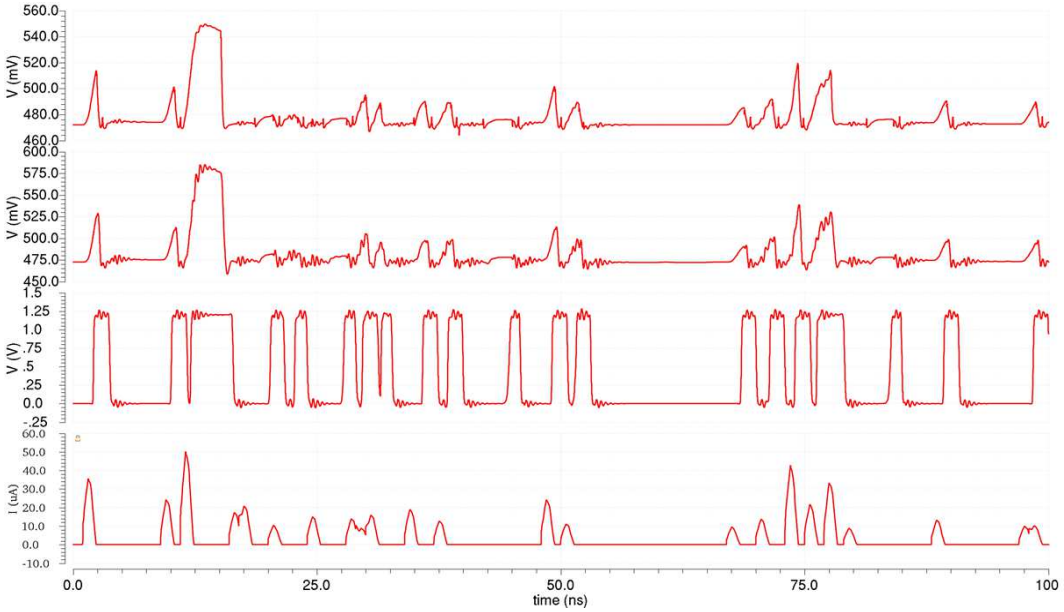

Fig. 3.91 Schematic simulation waveforms for 250 MeV protons at 200 MHz at 27 degC and I_{bias} = 7 mA. From the top: the CSA output voltage, the OTA buffer output voltage, the discriminator output and the input signal.

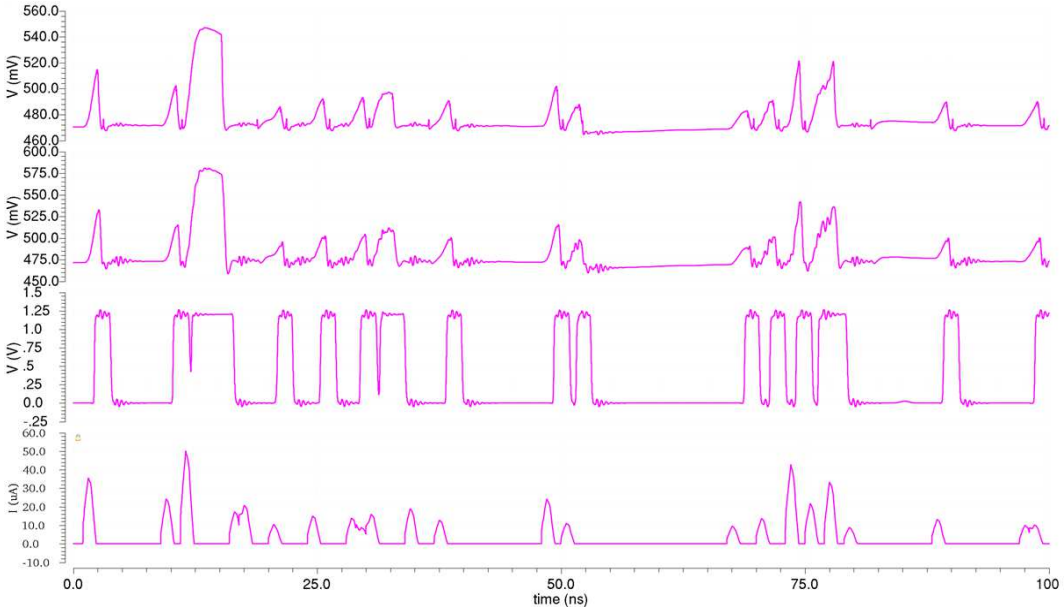

Fig. 3.92 Schematic simulation waveforms for 250 MeV protons at 200 MHz at 55 degC and I_{bias} = 9 mA. From the top: the CSA output voltage, the OTA buffer output voltage, the discriminator output and the input signal.

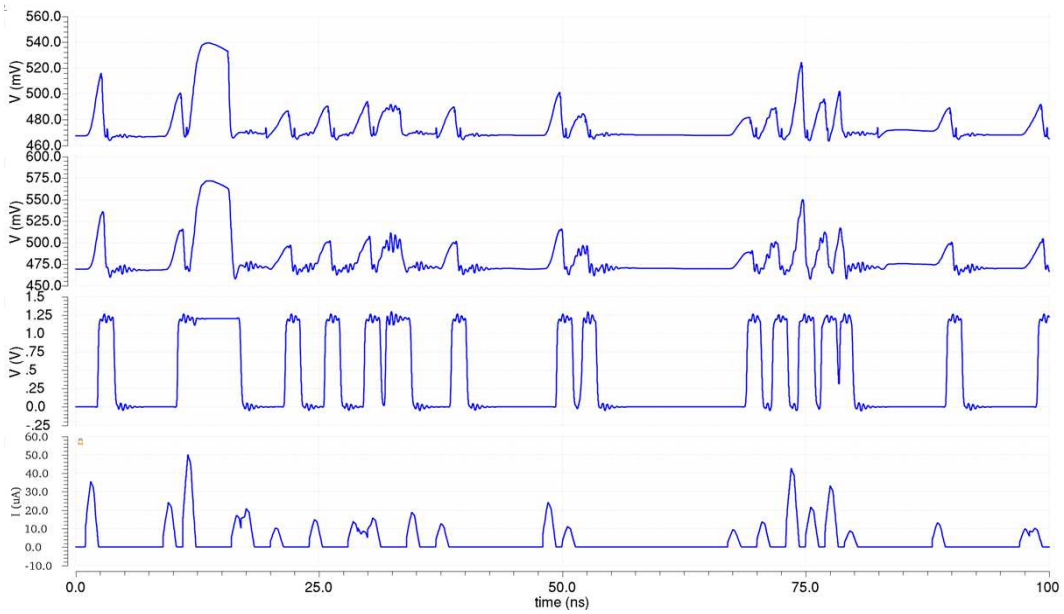

Fig. 3.93 Schematic simulation waveforms for 250 MeV protons at 200 MHz at 90 degC and I_{bias} = 10 mA. From the top: the CSA output voltage, the OTA buffer output voltage, the discriminator output and the input signal.

To control this phenomena it is possible to tune the bias current of the right branch of the telescopic cascode amplifier (in CSA configuration). The nominal value for this current is 9mA at 55deg C and it has to be reduced to 7mA at 27deg C or increased to 10 mA at 90deg C, as shown in Figures 3.91, 3.91, 3.91.

3.14 Summary

The goal of this project is to realize a single ion discriminator for particle therapy applications. Promising experimental results on UFSD sensors show an excellent time resolution of 35 ps and signal duration at nanosecond level with a high S/N value. Nevertheless critical aspects have to be managed from now on in order to minimize effects like pile-up and radiation damage. Sensors have then to be coupled with custom front-end electronics, called ABACUS, able to discriminate nanosecond wide pulses at 10^7 MHz rate, in a pulse charge range extending from 3fC up to 150 fC.
In Turin, the microelectronics team collaborating with MoVeIT, approached this challenge designing two different ASICs featuring the first a resistive feedback amplifier and the second one a capacitive feedback amplifier. The device that will appear more efficient in terms of discrimination efficiency, will be upgraded and integrated in a full ASIC embedding the counter and register logic.
This thesis macro-chapter focuses on the design of the capacitive feedback architecture. This 24-channel chip design has been extensively tested with GEANT4 provided files that replicates the amplitude and time distribution of the signals coming from the MoVeIT sensors; the pulse shapes have been modeled using TCAD and Weightfield2 which are dedicated silicon sensor simulation tools. Analyzing the schematic level simulation results, ABACUS behave properly, with a discrimination efficiency of the input files close to 100%. From post-layout simulations the quality of this result is degraded to 85% but it is still not possible to understand the mitigation effect due to the P-Well Blockage since both the substrate models and the dedicated simulator tool licenses were not available. Due to the fact that simulations at the ASIC top level require a huge amount of computing power ($\sim$ 24h for few nanoseconds of transient simulation), it is still not possible to predict the cumulative R-C effects for the sector to sector interconnections and the top level routing. In this application, these effects could be a benefit, smoothing the CML induced perturbations that propagate on ground and substrate paths. The simulation results are going to be compared soon with the on-silicon behavior, once the ASIC prototypes will be delivered.

References

[1] http://www.tifpa.infn.it/projects/move-it/

[2] https://fondazionecnao.it/it/

[3] http://www.tifpa.infn.it/

[4] https://www.lns.infn.it/it/applicazioni/catana.html

[5] N. Cartiglia, et al., Tracking in 4 dimensions, Nucl. Instrum. Meth. A, vol. 845, pp. 47-51, 2016.

[6] F. Cenna, et al., Weightfield2: a fast simulator for silicon and diamond solid state detector, NIM A 796 (2015) 149–153.

[7] Hartmut F-W Sadrozinski1,3, Abraham Seiden1 and Nicolò Cartiglia, 4-Dimensional Tracking with Ultra-Fast Silicon Detectors, IOP Publishing Ltd, Reports on Progress in Physics, Volume 81, Number 2, 2017. INFN, Torino, Italia

[8] J. Allison et al, "Geant4 developments and applications", NIMA 506 (2003) 250-303, "Geant4 developments and applications", IEEE Transactions on Nuclear Science 53 No. 1 (2006) 270- 278, "Recent developments in GEANT4", NIMA 835 (2016) 186-225.

[9] Synopsis, Sentaurus Device User Guide, Version D-2010.03 (2010) 344-348

[10] W. Maes, K. De Meyer and R. Van Overstraeten, "Impact Ionization in Silicon, a Review and Update", Solid-State Electronics, Vol. 33, No. 6, pp. 705-718, 1990.

[11] D.J Massey et al, “Temperature Dependence of Impact Ionization in Submicrometer Silicon Devices”, IEEE Transactions on Electron Devices, vol. 53, no. 9, September 2006.

[12] M. Petasecca et al, “Numerical simulation of radiation damage effects in p-type silicon detectors”, NIMA 563 (2006), 192-195.

[13] Crowell, CR and Sze, SM, Temperature dependence of avalanche multiplication in semiconductors, Applied Physics Letters 9 (1966) 242

[14] CMS upgrade project, https://cds.cern.ch/record/2020886?ln=en&asof=2099-12-31.

[15] H. F.-W. Sadrozinski, “Exploring charge multiplication for fast timing with silicon sensors”, 20th RD50 Workshop, Bari, Italy, May 30 – June 1, 2012, and references therein https://indico.cern.ch/conferenceOtherViews.py?view=standard&confId=175330.

[16] M. Carulla et al, “First 50 μm thick LGAD fabrication at CNM”, 28th RD50 Workshop, Torino, June 7th 2016.

[17] N. Cartiglia, et al., Design optimization of ultra-fast silicon detectors, Nucl. Instrum. Meth. A, vol. 796, pp. 141-148, 2015.

[18] A. Vignati et al., Innovative thin silicon detectors for monitoring of therapeutic proton beams: preliminary beam tests, Journal of Instrumentation, vol. 12, 2017.

[19] https://cividec.at/index.php?module=public.product& idProduct=33&scr=0

[20] http://teledynelecroy.com/oscilloscope/oscilloscopemodel.aspx?modelid=4778

[21] http://www.caen.it/csite/CaenProd.jsp?parent=14& idmod=651

[22] http://www.ptw.de/pinpoint_chambers1.html

[23] G. Kramberger et al., Radiation effects in Low Gain Avalanche Detectors after hadron irradiations, 2015 JINST 10 P07006

[24] M. Mandurrino et al., Numerical Simulatiohttp://1.1.1.1/n of Charge Multiplication in Ultra-Fast Silicon Detectors (UFSD) and Comparison with Experimental Data, NSS Proceedings, 2017.

[25] F. Cenna et al., TOFFEE: a full custom amplifier-comparator chip for timing applications with silicon detectors, Journal of Instrumentation, 2017.

[26] https://cividec.at/index.php?module=public.product& idProduct=33&scr=0

[27] http://teledynelecroy.com/oscilloscope/oscilloscopemodel.aspx?modelid=4778

[28] http://www.caen.it/csite/CaenProd.jsp?parent=14&idmod=651

[29] S Pommé et al., Uncertainty of nuclear counting, Metrologia 52 S3, 2015.

[30] F. Cenna et al., TOFFEE: a full custom amplifier-comparator chip for timing applications with silicon detectors, 2017 JINST 12 C03031

[31] F. Fausti *et al.*, Design and characterization of a 64 channels ASIC front-end electronics for high-flux particle beam detectors, NIM A (2017)

[32] http://www.europractice-ic.com/

[33] F. Cenna et al. TOFFEE: a full custom amplifier-comparator chip for timing applications with silicon detectors. In: Journal of Instrumentation 12 (2017).

[34] A. Rivetti, CMOS Front-End Electronics for Radiation Sensors, CRC Press, 2017

Chapter 4

The Mini-EUSO multi-level trigger integration and testing

4.1 Introduction

Ultra-High Energy Cosmic Rays (UHECRs) are still a mysterious physic phenomenon which origin and composition are mostly unknown. Recent data confirmed the UHECRs extragalactic nature but an increased effort in the observation at extreme energy limit (around 10^{20} eV) is thus needed to identify their primary source. Since at such extreme energy the flux of the UHECRs at Earth is very small (in the order of $\sim$ 1 particle/km^2 /century), UHECRs can be detected by measuring the extensive air showers (EAS) produced as interaction with these cosmic rays in the atmosphere. At ground level, physicist use air fluorescence detectors and large arrays of particle detectors, like in the Pierre Auger Observatory [1] [2] (Auger, which cover 3000 km^2) and Telescope Array [3] [4] (TA, which cover 700 km^2). The intrinsic geometrical limit of these ground-based experiments can be overcame reversing the direction of the exposition with a space detector looking at the Earth atmosphere to increase by orders of magnitude the FoV at ground. With this purpose the JEM-EUSO collaboration (Japanese Experiment Module, Extreme Universe Space Observatory) bases the study of UHECRs on the detection of the secondary light emissions induced by cosmic rays in the atmosphere, maximizing the statistics of these rare events through a wide field of view from space. Chasing such a goal, a good strategy is to plan and develop intermediate pathfinder experiments.

The JEM-EUSO collaboration scheduled the work flow of such ambitious project planning the realization of several intermediate pathfinder experiments. Mini-EUSO, along with EUSO-TA, EUSO-Balloon, and EUSO-SPB, forms the next step towards the observation of UHECRs from space. Mini-EUSO is a JEM-EUSO 1:10 in-scale telescope, designed to trace a UV-map of the Earth from the vantage point of the International Space Station (ISS) in low Earth orbit. A multi-level trigger logic operates data selection over different time scales thus optimizing the collection efficiency through a sort of on-line data analysis. Owing to the limited amount of storage available, an efficient trigger system has to be developed. Such a system has to be deeply and exhaustively tested both at software level and during the hardware integration in the processing system. This thesis chapter first introduces the Mini-EUSO scientific context and features and then focuses on the motivation behind the trigger logic. In the second half of the chapter the topic moves to the integration technicality and the testing phase.

