An Analog Pixel Front-End for High Granularity Space-Time Measurements

Lorenzo Piccolo

* * * * * *

Prof. Angelo Rivetti, Supervisor
Prof. Michele Goano Co-supervisor

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

. . . .
Lorenzo Piccolo
Turin, June 22, 2022

Summary

This thesis describes the results of the design and characterization work done to develop an analog front-end for a pixel front-end Application Specific Integrated Circuit (ASIC) with time measurement capability.

The work was carried out from 2018 to 2021, and has produced two prototype ASICs manufactured in a commercial 28 nm Complementary Metal–Oxide–Semiconductor (CMOS) technology. This includes the scheme design, simulation and implementation, as well as the experimental characterization of the prototypes and the design of the experimental setup employed. The main subject of the research is the analog part of the core channel of the ASIC. The main function of this circuit is to transduce the sensor signal into a format that best fits the input of the subsequent digitizer. Both of these functions play a crucial role in the time measurement process since they represent the first processing of the raw signal. Particular attention has been paid to minimize the time fluctuation contribution of this circuit to the overall system time fluctuation.

An important aspect to consider is that the front-end channel is part of a large multichannel system. This aspect influences the single channel design by imposing constrains arising from the pixel-matrix configuration. The channel must fit inside a limited pixel-area, while leaving room for the implementation of the other pixel dedicated circuits. The area limitation will also constraint the total power budget available to the analog front-end. Moreover, a certain degree of uniformity is required across the pixel matrix. These challenges have been addressed with novel architectural solutions. First, a new very front-end scheme has been designed in order to reduce the time fluctuations within the available power budget. Second, a discrete time technique has been implemented in order to equalize the channels while limiting the area and power consumption needed for this task.

The author's work was not only limited to the core channel. Given the system-like nature of the project, design and test work was also carried out on a system level. The author is one of the three lead designers of the ASIC. Therefore, a crucial role has been covered in the definition of the pixels and matrix architectures and floor-plans. The main contribution was the integration of the full analog part of the pixel-matrix, including all the service and configuration blocks used to operate it. As a result, the whole pixel matrix can be operated by providing only one reference

signal configured via a digital interface.

The main application for this type of ASIC is in future experiments in High Energy Physics (HEP). In this field, particles properties and trajectories are reconstructed by large detectors composed of many sensitive layers. The information gathered on the innermost layers is used to reconstruct the particle track in the process named tracking. This process is based on the particle hits positions and is used to date. However, HEP experiments such as the ones in Conseil Européen pour la Recherche Nucléaire (CERN), are planning to increase their accelerator nominal luminosity. In case of the Large Hadron Collider (LHC), this upgrade is scheduled in 2029. The luminosity boost will increase the chance of observing rare events by increasing the event rate. The downside of this approach is that the number of spurious events is also expected to rise, making current tracking techniques ineffective. This issue can be solved by adding a time measurement capability to the inner layers of the detector achieving what is called 4D-tracking. In principle, future detectors require a time resolution of at least 100 ps at the level of pixels tens of micrometers large. 4D-tracking demands therefore the research and development of new pixel front-end ASICs with timing measurement capability.

The construction of this type of detectors is beyond the sole scope of ASIC level design. In fact, in order to reach the required space-time resolution, the whole system must be tailored around this goal. The TimeSPOT project (Time and SPace real-time Operating Tracker) by the Italian institute for nuclear physics (Istituto Nazionale di Fisica Nucleare INFN) aims to research and develop a demonstrator detector suitable for high-luminosity HEP experiments. The demonstrator will be realized via a small-scale telescope that will include the pixel sensor matrix, the pixel front-end ASIC, and the readout electronics. The work presented in this thesis is actually part of the development of the Timespot ASIC family. The ASIC is researched in tandem with its sensor and readout in order to achieve the target specifications in terms of: space-time resolution, maximum event rate, data throughput, power consumption and radiation tolerance. The project time resolution specification has been defined on the basis of what is achieved by its sensor: 20 ps or better.

The 28 nm CMOS technology node was chosen over more conventional nodes for the field due to its superior jitter performance over power consumption. It also opens up the possibility to integrate more features in the same silicon area. The radiation hardness of this node is on par with the best results achieved in 65 nm and 130 nm nodes. The complexity of this technology has determined an additional challenge in terms of ASIC design.

Other applications for this ASIC can be found in other fields that require a granular space-time measurement such as: detector for space applications, medical equipments and in general detector for imaging applications.

The thesis will discuss the work done by starting with two introductory chapters and then moving to the description of the design and characterization work.

The introductory material is used to give context to the following chapter and to introduce concepts that will be extensively utilized throughout the whole thesis. Chapter 1 introduces the reader to the field of space-time measurement in HEP and its pixel front-end ASICs, including a small review of the ASICs currently in development. Chapter 2 presents the architecture, specifications and plans of the ASICs for the TimeSPOT project. Chapters 3 and 4 describe the design and characterization of Timespot0 analog front-end. Timespot0 is the first ASIC developed for the project. The design process is described in terms of the circuit architecture and implementation. A brief analytical derivation of the circuit operation is also included. Chapter 5 and 6 elaborates on the same topics regarding the second prototype: Timespot1. Timespot1 features a pixel-matrix with a thousand channels. The design description of chapter 5 is based on what was already presented in chapter 3 focusing on the changes between the two versions. This chapter also includes the system level work carried out specifically by the author. Finally, a brief conclusion is outlined at the end of the thesis.

Acknowledgements

I would like to acknowledge Ilaria Gramigna Policreti for the English proofreading.

I would like to dedicate this thesis to my colleagues from INFN Cagliari and INFN Milano. I hope that this thesis will serve as a reference for the future developments on our project.

Contents

List of Tables

List of Figures

xv

Chapter 1

Pixel Front-End ASICs with Timing

This chapter illustrates the state of the art in the field of pixel front-end ASIC with timing measurement capability. First of all, section 1.1 presents the state of the research in the field of timing measurement for HEP: starting from a discussion on the need to update the current detectors to include the time information, moving to the candidate sensors for this purpose and ending with a general presentation of the ASICs used in this field. Section 1.2 illustrates the general structures and concepts in common to the architectures of these type of ASICs. Section 1.3 will articulate on the goals and challenges of implementing the measurement of the time information in the pixel ASIC architecture. Finally, section 1.4 will present a review on the state of the art of the timing pixel front-end ASICs currently in development for HEP experiments.

1.1 Overview

This section presents the field of application of timing pixel front-end ASICs for HEP. The first part will elaborate on the needing of the time information in HEP, whereas the second one presents the major candidate as pixel sensors for this purpose. The pixel sensor represents a key constraint in the design of a front-end ASIC.

The second part describes the historical process which has led to the development of pixel front-end ASICs, and their advantage compared to other solutions for radiation detection.
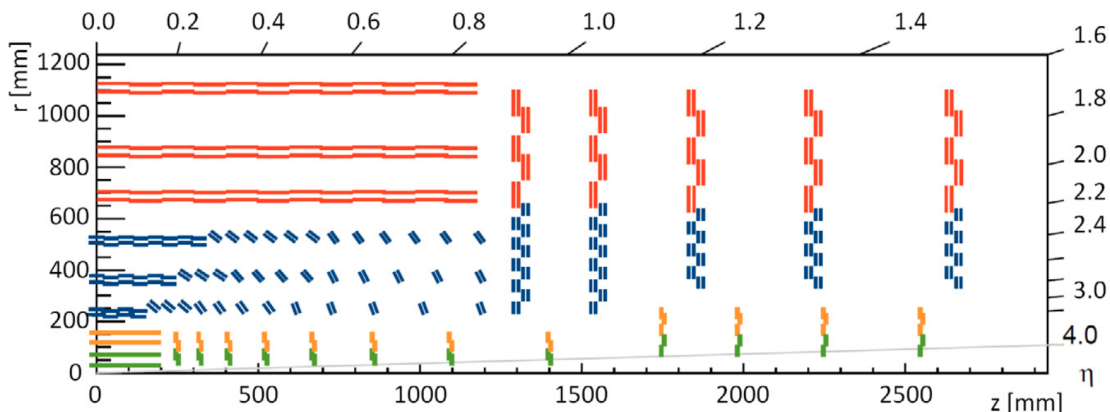
Figure 1.1: Layout of the new CMS Tracker. Taken from [25].

1.1.1 Space-Time Measurements in High Energy Physics

The goal of HEP is investigating the fundamental working principles of matter by studying the properties of its constituents: the elementary particles. This goal is achieved by the observation of exotic states of matter produced through the interaction of common particles (such as protons or electrons) accelerated to speeds close to the one of light. The particles produced with this process are statically determined, wherein the interesting states of matter constitute a small fraction of rare events. In particle colliders two bunches of accelerated particles are crossed at a constant rate, the interaction point of this reaction is positioned inside a particle detector which detects the fragments and secondary products of the generated particle. The typical particle detector consists of different sensitive layers used to identify the particle position or energy. Particle momentum is then reconstructed by tracking the position of signals matching the same particle. The innermost layers of the detector are called tracker [41] and feature a high granularity space measurement. An example of a particle tracker is shown in figure 1.1. The outer layers are constituted of larger sensitive units with a more coarse space measurement, but with different measurement of other event properties (such as energy in calorimeters [28]).

In the Large Hadron Collider (LHC) at Conseil Européen pour la Recherche Nucléaire (CERN) bunches of 10^5 protons accelerated to $6.5\,\mathrm{TeV}$ are condensed to an interaction point of $64\,\mu\mathrm{m}$ in size and crossed every $25\,\mathrm{ns}$ producing about 27 collisions each time. A key parameter in determining collision probability is the luminosity L $(\mathrm{cm}^{-2}\mathrm{s}^{-1})$, it expresses the probability of interaction per unit time, area and cross-section. In order to find more rare events in the same observation time, HEP experiments are planning to increase their nominal luminosity. In High-Luminosity LHC (HL-LHC) the nominal luminosity is planned to be increased from $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ to $10^{35}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ [7] with an expected hit-rate per unit area of

2
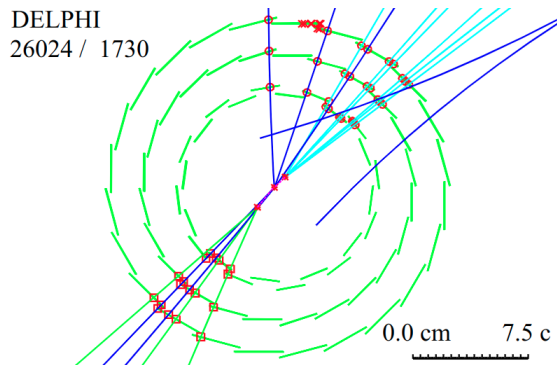
Figure 1.2: Example of particle tracking: a b-event with the DELPHI vertex detector. Taken from [41].

$3\,\text{GHz}\cdot\text{cm}^{-2}$. In this regime, the number of collisions per bunch crossing will increase to 200.

In this condition typical tracking techniques based on hit position will fail to reliably separate particles hitting the same area. An example of a reconstructed event is shown in figure 1.2. The tracking algorithms will occur in mainly two artifacts which will compromise the benefits of the luminosity increase: miss-tracks and ghost-tracks. Miss-tracks, as the name implies, are tracks which has been detected but were not recognized because they were overshadowed by the noise floor formed by the hits of other tracks. Ghost tracks, on the other hand, are fake tracks that the reconstruction algorithm improperly reconstruct from hits belonging to other tracks. An example of a high pileup event is shown in figure 1.3.

A possible solution to mitigate the pile-up problem is to supplement the position information with a high-resolution timing measurement[44]. The time information will not only provide a new dimension for implementing the track reconstruction algorithms, but it will also reduce the observation timeframe of the single snapshot of an event. The full evolution of the event will be observed, as opposed to be integrated over a large time. The advantages of the time measurement will start to become relevant for resolution better than 100 ps. The two candidate position inside the detector for the insertion of the timing measurement are the inner tracker or a single dedicated timing layer. The first case defines the 4D Tracking[76]: a high-granularity multilayer detector close to the interaction point (with a spatial resolution in the order of 10 µm). The alternative solution is to insert an additional layer inside the particle detector specifically for the timing information [114]. The advantage of this last approach is connected to the fact that both the radiation hardness and the granularity of the detector can be reduced (with a spatial resolution in the order of 1 mm), prioritizing the time resolution.

