ELECTRONIC SYSTEMS FOR RADIATION DETECTION IN SPACE AND HIGH ENERGY PHYSICS APPLICATIONS

PIERPAOLO VALERIO

SAPIENZA
UNIVERSITÀ DI ROMA

Ph.D. Thesis

SUPERVISORS
Prof. Augusto Nascetti
Dr. Rafael Ballabriga

October 2013

Pierpaolo Valerio: *Electronic systems for radiation detection in space and high energy physics applications,* Ph.D. Thesis, © October 2013

This was a triumph
I'm making a note here:
HUGE SUCCESS.
— GLaDOS from *Portal*

Ohana means family.
Family means nobody gets left behind, or forgotten.
— Lilo from *Lilo & Stitch*

ABSTRACT (ENGLISH)

This Ph.D. thesis focuses on the analysis and development of novel solution for electronics system for radiation detector, especially suited for space and high energy physics applications. The many blocks of a readout system were studied to develop complete systems, investigating where the performances can be improved over state of the art technologies.
Two different architectures, suitable for different applications, were studied: Fractional Packet Counting, for High Dynamic Range (HDR) integrating imagers and CLICpix, an example of high-accuracy hybrid photon counting detector. The main specifications of the two systems were anayzed and solutions were proposed and implemented to meet them.
A CLICpix prototype has been designed, fabricated and tested (characterization is still ongoing) using a commercial 65 nm CMOS technology. The technology used for the prototype has also been characterized and validated for High Energy Physics (HEP) use and radiation hard design.

ABSTRACT (ITALIANO)

Questa tesi di dottorato riguarda l'analisi e lo sviluppo di nuove soluzioni per sistemi elettronici per il rilevamento di radiazioni, in particolare per utilizzo in ambito spaziale o per la fisica delle alte energie. I principali blocchi di un sistema di acquisizione sono stati studiati, sviluppando sistemi completi, ricercando soluzioni che permettessero di superare le tecnologie disponibili allo stato dell'arte.
Due diverse architetture sono state studiate, ognuna adatta a una applicazione specifica: Fractional Packet Counting, per sensori ad integrazione ad alto range dinamico, e CLICpix, un esempio di un detector ibrido basato su photon counting. Le principali specifiche sono state analizzate e sono state proposte soluzioni in grado di rispettarle.
Un prototipo di CLICpix e' stato progettato, realizzato e testato (la caratterizzazione e' ancora in corso) utilizzando una tecnologia commerciale CMOS a 65 nm. Questa tecnologia e' stata anche caratterizzata e validata per l'utilizzo nel disegno di strutture rad-hard nella fisica delle alte energie.

PUBLICATIONS

Some ideas and figures have appeared previously in the following publications:

- Valerio, P., Alozy, J. A., Arfaoui, S., Ballabriga, R., Benoit, M., Bonacini, S., Campbell, M., Dannheim, D., De Gaspari, M., Felici, D., Kulis, S., Llopart, X., Nascetti, A., Poikela, T., Wong, W. S. (2013). A prototype hybrid pixel detector ASIC for the CLIC experiment. *Accepted as Oral presentation at TWEPP 2013*
- De Gaspari, M., Alozy, J. A., Ballabriga, R., Campbell, M., Llopart, X., Poikela, T., Valerio, P., Wong, W. S. (2013). Design of the analog front-end for the Timepix3 and Smallpix hybrid pixel detectors in 130nm CMOS technology. *Accepted as Oral presentation at TWEPP 2013*
- Bonacini, S., Valerio, P., Avramidou, R., Ballabriga, R., Faccio, F., Kloukinas, K., Marchioro, A. (2012). Characterization of a commercial 65 nm CMOS technology for SLHC applications. *Journal of Instrumentation*
- Wong, W. S., Anton, G., Ballabriga, R., Blaj, G., Böhnel, M., Campbell, M., Gabor, T., Heijne, E., Llopart, X., Michel, T., Ritter, I., Poikela, T., Sievers, P., Tlustos, L., Valerio, P. (2012). Electrical measurements of a multi-mode hybrid pixel detector ASIC for radiation detection. *Journal of Instrumentation*
- Wong, W. S., Anton, G., Ballabriga, R., Böhnel, M., Campbell, M., Heijne, E., Llopart, X., Michel, T., Münster, I., Plackett, R., Sievers, P., Takoukam, P., Tlustos, L., Valerio, P. (2011). A pixel detector asic for dosimetry using time-over-threshold energy measurements. *Radiation Measurements*
- Nascetti, A., Valerio, P. (2011). Use of fractional packet counting for high dynamic range imaging applications. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*
- Nascetti, A., Valerio, P. (2011). Design of pixel electronics based on asynchronous self-reset approach with floating-point output representation for high dynamic range imagers. *Journal of Instrumentation*
- Nascetti, A., Valerio, P., Caputo, D., De Cesare, G. (2010). Detection system based on a novel large area hybrid detector. *Microelectronics Journal*

With engineering, I view this year's failure
as next year's opportunity to try it again.
Failures are not something to be avoided.
You want to have them happen as quickly as
you can so you can make progress rapidly.

— Gordon Moore

ACKNOWLEDGMENTS

So, where to begin... I'm about to be a PhD! Looking back at my studies, it seems that some of what happened has been slightly *random*. I've studied mostly digital stuff at the university, but I worked characterizing a pixel detector for my thesis only because it was my opportunity to go abroad for a while. It turned out my work on pixel detectors helped me to come to CERN, which was something I'd never even thought about before.

Sure, I have plenty of people to thank for being here. Augusto, my university supervisor, for helping me when I was in Rome and supporting me when I went to Geneva, along with all the jolly good fellows who worked with me at the university. My parents who always supported me and especially Enrica, who still loves me despite my choice of pursuing a career abroad (or maybe because of it...). At CERN I've found plenty of awesome colleagues, starting from Rafa, my supervisor, who helped me a lot, not only at work. I have to thank all my "bosses": Lucie, Michael, Alessandro, who, contrary to how a boss is usually depicted, were in fact always kind and supportive. I'd like to mention the other colleagues who helped me and who put up with my stupid emails from time to time: Xavi, Winnie, Massimiliano, Tuomas, Sandro, Szymon, Jérôme, Davide, Lorenzo, Thanu, Daniele (who shared the apartment with me, by the way) and many others.

I also need to thank all the fantastic people I've met here in the *CERN Games Club* (they are too many to list, thank you all!) and our *Random Trip Club* (Hugo, Angela, Herta and Lena, who helped me go through some rough times: tack!). I still miss all my friends I've left in Italy, but I can say people here are awesome too. Speaking of friends in Italy, I need to list at least some: my three "DB pals", Nunzia, Simona and Silvia, and everyone who shared silly passions with me even while hundreds of kilometers away (fantasy football, videogames, *sillier things*).

Finally, I would like to thank all people I've met *once*, whose name I can't remember who gave me a laugh or shared a few minutes of their time with me. Traveling around I've found out that the world is full of diverse people who are wonderful, each in their own way.

Life is good, even as a PhD student. Hell, *especially* as a PhD student. Just enjoy the ride.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Part I

INTRODUCTION

This part presents the topic of the thesis and the motivations for researching radiation detectors. Hybrid pixel detectors and monolithic detectors, two of the technologies most used in HEP will be compared. The main characteristics of a detector system will also be discussed. Various applications for different detectors will also be presented, with state-of-the-art examples.

1 INTRODUCTION TO RADIATION DETECTION

1.1 MOTIVATION

The topic of this thesis is the study of electronics for radiation detectors. Radiation detectors are used to track, identify and measure particles and radiation in a multitude of environments and applications. The most common application of a radiation detector is digital imaging: examples of detectors are CCDs or CMOS pixel sensors normally found in digital cameras, which are detectors sensitive to a particular kind of radiation, visible light. These detectors are "pixellated": they are made up multiple of small elements (pixels), each acquiring data independently. The resulting image is formed by reading the data acquired by all pixels.

Sensor designed to detect different kinds of radiations are used in other applications. One of the environment for which there is need of state-of-the-art detector system is High Energy Physics (HEP). HEP is the science that studies the nature of particles that constitute the matter. Silicon pixel detectors were introduced in the HEP environment at the beginning of the 1990s (see Anghinolfi et al. [1]) as a replacement for strip detectors when a high spatial resolution is needed. Multiple layers of pixellated detectors are used in modern HEP experiments to "track" and identify particles passing through them. Pixel detectors are now widely used for this application and they constitute a fundamental block of the HEP experiments operating at the LHC at CERN, among others.

An important application for novel detectors is also space science. Pixel detectors are used for optical and X-ray astronomy, for planetary and solar science. Cameras or imaging devices are often mounted on satellites, both for space and earth observation. Another case in which a radiation sensor can be used in space is as a dosimetry device for astronauts, to measure their exposure to potentially harmful radiation (see Stoffle et al. [2]).

Technology developed for space and HEP applications, especially for X-ray imaging, can be applied in other fields too. One particular field of application is medical imaging, where pixel detectors are one of the main components of systems for Computed Tomography, Positron Emission Tomography (PET) and other diagnostic techniques.

1.2 MAIN CHALLENGES

One of the main requirement of a detector system is being able to work reliably in its operational environment. For many applications (space, most of HEP experiments) this means that the electronics must be able to cope with a highly radioactive environment, that can cause damage to the devices over time, or sudden failures. The effects of radiation damage are detailed in chapter 4, along with techniques to build radiation-robust systems.
Other challenges include meeting the requirement on the accuracy of the measurement. One of the most important parameter is the resolution, both spatially and of other properties (energy, time-stamp) if the system is designed for them. In order to achieve a high spacial accuracy the detector must feature small pixels, so that the pixel matrix has a smaller pitch. The need of smaller pixels requires the use of newer downscaled CMOS technologies, as explained in chapter 3. Other important characteristics for many applications are the dynamic range of the detector (the range of particle energy that can be correctly detected) or the speed of the front-end (how many particles can be detected in a unit of time).
Improvements of the technology, together with smaller pixels, opened the possibility to build "intelligence" in the pixels, in the form of digital circuits that can implement advanced features (like, for example, in Ballabriga et al. [3]). Some of the capabilities that can be included are calibration circuits, on-chip data processing or error correction algorithms.
All the features of modern detectors come, of course, at the cost of additional complexity and power consumption. In many applications a high power consumption is to be avoided, as heat dissipation can be a problem. In space, for example, detectors may be used in an airless environment, which reduces the cooling possibilities. In HEP experiments a cooling system may interfere with the measurement being performed, due to a requirement on material budget (see chapter 3). Novel techniques allowing to reduce the power consumption are thus needed.

1.3 AN ENGINEERING APPROACH

A number of different architectures can be used to design a radiation detector, each suited to different applications. One of the main difference between detectors is the way they measure the charge deposited by particles, by integrating it over time or by counting single events. In both systems charge is produced by the interaction of a particle with a material layer that can produce electron-hole pairs when exposed to radiation. In integrating systems the charge is collected in an analog front-end and integrated over time. Currents generated by

other sources (such as leakage currents) are integrated as well, producing a noise signal. The amount of charge collected in a specific acquisition time is then stored and measured and read out. An event-counting device, on the other hand, compares the collected charge with a threshold to detect single events. If the charge is above the threshold a counter is incremented, otherwise the signal is discarded. This approach allows for a higher Signal to Noise Ratio (SNR) since the contribution of leakage currents can be discarded. It also allows to perform additional measurements such as acquiring the time of arrival. Multiple thresholds can also be used to discriminate particles which deposited more or less charge, helping their identification. The main disadvantage of such systems compared to integrating devices is that the system must be ready to acquire a new particle after detecting one. This limits the number of particles it can detect if they arrive at a high rate (its "count rate" is limited).
Detection systems can also be divided in different categories according to the technology used for charge collection. Here the two main technologies used for HEP applications are presented, but there are many others used in other environments. For an overview of them, see for example Bergauer [4]. One way of collecting charge is using the Monolithic Active Pixel Sensors (MAPS) approach. These devices incorporate in the same substrate a thin layer of sensitive material (generating electron-hole pairs interacting with incoming particles) as well as the readout electronics, which can be built using standard CMOS technology. The signal is generated by the charge being collected by a diode formed between the n-well and a lightly doped, thin p-type epitaxial layer built between the CMOS processing layers and the substrate. The generated charge inside the epitaxial layer moves due to thermal diffusion since the layer cannot be fully depleted using standard CMOS voltages. The high doping concentration of the neighboring p-type substrate and p-wells confine the charge inside the epitaxial layer until it reaches the collection diode. The standard readout of a CMOS MAPS uses three transistors (3T) per pixel: one is used to select the pixel, one to buffer the signal coming from the diode (in a source follower topology) and another one to reset the pixel after events (and to compensate for leakage currents). One should avoid using any additional n-well in the pixels, as they would collect charge from the epitaxial layer instead of the collection diode, lowering the input signal. An effort to solve this limitation is being made using modern triple-well options available in many commercial technologies (see for example Bettarini et al. [5]). A diagram of how this architecture works can be found in figure 1.

Even though this technique allows more complex developments, with full CMOS readout systems, there are still two limitations. The

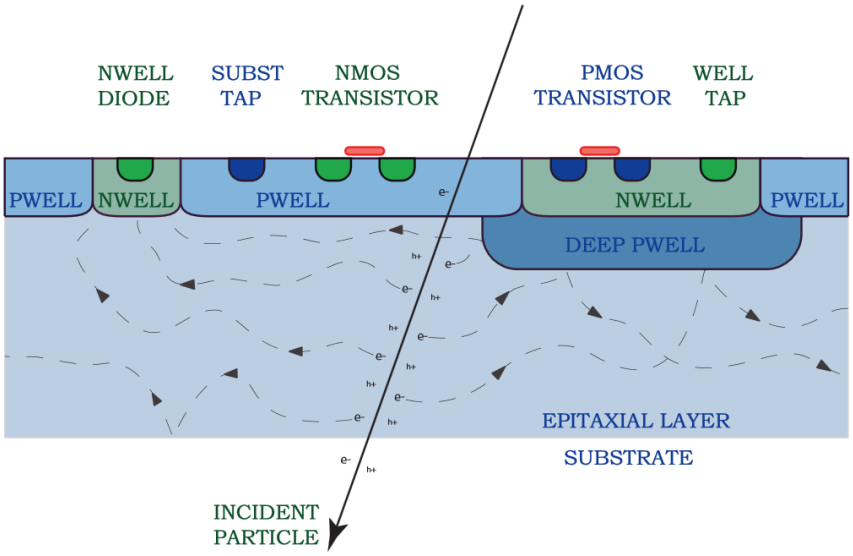


Figure 1: Schematic of a monolithic pixel detector, with a PMOS transistor using a deep n-well implantation technology.

first is the availability of the technology: the thickness of the epitaxial layer changes between foundries, so the design is very technology-dependent. The second one is the slow charge collection, due to the relying on charge diffusion, which limits the possibility of performing accurate timing measurements.
The other main technology used for radiation detectors (which is the one used for the project described in chapter 5) is the hybrid pixel detector architecture. In this kind of detector, the readout electronics is built separately from the sensitive material. The sensor is divided in pixels with the same pitch as the readout chip and the two are connected using flip-chip technology (see the diagram in figure 2). Since the two parts are produced separately, they can be optimized and designed independently from each other. Also, any standard CMOS technology can be used to design the readout electronics, so the advances in the lithographic process can be exploited to build more advanced systems, with smaller features and/or more features. The main disadvantage of this architecture is the cost of the flip chip process, especially for detectors with very small pixels.

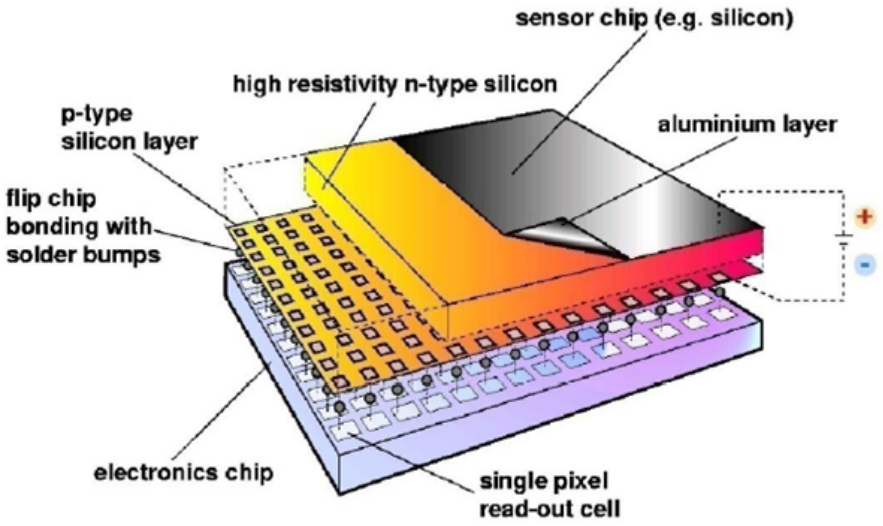


Figure 2: Schematic of a hybrid pixel detector, with the sensor on the top and the electronics at the bottom.

Part II

HIGH DYNAMIC RANGE DETECTORS

This part presents Fractional Packet Counting (FPC), a novel architecture for a High Dynamic Range (HDR) front-end for pixellated imaging systems, to be used in applications where Dynamic Range is important, such as space observation. Detailed simulations, comparing this architecture to other state-of-the-art solutions for HDR imaging will also be presented.

2

THE FRACTIONAL PACKET COUNTING ARCHITECTURE

2.1 MOTIVATIONS AND APPLICATIONS

The term Dynamic Range (DR from now) is used to address the capability of a detection system to measure radiation with a wide range of energies. It is calculated by dividing the highest signal the system can measure (before saturating) by the minimum detectable signal. DR is an important figure of merit in many imaging fields, in particular medical imaging or space observation, where a DR of 10^6 or greater is usually needed. This circuit can be therefore used in applications such as computed tomography (see Luhta et al. [6]), readout ICs for hybrid detectors or Thin Film on ASIC (TFA) imagers. The readout electronics is often the limiting factor in the design of High Dynamic Range imaging systems, often due to the constraints in terms of area occupation and power consumption imposed by the pixel size and pixel count of the imager itself. A novel readout architecture that can achieve a high dynamic range will be proposed in this chapter, along with other state-of-the-art solutions.

2.2 ARCHITECTURES FOR HDR DETECTORS

Several methods for extending image sensor dynamic range have been developed. An overview of many architectures commonly used in HDR detectors can be found in Kavusi and El Gamal [7]. One of the most common acquisition mode is integrating the input signal over a certain time and using a sample-and-hold circuit with an Analog-to Digital Converter (ADC) to measure the voltage at the end of the acquisition time. This principle works well for very low input signals, but it rapidly leads to saturation for a high input current. Another architecture is to integrate the input signal and compare it with a threshold, periodically resetting the integrator when its output is above the threshold. Counting how many times the integrating capacitance is reset in a certain period of time gives an estimation of the input current value. This principle, on the other hand, has a high minimum detectable signal, as signals which are so low that they don't cause any reset are not detectable. The main concept is using data coming from the two different measuring circuits at the same time, one designed to measure very small signals and one for very large signals, and use both of them to extend the DR as much as possible.
As an example, a possible architecture is the synchronous reset with

residue readout, first proposed in Rhee and Joo [8]. The photo current is integrated and converted into voltage $v(t)$, which is periodically compared to a reference voltage V_{max} . If $v(t) \geq V_{max}$, the comparator switches, the integrator is reset, and the counter is incremented. At the end of integration, the digitized value of output voltage and the reset count are combined to estimate the photo current.
The concept presented in this chapter is similar, implementing an architecture called fractional packet counting (FPC, see Nascetti and Valerio [9]). The principle behind the fractional packet counting is to integrate the current with an integrator, which is asynchronously reset when its output reaches a given threshold. After the end of the integration time, an ADC is used to measure the last incomplete charge packet. The digital result of the conversion is composed of both the number of resets and the charge collected in the last packet providing a fractional representation of the charge delivered in the integration time. A block diagram is shown in figure 3. The added benefit of FPC over other HDR architectures is working at a constant relative resolution, which means to use a floating point representation with a constant number of significant bits. This solution is implemented with a floating point logic that stores the output in a single register and adapts the resolution of the ADC to the input, so that the amount of data produced can be significantly reduced.

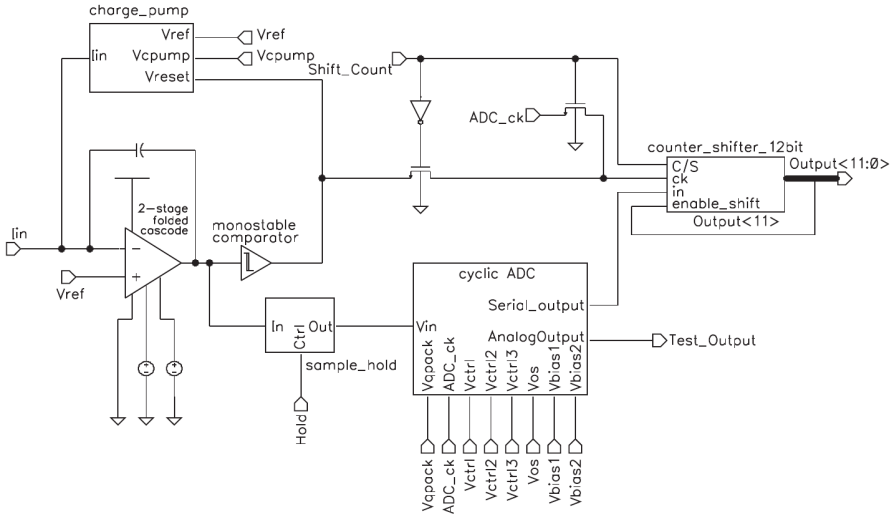


Figure 3: Block diagram of a circuit implementing the FPC architecture.

The compact output representation is achieved during the conversion itself, i.e. without any post-processing, by combining a successive approximation approach for the calculation of the fractional part and a sequential circuit that controls the output register by left-shifting its content at the end of the integration time until all the leading zeros are skipped. During this action, the bits of the fractional part, calculated by a successive approximation cyclic ADC, are shifted in

the register and a counter stores the position of the binary point to allow the correct interpretation of the output string. This approach allows to achieve a simple pixel-wise and frame-wise automatic tuning across the entire dynamic range of the circuit. The operation of the binary point logic is schematically described in figure 4.

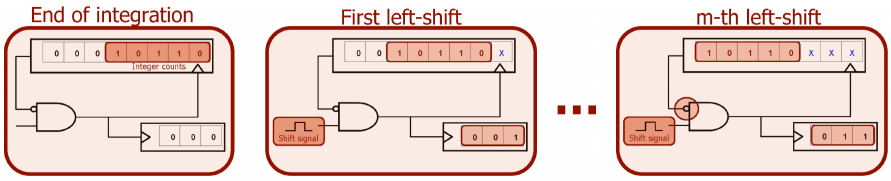


Figure 4: Binary point logic operation. When all the leading zeros of the integer counts value are skipped the logic gate prevents the execution of further shift commands and the least significant bits contain the fractional part.

2.3 MAIN BUILDING BLOCKS

The circuit implementing the scheme described above has been designed in a commercial 350nm technology. The integrator is an OTA with a feedback capacitance on which the charge is collected. The amplifier has a folded cascode architecture to increase the gain up to 106 dB to reduce charge loss due to offset. In order to prevent the loss of signal charge during the reset action, the reset mechanism has been implemented with a charge pump circuit, connected at the input note, which subtracts a constant amount of charge from the integration capacitance. In order to achieve a good linearity, the feedback signal that triggers the reset has to ensure the full discharge of the charge pump capacitor. This is achieved by using a monostable comparator, which supplies a reset pulse of constant duration. Since the input current is integrated continuously, even during the reset, the voltage at the integrator output node increases above the threshold and also it does not reach the initial reset condition. This effect increases with increasing the signal current as shown in figure 5. Since in the proposed scheme the resolution of the ADC decreases when the current increases, however, the resulting error is always lower than one LSB, so its effect on the measurement is not relevant. An integral non-linearity below 1 LSB is achieved.

The second stage is a 10-bit cyclic ADC (the schematic is shown in figure 6) that calculates the fractional part bits during the binary point position search algorithm. The cyclic working principle was chosen as its resolution can be easily controlled by the floating point logic, even though in principle any type of ADC can be used. The proposed ADC uses a switching matched capacitors approach to generate different

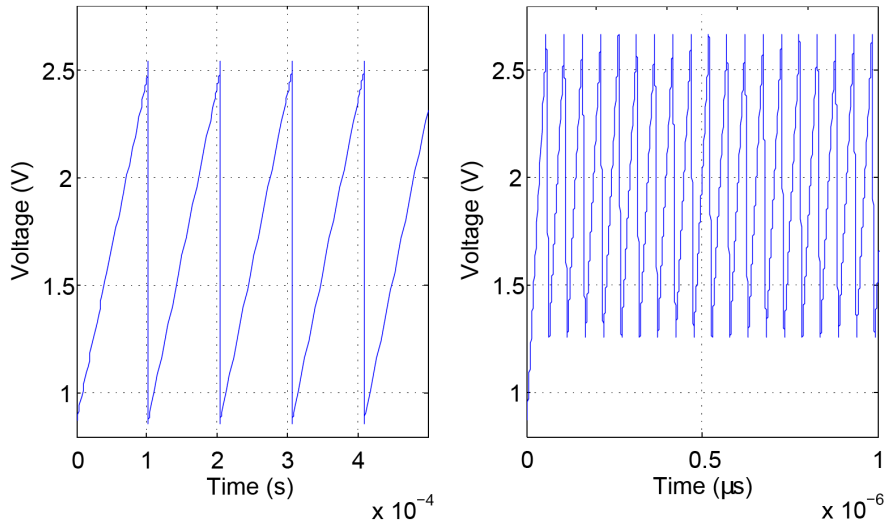


Figure 5: Integrator output at different currents. A decrease in the output dynamic range can be observed, but the effect is not relevant due to the reduced ADC resolution for higher reset counts.

thresholds at each cycle that are compared with the input voltage. The circuit is composed of four blocks: the charge sharing unit, the integrator that generates the new threshold voltage, a comparator to compare it with the charge packet counting output sampled at the end of the signal integration time and a D-latch that holds the comparator output for the next conversion step and provides the input to the output register. Using a matched capacitor architecture allowed designing a compact converter, requiring few control signals, suited for multi-channel readout circuits.

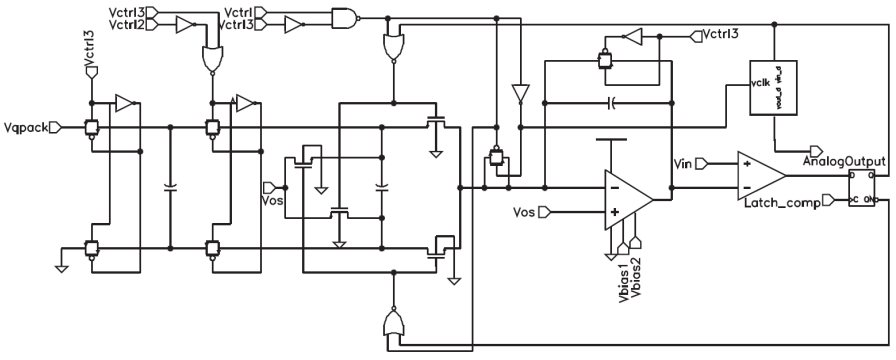


Figure 6: Detailed Cyclic ADC schematic.

For 10 bits operation, a minimum of 79 db open loop gain must be achieved (calculations can be found in Manen et al. [10]). In the same

work an estimation of the minimum value for the charge sharing capacitances is detailed, relating KT/C noise to the ADC characteristics:

$$U_{noise} = \frac{U_{max}/2^{N+1}}{\sqrt{12}} = \sqrt{\frac{KT}{C}} \tag{1}$$

where Umax is the input voltage swing (1.7 volts in this case) and N is the number of bits. The minimum value for capacitors used for generating the threshold is therefore only 0.144 fF which isn't a limiting factor for the capacitors size. The size of these capacitors, on the other hand, is limited by the pixel size and the mismatch ratio, which has to ensure the 10-bit resolution. The desired mismatch rate must be better than $1/2^{N+1}$, so less than 0.5%.
The comparator must have a gain high enough to get a full output swing for an LSB input, so in this case at least 76 dB. Its bandwidth must also be sufficient for a 2 KHz operation and it must work in a 400 ns time window, so a bandwidth of at least 1.25 MHz was chosen. Simulations performed on the overall circuit showed that it can reach 117 dB dynamic range with a resolution of 12 bits (for the counter) and 10 bits (for the ADC) working at a 2 KHz frequency. The minimum detectable signal is 24 fA, while the maximum nonsaturating current is 400 nA.