4.1.1 The author's contribution

The author's contribution on the work presented in Chapter 4, is related to the Mini-EUSO first level trigger logic. During one year work (six months in parallel with the TERA09 project and six months full-time), the author interacted with the the Mini-EUSO trigger responsible and the department technician that developed the firmware for the trigger algorithm. After this support and consulting phase, the author moved to the Skobelstyn Institute of Nuclear Physics of the Moscow State University (where the FPGA main developer works), to develop from scratch some trigger ancillary IPs like the Artificial Data Generator (ADG). The ADG allows to test the trigger stand-alone, after its hardware implementation, stimulating the trigger logic (suitable for both, first and second level trigger) with signals of different complexity levels; in this way it is possible to discriminate electrical problem from algorithm ones. Other IP blocks developed are a time stamp generator which attaches a time reference to trigger events and a pixel masking block that is suitable to reduce the wrong data saving, in case of bad pixel behavior. Further efforts provided by the writer, have been collaborating to both a readout chain test (from photomultiplier tubes to the main FPGA) with a pulser and the Mini-EUSO test in TurLab. The results of these studies, as well as the detector hardware and logic description, will be presented in the following sections.

4.2 Mini-EUSO in the behalf of the JEM-USO collaboration

Conventionally, UHECR refers to the extreme part of the detected particles energy spectrum. In the cosmic ray physics, recurrent fundamental questions are:

- how does the cosmic-ray spectrum continue beyond the existing data? is there a maximum energy?
- is there an anisotropy that indicates source regions?
- are UHECRs protons, nuclei, photons, neutrinos or exotic particles?

In the energy spectral region above $10^{19}eV$, the amount of available physics information is inversely proportional to the rate of the detected particles, leading to an extremely poor statistics above $10^{20}eV$. World wide research groups agree to the fact that space mission are needed to increase the acceptance as well the exposure and the statistics oriented to this kind of detection campaign. The expensive and heavy effort demanding nature of a space mission, favors as consolidate succeeding strategy of being carried on by international collaborations like EUSO.

EUSO has been an European Space Agency (ESA) mission, designed to be hosted on the International Space Station as an external payload of the Columbus. Even if EUSO successfully completed the "Phase A" study, programmatic but mainly financial constraints brought ESA to suspend the program in 2004. Although this problematic initial phase, the project goal was at a later time re-oriented as a payload to be hosted on board the JEM module nicknamed Kibo (hope, in Japanese), the Japanese science module for the ISS developed by Japan Aerospace Exploration Agency (JAXA). At that point the collaboration renamed the mission as JEM-EUSO. JEM-EUSO is the first space mission concept devoted to the investigation of cosmic rays (CR) and neutrinos of extreme energy ($E > 5\times 10^{19}eV$). Using the Earth's atmosphere as a giant gas detector, the detection is performed by looking at the streak of fluorescence produced when such a particle interacts with the Earth's atmosphere.

JEM-EUSO is currently studied by hundreds pf researcher in 95 Institutes from 16 Countries, with the support of the most important International and National Space Agencies, aiming for a flight after 2020. The proposed instrument consists of a set of three large Fresnel lenses of 2.65 m diameter (with top and bottom cut

off to reduce the minimum diameter to 1.9 m so that they fit in the resupply vehicle in which the instrument has to be launched) feeding a detector consisting of 137 modules each a 48x48 array of photomultiplier tubes. The imaging takes place in the 300-450 nm band (low energy UV through deep-blue), and photons are time-tagged with 2.5 microsecond precision [5].

4.3 The scientific scenario and motivations

The main scientific goal of Mini-EUSO is to produce a high-resolution Earth map in the 290 nm - 430 nm UV range; this radiation sensitive window has been chosen considering the nitrogen molecules abundance in atmosphere and the fact that excited nitrogen atoms have a 268 nm-546 nm second positive line in the emission spectrum. This map is crucial since the event detection is based on the signal discrimination from a background pattern and the event itself is indirectly detected looking at the UV emission as cosmic rays-atmosphere interaction result. With a spatial resolution of $\sim$5 km and a temporal resolution of 2.5 µs, Mini-EUSO is expected to find new information regarding UHECRs detection threshold from space, estimating the duty cycle of future experiments.
Even if the Mini-EUSO energy threshold is too high to directly detect UHECRs, laser-induced CR-like tracks will be provided by ground shooting, allowing the validation of the trigger logic principles and testing the instrumentation developed for UHECRs detection. Furthermore Mini-EUSO presents the opportunity to study a variety of other scientific phenomena, including atmospheric physics, strange quark matter, space debris detection and bio-luminescence and anthropogenic lights. Considering data resulting from previous missions, a minimum flux of $3\times 10^{11} photons \times m^{-2}\, s^{-1}\, sr^{-1}$ has been detected by the Tatiana experiment, with a 100 km spatial resolution [6]. This value came out as minimum for dark areas of the Earth during moonless nights. A factor 2-5 higher is expected over clouds or cities, whereas a factor 1-2 higher over aurora regions. Higher resolution measurements of the UV ground emission have been performed by pressure flying balloons like EUSO-Balloon and BABY [7]. These detection campaign with a 10 km spatial resolution were located at less than 40 km altitudes thus excluding high altitude peculiar events such air-glow emissions and aurora. Furthermore there are Transient Luminous Events (TLEs), typical for being upper-atmospheric, in the UV range and occurring

with high repetition rate [8], [9] (more detail in the next subsection). Characterizing TLEs like events helps to increase the UHECRs detection efficiency. In this rare-events region the flux is as low as 1 particle/km^2 x century and thus the effective area that can be observed by a detector is a key feature. The JEM-EUSO collaboration actually chose a space-based detector in order to increase the sensitive volume for a more complete statistics [10], [11]. Prior to reach this goal, intermediate pathfinder experiments have been developed by the collaboration. Mini-EUSO, as well as EUSO-TA [12], EUSO-Balloon [13], and EUSO-SPB [14], forms the pioneer set of detectors looking at UHECRs from space. Additionally, UV night glow measurements are currently being conducted with a spatial resolution of 5 km by the TUS experiment on board the Lomonosov satellite [15]. The UV radiation measured by satellite missions is highly variable due to the presence of clouds, cities, aurora and other factors in the moving field of view. Mini-EUSO and TUS have a similar spatial resolution $\sim$5-6 km and their space missions are going to catch unprecedented high resolution informations. The Mini-EUSO mission, originated as a joint project between Italy and Russia, was selected in Italy by the Italian Space Agency (ASI) and is supported by the National Institute of Nuclear Physics (INFN); then, under the name "UV atmosphere", it has been approved by the Russian Space Agency Roscosmos which included it in the long-term program of space experiments on the ISS. After the signature of a common Scientific Agreement, Mini-EUSO is now an established project between the participating countries of the JEM-EUSO collaboration. The Mini-EUSO launch is scheduled for the end 2018 (or early 2019) and the detector is going to be installed inside the International Space Station (ISS), placed at a nadir-facing, UV-transparent window on the Russian Zvezda module [16].

4.3.1 UHECRs: basic concepts

The detection of Cosmic Rays (CRs) with energies greater than 10^{18} and even exceeding 10^{20} eV is performed at Earth ground with almost random arrival directions, classifying these phenomena origin well far away the Solar System. A key role in the CRs study is played by the energy spectrum that falls off close to a power-law function (average power index of $\gamma \sim 3$). From 10^{14} eV and downward, the CRs flux is very low and forces to an indirect detection approach by measuring secondary particles, or the extensive air showers (EAS) produced by the primary

CR particles in the atmosphere. The ultra-high energy CR spectrum extends by over ten orer of magnitude and has been characterized by several experiments [17]. At energies of $\sim$5 x 10^{19}eV, there is a sharp cut-off in the cosmic ray flux to the level of $\sim$ 1 particle/km^2 x century [18]. This phenomena, currently named the GZK cutoff, has been independently predicted in 1966 by Greisen [19], Zatsepin and Kuz'min [20] whose theories explained it as a consequence of photo-pion production resulting form the interaction between the CRs with the low energy cosmic microwave background (CMB) photons. In the case of a pure proton composition, the explanation basis on the electron-positron pair-production from CR protons with the CMB photons interaction [21]. Considering a mixed composition, propagation effects are complicated by the fact that the primary nuclei also suffer interactions that cause fragmentations. Other alternative models point to a cut-off originated from acceleration mechanism features at the source [22]. Due to the rare event statistics, the observation of the cut-off in the energy spectrum requires detectors with a large effective area working for tens of years keeping a good energy resolution. AGASA (Akeno Giant Air Shower Array) [23] and Hi-Resolution Fly's Eye (HiRes) [24] have been the first cosmic ray detectors working in the energy spectrum of UHECRs above 10^{19} eV, as represented in 4.1. AGASA used an array of scintillation counters to detect EAS particles at the ground level, while HiRes was based on fluorescence detectors sensitive to fluorescence light emitted due to the energy deposition of the EAS particles in the atmosphere.

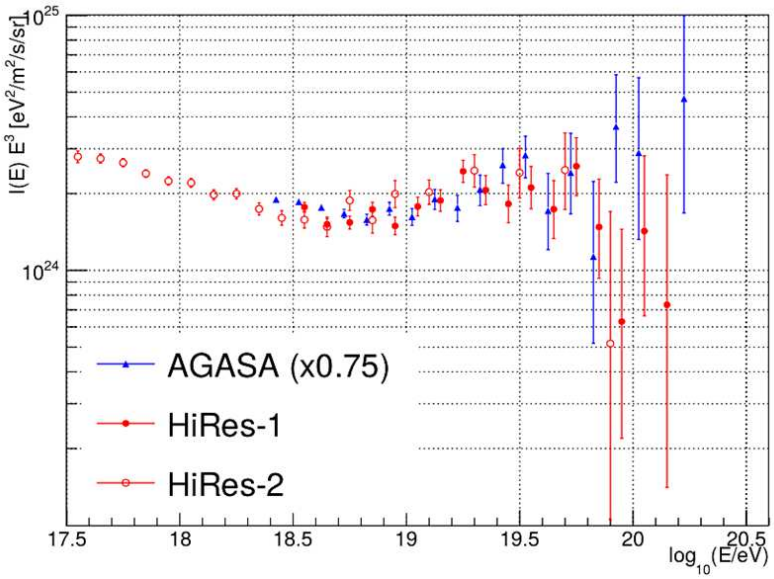

Fig. 4.1 Energy spectra measured by the AGASA and HiRes. The energy of AGASA data points is scaled down to 75% to match the position of the ankle with HiRes at $\sim 5 \times 10^{18}$ eV. Picture from [18]

Since the systematic uncertainties in determining the CR primary energy were $\sim$20% in both cases, in modern experiments a hybrid detection technique is adopted, where the CR EAS are simultaneously observed with fluorescence detectors (FDs) and surface detectors (SDs), allowing a very precise determination of the CR energies and arrival directions. The FDs measure fluorescence light emitted by the atmospheric molecules excited by the charged particles in the EAS, and observe the longitudinal development of the EAS using mirror telescopes coupled with clusters of photo-multiplier tubes. The FDs operate at a $\sim$10% duty cycle because the FD data can be collected only during nights with low moonlight background and with dry air and clear skies. The SDs, on the other hand, directly measure EAS particles at the ground level at a nearly 100% duty cycle, regardless of the weather conditions. Regarding the experiment setup, alternative possibilities respect to ground-based detectors are represented by atmospheric balloons and space orbiting telescopes, as the EUSO collaboration experienced during its evolution. From tens or even hundreds of km altitude, the FoV at ground is easily extended by order of magnitude moreover, orbiting detectors allow an improved statistics due to their periodic motion around the Earth. Although all these positives, a flying devices like Mini-EUSO has

to deal with a lot of technical difficulties due to its motion, eterogenesous luminosity conditions etc., as explained in this chapter.

4.3.2 TLEs: basic concepts

Since their discovery in 1989, TLEs became an hot topic in the atmospheric physics study, leading in the following years to various data acquisition campaigns by aircraft, satellite and even by space shuttle and ISS. The nature of these phenomenons is characterized by a bright milliseconds duration, a reduced spatial scale (1 km to tens of km) and mesosphere altitudes (TLEs generally appears above the clouds level, 50 km-100 km high and they rarely interest the ionosphere). In 1925 the Scottish physicists C.T.R. Wilson elaborated the first experimental evidence related to the connection with the thunder cloud discharge (light source) due to the classic breakdown mechanism as result of an electric field threshold limit crossing. More than 60 years later Robert Franz, Robert Nemzek and John Winckler, University of Minnesota scientists, experimentally discovered a red sprite in night time. In 1998 the first TLE has been detected in daytime by Mark Stanley Schneider [25]. Apart from red sprites, other types of TLE like blue jets, elves, halos and trolls have been observed and classified as reported in Figure 4.2. It has been estimated that the energy production by all types of TLEs in the atmosphere is about 700 MJ per minute [26]. Sprites and elves follow cloud to Earth discharges with a strong correlation between events-rate and thunderstorms-frequency.

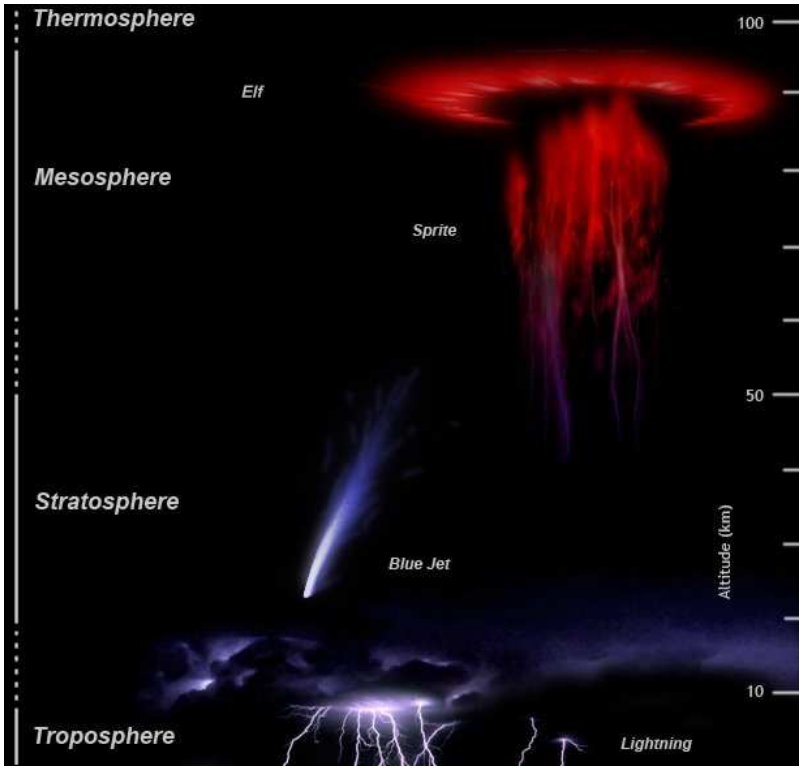

Fig. 4.2 A schematic overview of different types of the Transient Luminous Events. Aggiungere fonte

ELVES Emission of Light and Very Low Frequency perturbations due to Electromagnetic pulse Sources, are low ionosphere concentric rings, expanding at light-speed up to 200 km-500 km diameters in one millisecond. An hypothesis on the ELVES production considers a rapid electric field change as result of a strong lighting. This field gradient performs an upward electromagnetic pulse propagation showing as light emission once it reaches the ionosphere. From [26], the world wide ELVES occurrence rate is 35 per minute. The spatially averaged brightness of an elves is: 0.17 ± 0.08 MR = $1.36 \cdot 10^{14}$ ph sr^{-1} m^{-2} s^{-1}. The unit R stands for "Rayleigh" (unit of photon flux) typically used for these atmospheric phenomena. The observed characteristics of different types of TLE is depicted in Figure 4.3.

Sprites are low-luminosity massive events appearing at altitudes of 40 - 90 km. Like a giant storm clouds 1000 km and over widespread produce strong electric field in the mesosphere, characterized by a bright region at 65-75 km above which it extends a red glow up to 90 km whereas below, blue structures like filaments extend

downward to 40 Km. Usually two or more sprites are observed together for a typical milliseconds time duration. Measurements of microsecond time resolution may help to explain the mechanism behind the peculiar shaping phenomenon. The measured brightness is 1.5 $\pm$ 1.1 MR, which translates into some 3.48 $\cdot 10^{11}$ ph in total and 1.46 $\cdot 10^{11}$ in the wavelength region detected by Mini-EUSO. This number is the spatially integrated amount of photons recorded throughout the whole event duration.