The research activities in HEP are now focused on the development of this new generation of detectors from the point of view of the system, its sensors, and their
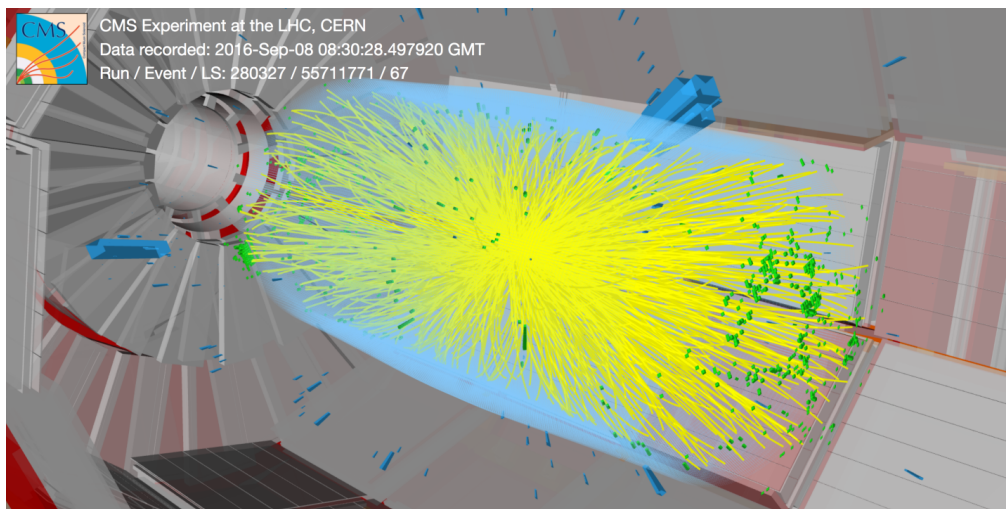
Figure 1.3: A high-pileup event with 86 reconstructed vertices observed by the CMS detector. Taken from [24].

read-out and front-end electronics. A front-end ASIC in this environment must reach the required time-resolution, and cope with the challenges related to the increased data-throughput and radiation hardness.

1.1.2 Pixel Sensors with High Time Resolution

The first requirement to develop a timing pixel detector is the availability of a high resolution sensor. A general pixel sensor consists of reversed-biased semiconductor junction [66]. Inside the sensor volume, an electric field is formed. When a particle interacts with the semiconductor lattice, it looses some energy that locally produces free charge carriers. An electrical signal is induced by the migration of these free carriers towards the sensor electrodes. In the case of pixel-sensors, the typical geometry is planar: both the electrodes are realized on the surface of the material, with the active area between them. In terms of their coupling with the front-end electronics, two variations of pixel sensors exist: monolithic detectors[103] and hybris detectors. In monolithic detectors the sensitive area is created in the same substrate of the pixel electronics, with the two sharing the total surface. In hybrid pixel-detector the sensor and the electronics are implemented in two different chips, with different process, and then coupled together via bump bonding. The two chips must feature a bump-bond matrix with equivalent pitch.

In order to adapt the conventional technology to obtain a timing detector, the sensor design must be improved in terms of its signal variability. In particular, a timing signal must be reliable and stable. The signal must provide the event timing regardless of its other properties, like interaction position or deposited charge. The strength of the field, for example, influences the drifting velocity of the carriers

4
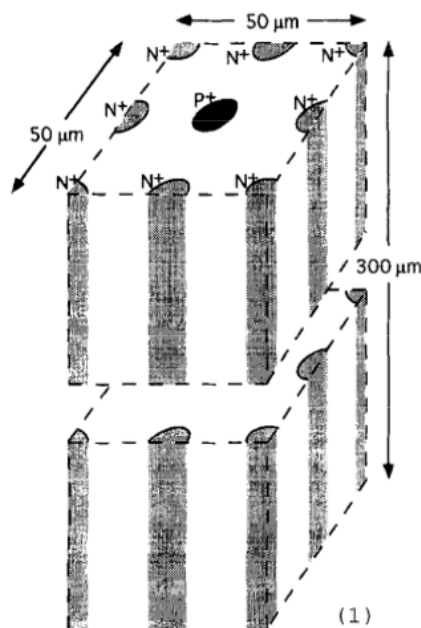
Figure 1.4: Concept of a 3D sensor. Taken from [81]

toward the electrodes. If the field is sufficiently low, the carrier will move by
diffusion, which greatly increases their collection time. In a sensor in which the
field is non-uniform, the arrival time of a signal is strongly related to the position
of the interaction point. The condition of full depletion is the one that maximizes
the field uniformity, therefore it is a characteristic necessary in timing sensors.

A sensor topology with promising timing properties is the so-called 3D sen-
sor[81]. In 3D sensors the electrodes are built in vertical columns normal to the
sensor surface. The electrodes and therefore the sensitive area, will extend deeply
on the whole bulk of the material. A schematic representation of this sensor topol-
ogy is shown in figure 1.4. The impinging particle will have a longer path to deposit
the charge, as a consequence it will deposit more charge compared to a planar sen-
sor with the same pitch. The reduced inter-electrodes distance will benefit the
timing for different reasons: it will create a more uniform field, reduce the proba-
bility of delta rays [1] and reduce the variability of the drifting path of the carriers.
The main draw back of this architecture is the larger capacitance between the elec-
trodes. Both the signal strength and the sensor capacitance will affect the total
system resolution because they influence the Signal to Noise Ratio (SNR) of the
very front-end.

In this regard, another approach to increase the time resolution of the system

---

[1]Secondary electrons produced inside the bulk that will interact with the sensor in a second
moment, or escape from the sensor entirely.

is to increase the SNR of the signal through the insertion of a gain layer inside the sensor. This is the case of an Avalanche Photo Diode (APD)[45]. In this sensor the internal gain is provided by a multiplication of the electron number due to the avalanche effect. The reason behind the development of the APD was related to the necessity to increase the detection limit for low energy signal. Due to the multiplication, however, the APD features a large noise and leakage current. By creating a sensor with lower gain, a Low Gain Avalanche photo-Diode (LGAD)[71], these two drawbacks can be avoided. With an LGAD with a gain factor from 10 to 50 it is possible to improve the system SNR and therefore its time resolution. The main disadvantage of this approach is the presence of an additional variability source connected to the multiplication process.

One last approach is to leverage the intrinsic advantage of the monolithic sensor to obtain a high time resolution. Monolithic sensors can be suitable for timing since the absence of coupling reduces the sensor capacitance and allows producing these sensors with more aggressive pitch[54]. Moreover, advancements in their manufacturing process has opened the possibility to fully deplete them. The main disadvantage of this sensor is its relatively lower radiation-hardness.

1.1.3 Pixel Front-End ASICs

Nowadays pixel front-end ASIC based detectors represent a fundamental instrument for key applications in radiation detection such as HEP, x-ray imaging, medical instruments and dosimetry. This kind of detector opens up the possibility to perform a spatially resolved measure by using a local micron-sized electronic chain. In particular, in HEP, pixel detectors have been developed and used since the beginning of the nineties[43].

The concept of an active pixel sensor comes as a natural evolution of the Charge-Coupled Device (CCD)[15] based sensor. In CCDs the charge deposited by an impinging particle is stored inside an array of MOS capacitors coupled together. The stored information can be read serially by transferring the charge from one pixel to the next, like in a shift-register. This is performed by acting on the common bias voltages. This configuration creates a limit for the maximum readout speed of a full matrix, making it impossible to observe events at a high rate which will be inevitably integrated by the CCD.

In order to overcame this limitation, the information needs to be processed locally and converted to a format more versatile from the point of view of data transmission. Therefore the basic pixel architecture must include at least the front-end electronics as well as a transmission interface. The development of such detector has become possible in the mid-eighties due to the advancement in the CMOS manufacturing process as well as the development of reliable techniques for sensor coupling via bump-bonding. From the point of view of the sensor, the timing was also ideal due to the maturity in the development of micro-strip sensors [4].
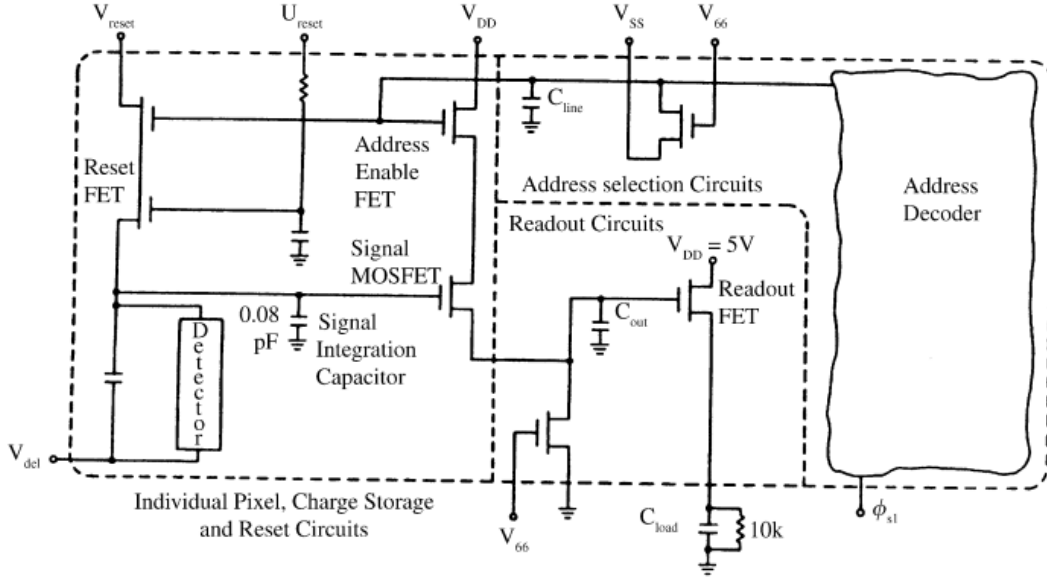
Figure 1.5: Pixel and readout architecture for the first ever pixel front-end ASIC: the readout chip for the focal plane imaging sensors developed at Hughes Aircraft Co. Taken from [34].

Micro-strips feature the desired topology of a planar sensitive area with micron-sized patterning. The sensitive area has its characteristic elongated aspect-ratio as a consequence of the need of realizing the wire-bond pads on the perimeter of the chip. Pixel sensors does not have this limitation and, moreover, the smaller size will grant a low leakage current which in turn will enhance the detector SNR. The bias voltage required to bias the sensor will also scale quadratically with the inter-electrode distance, enhancing the ease of use of the system. Moreover, another advantage of the pixel topology comes from the fact that performing imaging with strip sensors requires multiple parallel sensitive planes. The image reconstructed with coincidences of the hits on these planes will tend to become ambiguous in case of an excessive particle flow.

The first ever pixel front-end ASIC was developed in these times is the readout chip for the focal plane imaging sensors developed at Hughes Aircraft Coonment in 1984 [34]. The pixel and readout architecture of this chip is presented in figure 1.5. This ASIC was developed for x-ray imaging. The pixel architecture implements a binary front-end composed of only four transistors, it was directly connectable to the sensor via bump-bonding. The term binary front-end refers to the fact that the circuit is only capable to register the occurrence of an event, with no additional information. Each pixel integrates the signal into its capacitor whereas the interface is used to both transmit the state of the pixel and reset its stored information.

Following this chip, a research effort was carried out in the HEP field in order
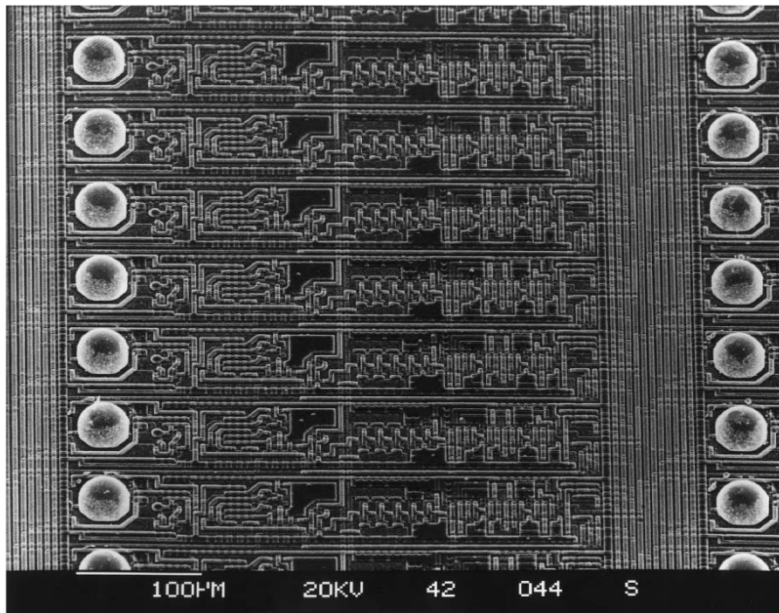
Figure 1.6: SEM image of the Omega2 chip. For reference, the bump-bond ball has a diameter of 38 µm. Taken from [43].

to develop a suitable pixel ASIC for the RD19 project[12]. This R&D produced the Omega ASIC family, with Omega2[22] being the first large-area pixel ASIC for HEP in 1994. The chip has been developed using a 3 µm Self Aligned CMOS technology. This ASIC was designed to cover an area of 5 cm × 5 cm with each chip having 1006 pixels of 75 µm × 500 µm organized in a matrix of 64×16 pixels. A Scanning Electron Microscope (SEM) imagine of a set of Omega2 channels is shown in figure 1.6. The Omega front-end is binary with a continuous sensitive input amplifier and an asynchronous discriminator with a total power consumption of just 30 µW per channel. In this way the channel does not latch the information, but instead it registers it on a dedicated register which is read via an external trigger. The external trigger is provided as a per-column strobe signal. By synchronizing the delays of the strobe along the column, it is possible to read the column state with a 16 bits serial bus. The initial timing requirement in HEP was to be able to read the data inside a 25 ns period in order to assign the event to the correct LHC bunch crossing which is operated at 40 MHz. Table 1.1 shows a comparison between the ASICs operated in experiments in the first years of this field.