2.4 QUANTITATIVE ANALYSIS

The dynamic range of the circuit can be calculated as the ratio of the largest non-saturating photo current to the smallest detectable signal. For an integrator circuit, it is:

$$DR = \frac{i_{max}}{i_{min}} = \frac{Q_{max}}{\sigma_{readout}} \tag{2}$$

where $i_{max} = q \cdot Q_{max}/t_{int}$ and $i_{min} = q \cdot \sigma_{readout}/t_{int}$, being Q_{max} the maximum charge stored in the integrating capacitance before the reset, while $\sigma_{readout}$ is the equivalent noise of the conversion. In the FPC architecture, the current is integrated and converted into voltage, which is compared to a reference voltage. When the comparator switches, the integrating capacitance is reset and the counter is incremented. At the end of the integration period, the residue value is sampled and combined with the count number to obtain the final result. Following the same approach detailed in Kavusi and El Gamal [7], the total average distortion error (here referred to as $\sigma^2_{distorsion}$) can be expressed as being due to the finite time needed for the integrator to reset.

$$\sigma^2_{distorsion} = \frac{i \cdot T_{switch}}{T_{switch} + t_{rise}} \tag{3}$$

which depends on the speed of the comparator and of the reset circuit. The usage of a charge pump reset mechanism should in principle cancel this limitation, by integrating the charge also while resetting. Due to non-ideal devices, though, there is a remaining error which can still be modeled with a finite reset speed (estimated to be ~4 ns in simulations). More detailed calculations can be found in Nascetti and Valerio [11]. To find the total noise we must add contributions from shot noise, ADC noise and gain FPN (which is in part due to reset offset variation):

$$\sigma_{tot}^2 = \sigma_{distorsion}^2 + \frac{q \cdot i}{t_{int}} + (n_{reset} + 1)\left(\frac{q \cdot \sigma_{reset}}{t_{int}}\right)^2 + \\ + \left(\frac{q \cdot \sigma_{readout}}{t_{int}}\right)^2 + \left(\frac{q \cdot \sigma_{offset} \cdot n_{reset}}{t_{int}}\right)^2 + (\sigma_{FPN} \cdot i)^2 \qquad (4)$$

where $\frac{q \cdot i}{t_{int}}$ is the integrated shot noise, σ_{reset} is the comparator threshold offset, $\sigma_{readout}$ is the ADC quantization error, σ_{offset} is the integrator offset variation and σ_{FPN} is the fixed pattern noise.
The model described above can be used to derive the SNR decrease due to analyze the impact of the fixed relative resolution on the SNR. By devoting less bits to the fractional part we are basically reducing the ADC effective resolution. This can be included in the SNR formula by multiplying the $\sigma_{readout}$ noise source by a factor of 2 for each lost bit. The resulting value for the ADC noise is therefore $\left(\frac{q \cdot \sigma_{readout} \cdot 2^B}{t_{int}}\right)^2$, where $B = \lceil \log_2 (n_{reset} + 1) \rceil$. The result is plotted in figure 7. It can be clearly seen that this solution has almost no impact on the SNR.

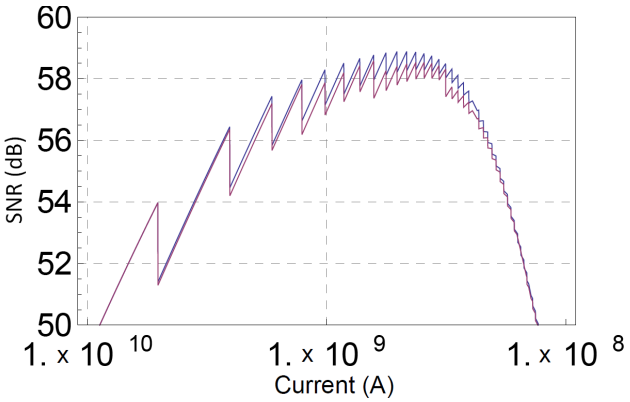


Figure 7: Comparison between the SNR of a fixed point notation (blue curve) and a floating point solution (red curve) using the FPC architecture.

Additionally, looking at the various noise contributions, a plot of the SNR accounting only one noise source at a time could be derived

(as depicted in figure 8). At lower currents, the limiting factors are the shot noise (red line), the variability of the discriminator threshold (green line) and the ADC resolution (purple curve). At higher currents the reset mechanism (blue line) is the main cause of noise, while the horizontal orange line describing the fixed pattern noise sets an upper limit to the SNR.

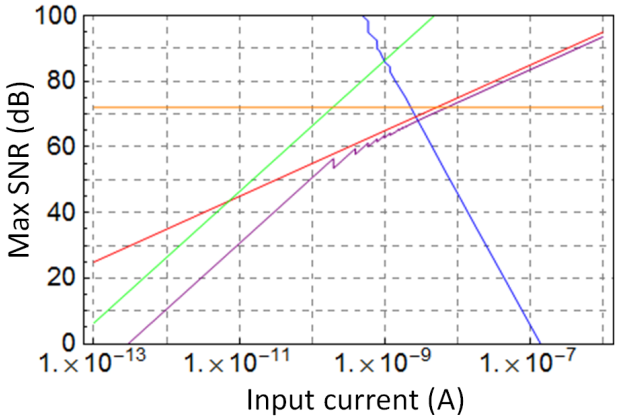


Figure 8: Comparison between the different noise sources in the FPC architecture.

The results of this analysis show that the FPC architecture is suitable for an high dynamic range front-end, allowing a reduction in the data bandwidth (since the number of bits to transmit is reduced) with a negligible impact on system performances. This improvement is particularly useful for applications such as space observation, where data rate is an important factor in the system design.

Part III

ELECTRONICS FOR HIGH ENERGY PHYSICS

This part focuses on the work carried out to design, prototype and test CLICpix, a pixel detector specifically designed to be used as the vertex detector for CLIC, a novel particle accelerator being currently studied at CERN. All the phases of the project, from the feasibility study to the testing will be presented, explaining both technological and design problems faced during the work and how they were approached and solved.

3

THE CLIC EXPERIMENT

3.1 A NEW PARTICLE ACCELERATOR

This part of the thesis work was carried out at CERN, in Geneva, as part of a Doctoral Student program, working for the CLIC project.

3.1.1 *Physics of the CLIC linear accelerator*

CLIC (acronym for Compact LInear Collider) is a linear particle accelerator currently under study at CERN (see Aicheler et al. [12]). A linear accelerator is a type of particle accelerator which uses a linear beam-line, as opposed to other circular accelerator, such as the Large Hadron Collider (LHC) which is being operated at CERN. The basic principle of a linear accelerator is to use a series of oscillating electric potentials (usually using RF cavities) to increase the velocity of charged particles along the beam-line (see Ising [13]). A very common, albeit simple, example of a linear accelerator is the cathode tube used in old television sets. CLIC is an electron-positron (e^+e^-) collider, i.e. it accelerates electrons and positrons on opposite directions to make them collide at the center of the beam-line, where a complex series of detectors is placed to acquire data on the collisions. The need for an e^-e^+ accelerator comes from its capability of allowing precise measurements on collisions at energies from a few hundreds GeV to 3 TeV. This allows for precision measurements of the Higgs boson and the testing of such models as supersymmetry, Higgs strong interactions, contact interactions and extra dimensions.
CLIC uses room temperature RF cavities, as this design allows for a relatively short beam-line length compared to superconducting cavities (see Jensen [14]). Since no conventional RF source can drive the necessary power for the beam (at the needed frequency, which would be 12 GHz), a two-beam acceleration scheme has been designed, where a high-current low-energy beam is used as an RF power source for the main particle beam (low-current, high-energy). The main beams are brought into collision at the center of the accelerator, where the detector is installed (as shown in figure 9).

3.1.2 *Comparison with the LHC*

There are significant differences between CLIC and LHC, the particle accelerator that is currently used for the main experiments at CERN.

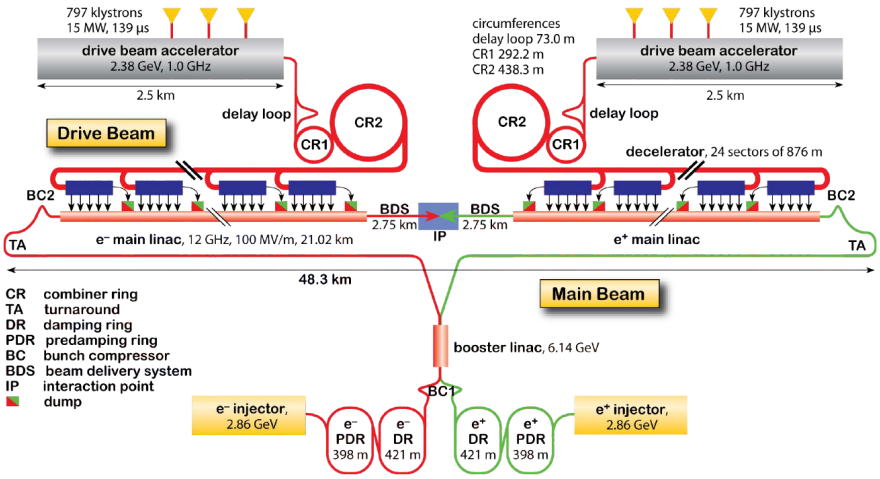


Figure 9: Double beam accelerator scheme, taken from Aicheler et al. [12].

LHC is a hadron collider and its physics program consists mainly of proton-proton collisions, with the possibility of running heavy-ion collisions. This allows to investigate a very large energy spectrum, going as high as 14 TeV when running at the intended operating point and it is indeed leading to very interesting and important results (see [15]). Protons, however, are not elementary particles and their center-of-mass cannot be determined accurately, making the analysis of data from the LHC experiments very challenging. A linear accelerator such as CLIC can accelerate electrons and positrons, which are elementary particles, on the other hand, so it can be used to achieve a higher accuracy than any proton collider. Moreover, CLIC operating at 3 TeV can allow the exploration of energy ranges not covered by existing accelerators, leading to the possible detection of new particles and the testing of theories such as supersimmetry (see Linssen et al. [16]). The main reason why such accelerator cannot be circular like LHC is because the accelerated particles have a much lower mass and thus they would lose much more energy due to synchrotron radiation.

3.1.3 *Some specifications*

The linearity of the machine means it is possible to upgrade the machine by making it modular: a shorter linac can be initially built running at lower energies, making it longer to increase the collision energy in a second phase. An initial version running at 500 GeV (using a 13.2 km long tunnel) has been proposed, which would then be upgraded up to the full 3 TeV (with a 48.3 km long tunnel) at a later stage. The accelerator is anyway being designed according to the specification of the higher energy version from the beginning, to simplify the upgrade process.

Parameter	Unit	Stage 1	Stage 3
Center-of-mass	GeV	500	300000
Repetition frequency	Hz	50	50
Number of bunches per train		354	312
Bunch separation	ns	0.5	0.5
Accelerating gradient	MV/m	80	100
Luminosity	$10^{34} cm^{-2} s^{-1}$	2.3	5.9
Main tunnel length	km	13.2	48.3
Power consumption	MW	272	589

Table 1: Parameters for two CLIC energy stages (taken from Aicheler et al. [12]).

The two-beam scheme works makes it necessary to have a very small duty cycle, with bunch trains being produced at a frequency of 50 Hz. Every bunch train is composed of 312 bunches and lasts for 156 ns (so that bunches are separated by 0.5 ns). The "off" time, when no collision is produced, can be used to read the data out from the detectors. Also, the low frequency of bunch trains allows a *trigger-less* scheme to be used, as the start of the data acquisition can be synchronized globally and every chip can be read out between two bunch trains. In LHC, for example, a *trigger* signal must be produced by the detectors in order to determine when to acquire data. The *trigger-less* architecture simplifies the design of the detectors, as they can rely on an external signal.
CLIC is designed to have two experiments sharing a single collision point (as it's impossible to have multiple collision points like it is done in a circular collider), the CLIC_ILD and CLIC_SiD. The detectors will be moved using a "push-pull" system. Both detectors have a length of 12.8 m along the beam line and a diameter of 14 m, with the collision point situated in the centre. Surrounding the collision point there is a succession of particle detectors of various types, aiming at collecting a maximum of information on the particles that are produced in the collision. This allows to measure, for example, the momentum, charge, energy and particle type of the particles. The main difference between the two experiments is that the CLIC_ILD concept is based on a Time Projection Chamber, which provides a highly redundant continuous tracking with relatively little material in the tracking volume itself. The CLIC_SiD concept has a compact all-silicon tracking system, which has the advantage of fast charge collection.
A summary of the parameters of CLIC (for two energy stages) can be found in table 1

3.2 REQUIREMENTS FOR THE VERTEX DETECTORS

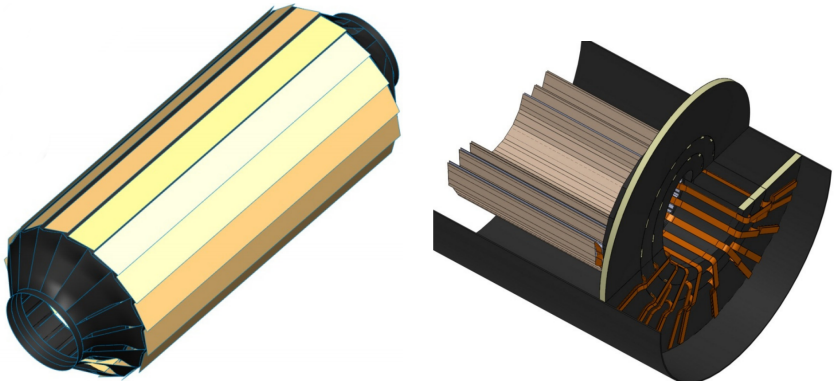
3.2.1 *General requirements for the vertex detector*

The work described in this part focuses on the vertex detector[1] of the CLIC experiments. The vertex detector specifications are the same for both the CLIC_ILD and CLIC_SiD concepts. In order to acquire data with sufficient accuracy and efficiency, the position of the passage of a particle in the silicon sensor layers of the vertex detector needs to be measured precisely. Simulations show that a spatial resolution of ~3 µm is needed to avoid deterioration of the accuracy due to multiple scattering. The material budget should be be kept smaller than 0.2% X_0[2] for the beam pipe and each of the detection layers. The requirement of a very low material budget reflects directly in other requirements for the detector. The most important one is the need to use air cooling (with a forced flow) to dissipate the heat produced by the electronics, as it would not be possible to meet the material budget specification with a cooling liquid running in pipes. This limits the amount of power the electronics can use without the risk of overheating and, as it will be explained in the next section, it is one of the main challenges of the design of the pixel detector. Mechanical support structures and power and signal distribution systems must also face this restriction and they are being designed using innovative architectures and materials to keep the material budget as low as possible (see Linssen et al. [16]). In order to cover the needed area, "modules" of multiple chips will be built and arranged in a circular "barrel". Multiple concentric barrels will form the inner layers of the detectors (as explained in Linssen et al. [16]). Given the pixel size, each module will be composed of 24 chips arranged in a ladder, (as shown in figure 10).

Another requirement for an accurate data reconstruction is having a Time-of-Arrival (TOA) measurement for the particle hit, allowing to identify to which bunch every hit belongs (with some uncertainty, a 10 ns accurate time-stamp is enough, according to simulations). Si-

1 A vertex detector in a collider experiments is a particle detector positioned as close as possible to the collision point. The goal of a vertex detector is to measure particle tracks very close to the interaction point

2 "Material budget" is used to define the thickness of the material used in the detector layers, especially the inner ones. Interactions of particles with the detector itself (including the mechanical support structures, cabling, etc.) lead to energy loss from bremsstrahlung or other effects, reducing the accuracy of the measurements. Thus, it's important to reduce the energy losses, by making the detector layers as thin as possible. The material budget is usually measured in radiation lengths (X_0), which is a characteristic of a material, relating the energy loss of particles when interacting with that specific material. It is defined as the mean distance over which a high-energy electron loses all but $1/e$ of its energy bremsstrahlung and for silicon its value is about 10 cm.



(a) External view. (b) Section view.

Figure 10: The inner CLIC detector barrel.

multaneously, an energy measurement must be performed, to estimate the amount of charge deposited by each particle and to allow sub-pixel spacial resolution.

3.2.2 *Specifications for the pixel detector*

The requirements of material budget translate directly in a specification on the thickness of the silicon sensor, in addition to other structures. The allowed material assigned to the electronics is equivalent to less than 200 µm of silicon. There are two different approaches to designing such a detector, a monolithic one or a hybrid pixel detector, with a silicon sensor bump bonded to the electronics (as already discussed in section 1.3). The monolithic solution would allow for a very thin detector, as it doesn't need an additional die for charge collection. The hybrid solution, on the other hand, can achieve a better time resolution and allows more "intelligence" to be implemented in the pixel area while still using a standard CMOS process. In this thesis a hybrid detector solution has been studied and developed. In order to keep the material budget low, very thin substrates must be used. The possibility of building a 50 µm thick silicon sensor for charge collection has been already demonstrated (such sensors are available, for example, from Micron Semiconductor, as shown in [17]), and the process of thinning a chip substrate to 50-100 µm is also available (see, among the others, [18]). The solution of a hybrid assembly with an ultra-thin sensor bonded to a thinned silicon die was deemed then to be the most suitable for the CLIC application.
In order to measure the Time-of-Arrival and the energy of the incoming signal at the same time, a similar approach to the one used in the Timepix chip was explored (see Llopart et al. [19]). One way of measuring the energy of the signal is using a Time-over-Threshold (TOT) measurement. The TOT is calculated by measuring the amount of time the incoming signal (after being converted to a voltage and

shaped) stays above a certain voltage threshold: the longer it takes for the input pulse to return to the baseline value, the larger the energy deposited on the pixel is. This approach allows using a very high gain input amplifier, reducing the equivalent noise of the front-end. A more complete explanation, together with simulation results, will be shown in figure 11.

In order to have a sufficient accuracy for the TOA measurement, a 4-bit counter is needed (the bunch train is 156 ns long, so having 16 time-slices will ensure an accuracy slightly better than 10 ns). According to physics simulations, a 4-bit TOT counter is also enough to give sufficient accuracy for the energy measurement. Each pixel, then, needs to acquire, store and readout 8 bits of data per bunch train (see Aicheler et al. [12]).

Another very strict requirement is spatial accuracy. In order to achieve an accuracy of approximately 3 µm pixels don't need to be 3 µm wide, as the position of the hit can be interpolated using also the information on charge deposited in a cluster of pixels for a particle traversing the detector due to charge diffusion. According to physics simulations, taking into account the charge diffusion in a 50 µm thick silicon sensor, a pixel of 25 by 25 µm is sufficient to achieve the needed spatial resolution.

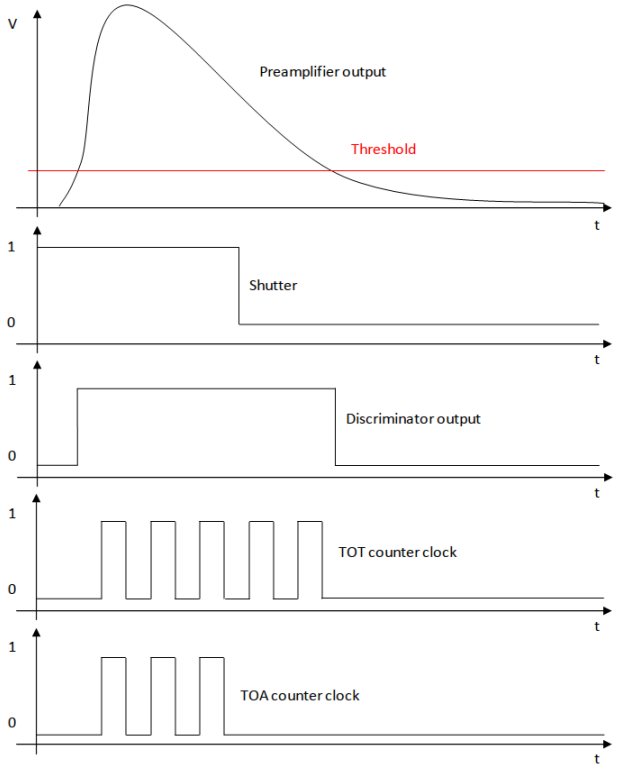


Figure 11: Example of an incoming pulse and how the counting clocks for TOT and TOA measurements are generated from the discriminator output and shutter signal.

The last requirement with a heavy impact on the choice of the architecture of the chip is the cooling, which translates directly in a power consumption requirement. The amount of heat that can be dissipated with a forced air flow system is around 50 mW/cm^2. It became soon clear (see section 5.4) that some analog structures cannot not be designed to use so little power if running continuously (a total of ~312 nW per pixel). Fortunately, the time between bunch trains (the trains are produced with a 50 Hz frequency) can be exploited to lower the average power consumption of the chip by turning off certain parts of the chip between bunch trains. A "power pulsing" scheme, then, allowing a quick power cycling is being developed for the application. A brief list of the main detector requirements can be found in table 2

Parameter	Unit	Value
Counter Depth (TOT)	bit	4
Counter Depth (TOA)	bit	4
Pixel size	µm	20-25 by 20-25
Power consumption	mW/cm^2	< 50 (after power pulsing)
Assembly thickness	µm	100
Pixel occupancy		2-5%

Table 2: CLIC pixel detector main requirements.

3.3 TECHNOLOGY AND DESIGN CHOICES

3.3.1 *State-of-the-art*

Work on the pixel detector has been carried out at CERN within the framework of the Medipix project. Medipix1, Medipix2 and Medipix3 are international collaborations, which bring together many universities and research centers from around the world to develop radiation detectors and their applications in many different fields of science. The Medipix team developed several detectors, the latest ones being Medipix3 (Ballabriga et al. [3]), Timepix3 (Poikela et al. [20]) and Dosepix (Wong et al. [21]), each with different specifications and tailored to different applications. Despite them being developed as separate projects, the group exchanges know-how and experience so that blocks and ideas can be reused among different chips. Being part of the Medipix family, the detector described in this part has been named CLICpix.
The CLICpix project has many similarities with the Timepix3 project, which is being developed by the Medipix group at the time of writing this manuscript. Both chips have simultaneous TOT and TOA measurements, even though Timepix3 can be programmed to oper-

ate in many additional modes. A small list of the main specifications of Timepix3 can be found in table 3

Parameter	Unit	Value
Counter Depth (TOT)	bit	10
Counter Depth (TOA)	bit	14
Pixel size	µm	55 by 55
Power consumption	mW/cm^2	~400 (before power pulsing)
Readout architecture		Data driven
CMOS Technology		130 nm

Table 3: Timepix3 specifications.

There are a few key differences with the CLICpix chip. The surface is more than 4 times bigger in Timepix3 compared to CLICpix. The Timepix3 chip has more complex modes of operations (being developed for a wider range of applications) and deeper counters for TOT and TOA. The analog front-ends of the two chips have a similar architecture, using the same functional blocks, so only the digital area would be reduced. In order to reduce the pixel size to 25 by 25 µm the size of both logic and analog blocks needed to be scaled, using a more downscaled technology, reducing substantially the size of features (especially digital cells, as depicted in figure 12). Details about the choice of the technology are provided in the next section.

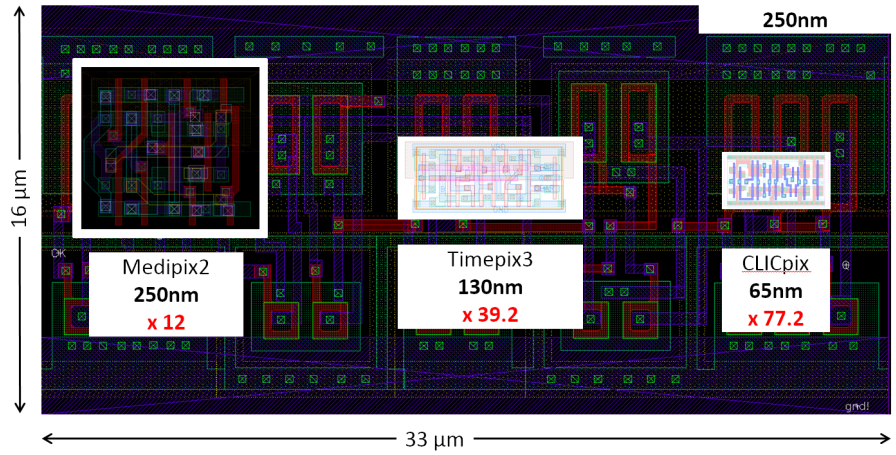


Figure 12: Size of a standard cell Flip-Flop in different CMOS technologies. The bigger cell in the background uses a special layout to increase its radiation hardness.

The timing resolution of CLICpix is also going to be different. The Timepix3 chip can achieve a higher TOA accuracy by using local oscillators running at a higher frequency than the clock being distributed

to the pixels. The 10 ns accuracy required by CLICpix (also, leveraging the faster logic allowed by a more downscaled CMOS technology) requires the clock to be distributed at full speed to whole pixel array. This choice increases the power consumption of the array, as a faster clock must be distributed, but it reduces the pixel complexity and it saves pixel area.

Another difference is the amount of data to read out from each chip per frame, which is tied to the expected occupancy. Timepix3 uses a packet-based readout (see Poikela et al. [20]), reading out only pixels with valid hits. This is efficient for low occupancy rates, as it allows to skip readout of empty pixels altogether. For CLICpix, knowing the expected occupancy of 2% to 5%, another readout scheme was chosen, reading full frames with an on-pixel data compression algorithm. As discussed in section 5.5, this solution was simulated and it was found to be more suited for the CLIC occupancy range. From the simulated amount of data to read out, 320 Mbps is sufficient to read out a whole 24 chip module using a single data line (again, to keep the material budget due to cabling as low as possible) in the 20 ms between two consecutive bunch trains.

3.3.2 *Choosing a CMOS technology*

As discussed in the previous section, the move to a more downscaled technology is needed in order to achieve a smaller pixel area. Timepix3 (like Medipix3 and other chips from the Medipix group) uses a commercial 130 nm CMOS technology; commercially available CMOS technologies with smaller features include several nodes, from 90 nm down to 32 nm.

The decision on which technology to use is both a design choice and an organization choice. This technology must in fact be supported by CERN on many levels (software and tool support, design courses, legal and administrative issues) and it cannot be used for the CLICpix project alone. The choice must go to on a technology which can then be used by other microelectronic projects at CERN.

While very deep downscaled technologies (45 nm, 32 nm) would allow for an easier design being more area efficient, other issues must be taken into account. The two main concerns on using such processes are cost and radiation hardness. In an industry environment the cost per chip of a project goes down with newer technologies, mainly because of the possibilities of having more chips on a wafer because of larger wafers and smaller chip areas. This is not true, however, for non recurrent costs (commonly referred to as NRE, Non-Recurring Engineering) including mask printing, foundry access and design time. For projects with small production volumes, such as most of the designs in the HEP community, including CERN, the total cost increases substantially when using more downscaled tech-

nologies. The other important issue is the robustness of the technology to radiation induced effects. While this is not a critical factor for CLICpix, many chips developed by the HEP community have to work in environments with very high radiation levels, such as LHC. Effects of radiation dose on microelectronics were thoroughly studied for technologies used before at CERN. These effects change quantitatively with the technology scaling, but the physics behind them are well understood (see Faccio and Cervelli [22]). New highly downscaled technologies (some commercial 45 nm, but mostly 32 nm ones) use high-K materials as gate dielectrics, as opposed to the standard silicon oxide used for older technologies. While the radiation effects on these materials have been analyzed (see Zebrev et al. [23]), no test could be found studying the response of these technologies to extremely high doses[3].
The first choice, taking all into account, fell on a commercial 65 nm technology, which could provide a substantial gain in terms of area and power consumption, while keeping the costs from increasing too much[4]. A test campaign to validate the robustness of the chosen technology with respect to radiation was, moreover, needed before a final choice could be made. Details on this characterization will be shown in chapter 5.

The chosen 65 nm low-power CMOS technology was developed for logic and mixed-signal/RF circuits, and allows multiple supply voltages for core and I/O. Its nominal supply voltage is 1.2 V and it features a high-resistivity epitaxial substrate process, shallow trench isolation (STI), two gate oxide options (1.2 and 2.5 V), nickel-silicided low-resistance n+ and p+ polysilicon and diffusion areas. Its device options contain nMOS and pMOS with several different threshold values. High-voltage 5V-drain-tolerant devices are optional. The back-end offers 3 to 9 copper metal layers for interconnection plus 1 top aluminum layer for wire-bond/flip-chip pad, pad redistribution layer and laser fuses. Low-k dielectric is used as inter-metal insulator in thin metal layers.

3 For some LHC applications robustness up to hundreds of Mrads or even a few Grads is required.

4 Unfortunately, actual costs for foundry access cannot be disclosed publicly

4

TECHNOLOGY CHARACTERIZATION

4.1 RADIATION INDUCED EFFECTS ON ELECTRONICS

Most semiconductor electronic components are susceptible to radiation damage: that is, their behaviour is modified, in terms of both reliability and performances, to the exposure of ionizing radiation (high energy electromagnetic radiation or particle radiation), a common hazard in environments such as outer space, high-altitude flight or near particle accelerators. These effects are particularly important for radiation detectors, as their role is to perform measurements on those particles or radiations which are responsible for the damage and because they are exposed to higher radiation fluxes. It is, thus important to understand the different types of radiation effects and how to properly design electronics to minimize their impact on the performances of the systems. The performances of the devices degrade because of cumulative effects, caused by the radiation dose received throughout the device operative life. These effects can be measured over a long time and they are mostly deterministic. The second kind of radiation damage effects are due to single events, which can cause an upset in a memory element or cause a latch-up. These effects are stochastic, as they are dependent on the hit of single high-energetic particle with specific circuit nodes.