Blue Jets differ from sprites in that they are optical ejections from the top of the electrically most active regions of the cumulonimbus above a thunderstorm. a Blue Jets propagate with a 100 Km/s vertical speed with about 15 degrees full width narrow cone, 400 m large at the base that disappears at about 40 - 50 km altitude. Brighter than sprites, Blue Jets are more rare and blue colored events not related with lighting occurrences but triggered by the cloud inner electric field. The brightness of blue jets is estimated as 0.5 MR [27]. They were firstly recorded on October 21, 1989, on thunderstorm video on the horizon taken from the Space Shuttle as it passed over Australia.

Giant Jets are originated from the cloud top thus developing in a cascade until the lower ionosphere. As the name claims, Giant Jets are much more extended events than Blue Jets and they propagate at the same speed of the former.

Type of TLEs	Altitude Regime	Transverse Dimensions	Spatial Characteristics	Apparent Motion	Duration
Sprites	$\sim 50-90$ km	$\sim 1-20$ km	Top (>80 km) diffuse Bottom (<70 km) structured	Top – upward Bottom – downward	few ms
Elves	~ 90 km	>100 km	Diffuse	Lateral expansion	$\sim$ ms
Blue jets	$\sim 18-45$ km	few km	Structured	Upward	hundreds ms
Giant Blue jets	$\sim 18-75$ km	few km	Structured	Upward	hundreds ms
Hallos	~ 75 km	50 km	Diffuse	Downward	$\sim$ ms
Trolls	$\sim 60-70$ km	kms	Structured	Upward within decaying sprite tendrils	hundreds ms

Fig. 4.3 Principal TLEs types in the upper atmosphere [28]

4.4 The Mini-EUSO telescope: Instrument overview

The Mini-EUSO telescope consists of three main sub-systems: the Fresnel-based optical system, the Photo Detector Module (PDM) and the readout electronics. The idea behind the Mini-EUSO project is to test a single JEM-EUSO detection unit, the PDM, consisting of 36 Hamamatsu Multi Anode Photo Multiplier Tubes (MAPMT Hamamatsu R11265-M64 [29]), 64 independent pixels each, for a total of 2304 pixels. Two double sided flat, UV transparent Fresnel lenses made of PMMA (polymethyl methacrylate) and 25 cm diameter each form the optical system, which focuses the light onto the PDM Focal Surface reaching a $44^\circ \times 44^\circ$ Field of View (FoV) at r = 85 mm on the focal surface. The short diameter lenses and their 11 mm thickness, reduces the optics mass for a light and compact system ($\sim$0.3 kg/lens), as required in space applications. The Mini-EUSO optics has a low focal number F= 0.6, and the effective focal length is 150 mm. As previously reported, the PDM detects UV photons (300 nm - 400 nm) with a readout sampling period of 2.5μs, having a spatial resolution at ground of $\sim$6 km. A custom software code has been

developed for the photon collection efficiency (PCE) study. This model includes and takes into account several loss factors as surface reflection and roughness, material absorption, Fresnel facet back cut and support structure obscuration. The resulting data is a $\sim$45% PCE.

In addition to the main detector, Mini-EUSO includes two ancillary cameras for complementary measurements in the near infrared (1500 nm - 1600 nm) and visible (400 nm - 780 nm) range with the main task of performing atmospheric monitoring measurements (as the thermodynamic phase detection of clouds) [30]. Moreover cameras will help in the measurements of the emissions of the Earth and the study of transient phenomena. Mini-EUSO will be housed in a space qualified mechanical box (Al Ergal). Fig. 4.4 shows the CAD design of the container box with an inner view of the detector components.

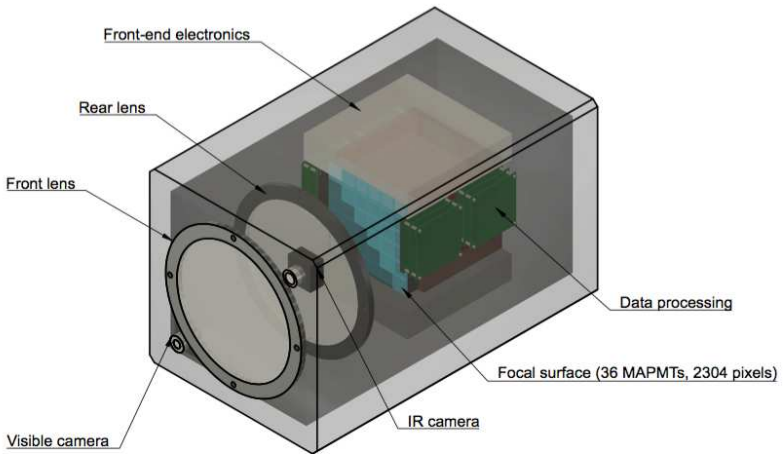

Fig. 4.4 A CAD representation of the Mini-EUSO instrument. The main sub-systems are shown: the two double-sided Fresnel lenses, the PDM and the readout electronics. The near infra-red and visible cameras are mounted at the level of the first lens, outside of the optical system. The dimensions of the instrument are $37 \times 37 \times 62\ cm^3$.

A special adapter provides the interface connection with the nadir looking, UV-transparent window of the Russian ISS module Zvezda.

4.4.1 The digitization data path

The data acquisition system,incorporating the functionality of several subsystems into one single board, is an evolution of the system adopted and validated on-field

in the previous EUSO pathfinders, such as EUSO-TA, EUSO-Balloon and EUSO-SPB. The reference module, called PDM consists of an array of 3x3 Elementary Cells (ECs), each of which has 4 MAPMTs. The single MAPMT is readout by a SPACIROC3 ASIC [31], [32], a 64 channels chip in AMS 0.35 μm SiGe technology. Each channel is related to one MAPMT pixel. Every EC has a dedicated PCB called EC ASIC Board that works in LVDS logic managing the 40 MHz SPACIROC3 throughput. The MAPMTs generates either pulsed signals, elaborated in a discrete single photon counting mode or DC signals, elaborated by means of a charge-to-time conversion, for photon flux intensity measure. This data is digitized every 2.5 µs acquisition window which is hereafter referred as a gate time unit (GTU). The output of the SPACIROC3 ASIC is then passed to the photo detector module data processing system (PDM-DP) that was specially developed for the experiment. The PDM-DP consists of three boards, the Cross Board, the Zynq Board and the Power Board as shown in Figure 4.5. The Cross Board contains three synchronized Xilinx Artix7 FPGAs ([33]) to perform data gathering from the EC, pixel mapping and data multiplexing. The three Artix7 placed in a row are synchronized and organized with the central FPGA working as the master and the lateral ones as slaves. Data leave the Cross Board in a 48 $\times$ 48 bits format transferred at 200 MHz, using a 100 MHz clock in DDR3 mode, reaching the ZYNQ Board. As well the SPACIROCs are synchronized from Artix7 FPGAs (that pass a DDR clock signal), the ZYNQ Board is synchronized by Artix7 master too.

The Zynq Board interfaces the cross board and contains a ZYNQ XC7Z030 system of programmable logic (PL) Xilinx Kintex7 FPGA, with an embedded dual core ARM9 CPU Processing System (PS). The Zynq Board does the majority of the data handling including data buffering, slow configuration of the SPACIROC3 ASICs, triggering (both Level I trigger, Level II trigger are managed by the ZYNQ Board), synchronisation, and interfacing with the separate CPU system for data storage. In addition to these tasks, the high-voltage control to the MAPMTs is also under the Zynq Board control. The Power Board provides the necessary voltages to the system. Figure 4.6 summarizes the digitization data path. A multi-level trigger [34] is implemented in the Zynq Board for the mini-EUSO instrument to perform measurements in various time scales (temporal resolutions of 2.5 μs, 320 μs, 40 ms) in order to maximize the scientific output efficiency of the instrument while limiting the volume of the stored information. This trigger was successfully implemented and tested during laboratory trials. The CPU, a PCIe/104 form factor, performs the

control of the instrument sub-systems as well as the data management and storage, housekeeping, switching between operational modes and collecting data from the infrared and visible cameras. Estimating the storing flux, we know that the ZYNQ Board receives data from an EC Board hosting 4 SPACIROC3s, multiplied by 9 EC Boards thus 36× 400 KHz (2.5 μs GTU windows) x 64 pixels (SPACIROC3s channels) x 8 bits/pixel = 921.6 MB/s. Such amount of data is managed by internal BRAMs (Block RAM), 2 MB deep. Therefore the system is able to store 2304 Bytes for GTU×128 GTUs x 4 = 1.2 MB (in case of multiple triggered events storage, as better explained in later sections). In case the storage of more than 2 MB data is needed, PL DDR memories should be used (where PL is the Programmable Logic embedded into the ZYNQ chip). These 16 bits memories are 512 MB deep, with a 533 MHz x 2 (DDR3 format) x 2 (2 bytes for the 16 bits) bandwidth. The incoming data speed is about 1GB/s therefore the 512 MB memory has a 0.5 sec autonomy.

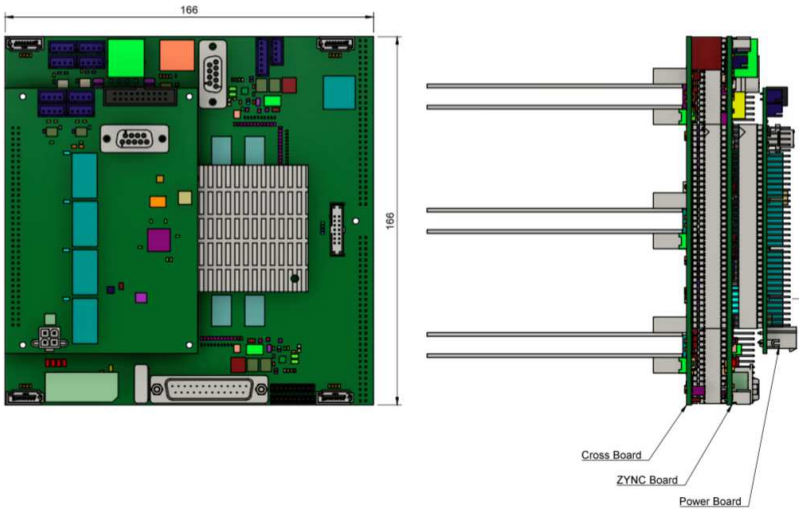

Fig. 4.5 The PDM-DP is shown with dimensions in mm. The 3 separate boards are shown with the mechanical support for the SPACIROC3 ASICs on their left.

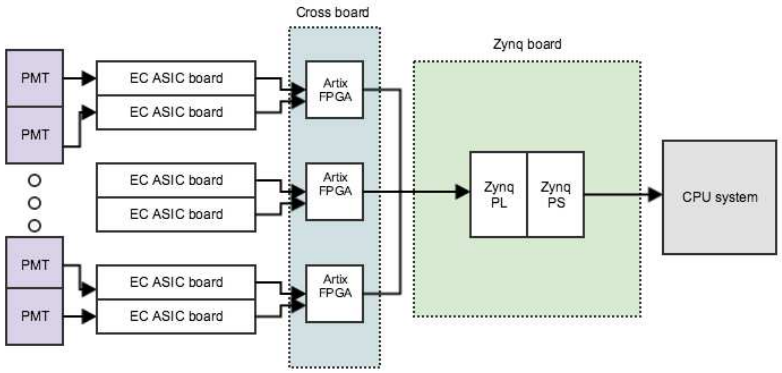

Fig. 4.6 A schematic representation showing the detecting system digitization data path.

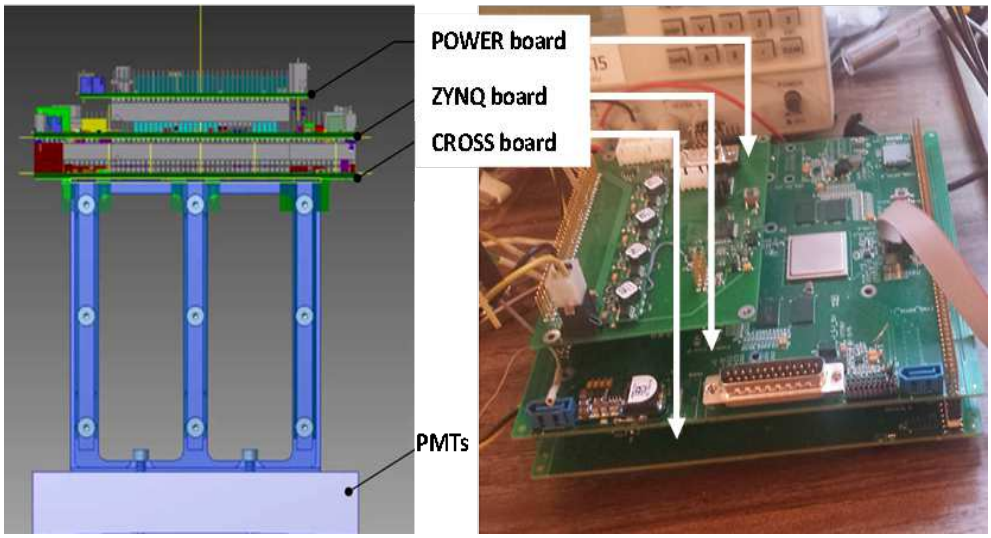

Fig. 4.7 CAD modeling of a part of the Mini EUSO mechanical structure (left). Picture of the Power Board, ZYNQ Board and Cross Board, assembled in a PCBs commpact block (right).

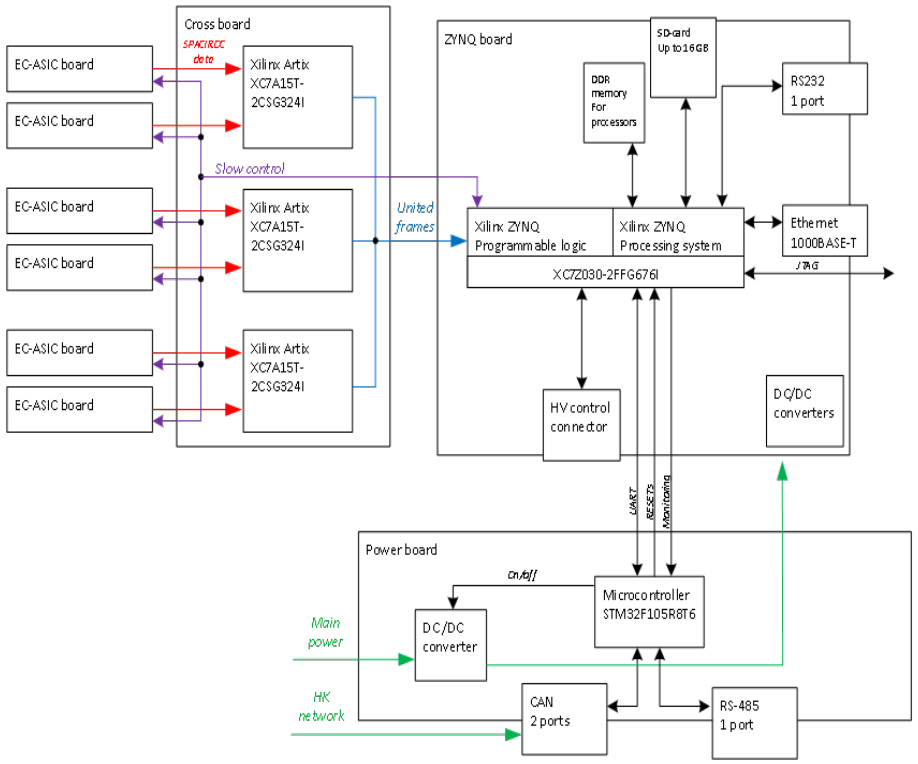

Fig. 4.8 A schematic representation showing the interconnection among the PCBs whit inner components details.

4.4.2 The multi-level trigger

Looking for UHECR-like events, Mini-EUSO is also capable of capturing a variety of both atmospheric and terrestrial UV phenomena such as TLEs, meteors, space debris, strange quark matter (SQM), bio-luminescence and anthropogenic lights. The duration of such events varies on the order of 10^6 s, motivating a multi-level trigger system to maximize the scientific return, given constraints on the duty cycle and data storage. The Mini–EUSO trigger logic is implemented in VHDL inside the PL of the ZYNQ Board and consists of two levels, level 1 (L1) and level 2 (L2), that work with different time resolution. Each level is dedicated to a specific category of events that will be observed by Mini-EUSO. The motivation behind the trigger algorithm is to capture different events of interest on short timescales, but also to provide continuous imaging on slower timescales as Mini-EUSO orbits around the Earth. In order to achieve this target efficiently, 3 different types of data are stored with different time resolution.