In the following years down to the present day, pixel ASICs have evolved iteratively by integrating more features both at pixel and chip level in order to achieve cutting-edge performances in radiation detection. The advancement on the sub-micron CMOS process has been critical for these progresses, enabling the integration of more features in the pixel area. The latest ASICs in the field are being developed in 65 nm and 28 nm. Front-ends can be realized in more complex

Comparison of pixel detectors operated in experiments

	WA97/NA57 (RD19)	DELPHI VFT	NA50
First year	1994	1996	1997
Last year used	> 2000	2000	1998
# kpix	1093	1226	67
Area (cm^2)	372	1335	18
# Chips installed	792	2432	$\sim$ 200
# Readout wafers	$\sim$ 60 $\varnothing$ 100 mm	nn $\varnothing$ 100 mm	none
	$\sim$ 50 $\varnothing$ 150 mm		(From WA97)

Table 1.1: Review table of the most important pixel front-end ASICs operated in the first years of LHC presented in [43]

architectures than the binary one in order to extract more information from the incoming particle, such as the deposited charge and the precise time of hit arrival. Thus, more information on the particle energy and momentum can be inferred, enhancing the overall quality of the event reconstruction. The measured value can be transmitted to a readout circuit in an analog format or directly digitized at the pixel level. Moreover, a local memory can be integrated inside the pixel, making the whole pixel matrix an event buffer. In this way the readout mode can divert from a trigger based one to the one most suited for the application. The active pixel can also process and filter the local signal in order to compress it or discard spurious events. In this way the readout circuit is offloaded of these tasks, enhancing the ASIC and thus the detector throughput. Finally, the ASIC can be made more flexible by integrating configurable circuits; and more autonomous by adding service blocks for local reference generation.

The next section illustrates various typical structures in pixel front-end ASICs that are commonly found in this type of chips.

1.2 Architectures and Concepts

This section presents the architecture and concepts typical of a pixel front-end ASIC. The aspects outlined in this section are of a general nature, they represent practical implementations depending on the application. The general architecture will be firstly described starting from the electronics chain connecting the sensor to the ASIC output. After this, the section will describe the topological implication of implementing this chain in a pixel-matrix structure. The end of the section presents the measurable quantities and the data format and readout modes used to transfer the measured information outside the chip.
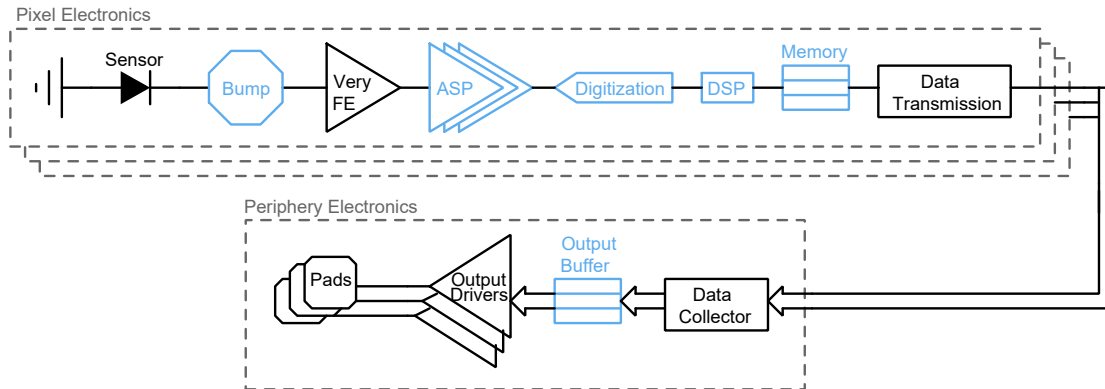
Figure 1.7: Schematic representation of the general electronics chain of a pixel front-end ASIC. The scheme follows the signal and data path. Optional blocks are drawn in light blue.

1.2.1 The Electronics Chain

A pixel front-end ASIC solves basically two functions: processing at pixel level the signal coming from a sensor and transmitting the data generated from the pixel matrix to outside the chip. From the point of view of the signal and data path, the architecture consists of the following elements:

1. The Sensor: this is the area in which the sensor is realized in monolithic detectors. Depending on the technology, it can also be possible to integrate the front-end transistors inside this area.

2. The very front-end: it consists in the first stage of the electronics chain which is directly connected to the bump-bond pads in case of an hybrid detector, or more generally, to the sensor. It has the function of sensing the input signal and adapting it to the next stage.

3. An Analog signal processing block (optional): it is the set of analog stages which alter the signal to properly suit the measurement process. Some examples of possible stages can be a filter for noise reduction, a shaper stage or a simple discriminator used to produce binary signals. The discriminator also solves the function of rejecting uninteresting signals.

4. A Digitizer (optional): in this stage the signal is converted to the digital domain. This stage is present whenever the readout interface is digital. In case of binary front-ends it can consist of a simple buffer or latch, but in case of more complex architectures it can include converters such as an Analog to Digital Converter (ADC) for amplitude measurements or a Time to Digital Converter (TDC) for direct time measurement. Ideally, it is also possible to insert a full signal sampler at this stage.

10

5. A Digital Signal Processing (DSP) (optional) block: it can be used as in point 3 to modify the signal properties depending on the needs. For example, it can be used to enhance the signal SNR, filter spurious events or compress the data. Performing these operations locally can increase the data throughput and offload the subsequent data processors of some tasks.

6. A local memory (optional): it can be used to store locally the data produced by the front-end. This can consists in a single event memory for a basic trigger-based readout. The depth of this memory is greatly constrained by the available space on the pixel area. Usually at least a simple buffer memory is required due to bandwidth bottlenecks on the data transmission interface.

7. The local data transmission interface: this block includes all the circuits which operate the data transmission from the pixel to the rest of the chip. It can consist of a serializer or deserializer if the data are stored in a digital format, or a form of line driver in case of analog data. In any case, it must provide a form of address identification and, in case of shared busses, a method for arbitration. This stage can also implement additional features such as zero suppression, signaling the presence of event pile-up or referring the data to a specific time reference.

8. The data collection interface: this block collects the data from many pixels in order to distribute them to the output buffers. The data words coming from this stage must include a geographical address and an eventual time reference. Both these functions must be provided by this block. If the transmission data bandwidth is insufficient, a de-randomization step is required. This can be implemented, for example, with a simple First In First Out (FIFO) stage.

9. An output buffer (optional): the data formatted by the collector can be stored in this memory. This stage is necessary in cases in which the output driver bandwidth is insufficient. Again, if an arbitration of the output resources is required, a de-randomization step must also be inserted.

10. The output transmission drivers: these are the transmitters responsible for outputting the data generated by the whole ASIC. The driver must sustain the target analog bandwidth while driving the external impedance. Their number is usually limited by the pads availability or the power consumption budget. Output drivers can be realized according to various standards such as: full CMOS signaling, Low-Voltage Differential Signaling (LVDS), Scalable Low-Voltage Signaling (SLVS), Current Mode Logic (CML), etc.

A schematic representation of this electronics chain is presented in figure 1.7.

1.2.2 Topology

From a topological point of view, these blocks must be organized to accommodate the data flow from the matrix configuration of the sensor to the output drivers. It is evident that every circuit described from point 1 to 7 must be implemented locally, near the sensor or bump-pad; whereas the others are bounded to be integrated in proximity of the output connections.

Typically, the output drivers are connected to the outside of the chip via wire-bonding. Wire-bond pads are usually realized on the perimeter of the chip, and in some cases, they can be realized in multiple parallel lines. The number of parallel lines is limited by the difficulty to connect the wire-bonds to the innermost pads. Moreover, in the case of a hybrid detector, the chip top surface is already occupied by the bump-bond pads. For these reasons, it is usually practical to create at most two parallel lines of staggered pads. Redistributing the data in such configuration means interfacing a planar topology to a linear one. This creates an intrinsic problem in the ASIC floor-plan: the number of pixels will scale quadratically with the chip size, whereas the number of output drivers will scale linearly. As a consequence, large-area ASICs will suffer from a congestion in data lines and a bottleneck in the data redistribution bandwidth.

Moreover, the pixel matrix area must accommodate not only the pixels own electronics but also the data lines. If the data lines consist of per-pixel dedicated buses, the congestion will grow from the innermost pixel to the external ones. This problem can be circumvented by having a common bus or a daisy-chain connection between the channels. The first solution will, however, increase the data transmission latency requiring more bits to refer the event to the correct reference period (time-stamping), thus reducing the overall maximum processable event rate. In the same way, the second solution will require some form of arbitration which will increase again the average latency and reduces the data bandwidth to the one of the shared bus.

The floor-plan must also accommodate for redistribution of the power net, the ground net and all the reference signals used to operate the individual pixels. These references can be analog voltage levels and current biases or digital configurations, clock nets and dynamic digital signals (e.g. a trigger). The path of these nets will follow inversely the data path, encountering the same geometrical problems and bottlenecks. Therefore, large-area pixel ASICs shows a challenging power distribution. Additionally, the majority of these references can be generated inside the chip using dedicated blocks. For instance, the analog voltages and currents can be provided by integrating a dedicated Digital to Analog Converter (DAC) coupled with a Band-gap, while the reference clock can be generated internally and phased with an input one using a Phased Locked Loop (PLL). The area not dedicated to the pixel must also accommodate the necessary configuration registers.

By taking into account the aforementioned considerations, the general topology
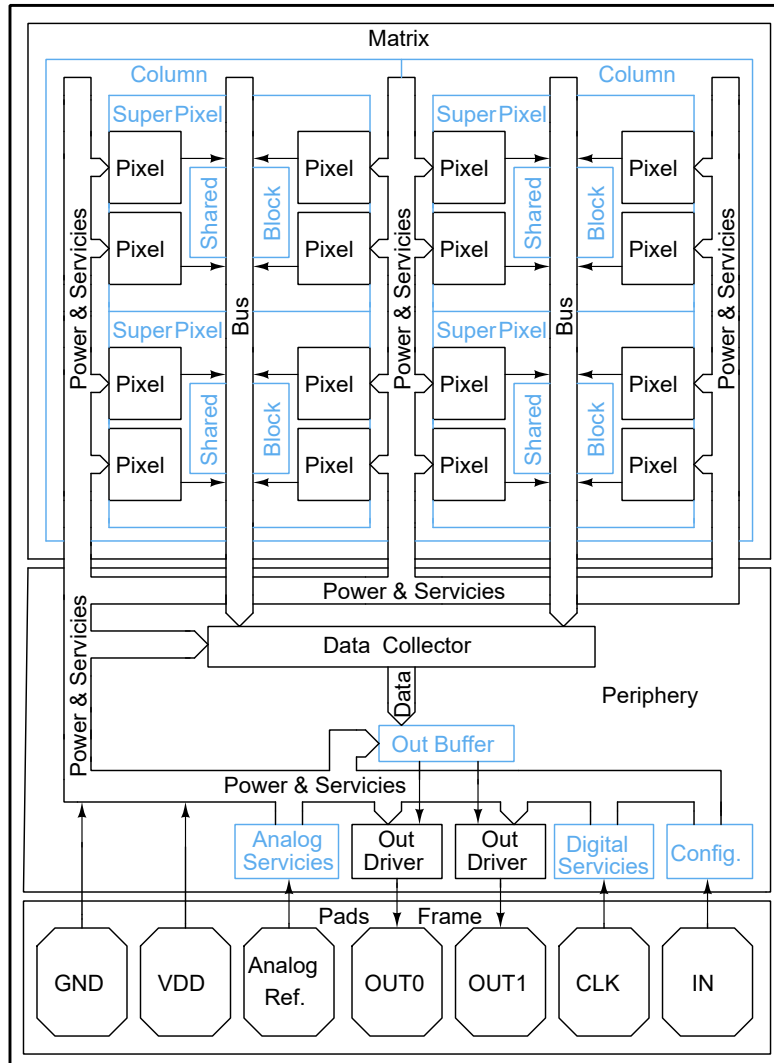
Figure 1.8: Example of a possible floor-plan of a pixel front-end ASIC. The scheme is not in scale. In this example the pixel matrix is composed of 4×4 pixels grouped together in two double-columns of 2×4 pixels, every 2×2 pixels shares part of their electronics in a super-pixel. Optional blocks are drawn in light blue.

of a pixel ASIC with IO and power connection realized with wire bonding consists of the points below. The structures are ordered in a bottom-up order from the pixel to the chip.

1. Pixel: the individual repeated unit connected to the sensor. Each pixel can include all the circuits described in the previous list from point 1 to 7 and some local services. Some examples of local services can be: local configurations or per-pixel calibration circuits.