4.1.1 *Total dose effects*

With the term "Total Ionizing Dose" (TID) a range of effects are indicated which are due to a prolonged exposure to a radiation source, which causes a degenerative process that impacts on the performances of an electronic device, such as a transistor. Usually in the radiation hard electronics community the radiation dose is specified in "rads", a unit of measure defined as 0.01 Gy, or 0.01 J/kg in SI units, which quantifies the amount of energy absorbed per unit of mass. After some radiation dose the performance of a circuit gets generically worse because of the damage the radiation causes in the semiconductor. Commercially available radiation-hardened microchips are usually designed and tested to be robust up to a few Mrads (see Velazco [24]). These values are referred to a 130 nm CMOS technology, but they are highly technology-dependent, as explained below. The environment of a particle accelerator such as LHC is extreme from this point of view. For example, in CMS the inner pixel detectors are exposed to a dose of ~10 Mrads/year, so the electronics was designed

to be robust up to more than 40 Mrads (see Cerati et al. [25]). There is, then, the need to carefully characterize the technology used for CMOS design in a HEP environment with respect to radiation damage. Tests to validate the performances of the CMOS process up to very high doses (200 Mrads was chosen as an upper limit taking into account the needs of other HEP projects) were thus needed.
Effects on semiconductor electronics due to total dose are known since long ago (a good summary of them can be found in Nichols [26]). The basic damage effects of ionizing radiation result from the generation of electron-hole pairs in the semiconductor substrate. The most prominent effect can be found when the pairs are generated in the oxide, being it the gate or the field oxide in a MOS transistor. At room temperature, both electrons and holes will drift under the influence of local fields (which can exist even in the absence of an applied bias) but the electrons are much more mobile so that a net build-up of trapped holes occurs. The positive trapped charge induces a negative shift in the threshold voltage and it also causes an increase in leakage currents. Since n-channel devices are operated with a positive bias on the gate, the trapped oxide charge lies much nearer to the $SiO_2 - Si$ interface than to the gate. In p-channel devices the opposite is true. This positional difference in location of trapped charge requires that a much greater change in bias is required to neutralize the trapped charge in n-channel MOSFETs so the effective change in radiation-induced threshold voltage is also much larger for NMOSFETs than for PMOSFETs. This strong bias dependence of TID effects also implies that non-powered devices are much more resistant to ionizing radiation than devices under bias voltage. At high doses, also, interface states can be generated at the oxide-silicon boundary, causing a modification of the MOS device characteristics such as a decreased mobility and an increase in propagation delays.
For deep submicron CMOS technologies the radiation effects changes, due to a much narrower oxide, and charge trapping at the edge of the transistors becomes prominent (see Faccio and Cervelli [22]). On one hand the narrow gate oxide makes it so the charge is less likely to get trapped in it and holes reach the substrate or the gate more easily. On the other hand, the Shallow Trench Insulation (STI, see figure 13) used in modern technologies doesn't scale as much as the channel length or the gate oxide thickness, so charge trapping in the STI oxide becomes statistically more and more relevant, limiting the radiation tolerance of conventional CMOS circuits. The charge in the STI at the transistor edge can cause a channel to be formed between source and drain, activating a parasitic "lateral" transistor, substantially increasing the channel leakage of the transistor. In NMOS transistors, the negative charge trapped in interface states only starts to compete with the oxide-trapped charge with some delay, giving origin to a "rebound" effect, in which the leakage current increases up to a cer-

tain dose (dependent on the technology) and then decreases again for higher doses.
For narrow channel transistors, the charge balance at the transistor edges not only determines the accumulation, depletion or inversion condition of the parasitic lateral transistor, but also influences the electric field of the main transistor. This effect is known in CMOS technologies as "narrow channel effect", and it is observable in any deep submicron process as a decrease of V_{th} with transistor width.

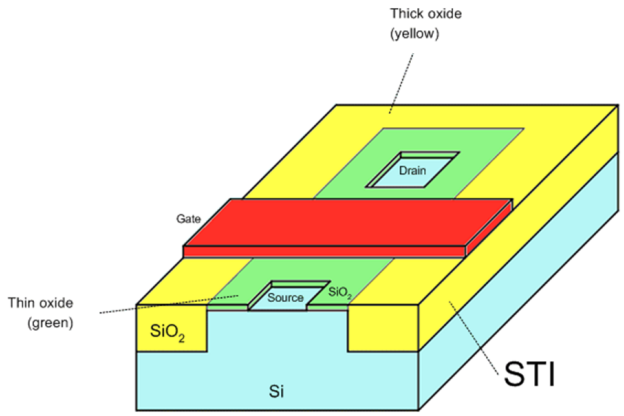


Figure 13: Diagram of a transistor with STI insulation (figure from Bonacini and Kloukinas [27]).

It should be noted that the effect described above are due to the trapping of charge inside oxide layers. Thermal annealing can, partially or totally, cause the chip to recover the original performances by detrapping those charges and restoring the damaged interface states of the gate oxide.

4.1.2 *Single event effects*

Radiation damage can also have effects other than a degradation in performances due to TID. As discussed in May and Woods [28] a "soft error" can be generated in memory elements due to accumulated charge in storage nodes. High energy ionizing particles can create electron-hole pairs close to sensitive nodes and, when the accumulated charge is enough, they can cause the value stored in a latch or flip-flop to be "flipped". These errors are random, nonrecurring and not permanent: the charge injected in a node is in fact only causing a bit error and the circuit functionality is completely recovered when the affected storage element is re-written. Such error can, though, have a very large impact on the operation of digital circuits, as it can lead to invalid data or complete system failure due to upsets in the state bits of state machines. In some cases, a parasitic transistor can be activated by the injected charge, causing latch-up, which can lead to permanent damage. Such errors are commonly referred to as

Single Event Effects (SEE) or Single Event Upset (SEU).
The probability of a particle causing one (or more) bit upset is dependent for the most part on two things: how much energy the particle is able to transfer to the silicon and the minimum charge needed for a storage element to flip state. The energy transferred is defined as Linear Energy Transfer (LET), a measure of the energy transferred to the device per unit length as an ionizing particle travels through a material. The common unit is $MeV \cdot cm^2/mg$ of material (Si for MOS devices). The minimum LET to cause a detectable effect in a node is called LET threshold (LET_{TH}). Experimental tests can be conducted to calculate the "cross section" of a device, which is a measure of the response of the device to the radiation. For a given LET, the cross section is the number of errors divided by the incoming particle fluence ($\#particles/cm^2$).
SEU robustness is strongly dependent on the CMOS technology. More downscaled technologies allow smaller feature and thus smaller memory cells. This leads to a reduction in the cross-section, as it's less likely for a particle to hit a sensitive node because of its area. On the other hand, these nodes have a smaller parasitic capacitance associated to them, so it requires less energy to cause a bit upset: the LET_{TH}, thus, reduces too with the technology scaling. Another effect of submicron technologies is that by having a denser layout it becomes more likely that a single particle upsets multiple bits at ones, as sensitive nodes are closer and less charge is needed to upset them. The impact of Multiple Bit Upsets (MBU) can be tested as well, even though it's highly dependent on the circuit layout in addition to the technology.

4.1.3 *Radiation hard electronics by design*

There are techniques which allow to design circuits that can resist to radiation damage, leading to *radiation-hard* ASICs. The two different effects described above need different solutions, as they cause fundamentally different problems; they can be faced at both a layout and a schematic level. Some of these techniques are detailed in Velazco [24]
TID effects cause a shift in the threshold voltage and a large increase of leakage current of MOS transistors. Care should thus be taken in using circuit architectures which can work with high channel leakage and which don't rely on precise threshold voltage values. This is unfortunately not always practical, as it can lead to a larger area or higher power consumption. A commonly used layout technique to reduce the problem is making sure the STI oxide doesn't touch both the ends of the channel of a transistor, forming a parasitic channel where leakage current could flow. In the case of NMOS transistors, this is possible by surrounding one of the n+ diffusion with gate oxide, drawing what's called an Enclosed Layout Transistor (ELT, see

figure 14).

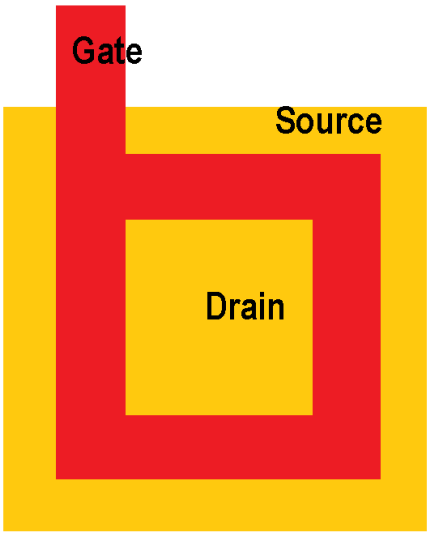


Figure 14: Layout of an Enclosed Layout Transistor.

Injected charge can still lead to leakage currents between different n+ diffusions at different potentials (between adjacent transistors). A p+ doped guard-ring can be drawn between ELT transistors to prevent this effect. ELT transistors have been commonly used in HEP environments, including CERN, where a commercial 0.25 µm technology has been used for the LHC experiments (see Anelli et al. [29]). ELT transistors, though, pose a number of issue to the designer: first of all there is a noticeable limitation in the W/L ratio that can be achieved with this geometry. There is also a problem with the transistor model, as the enclosed geometry is not usually recognized by standard EDA tools (although this can be worked around by writing custom rules) and the there can be discrepancies in the extracted parasitics of the ELT compared to the one of a standard transistor of the same equivalent size. The lack of symmetry can also be a problem for some layout-sensitive circuits.
Single event upsets need a different approach. On a circuit level, cells that are more robust to injected charge in sensitive nodes can be designed. The simplest way of achieving this is to increase the capacitance of the sensitive nodes, in order to increase the minimum charge needed to upset the stored value. By accepting an area and power consumption penalty, the error rate can be decreased by more than one order of magnitude. Another solution is to create structures that have multiple nodes that must be upset at once to change the stored value: if the two nodes are spaced out in the layout, the probability of both of them being hit by a particle at the same time is drastically reduced. An example of this is the DICE cell, as described in Naseer [30]. Another approach to protect the circuit from SEU is to add redundancy

to the stored information. This can be done both by triplicating the number of cells storing the information and then having a voter, so that a single upset will not corrupt the data (this technique is typically called Triple Modular Redundancy, TMR, as detailed in figure 15). The same effect can be achieved by storing the data using an error detection and correction code. For both these techniques, layout plays also an important role, as redundant elements should not be placed close to each other, to avoid the possibility of a single particle causing upsets in more than one memory cell at once.

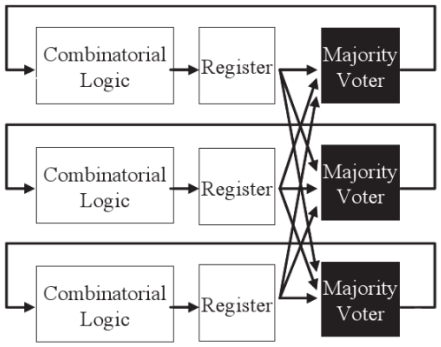
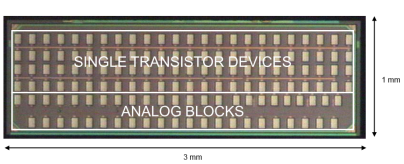


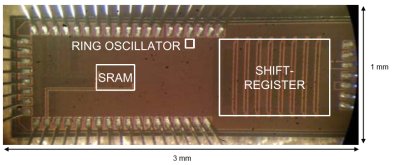
Figure 15: SEU hardening with redundancy: an example of a triplicated cell (figure from Velazco [24]).

4.2 DESIGN OF A TEST CHIP FOR RADIATION MEASUREMENTS

In order to validate the radiation performance of the chosen 65 nm technology, two test chips have been designed (visible in figure 16), containing structures specifically aimed at performing radiation damage measurements. The test chips contain single transistor devices as well as digital and analog prototype circuits. One test chip contains digital blocks and another one analog structures and single transistor devices. The chosen metal stack for the reticle consists of 6 metals (thin M1-M4, thick M5-M6). Both chips are 3×1 mm^2 in size.



(a) Chip with single transistors and analog blocks.



(b) Chip with 64-kbit shift-register, 56-kbit SRAM and 1025 elements ring-oscillator.

Figure 16: The two test chips.

4.2.1 *TID test structures*

In order to check how device performances would change with radiation exposure, an array of single devices were designed in one of the chips, to be tested individually. Each transistor had all its terminals connected to external pads, in order to be biased and fully characterized. Each transistor has a guardring around it (p+ for NMOSFETs, n+ in nwell for PMOSFETs).
The single transistor devices were designed as four columns of 24 pads each,with of NMOS and PMOS transistors of two types: core transistors for 1.2 V applications, I/O transistors with thicker oxide (or dual-gate, DG) available for applications using 2.5 V supply. Several transistor sizes were chosen for core transistors: an array with length L=60 nm and width W from 120 nm to 1 µm, two transistors with W=10 µm (L=1 and 10 µm). The sizes were chosen to be representative of the ones used in pixel designs. All these transistors were designed with the conventional regular layout. Additionally an annular ELT transistor with the minimum allowed size of 1480×60 nm was included together with a series of Field Oxide Transistors (FOXFETs) whose source and drain were either n+ diffusions or n-wells. A smaller set was chosen for I/O transistors (W×L = 400×280, 800×280, 2000×280 nm, and 10×1, 10×10 µm). The large range of sizes for the transistors are meant to be able to analyze the dependence of radiation damage effects on the W and L of the devices. It should be pointed out that ELT transistors violate DRC rules in this particular technology, as non-rectangular gates are not allowed (in this case, the DRC rule was waived). Despite this, the transistors in the test chip worked normally.
In addition to transistors, a few other devices were added to the test chip, including a p+ diffusion diode, a polysilicon resistor and a standard ESD protection structure.

4.2.2 *Digital test structures*

The digital prototype circuits are namely a 64-kbit shift-register, a 56-kbit SRAM, and a 1025-elements ring-oscillator. All digital blocks were assembled with the IP available from the foundry, including libraries of standard cells and I/O pads, and an SRAM compiler. All the test structures are not designed to be rad-hard (they have no triplication or any other redundancy), in order to be able to correctly measure the cross-section of the circuit correctly.
The purpose of these structures are both testing the degradation in performances due to TID (by measuring the oscillation frequency of the ring oscillator and the power consumption of all structures) and the impact of SEU errors in memory blocks, both flip-flops and SRAM.

4.3 TID AND SEU MEASUREMENTS RESULTS

4.3.1 *TID testing setup*

In order to examine the devices' response to TID, irradiation of the devices was performed using a calibrated 50-kV 3-kW X-ray generator (SEIFERT RP149, in figure 17) up to 200 Mrad (in SiO_2).

Figure 17: Bench with SEIFERT RP149 X-ray generator and probe station.

The ambient temperature during irradiation was 25° C. Measurements have been performed with the test chips mounted on a Karl-Suss PA200 probe station with a custom probe card installed within X-ray irradiation cabinet. The dose rate was about 41.5 krad/min. Devices were kept under worst-case bias during irradiation, therefore all terminals of the transistors were grounded, except the gate of the transistors, which was kept at the nominal power supply voltage V_{dd} (1.2 V for core transistors and 2.5 V for I/O transistors). The custom probe card has 32 probe tips (two columns of 16) which match the size and pitch of the pads in the test chip. A semiconductor parameter analyzer HP4145 was used to perform the static transistor measurements, applying and measuring currents and/or voltages. A Keithley 707 switching matrix connected the measuring channels of the HP4145, or the output of the voltage source, to the appropriate probe tip channels. All the instrumentation and X-ray generator was controlled by a PC running Labview, in particular the full measurement and irradiation of the transistors could be done sequentially and fully automatically. This unique test setup enabled to perform all the characterization without the need for any manipulation, hence ensuring that the thin gate oxide of the transistors under test was not damaged by electrostatic discharge.
MATLAB scripts were used to automatically extract relevant figures from the I-V measurements made with the HP4145 at different total doses and monitor the changes in the performances of the devices.

4.3.2 *TID test results*

The test results were presented in Bonacini et al. [31]. Results on single transistor measurements come from two identical sets of devices, as the extremely fragile MOSFETs were damaged very easily during the measurement process. The drains of the devices, in fact, had no protection against ESD damage, in order to achieve a better accuracy on the drain current measurement (for most transistors, the ESD structures would have carried far more current than the device under test). The switching of relays and other sources of noise in the test setup, especially due to long cables leading from the probe card to the measurement instrumentation, very often led to ESD damage. Moreover, in order to maximize the number of transistors to be tested the same device types shared the gate connections, so a single failure in the array meant the failure of the whole chip.
The threshold voltage of the core NMOS devices changes very little (up to 20 mV) in the explored TID range, as can be seen from figure 18. A local minimum for the threshold voltage is visible only for narrow devices around ~10 Mrads, most likely due to the different contribution of the two opposing effects: charges trapped in the oxide (decreasing the threshold voltage) and ones trapped in the $SiO_2 - Si$ (causing an increase in threshold voltage). The subtreshold slope does not change significantly.

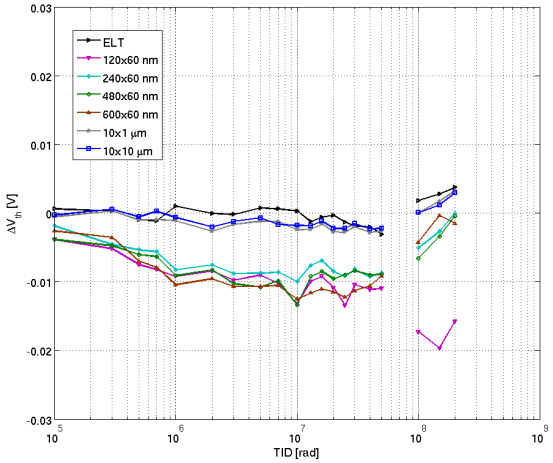


Figure 18: Threshold voltage shift of core NMOS devices. The curves are separated in two because the measurements come from two different samples, as the test setup didn't allow to perform enough measurements on the devices without damaging the gate oxide.

The curve for the ELT device demonstrates a high TID tolerance of the gate oxide of the studied technology, since its threshold voltage shift is practically unnoticeable, as it is for the increase in leakage

current, visible in figure 19. This behaviour is, similar, though, to standard transistors of similar sizes (W > 1 µm). In order to assure a limited (within an order of magnitude) increase in the leakage current in fact, no special layout is needed, but narrow transistors (W < 1 µm, for minimum L devices) should be avoided.

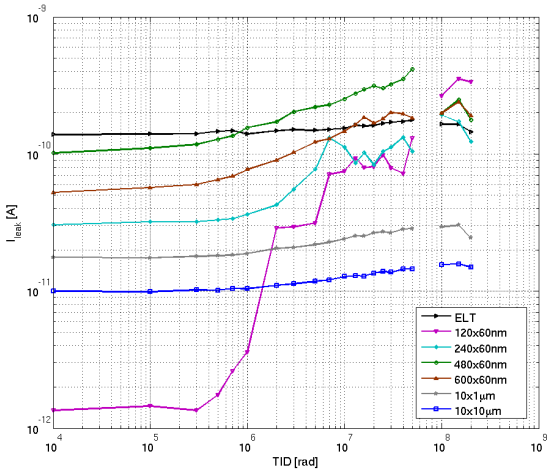


Figure 19: Variation of leakage current of core NMOS devices.

The increase in leakage current for the narrower devices is more than two orders of magnitude over 200 Mrads, but devices with a width larger than 360 nm show an increase of less than a factor 10. This result is an improvement over what was measured for the 130 nm technology which is being used for the LHC upgrades (see Faccio and Cervelli [22]).
The threshold voltage shift of PMOS devices is limited to 60 mV for the narrower devices, as shown in figure 20, and even less than 10 mV for wide devices (W>1 mm). The leakage current, in this case, monotonically decreases with radiation.

The maximum drive current of PMOS devices degrades with radiation by 50% for the narrowest device, but this is only partly due to the change in V_{th}. Most of the degradation comes instead from a reduction in the transconductance (g_m) in the strong inversion region as visible in figure 21. This value was calculated from a Id-Vg plot, extracting g_m in the point of maximum slope of the curve. The reduction depends again on the width of the transistor, the wider transistors having a smaller decrease. The degradation in drive current can influence the speed of digital logic. Measurements performed on NMOS devices, on the other hand, did not show any significant decrease of transconductance.

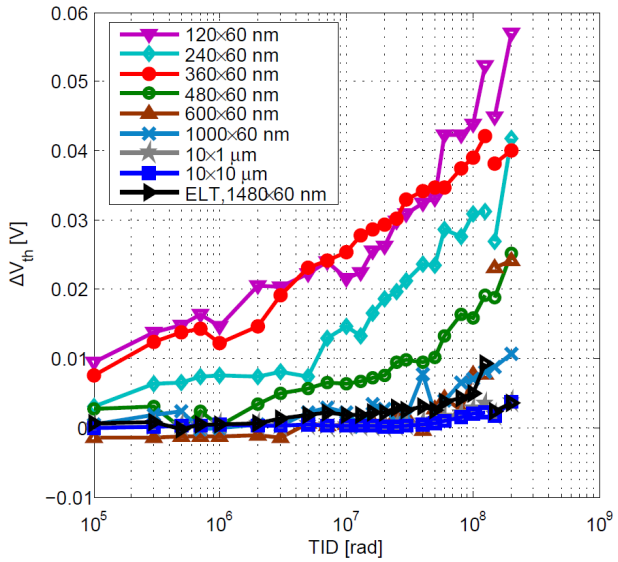


Figure 20: Threshold voltage shift of core PMOS devices.

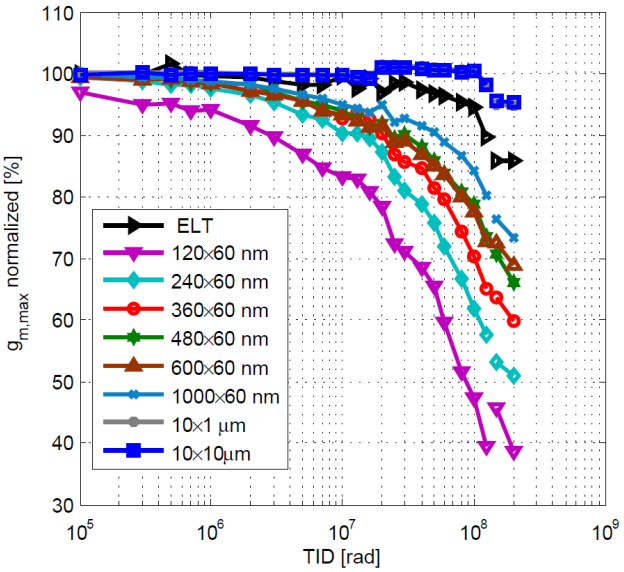


Figure 21: Maximum g_m (in strong inversion) for core PMOS devices, normalized to pre-rad measurements.

The test array included also a limited number of high-Vt devices. These devices are very important for applications such as pixel detectors, as their channel leakage current is very low (at the expense of speed) so it's possible to build circuits with very low power consumption using them. They were widely used in the CLICpix project, as explained in chapter 5 for their low leakage current for circuits that have to drive currents in the range of the nA and in which speed was not critical. The measurements showed a V_{th} shift which is comparable with similar-sized standard-Vt transistors. Results can be found for both NMOS and PMOS devices in figures 22 and 23.

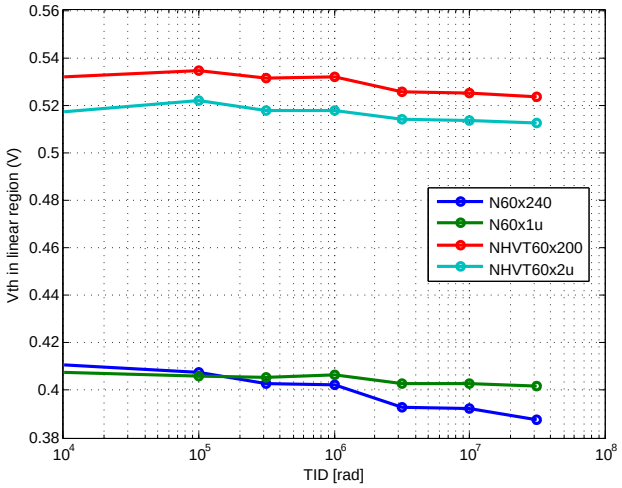


Figure 22: Threshold voltage shift of high-Vt NMOS devices ($L = 60nm$).

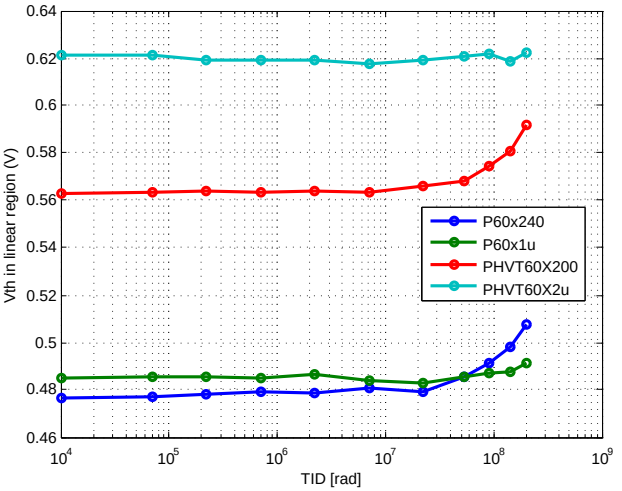


Figure 23: Threshold voltage shift of high-Vt PMOS devices ($L = 60nm$).

It should be noted that measurements on the high-Vt NMOS devices go up to 30 Mrads since all efforts trying to perform measurements up to 200 Mrads like other samples failed due to the limits of

the testing setup. Leakage current variation is also comparable with standard-Vt transistors. Normalized plots can be found in figures 24 and 25.

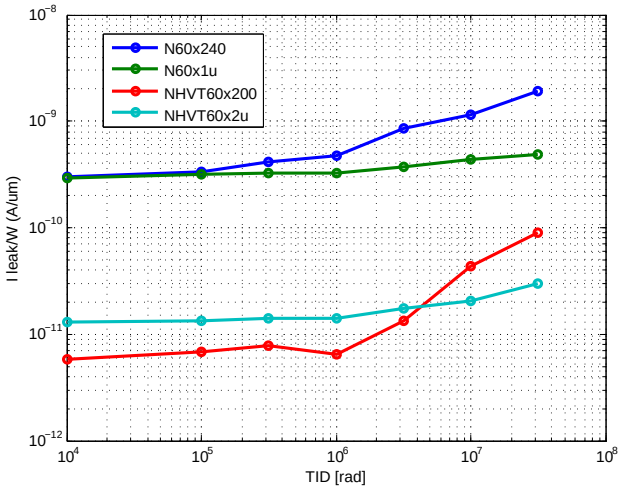


Figure 24: Variation of leakage current of high-Vt NMOS devices ($L = 60nm$).

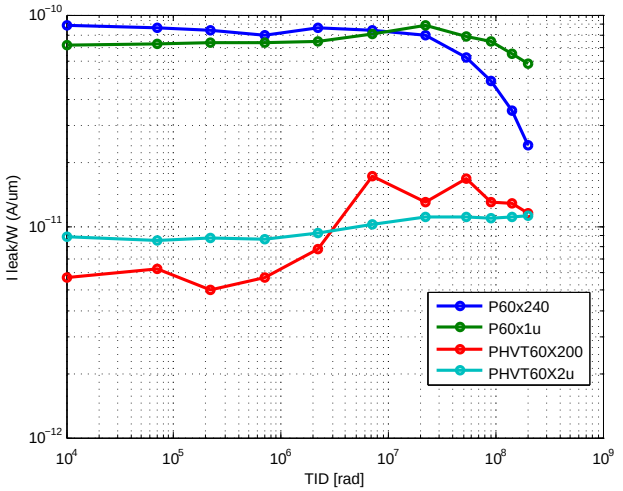


Figure 25: Variation of leakage current of high-Vt PMOS devices ($L = 60nm$).

Measurements on I/O (thick oxide) devices were also performed. The threshold voltage shift measured for I/O NMOS devices is within 200 mV, which is about 40% of the typical Vth of I/O transistors, and positive for all devices. The leakage current of the devices increases by 2 orders of magnitude. Most of the change happens at about 1 Mrad. This result suggests that ELT should be employed in I/O design. Once again, the performance of the I/O NMOS transistor in this technology is superior to the 130 nm previously referenced, where

I/O leakage peaks at 1 mA for similarly-sized devices. PMOS I/O transistors have a strong degradation in their performance, having a considerable shift of the threshold voltage: up to 800 mV, which accounts for 160% of the typical pre-rad value. This shift is more pronounced for narrow devices, which should be avoided in the design of critical blocks. The decrease in transconductance is also bigger compared to core devices, which could lead to slower I/O circuits. More detailed results on these devices, which were not used in the CLICpix prototype described in chapter 5, can be found in Bonacini et al. [31]. Digital structures were also tested for TID effects, using the same X-ray setup used for the array of single devices. Although functionality wasn't compromised by the radiation, a noticeable shift in speed and power consumption of the blocks were measured, as shown in figures 26 and 27. The very large increase in static power consumption of the SRAM block is probably due to the use of ultra-narrow transistors (80 nm wide). The annealing step restores the values to almost pre-rad measurements. The ring oscillator speed and power consumption decrease by 13% in the explored TID range. This is due to the reduced driving capabilities of the PMOS transistors, as discussed before. A safety margin for the timing checks should thus be included in the design when simulating circuits meant to work in radiation environment.