The L1 trigger returns data with a time resolution of 2.5 µs and looks for signal over threshold on a timescale of 20 µs, as this corresponds to the timescale of UHECR–like events. Each pixel is considered as independent, motivated by the fact that the field of view at ground is ∼6 km, so light takes ∼20 µs to cross one pixel FoV at ground.

The first level trigger has three tasks:
a) it searches for signal exceeding the threshold on a pixel integrating 8 consecutive GTUs.
b) it accumulates data in bunches of 128 GTUs (320 μs).
c) it sends data to memory to be accumulated at steps of 2.5 μs.

A brief description of the three macro steps is hereafter reported.

a The SPACIRO3 ASIC output data, moving through the Artix Boards, is accumulated in a memory with 8 GTU depth. Starting from the 8th GTU, for each GTU a check is performed to verify if one pixel has accumulated in the past 8 GTUs a signal higher than a prefixed threshold, which is set at 8 sigma above the average background calculated on bunches of 128 GTUs. If 1 pixel has a signal higher than $8\sigma \times$ the background threshold, a L1 trigger event signal (L1Event) is issued, the whole focal surface is read out and a packet of 128 GTUs are saved in memory.

b Data coming from the ASIC is accumulated in a register. This register stores the integral of 128 GTUs (μs) for each pixel. These values are used to set the L1 trigger thresholds at pixel level. In parallel, every 320 μs these data are transferred to the L2 trigger logic.

c Data coming from SPACIROC is sent to the memory for storage. These data can fill only a memory depth of 128 GTUs, therefore data is overwritten. When L1 trigger occurs, a write-disable signal is sent to the memory after some latency (64 GTUs after, if we want a trigger-event-centered data storage) and data cannot be written anymore in the memory. L1 logic continues performing point b), which is the creation of 320 μs-long GTUs.

The L2 trigger receives the integration of 128 GTUs (= 1 L2 GTU, or GTU_{L2}) as input from the L1. It operates with a similar logic, but with a time resolution of 320 µs, well-suited to capture atmospheric events, such as TLEs and lightning, which have timescales of <1 ms. Background is set by integrating 128 GTU_{L2}, which is also stored as the level 3 (L3) data, or 1 GTU_{L3}. An L2 trigger occurs when the signal in 8 GTU_{L2} is greater than 4 times the background level, and the event is stored.

After the accumulation of 128 GTU_{L3}, or every 5.24 s, all stored events from L1, L2 and the continuous storing (named L3 data) are transferred to the CPU for formatting and storage on the disk. If no L1 or L2 events are triggered, the last 128 GTU_{L2} present in the buffer are read out. In this way, a continuous and controlled readout is achieved with a resolution of 40.96 ms whilst also capturing interesting events at faster timescales. This 40.96 ms "movie" will be used to search for meteors, space debris and strange quark matter using off-line trigger algorithms, as well as for the mapping of the Earth in UV. The L1 and L2 trigger algorithm is summarized in Figure 4.9. In principle, key parameters such as the threshold and the duration of signal integration can be altered to optimize the trigger performance.

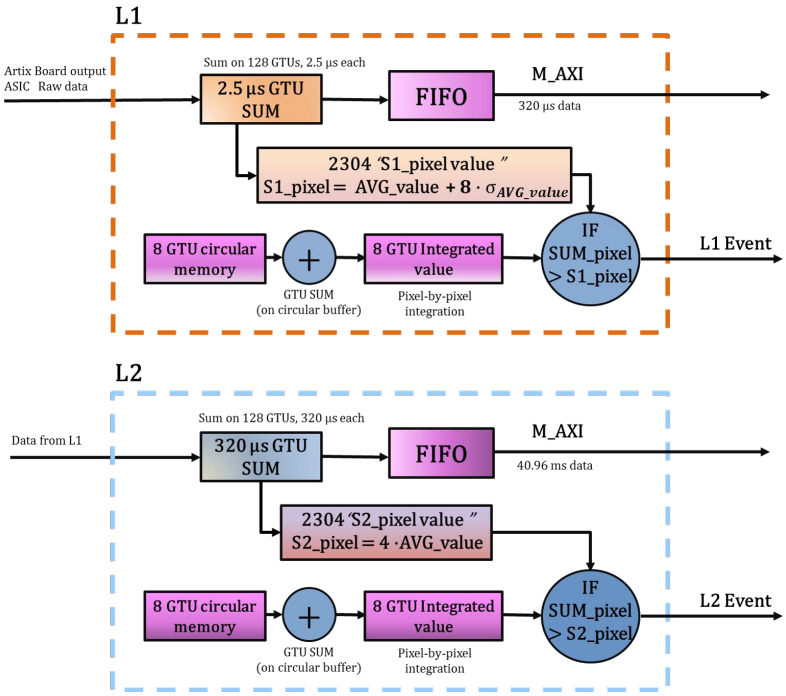

Fig. 4.9 A block diagram summarizing the trigger logic. Top: L1, bottom: L2. The trigger outputs 3 separate types of data with time resolutions of 2.5 µs, 320 µs and 40.96 ms.

The memory buffer will, therefore, contain:

- the latest 320 μs (128 GTUs) of the 5.24 s sequence with 2.5 μs time resolution if no L1 trigger occurs. In case of L1 trigger, the 128 GTUs around the over threshold event.
- the latest 40.96ms of data of the 5.24s sequence with 320μs time resolution if no L2 trigger occurs. In case of an L2 trigger, the 128 frames of 320 μs each, for a total of 40.96 ms around the L2 trigger (e.g. lighting events) will be stored.
- 128 frames of 40.96 ms each of the last 5.24 s which represent a continuous movie taken by MINI-EUSO from the ISS.

The 40.96 ms time frames are similar to what is typically done from ground for meteors (30 frames/s). A data analysis will be performed on ground looking for meteors, nuclearities, and space debris as well as for reconstruct UV maps, search for bio-luminescence, etc. The 320 μs long frames will record lightnings, TLEs and other types of events which last several ms and in which a few hundreds μs resolution is enough. The short 2.5 μs resolution frames should contain events which last a few GTUs (tenths of μs), therefore on the cosmic ray time scale.

As previously reported, L1 and L2 thresholds are set to trigger, on average, at a rate lower than 1 event per 5.24 s. Assuming that 3 byte/pixel are recorded, the presented trigger algorithm gives a data readout of 507 kB/s. Assuming an optimistic duty cycle of 50%, this results in a data storage requirement of 660 GB/month. Assuming some ancillary data from the camera and housekeeping systems, it is still reasonable to estimate a maximum data output of 1 TB/month.

4.5 Verification of the trigger algorithm

Prior to the implementation of the trigger algorithm in hardware, the logic has been tested extensively using both simulated data and data taken at the TurLab facility in the Physics Department of the Turin University.

4.5.1 L1 trigger tests at TurLab

The EUSO@TurLab project is an ongoing activity aimed to reproduce atmospheric and luminous phenomena that the JEM–EUSO and EUSO style telescopes will observe from Earth orbit. TurLab is a laboratory equipped with a 5 m diameter rotating tank and located 15 m below ground level. Therefore, without artificial illumination, the room is darker than the night sky by several orders of magnitude. The EUSO@TurLab project makes use of the TurLab rotating tank with a series of different light configurations to reproduce the UV emission of the Earth. The Mini–EUSO detector is represented by one elementary cell (EC) unit of 4 MAPMTs and the necessary readout electronics. The detector is suspended from the ceiling and looks down on the rotating tank to mimic the observation from orbit (see Figure 4.10).

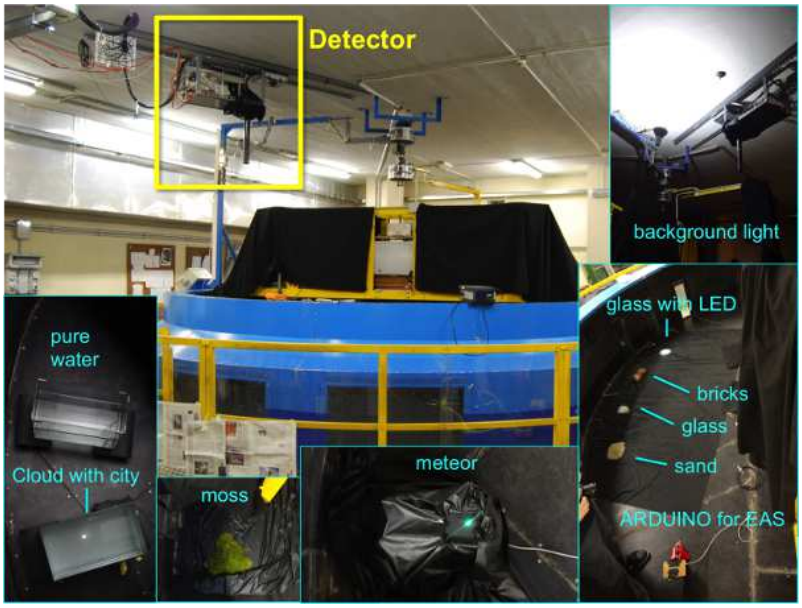

Fig. 4.10 The TurLab rotating tank. The black tube on the ceiling shows the collimator of the experimental setup used to mimic the Mini-EUSO telescope. Light sources and materials used to mimic other UV sources are also shown.

The capability of controlling the tank rotation speed (3 s -20 minutes per turn) allows for the reproduction of events of different duration and spatial extent, as seen from ISS, with the same configuration.

Vital to the testing of the trigger algorithm in this setup is the choice and variety of light sources. There are two types of light source: 1) direct light emitting sources; 2) materials reflecting ambient light. A range of different light sources are used, with

the intent of reproducing different kinds of phenomena: a) LEDs inside tubes of different dimension, in order to reproduce extended intense light directly pointing towards the MAPMT: this is to represent urban areas; b) an oscilloscope generating Lissajous curves for events such as meteors; c) LEDs driven by a pulse generator for fast luminous events such as lightning; d) LEDs or optical fibers driven by and Arduino board for light pulses with μs duration. A more detailed discussion of the setup is reported in [35].

For the EUSO@TurLab measurements, the apparatus consists of one fully-equipped EC unit, similar to those used in Mini-EUSO, with a 1 inch focusing lens (50 cm focal length) placed directly in front of the MAPMTs. A test board is used to retrieve data from the EC ASICs and a LabView program is used for data acquisition. The main differences between the TurLab setup and Mini–EUSO are that data is acquired in packets of 100 GTU instead of the nominal 128 GTU, and the system has a $\sim$50 ms delay between two consecutive acquisitions of 100 GTU. This condition slows down the measurements and introduces artificial discontinuities in the recorded light between two acquisitions. To avoid them, 200 simulated data packets were added between two experimental packets in order to smooth out such discontinuities. In this way it was possible to extend the 8.2×10^5 collected GTUs in about 7 minutes rotation, to a total number of 1.6×10^8 GTUs used to test the L1 trigger off-line.

Figure 4.11 shows an example of the performance of the trigger logic described in Section 4.4.2 for one EC unit. The figure is divided in 4 different blocks. In each block the top plot shows the average number of counts per pixel normalized at PMT and packet level as a function of time for one PMT, while the bottom plot indicates the location in time when the L1 triggers were activated. The different letters (from A to I) in the plot of PMT 2 indicate different types of light surface or reflective source present in the tank which are responsible for a different light intensity seen by the PMTs. The same pattern is apparent in all 4 PMTs but with different intensities and slightly shifted in time due to the movement of the tank and the size of the light source.

Pictures of these sources are displayed in Figure 4.10. A represents clouds; B and D represent the response to ground glass in which D looks brighter because the glass is illuminated by a led; C, E and F is the reflection from sand, brick and moss, respectively; G is due to meteor-like signals, while H to Arduino-emulated cosmic ray and I to the reflection of pure water. The Arduino event looks quite dim

compared to other signals because the track is limited to a few pixels, therefore, it is almost overwhelmed by the total number of counts in the PMT.

Despite the presence of several light sources of different intensity, duration and extension, most of the triggers occur in coincidence of the Arduino UHECR–like signal transit in the field of view of the telescope for all four PMTs. The rate of spurious triggers is $\sim$0.2 Hz, therefore, compatible with the acquisition logic.

The most significant portions of these data were tested also with the VHDL code implemented in FPGA. and same results were obtained. These result demonstrate that the L1 trigger is sensitive to the presence of UHECR–like light signals.

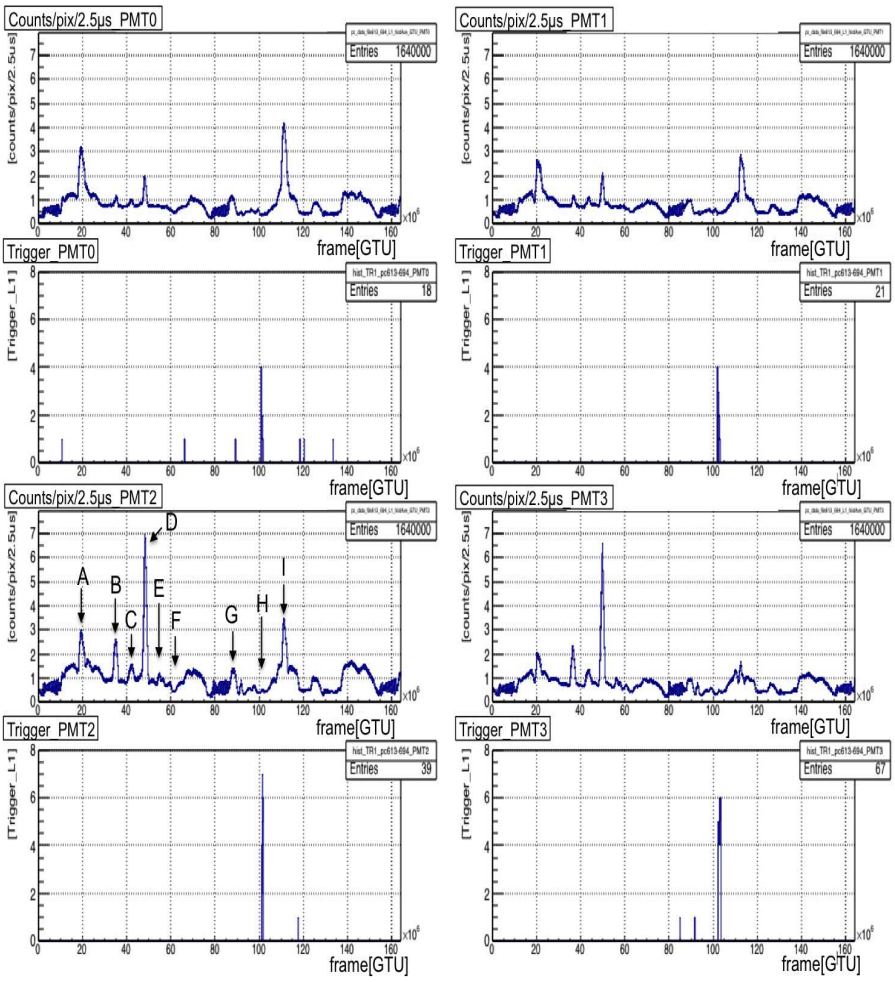

Fig. 4.11 The figure is divided in 4 different blocks. In each block the top plot shows the average number of counts per pixel normalized at PMT and packet level as a function of time for one PMT. The bottom plot indicates the time of L1 trigger activation. See text for details.

4.5.2 L1 trigger tests with ESAF

The main objective of the TurLab tests was the verification of the capability of the L1 trigger logic and the optimization of the trigger thresholds with the variations of light intensity. This is important in order to keep the rate of false triggers at an acceptable level. The logic demonstrated the capability of recognizing and triggering on EAS-like signals. Events of longer duration such as meteors, city lights, clouds, etc. do not generate triggers, as required.

In order to evaluate the trigger performance for UHECR observation, simulations using the ESAF code were performed. The EUSO Simulation and Analysis Framework (ESAF) [36] is currently used as the simulation and analysis software for the JEM-EUSO and its pathfinder missions. ESAF performs the simulation of the shower development, photon production and transport in the atmosphere, and detector simulations for optics and electronics. Furthermore, algorithms and tools for the reconstruction of the shower properties are included in the ESAF package [37]. Recently, all the necessary steps were taken to implement the Mini-EUSO mission configuration, including the L1 trigger logic, in order to assess its performance.