2. Super-pixel (optional): a set of pixels sharing part of their electronics. Sharing electronics enables to save chip area and power at the cost of the arbitration of these resources.

3. Column or row (optional): a repeated structure of pixels comprising their transmission lines as well as the reference and power redistribution. They usually identify a subset of geographical addresses. The column or row name is conventional and interchangeable: the name comes from the convenience of building them in a straight path. Often a double column configuration is preferred since it enables to integrate the data and power paths in a shared area.

4. Matrix: the whole set of pixels with their interfaces.

5. Periphery: the set of circuits connected to the matrix which are used for IO purposes, configuration and services. Its portion directly connected to the columns is often called End of Column (EoC). The number of these units does not scale with the number of pixels.

It must be noted that inside a single ASIC both the matrix and the periphery can be subdivided into some repeated blocks, making them appear as sub-matrices from the rest of the system.

Most of the topological problems described above can be avoided by realizing the IO and power interconnection leveraging the whole ASIC area. Using Through Silicon Via (TSV)[75], for instance, it is possible to create a dense pattern of vertical interconnections from the chip bottom face down to the device level. In this way, the ASIC can be connected with a matrix of bump-bonds to the sensor and with a matrix of TSVs to the carrier board. Moreover, with 3D integration, it is theoretically possible to cover a large area by tiling together multiple chips without dead-area. Without the need of redistributing the nets to the perimeter of the chip, it is ideally possible to create a periphery-less ASIC in which all the IO and service blocks will be integrated seamlessly inside the pixel matrix.

1.2.3 Readout Modes, Measurable Quantities and Data Formats

The pixel ASIC can be seen as a data source with a specific mode and protocol. In each possible implementation of the protocol, the possibility that a hit arrives when the dedicated memory is full must be foreseen. Some measures are thus applied to address the data-loss: the new event can be ignored or one of the stored data can be overwritten. Additionally, a pile-up flag can be raised for signaling the occurrence of this event. Three major categories of readout modes can be identified:

- Frame based: the state of the ASIC is continuously readout for the entire matrix. A frame is a single snapshot of the entire matrix. The refresh-rate identifies the rate at which a full frame can be acquired. This readout mode requires a straightforward readout electronics and is suitable for applications in which the available memory is sufficient to buffer all the hits arriving during the refresh period.

- Trigger based: the data is stored inside the ASIC waiting to be readout. An external trigger signal initiates the acquisition of the stored data for the entire matrix or a subset of it. The maximum sustainable per-pixel hit-rate is limited by the time required to read the full trigger set. Therefore, a trigger based readout is not suitable for an application which involves a constant monitoring of a high hit-rate system. The main advantage of this mode is the possibility to discard spurious events a-priori, without overloading the following readout by leveraging the trigger system.

- Data driven: individual pixel data are continuously transmitted off-chip as they are produced. As described in the previous section, it is unrealistic to expect a per-pixel transmission line with sufficient bandwidth to serve the full matrix. For this reason, this type of system has heterogeneous transmission lines which involve arbitration and de-randomization. The readout of this type of ASIC tends to be more complicated compared to the previous ones, but offers the most flexible solution. This readout-mode is most suited for sparse events with unpredictable or uneven space and time phases.

The relation between the impinging particle properties and the measurable quantities is strongly related to the operating principles of the sensor. As explained in section 1.1.2, the typical sensor response is a current signal induced by the drift of the charges deposited by a particle from the deposition point to the sensor electrodes. More specifically, the particle deposits the charge in multiple points along its tracks. The deposited charge, and therefore the integral of the current signal, is proportional to the kinetic energy lost by the particle during the interaction. Even though the charge represents the primary information, additional information on the particle-sensor interaction can be derived from the time development of the signal. In any case, the pixel structure will naturally provide the spatial information by the means of the centroid of the activated pixels. From the point of view of the pixel electronics the measured quantities can be classified as:

- Particle passage: the occurrence of an event during a time period. This quantity represents the most basic information on the interaction and can be provided by a simple binary front-end. It must be noted that even in the most simple front-end, a discrimination process will take place; otherwise the front-end be triggered inevitably by noise. As an indirect effect of the noise

discrimination, the front-end will still produce an indication that the signal amplitude exceeds a certain amplitude. Moreover, by integrating a local counter, the number of occurrences on a certain time frame can be provided.

- Charge: the deposited charge can be retrieved by integrating the whole signal. This can be done simply using a capacitor or an active Charge Sensitive Amplifier (CSA) stage. The integrated amplitude can be digitized by with a local ADC. The digitization process must be triggered after the full integration time. In systems featuring a discrete-time discharge it is trivial to implement this trigger. To clarify, a discrete-time discharge front-end is typically implemented with a discharge switch is parallel to the integrating capacitor, the switch is toggled periodically in order to transit from an infinite resistance loop (integration) to a zero resistance one (discharge).

  A smart solution for charge measurement is to leverage a TDC to perform it. In particular this can be done in systems in which there is a monotonic and reliable relation between the integrated amplitude and the signal time development. The simplest example is the case of a CSA with constant current discharge. If the input signal time development is shorter than the CSA integration time, the total signal length is proportional to its amplitude and therefore to the input charge. By using a discriminator with a programmable threshold it is possible to produce a digital pules with a certain Time over Threshold (ToT). By digitizing it, the charge information can be acquired.

- Time: this measurement can be performed by directly measuring the timing characteristics of the very-front signal after a discrimination process. The key parameter is represented by the Time of Arrival (TA) of the signal, which is the measure of the elapsed time between the event to be measured and a reference signal. Typically the rising or falling edge of these two signals is used to initiate and end the measurement. Another parameter is the ToT.

  The analog front-end has only the function of providing the signal to be digitized. If it is implemented with a Trans Impedance Amplifier (TIA) with sufficient bandwidth, the timing characteristics of the signal will remain unchanged. However, if a form of integration takes place the TA will depend on the input charge. The limit case is represented by the CSA in which the rising time is independent from the signal charge, whereas the slew-rate and therefore the TA will depend inversely on the input charge. This phenomenon is named time walk: the drifting of the TA measurement as the charge decreases. The time-walk represents a systematic effect which can be corrected by knowing the input charge. A convenient time-walk correction method can be based on the ToT. An alternative solution to the time-walk correction is to utilize Constant Fraction Discriminator (CFD)[35][68] combined with a CSA which automatically moves the threshold in order to contrast the time walk.

The digitization is carried out using a TDC which can be implemented in various architectures. The most basic TDC can be implemented by counting the number of oscillations of a local Voltage Controlled Oscillator (VCO)[113] or Digital Controlled Oscillator (DCO)[84], which are elapsed during the time to be measured. Higher resolutions can be obtained by using a delay chain with the purpose of registering its state at the end of the measuring process. In this way the resolution limit is determined by the delay of a single element. This last configuration can be looped in a Delay Locked Loop (DLL) combining both of the two aforementioned methods[84]. Furthermore, using a Vernier configuration, two parallel lines of this kind can be used to encode the information on the phase difference between them[21][52]. A time amplification technique can also be used to increase the time to be measured[72][73], which is then digitized using one of the previous architectures. Another completely different solution is the pulse-shrink architecture [20][47]: the input pulse is iteratively shirked until it surpass the detection limit; the number of iterations provides the measurement. Finally, the timing characteristic of a signal can be directly converted to an analog amplitude using a Time to Voltage Converter (TVC) [79], and then digitized with an ADC.

- Amplitude: this parameter can be measured at the output of the pre-amplifier. In the case of a TIA, the relation between its output voltage and the input current amplitude is linear and well known. The digitizer (e.g. an ADC) must sample the signal at the correct point of its development, therefore the front-end must also provide a signal that triggers the digitization. For example, the peak of a signal can be triggered using a differentiator stage in parallel to the main channel. Similarly, a CFD can trigger the digitization of a certain relative threshold.

- Full signal development: the full signal can be recorded using a Sample and Hold (S/H) process. This measurement is the most power and area consuming. Since it samples the signal it produces a really flexible information at the cost of a higher bit length per event. Moreover, it requires a S/H implementation with a sample rate adequate to measure the signal. For these reasons, it is often unfeasible to insert a full S/H circuit in a pixel front-end for HEP.

The typical data format of a word produced by a pixel front-end ASIC with digital readout is composed of:

- Measurement: the output bits of the digitizer corresponding to the measured quantity described in the previous paragraph.

- Address: an index or couple of indices that identify the pixel position.

- Time-stamp: an identifier, usually an integer, which correlates each measurement with the reference clock period in which the measurement took place. This procedure constitutes a form of coarse time measurement. Time-stamping may be required in cases in which the data cannot be readout within the period of the reference clock.

- Additional information: additional bits are often reserved for some utilities. For example, these bits can be used for the signaling of a pileup event, as described above.

1.3   Pixels with Timing

This section explains the research challenges concerning the addition of a high resolution time measurement to a typical pixel detector. The topic will be introduced by analyzing the general components which affect the time resolution in this type of detector. After this, the section will discuss the challenges of implementing a high time resolution electronics in the specific case of a pixel front-end ASIC.

1.3.1   Resolution Contributions

The resolution of a timing system can be quantified with the standard deviation of a set of repeated measurements of different events occurring at the same time. Therefore the timing system must be insensitive to the other properties of the event such as its charge. Using the probabilistic definition of standard deviation and imposing the average value equal to zero, the timing resolution becomes:

$$\sigma_{t,res} = \sqrt{\int_{-\infty}^{+\infty} P(t) t^2 \, dt} \qquad (1.1)$$

where $P(t)$ is the Probabilty Density Function (PDF) of the measured time t.

The most general system is composed of a sensor, a very front-end and a digitizer (see 1.2.1). This kind of system is able to produce a digital code corresponding to the time of arrival of a certain event. As explained earlier, the function of the very front-end is to adapt the sensor signal to the digitizer input. Considering these contributions to be independent one from the other, the resolution can be expressed as:

$$\sigma_{t,res} = \sqrt{\sigma_{t,sens}^2 + \sigma_{t,ana}^2 + \sigma_{t,digi}^2 + \sigma_{t,corr}^2} \qquad (1.2)$$

where $\sigma_{t,sens}$ is the sensor intrinsic contribution ot the total resolution, $\sigma_{t,ana}$ is the contribution from the very front-end and $\sigma_{t,digi}$ is the contribution from the digitization process.

The last term, $\sigma_{t,corr}$, represents the error related to an eventual digital correction of systematic effects. This is not part of the core time measurement process, but it is able to enhance the overall measurement resolution. The tradeoff of correcting a measurement is connected to the error generated from this operation.

Additionally, by multi-sampling the same event, the resolution can be further enhanced. In particular, all the random contributions non-related to the sensor variations ($\sigma_{t,ran}$) can be reduced with the square root of the number of samples N_s:

$$\sigma_{t,res} = \sqrt{\sigma_{t,sens}^2 + \frac{\sigma_{t,ran}^2}{N_s} + \sigma_{t,sys}^2} \tag{1.3}$$

Where $\sigma_{t,sys}$ are all the systematic contribution to the time variation. Again in this case, the systematic component can be eliminated using a digital correction:

$$\sigma_{t,res} = \sqrt{\sigma_{t,sens}^2 + \frac{\sigma_{t,ran}^2}{N_s} + \sigma_{t,corr}^2} \tag{1.4}$$

All the components will be now analyzed one by one.

Sensor contribution

The sensor current signal can both vary in its total charge and in its time development. To be accurate, the whole signal shape will inevitably affect its timing. It is convenient to disentangle the two components since the sensor acts as transducer for the energy lost by the particle in the interaction to a charge. Therefore $\sigma_{t,sens}$ accounts for all the non charge related signal variations. The major time variability source is related to the geometrical non-uniformity of its field. This component is in principle independent of the impinging particle charge, and it is related only to the interaction position and angle of incidence. Therefore it can be assumed as timing variability on the charge generation $\sigma_{t,gen}$. Another component can be introduced by an eventual gain layer $\sigma_{t,gain}$, and it is connected to the characteristic time of the process. In any case the two components can be assumed to be independent, resulting in a total sensor contribution expressed as:

$$\sigma_{t,sens} = \sqrt{\sigma_{t,gen}^2 + \sigma_{t,gain}^2} \tag{1.5}$$

The distribution of the first term is complex and it depends on the detector and sensor geometries, whereas the second one typically follows a Landau distribution.

The other source of variability is the quantity of charge deposited by the particle q_i. This quantity will vary with a certain PDF $P_{dep}(q_i)$ which is directly connected to the radiation-matter interaction. $P_{dep}(q_i)$ can be assumed to be a Landau distribution. Even though this component will not affect the intrinsic timing properties of the sensor ad its signal, it will in turn influence the response of the very front-end.