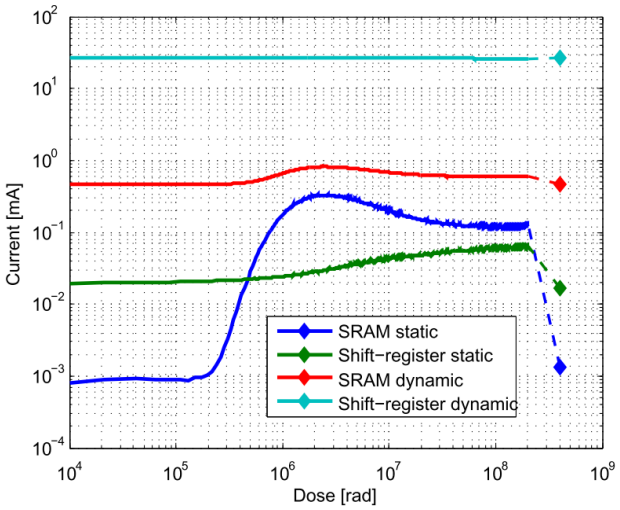


Figure 26: Variation of power consumption of memory blocks when irradiated. The last point is after annealing at 100 °C for one week.

The results suggest the possibility of using normal-layout transistors without special protection techniques against TID, even in heavy radiation environment, just by avoiding the usage of narrow transistors. The decrease in speed of the logic should also be taken into account. Appropriate safety margins must be included in timing calculations and simulations.

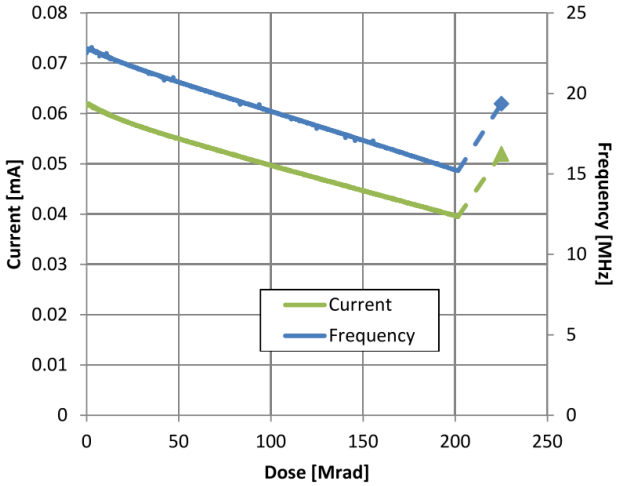


Figure 27: Variation of power consumption and oscillation frequency of the ring oscillator when irradiated. The last point is after annealing at 100 °C for one week.

4.3.3 *SEU testing setup*

Samples of the digital prototypes were irradiated in the heavy-ion beam facility in Louvain-La-Neuve (Belgium), in order to assess the SRAM and shift-register SEU tolerance. The same test board as for the X-ray irradiation was used for the SEU test, placed in a vacuum chamber and connected with a host computer outside the chamber (see figure 28). High-penetration ions were chosen for this test. Two kinds of tests were done for the shift-register: static and dynamic. Only the static test was run for the SRAM. In the static test the data retention of the SRAM and shift-register storage was studied. In the dynamic test, the robustness of the registers to SEUs when the clock is running was investigated. For the static test the procedure is as follows: a pattern is loaded in the shift-register and SRAM while the beam is off; the clock is stopped and all Device Under Test (DUT) input signals are frozen; the beam is turned on for a specific fluence and then turned off; the clock is run again and the output of the DUT is compared with the original (expected) one. For the dynamic test the procedure is: an indefinitely long configuration bitstream is loaded in the shift-register chain; at the same time the output bitstream is continuously compared with the original one; the beam is turned on for a specific fluence and then turned off; the clock is stopped. The SEU cross-sections are derived by dividing the total number of errors observed for each ion Linear Energy Transfer (LET) by the total fluence of the ion beam. A variety of ions were used, ranging from a LET of 3 to 20.4 $MeVcm^2/mg$. Additional other LET points are obtained by tilting the beam with respect to the chip surface (45° and 60°). The test was run at two different power supply settings, the nominal 1.2V and a reduced 0.9V, in order to evaluate the impact

of the reduced power supply on the cross-section, in the vision of using a reduced supply for non-critical digital blocks for low-power applications.

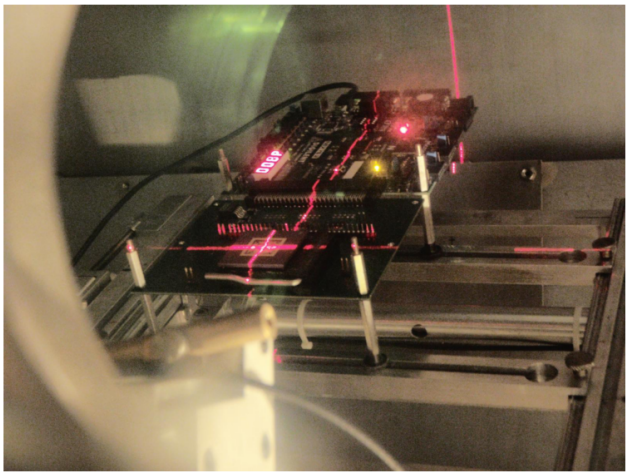


Figure 28: Test card in the vacuum chamber during targeting, ready for SEU testing.

4.3.4 *SEU test results*

The cross-section per bit of the SRAM and shift-register blocks are presented in figure 29 (for both 1.2 V and 0.9 V power supply). No substantial difference is visible between the static and dynamic tests run on the shift-register. The SRAM cell has an area about 13 times smaller with respect to the flip-flop used in the shift-register, therefore it has a lower cross-section (calculated as the the ratio of the number of upsets to the particle fluence), though not strictly proportional to the area. The LET threshold is lower than 1.1 $MeVcm^2/mg$ for both the digital cells.

Multiple Bit Upsets (MBUs) have a strong contribution to the cross-section in the SRAM sensitivity. Figure 30 shows the importance of simultaneous errors for the SRAM powered at 1.2 V. Most of these errors occur in adjacent cells along the direction of the n-wells present in the SRAM and which run all the way through the block. The shift-register also showed an evidence of particle hits on the clock tree, which caused thousands of errors simultaneously. Even though the cross-section of these events is too low to be plotted, designers should take this effect into account when building SEU-robust blocks.

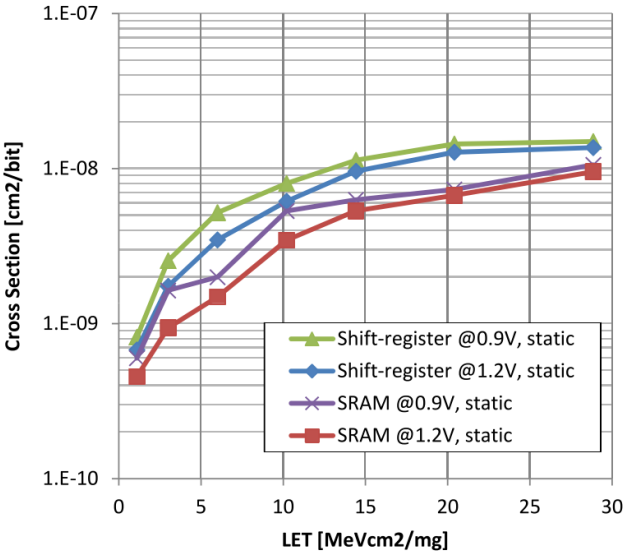


Figure 29: Cross-section at different LET values for digital blocks.

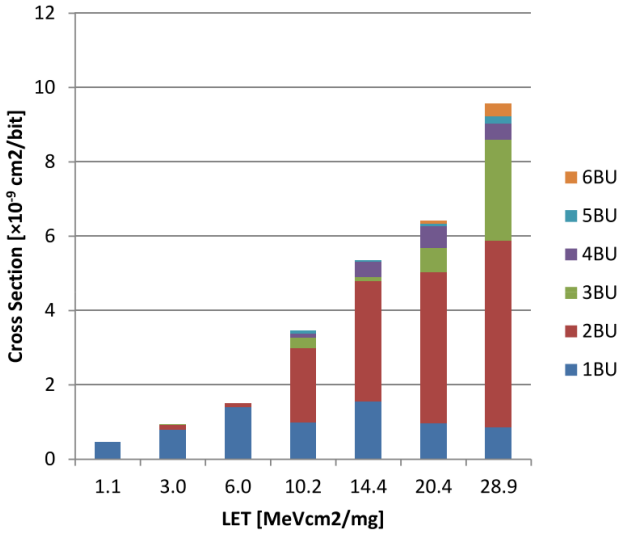


Figure 30: Multiple Bit Upset occurrence in the SRAM block (while powered at 1.2 V).

4.4 TESTS ON ANALOG FRONT-END BLOCKS

The analog blocks included in the test chip are common building blocks of the analog front-end of a hybrid pixel detector readout chip, designed for the technology under test. They were included to test both their radiation hardness and the general performances of these circuits, since they were the first attempts at designing a front-end with this technology. All blocks are connected to pads which allow to monitor their inputs and outputs as well as setting their bias voltages and currents. Since this design work was chronologically done before finalizing the specifications of CLICpix, some characteristics (dynamic range and area occupied for example) do not match the specifications described in chapter 5, although the circuits are very similar to their final versions.
A block diagram of the front-end can be found in figure 31.

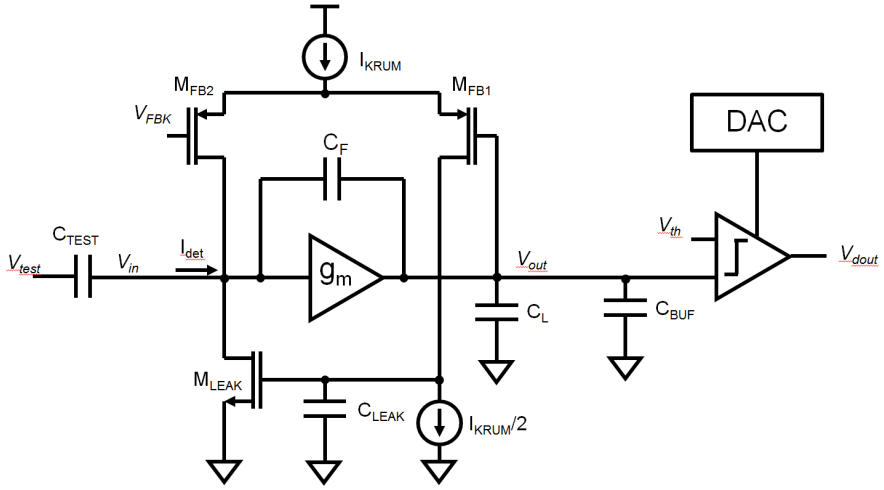


Figure 31: Simplified schematic of the Front-End included in the test chip. For testing purposes, the DAC is not connected to the discriminator in the test chip, but it is implemented as a standalone circuit.

The analog test chip contains three pixel front-ends, each one including a single ended preamplifier (with a Krummenacher feedback architecture, as described in Krummenacher [32]) and a discriminator. In order to test the pixel calibration system and to determine the matching performance of the technology, two DACs (Digital-to-Analog Converters) are also included with two different architectures: a binary-weighted 6-bit DAC and a sub-binary radix DAC with 6 bit resolution. They are not connected to the discriminator in this prototype, but the aim of the DACs is to digitally compensate the offset in the preamplifier and discriminator to allow pixel-to-pixel calibration. An explanation of how the circuit works can be found in the following subsections.

4.4.1 *Preamplifier block*

The input of the pixel is a Charge Sensitive Amplifier (CSA) that integrates the charge induced in the input pad. Its input is normally connected to the sensor. In this particular prototype chip no sensor was used, as due to having used a MPW submission, only single dice were available. A test capacitor is included to inject charge using a voltage pulse from an external pin. A detailed schematic of the circuit can be found in figure 32.

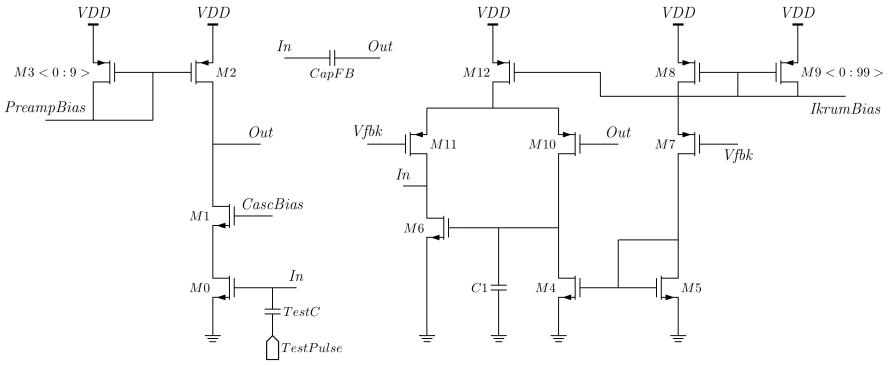


Figure 32: Schematic of the preamplifier with Krummenacher feedback network.

The original schematic of the preamplifier was proposed in Krummenacher [32]. It both integrates the charge on the feedback capacitor and compensates for DC currents on the input pin, acting as a band-pass filter. The circuit is made of a single-ended NMOS amplifier (with a cascode to increase its gain) and a feedback network (M4-12 in figure 32). M11 is connected to the input of the amplifier and acts as a first feedback loop together with M10 and M12. The effect of this loop is similar to adding a resistance of $2/g_m$ in parallel to the integration capacitance $CapFB$. Another feedback loop is made by M10, M4 and M6: the current flowing in M10 is integrated on the C1 capacitor and controls the gate voltage of M6. This loop is equivalent to an inductor in parallel to $CapFB$ and it makes so that a positive leakage current from the input flows in M6. If we call I_{krum} the biasing current flowing in M12, the circuit can compensate for a maximum negative leakage current of $I_{krum}/2$, while the positive value is not limited by the applied biasing.
Due to area constraints, this circuit is the only shaping of the signal, its output being directly connected to a fast discriminator, so the preamplifier noise must be studied carefully as there is no additional stage to reduce its bandwidth. The DC output of the preamplifier is set by the Vfbk biasing voltage, although is susceptible to variations due to transistor mismatch.

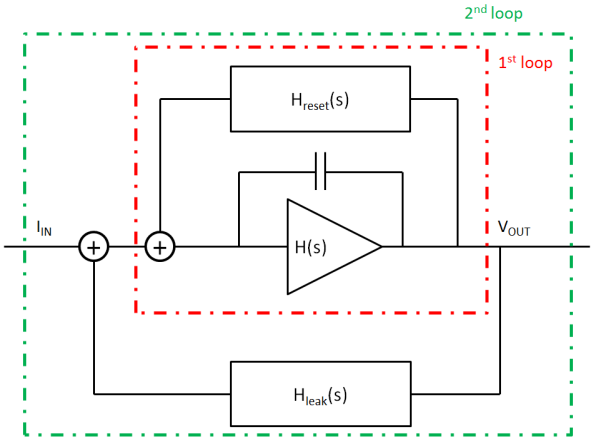


Figure 33: Diagram showing the two feedback loops of the preamplifier with Krummenacher network.

In-depth calculations on the transfer functions of the system can be found in Ballabriga et al. [33], a block diagram is found in figure 33. The CSA is implemented as a single-ended amplifier to minimize area and power consumption. The small signal transfer function of the CSA by itself has one pole and one zero, assuming a large output resistance of the OTA:

$$H(s) = \frac{V_{OUT}}{I_{IN}}(s) = -\frac{1}{sC_{FB}}\frac{1 - s\tau_z}{1 - s\tau_p} \tag{5}$$

$$\tau_z = \frac{C_{FB}}{g_m} \tag{6}$$

$$\tau_p = \frac{C_T}{g_m} \tag{7}$$

$$C_T = C_{FB}C_{IN} + C_{FB}C_{LOAD} + C_{IN}C_{LOAD} \tag{8}$$

where C_{FB} is the feedback capacitance, C_{IN} the input capacitance and C_{LOAD} the load capacitance. The transfer function of the reset network can be calculated for low frequencies as:

$$H_{reset}(s) = \frac{g_{m_{FB}}}{2} \tag{9}$$

where $g_{m_{FB}}$ is the transconductance of the M10-M11 transistors. Finally, the leakage compensation transfer function can be expressed as:

$$H_{leak}(s) = -\frac{g_{m_{FB}} \cdot g_{m_{LEAK}}}{2G_o(\frac{s \cdot C_L}{G_o} + 1)} \tag{10}$$

where $g_{m_{LEAK}}$ is the transconductance of the M6 transistor and G_o is the output conductance of the M4 transistor.

The transfer function of the first loop can be calculated as

$$H_{1st}(s) = \frac{H(s)}{1 - H(s)H_{reset}(s)} \backsimeq -\frac{g_{m_{FB}}}{2} \frac{1 - s\tau_z}{(1 - s\tau_{p1})(1 - s\tau_{p2})} \tag{11}$$

$$\tau_z = \frac{C_{FB}}{g_m} \tag{12}$$

$$\tau_{p1} = \frac{C_T}{g_m C_{FB}} \tag{13}$$

$$\tau_{p2} = \frac{2C_{FB}}{g_{m_{FB}}} \tag{14}$$

Finally, the transfer function of the second loop is:

$$\begin{aligned} H_{TOT}(s) &= \frac{H_{1st}(s)}{1 - H_{1st}(s)H_{leak}(s)} \backsimeq \\ &\backsimeq -\frac{2G_o}{g_{m_{leak}} g_{m_{FB}}} \frac{(1 - s\tau_{z1})(1 - s\tau_{z2})}{(1 - s\tau_{p1})(1 - s\tau_{p2})(1 - s\tau_{p3})} \end{aligned} \tag{15}$$

$$\tau_{z1} = \frac{C_{FB}}{g_m} \tag{16}$$

$$\tau_{z1} = -\frac{C_{leak}}{G_o} \tag{17}$$

$$\tau_{p1} = \frac{C_T}{g_m C_{FB}} \tag{18}$$

$$\tau_{p2} = \frac{2C_{FB}}{g_{m_{FB}}} \tag{19}$$

$$\tau_{p3} = \frac{C_{leak}}{g_{m_{LEAK}o}} \tag{20}$$

According to simulations the peaking time of the circuit is 50 ns and the minimum return to baseline time is 400ns for a 10ke- input. The bias point can be changed in order to work with both hole and electron collection. A summary of the amplifier simulated performances are shown in table 4. The test chip included three different channels:

one has regular MOSFETs, the second one has Enclosed Layout Transistors (ELT) for the NMOS part of the Krummenacher feedback to improve radiation hardness and the third one has regular transistors, segmented in length, which are intended to improve transistor matching (as proposed in Tuinhout et al. [34]).

Preamplifier gain	30 mV/ke^-
Preamplifier rise time	50 ns
Preamplifier current	1.5 µA
Feedback current	3 to 15 nA
Front-end non-linearity	5% at 16 ke^-
Equivalent Noise Charge	55 e^-
Preamplifier DC output variation	4.6 mV (130 e^-)

Table 4: Main characteristics of the simulated preamplifier with feedback network.

At the time of designing the test structures, radiation damage measurements had been performed on 130 nm and 250 nm technologies previously used in other projects. For such technologies, the most susceptible devices are the NMOS transistors, which is why the NMOS part of the feedback network has been designed also with a ELT layout. From subsequent measurements, the PMOS transistors were discovered to be more susceptible to radiation damage, significantly decreasing their transconductance.

4.4.2 *Discriminator block*

The output of the preamplifier, a shaped voltage pulse with an amplitude proportional to the input charge, is sent to the input of a discriminator, where it is compared with a fixed threshold. The output of this discriminator is used to identify whether the pixel was hit or not and is processed by the digital circuitry, which can possibly perform additional measurements such as quantifying the pulse energy (by its TOT) or giving it a time-stamp. The discriminator, built with a series of two amplifiers working in non-linear region (see figure 34), has a 3.3 ns delay time and its sensitivity is 0.6 mV (measured as the input voltage difference needed for the output to switch). Its output is connected to a digital inverter to increase the gain of the circuit. The power consumption of the discriminator is around 5 µW, although the biasing current can be changed, allowing to tune its power consumption and its delay.

Since both the discriminator threshold and the DC output of the preamplifier are subject to random variation due to transistor mis-

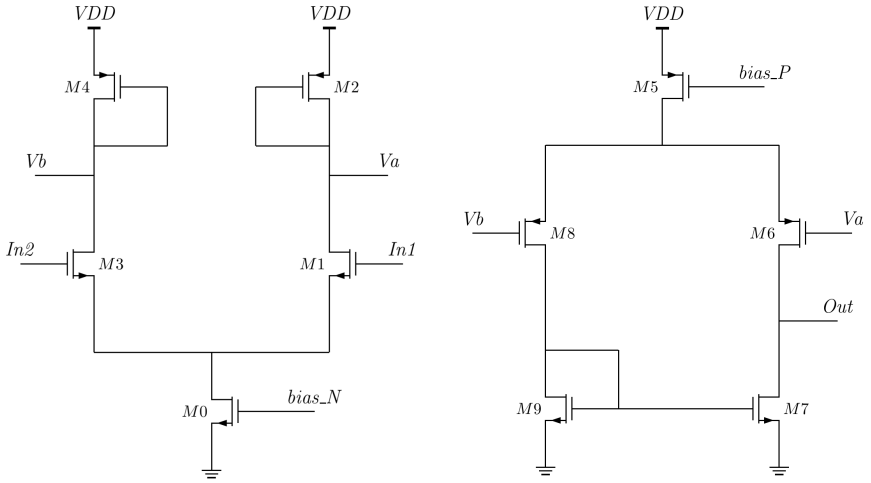


Figure 34: Schematic of the discriminator included in the test chip.

match, replicating this front-end in a pixel array leads to an effective pixel-to-pixel threshold mismatch (not every pixels having the same threshold), which distorts the measurement results. To solve this problem a calibration scheme must be introduced, to equalize the threshold of all the pixels. The topology used for the discriminator allows to easily implement a calibration system by means of a Digital-to-Analog Converter (DAC) with a differential current output connected to the nodes *Va* and *Vb*. The DAC outputs can be used to "unbalance" the differential pair of the first stage of the discriminator, changing the point at which the discriminator triggers, modifying its effective threshold. The dynamic range of the DAC must be chosen according to the expected mismatch of the front-end (a larger dynamic range is needed to compensate a larger mismatch). Two possible implementations for a calibration DAC are discussed in the following sub-sections. In the test-chip described in this chapter, these DACs were not connected to the discriminator, but included as separate circuits to be characterized independently.

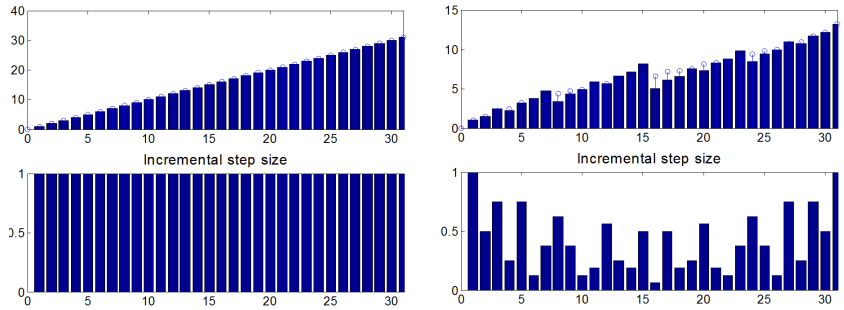
4.4.3 *Calibration DACs*

The output current of a generic DAC (using n bits) can be written as:

$$I = \sum_{i=1}^{n} b_i d_i \tag{21}$$

where $d_i \in [0, 1]$ and $b_i = r^{i-1}$.
Setting r = 2 is the traditional way of designing a DAC. However the ratio between these current sources (r, also called "radix") is a design choice and does not necessarily have to be set at 2 ("binary-radix"), but a smaller number can be used ("sub-binary-radix"). In binary-radix DACs the weight of each bit needs to be precisely set to avoid

missing codes, so their transistors must have a good matching. For this reason, they usually occupy an important circuit area, as the mismatch of transistor gets worse reducing their size. The area penalty is larger the more bits in the DAC, as the matching needs to be more accurate.
Sub-binary radix DACs introduce redundancy in their transfer function reducing the risk of missing codes at a reduced circuit area. This aspect is particularly important for projects in which the area constraints given by the small pixel size do not allow the use of large transistor to improve matching characteristics. A sub-binary-radix DAC has a non-linear and non-monotonic transfer function, as the one depicted in figure 35. A drawback of using this architecture is that the non-monothonic characteristic makes the calibration procedure more complicated, as each DAC has to be fully measured and its transfer function cannot be interpolated from a low number of measurements.

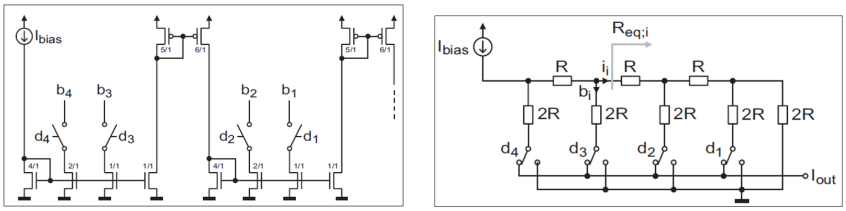


(a) Binary-radix DAC. (b) Sub-binary-radix DAC.

Figure 35: Ideal characteristic of a binary-radix DAC (left) and a sub-binary-radix one (right).

The binary DAC included in the test chip is designed using multiple proportional current sources, where the current source corresponding to each of the 6 bits provides double the current of the previous one. This structure is particularly sensitive to mismatch, as an error in one current source does not affect the others. In order to increase the number of bits, though, the area occupied increases exponentially. For the sub-binary radix DAC, more bits are needed to recover from the lost dynamic range due to the non-monotonic characteristic. A more compact topology was thus chosen (see Pastre and Kayal [35]), leveraging the redundancy introduced to compensate for a less mismatch-robust implementation, a ladder-like dividing network (both schematics can be found in figure 36).

This circuit takes an input current and provides several outputs (one per bit). The different impedance of the output nodes makes the current divide according to the desired function. Output switches control whether the current is driven to the direct or the complemen-



(a) Multiple current sources DAC. (b) Current-dividing ladder DAC.

Figure 36: Schematics of the two different DAC implementations included in the test chip (figures from Arthur van Roermund [36]).

tary output. This solution is well known in DAC design for the R/2R network, which is a fully passive network commonly used to build a binary-radix DAC. By using transistors in linear region instead of actual resistors and by changing slightly the topology of the net, it can be extended to work with an arbitrary radix. Two solutions were explored in order to design the topology of the net. The first one was to use the same implementation of the R/2R net, with a "dummy" branch after the last bit. By doing some calculations, a general rule for sizing the resistances (and therefore the transistors) can be extrapolated. Giving a nominal value of R to the resistances in the output branches (the 2R resistor of the R/2R net), the ones connecting them have to follow this rule:

$$R_i = \frac{r^i\,(r-1)}{\sum_{k=1}^{i} r^k + 2} \tag{22}$$

where r is the chosen radix and i is the bit number, from 1 (for the LSB) to n – 1. Using these values the output will be precisely divided in the desired way. This solution was simulated and found to work, but in the end was discarded because the use of different-sized transistors in the ladder led to inaccuracy in the resistances (because of parasitic resistances in addition to the transistor) and potential problems in the layout. The other possibility was to use a ladder using only unity-sized transistors and already studied in literature (see for example Arthur van Roermund [36]): the M/3M ladder (or R/3R, if using resistors). This net provides a theoretical radix of 1.77, so 8 bits are necessary to achieve the same dynamic range as a 6-bit binary radix DAC. The accepted mismatch for the impedance of the transistors, without letting any of the radices to go above 2 (and thus cause missing codes) is 13% r.m.s., so small transistors can be used. In addition, this ladder does not have a "dead" branch disconnected from the output, but uses its termination as any other bit (although with a different resistance). In order to use only one size for all transistors, the termination could not be set to the desired value of 2.3R, but was set to 2 instead. This leads to a slight inaccuracy for the first radices (especially the second one can be as low as ~1.4), but simulations confirmed that this behavior can be accepted without reducing the

system performance. According to simulations, the two DACs have similar integral non-linearity, but the difference in area is very large, with the sub-binary radix DAC being a fraction of the binary radix one, as shown in figure 37. The area gain is on the other hand reduced for DACs with fewer bits and it comes at the cost of a more complicated calibration procedure due to the non-monotonic characteristic. A measured DAC transfer function can be found in figure 38.