Figure 4.12 shows the expected track (left) and light curve (right) of an Extreme Energy Cosmic Ray with energy $E = 10^{21}$ eV. Figure 4.13 right plot shows the trigger efficiency curve for Mini–EUSO adopting the L1 trigger logic here described.

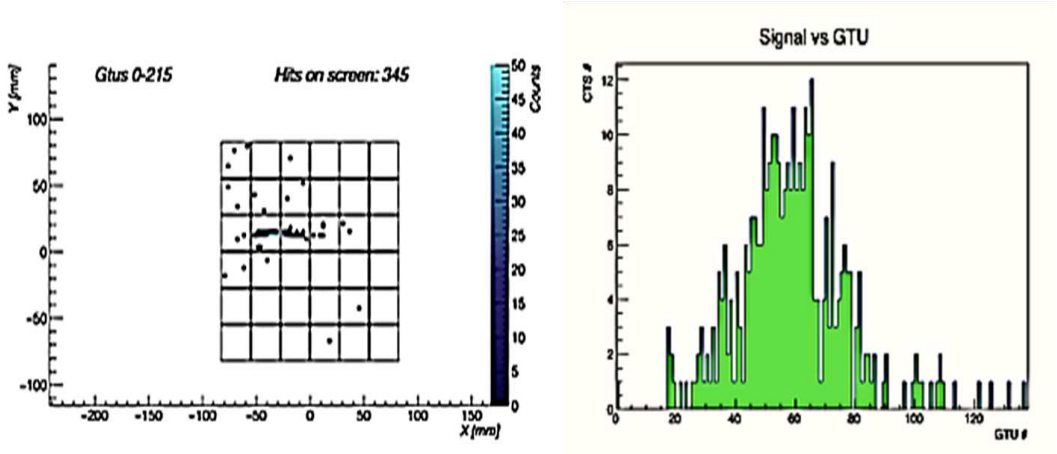

Fig. 4.12 Left: Photon counts observed in the Mini–EUSO focal surface for a simulation of a $E = 10^{21}$ eV event with an inclination of 80° to the nadir (background is not included in the simulation). Center: Light curve for the same event. The x-axis shows time in units of GTU (1 GTU = 2.5 μs)

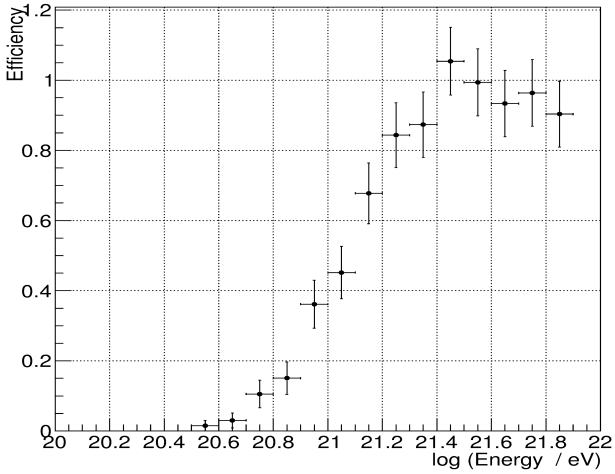

Fig. 4.13 Trigger efficiency curve as a function of the EAS energy.

Despite its energy threshold is too high for cosmic ray detection ($E_{thr} \sim 10^{21}$ eV), with its annual exposure of $\sim$ 15,000 km^2 year sr, Mini-EUSO will provide a significant contribution in estimating an absolute limit on the cosmic ray flux above such energies for a null detection.

As ESAF allows the simulation of phenomena or longer durations such as TLEs, meteors, cities, etc., a few examples of these classes of events were generated. These simulations confirmed the capability of the L1 trigger logic to avoid triggers on meteors and cities, while in case of TLEs it was verified that the L1 would trigger if the derivative of the light curve in the rising phase is so steep that the adaptation of the trigger thresholds at steps of 320 μs is too slow to follow the light increase. Even though the detection of TLEs and lightning is one of the main objectives of the L2 trigger logic, the L1 will allow recording with much higher time resolution the rising phase of the brightest and fastest signals.

4.5.3 L2 trigger tests

As described in section 4.4.2, the L2 trigger operates on integrated packets of 128 GTU generated by the L1 trigger. Triggering is performed on the timescale of $\sim$ 40 ms with a time resolution of 320 μs , designed to capture the range of transient luminous events (TLEs) in the Earth's atmosphere that will be visible to Mini-EUSO. TLEs are important to study as they are part of the UV background that will be

encountered by future instruments looking to study UHECRs from space. However, the high time and spatial resolution of Mini-EUSO means that it will also be possible to make unique observations of these atmospheric events, complementing those of other dedicated instruments scheduled to fly in Earth orbit during the same period (e.g. TUS [9], ASIM [38]).

In order to test the algorithm, the ESAF simulation software was used to generate a range of typical TLE events (namely blue jets, elves and sprites), as would be seen by the Mini-EUSO focal surface. Background was superimposed onto the simulated data packets, assuming a Poisson distribution of background events centered on 1 photon/pixel/GTU. Examples of the TLE events considered are shown in Figure 4.14. The L2 trigger was then run on this simulated data to test its performance. Two key parameters, the threshold level and the persistence, were varied to investigate their effect on the trigger efficiency. The threshold is simply the level at which the signal is triggered and the persistence is the time frame used to compare the instantaneous signal to background.

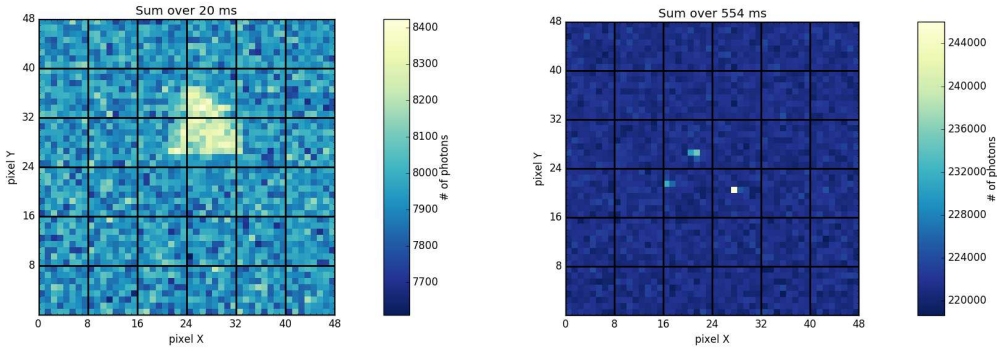

Fig. 4.14 Left: A typical integrated frame showing a diffuse elf event which brightens the whole PDM. Right: A similar integrated frame showing 3 localized blue jet events summed over 554 ms. The x and y axes represent the pixel grid of the Mini-EUSO focal surface and the colormap shows the number of photons counted by each pixel. All events are simulated with a Poisson statistic background distribution centered on 1 photon/pixel/GTU.

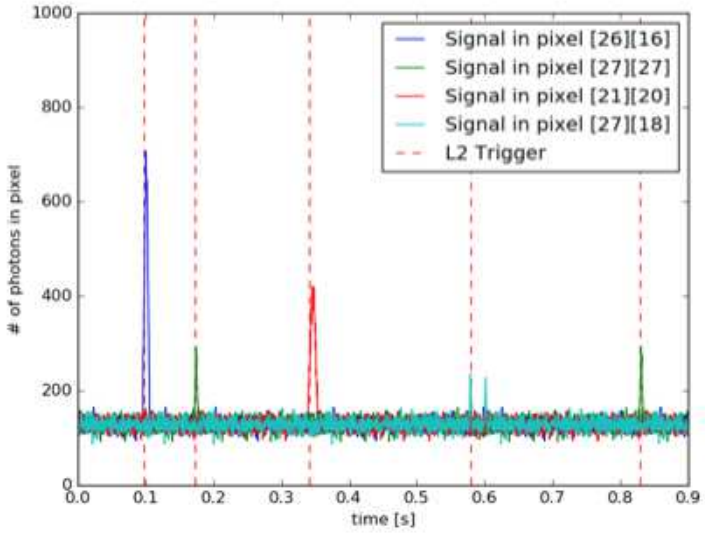

Fig. 4.15 The result of running the L2 trigger algorithm on the simulated data packet. The included events (from left to right) are a blue jet, a sprite, another blue jet, an elf and a final sprite. The events were spread over different areas of the focal surface. The red dashed line marks an L2 trigger. This result was achieved following the optimization of the trigger parameters and a trigger efficiency of 100% is obtained.

The testing of the trigger algorithm confirmed its ability to distinguish events of interest from typical background levels and also allowed approximate lower limits to be set on the magnitude of the TLEs that Mini-EUSO will be able to detect (for typical sprites and blue jets, an absolute magnitude of $\sim$3, and for elves an absolute magnitude of $\sim$1). Figure 4.15 shows the trigger response to five different simulated TLEs. The optimal trigger parameters were a threshold of 5σ above background level and a persistence of 8 frames of 320 µs. It should be noted that a longer persistence increases the sensitivity of the algorithm to the more diffuse elves, but at the expense of the detection of the more localized blue jets and sprites. The final implementation of the L2 trigger should allow for some compromise here and ideally have parameters which are adjustable in-flight.

4.6 Trigger implementation in the Engineering Model hardware

4.6.1 First tests

As already anticipated, the Mini EUSO trigger is coded in VHDL. The used FPGAs are Xilinx type and the software is Vivado (up to the 2016.2 release [3]). As a common procedure related to an FPGA code implementation, it moved through the code test-bench simulations, post-synthesis simulations and then the bit stream has to be exported on hardware for the place and route. After this phase, the trigger logic has been tested by means of data coming from software simulations and real data from atmospheric balloon campaigns. Moreover, a first and more realistic test for the trigger logic consisted in artificially stimulating the hardware chain from the EC ASIC Board to the ZYNQ Board. The resulting L1 events were consistent with the set pulse generator rate. Some details about this last test are hereafter reported.

4.6.2 Pulse generator as signal source: setup and outcomes

The hardware chain tested consists in the following: a wave function generator in pulse mode, connected whit a proper Kapton cable to the PMT connection point on the EC ASIC Board that is packaged with the Cross Board, ZYNQ Board and Power Board. Through the pin configuration in the Vivado Layout view, the L1 Event signal has been associated to one ZYNQ Board physical output pin. From that pin, an oscilloscope probe, detected the trigger results.

Agilent wave functions generator setup:

- Burst period 1 ms;
- Charge input: pulse width = 8 ns, pulse height = 100mV $\approx$ 2 photon electron charge equivalent (PE), for a 5*10^6 PMT gain.

[3] www.xilinx.com/products/design-tools/vivado.html

Tektronix Digital Osciloscope setup:

- Monitor input pulse and trigger output by probe;
- Measured trigger rate during 1min.

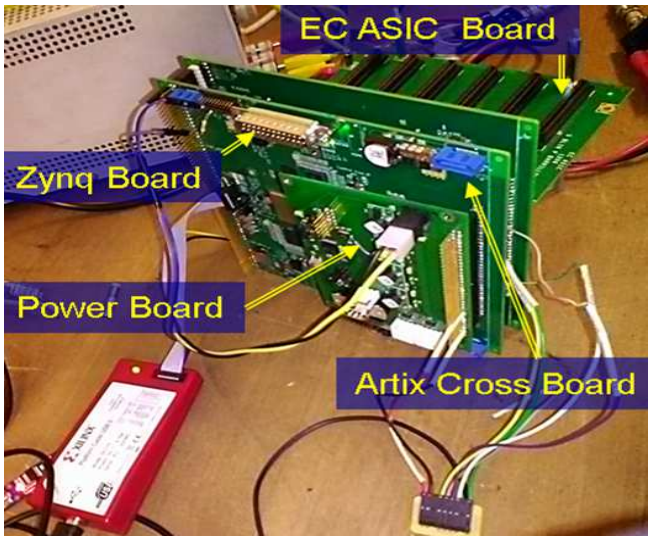

Fig. 4.16 Picture of the PDM module.

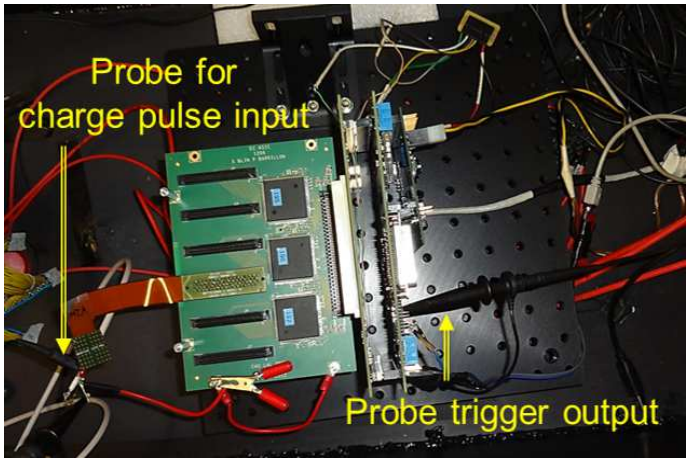

Fig. 4.17 Picture showing the input and output test probes.

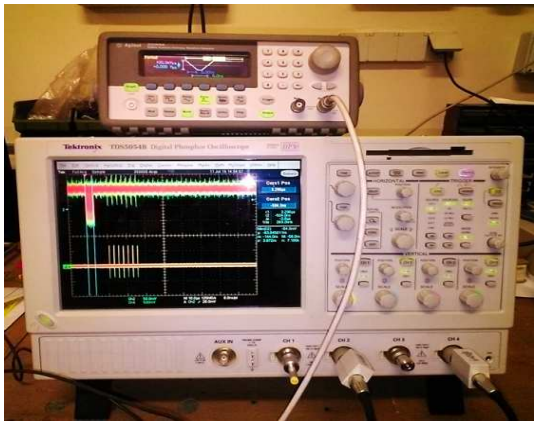

Fig. 4.18 In the upper part, the wave functions generator. The Digital Oscilloscope screen, reporting 8 consecutive L1 Events (lower part).

The charge input pulsed by Agilent function generator was a 100 mV amplitude signal with 8ns pulse width. This signal moved through an RC circuit with Rin = 500 Ohm, C =100 nF:
Qin = 100 mV *8 ns /500 Ohm = $1.6 * 10^{-12}$ (corresponding to roughly 2 PE charge for the PMT gain 5*10^6). After that configuration, we changed the pulse generator set to a burst mode: pulses were fired every 1s, testing the ASIC Channel number 49, at the pixel 17.

In the table below (Figure 4.19), the results for different number of pulses sent are reported, specifying the burst time duration. In this setup, the EC ASIC Board had the SPACIROC1 ASICs and in the Figure 4.19 specific case, the DAC threshold for the ASIC discriminator, was set to 150. Currently the migration form SPACIROC1 to SPACIROC3 ASIC is in progress and the later one will be the Mini EUSO official readout chip.

Measurement1, DAC150, fixed interval between pulses (100ns), burst rate 1Hz			
N° of pulses	N° of trigger/min	trigger eff. [%]	burst [μs]
40	61	102	4
38	60	100	3.8
36	60	100	3.6
34	42	70	3.4
32	37	62	3.2
30	37	62	3
20	10.3	17	2

Fig. 4.19 L1 trigger testing in hardware. The hardware chain was stimulated with a pulse burst with a burst rate of 1 Hz. Changing the number of pulses within the pulse changes the trigger efficiency as expected.

4.6.3 Trigger ancillary IP blocks

Another important test source is the Artificial Data Generator, a block coded and implemented in hardware, that allows to provide a trigger stimuli in the case of a completely assembled detector and even with a stand alone ZYNQ Board. Others custom-realized IP blocks are the Pixel Masking module and the Time Stamp Generator module; in a data-flow scheme those FPGA-integrated components are placed respectively before and after the trigger, as represented in Fig.4.20.
During the FPGA design phase with Vivado, has been adopted a Block Design level based on IPs and on the classical Master to Slave paradigm. This means that our coded components should be properly packaged thus to create a standard communication protocol between them. The chosen format is the AXI4_Stream one [40]. An AXI DMA standard IP, provided by Vivado, is needed to interconnect the PL part with the processor since the first uses the AXI Stream format while the second one, needs the AXI Memory Mapped type. The ZYNQ PS has the DRAM controller embedded. The IP block essential notes are hereafter reported. All these blocks, share with the L1 and L2 trigger blocks the same clock signal and the same Frame signal, a single bit signal that is high when there are sensitive GTU-data.