Very front-end contribution

The very front-end response to the input signal can be generally expressed as a function of the sensor current signal $i_i(t)$. For the purpose of describing only the very front-end contribution to the total timing the charge development is considered to be instant, and therefore the input signal will be represented with a delta-like pulse with a certain charge:

$$i_i(t) = \delta(t)q_i \tag{1.6}$$

Where q_i is the total deposited charge. The front-end output t_{ana} will be a time value representing the arrival time of the input signal. Its operation can be generally represented as:

$$t_{ana} = T_{q \to t}(q_i) + \chi_{jit}(q_i) \tag{1.7}$$

where $T_{q \to t}$ is the charge to time transfer function of the analog processor and $\chi_{jit}(q_i)$ is a stochastic term representing the output jitter contribution added by the circuit intrinsic noise. The decision to represent this stage as a charge dependent one is not related to the very front-end own architecture, but again only a mean of separating the input variations from the ones introduced by this stage. In this regard, $T_{q \to t}$ can assume any kind of non-linear form. Going into detail, a timing very front-end will convert through a two stage process: firstly it will convert the signal into an analog amplitude signal $a_o(t)$ (namely a current or a voltage) and then it will convert it to a digital pulse with a well-defined transition. In this way $T_{q \to t}$ can be decomposed as the convolution of the two transfer functions:

$$a_o(t) = T_{q \to a}(\delta(t)q_i) \tag{1.8}$$
$$t_o = T_{a \to t}(a_i(t)) \tag{1.9}$$
$$t_o = T_{q \to t}(q_i) = [T_{a \to t} * T_{q \to a}](q_i) \tag{1.10}$$

$a_o(t)$ is the time development of the front-end output signal and $T_{q \to a}$ is the transfer function which relates it to the input charge development. t_o is the output timing signal and $T_{a \to t}$ represents the discrimination process. It is safe to assume that $T_{q \to a}$ is monotonically increasing and $T_{a \to t}$ is monotonic decreasing.

As a result, the first term of (1.7) will generally produce a time-walk effect. The most probable value of the output time $\hat{t}_{ana}$ can be computed on the basis of $P_{dep}(q_i)$:

$$\hat{t}_{ana} = \int_0^{+\infty} P_{dep}(q_i) \cdot T_{q \to t}(q_i) \, dq_i \tag{1.11}$$

This value is computed in the whole charge range, even though $P_{dep}(q_i)$ is non-zero only in a limited range. Likewise, the standard deviation of this parameter is computed as:

$$\sigma_{tw} = \sqrt{\int_0^{+\infty} P_{dep}(q_i) \cdot (T_{q\to t}(q_i) - \hat{t}_{ana})^2 \, dq_i} \tag{1.12}$$

This term can be interpreted as the time-walk dependent component of the analog electronics time variation. In general the PDF resulting from this conversion $P_{tw}(t)$ is a transformation of a random variable [101] from the charge to a time:

$$P_{tw}(t) = P_{dep}(T_{q\to t}^{-1}(t)) \cdot \left| \frac{d}{dt} T_{q\to t}^{-1}(t) \right| \tag{1.13}$$

Therefore it has a shape related to $P_{dep}(q_i)$.

Moving to the front-end intrinsic jitter, it can be usually expressed as:

$$P_{jit}(t|q_i) = \phi(t|\sigma_{fe}(q_i)) \tag{1.14}$$

where $\phi(t|\sigma_{fe}(q_i))$ is a zero centered Gaussian PDF defined as follows:

$$\phi(x|\mu,\sigma) = \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \tag{1.15}$$

This distribution is related to the PDF of the electronics intrinsic noise, in which the main contributor is the thermal noise [50]. This component can be altered by the presence of other noise components (such as $1/f$) or by the presence of noise filtering stages. The standard deviation of the distribution can be generally connected to the stage output noise as σ_{na}:

$$\sigma_{fe}(q_i) = \frac{\sigma_{na}}{a_o'(q_i)} \tag{1.16}$$

In which $T_{q\to a}'(q_i)$ is the derivative of the charge to amplitude transfer function at the sampling point:

$$a_o'(q_i) = \frac{d}{dt}[a_o(t)] = \frac{d}{dt}[T_{q\to a}(\delta(t)q_i)] \tag{1.17}$$

The amplitude noise is converted to a time one by the means of amplitude signal transient $a_o'(q_i)$ (namely its slew-rate) which connects the magnitude of this variation to the input charge. Since the resulting distributions will be all centered in zero, the standard deviation of the jitter across the whole input range can be computed as:

$$\sigma_{jit}(t) = \sqrt{\int_0^{+\infty} \sigma_{fe}^2(q_i) P_{dep}(q_i) \, dq_i} \tag{1.18}$$

In order to compute the total PDF of the two components, it is helpful to write σ_{fe} as a function of the output time using $T_{q\to t}(q_i)$:

$$\sigma_{fe}(t_0) = \frac{\sigma_{na}}{T'_{q \to a}(T^{-1}_{q \to t}(t))} \tag{1.19}$$

Using this expression in (1.14) in combination with (1.13), the total PDF $P_{t,ana}(t)$ can be computed with:

$$P_{t,ana}(t) = \int P_{jit}(t - t_0 | \sigma_{fe}(t_0)) \cdot P_{tw}(t_0) \, dt_0 \tag{1.20}$$

The resulting distribution is an interplay of the two contributions, in which the total standard deviation can be computed using the general relation:

$$\sigma_{t,ana} = \sqrt{\int_{-\infty}^{+\infty} P_{t,ana}(t) \cdot (t - \hat{t}_{ana})^2 \, dt} \tag{1.21}$$

It can be easily verified that:

$$\sigma_{na} \to 0 \Rightarrow \sigma_{t,ana} \to \sigma_{tw} \tag{1.22}$$
$$\sigma_{tw} \to 0 \Rightarrow \sigma_{t,ana} \to \sigma_{jit} \tag{1.23}$$

The first case represents a very front-end with no intrinsic noise, whereas the second represents a very front-end with no time-walk. The second case can actually be found in a system in which the time-walk is actively compensated with a discrimination process insensitive to the time-walk (for example with a CFD). In those cases $T_{q \to a}$ and $T_{a \to t}$ contrasts each other cancelling their respective distributions.

In summary, the best very front-end for timing must firstly minimize its output jitter. This can be done by minimizing the stage intrinsic noise while maximizing the time derivative of its amplitude. The time-walk component can be minimized by choosing a charge to time transfer function which minimizes the dependency from the input charge distribution. This can be achieved at the amplitude level by having a steeper conversion relation or with a limited bandwidth system. From this point of view, the jitter and time-walk components can be reduced together. Finally, the time walk can be also be limited with specific circuital solution on the discrimination process.

Digitization contribution

The digitizer solves the function to convert the analog time signal to a digital code. It can use various mechanisms (usually a TDC), all of them consist in quantizing the timing interval between the input signal and a certain reference. The quantization process will determine an intrinsic source of variability σ_{quant} correlated to its quantization error. The timing reference will also present a certain fluctuation around its mean value σ_{ref}. Finally, the whole process will feature a random variation correlated to its Differential Non-Linearity (DNL), which will be named σ_{DNL}. All these components can be assumed to be independent one to

the other because: the first one is a statistical consequence of the quantization, the second one is originating from the reference, and the third one is intrinsic to the digitizer. So, to summarize, the digitization related component $\sigma_{t,digi}$ can be expressed as:

$$\sigma_{t,digi} = \sqrt{\sigma_{quant}^2 + \sigma_{ref}^2 + \sigma_{DNL}^2} \tag{1.24}$$

$$= \sqrt{\left(\frac{LSB}{\sqrt{12}}\right)^2 + \sigma_{ref}^2 + \sigma_{DNL}^2} \tag{1.25}$$

In the last expression, the quantization variation has been expressed as a function of the digitizer Least Significant Bit (LSB). It must be noted that the digitizer can introduce a systematic error in the case its response is non-uniform inside the range.

In order to minimize the digitization component, it is equally important to find a reliable digitization process which employs a low-jitter reference. while minimizing the LSB.

Digital correction contribution

This optional stage is used to correct and eliminate the systematic components introduced by the previous stage. This correction is digital from the point of view that it acts on the digitizer output. Usually this requires an auxiliary measurement of another quantity x, with its related variation σ_{aux}. Then, the error is corrected using a correction function $f_c(x)$ which outputs the correction code. The granularity of this code is ultimately limited by the LSB of the time measurement. Assuming, for sake of simplicity, that $f_c(x)$ will simply convert σ_{aux} in the time domain, the total correction contribution can be written as:

$$\sigma_{t,corr} = \sqrt{f_c^2(\sigma_{aux}) + \left(\frac{LSB}{\sqrt{12}}\right)^2} \tag{1.26}$$

An example of this kind of correction is reducing the time-walk by measuring the input charge and using its inverse input-output relation as the correction function.

1.3.2 Goals and Challenges

Research and development efforts in the field of Front-End ASIC for high resolution space-time measurements are driven by the specifications outlined in the upgrade plans of major high energy physics experiments [107][106][108]. Space and time resolutions represent the key requirements for detectors. The values range from 100 ps to 30 ps for the time resolution and down to 20 μm for the spatial one. Another important factor is related to the rate of event that the ASIC is able to

process and transmit to the read-out system. This aspect defines the maximum hit rate sustainable by the ASIC for a short period of time and the average hit-rate that it can process indefinitely without loosing data. These performance levels must be achieved while respecting the specifications of power consumption posed by cooling and power-delivery systems. This specification translates to a power consumption per unit area ranging from $0.1\,\mathrm{W/cm^2}$ to $1.3\,\mathrm{W/cm^2}$[23]. Lastly, the architecture must be sufficiently radiation tolerant in order to sustain a radiation dose compatible with $250\,\mathrm{fb^{-1}}$ per year of integrated luminosity.

The interplay of these specifications imply to various challenging trade-offs that must be overcome. The spatial resolution can be improved by reducing the pixel pitch. The absolute limit is dictated by the smallest area which allows integrating the desired features. However, the increase in channel density will affect the power consumption per unit area and produce more local data that will tend to bottleneck the data redistribution net. In turn, a larger pixel will intercept more hits for a given time, increasing the amount of signals that must be processed by the single channel. Moreover, depending on sensor type and geometry, the sensitive area size will affect signal strength and input impedance, which will in turn affect the very-front-end SNR.

Another challenge is related to improving the time resolution. The two main factors in the design of a high-resolution timing front-end are the input signal SNR and the available power budget. Analog circuits SNR and bandwidth as well as digital circuit operating frequency are key values for the digitisation of the timing properties of the signal. As explained earlier, these aspects are tightly related to the channel area and thus to the sustainable hit-rate. Another aspect to point out is that in order to reach the desired time resolution, additional circuits for calibration and correction may be required. These additional features will take more pixel-area. Overall, the increase in pixel density, complexity and precision will require longer data words that will in turn affect the ASIC throughput. One last challenge is posed by the radiation hardness of the architecture. Radiation tolerant techniques such as radiation-aware-layout and Triple Modular Redundancy (TMR) come with their area and power costs.

Since the particle tracking systems currently in use are not based on space-time measurement, front-end ASICs have not yet met the requirements mentioned above. The different r&d projects focus on different specifications tied to the target experiments or role inside the vertex detector. Timing front-end ASICs can be roughly classified in terms of their granularity and sensor type. High granularity ASICs are characterized by a pixel pitch in the scale of tens of micrometers, whereas low-granularity ones will target resolution around one millimeter. The sensor type will also affect the input signal characteristics and the input impedance. It defines the minimum achievable pitch and radiation hardness level. Monolithic sensors, due to the absence of wafer coupling, exhibit a lower input impedance and can be realized in tiny form factors. The downside of this approach is a comparatively

lower radiation-hardness. Hybrid sensors enable to couple a CMOS front-end chip with a sensor matrix built with different processes and technologies. This will enable to connect sensors made in materials alternative to silicon, such as diamond. Additionally, if required, the ASIC can also incorporate a method to compensate the signal time walk such as the usage of a constant fraction discriminator or a time over threshold (ToT) correction. This second method would require to implement ToT measurement in addition to the Time of arrival (TA) one.

1.4   Review on Timing Front-End ASICs

This section presents an overview of the state of the art in the field of pixel front-end ASIC with timing. Some key ASICs are illustrated in details below. The table 1.2 summarizes the measured results of the specifications on the major ASICs researched in the field. The table is shown at the end of the section. Timespot1, the pixel ASIC developed during the PhD research described in this thesis, is also present for comparison.

1.4.1   Timepix 3

Timepix3 [90] is an ASIC developed in 2013 by the Medipix collaboration [69][8] as a successor to Timepix1 [60]. Like its predecessor, this ASIC is aimed at applications of X-rays imaging, particle tracking and space dosimetry. It implements the front-end electronics for a hybrid detector based on silicon or gas sensors. The ASIC is implemented in a 130 nm CMOS, and it aims to increase by a factor six the time resolution of its predecessor. Timepix3 has been built to be connected to a test board by using both conventional wire-bonding and TSV.