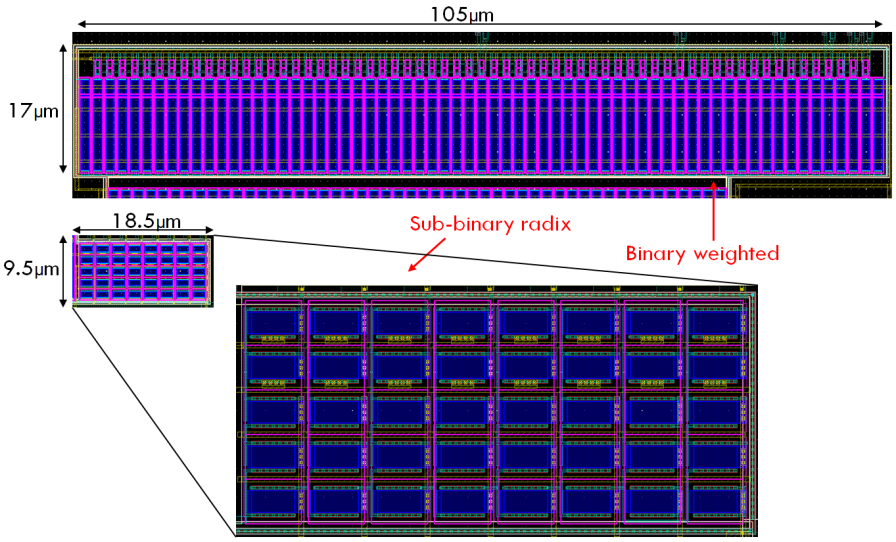


Figure 37: Layout of the two DACs included in the test chip (to scale).

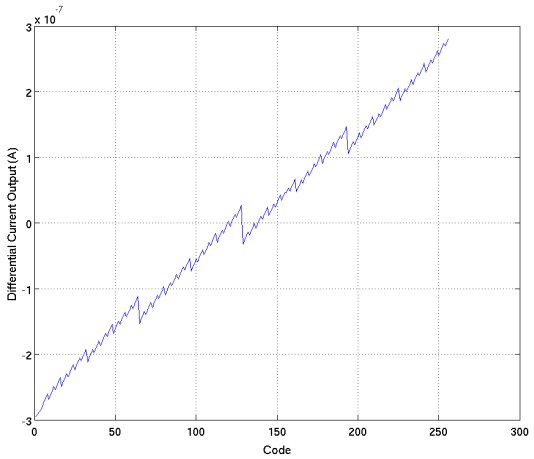
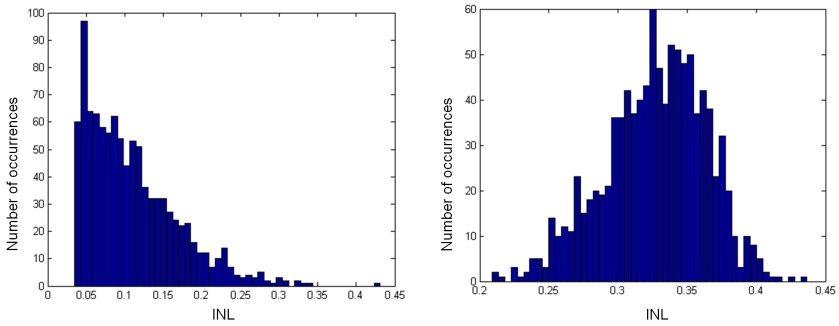


Figure 38: Measured trans-characteristic of the sub-binary radix DAC.

Monte Carlo simulations were performed to calculate the Integral Non-Linearity (INL) of the DACs and make sure INL < 0.5 for every code. The results are shown in figure 39. The two plots have different shapes, with the binary-radix DAC having a maximum for INL values close to zero, while the sub-binary radix one has a peak for bigger values. Due to the non-linear characteristic, even in an ideal case (no

transistor mismatch), the sub-binary radix DAC would have an INL greater than zero.

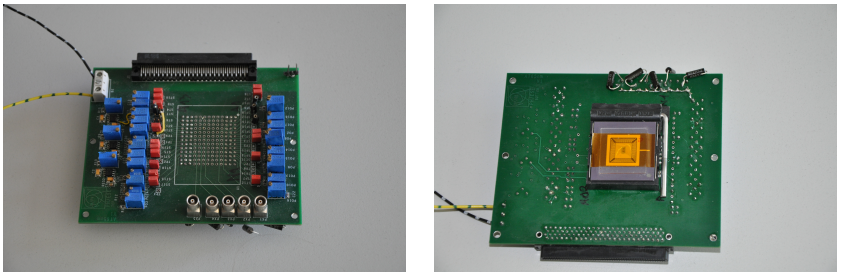


(a) Binary-radix DAC. (b) Sub-binary radix DAC.

Figure 39: INL of the two different DAC implementations, obtained with a 1000 points Monte Carlo simulation.

4.4.4 *Testing setup*

The test chip containing the analog blocks was mounted in a PGA100 carrier and an ad-hoc test board was developed containing voltage regulators and trimmers to control the biasing of the various circuits (shown in figure 40). The chip was connected through a socket, in order to be able to change it easily in case of radiation damage. All the measurements were carried out using standard lab equipment (digital multimeters, oscilloscopes, power supplies) being controlled by a PC via GPIB interface. A C layer was developed to communicate using the GPIB standard and was then called from MATLAB scripts which were used to automate the measurements. The chip was also irradiated using the same X-Ray setup described in the previous section to check the variation in performances with respect to the TID.



(a) Front view. (b) Back view.

Figure 40: Test board built to perform measurements on the analog blocks in the test chip

In order to test the DACs, a shift register containing their digital input was included in the chip, along with pads to be able to program it. This register was programmed from the PC using a Xilinx Spartan3 FPGA test board, implementing a USB interface to be controlled by the PC (as with the GPIB, from a MATLAB script).

4.4.5 *Measurement results*

Functionality tests were performed on each analog structure. The front-end (preamplifier, reset network and discriminator) was tested by using a pulse generator connected with a test capacitor to reproduce an incoming event (see figure 41) by injecting a controlled amount of charge in the preamplifier.

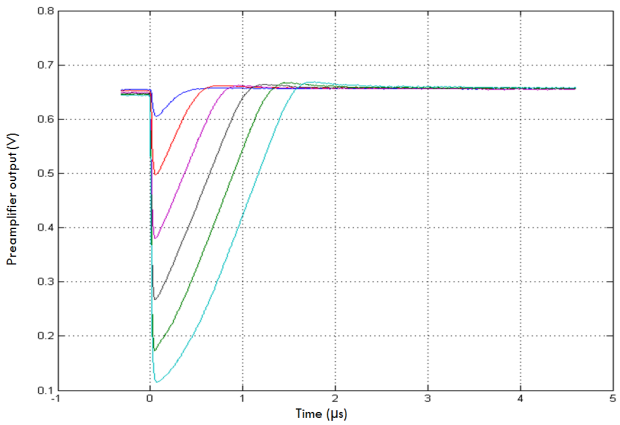


Figure 41: Output of the preamplifier for different test pulses, from 1 to 20 ke^-

Measurements were performed using various bias points to determine the relationship between the different parameters and the main figure of merit of the front-end. The behavior of the preamplifiers was found to match closely the simulations and all three implementations showed very similar performances, within the variation expected process variations. The measured rise time is ~65 ns and the preamplifier shows an output linearity < 5% for both electron and hole collection up to 15 ke^- (shown in figure 42). The linearity of Time over Threshold goes beyond 20 ke^-. This is the maximum tested value; higher values were not tested because ESD protection prevented to inject bigger charges through the test capacitance. The TOT count is linear up to much high energies, even if the preamplifier saturates. This is due to the integration capacitance being discharged by a constant current. The gain corresponds to the simulated value of 30 mV/ke^- and is constant despite any change in the biasing scheme, as it's determined only by the value of an integrated feedback capacitor. The gain is subject to the uncertainty in the value of the feedback capaci-

tor. Being implemented as a metal-to-metal capacitor using the routing metals, this is mostly due to the uncertainty in the inter-metal dielectric thickness, which gives a tolerance of about ±20%.The fall time of the pulses matches the simulations as well, being inversely proportional to the feedback current in the Krummenacher network.

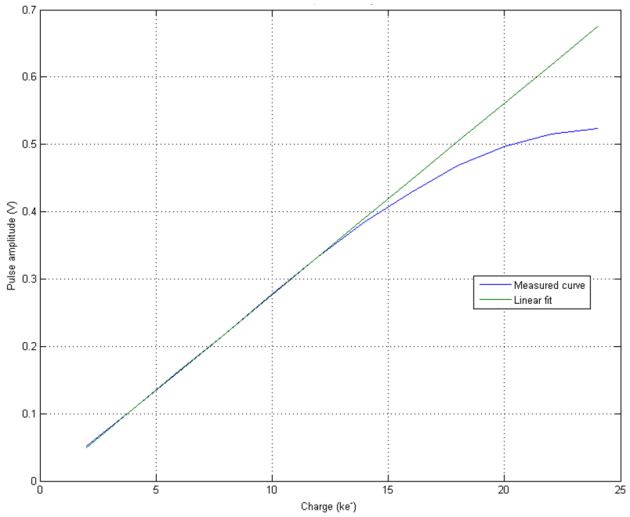


Figure 42: Linearity of the output amplitude of the preamplifier with respect to the input charge.

Noise measurements have been performed using the S-curve method (see figure 43). This algorithm consists in injecting a known number of pulses into the front-end and plotting how many of them make the discriminator switch for any given threshold. This plot produces a Gauss error function, as for thresholds very close to the pulse peak only some pulses are counted, depending on the system noise. From the error function, the noise parameters can be derived. The Equivalent Noise Charge (ENC) was found to be 60 electrons r.m.s. for the standard front-end and the one with segmented transistors, while it was slightly higher for the front-end using ELT transistors (70 electrons r.m.s.). This effect is due to the different size of ELT transistors, one of which contributes substantially to the noise because its drain terminal is directly connected to the input of the preamplifier. Results are, nevertheless, compatible with the simulated value of 55 electrons r.m.s.

Measurements on the DACs showed a non-linearity compatible with Monte Carlo simulations, although not enough statistics were collected to confirm the matching properties.
All analog structures were irradiated up to 200 Mrad as for the test devices. All measurements done before irradiation were repeated after each TID step, to monitor the behavior of the various figures of merit changing the total dose. Most of the measured parameters did

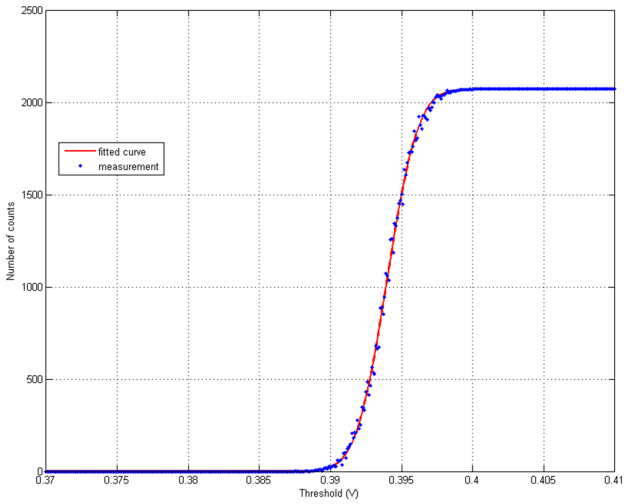


Figure 43: S-curve measurement used to calculate the Equivalent Noise Charge of the preamplifier.

not exhibit any noticeable change during the whole irradiation process. Gain in particular remained constant, as it depends on a metal-to-metal capacitor (the *CapFB* in figure 32), which is insensitive by radiation dose. The non-linearity of the circuit, the DC output voltage and the timing (pulse duration, rise time) of the front-end were practically unaffected as well. The only noticeable effect of radiation was an increase of the measured noise of the front-end implemented with standard transistors (up to 80 electrons at 200 Mrad).
A summary of measurements on the front-end can be found in table 5.

	Simulations	Measurements
Rise time	50 ns	65 ns
Gain	30 mV/ke^-	29.1 mV/ke^-
Non-Linearity	5% at 16ke^-	5% at 15ke^-
Equivalent Noise	55 e^-	60 e^- (70 e^- for ELT)

Table 5: Main characteristics of the front-end, measured on the test chip.

The test results show that this technology can be used to design compact and low-power circuits suited for highly integrated pixel detectors. Most of the work done designing the test structures could be used as a starting point to design CLICpix, as described in the next chapter.

5 THE CLICPIX DESIGN

5.1 THE CLIC VERTEX DETECTOR SPECIFICATIONS

After the validation of the technology, having characterized its performances in terms of both radiation hardness and the feasibility of an analog front-end, the focus switched to designing a fully-featured prototype for the CLIC vertex detector, hereby called CLICpix (as explained in chapter 3). The designed prototype contains a 64 by 64 pixel matrix and it is a fully featured but smaller version of the final detector. The prototype can give a good indication on the performances of the system while keeping the prototyping cost low, as it could be submitted in a Multi Project Wafer (MPW).

5.1.1 *Final specifications of the analog front-end*

The requirements of the analog front-end are dictated mainly by the accuracy of the measurements needed by the CLIC experiment (see chapter 3). The experience acquired designing the front-end in the test chip (see chapter 4) has been used to define the target area of the circuit and thus the pixel size, as well as agreeing on feasible specifications (detailed in table 6).

Parameter	Unit	Value
Counter Depth (TOT)	bit	4
Counter Depth (TOA)	bit	4
Pixel size	µm	25 by 25
Prototype array size		64 by 64
Power consumption	$\mu W/cm^2$	< 50 (after power pulsing)
TOT dynamic range	ke^-	up to 40

Table 6: CLICpix prototype main pixel requirements.

By considering the layout used for the test structures described in chapter 4, a goal of 25 by 25 µm pixel size was decided. While the size of the test structures were bigger, a more careful layout was deemed sufficient to reduce their area enough (there were no hard area constraint in the TID test chip). The very small available standard cells for the digital circuits allowed also to use most of the pixel area for analog structures, which don't scale with the same factor as digital structures when using more downscaled technologies.

The power consumption constraint resulted directly in a specification on the current the analog front-end could use. Using a power pulsing scheme, considering the power consumption of the front-end to be negligible when it's turned off, the maximum power used by it is:

$$P_{pixelON} = \frac{P_{matrix}}{\#pixels} \cdot \frac{1}{DutyCycle} \tag{23}$$

where $DutyCycle$ is the ratio between the time the front-end is active and the bunch train period (20 ms). By estimating a T_{on} of ~25 µs (simulations discussed in 5.4 show that this number is realistic) the resulting maximum power consumption is 24 µW per pixel. Part of this budget must be assigned to the digital part, so assuming it can be divided in half, it leaves the analog front-end with 12 µW. So, while designing the analog circuits, a priority on minimizing their power consumption was considered, using 12 µW as an upper limit.
The dynamic range for the energy measurement specification comes from the expected energy deposited in the sensor by any given particle. The CLIC vertex detector is aimed at detecting Minimum Ionizing Particles (MIPs) in a 50 µm silicon sensor (see Aicheler et al. [12]). A MIP passing through the sensor will gradually lose energy forming electron-hole pairs. How many pairs are formed depends on the sensor material and on its thickness. Studies (such as Friedl [37]) have shown that the average energy deposited in silicon is around 80 $e^-/\mu m$. A plot showing the distribution of energy deposited in a 300 µm thick silicon sensor can be found in figure 44.

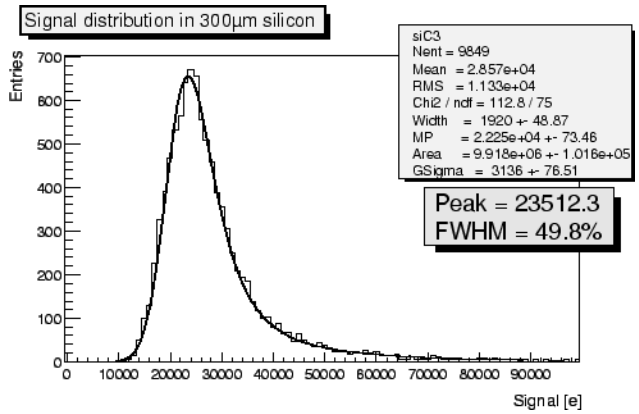


Figure 44: Distribution of energy deposited by a MIP in a silicon substrate, taken from Friedl [37].

In order to cover the whole energy distribution, a large safety factor was included in the choice. The average energy deposited in the sensor is in fact ~ 4 ke^- and the chosen maximum dynamic range 40 ke^-. Moreover, due to the charge not being collected by a single pixel ("charge sharing"), the resulting energy distribution will be distorted,

with a higher chance of lower energies due to pixels collecting only part of the total deposited charge. Due to these concerns, an effort was made in making the dynamic range for the TOT measurement modifiable on chip, using a programmable clock divider.
Other specifications were already discussed in 4.

5.1.2 *Timing and other global specifications*

The TOA resolution is attainable by choosing the right clock speed. A 100 MHz counting clock was chosen because a period of 10 ns is sufficient to have the required accuracy with a 4-bit counter. One limitation of the TOA measurement is Time Walk, i.e. the dependency of the time it takes for the preamplifier output to reach the threshold on the pulse energy (especially noticeable for low-energy pulses). Time walk will be discussed in section 5.4, along with relevant simulations.
Another important specification is the readout time. To minimize the number of connections required, a chip module is designed to be read out with a single data line, daisy-chaining its ASICs. This means that each ASIC has only a given amount of time available between consecutive bunch trains to be fully read out. Given the number of chips in a module (24, see chapter 3), each chip readout must take less than 800 µs. Reading out 8 bits of data per pixel from a full 512 by 512 pixels array means that the data rate should be higher than 2.6 Gbps. This very high value, which would lead to a very high power consumption, led to the implementation of a compression scheme to reduce the amount of data to be read during every readout phase, leveraging the low occupancy which is expected from simulations (2-5%). Various types of compressed data readout architectures were considered, as shown in section 5.3. A full frame data readout with zero compression was chosen.
One more requirement is the power-on and power-off times. In principle it's possible to turn the whole pixel matrix on and off at once, but it would require a powering scheme with a very large bandwidth to cope with the sudden increase in current consumption (see Blanchot and Fuentes [38]). For this reason a system which allowed to turn on the chip gradually was implemented. Details on its implementation can be found in section 5.6. Moreover, it is important to design the front-end so that it can acquire data as soon as possible after it has been turned on, in order to reduce the time the front-end needs to be powered (and thus, its average power consumption). Simulations shown in section 5.4 led to the conclusion that a wake-up time of ~15 µs was achievable.
A summary of the timing specifications can be found in table 7.

Parameter	Unit	Value
Time-stamp resolution	ns	10
Wake-up time	µs	~15
Readout time (for the full matrix)	µs	< 800
Pixel occupancy		2-5%

Table 7: Timing and global requirements.

5.2 CLICPIX ARCHITECTURE

In this section the general architecture of the chip will be discussed. Additional details on the various circuits implemented will be provided in sections 5.4, 5.5 and 5.6.

5.2.1 *Pixel architecture*

A block diagram of both the analog and digital circuits is detailed in figure 45.

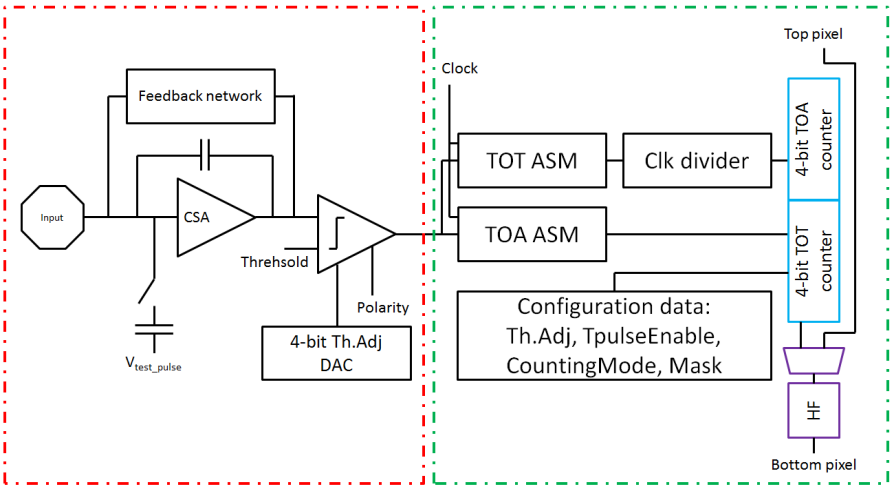


Figure 45: Block diagram of the pixel circuits in the CLICpix protoype. The red part is the analog front-end, the green part the digital one.

Each pixel implements an analog front-end which uses the same building blocks tested during the work described in chapter 4, modified to fulfill the updated specifications, especially in terms of area. The front-end block diagram is described in figure 31. Current pulses coming from the sensor (connected to a metal pad physically located on top of the pixel) or from a test capacitor are amplified and shaped by the preamplifier and feedback network and compared to a global threshold, locally adjusted with a 4-bit DAC to compensate for pixel-to-pixel threshold mismatch. The result of the comparison is used in-

side the pixel logic as an enable signal for the counting clocks of both the TOT and TOA counters. Local state machines are implemented in order to decide when to stop counting: for the TOA counter, it is tied to a global shutter signal, which is synchronously distributed to the whole pixel matrix and used as a timing reference. For the TOT measurement, the counting will stop as soon as the discriminator signal goes down. A diagram of how the counting works can be found in figure 11.The TOT counting clock frequency can also be divided by a programmable amount, to adjust the dynamic range to the measurement. The content of the 4-bit TOA and 4-bit TOT counters forms the 8-bit data acquired by each pixel with a valid hit. Pixels without incoming pulses that go over the selected threshold will have empty counters, as no counting clock will be enabled to the counters.
In order to minimize the pixel area and the amount of biasing lines, pixels are arranged in double columns, with the left pixels having the analog parts on their left and digital part on their right, while right pixels having the parts in the reverse order (see layout in figure 46). This allows the digital part of adjacent pixels to be merged (as discussed in 5.5 it was synthesized automatically) and share common resources, such as clock buffers. The layout of the analog part was not mirrored between the left and right pixels, in order to minimize the impact of this pixel arrangement on the matching of the analog structures. The layout of part of the regular structure in the pixel matrix can be seen in figure 46 (depicted horizontally for space reasons).

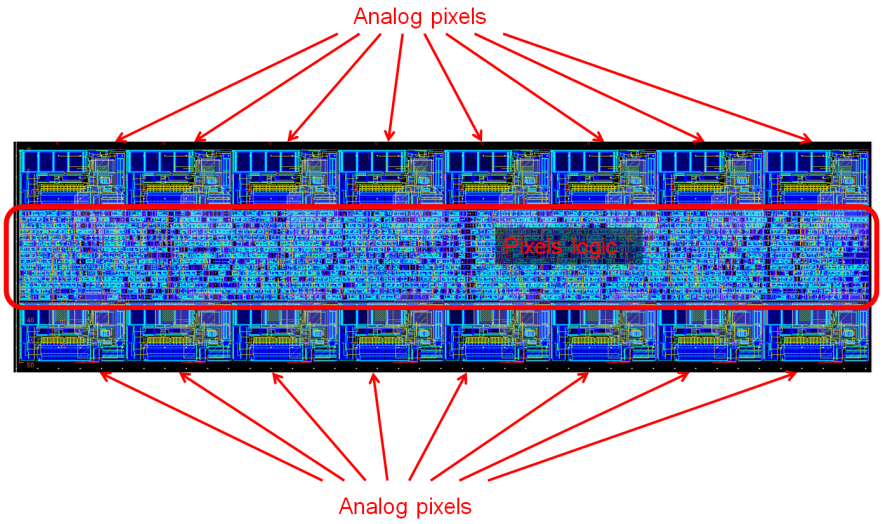


Figure 46: Layout of 16 pixels organized as a double column (depicted horizontally for space reasons).

5.2.2 *Global chip architecture*

When the chip is being read out, the two counters in the pixels are connected together as an 8-bit shift register. Pixels in every double column are connected in a "snake-like" ladder, from the top to the bottom (as shown in figure 47, rotated by 90 degrees)

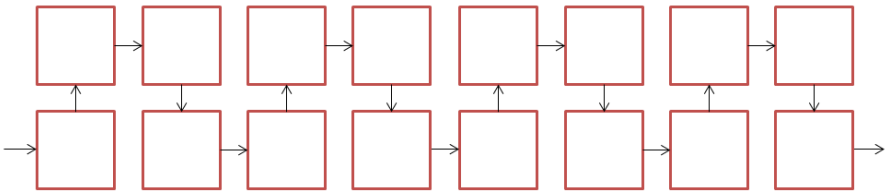


Figure 47: Scheme of pixel connections during readout in the regual structure shown in figure 46. The arrows indicate the fata transfer during chip readout. The periphery of the chip is located at the right hand side.

A block diagram of the whole chip can be found in figure 48.

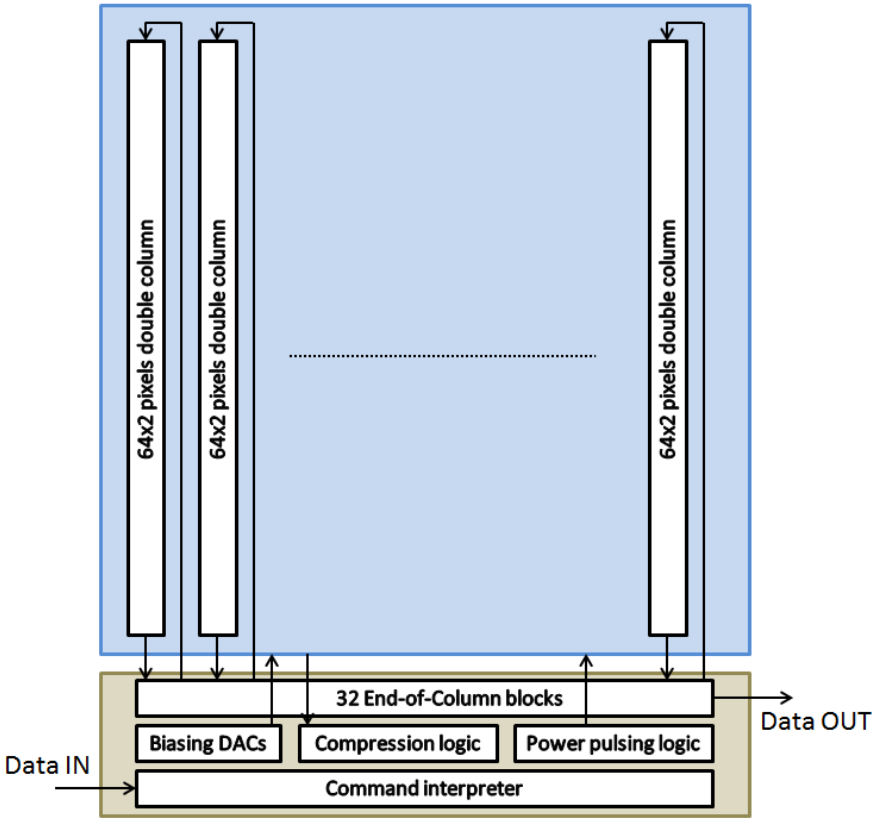


Figure 48: Block diagram of the CLICpix prototype, detailing the periphery area (in brown) and pixel matrix (blue).

Data is then shifted out one double column at a time, one bit per clock cycle, starting with the leftmost double column. Each pixel shifts the data to the next one making the counters work as a long shift register. When all the data from one column have been shifted out of it, the readout continues with the next column, until all the columns have been read out.

The chip uses two different clock signals: a 320 MHz clock for the readout and I/O and a 100 MHz clock for the acquisition. In the final CLICpix project an internal PLL will generate the two different clock frequencies, but the lack of an available IP from the foundry let to the choice of having two inputs for external clocks in the prototype chip. The I/O clock is distributed to the periphery logic using a perfectly-balanced clock tree (automatically generated with synthesis tools). A multiplexer placed at the bottom of each double pixel column allows to select which one of the two clocks will be distributed to the pixel matrix. The distribution of the clock to the pixels doesn't use the same kind of clock tree that is used in the periphery. Here the clock is sent from the bottom to the top of each double column through a series of buffers, one each 16 pixels, so the clock arrives with a slightly different delay to each pixel.

This delayed (not in phase) clock distribution has multiple effects on the pixel array. First of all it helps minimizing the total area, as it reduces the number of clock buffers that need to be implemented. Total power consumption is reduced by not having many clock buffers switching activities. Also, by not making all pixels logic gates switch at the same time, it reduces the instantaneous power consumption (it "averages" the power consumption over time instead of having peaks during clock transitions). The impact of receiving the clock with a different phase is negligible during data acquisition, as input pulses are not in-phase with the clock anyway and pixels don't need to communicate with each other. During data readout, the clock delay modifies the timing of the circuitry, as it means two consecutive flip-flops in the shift register formed between adjacent pixels can receive clocks with a slight delay. This effect has been simulated and it has been found to be negligible because the phase difference that can occur is small, being caused by at most a single clock buffer.