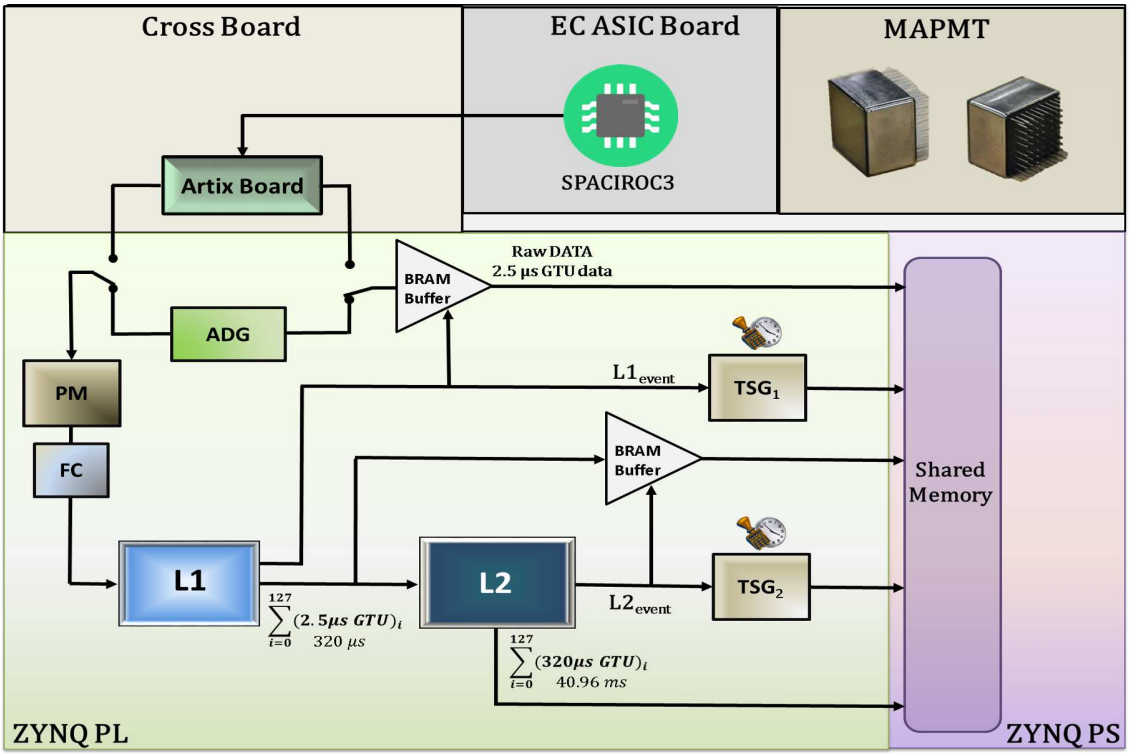

Fig. 4.20 Schematic representation of the trigger logic surroundings. Blocks names: ADG = Artificial Data Generator, PM = Pixel Masking, FC = Format Converter, L1 = First Level Trigger, L2 = Second Level Trigger, TSG = Time Stamp Generator.

The Artificial Data Generator In order to use the Artificial Data Generator IP (ADG IP) the architecture is equipped with a real data - artificial data switch. The Artificial Data Generator receives a clock, a start/stop command and a mode of operation instruction. With the operation mode, is it possible to set a different behavior, among the following ones, consistent in terms of format to the Cross Board output (timing constraints, input data, frame and clock signals correspondence).

ADG IP mode:

- Mode 00 is the coarse trigger test (interconnection test): all zero except one high pixel-value that is able to rise up the L1_Event, output signal. This mode is a kind of electric test or a component interconnections test.
- Mode 01 is the AXI DMA test (data transfer PL-PS): 128 GTUs are transferred between the FPGA PL and the CPU.

- Mode 10 is the L1 mode: one pixel-value increasing by one count every 128 GTUs, remains high for all the GTUs thus replicating the suitable condition to trigger only with the L1 algorithm.
- Mode 11 is the L2 mode: one pixel-value increasing every 128 GTUs, remains high only for the first GTU, replies the suitable condition to trigger only with the L2 algorithm.

For the coarse mode, the generator provides a 2304 pixels matrix, which are all zero except one, set with an high value (e.g. 50, considering that the minimum fixed values for the L1 threshold is 15 counts), thus assuring trigger events. This is useful mainly to have a direct connectivity test. Whereas for the more complex modes, it is possible to generate values at pixel level, with a steep or smooth increase according to the level of trigger (L1 or L2) which is under test.

The Pixel Masking A pixel masking module is a reliable way to exclude those pixels which behavior is not the expected one (e.g. too noisy and/or a too high number of induced fake triggers). The easiest way to do this is using a pre compiled text file, a sort of 2304 bits binary file, that declares with ones and zeros the pixels that respectively need to be masked (with "1", the associated pixel is fine whereas, with "0", the 8 bits associated to the pixels, have to be replaced with 8 zeros). This IP allows to reject the corrupted data before saving them, once the user knows the detector behaviour (which has to be checked in advance). Fake triggers due only to the fact that the hardware have broken pixels would mean a huge waste of storage space and computing power.

Time Stamp Generator We need to recognize the GTU in time relation to the trigger, especially when both L1 and L2 events occur. The output data storage format is: DatYYMMDDHHmmSSGGGGGG.ready (G stands for GTU and 6 characters are needed to represent the 399999 2.5 μs GTUs, in each second). We can summarize the time stamp procedure in four steps:
(I)Starting data boot, (II)Time counter, (III)Time-latched Trigger event and (IV)Count-to-string: output data format.
I) Whenever the Counter set is high, the preloaded data referred to Year, Month, Day, hour, minute and second are booted into a counter module. From that moment, the

frame signal (i.e. the GTU signal), which is generated from the Artixs (Cross Board), controls the time flow.
II) The counter module works in a cascade way: as a first level, there is a Frame counter (2.5 us GTU), that resets itself after 399999 counts; this reset implies the first count for the level number 2, i.e. the Second counter. This second level resets itself as soon it reaches 59 counts to increment a third counter, the Minute counter. The same procedure is used to increment the hour and the date counters.

III) The trigger signals is then latched with the time reference.
IV) The data output has to be rearranged as a string file like this: DatYYMMDDHHmmSSGGGGGG.ready. Since each character is represented by 8 bits, the string YYMMDDHHmmSSGGGGGG , needs a 144 bits vector. Currently the Time Stamp Generator is a Vivado custom IP and it has been implemented in the ZYNQ FPGA, and tested with the SDK (Software Development Kit, a Vivado tool that provides the instruments for CPU programming), PS (Processing System) interaction. The Figure 4.21 gives a block representation of the IP.

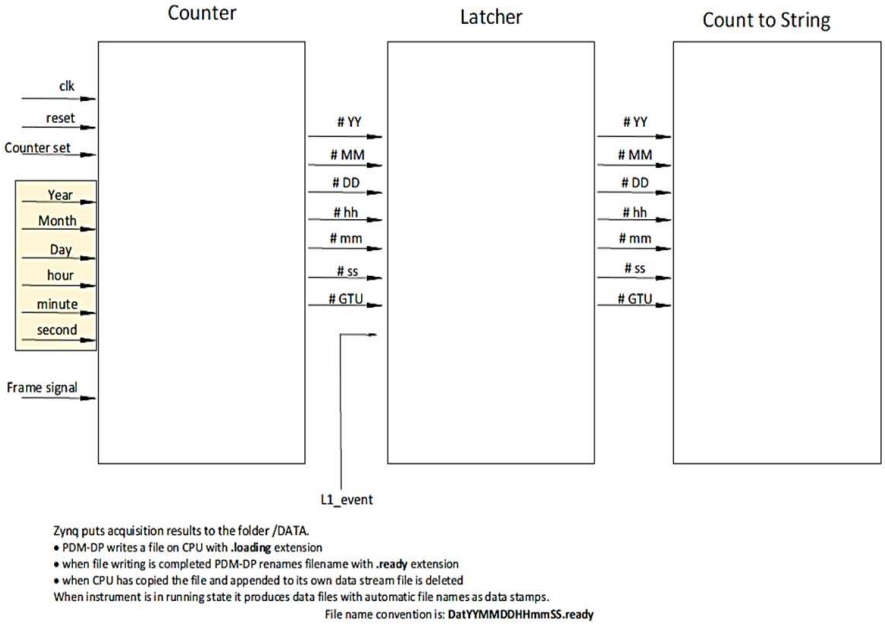

Fig. 4.21 Schematic representation of the Time Stamp Generator logic.

4.7 Trigger logic improvements

4.7.1 Dynamic threshold setting

Concerning the trigger logic principle, the L2 has been improved underway. Let's have a look at the Figure 4.22. In such scheme we decide to have a trigger if S2_pixel is > 4 x S2_pixelvalue, in other words if the signal in a pixel is 4 times higher than the average background calculated on the previous 128 L2GTUs (1 L2GTUs=320 μs) and re-normalized to 8 L2GTUs blocks.

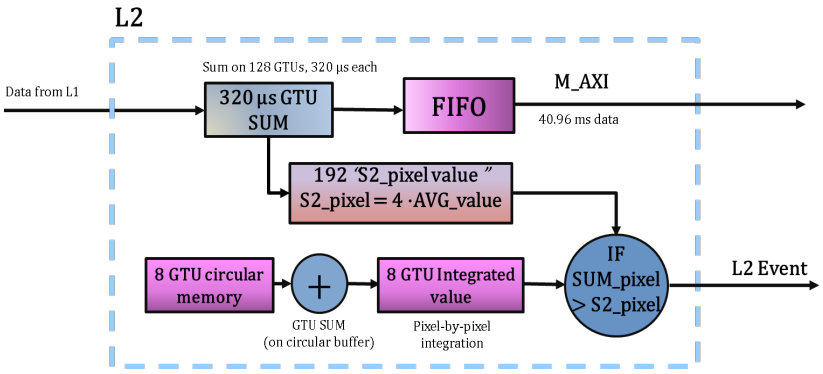

Fig. 4.22 L2 logic block diagram.

In order to increase the degrees of freedom and thus the system flexibility, we define as external parameters 2 values called N and P. N indicates how many times the signal should exceed the background to trigger an event. P indicates how many L2GTUs we integrate to calculate the average background and its sum. As an example in L1 we calculate the excess on blocks of 8 GTUs, therefore P = 8. This number is not needed to be the same also for L2 since signals might last longer in one pixel in case they do not move at speed of light; moreover, lightnings are events with a large variation in their characteristics. Therefore, the new formula to calculate the background would be: S2_pixel = N $\times$ (SUM_PIXEL / 128) $\times$ P
where SUM_PIXEL = sum of counts recorded by the pixel in 128 L2GTUs, N and P can vary between 1 and 16 (even though N = 1 would be meaningless). By default, N=4 and P=8. The concept is represented in the block diagram of Figure 4.23. N and P values should be defined (write) and then read from a FIFO at the beginning of the acquisition. The choice of N as number of times the signal is higher than the background, instead of using the standard deviation, is due to the fact that with

very high numbers the Poisson number of sigma is not a good parameter, since the interfering fluctuations become Gaussian. This new version of L2 has been coded in VHDL and deeply tested as well the previous one.

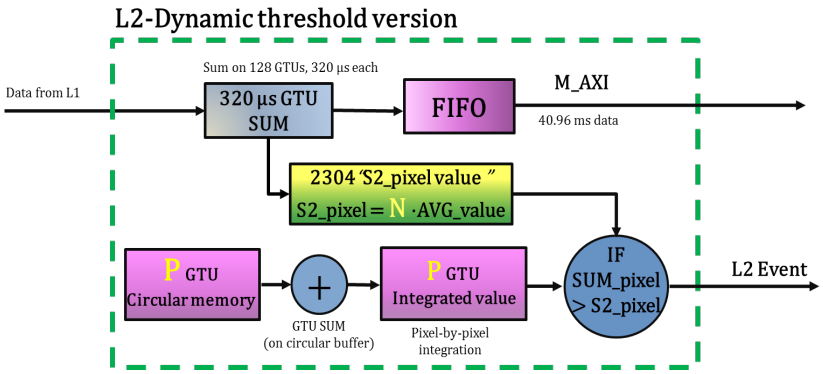

Fig. 4.23 Block diagram representing the L2 trigger logic modified with a threshold adjustment capability

4.7.2 High Level Synthesis trigger implementation

As previously mentioned, the FPGA-side of the Mini-EUSO project is being developed using the Xilinx tools [33]. The project started with the ISE Webpack [39], and then migrated to the Vivado Webpack [41] once we chose the Xilinx 7th serie devices (Artix7 and Kintex7). For the whole project the FPGA firmware has been coded in VHDL thus obtaining the Register Transfer Level (RTL) description in a classical way. In digital circuit design, RTL is a design abstraction which models a synchronous digital circuit in terms of the flow of digital signals (data) between hardware registers, and the logical operations performed on those signals. RTL abstraction is described in Hardware Description Languages (HDLs) like Verilog and VHDL to create high-level representations of a circuit, from which lower-level representations and ultimately actual wiring can be derived. Design at the RTL level is typical practice in modern digital design [42]. The Xilinx platform is rich of very powerful tools, giving the user the chance of conveniently adapt the strategy to its task, in order to maximize the efficiency. A smart example of this is represented by the Vivado High-Level Synthesis compiler (HLS). Vivado HLS enables C, C++ and SystemC programs to be directly targeted into Xilinx devices without the need to manually create RTL. To an FPGA user this should sound like a great step forward that is even greater if the logic that has to be integrated in hardware came from a

scientific software like the Mini-EUSO trigger, born as a C-like code. The Vivado HLS concept flow is represented in Figure 4.24.

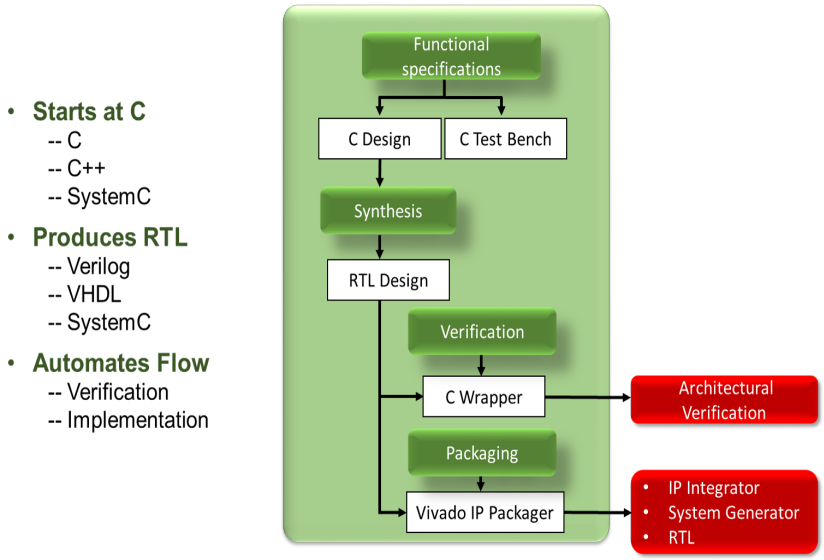

Fig. 4.24 Vivado High Level Synthesis flow, represented as a block diagram.

Differently from the classical HDL approach, currently the HLS doesn't allow the user to deeply control every step from the code to the hardware implementation. The design flow asks the user to rely on some black-box steps. For this reason, the HLS reliability has to be proved in our specific scenario. The Mini-EUSO trigger integration group decided to test and use the HLS approach for a spare version of the L2 HDL-based. Since the L2 algorithm is more complex than the L1 one, a direct passage from C++ to net-list greatly simplify the engineer work. With this HLS-L2 version tested and implemented in the ZYNQ Board, in parallel with the HDL one, the Mini-EUSO Engineering Model is useful for an additional test. Moreover, a redundant trigger gives a more severe crosscheck to the idea behind the code.

4.7.3 Further upgrading idea: multiple events recording

The trigger integration is a matter of memory capability and data throughput that has to suit the higher level system gear. its integration moves through technical boundary that are primary hardware and than logic dependent. On this premise we can think about improving possibilities that might be soon adopted. Hereafter is presented a technical solution for a multiple event recording capability.