The ASIC can operate in two modes: data-driven and trigger based. The data-driven readout was selected in favor of the frame-based of its predecessor since it enables to achieve a fast readout time for the desired pixel occupancy. Timepix3 has been developed, in fact, for a sparse-readout occupancy corresponding to a hit-rate per unit area of 40 Mhits/s/cm². The acquisition can be set in three modes: charge (ToT) and time (TA), time only (TA) and event counting. A zero-suppression mechanism is also adopted. The triggered readout mode is used for the event counting acquisition, each double-column can be triggered individually.

The pixel matrix is composed of 256×256 pixels with a 55 µm × 55 µm pitch. The pixels are organized in double-columns which are in turn segmented into super-pixels formed by 2×4 pixels. Each super-pixel integrates the control and data transmission logic which is shared among the 8 pixels. In this way the pixel area is utilized in a more efficient way. The layout of the super-pixel is presented in figure 1.9 while a schematic block representation of its architecture is presented in figure 1.10.
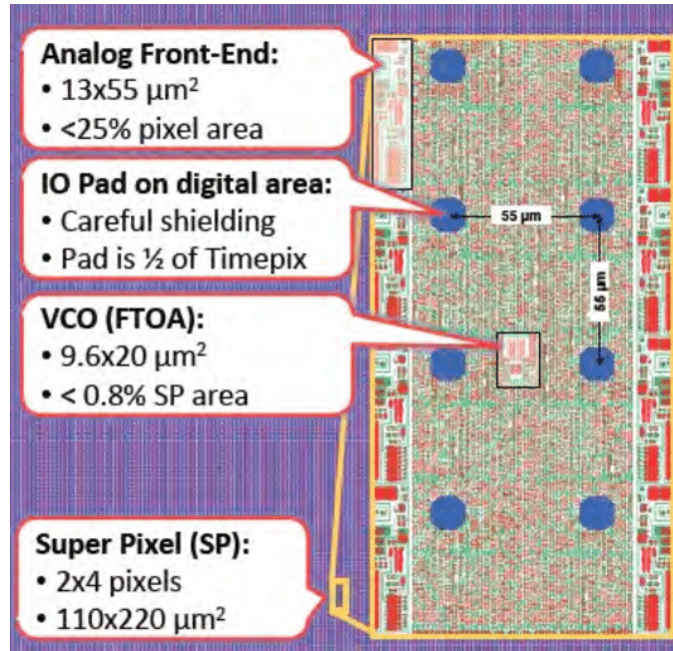
Figure 1.9: Timepix3 super-pixel layout. Taken from [9].
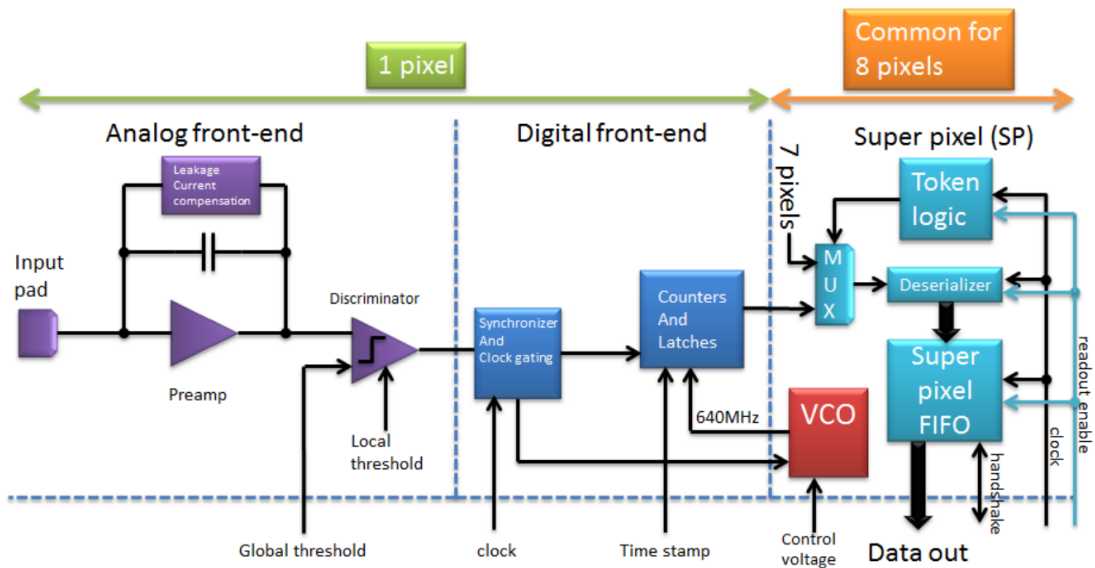
Figure 1.10: Schematic block representation of the Timepix3 super-pixel architecture. Taken from [90].

Each pixel features a CSA with a constant current discharge as its input stage followed by leading-edge discriminator. A global threshold is distributed among the pixels whereas a local one is used to compensate per-channel variations. From

26

a digital point of view, each pixel integrates a dedicated hit-counter. The actual time measurement is performed by VCO oscillating at 640 MHz integrated at the super-pixel level. The measurement consists in singe-edge counting of the VCO oscillations. The discriminator triggers the VCO oscillation and therefore the count starts, while the next reference edge of the 40 MHz clock stops the VCO. In this way it is possible to obtain a TA measurement with a 1.563 ns binning. The ToT measurement is carried out by asynchronously counting the reference clock, resulting in a time binning of 25 ns. The super-pixel configuration creates a bottle-neck for the average per-pixel hit-rate caused by the VCO sharing.

The data-word generated at super-pixel level in charge-time mode is composed of 4 address bits, 14 coarse TA bits, 10 ToT bits and 4 fine-time bits. Both the TA coarse and ToT words are generated by reference clock counts encoded in grey. Each super-pixel buffers the data in a 2 words deep FIFO. This data are transmitted to the EoC with a 2 phase hand-shake. The bus arbitration is carried out using a token-ring arbitration. This arbitration is synchronous and it takes 64 cycles to complete a full token circulation.

The chip features 8 links for the data transmission with a total throughput of 5.12 Gb/s. Each link is a SLVS driver operating at 320 MHz DDR featuring a 8b/10b encoding. The analog side of the front-end can be power pulsed in order to limit its power consumption when inactive. In this way, it can be turned-on in 800 ns.

1.4.2   TDCpix

TDCpix [78] is the pixel front-end ASIC for the hybrid detector of the NA62 GigaTracker [39]. This ASIC has been developed in a 130 nm CMOS technology in 2013. The front-end has been designed to be coupled with a sensor which features 2.4 fC of most probable charge and signal time development shorter than 5 ns. The application requires a time resolution of 200 ps Root Measn Square (RMS) and a 99 % readout efficiency in the frequency range between 800 MHz and 1 GHz. The chip is tileable on three sides, making it able to efficiently cover a sensitive area when organized in a double-row configuration. The ASIC has a data driven interface and can process up to 210 Mhits/s.

The pixel matrix is composed of 40 ×45 pixels with a pitch of 300 µm × 300 µm. Each pixel integrates only the analog part of the front-end. The signal is formed with a peaking time lower than the 5 ns of the input signal development. The signal is then compared with a discriminator in order to produce the timing signal to be digitized. This signal is directly sent to the EoC which integrates 360 dual channel TDCs in total. A schematic of the TDCpix architecture and its floor plan are presented in figure 1.11. Each TDC serves asynchronously five front-end channels and provides both TA and the ToT measurements. The channels priority is regulated with an arbiter. If the TDC is busy when a channel requests the

27
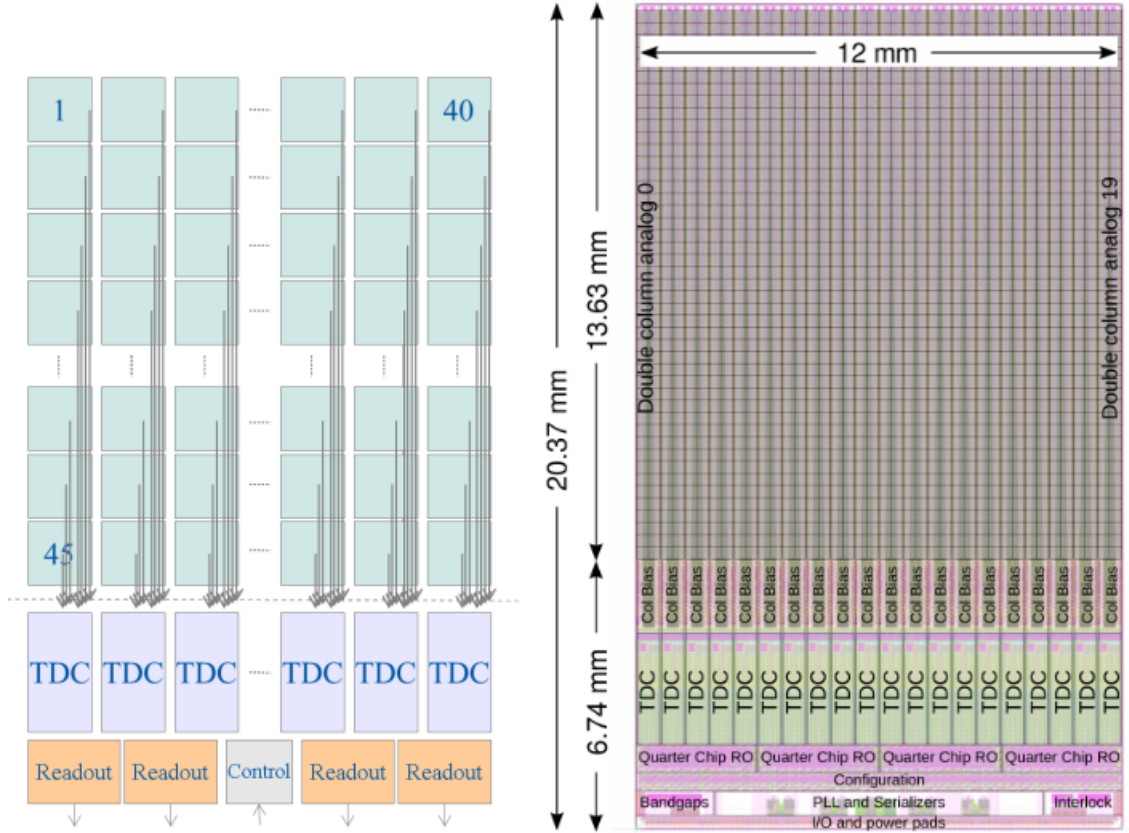
Figure 1.11: TDCpix architecture and floor plan. Taken from [78]

measurement, the event is flagged as a pile-up event. The ASIC integrates 360 dual channel TDCs in total.

The TDC is implemented with a DLL formed by 32 elements. The measurement is carried out by a first coarse stage with a 12 bit grey-counter on the loop completion. This counter length guarantees a 6.4 µs range. For the fine measurement, the state of each element of the DLL is registered in its 32 bit register. The resulting TDC LSB is 97 ps.

After the TDC, a 5bit address is added to the word for the positional information, the data is then inserted in a FIFO for de-randomization. Output data transmission is carried out with four 3.2 Gb/s serializers, each one serving a block of 10 columns. At these blocks level additional 28 bits can be added to extend the time range to 1700 s if needed.

In order to reduce the probability of SEU, a TMR is implemented for every configuration register and state memory. Tests on the analog front-end shows that the RMS jitter is 60 ps. The full chain resolution has been measured to be 70 ps rms.
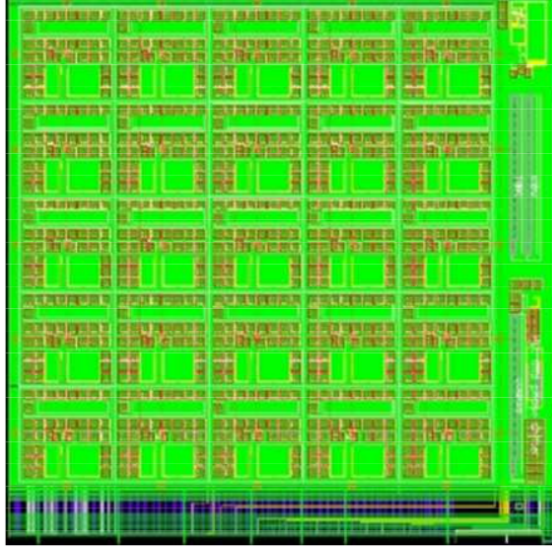
Figure 1.12: ALTIROC1 layout. Taken from [1].

1.4.3   ALTIROC

The read-out mode is trigger-based with an expected trigger rate of 1 MHz.