5.2.3 *Configuration and operation*

One of the main focus of the design was to make many aspects of the prototype configurable, so that it could be better characterized and more insight on the performances of its architecture and its technology could be learned. All the analog structures in the pixels have configurable biases, set using DACs placed in the periphery area, at the bottom of the matrix and sent to each pixel. Other global configuration options can be changed, such as enabling/disabling power

saving features or data compression (described in the next section).
Commands to the chip are sent using a single data line (with additional strobe and ready signals for synchronization purposes). A specific state machine in the periphery reads a 13-bit command, composed of a 5-bit op-code, specifying the operation to be performed, and an 8-bit data word.
Each pixel requires a configuration word, to set its equalization DAC code as well as other options (activate the pulse injection test capacitor or disable the pixel entirely). These configuration bits are stored locally in the pixels using latches. These latches have their inputs connected to the pixel counters, so the configuration routine includes writing in the counters and then "latching" the configuration data to store them. In order to program the pixels a special command can be issued to the chip, that shifts data into the counters, in 32-bit words (one bit per double column) from the top of the matrix. When all the counters of all pixels are filled with the desired configuration data, a store signal is sent to the latches in order to store the configuration data. Since the counters are now filled with data, the matrix needs to be readout once before starting an acquisition (this will reset the counters).

5.3 NOVEL FEATURES

5.3.1 *Power pulsing*

As discussed in section 5.1, the average power consumption of the chip must be lower than 50 mW/cm^2 and this justifies the use of a power pulsing technique. The power pulsing is implemented by having two DACs in the periphery for each state of the most power-consuming pixel analog blocks (the preamplifier and the discriminator). One DAC is used to control the nominal biasing used during the acquisition. The other one is designed to provide a stand by biasing current within a range of values several times lower than the first one. A multiplexer for each double column in the periphery can switch the biasing of the pixels from the nominal value to this "low-power" state. In this state, the analog circuits are not fully functional, but it is possible to wake up the pixel and start an acquisition in a very limited time (a few µs, more simulations can be found in section 5.4).
The power-on/power-off transition is managed by a specific state machine in the periphery that controls the multiplexers used to switch pixels to a different state, one column at a time. The amount of time this transition takes before the full chip is powered on or off is programmable using a configuration register. The chip is not designed to be powered on or off all at the same time in order not to cause abrupt changes in power consumption. This state machine is controlled with an external input signal in this prototype, to accurately characterize

the wake-up times.
The chip is designed to retain all its configuration during the low-power state, even if the analog power was physically disconnected. It is possible, so, to design the power supply to stop providing the analog supply voltage (as suggested in Blanchot and Fuentes [38]), although it would take more time for the system to wake up and start acquiring data.

5.3.2 *A data compression scheme*

In section 5.1 the need for data compression was also discussed. Most other chips in the Medipix family use a frame-based readout, meaning that after each acquisition the counters of the entire pixel matrix are read out, regardless of their content. In this case the bits of data to read out (and thus the readout time) are constant and independent from the pixel occupancy (percentage of pixels which have been hit during an acquisition). Timepix3 uses a different technique, a packet-based (or "zero-suppressed") readout. Pixels communicate with the periphery through a shared bus and send a "packet" of data after the pixel was hit (provided the bus is free). The data packet is formed of some synchronization data, the content of the pixel counters and an identifier of the pixel which sent it. In this scheme the amount of data to readout is linearly dependent on the occupancy, ideally going to zero if the entire chip wasn't hit by any particle.
Another way to reduce the amount of data being read out of the chip is to read the content of every pixel (like in a frame-based readout) but compressing the data coming from pixels which were not hit. This is done by adding a "flag" for every pixel, that is set to 1 if the counters are not empty and is left to 0 if the counters are empty. A multiplexer controlled by this flag allows the empty pixels to be "skipped" during the readout phase. This means that each pixel will have one additional bit that needs to be read out (9 bits instead of 8 in the case of CLICpix) but pixels without a valid hit will only use a single bit of data.
In order to further reduce the readout time, the zero-compression architecture can be improved by allowing to skip the readout of not only pixels, but also groups of pixels. In each double column, pixels are grouped in "super-pixels" of 16 (2 by 8), which , as described before, share the digital logic. Each super-pixel has one additional hit flag (which is simply the logic OR of the 16 pixel flags) that allows the super-pixel to be skipped during readout if none of its 16 pixels were hit. An additional hit flag per each double-column was also added in the periphery, to skip the readout of entire columns if no pixels in theme have valid data. The improvement of these two additional layers of compression are also shown in figure 49. Two global configuration bits allow different layers of compression to be turned on and

off, according to table 8.

Configuration bits	Pixel skipping	Super-pixel and column skipping
00	NO	NO
10	YES	NO
11	YES	YES

Table 8: Configuration of the compression scheme.

This approach is more efficient than the other two solutions unless the occupancy is very high or very low. For very high occupancies, the full-frame readout saves data because the overhead of one additional bit per pixel is not compensated by the data saved because of empty pixels. For very low occupancies, the packet-based readout is more efficient, as the used data goes asymptotically to zero, compared to the "pixel-skipping" technique that needs to read at least one bit per pixel. Simulations have been run using a realistic CLIC background to calculate the amount and position of hits in the chip and the zero-suppression with pixel skipping was found to be the best option for the range of expected occupancies (2% to 5%). A plot of the readout time (using a 320 MHz clock) at various occupancies can be found in figure 49.

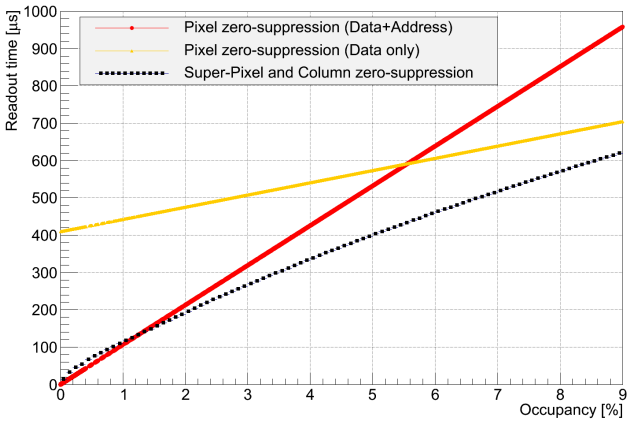


Figure 49: Comparison of various data compression solutions.

5.4 THE ANALOG FRONT-END

In this section the details about the design of the analog front-end will be given, including relevant simulation results.

5.4.1 *Preamplifier and feedback network*

As discussed before, the preamplifier uses the Krummenacher feedback architecture (see Krummenacher [32]). The schematic is the one described in figure 32. Compared to the circuit already described in chapter 4, the layout has been completely reworked, in order to fit, together with the other analog circuits, in the target area of 14 by 25 µm.
Feedback and test capacitors are implemented using vertical natural capacitors (using the parasitic lateral metal to metal capacitance) in order to have good accuracy at the expense of a larger area. The leakage compensation capacitor is implemented with a PMOS gate because of its high value of capacitance (~200 fF). NMOS transistors in the preamplifier are implemented as Deep N-Well (DNW) transistors to isolate the front-end from the substrate. No special layout technique was used to enhance the radiation hardness of the circuit, since it wasn't a requirement of the CLIC application. The vertical natural capacitor is considered by the CAD tool as a parasitic capacitance, as it's due to coupling of routing metals. All simulations, for this reason, have been done on an extracted view, where the EDA tool extracted parasitic capacitances of the layout. This functionality also helped with the layout design, by showing the coupling between different lines and other unwanted effects that could arise due to a wrong placement of devices. A MOS switch is used to disconnect the test capacitor from the input when the test pulse circuitry is disabled, to reduce its noise contribution to the front-end. Two additional switches are used to switch the voltage at one end of the test capacitor in order to inject charge into the system. The test capacitance is 10 fF (with a 20% uncertainty): this value was chosen to cover the desired energy dynamic range of the front-end by switching the voltage on its end by less than 1 V.
The CSA transfer function has been simulated to test for circuit stability, using a closed loop AC simulation. An amplitude plot at different I_{krum} values is shown in figure 50. The poles of the closed loop were calculated from the simulation and are shown in table 9, along with their quality factor[1].

1 For a pole $p = \alpha + \jmath\beta$ the Quality Factor is defined as $\frac{\sqrt{\alpha^2+\beta^2}}{2\alpha}$.

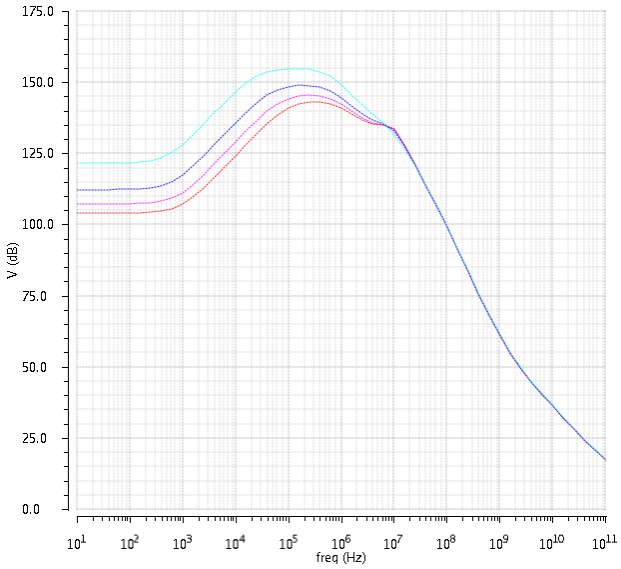


Figure 50: Amplitude plot of the closed loop response for I_{krum} values from 3 nA (cyan curve) to 12 nA (red curve).

Real	Imaginary	Quality factor
-1.05656e+01	0.00000e+00	5.00000e-01
-1.76987e+04	0.00000e+00	5.00000e-01
-2.62171e+06	0.00000e+00	5.00000e-01
-3.74135e+06	0.00000e+00	5.00000e-01
-4.31482e+08	0.00000e+00	5.00000e-01
-1.60688e+11	+/- 6.36177e+09	5.00392e-01

Table 9: Poles of the front-end closed loop response. The first 5 real poles and the first complex couple are shown. The stability is assured by the real part of the poles being negative and the complex poles having a low quality factor.

One of the biggest change compared to the circuit already tested is the gain, together with the choice of the nominal value for the feedback current I_{krum}. Using a clock divider to reduce the frequency of the TOT counting clock by a factor 8 and using a nominal I_{krum} of 8 nA, the 4-bit TOT counter is saturated in about 1.28 µs, corresponding to ~40 ke^-. Further adjustment of the I_{krum} value can be made in order to control the desired dynamic range. A value of 3.2 fF (with a 20% uncertainty) for the feedback capacitor gives a gain of ~44 mV/ke^- (although the front-end amplitude is quickly saturated even for the input charges of less than 10 ke^-, the energy measurement is done exclusively with the TOT counter).
The linearity of the TOT measurement was simulated when designing the front-end and is shown in figure 51. The figure plots the pulse length at the output of the preamplifier as a function of in-

put charge. The conversion to a digital value (the discriminator and counter blocks) doesn't introduce additional non-linearity. The non-linearity at 40 ke^- is smaller than 8.9%. Simulations show that non-linearity of the amplitude increases with the input charge, but the pulse length continues to monotonically increase up to an input charge of about 150 ke^-.

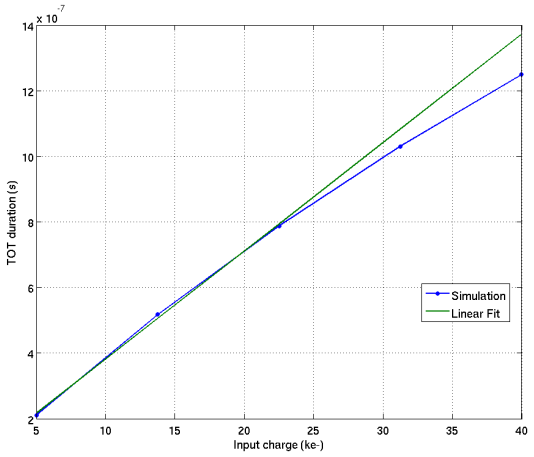


Figure 51: Linearity of the TOT measurement with respect of the input charge.

The dependency of the pulse length (and thus of the TOT measurement) from the value of I_{krum} was also investigated with simulations. I_{krum} is, in fact, the main way of controlling the speed with which the front-end output returns to the baseline voltage, so it can be used to adjust the dynamic range of the TOT measurement. This dependency is shown in figure 52 for an input charge of 5 ke^-.

The front-end is designed to work in both polarities, collecting negative or positive charges according to the type of sensor bonded to the chip (and its biasing). All plots in this chapter were done assuming hole collection. In order to simplify the logic design, the way the polarity switch is implemented is a global configuration bit that controls a set of switches in the front-end that connect the output of the preamplifier to the discriminator. In one polarity the output of the preamplifier is connected to the positive input of the discriminator and the global threshold is connected to the negative one, while in the other polarity connections are inverted. In this way the discriminator and the digital logic work the same regardless of the polarity. The preamplifier was simulated for both input polarities and was found to behave slightly differently in the two modes. Linearity and other characteristics were unchanged, but the TOT dynamic range was slightly different. This difference can be compensated by changing the I_{krum} value from the nominal 8 nA to 10 nA when working in

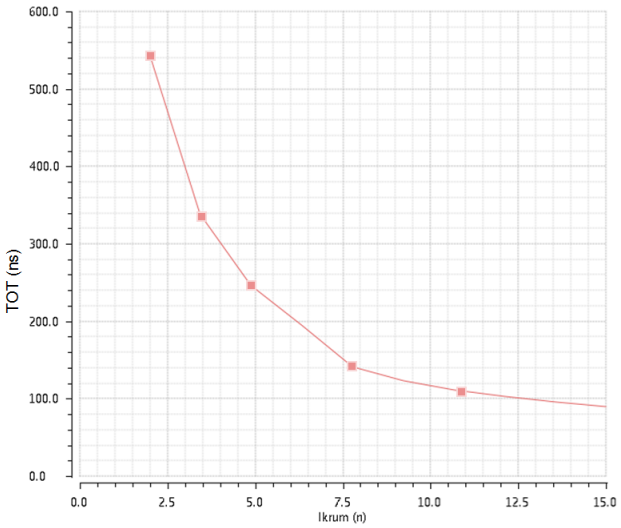


Figure 52: Simulated dependency of the TOT measurement on the I_{krum} value for a 5 ke^- input charge.

electron collection mode.
The noise performances of the front-end were also simulated. From an analytical point of view, the system can be seen as a two port network, with input noise generator (a parallel current generator and a series voltage generator) associated to it (as explained in Sansen [39]). The input series (voltage) generator refers to the preamplifier input transistor and can be expressed as the sum of thermal noise and flicker ($1/f$) noise:

$$\frac{\overline{V}^2_{series}}{\Delta f} = \frac{4kT\gamma}{g_m} + \frac{K_f}{C^2_{ox}WLf} \tag{24}$$

where the first term of the sum is the thermal noise and the second one is the flicker noise. k is the Boltzmann constant, T is the absolute temperature of the circuit (in Kelvin degrees), γ is a complex function of the technological parameters and bias conditions (it is higher for strong inversion bias, lower for weak inversion). K_f is a technological parameter, W and L are the width and length of the transistor, C_{ox} is the gate capacitance per area unit. The contribution of the bulk resistance was neglected because the transistor is isolated from the bulk via a Deep N-Well implant. The parallel noise generator takes into account the equivalent feedback loop resistance and the input current:

$$\frac{\overline{I}^2_{parallel}}{\Delta f} = \frac{4kT}{R_F} + 2qI_{DET} \tag{25}$$

where R_F is the equivalent feedback resistance of the feedback network, q is the electron charge and I_{DET} is the sensor leakage current.

In order to calculate the total integrated RMS output noise these contributions need to be integrated and multiplied by the transfer function of the system. Simulations show that the main contributors to the noise of the front-end are the three transistors connected to the input node (referring to figure 32, M1 series noise, M11 and M6 parallel noise).
In order to calculate the Equivalent Noise Charge (ENC, defined as the nput charge for which the Signal-to-Noise ratio is 1), the noise spectrum must be multiplied by the squared transfer function and integrated:

$$ENC^2 = \frac{\int_0^\infty \frac{\overline{V}^2_{noise}}{\Delta f} |H(2\pi jf)|^2 \, \mathrm{d}f}{\frac{V_{peak}}{Q}}, \tag{26}$$

where $H(2\pi jf)$ is the transfer function of the circuit, $\frac{\overline{V}^2_{noise}}{\Delta f}$ is the noise spectrum (it must be multiplied by the input impedance in case of parallel noise) and $\frac{V_{peak}}{Q}$ is the maximum amplitude of the output pulse divided by the input charge. Using the previous equations, the ENC can be expressed as:

$$ENC^2 = \frac{\int_0^\infty |H(2\pi jf)|^2 \left[\left(\frac{4kT\gamma}{g_m} + \frac{K_f}{C_{ox}^2 WLf} \right) \left(\frac{C_{IN}}{C_{FB}} \right)^2 + \frac{\frac{4kT}{R_F} + 2qI_{IN}}{(2\pi f)^2} \right] \mathrm{d}f}{\frac{1}{C_{FB}^2}} \tag{27}$$

By assuming the first pole of the transfer function (corresponding to the inverse of the peaking time) to be dominating the others, the contributions can be derived to be:

$$ENC^2_{thermal} \propto \frac{C_{IN}^2}{g_m \tau} \tag{28}$$

$$ENC^2_{flicker} \propto \frac{C_{IN}^2 K_f}{C_{ox}^2 WL} \tag{29}$$

while the contribution of the parallel generator is proportional to:

$$ENC^2_{parallel} \propto \left(\frac{4kT}{R_F} + 2qI_{DET} \right) \tau \tag{30}$$

where, τ is the peaking time. The dependency of noise from the peaking time is different for the different contributions, with the thermal noise decreasing, shot noise increasing and flicker noise being constant for increasing peaking times. It is possible to minimize the equivalent noise by using slow return to baseline times (low I_{krum}, to

increase R_F) and a small input capacitance C_{IN}. The detector capacitance is an inherent feature of the geometry of the sensor diodes and the bonding technique, as well as pixel size. Care has been taken in order to avoid increasing the input capacitance, by disconnecting the test capacitor with a switch when it is not used. A simulation showing the relation between the ENC and I_{krum} is shown in figure 53, while in figure 54 the relation between ENC and input capacitance.

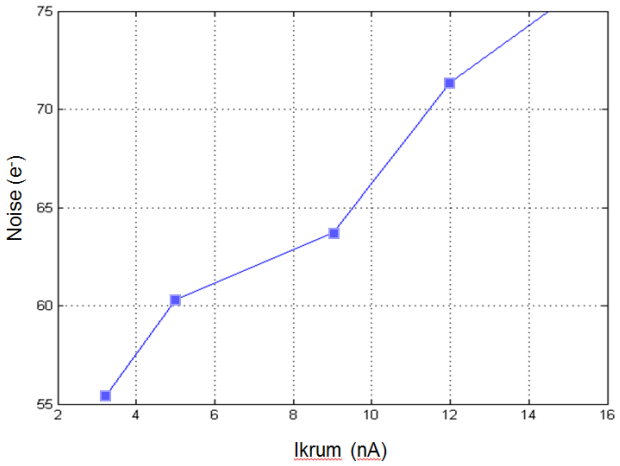


Figure 53: Front-end equivalent noise charge (in e^-) as a function of the I_{krum} feedback current.

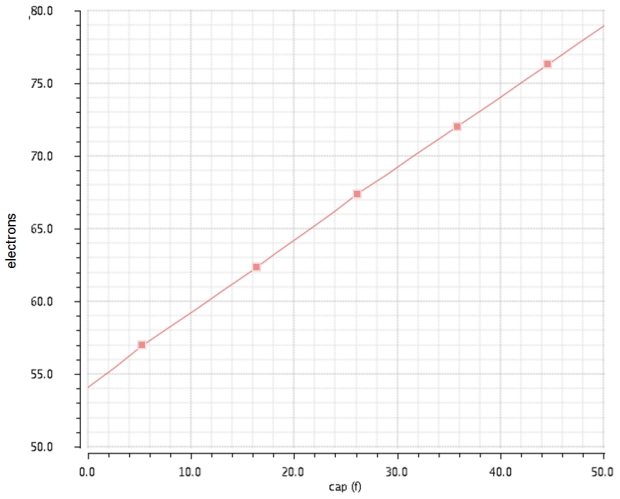


Figure 54: Front-end equivalent noise charge (in e^-) as a function of the input capacitance. The input capacitance here is intended as the additional capacitance due to the sensor and the bump bonding: the extracted capacitance of the metal input pad is already included in simulations.

The front-end noise can be used to estimate the jitter in TOA and TOT measurement. Approximating the pulse shape as triangular, the rms noise on the TOA measurement is

$$jitter_{TOA} \propto ENC \cdot \tau \tag{31}$$

According to the simulations, this value is ~300 ps (for a pulse whose amplitude is above the threshold by 3σ) , which is very low compared to the 2.9 ns due to the quantization error ($\frac{LSB}{\sqrt{12}}$). In the same way the TOT jitter is proportional to

$$jitter_{TOT} \propto \frac{ENC}{I_{KRUM}} \tag{32}$$

Simulations give a value of ~670 ps, compared to the 23 ns[2] due to the quantization error in nominal operating conditions.
The front-end was also simulated with different input charges to study the impact of "time walk" on the TOA measurement. "Time walk" is the effect of the finite rise time of the amplifier, which leads the discriminator to trigger earlier for larger input charges compared to smaller ones (that will trigger closer to the pulse peak). A simulation of the total delay of the discriminator output is shown in figure 55.

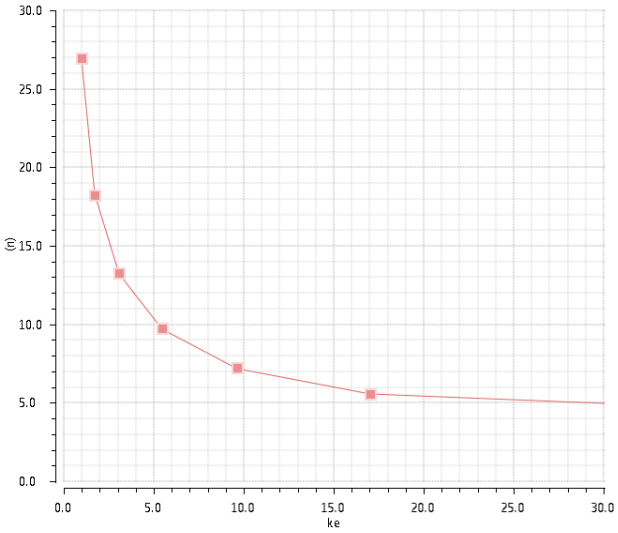


Figure 55: Front-end output delay ("time walk") as a function of the input charge.

The specification of 10 ns TOA accuracy can be met by using an "offline" correction for time walk: during data analysis, the delay can be corrected by knowing the TOT measurement. The nominal charge

2 This number is the standard deviation of the TOT measurement due to the quantization error, calculated as $\sigma = \frac{TOT_{LSB}}{\sqrt{12}}$, where TOT_{LSB} is 80 ns (due to a 10 ns clock period multiplied by 8 by the clock divider).

corresponding to a TOT measurement of 1 is 2.5 ke^-, so by correcting the delay for the hits with 1 or 2 counts, the desired accuracy can be reached.
The nominal bias current for the single-ended amplifier is 1.5 µA, but the current can be modified using an 8-bit DAC in the periphery. Details about the biasing system can be found in section 5.6. The preamplifier current is the main contribution to the power consumption of this block. The I_{krum} current is two orders of magnitude lower, so it's negligible compared to the preamplifier. For this reason, the power pulsing scheme doesn't implement a low-power bias for this current.
Monte Carlo simulations were also run to estimate the pixel-to-pixel mismatch of the DC output of the front-end, to correctly design the calibration circuit that compensates for it. DC simulations were done using the transistor mismatch models, also by taking into account the variation of critical transistors one at a time, to equalize their contribution to the total mismatch. The standard deviation of the DC voltage output was calculated to be ~5 mV, before any calibration. This number needs to be added quadratically to the mismatch of the discriminator stage, as explained below. In order to achieve this result, the NMOS transistors in the feedback network had to be large, so that their mismatch were reduced: in the resulting layout this part of the circuit takes around 30% of the available space.

5.4.2 *Discriminator*

The discriminator is a two stage open-loop amplifier, with an additional digital inverter connected to its output to increase its gain. The circuit schematic is the same as the one shown in figure 34, with the addition of a connection of the $Out1$ and $Out2$ nodes to a calibration DAC, described below. A plot of the transfer function of the circuit for different values of calibration current is shown in figure 56.

The discriminator output has a dynamic range of 0.2 V to 1 V with a 0.5 mV input swing. The delay is less than 5 ns (it is slightly faster for the zero to one transition than the opposite). The two stages were sized to minimize the effect threshold dispersion due to transistor mismatch, while being compatible with the available area. The threshold standard deviation due to the discriminator is ~5.2 mV, which combined to the dispersion due to the preamplifier gives a total standard deviation of 7.2 mV, corresponding to 160 e^-. The calibration system was thus designed to cover the $\pm 3\sigma$ range, as detailed below. The current consumption of the discriminator is nominally 4 µA, 1.5 µA for the first stage and 2.5 for the second one. The two bias currents can be adjusted using two 8-bit periphery DACs. The current in the first stage has a lower limit of about ~5 times the current that

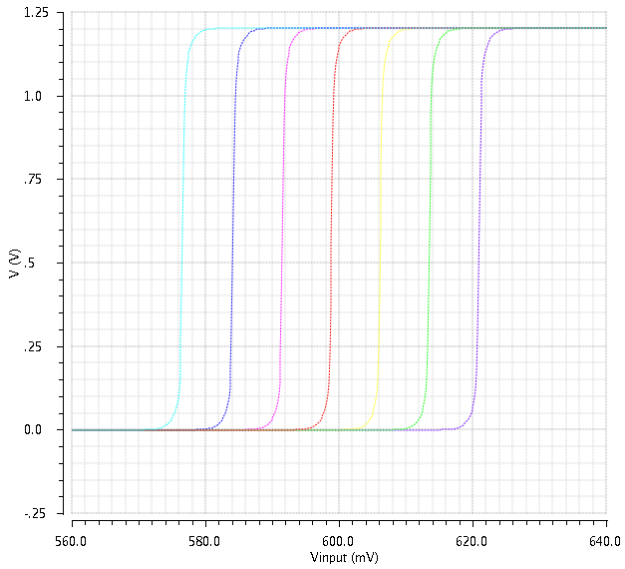


Figure 56: Discriminator transfer functions for different calibration currents (from -200 nA to 200 nA).

gets drawn by the calibration DAC. The DAC works, by unbalancing the two branches of the first stage to modify the effective threshold voltage. If the current drawn by the DAC in either branches is a considerable portion of the bias current, the effect on the threshold is not linear. The linearity between the calibration current and the effective threshold is shown in figure 57.

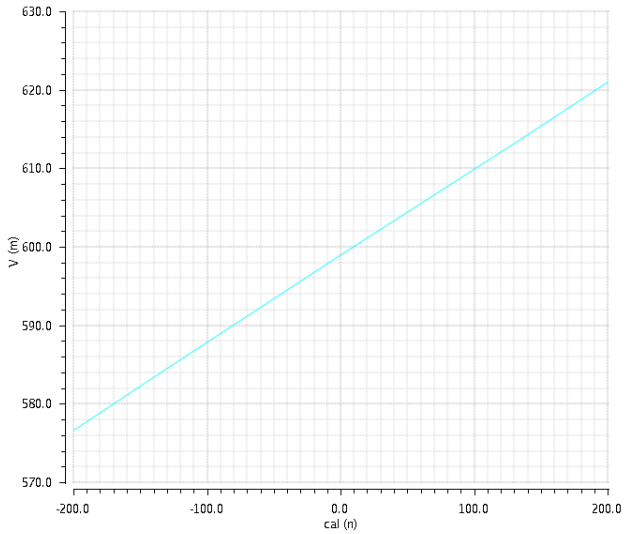


Figure 57: Linearity of the dependency of the effective threshold voltage of the discriminator on the calibration current in either branches of the first stage.

Simulations were also made to test the wake-up time of the whole front-end when going from a low-power state to an acquisition state.

During the low-power state the preamplifier and the discriminator are biased with very low quiescent current, but they still retain their functionality. This ensures that the inner circuit node voltages don't change significantly from one state to the other, so that during the wake-up they are already "preloaded" to the correct value. According to simulations, less than 15 µs are needed for the system to start acquiring again. The analog power supply could also be completely removed during the low-power state, without any loss in functionality of the digital part.