As first implementation, the output data from the L1 Trigger logic (Figure 4.25) is the continuous sum over 128 GTUs (2.5 μs GTU) and the L1 Event signal. Following what has been previously explained, whenever the First L1 Event occurs, the procedure just wait until the end of the 5.24 s period. In this situation it should be clear that it might happens to waste more than 5 seconds of data for just 320 μs of useful informations. Moreover, the entire data section could be related to a fake trigger event. When we have an L1 trigger, we would like to have an information on which L1 packet triggered inside the 5.24s (every 5.24s the data are transferred to disk); the situation would be even more complicated in case L1 and L2 trigger in the same time range. A case could be a TLE (Transient Luminous Event) which triggers firstly L1 and then L2 since the two events are correlated. It is thus useful to understand from the time stamp or a packet number, that the two triggered events are indeed the same one triggered at different time scale. An improvement in this sense, is obtained implementing the possibility to store more than one event, for example 4 events, before stopping the write data process. Another free parameter is the number of GTUs that we want to keep in memory. Although t_1 has to be fixed to 2.5 μs (raw data from the ASICs), t_2 , the time resolution of the L2 Trigger scales, can vary following t_2 = t_1 * N*GTU . Let assume to have 300 or 500 GTUs blocks, and multiple event storage capability; in these conditions the dead time is reduced and it becomes possible saving data related to lightnings, with the maximum time resolution (2.5 μs). Suppose having a four slots memory, where each slot has a referred address as presented in Figure 4.25

.

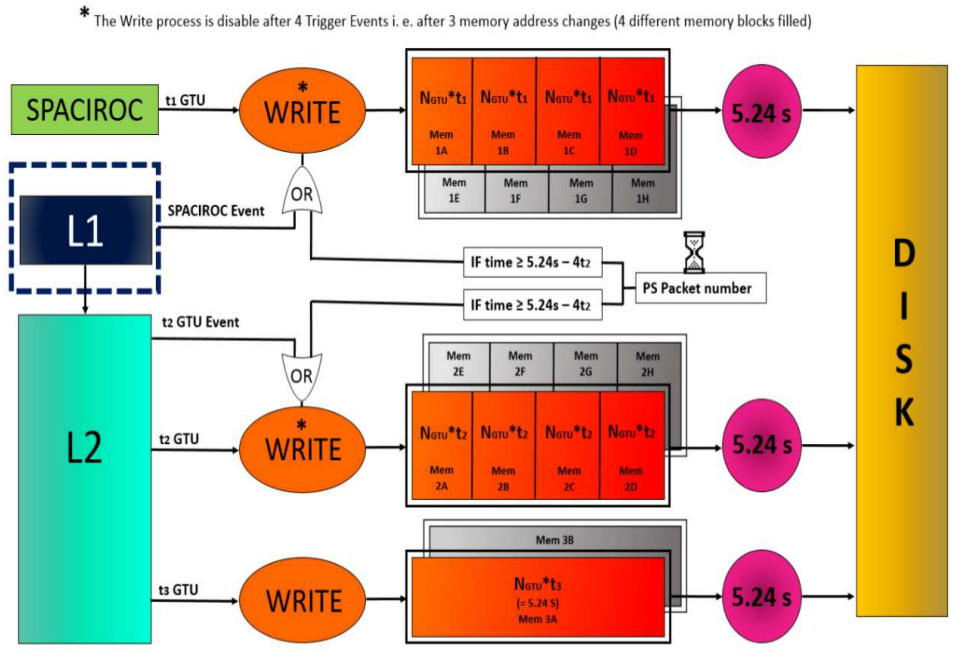

Fig. 4.25 Trigger logic improvements for L1 and L2 multiple events data storage.

Until the first L1 Event is not triggered, the pointed slot of the memory will be the first one (1A in Figure 4.25). In this memory position the sum over N GTU are continuously overwritten. When the first L1 Event occurs, the PS has to change the pointed memory location to the second one. This happens only after a certain delay which allows to put the event in the middle of the N GTU block. The events that occur during one event registration are rejected. After the fourth event has been recorded, the writing process is stopped until the end of the 5.24s period (the address changing request will be rejected).

The same architecture is repeated for the L2 level. This latter level is moreover responsible to the no-event case behavior. We want to extend the records of useful data therefore we would like to save something even in the case of no-event during the entire 5.24 s period. More in detail, we want to always have 4 recorded events and this can be done inducing fake events or simply forced data acquisition up to fill the four memory slots related to L1 as well the L2 level. To make it easier, we can use the data as they were defined in Figure 4.22. From the second level trigger we have sums every 40.96 ms. A packet counter is needed to alert the system about the situation in which the time equivalent to the packet number is 5.24 s- 163.84 ms (40.96 * 4); at this point a memory-address-change-request is sent. If the actual

memory slot is not the fourth one, the request is accepted and the new memory position is fulfilled. This happens for both L1 and L2. In this case, instead of having no data at the end of the 5.24 s period, we can save informations related to the sums, in order to enrich the data statistics. It is true that, since we force also the L1 memory slots, we will waste 163.84 ms but this is only the 3% of data over the 5.24 s period. As previously mentioned, once arrived the memory slot number four, the address changing request will be rejected. At the end of the 5.24 s period, the system moves to a second memory blocks level 1E, 1F, 1G, 1H, for L1 and 2E, 2F, 2G, 2H for L2, (represented in gray in Figure 4.16). This two memory levels swipe allows a death time free data transfer.

With the new configuration, we add new informations to the data acquisition. The Mini EUSO FoV is 19 ° i.e. a 275 Km diameter at ground (tan 19*400 Km*2*). The ISS speed is 7.6 Km/s thus it needs 36 s to cover the 275 Km length. The Mini EUSO trigger logic works with 5.24 s steps, plus some dead time to transfer data i.e. approximately 6s steps. For each 6s step the system catch 4 L1 events (plus 4 L2 events). In 6s, the ISS moves by 45 Km at ground. This means that the PMT FoV is 45 Km and 45 * 6 = 270 Km thus, if we are in the main axes position, 6 PMTs will see the same situation. This is of course different for the perimeter PMTs but we can say that the systems has a sort of redundant data acquisition, which allows to extrapolate informations about PMTs gains and efficiencies.

The new requirement would require the following effort:

- L1 trigger: 4 events/5.24s * 512 GTUs /event * 2304 pixel * 1byte/pixel = 900 kB/sec
- L2 trigger: 4 events/5.24s * 128 GTUs/event * 2304 pixel 2bytes/pixel = 450 kB/sec
- movie: 1 event/5.24s * 128 GTUs/event * 2304 pix * 2bytes/pixel = 110 kb/sec
- TOTAL: about 1.5 MB/s

Assuming to have a 500 GB disks:
500 GB / 1.5 MB/s = 333333 s = 92.6 h. Assuming now 8 h/day = 11.6 days/month. Taking into account also a lower duty cycle because we share the window with another experiment the data transfer will be in the order of 1 TB/month. This solution would allow to have more or less 1 event/1.3 s for both L1 and L2. Therefore, in

case there are TLEs we could record 1.28 ms/event during the rising phase of the TLE with a 2.5 μs resolution and then the entire duration of the event with a 320 μs resolution.

4.8 Project status summary

In the last 12 months the JEM-EUSO Collaboration members concentrated the efforts on successful tests for Mini-EUSO. These activities took place in world wide locations, shared among the research groups following the subsystem:

- Japan for the lenses
- France for the front-end electronics
- Russia for the Data Acquisition system and the mechanical interfaces to the ISS
- Italy for the CPU, the mechanics, the trigger and the LVPS, the ancillary cameras and the SiPM
- Poland for the HV system
- Sweden for the software
- Mexico for the Housekeeping
- and several other countries for the MAPMTs acquiring.

In summer and autumn 2017 the final integration and calibration of the Engineering Model (EM) occurred at the Physics Department of the University of Rome Tor Vergata 4.26. After this step, thermo-vacuum test, vibration test, electromagnetic compatibility check and out gassing tests are going to be performed in spring 2018, at the Kayser Aerospace Company in Italy. The final step before the launch will be the acceptance tests in Russia. In the meanwhile the Flight Model (FM) has to be prepared waiting for the succeeding milestone of the Russian acceptance test. This step is needed to optimize the time schedule for the FM integration in Rome and the delivery to Russian Space Agencies Roscosmos and Energia for the final phases prior the space mission start. ASI and Roscosmos are setting up the final agreements

to choose one of the next launch increment mission to ISS with the Progress carrier and select and train the astronauts who will operate Mini-EUSO on board.

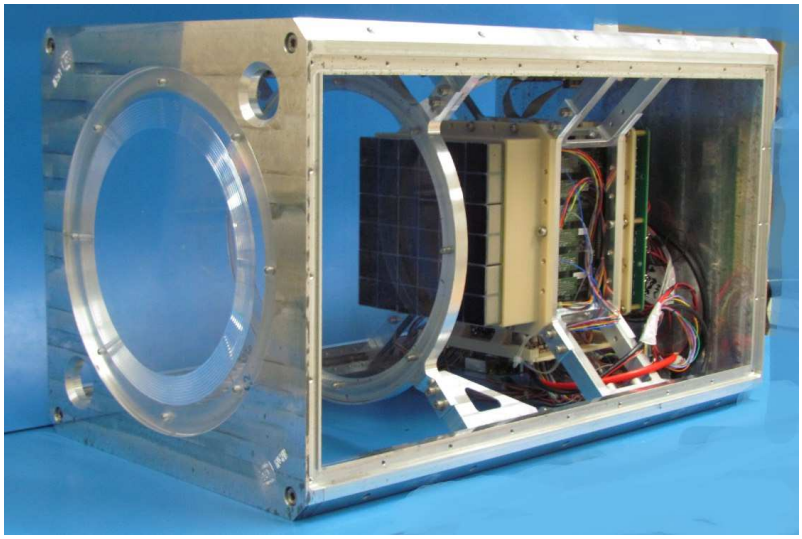

Fig. 4.26 Mini-EUSO Engineering model photo, during its integration phase, June 2017.

4.9 Summary

The Mini-EUSO trigger algorithm has been integrated in the Zynq Board FPGA. Before this integration, the trigger algorithm was tested successfully using simulated data and data generated as part of the EUSO@TurLab project. Once integrated in the hardware, the trigger was then tested using a pulse generator and the complete data acquisition chain. The artificial data generator implemented in the Zynq board will allow for stand alone testing of the trigger logic. At this point, the logic allows a reliable FPGA pixels masking, just booting a binary input vector as pixels correspondence ("1" = fine pixel, "0" pixel to be masked). The trigger events are time-stamp latched to save data in time order, moreover allowing events discrimination in situations with multiple levels triggers firing. Following the trigger implementation the PDM-DP system has been integrated with the Min-EUSO Engineering Module for end-to-end testing of the data acquisition system.

References

[1] The Pierre Auger Collaboration, Nucl. Instrum. Meth. A 798 (2015) 172.

[2] P. Ghia, for the Pierre Auger Collaboration, in Proc. 34th ICRC 2015, The Hague, The Netherlands, PoS (ICRC2015) 034.

[3] M. Fukushima et al., Prog. Theor. Phys. Suppl. 151 (2003) 206.

[4] C. Jui, for the Telescope Array Collaboration, in Proc. 34th ICRC 2015, The Hague, The Netherlands, PoS (ICRC2015) 035.

[5] Adams, J. H. A*et al.*, An evaluation of the exposure in nadir observation of the JEM-EUSO mission. Astroparticle Physics, 44–76 (2013).

[6] G.K. Garipov *et al.*, UV radiation from the atmosphere: Results of the MSU "Tatiana" satellite measurements. Astroparticle Physics, 24(4-5): 400–408, Dec. 2005.

[7] G. Catalano, *et al.* The atmospheric nightglow in the 300 – 400 nm wavelength Results by the balloon-borne experiment BABY. Nuclear Instruments and Methods in Physics Research Section A, 480(2-3):547–554, Mar. 2002.

[8] G.K. Garipov A*et al.*, Program of transient UV event research at Tatiana-2 satellite. Journal of Geophysical Research: Space Physics, 115(A5), 2010. ISSN 2156-2202. A00E24.

[9] M. I. Panasyuk A*et al.*, Transient luminous event phenomena and energetic particles impacting the upper atmosphere: Russian space experiment programs. Journal of Geophysical Research: Space Physics (1978–2012), 115(A6) (2010).

[10] J. H. Adams, A*et al.*, The JEM-EUSO instrument. Experimental Astronomy, 40(1):19–44 (2015).

[11] V. Olinto, E. Parizot, M. Bertaina, and G. Medina-Tanco. JEM-EUSO Science. In Proceedings of the 34th International Cosmic Ray Conference (2015).

[12] Y. Kawasaki *et al.*, for the JEM-EUSO Collaboration. Ground-based tests of JEM-EUSO components at the Telescope Array site, "EUSO-TA". Experimental Astronomy, 40(1):301–314 (2015).

[13] V. Scotti and G. Osteria. EUSO-Balloon: The first flight. Nucl. Instrum. Meth., A824:655–657, 2016.

[14] L. Wiencke, A*et al.*, The EUSO-SPB Mission. In Bulletin of the American Physical Society, 2016.

[15] J. H. Adams, S. Ahmad, J. N. Albert, et al. Space experiment TUS on board the Lomonosov satellite as pathfinder of JEM-EUSO. Experimental Astronomy, 40(1): 315–326, Nov. 2015.

[16] M. Ricci, *et al.* Mini-EUSO: a precursor mission to observe Atmosphere and Earth UV emission from the International Space Station Proc. 35th ICRC, 2017, these proceedings

[17] V. Verzi, Rapporteur Report of 34th ICRC 2015, The Hague, The Netherlands, PoS (ICRC2015) 015.

[18] V. Verzi et al. Measurement of energy spectrum of ultra-high energy cosmic rays, Progress of Theoretical and Experimental Physics, Volume 2017, Issue 12, 12A103

[19] K. Greisen, Phys. Rev. Lett. 16 (1966) 183

[20] G. T. Zatsepin and V. A. Kuz'min, Sov. Phys. JETP Lett. 4 (1966) 114.

[21] V. Berezinsky et al., Phys. Rev. D 74 (2006) 043005.

[22] R. Aloisio et al., J. Phys. Conf. Ser. 337 (2012) 012042.

[23] M. Takeda et al., Astropart. Phys 19 (2003) 447; K. Shinozaki et al., Nucl. Phys. B (Proc. Suppl.) 151 (2006) 3.

[24] R.U. Abbasi et al., Phys. Rev. Lett., 100 (2008) 101101.

[25] D. Schenider: Sci. Am. 88(4), 1 (2000))

[26] Kuo, C.L.A*et al.*,: Radiative emission and energy deposition in transient luminous events. J. Phys. D: Appl. Phys. 41, 234014 (14pp) (2008)

[27] Wescott A*et al.*, Preliminary results from the Sprites94 aircraft campaign: 2. Blue jets. Geophys. Res. Lett. 22(10), 1205–1208 (1995)

[28] Sentman, D.D.: Overview of transient luminous events, University of Berkeley, California, 2005

[29] http://www.hamamatsu.com

[30] S.Turrizziani *et al.* - JEM-EUSO Collaboration, "Ancillary detectors for the Mini-EUSO telescope: control software development and expected science" Proc. 35th ICRC, 2017, these proceedings

[31] H. Miyamoto *et al.*, - JEM-EUSO Collaboration, "Performance of the SPACIROC front-end ASIC for JEM-EUSO" Proc. 33rd ICRC, 2013, these proceedings

[32] S. Ahmad *et al.*, - JEM-EUSO Collaboration, "SPACIROC: A Front-End Readout ASIC for Spatial Cosmic Ray Observatory" Proc. 32nd ICRC, 2011, these proceedings

[33] https://www.xilinx.com/

[34] M. Bertaina *et al.*, - JEM-EUSO Collaboration, "The Mini-EUSO multilevel trigger algorithm and its performance" Proc. 35th ICRC, 2017, these proceedings

[35] M. Bertaina, *et al.*, EUSO@TurLab: An experimental replica of ISS orbits, EPJ Web of Conferences 89 (2015) 03003.

[36] C. Berat *et al.*, Astrop. Physics, vol. 33/4 pag. 221-247, 2010.