The pixel architecture, shown in figure 1.13, consists in RF pre-amplifier followed by a discriminator and two TDC. One TDC preforms the TA measurement while the second one performs the ToT one. Since the TA information is more crucial, it features a smaller LSB compared to the ToT one: 20 ps compared to the 40 ps of the ToT. The TA-TDC is implemented with a Vernier-DLL in which the time bin is obtained by the difference between the delay values of the elements of two DLLs. In this case the two unit delays are 120 ps and 140 ps. The number of delay cells determines the range of the TA measurement. Each delay is adjustable via a shunt capacitor voltage-controlled delay cell for calibration purpose. The ALTIROC TDC includes 128 cells obtaining a range of 2.56 ns which is sufficient to cover the measurement window on the ATLAS experiment. The data word generated at the pixel level is composed by 7 bit of TA data, 9 bit of ToT and one data valid bit. A local SRAM is also integrated in the pixel with a memory depth of 10 µs. Data waiting for the trigger are stored locally in these memories.

Electrical tests on the standalone ASIC show that it is possible to achieve a 15 ps time resolution with 10 fC of input charge, and 35 ps with a 4 fC input charge. Tests with the actual sensor have shown a time resolution of 55 ps with a 4 fC input charge.
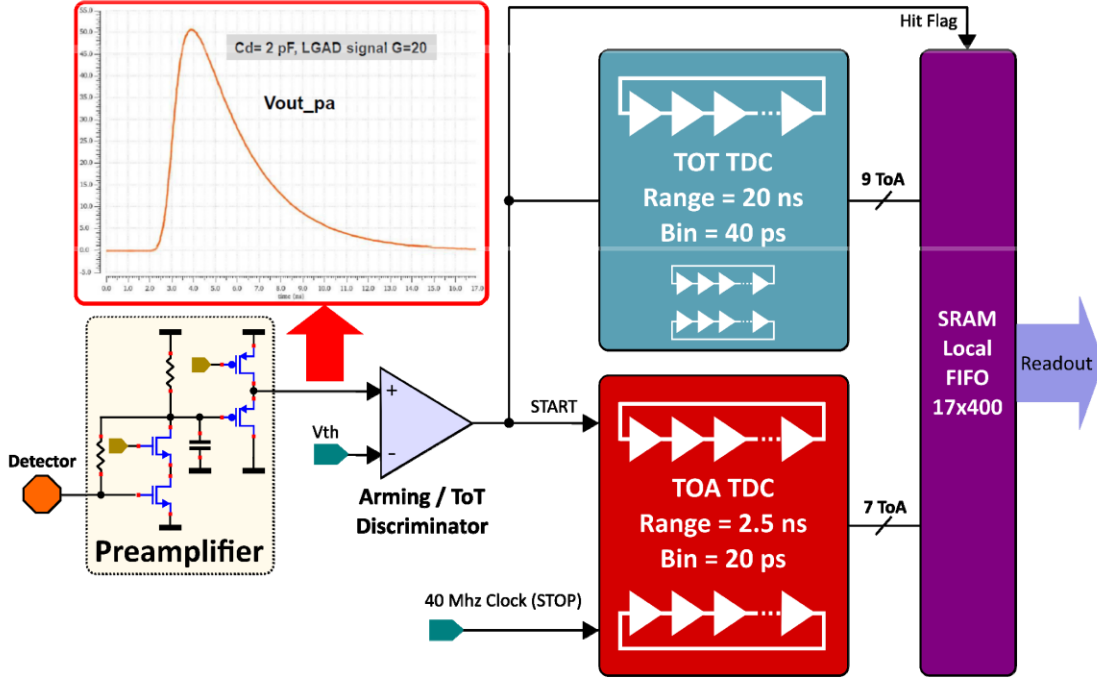
Figure 1.13: Architecture of the ALTIROC1 pixel. Taken from [67].

1.4.4 Timepix 4

Timepix4 [61] is the latest ASIC in the Timepix family. This front-end chip has been developed in 2020 in a 65 nm CMOS technology with the goal of improving in every aspect on its predecessor, Timepix3. First of all, the target specification for the timing bin has been set below the nanosecond mark, namely to 200 ps. The overall sensitive area was also drastically increased to include a 448 ×512 pixel matrix with the Timepix standard pitch of 55 µm × 55 µm. The overall chip size is 24.7 mm × 30 mm, making it extremely efficient in covering a large sensitive area with a large fill-factor. Moreover, the totality of the chip surface can be covered with the sensor, including the chip periphery, leaving no dead area on the chip surface. Additionally, if the TSV are used for IO and power connections, this ASIC can be tiled on 4 sides which will result in the removal of the interconnection area occupied by the wire-bonding. The standard chip includes an interface area on the south side for wire bonding, this section can be diced when the chip is used in TSV configuration.

As mentioned above, the ASIC can be coupled with a sensor of the same dimension. In order to do so, the sensor bump-bond pads are realized on the totality of the chip bottom face with pitch of 55 µm × 55 µm. However, the actual pixel electronics matrix covers a smaller fraction of the chip: the electronic pixel pitch is reduced in the y-axis making its aspect-ratio 55 µm × 51.4 µm. In this way two
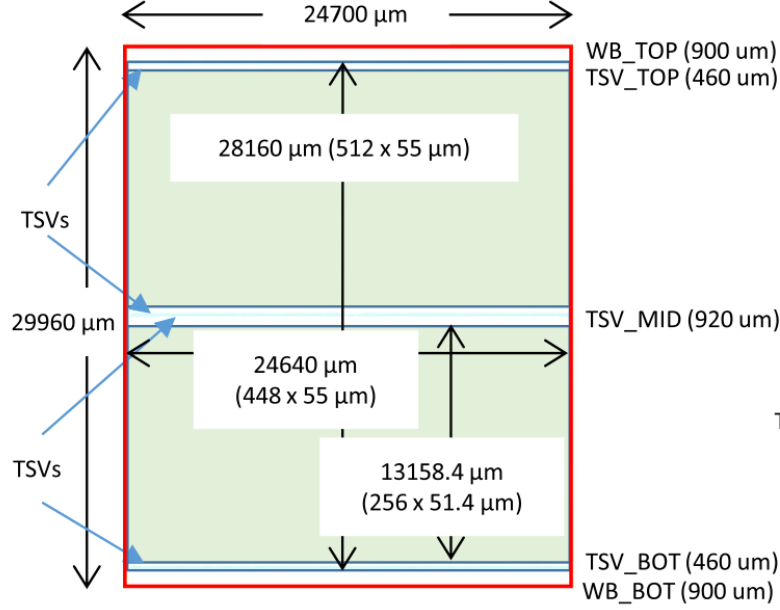
Figure 1.14: Timepix4 floor plan. Taken from [61]

corridors can be derived above and below the matrix, leaving the space to integrate the service electronics and realize the TSV. The ASIC is organized in two sub-matrices of 256×448 pixels, with three service corridors: a north and a south one $460 \, \mu m$ wide, and a central one $920 \, \mu m$ wide. The floor plan of the ASIC detailing this organization is presented in figure 1.14.

From the point of view of the analog front-end electronics the scheme consists in a CSA connected to a discriminator. The CSA is realized with a Krummenacher filter [53] and features a programmable charge-to-voltage gain. Similarly to its predecessor, the discriminator threshold is adjusted locally using a 5 bit DAC. The total jitter introduced by the analog-front end has been quantified to be $75 \, ps$ RMS [42].

As in Timepix3, the matrix is organized in super-pixel blocks of 8 pixels with shared electronics. A schematic of the pixel, super-pixel and column architecture is presented in figure 1.15. The TDC is part of these shared blocks and is realized with $640 \, MHz$ VCO formed by 4 inverters. The TA measurement is performed with a coarse and fine steps. The coarse step is implemented similarly to what was done in Timepix3 by counting the VCO oscillations, it accounts for 16 bits. The fine measurement is provided by 4 bits of the inverters state which are latched when the stop signal arrives. The ToT measurement is preformed by a combination of counts on the $40 \, MHz$ reference clock and VCO one. The slow counter is used to cover a large range and the fine one for the residuals. In this way the high power consumption time of the running VCO is limited to the residuals measurements.
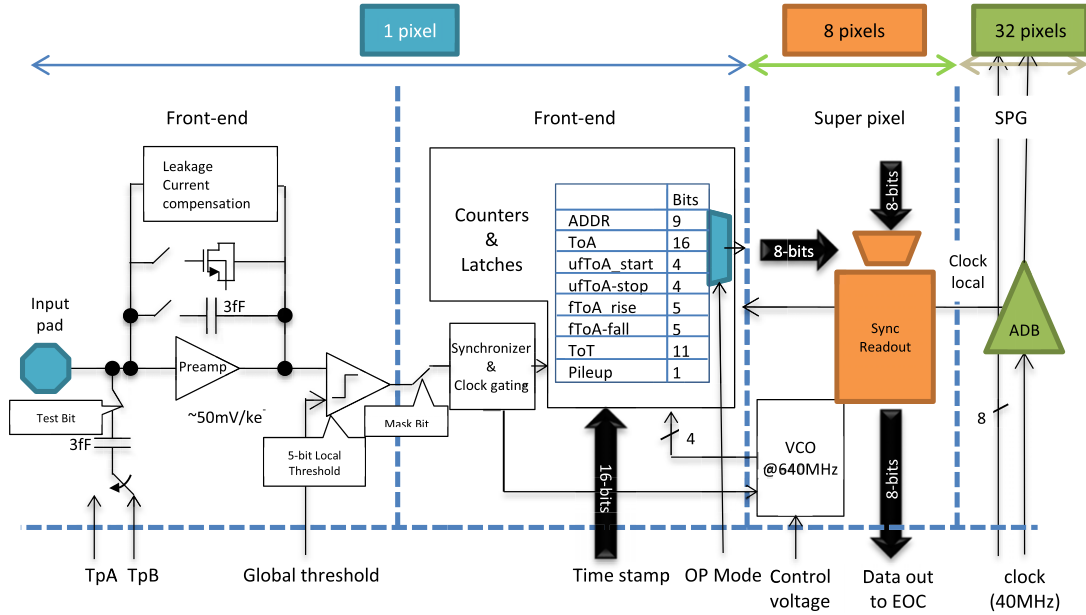
Figure 1.15: Timepix4 pixel, super-pixel and column architecture. Taken from [61]

The total power consumption per unit area is inferior to $1\,\mathrm{W/cm^2}$.

The data transmission to the EoC is also implemented at the super-pixel level. Each super-pixel transmits via a priority encoder its data to the next until the EoC is reached. In this phase a 9 bits vertical address is added whereas the 8 bits column address is added at the EoC resulting in event word 65 bits long. Finally a multiplexer distributes the data among the output links. Timepix4 integrates 16 output serializes for a total bandwidth of $10\,\mathrm{Gb/s}$.

The reference clock distribution along the columns is carried out by using a chain of adjustable delays. Since it is unfeasible to reliably distribute the clock among the pixel with a small skew, the clock is distributed with a controlled phase. Across the column, a delay loop is created. 2 adjustable delay units are implemented every 4 super-pixels, one transmits the clock to the next pixel whereas the other is used in the loop-back. In the EoC a controller sets the total loop-phase to $25\,\mathrm{ns}$ by adjusting the delays with a global code. In this way the total clock skew between the first and last pixel of the column is limited to half of the clock period. The actual phase from the reference must then be removed in a post-processing phase.