5.4.3 *Calibration DAC*

The threshold calibration DAC is a 4 bit binary weighted current DAC. Four bits are sufficient because, the needed dynamic range needs to compensate for ± 3 standard deviations of the threshold value due to pixel-to-pixel mismatch. This leads to a total dynamic range of about 43 mV, which corresponds of an LSB of 60 e^- for a 4-bit accuracy, which is comparable with the equivalent input noise.
Both architectures described in chapter 4 were studied and the binary radix one was chosen. The main advantage of a sub-binary radix DAC was area, but for a 4-bit converter the difference was negligible. Also, a sub-binary radix solution would have have needed more configuration latches to store the configuration code and a more complex implementation of the calibration procedure due to its non monotonic characteristic. The DAC uses multiple current mirrors in a scheme similar to figure 36 (right picture). The schematic is shown in figure 58. Each branch is made up by two transistors forming a cascoded current mirror (MX and MXB) and two switches (MXC and MXD) connected to the corresponding configuration bits, which divert the current of that branch into the output or the complementary output. Each branch is doubled compared to the previous one (physically replicated in multiple copies), so that it carries twice the current, making every bit of the code "binary weighted". An NMOS output stage with two high-swing cascoded current mirrors (not shown in figure 58) is included to make the voltage at the output nodes of the current mirrors of the DACs more stable, as the nodes to inject the current in have a large voltage swing, being the outputs of the first stage of the discriminator.

The current mirrors were sized to reduce their statistical mismatch in order to have a reduced integral non-linearity error of less than 1/2 LSB, while still fitting in the available area. Monte Carlo simulations were done to calculate the INL, resulting in an average INL of $11.8 \cdot 10^{-3}$ with a standard deviation of $7 \cdot 10^{-3}$.
In the nominal condition, the LSB is 12.5 nA, for a total full-scale current in either output of 200 nA. This value covers the desired dynamic

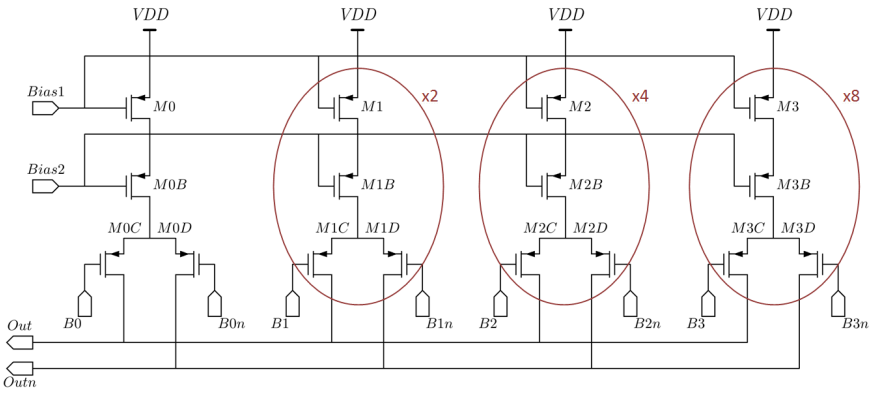


Figure 58: Schematic of the 4-bit threshold calibration DAC (without the NMOS output stage).

range of a threshold dispersion of 43 mV. The total power consumption is not dependent on the DAC code, as the sum of the currents in the two outputs is always equal to the full scale value. This value is negligible compared to the power used by the preamplifier and discriminator blocks, so no low-power state was implemented for this block.

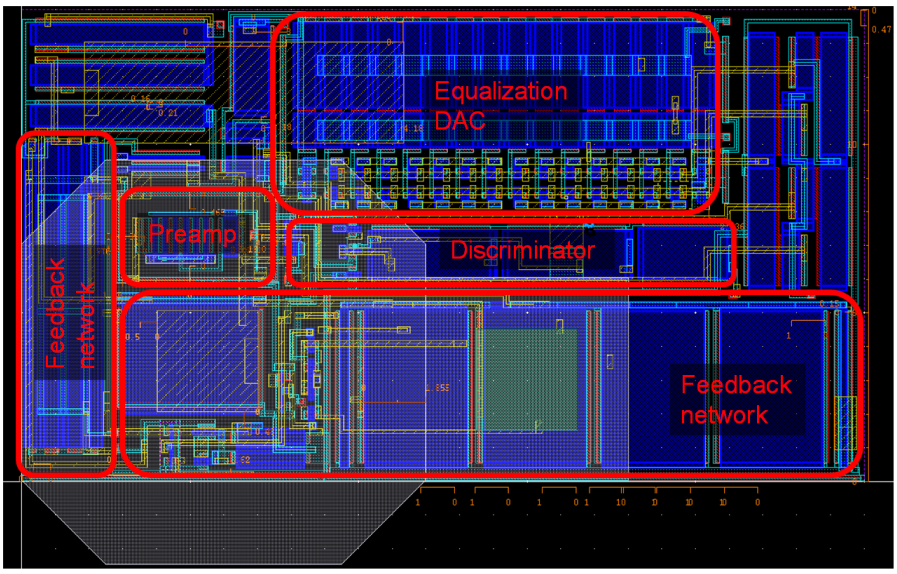


Figure 59: Layout of the analog part of the pixel. Physical size is 25 µm by 14 µm.

The layout of the entire analog part of the pixel can be found in figure 59.

5.5 PIXEL LOGIC

A block diagram of the pixel logic can be found in figure 45. All the logic was designed with an automatic synthesizer, place and router using standard IP cells from the foundry. Among the various different IP libraries available, the one with the finest pitch was used, as area was the main constraint of the design. Only cells using high-Vt (low-power) transistors were used, in order to reduce the pixel matrix power consumption due to static leakage currents, powered at 1.0 V. Simulations were performed to ensure that the circuit could meet timing requirements while using this power supply.

5.5.1 *Digital pixel blocks*

Each pixel is designed to measure TOT and TOA simultaneously. The output of the discriminator is used as an input to two state machines, which generate counting signals for two 4-bit counters. The counters are implemented as 4-bit linear feedback shift registers (LFSR) used to count TOT and TOA. The LFSR architecture was chosen because it uses less area than a standard binary or Gray counter and it is easier to switch between a counting state to a shifting state during readout. An off-chip conversion to a binary value is needed to analyze the data, but for 4-bit registers a look-up table with only 16 values is needed. An overflow control logic is included to stop counting when the stored value reach its maximum value).
An additional "hit flag" bit is added after the counters to allow data compression. It is set to 1 as soon as the pixel is hit, by checking the values of the first bits of the two counters. This value is latched at the beginning of the readout phase and it is used to control a multiplexer (also shown in figure 45) that will allow the pixel to be skipped during the readout, reading 1 bit instead of 9 (8 bits for the counters plus the flag bit) for pixels without a valid hit.
Each counter can work in two modes. During acquisition, it receives a counting clock for the TOT or TOA measurements. During readout, a multiplexer is used to connect both counters together as an 8-bit shift register, in order to shift the data down the column and off-chip.
An additional mode of operation is available, event counting. In this mode the TOA counter is used to count the number of times the discriminator output goes high during an acquisition. This mode is not used during a normal acquisition, but it can be used to equalize the pixels-to-pixel threshold mismatch, as it will be shown in chapter 6.
Latches for local configuration are included: 4 bits for the threshold adjustment DAC, one for pixel masking one to allow the injection of test pulses and one to configure the pixel for event counting mode. The mask bit makes the the pixel to be always skept during readout.

These latches have their inputs connected to the counter flip-flops. In order to program them, configuration data are shifted in the counters from the periphery and a global configure signal ("Load_conf", in the figure) is sent to latch their value.
The logic was simulated at a target speed of 320 MHz during the readout phases and a 100 MHz during acquisition phase. The switch between the two clock signals is performed in the periphery, as detailed further below.
A summary of the characteristics of the pixel logic can be found in table 10.

Cells	Foundry low-power standard cells IP
Area	25 µm by 11 µm
Acquired data	4 bits TOT + 4 bits TOA + 1 bit hit flag
Configuration latches	7 bits in total
Target Clock Speed	100 MHz (acquisition), 320 MHz (readout)
Acquisition Type	Sequential acquisition-readout

Table 10: Summary of the pixel logic characteristics.

5.5.2 *Superpixel architecture*

Pixels are grouped together in 2 by 8 pixel clusters (or "superpixels"). In the CLICpix demonstrator there are 8 superpixels per double column. Superpixels were synthesized together (flattening the design) to reduce their area. This was needed as the digital layout is very dense, with cells using ~95% of the available space. Synthesizing the different blocks hierarchically would not have met the area requirement. In addition to this, some blocks are shared between pixels in the same superpixel. One of such blocks is the clock buffer that propagates the clock signals to the various pixel blocks and to the next superpixel.
Another block implemented on a superpixel level is an additional hit flag flip-flop to allow superpixel skipping. It works in the same way as the hit flag in the pixel, by having its value stored in a latch at the beginning of the readout and this latch controlling a multiplexer which allow the superpixel to be skipped during the double column readout. In this case in order to check if a pixel in the superpixel was hit, the value of the hit flag is calculated with a logic OR among the 16 hit flags of individual pixels. A block diagram of this scheme can be found in figure 60.

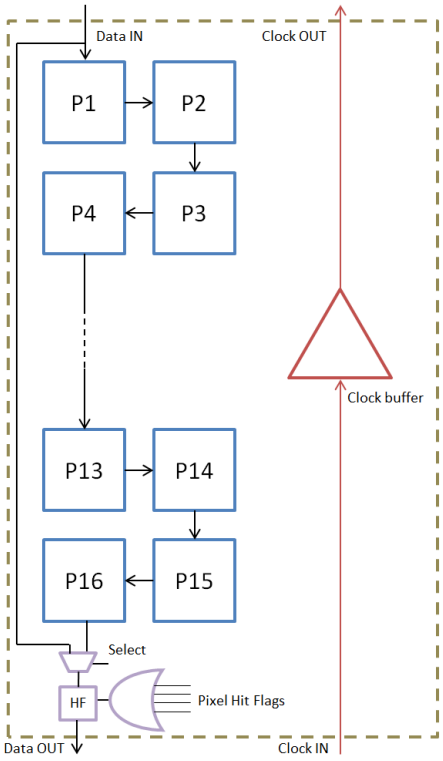


Figure 60: Block diagram of the superpixel logic (with the hit flag logic at the bottom).

5.5.3 *Asynchronous state machines*

Each pixel includes two state machines that control the generation of the counting signals for the TOT and TOA counters. These state machines are implemented as Asynchronous State Machines (ASM), which means they don't use memory elements to store the states. The current state is stored exploiting the delay of the logic gates and it is updated as soon as an input changes. Details on this particular kind of design technique can be found in Roth [40]. These state machines, unlike most of the logic, were manually designed on a gate level, using automatic tools only for the place and route phase. This choice allowed to define the states in order to minimize the switching activities of gates due to the clock when the counters are not enabled.
The two state machines generate an enable signal for a clock gating block, using the discriminator output and a global Shutter signal as inputs. The shutter is provided from outside of the chip using an I/O pad in this prototype, for debug purposes. In the final CLICpix chip, this signal can be generated in the chip periphery. The Shutter signal is sent to the pixel matrix using large buffers in the periphery, so that the signal is synchronous among all pixels and can be used as a global timing reference. The pixels are sensitive to hits arriving only when the Shutter signal is high. The TOA enable signal is set when the discriminator output goes high and it is reset to zero when

the shutter goes down. In this way only the TOA counters associated with the pixels that were hit receive a clock signal, leveraging the low expected occupancy to minimize the power consumption. The TOT state machine sets the enable signal when the discriminator output goes high and resets it when it goes low, in order to measure the Time over Threshold. The enable signal is not reset when the shutter signal goes down, as in the nominal operating conditions a pulse at the upper limit of the energy dynamic range can last for more than 1 µs, while the Shutter is high only for the duration of a pulse train (156 ns). It is therefore normal for the discriminator output to go to zero after the Shutter signal.
Each pixel also features a frequency divider that is used to adjust the TOT dynamic range by changing the frequency of the counting signal connected to the TOT counter. A chain of four flip flops, each toggled by the output of the previous one, implements the divider. A multiplexer controlled by a global configuration register in the periphery selects which of the flip-flop's output is used as a TOT counting signal, in order to choose the dividing factor.

5.6 PERIPHERY AND OTHER BLOCKS

5.6.1 *End-of-column block*

As already detailed, the readout of the chip is performed serially, one "double column" at a time, so two columns share a single "end of column" block, whose block scheme is depicted in figure 61.

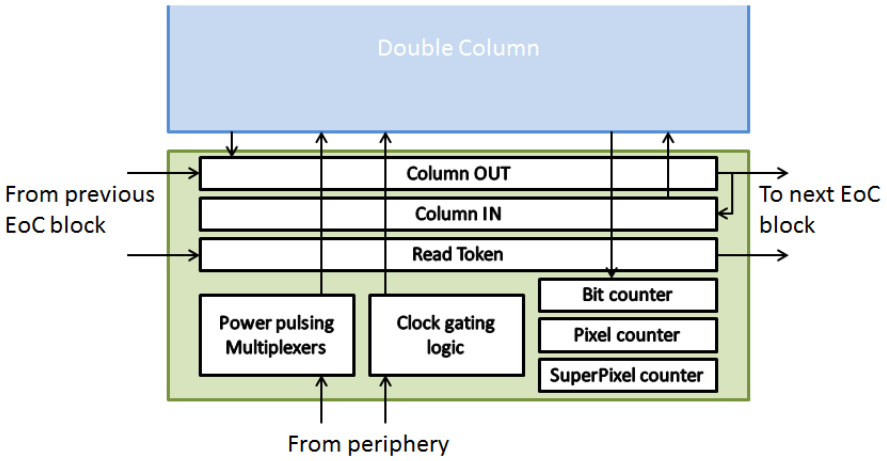


Figure 61: Block diagram of the end-of-column block.

Each end-of-column has a state machine that counts the number of pixels being read out (with multiple counters, taking into account skipped pixels and skipped clusters) to be able to send a start signal to the next column once the full column has been read out. A

"token" bit is moved from column to column to indicate which column is being read out. This token bit is also used to control the clock gating block, so that the readout clock is sent only to the columns being read out. The end-of-column block includes two flip-flops to hold data: one to shift data out (connected to the next column) and another one to buffer the configuration data and connected to the top pixel of the double column. Each double column also features another hit flag logic, working in the same way as the one included in the superpixel. In this case, the logic OR of the hit flags of the superpixels is stored, to check if any pixel in the double column was hit and skip the readout of the entire double column if necessary.
During the matrix programming phase, a complete word (1 bit per double column) is shifted in the data-line and then stored in a buffer register connected to the top pixel of each column. Between two consecutive words, the clock is sent to the array to shift the stored word into the array from the top, two columns at a time, to reduce the peak power consumption while still leaving enough time for hold and setup time of the flip-flops.

5.6.2 *State machines and control logic*

The periphery includes four digital blocks in addition to the end-of-column blocks. The first one is the command interpreter, which interfaces the chip to the outside world. Using a two-way handshake scheme, the periphery reads a 13-bit command, composed of a 5-bit op-code, specifying the operation to be performed, and an 8-bit data word. The command is then read by one of two additional blocks, either the readout control block, which starts the readout/programming by sending the appropriate command to the end-of-column blocks or the configuration state machine, which takes the 8-bit word and stores it in the configuration register specified in the op-code. It is used to program periphery DAC values or configure global configuration bits, such as the polarity of the pulses or the enable/disable signal for the compression scheme. The data word is ignored for the readout and matrix programming commands. The list of all commands that can be issued to the chip can be found in table 11.
Another state machine (not tied to the command interpreter) manages the power pulsing routines, by serially powering on (and off) the various columns one at a time when an external power signal is issued. The amount of clock cycles it takes for the full array to be powered up is configurable by setting a delay between the power-on of different double columns, using the configuration state machine described above.

Op-code	Operation
01000	Readout
01001	Array Programming
10000	Program Discriminator 1st stage DAC
10001	Program Discriminator 2nd stage DAC
10010	Program I_{KRUM} DAC
10011	Program Preamplifier DAC
10100	Program Threshold adj. DAC
10101	Program buffer DAC
10110	Program Preamplifier cascode DAC
10111	Program Threshold adj. cascode DAC
11000	Program DAC output cascode DAC
11001	Program Vfbk DAC
11010	Program Threshold DAC (MSBs)
11011	Program Threshold DAC (LSBs)
11100	Program Global Config register
11101	Program power-on speed
11110	Program power-off speed
11111	Program DAC output mux
00100	Program Discriminator 1st stage DAC (off state)
00101	Program Discriminator 2nd stage DAC (off state)
00110	Program Preamplifier DAC (off state)

Table 11: List of valid op-codes for controlling the CLICpix demonstrator.

5.6.3 *Global configuration and biasing*

The periphery includes 13 8-bit DACs and one 12-bit DAC (for the threshold voltage) to provide the biasing voltages to the analog blocks. DACs providing bias voltages to pixels have their output buffered to cope with gate leakage. Due to the lack of a bandgap voltage reference in the current prototype, the DACs themselves are biased with an external input. The LSB current in the binary weighted current source matrix is 50 nA but the dynamic range of the DAC is adjusted with its output stage. The DAC output is also sent to a multiplexer which selects one of them (which one can be chosen with a configuration register) to be sent to an external pad for debugging purposes.

DAC	Dynamic range
Discriminator 1st stage	0-12 µA
Discriminator 2nd stage	0-12 µA
Ikrum	0-50 nA
Preamplifier	0-4 µA
Threshold adj.	0-200 nA
Buffers	0-12 µA
Discriminator 1st stage (off state)	0-120 nA
Discriminator 2nd stage (off state)	0-120 nA
Preamplifier (off state)	0-40 nA

Table 12: Dynamic range of periphery DACs.

The current switching matrix of all the binary weighted DACs are identical, with the 12-bit one being made of a course 4-bit DAC and a fine 8-bit DAC. They are binary radix and use the same concept of multiple binary weighted current sources already explained in section 5.4. Their layout is also the same, with transistors placed using a common centroid. MOSFETs belonging to different current sources are placed symmetrically from the center and divided equally on both sides, in order to minimize the effect of gradients due to processing (see Hastings [41]). Simulations were performed to ensure that the INL of the DACs were lower than 0.5 LSB.
All the DAC outputs feature a buffer in order to increase the driving capabilities, as they are connected to a large load and their outputs can draw a non-negligible current due to gate leakage (they are connected to every pixel in the matrix).

5.6.4 *Power distribution*

The demonstrator uses two main power supplies, one for the analog structures and one for the digital ones. The analog power supply voltage is 1.2 V, the digital one is 1.0 V, as according to post-layout simulations the logic can meet the timing requirements even working with a reduced voltage. The power is distributed from the bottom of the chip to the top, with two metal lines (one per each power supply, using the thickest metal available in the technology) running along each double column.
One potential problem with this power distribution scheme is the voltage drop caused by the resistivity of the metal line, especially for the analog circuits that are more sensitive to a change in the power supply. In particular, the I_{KRUM} current source is sensitive to the power supply voltage and it can cause a non-uniformity in the TOT measurement based on the position of the pixel. It can be shown that the

voltage drop across a column can be calculated using the following formula:

$$\Delta V = \#pixels \cdot (\#pixels + 1) \cdot I \cdot r \cdot p \tag{33}$$

where I is the current absorbed by a single pixel, r is the resistivity of the metal line per unit of length and p is the pixel pitch. The $\#pixels$ is the number of pixels in a column (64 in this case, the "double column" architecture is already included in the formula), so the voltage drop scales quadratically with the matrix size. The resistivity of the metal line is dependent on the technology. The thickest available metal was used (~3 µm thick), making the line as wide as allowed by the design rules, while still leaving space for the connections to the input pads. The calculated drop for the demonstrator is less than 3 mV across the whole column. This value grows with the matrix size and it reaches ~200 mV for a 512 by 512 pixel matrix, which would prevent the circuits from working correctly. While this solution works for the demonstrator, so, in order to scale the matrix to bigger sizes for the final chip, another power distribution scheme must be found. The chosen technology does not feature a metallization option with more than one 3 µm thick copper layer, so the line resistivity cannot be lowered significantly by using more than one layer. A possible solution would be using Through-Silicon Vias (TSVs, see Motoyoshi [42]) in order to power the chip from multiple sides. Using TSVs would also have the advantage of not requiring wire bonding pads for I/O, so the non sensitive area of the chip would be smaller.

5.6.5 *I/O interfaces*

The chip is contacted using wire bonding pads, placed in an semicircle on one side of the die. A clearance of about 500 µm was left between the pixel array and the closest bonding pad, in order to leave space for a bump bonded sensor. Three different interfaces for I/O are used: one for analog signals, one for fast (high frequency) digital signals and one for slow digital signals. Other pads are included for the various power supplies.
The analog signals are connected to the internal circuits with bond pads including electrostatic discharge (ESD) protection structures. Output signals are buffered using two-stage Miller amplifiers in order to drive the pad capacitance. The slow digital signals, all of which are inputs, use foundry IP input pads, with input buffers using 2.5 V transistors. These buffers use a separate 2.5 V power supply voltage, used only for the pads. The voltage level was chosen to be compatible with the FPGA board used in the test setup (described in chapter 6). High frequency digital signals are implemented as differential low-voltage lines. Standard I/O pads, in fact, are designed for frequencies

smaller than 100 MHz, while the readout clock signals and the data lines require a faster interface. The Current Mode Logic (CML) standard was chosen, DC coupled and with a common mode voltage of 1.1 V. Each CML driver uses a constant current of 4 mA, so they are powered via an additional power pin, in order to measure the chip power consumption more independently from the pads. A detailed list of the I/O pins is shown in table 13. Additional test points (or "micro pads") on the chip were added to monitor internal signals for debug purposes.

#	Name	Interface
0	320 MHz clock P	CML IN
1	320 MHz clock N	CML IN
2	100 MHz clock P	CML IN
3	100 MHz clock N	CML IN
4	Data IN P	CML IN
5	Data IN N	CML IN
6	Data Ready P	CML IN
7	Data Ready N	CML IN
8	VSS CML	Analog IN
9	VDD CML	Analog IN
10	Data Strobe N	CML OUT
11	Data Strobe P	CML OUT
12	Data OUT N	CML OUT
13	Data OUT P	CML OUT
14	320 MHz clock OUT N	CML OUT
15	320 MHz clock OUT P	CML OUT
16	Analog VSS	Analog IN
17	Analog VDD	Analog IN
18	Power pulsing	CMOS IN
19	CML bias voltage	Analog IN
20	Digital VDD	Analog IN
21	Analog VDD	Analog IN
22	Analog VSS	Analog IN
23	Periphery DAC bias voltage	Analog IN
24	DAC output	Analog OUT
25	Test pulse voltage	Analog IN
26	Digital VSS	Analog IN
27	Digital VDD	Analog IN
28	Reset	CMOS IN
29	Shutter	CMOS IN
30	Test pulse switch	CMOS IN
31	Digital pads VDD	Analog IN
32	Digital VSS	Analog IN

Table 13: List of I/O pads on the CLICpix demonstrator.

6 CLICPIX ELECTRICAL CHARACTERIZATION

6.1 MEASUREMENT SETUP

A custom testing setup was developed to test and characterize the CLICpix prototype, consisting of a mezzanine board holding the chip, an FPGA development board and a command line interface program running on a Linux PC. A picture of the test setup is shown in figure 62. A picture of the chip taken with a microscope is in figure 63.

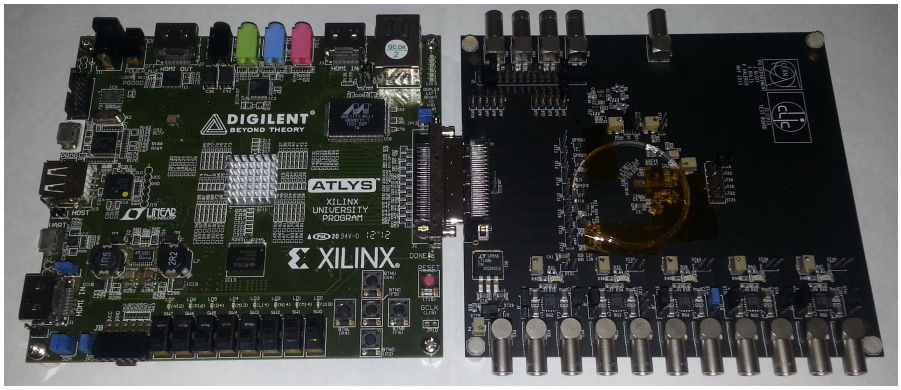


Figure 62: Custom PCB holding the CLICpix prototype (on the right) connected to an FPGA development board (on the left).

For sending test signals and reading out data from the chip, a Xilinx Spratan-6 FPGA was chosen, because it implements serializers and deserializers interfaces allowing data transmission rates of up to 1 Gbps. The FPGA is mounted on a commercial test board providing a Gigabit Ethernet connection which is used for the communication with the PC. The FPGA firmware implements a set of standard commands to be sent to the chip (the ones detailed in table 11), along with a data decoder to visualize the content of the chip counters in a readable format. It is also used to perform routines such as DAC or threshold scans, which consist of a large number of consecutive measurements.
All data are sent to a PC which is used to control the FPGA. A command-line-interface program written in Python is used to connect to the FPGA and send commands to it. This program is also used to format the output data, presenting the pixel states in an array form.

The chip is wire bonded to a custom PCB hosting the electronics needed for the chip operation and testing. Five voltage regulators

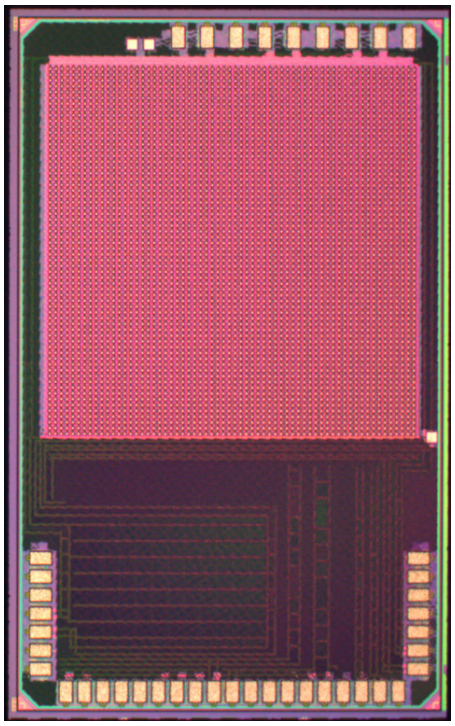


Figure 63: A picture of the CLICpix prototype chip taken with a microscope. The die size is 3 mm by 1.85 mm.

are included to provide the power supplies to the chip, with each regulator being used for different parts of the chip in order to monitor their power consumption independently from each other. Trimmers are used to set the global biasing current of the analog structures (due to the lack of a band-gap inside the chip) and level translators convert I/O signals from the CML standard used by the chip to the LVDS standard used by the FPGA. A number of test points are included as well to allow monitoring some signals with a scope if necessary during the first testings.

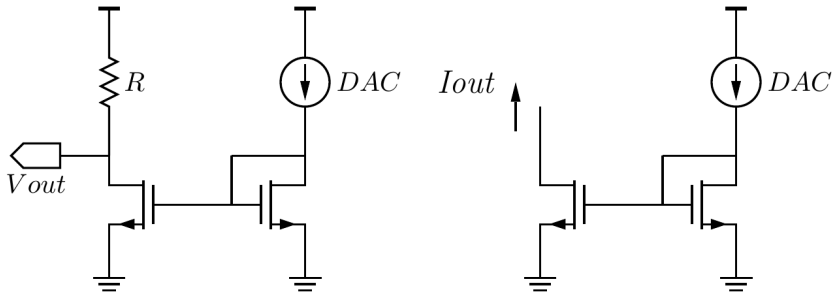
6.2 TEST RESULTS

The chip testing started in May, 2013. The characterization is still in progress, so results presented here should be considered as preliminary. All measurements were done using a 50 MHz readout clock, further tests to check the maximum speed of the chip are pending. The acquisition clock was run at the nominal 100 MHz.

6.2.1 *Periphery measurements*

The first blocks to be tested were the periphery DACs, whose output voltage can be monitored from an I/O pad. The prototype correctly recognized commands sent to it and it showed the expected behavior

in the DAC programming. All codes were scanned, to measure the characteristic of the periphery DACs. The DACs are designed to produce a voltage or a current output. Both types use the same current mirror array architecture but their output stage (described in figure 64) is different.



(a) Voltage DAC (b) Section view.

Figure 64: Schematic of Voltage (left) and Current (right) DAC output stages.

The scan of the 8-bit voltage DACs can be found in figure 65.

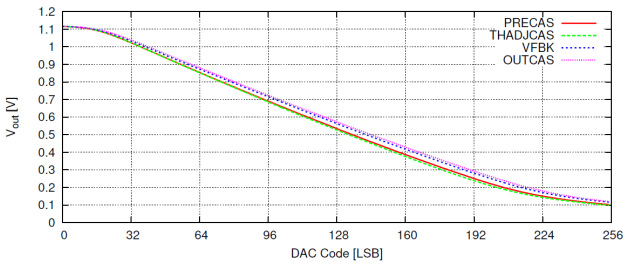


Figure 65: Scan of the characteristics of the voltage DACs in the CLICpix periphery.

The non linearity for values outside the range of 0.2 V to 1.1 V was expected and it is mainly due to the buffers at the DAC outputs not being designed as rail-to-rail. This doesn't have an impact on the operation of the chip, since no biasing voltage needs to be set outside this interval. Current DACs were scanned as well, but their linearity cannot be tested as only their voltage output could be measured. Their characteristics (shown in figure 66) were therefore compared to simulations and were found to be consistent with them within the uncertainties due to process variations.