[37] M. Bertaina, *et al.*, Performance and air-shower, 315 reconstruction techniques for the JEM-EUSO mission, Advances in Space Research.

[38] T. Neubert, *et al.*, ASIM—an Instrument Suite for the International Space Station, COUPLING OF THUNDERSTORMS AND LIGHTNING DISCHARGES TO NEAR-EARTH SPACE, Corte (France) 1118 (1) (2009) 8–12.

[39] https://www.xilinx.com/products/design-tools/ise-design-suite/ise-webpack.html

[40] https://www.xilinx.com/products/intellectual-property/axi4-stream_interconnect.html

[41] https://www.xilinx.com/products/design-tools/vivado/vivado-webpack.html

[42] F. Vahid, "Digital Design with RTL Design, Verilog and VHDL (2nd ed.)", John Wiley and Sons. p. 247., (2010).

Chapter 5

Conclusions

This thesis document reports the author's Doctoral research activity in the period November 2014-January 2018. The activity concerns the development of custom microelectronics for innovative particle physics detectors. The author worked on the design of front-end and readout electronics, mainly devoted to the projects run by the medical physics group of the Turin University. This group has a decennial history in the field of particle therapy, developing detector and the related analysis software that both have been integrated into clinical centers, as successful technology transfers. During his collaboration in this group, the author took part in two projects: TERA09 and MoVeIT. The TERA09 project consisted in the design and characterization of a front-end and readout electronics ASIC, to be coupled with ionization monitor chambers working with high-intensity pulsed-particle beams. These new clinical detectors have to deal with pulses 5-20 μs long, with a 0.1-1 kHz frequency, 10^8 proton per pulse and an instantaneous intensity of 10^{12}-10^{13} protons per second, corresponding to an average current during the pulse, in the range 1 nA-25 μA.

The TERA09 chip is a 64-channels current-to-frequency converter managing 100 pA -750 μA input currents with a linearity deviation in the conversion in the order of few percents ($\sim$3%). The TERA09 project started with the feasibility study based on testing the previous version of the chip, TERA08 (current limited at 4 μA), with a custom upper board that parallel interconnected the ASIC channels. From this point, the group developed the idea behind the architecture of the TERA09 chip that has been patented and then designed with a 350 nm VLSI technology node (the author is one of the inventors and he took part in the design). The TERA09 ASIC

has been fully characterized both in its technical features (dynamic range, conversion linearity and gain homogeneity among the channels) and with a radiation damage test. From this analysis, it resulted that TERA09 satisfies the technical requirements of compatibility with the current and standard version of IC chambers and it can be integrated into the new-era of monitor chambers. From the Single Event Upset test, it has been possible to predict the TERA09 expected error rate in a typical clinical treatment room. This number, $\sim$70 SEU/year, is absolutely under control, considering that each monitor chamber has at least one redundant detector. For instance, the probability that both the main and the redundant detector incur in SEU simultaneously is dramatically lower than the estimated number mentioned before. The ASIC characterization has been performed using a custom test board with a chip socket. A second PCB has been realized, carrying two ASICs for the integration of the front-end and readout electronics into the gas detector. This second board has an improved management of the critical signals (e.g. the 250 MHz differential clock), based on the ground metal layers packaging and a careful routing of the fast digital signals. From spring 2018, TERA09 is under test at GSI, Darmstadt. From these tests it resulted that a dedicated ring biased at the preamplifier reference voltage and shielding the input channels, would be needed to improve the channel-by-channel counting homogeneity.

The MoVeIT project is an Italian national collaboration working for a cohesive upgrade in Modeling and Verification techniques for Ion beam Treatment planning, in particle therapy. The Turing medical physics group is in charge of developing a detector prototype based on thin Low Gain Avalanche silicon detectors, for single ion discrimination in particle beams. Both the sensor and the front-end microelectronics designs have been developed in Turin. Even if the collaboration milestone is related to radio biology beams with fluxes up to $10^8\ cm^{-2}\ s^{-1}$ particles, the same community is interested in applying the same principle of ion level discrimination and particle counting to particle therapy beams. Although moving from ionization chamber to such a precise silicon device would lead to a revolution in this field, the technical issues make this goal very challenging. The system realized coupling this Ultra Fast Silicon Sensors and the front-end electronics have to deal with a wide charge range (3 - 150 fC) and a $\sim 10^9$ Hz particle fluence. Concerning the signal rate, the sensor prototype is a 3x3 cm^2 silicon area segmented in strips and this segmentation lowers the single channel counting frequency constraint. The front-end channel has to reach a high counting efficiency starting from a 100 MHz rate upward (radio biology limit).

The author worked on the design of a front-end chip prototype, named ABACUS equipped with 24 identical channels capable to deal with the entire charge range up to 250 MHz, keeping an efficiency close to 100%, form schematic level simulations. Simulating the post-layout design, the counting efficiency value is degraded to 85%. Nevertheless, the real counting efficiency on silicon could be higher, due to the fact that some layout techniques adopted to mitigate the propagation of electronics noise from the digital to the analog domain could not be simulated, since the related model was not available from the Process Design Kit. Sixty ABACUS prototypes have been delivered on the 21^{st} of June 2018. A dedicated test board where to glue the strip sensors and the ASIC before coupling them through wire bonding has been designed. The PCB production is ongoing and a new production of silicon sensors with different design approach is planned for the next autumn.

Apart from the activities related to the medical applications, the author collaborated for one year period (six months part-time and six months full-time), in the JEM-EUSO international collaboration for the Mini-EUSO development. Mini-EUSO is a UV-telescope that is going to be installed inside the Zvezda module of the International Space Station (installation foreseen in 2019), with the aim to map the Earth background in the ultraviolet spectrum; this approach is needed for the incoming JEM-EUSO experiments looking for Ultra High Energy Cosmic Rays (energies up to $10^{21}eV$) from space, using the Earth atmosphere as a giant gas detector. Furthermore, this detector is suitable for the detection of various kind of events like meteors, fluorescence phenomena, sprites, air-glows..
The telescope integrates multi-anode photomultiplier tubes and a 2034 pixels array. In order to discriminate among relevant events through a very large amount of raw data, a multi-level triggering system is needed. The astrophysics Turin group is in charge of the trigger algorithm logic and implementation in hardware. Working on this topic, the author collaborated on the first level trigger (L1 trigger) coding in VHDL and hardware implementation in FPGA. The development of trigger ancillary blocks like a time stamp generator, an artificial data generator and the pixel masking block have been author's duties which have been coded, integrated into FPGA and tested with the L1 trigger logic, resulting precisely working. A preliminary and successful test for the L1 trigger test has been performed using a pulse generator to stimulate a photomultiplier tube interconnected with the front-end and readout system. The main FPGA, named ZYNQ board, manages the data flow and integrates the trigger logic. Further tests on the Mini-EUSO entire device have been performed

in the Turin TurLab facility, with the telescope hanged up from a rotating tank that contained different light stimuli; this study was intended to emulate the detector space orbit and its capability in discriminating the events in various conditions, avoiding to trigger fake events (e.g. the artificial light pollution from a city).
Although the timescale of a space experiment is generally quite broad, the author experienced breakthroughs in the Mini-EUSO L1 trigger implementation in hardware and testing. The algorithm has been extensively verified through software simulations before to be hardware-integrated and then gradually tested with: artificial stimuli generated inside the FPGA, voltage pulses through the front-end and readout system and with the entire Mini-EUSO telescope in TurLab.

Appendix A

High speed signal transmission formats

A.1 Current Mode Logic (CML)

CML strength points:

- commonly used in high speed data transmission (e.g. laser driver, DVI and HDMI data transmission modules)
- simple output stage (Figure A.1)
- compatible with CMOS power domains (e.g. 1.2 V)
- voltage swing can be easily controlled by current source (as well the power)

CML drawbacks:

- asymmetric drive strength for rise and falling edges
- not fully-differential transmission (the mirror current is carried by the shield and not by the second conductor, Figure A.2)

The not fully-differential transmission limits the CML utilize to short connection on PCB (not suitable for cable differential signaling).

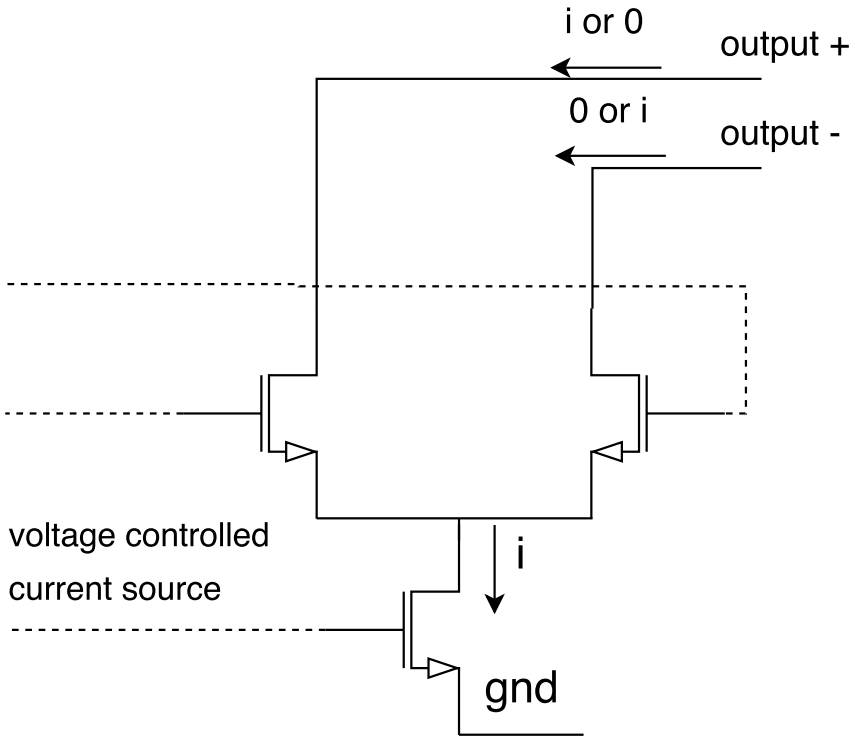

Fig. A.1 Typical structure for a CML output stage

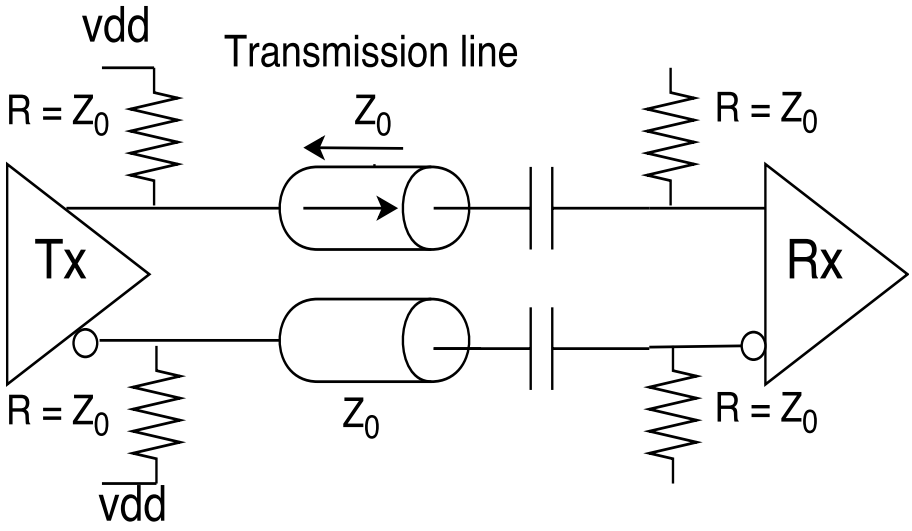

Fig. A.2 CML driver for AC-coupling.

A.2 Low Voltage Differential Signaling (LVDS)

LVDS strength points:

- is a popular format adopted for differential signal transmission, especially for long cable connections
- symmetric drive strength for rise and falling edges
- fully differential transmission (the mirror current is carried by the second conductor, the shield does not carries transient currents, Figure A.4
- floating termination
- excellent as standard digital I/O for sensitive analog ASICs

LVDS format drawbacks:

- more complex output stage (Figure A.3)
- not compatible with low power CMOS supply voltages (2.5 V required)
- voltage swing on 100 Ohm is equal to $2 \cdot 300$ mV, with a 1.2 V commod mode voltage

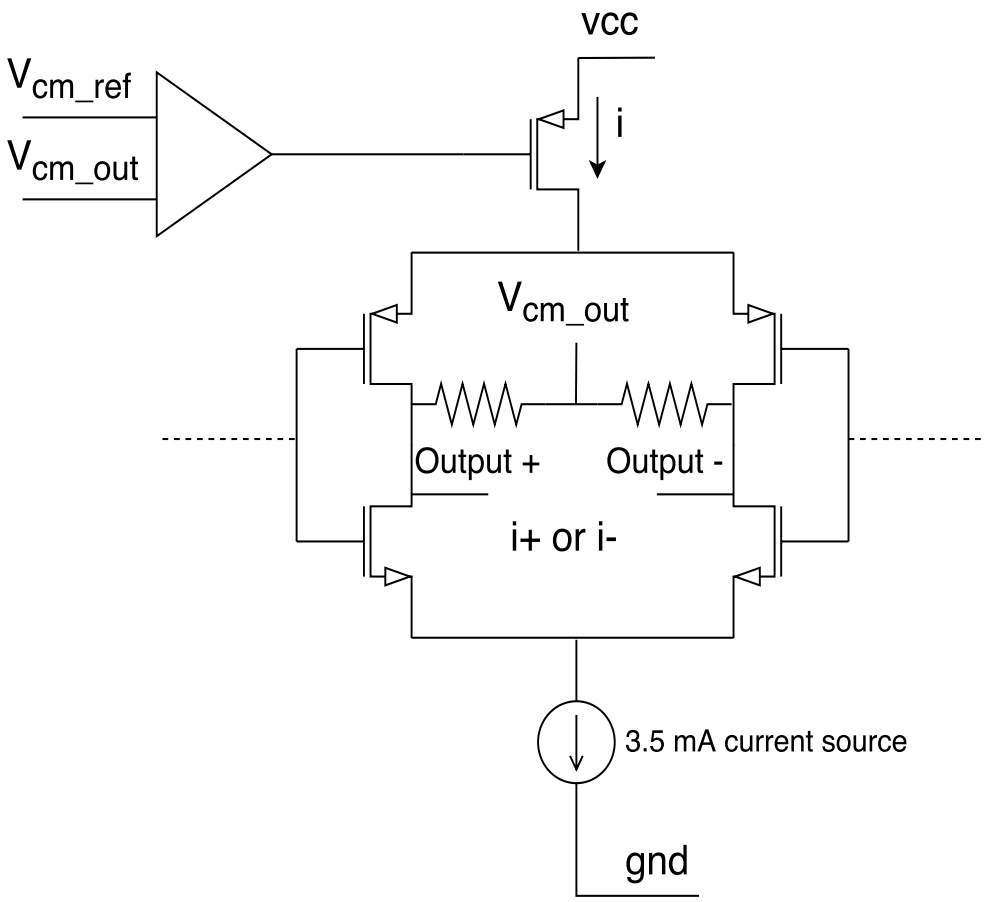

Fig. A.3 Typical structure for a LVDS output stage

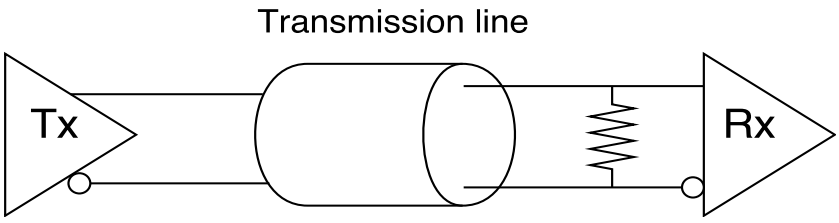

Fig. A.4 LVDS driver for DC-coupling.

A.3 Scalable Low Voltage Signaling (sLVS)

sLVS strength points:

- popular in electric high speed signaling for images and portable devices

- compatible with low voltage CMOS power domains (common mode = 200 mV)
- voltage swing can be controlled with current sources (0.5 mA - 2 mA)
- symmetric drive strength for rise and falling edges
- fully differential transmission (the mirror current is carried by the second conductor, the shield does not carries transient currents)

The sLVS has a 100 mV - 400 mV differential voltage over the 100 Ohm termination.