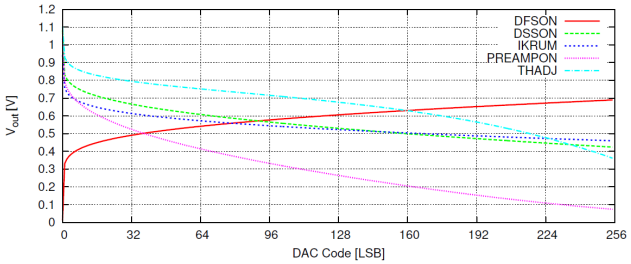


Figure 66: Scan of the characteristics of the current DACs in the CLICpix periphery. The threshold adjustment DAC transfer function changes slope because it uses a cascode configuration and it was measured without changing its cascode voltage.

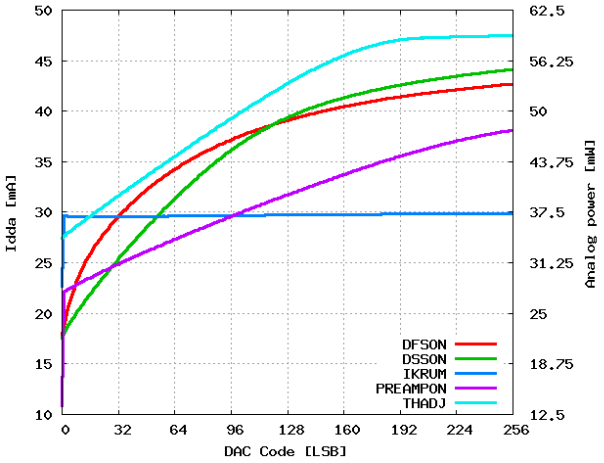


Figure 67: Variation of the chip power consumption by varying the bias currents of different blocks.

The power consumption of the pixel matrix was analyzed by varying the bias currents of the analog blocks. Results are shown in figure 67. The power consumption is basically not affected by the I_{KRUM} current, as its value is more than two orders of magnitude lower than the total current used by the analog pixel. The curve can thus be used as a baseline to compare the other scans. By subtracting the nominal value by the one corresponding to a code of 0 (which corresponds to a negligible current) and dividing the result by the number of pixels, the current in a particular block can be estimated. The calculated currents are matching the expected values, given the uncertainty on the current biasing the DACs provided by the PCB.

The threshold DAC was scanned as well, its characteristic can be found in figure 68. This DAC is implemented as two 8-bit DACs with overlapping characteristics, in order to achieve a 12-bit accuracy. All 65536 codes were tested and a software correction was applied, to reorder the codes and obtain a linear characteristic. The resulting DAC has a 13-bit accuracy (shown in figure 69), which is more than the

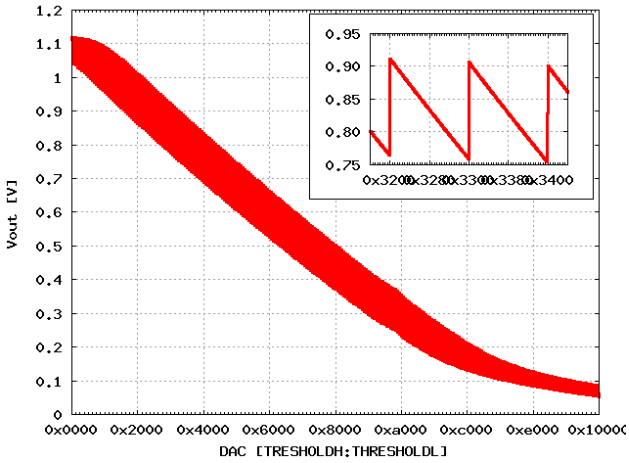


Figure 68: Scan of the characteristics of the threshold DAC in the CLICpix periphery.

simulated value.

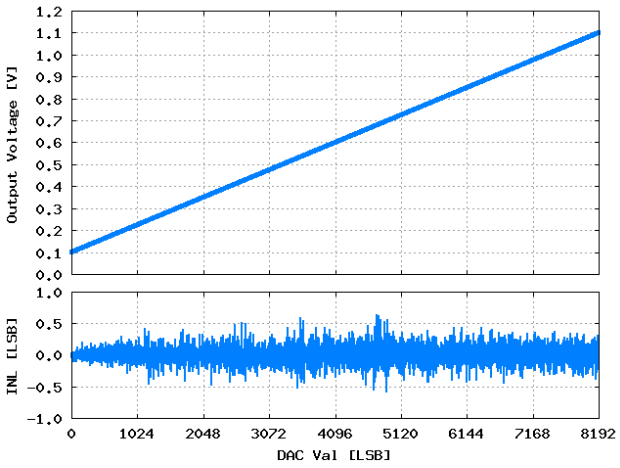


Figure 69: Scan of the characteristics of the threshold DAC in the CLICpix periphery, corrected via software. The deviation from a linear fit is shown in the bottom plot.

Other measurements on the periphery blocks included the power pulsing control system, which works according to specifications. By switching to a low-power state, the power consumption of the pixel matrix is reduced by the expected amount (~20 mA in the nominal conditions). The power-on and power-off times can be programmed using global configuration bits and the measured times match the programmed values.

6.2.2 *Pixel characterization*

The first tests on the pixel matrix were purely digital, implementing a routine to program the array using random data checking that the chip sends the same bit stream during a read out. This check was performed successfully. Measurements using test pulses were then performed and the chip was able to correctly detect pulses and stream out the measurement results. Both the TOT and TOA counters count correctly (meaning that the state machines producing their clock signals behave as expected, including the TOT clock divider). The photon counting mode was also tested to perform threshold scans, finding the baseline voltage for the pixels.
The on-chip zero compression algorithm was tested and test readouts were completed using it. The FPGA was used to decode the output stream and successfully compare it to an uncompressed readout.
Tests on the TOT counting were performed injecting a controlled amount of charge in the preamplifier, using the internal test capacitor. A plot showing the dependency of the TOT on the value of the I_{KRUM} current (which determines how fast the pulse returns to baseline) is shown in figure 70. The linearity of the TOA measurement was also tested, by sending test pulses with different delays from the shutter signal.

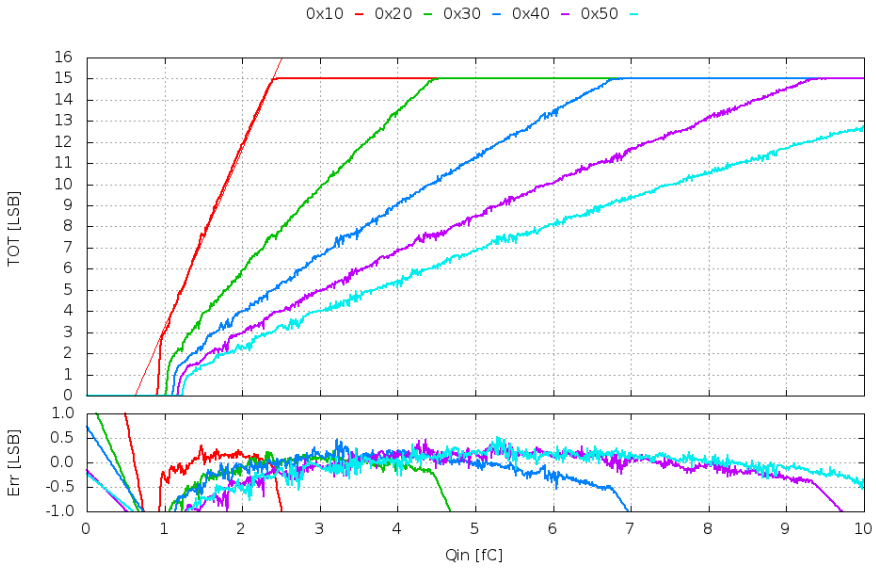


Figure 70: Scan of the average TOT counts of one pixel for different I_{KRUM} values, calculated on 256 samples per point. The deviation from a linear fit is shown in the bottom plot.

Tests on analog performances of the pixel front-end were performed using the pixel counting mode (summing the results of multiple measurements, in order to work around the 4-bit saturation of the counter). An S-curve measurement was performed (explained in chapter 4) by

sweeping the threshold for various input charges, as shown in figure 71. This allowed to estimate the front-end noise and the pixel gain (by measuring the shift between curves corresponding to different input charges). Results match the simulations closely, taking into account the uncertainty on the input charge due to the process variations on the value of the test pulse input capacitors

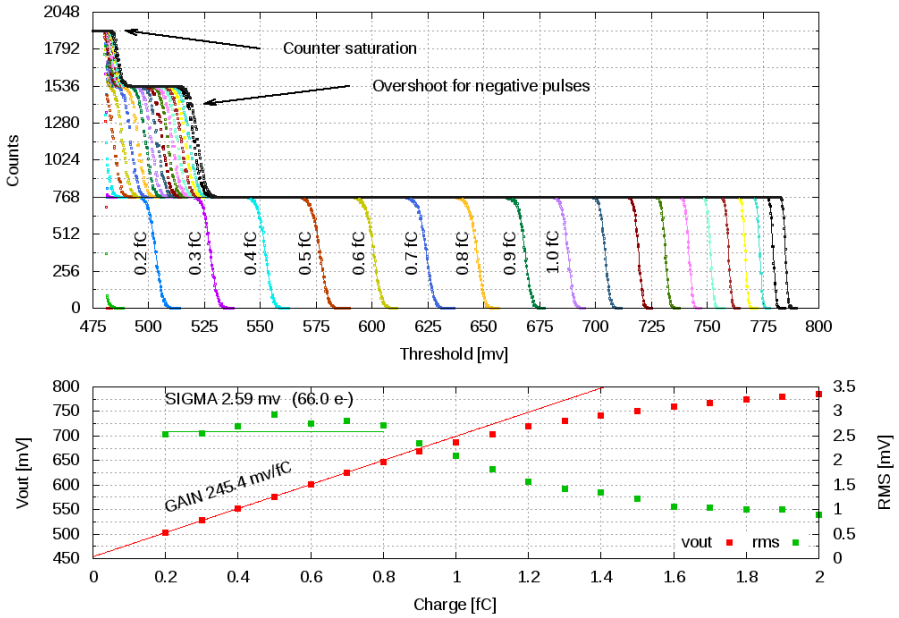


Figure 71: S-curve measurement on one pixel for various input charges. In the bottom plot, an estimation of the gain and noise rms value can be found.

A similar measurement was used to find the baseline voltage of each pixel's front-end, by localising the threshold value corresponding to the peak of the gaussian obtained by counting hits only due to noise. This allowed developing a calibration algorithm to equalize the baseline of pixels across the entire matrix. The threhsold value corresponding to the lowest and highest calibration DAC code in every pixel was measured, so that the DAC characteristic for every pixel could be interpolated. From this measurement, a code for each pixel was chosen in order to equalize the effective threshold. The results of the measurement, before and after calibration, can be found in figure 72.

A vertical stripe pattern is visible in the uncalibrated maps, due to the slightly different layout of pixels at the two sides of the double column. The variation is, however, within the simulated mismatch, so it can be equalized normally.
A summary of measurement results compared with simulations is

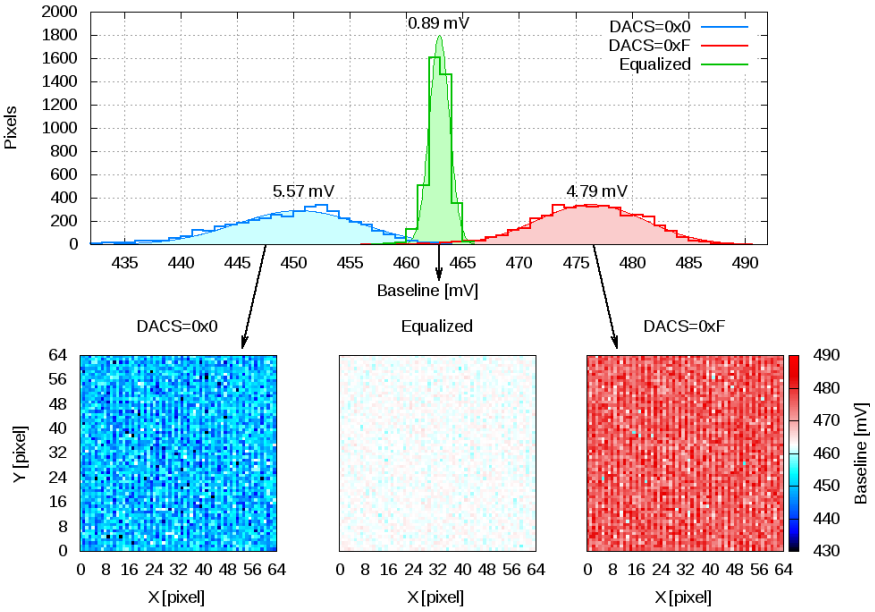


Figure 72: CLICpix threshold map, before (left and right plots) and after equalization (middle plot).

presented in table 14.

Parameter	Simulated Value	Measured Value
TOA Accuracy	< 10 ns	< 10 ns
Gain	44 mV/ke^-	40 mV/ke^-
Dynamic Range	up to 40 ke^-	up to 45 ke^-
Equivalent Noise	$\sigma = 60\,e^-$	$\sigma = 66\,e^-$
DC Spread (uncalibrated)	$\sigma = 160\,e^-$	$\sigma = 128\,e^-$
DC Spread (calibrated)	$\sigma = 24\,e^-$	$\sigma = 22\,e^-$
Minimum threshold	388 e^-	417 e^-
Power consumption	6.5 µW	7 µW

Table 14: Summary of results on analog measurements on the CLICpix prototype, compared to simulations, for nominal conditions.

Tests on other specifications (i.e. temperature and time stability, maximum clock frequency tests and radiation hardness characterization) will be carried out in the following weeks. The bump bonding of the chip to a silicon sensor is also planned, so that measures can be taken using radioactive sources rather than the internal charge injection mechanism. This will allow for more accurate characterization, as the value of the test capacitance can be ignored.

Part IV

CONCLUSIONS

In this thesis the design of pixel detectors suitable for space and High Energy Physics applications has been described, with the main focus on two different projects: the Fractional Packet Counting architecture and the CLICpix chip. The two projects are suited for different applications, with the first one being an integrating detector and the second one working in single event counting, showing the differences in the specifications and design techniques between the projects.

The FPC architecture was analyzed, simulated and compared to other HDR front-end architectures. Its performances were found to be comparable to state-of-the-art solutions, while using fewer bits of data, allowing for lower data-rates which can be critical in applications such as satellite systems, where the bandwitdth of the communication system is limited.

The second part of the thesis work was carried out at CERN within the framework of the CLIC vertex detector experiment. A commercial 65 nm CMOS technology was tested and characterized for HEP applications. Results on performances in terms of radiation hardness were positive and led to useful data to design radiation tolerant electronics. This technology was then used to design a prototype for CLICpix, a proposed hybrid pixel detector to be used as the vertex detector for CLIC, featuring a 25 µm pixel pitch, simultaneous Time-over-Threshold and Time-of-Arrival measurements, data compression and power pulsing capabilities. The prototype was tested and results meet the specifications, although the full electrical characterization is still in progress.

Future developments include completing the characterization of the prototype (also using radioactive sources). Finally, the many lessons from designing and characterizing the technology and the prototype chip should lead to the design and optimization of a full version of the CLICpix chip.

BIBLIOGRAPHY

[1] Francis Anghinolfi, P Aspell, K Bass, Werner Beusch, L Bosossio, C Boutonnet, P Burger, M Campbell, Enrico Guido Chesi, C Claeys, J C Clemens, M Cohen-Solal, I Debusschere, P A Delpierre, D Di Bari, B Dierickx, C C Enz, E Focardi, F Forti, Y Gally, M Glaser, Thierry Gys, M C Habrard, Erik H M Heijne, L Hermans, R B Hurst, P Inzani, J J Jaeger, Pierre Jarron, F Krummenacher, F Lemeilleur, V Lenti, V Manzari, G Meddeler, M Morando, A Munns, F Nava, F Navach, C Neyer, G Ottaviani, F Pellegrini, F X Pengg, R Perego, M Pindo, R Potheau, Emanuele Quercigh, N G Redaelli, L Rossi, D Sauvage, G F Segato, S Simone, G Stefanini, G Tonelli, G Vanstraelen, G Vegni, H Verweij, Gert M Viertel, and J Waisbard. A 1006 element hybrid silicon pixel detector with strobed binary output. (CERN-ECP-91-026. CERN-ECP-91-26):10 p, Dec 1991. (Cited on page 3.)

[2] N. Stoffle, L. Pinsky, D. Turecek, E. Semones, J. Hauss, C. Amberboy, K. Lee, N. Zapp, A. Bahadori, Z. Vykydal, J. Jakubek, and S. Pospisil. A Timepix Based Radiation Environment Monitor (REM) for Application to Space Radiation Dosimetry . *22nd Annual NASA Space Radiation Investigators' Workshop*, 2011. (Cited on page 3.)

[3] R. Ballabriga, M. Campbell, E. Heijne, X. Llopart, L. Tlustos, and W. Wong. Medipix3: A 64 k pixel detector readout chip working in single photon counting mode with improved spectrometric performance. *Nucl. Instr. and Meth. A*, June 2010. (Cited on pages 4 and 27.)

[4] T. Bergauer. Silicon Detectors in High Energy Physics . 2011. URL http://particles.ipm.ir/conferences/DetectorSchool2011/lectures/Thomas2.pdf. (Cited on page 5.)

[5] S. Bettarini, A. Bardi, G. Batignani, F. Bosi, G. Calderini, et al. Development of deep N-well monolithic active pixel sensors in a 0.13-mu-m CMOS technology. *Nucl.Instrum.Meth.*, A572:277–280, 2007. doi: 10.1016/j.nima.2006.10.302. (Cited on page 5.)

[6] Randy Luhta, Marc Chappo, Brian Harwood, Rod Mattson, Dave Salk, and Chris Vrettos. A new 2d-tiled detector for multislice ct. pages 61420U–61420U–12, 2006. doi: 10.1117/12.653245. URL +http://dx.doi.org/10.1117/12.653245. (Cited on page 11.)

[7] Sam Kavusi and Abbas El Gamal. Quantitative study of high-dynamic-range image sensor architectures. pages 264–275, 2004.

doi: 10.1117/12.544517. URL +http://dx.doi.org/10.1117/12.544517. (Cited on pages 11 and 15.)

[8] Jehyuk Rhee and YoungJoong Joo. Wide dynamic range cmos image sensor with pixel level adc. *Electronics Letters*, 39(4):360–361, 2003. ISSN 0013-5194. doi: 10.1049/el:20030246. (Cited on page 12.)

[9] A. Nascetti and P. Valerio. Design of pixel electronics based on asynchronous self-reset approach with floating-point output representation for high dynamic range imagers. *Journal of Instrumentation*, 6(01):C01070–C01070, January 2011. ISSN 1748-0221. doi: 10.1088/1748-0221/6/01/C01070. URL http://stacks.iop.org/1748-0221/6/i=01/a=C01070?key=crossref.e5d88d24d7796735e537626352e40e5d. (Cited on page 12.)

[10] S. Manen, L. Royer, and P. Gay. A custom 12-bit cyclic ADC for the electromagnetic calorimeter of the International Linear Collider. In *IEEE Nuclear Science Symposium Conference Record*, Dresden, Allemagne, 2008. IEEE. URL http://hal.in2p3.fr/in2p3-00344387. (Cited on page 14.)

[11] A. Nascetti and P. Valerio. Use of fractional packet counting for high dynamic range imaging applications. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 648:S146–S149, August 2011. ISSN 01689002. doi: 10.1016/j.nima.2010.12.060. URL http://linkinghub.elsevier.com/retrieve/pii/S0168900210028019. (Cited on page 16.)

[12] M. Aicheler, P. Burrows, M. Draper, T. Garvey, P. Lebrun, K. Peach, N. Phinney, H. Schmickler, D. Schulte, and N. Toge. A Multi-TeV linear collider based on CLIC technology: CLIC Conceptual Design Report. 2012. URL http://clic-study.org/accelerator/CLIC-ConceptDesignRep.php. (Cited on pages 21, 22, 23, 26, and 62.)

[13] G. Ising. Prinzip Einer Methode Zur Herstellung Von Kanalstrahlen Hoher Voltzahl. *Arkiv för Matematik, Astronomi och Fysik*, 1928. (Cited on page 21.)

[14] E. Jensen. RF Cavity Design. *CERN Accelerator School*, 2011. URL http://cas.web.cern.ch/cas/Greece-2011/Lectures/Jensen.pdf. (Cited on page 21.)

[15] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, August 2012. doi: 10.1016/

j.physletb.2012.08.020. URL http://arxiv.org/abs/1207.7214. (Cited on page 22.)

[16] L. Linssen, A. Miyamoto, M. Stanitzki, and H. Weerts. PHYSICS AND DETECTORS AT CLIC. 2011. URL http://clic-study.org/accelerator/CLIC-ConceptDesignRep.php. (Cited on pages 22 and 24.)

[17] Micron Semiconductor. Available products. URL http://www.micronsemiconductor.co.uk/product.asp. (Cited on page 25.)

[18] CORWIL. Precision Wafer Thinning and Wafer Polishing. URL http://www.corwil.com/services/waferthinning/pdfs/CORWIL-WaferThinandPolish-Datasheet.pdf. (Cited on page 25.)

[19] X. Llopart, R. Ballabriga, M. Campbell, L. Tlustos, and W. Wong. Timepix, a 65k programmable pixel readout chip for arrival time, energy and/or photon counting measurements. *Nucl. Instr. and Meth. A*, pages 485 – 494, October 2007. (Cited on page 25.)

[20] T Poikela, J Plosila, T Westerlund, J Buytaert, M Campbell, X Llopart, R Plackett, K Wyllie, M van Beuzekom, V Gromov, R Kluit, F Zappon, V Zivkovic, C Brezina, K Desch, X Fang, and A Kruth. Architectural modeling of pixel readout chips velopix and timepix3. *Journal of Instrumentation*, 7(01):C01093, 2012. URL http://stacks.iop.org/1748-0221/7/i=01/a=C01093. (Cited on pages 27 and 29.)

[21] W. Wong, G. Anton, R. Ballabriga, M. Böhnel, M. Campbell, E. Heijne, X. Llopart, T. Michel, I. Münster, R. Plackett, P. Sievers, P. Takoukam, L. Tlustos, and P. Valerio. A pixel detector asic for dosimetry using time-over-threshold energy measurements. *Radiation Measurements*, 46(12):1619–1623, December 2011. ISSN 13504487. doi: 10.1016/j.radmeas.2011.06.061. URL http://linkinghub.elsevier.com/retrieve/pii/S1350448711003209. (Cited on page 27.)

[22] F. Faccio and G. Cervelli. Radiation-induced edge effects in deep submicron CMOS transistors. *IEEE Transactions on Nuclear Science*, (06):2413 – 2420, December 2005. ISSN 0018-9499. doi: 10.1109/TNS.2005.860698. URL http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=1589217. (Cited on pages 30, 32, and 40.)

[23] G. I. Zebrev, M. S. Gorbunov, and V. S. Pershenkov. Radiation induced leakage due to stochastic charge trapping in isolation layers of nanoscale MOSFETs. *SPIE, Micro- and Nanoelectronics*, April 2008. doi: 10.1117/12.

802480. URL http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=792241. (Cited on page 30.)

[24] R. Velazco. Radiation Effects on Embedded Systems. *Springer*, 2007. (Cited on pages 31, 34, and 36.)

[25] G.B. Cerati, M.E. Dinardo, A. Florez, S. Kwan, and A. Lopez. Radiation tolerance of the CMS forward pixel detector. *Nucl.Instrum.Meth.*, A600:408–416, 2009. doi: 10.1016/j.nima.2008.11.114. (Cited on page 32.)

[26] D. K. Nichols. A Review of Dose Rate Dependent Effects of Total Ionizing Dose (TID) Irradiations. *Nuclear Science, IEEE Transactions on*, 27(2):1016–1024, 1980. ISSN 0018-9499. doi: 10.1109/TNS.1980.4330968. (Cited on page 32.)

[27] Sandro Bonacini and Kostas Kloukinas. *Design of two digital radiation tolerant Integrated Circuits for High Energy Physics experiments data readout*. PhD thesis, UniversitÃ di Modena e Reggio Emilia, Geneva, 2003. Presented on 21 Mar 2003. (Cited on page 33.)

[28] T. C. May and M. H. Woods. Alpha-Particle-Induced Soft Errors in Dynamic Memories. *Electron Devices, IEEE Transactions on*, 26 (1), January 1979. (Cited on page 33.)

[29] G. Anelli, M. Campbell, M. Delmastro, F. Faccio, S. Floria, A. Giraldo, E. Heijne, P. Jarron, K. Kloukinas, A. Marchioro, P. Moreira, and W. Snoeys. Radiation tolerant VLSI circuits in standard deep submicron CMOS technologies for the LHC experiments: practical design aspects. *Nuclear Science, IEEE Transactions on*, 46(6):1690–1696, 1999. ISSN 0018-9499. doi: 10.1109/23.819140. (Cited on page 35.)

[30] R. Naseer. The DF-DICE storage element for immunity to soft errors. In *Circuits and Systems, 2005. 48th Midwest Symposium on , August 7-10, 2005 Page(s):303 - 306*, 2005. (Cited on page 35.)

[31] S. Bonacini, P. Valerio, R. Avramidou, R. Ballabriga, F. Faccio, K. Kloukinas, and A. Marchioro. Characterization of a commercial 65 nm CMOS technology for SLHC applications. *Journal of Instrumentation*, 7(01):P01015–P01015, January 2012. ISSN 1748-0221. doi: 10.1088/1748-0221/7/01/P01015. URL http://stacks.iop.org/1748-0221/7/i=01/a=P01015?key=crossref.4192548a65603a5bcea0e76ca1f7b90f. (Cited on pages 39 and 44.)

[32] F. Krummenacher. Pixel detectors with local intelligence: an IC designer point of view. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 305

(3):527 – 532, 1991. ISSN 0168-9002. doi: 10.1016/0168-9002(91)90152-G. URL http://www.sciencedirect.com/science/article/pii/016890029190152G. (Cited on pages 48, 49, and 71.)

[33] Rafael Ballabriga, Michael Campbell, and Xavier VilasÃs-Cardona. *The Design and Implementation in 0.13ÎŒm CMOS of an Algorithm Permitting Spectroscopic Imaging with High Spatial Resolution for Hybrid Pixel Detectors*. PhD thesis, Ramon Llull U., Barcelona, Geneva, 2009. Presented on 24 Nov 2009. (Cited on page 50.)

[34] H. Tuinhout, N. Wils, M. Meijer, and P. Andricciola. Methodology to evaluate long channel matching deterioration and effects of transistor segmentation on MOSFET matching. In *Microelectronic Test Structures (ICMTS), 2010 IEEE International Conference on*, pages 176–181, 2010. doi: 10.1109/ICMTS.2010.5466824. (Cited on page 52.)

[35] Marc Pastre and Maher Kayal. Digital compensation circuits and sub-binary digital-to-analog converters. In *METHODOLOGY FOR THE DIGITAL CALIBRATION OF ANALOG CIRCUITS AND SYSTEMS*, volume 870 of *The International Series in Engineering and Computer Science*, pages 23–91. Springer Netherlands, 2006. ISBN 978-1-4020-4252-2. doi: 10.1007/1-4020-4253-1_3. URL http://dx.doi.org/10.1007/1-4020-4253-1_3. (Cited on page 54.)

[36] Pieter Harpe Arthur van Roermund, Hans Hegt. Smart AD and DA Conversion. *Springer*, 2010. doi: 10.1007/978-90-481-9042-3. URL http://www.sciencedirect.com/science/article/pii/016890029190152G. (Cited on page 55.)

[37] M. Friedl. PSI-Testbeam results, Dec. 2000. URL http://www.hephy.at/user/friedl/psi/. (Cited on page 62.)

[38] G. Blanchot and C. Fuentes. Power pulsing schemes for vertex detectors at CLIC. *Journal of Instrumentation*, 8(01): C01057, 2013. URL http://stacks.iop.org/1748-0221/8/i=01/a=C01057. (Cited on pages 63 and 69.)

[39] W. M.C. Sansen. Low-Noise Wide-Band Amplifiers in Bipolar and CMOS Technologies. *Springer*, 1991. (Cited on page 74.)

[40] C. H. Roth. Fundamentals of logic design. *Cengage Learning*, 2009. (Cited on page 84.)

[41] Alan Hastings. An Introduction to Matching and Layout. URL http://www.ece.arizona.edu/~brew/ece304spr07/PowerPoint/A_Hastings(4_4_05)[1].ppt. (Cited on page 88.)

[42] M. Motoyoshi. Through-silicon via (tsv). *Proceedings of the IEEE*, 97(1):43–48, 2009. ISSN 0018-9219. doi: 10.1109/JPROC.2008.2007462. (Cited on page 89.)

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography *"The Elements of Typographic Style"*. `classicthesis` is available for both LaTeX and LyX:

http://code.google.com/p/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured at:

http://postcards.miede.de/

Final Version as of October 3, 2013 (`classicthesis` version 4.1).