Timepix4 can also be operated in a frame-based counting mode. In this mode the VCO is disabled. Two ranges can be selected: an 8 bit one and a 16 bit one. Each pixel integrates a double buffer in order to prevent data loss. By selecting the lower count mode, it is possible to obtain a higher refresh rate of $90\,\mathrm{kframes/s}$, whereas in 16 bit mode the refresh rate is $45\,\mathrm{kframes/s}$. The highest obtainable hit-rate is $5\,\mathrm{Ghits/mm^2/s}$ in 8bit counting mode.
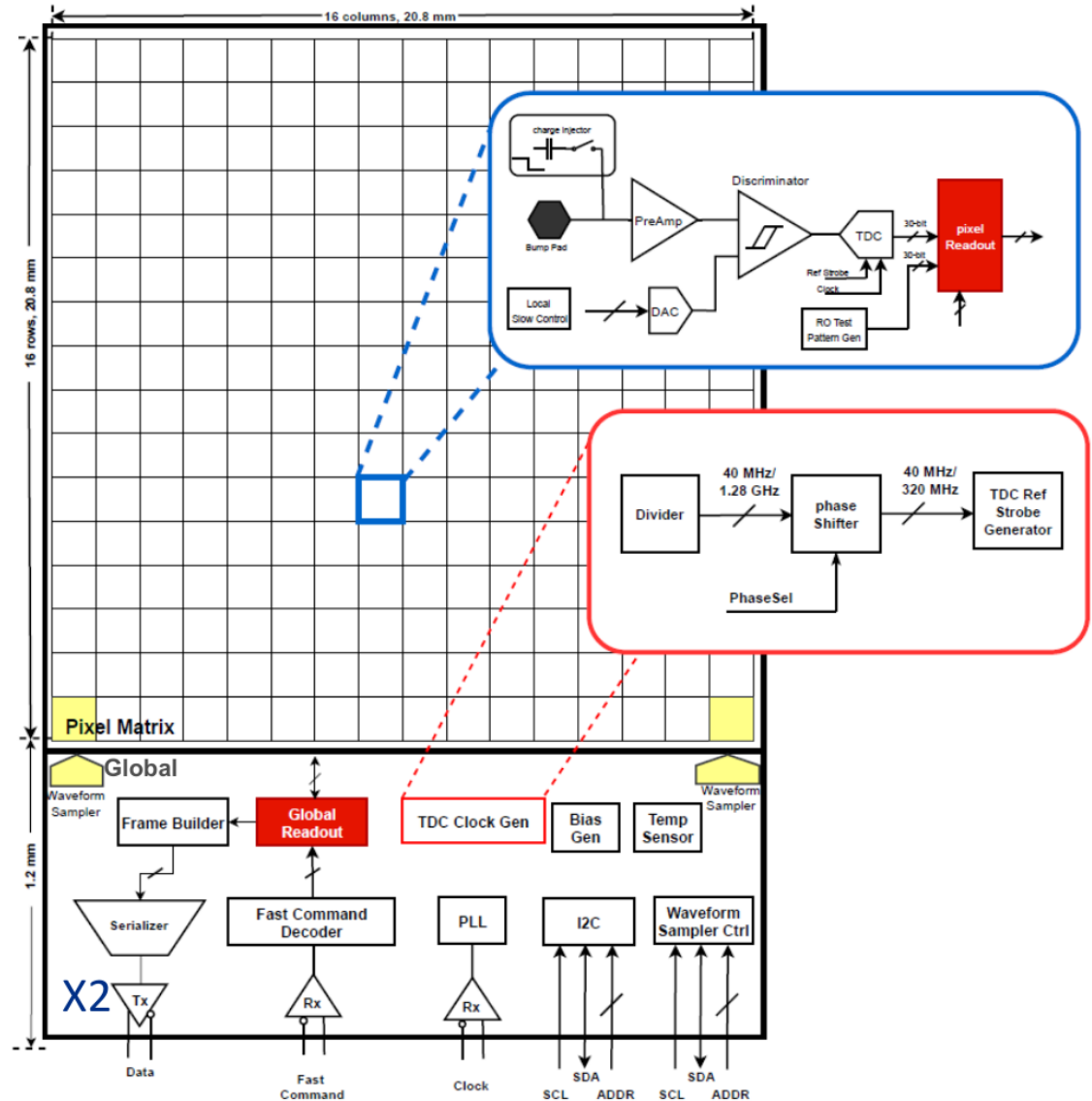
Figure 1.16: ETROC1 architecture. Taken from [59].

1.4.5 ETROC

ETROC is the pixel front-end ASIC developed for the future upgrade on the End-cap Timing Layer (ETL) [104] of the CMS experiment. This detector is designed to operate in the Hi-luminosity Phase of LHC in which the application requires 50 ps of time resolution per hit. Two prototype chips in 65 nm CMOS technology have been developed for this project: a single channel one named ETROC0[105], and a 16 ×16 pixels one named ETROC1[59]. The front-end is designed to read a 1.3 mm × 1.3 mm LGAD sensor with a capacitance of 1.5 pF. The sensor matrix is
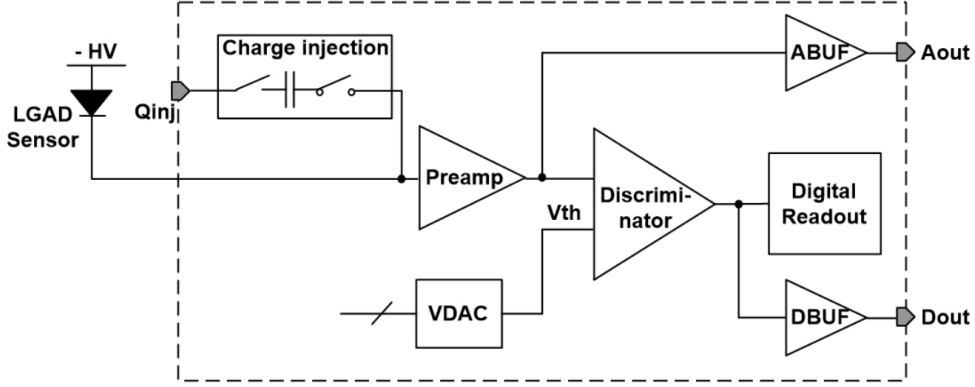
Figure 1.17: ETROC analog pixel architecture. Taken from [105].

coupled with the front-end via bump-bonding. The input charge ranges from 10 fC to 25 fC with a 15 fC MPV. The overall ASIC architecture is presented in figure 1.16.

The pixel architecture is comprised of a pre-amplifier, a discriminator and a TDC. In the pixel area, other blocks have been implemented: a DAC for threshold fine-tuning, a charge injection circuit for channel testing and a local RAM for data storage. A schematic representation of the analog front-end is presented in figure 1.17. The pre-amplifier is a TIA with a programmable passive resistive path for gain setting. The feedback resistor can be selected in a range of resistance ranging from 4.4 kΩ to 20 kΩ. The TIA peaking time is on the order of 2 ns whereas the average ToT is 6 ns. The discriminator is formed by 3 differential gain stages followed by a comparator with hysteresis. The chain is terminated with a digital buffer. The gain stages are able to obtain a 35 dB gain factor on a 3 mV input signal. The comparator is composed by a first differential cell with hysteresis followed by single ended gain stage. The threshold is set-per channel with the 10 bit DAC. This DAC can provide a LSB of 0.4 mV, the noise is limited using a first order RC-filter.

The TDC is implemented with different sets of DLLs [118]. One is used for ToT measurement whereas two are used for the TA one. The TA DLLs differ from one to the other in terms of granularity of the delay of each element, in this way the measurement is dived in a coarse and a fine step.

The power consumption is under 1 W of power per chip. The analog part of each front-end contributes with 1.53 mW from the TIA and 0.84 mW from the discriminator. Additionally, the TIA power can be set in a low-power mode reducing its consumption to 0.74 mW. The TDC consumes 0.94 μW of power.

ETROC0 allows direct testing of the analog front-end. Tests of this prototype show that the RMS jitter of a MPV signal is 16 ps for the TA and 4.5 ns for the ToT. When the whole chain is tasted in ETROC1 the TA RMS jitter becomes 43 ns

for the MPV signal.

Table 1.2: State of art of Timing Front-End ASICs

ASIC	year	node [nm]	Time Res [ps]	pixel size [µm]	# of pixels	power per channel [µW]	power per unit area [W/cm²]	sensor type	input capacitance [fF]	MIP [fC]	Hit rate [GHits/cm²]	TW correction
Timespot1[a][b]	2021	28	50	55×55	1024	40	1.32	3D-Si 3D-Diam.	45	1.8	99	ToT
Timepix4[b] [61]	2020	65	135	55×55	229×10^3	18	0.60		45	1.6	0.36	ToT
FASTPIX[c][16]	2021	180	142	20 hex	68	n.a.	n.a.	monolithic	n.a.	n.a.	n.a.	ToT
FAST2 [29]	2022	110	36	500×500	32	3×10^3	1.2	UFSD	3.4×10^3	16	120	soft CFD
FastIC[c][30]	2022	65	107	1500×1500	8	12×10^3	0.53	MCT SiPM PMT	n.a.	n.a.	n.a.	ToT
DIAMASIC[37]	2022	130	80	n.a.	n.a.	1.5×10^3	n.a.	Diamond Strip	1×10^3	10	n.a.	no
TOFHIR2X [77]	2021	130	24 BoL 55 EoL	3000×3000	32	12.4×10^3	0.11	SiPM	n.a.	n.a.	2.8×10^{-2}	amplitude
ETROC1[c][59]	2022	65	29	1300×1300	16	2.37×10^3	0.14	LGAD	3.5×10^3	6	2.3	ToT
ALTIROC1[1]	2020	130	35 BoL 70 EoL	1300×1300	25	4.4×10^3	0.26	LGAD	6×10^3	4	n.a.	ToT
FCFD0[b][c][91]	2021	65	30	n.a.	3	n.a.	n.a.	LGAD	3.4	5	n.a.	CFD

The ASICs presented in this table features a time resolution better than 200 ps experimentally measured. The tests have been performed with an actual sensor and radiation source or with an electrical auto-test. In case results from auto-test the circuit input capacitance is presented; in case of test with sensor, the total capacitance at the input node is presented. Data are flagged as n.a. when either not available or not applicable.

[a]This work, [b] Electrical auto-test only.,[c]no TDC

Chapter 2

The TimeSPOT ASICs

This chapter describes the ASIC developed as part of the TimeSPOT project. The work of this thesis was to develop and test the analog front-ends integrated into these two prototypes: Timespot0 and Timespot1. This chapter, therefore, illustrates the general content and structures of the two ASICs, whereas a detailed description of the design and test of circuits developed as part of this work will be discussed in the next chapters. Nevertheless, the author had a crucial contribution in defining the final scheme of the pixel front-and ASIC architecture.

The first section of the chapter presents the TimeSPOT project in terms of its goals and method, with particular attention to the structure of the final demonstrator and the characteristics of its sensors. This section also presents the 28 nm CMOS technology used to develop the two ASICs.

Section 2.2 describes in detail the contents and architectures of the two prototype ASICs.

2.1 The TimeSPOT Project

The Time and SPace real-time Operating Tracker (TimeSPOT) [55] is a project of the Italian institute for nuclear physics (Istituto Nazionale di Fisica Nucleare INFN[48]). It aims to develop a small scale 4D tracking detector suitable for high
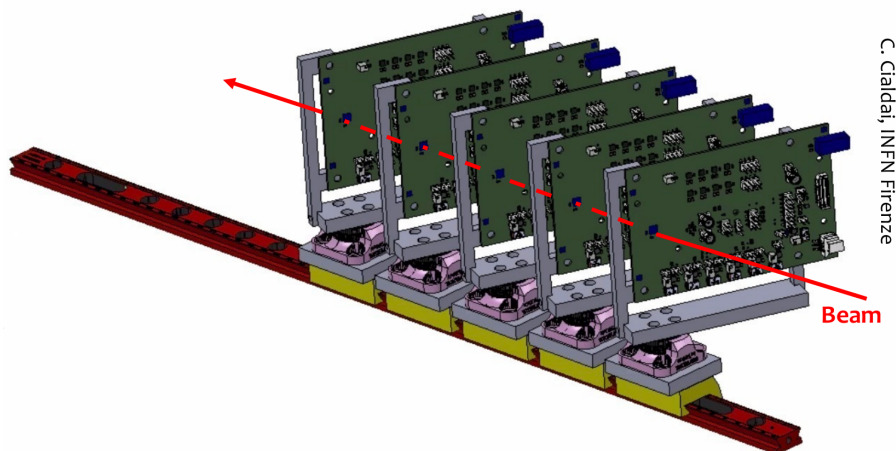
Figure 2.2: Rendering of the TimeSPOT demonstrator telescope. The long side of each PCB is 12 cm.

luminosity HEP experiments. The final demonstrator building blocks will include a pixel sensor, a front-end ASIC and a full readout chain.

The chosen detector configuration is a hybrid one: the sensor and its front-end electronics are thus realized in two separated dies and then connected via bump bonding. This topology was adopted because it offers higher radiation tolerance and better timing performance than the monolithic alternative. The sensor matrix is formed by $55\,\mu m \times 55\,\mu m$ featuring 3D pixels which are built in two different substrate materials: silicon and diamond.

The front-end ASIC is developed in a commercial 28 nm CMOS technology. The very font-end must be designed to be compatible with both the two sensor variants. The objective in terms of time resolution is to obtain a per-pixel RMS resolution of 100 ps or better, with the optimal target defined by the sensor intrinsic resolution: around 25 ps. The ASIC must sustain an event rate per unit area of $3\,GHz \cdot cm^{-2}$, for this reason the time measurement must be digitized locally via a per-pixel TDC. The target power consumption per unit area will be around $1.3\,W/cm^2$: compatible with the current state of the art in cooling systems [23].

The read-out electronics will be realized using Field Programmable Gate Array (FPGA) boards. The read-out firmware must sustain the data throughput of multiple front-end ASICs while also processing the data in order to facilitate the subsequent track reconstruction.

The TimeSPOT demonstrator will be composed of four to eight layers of hybrid detectors. Each unit will be composed of a small-scale matrix of around a thousand pixels. The hybrid will be mounted on a dedicated compact board attached to a rail bar in order to precisely set the inter-plane distance. The configuration of the TimeSPOT demonstrator telescope is shown in figure 2.2.

The project started in 2018 with the objective to test the demonstrator in 2022. The work was divided in six Work Packages (WP) among 10 INFN headquarters. Each WP is dedicated to:

(WP1) : 3D silicon sensor design and characterization.

(WP2) : 3D diamond sensor design and characterization.

(WP3) : Front-end ASIC design and characterization.

(WP4) : Read-out system design.

(WP5) : Reconstruction algorithm design.

(WP6) : Assembly and characterization of the demonstrator.

This thesis work is part of the WP3.

2.1.1   The TimeSPOT Sensors

(a) 3D-Silicon sensor

(b) 3D-Diamond sensor
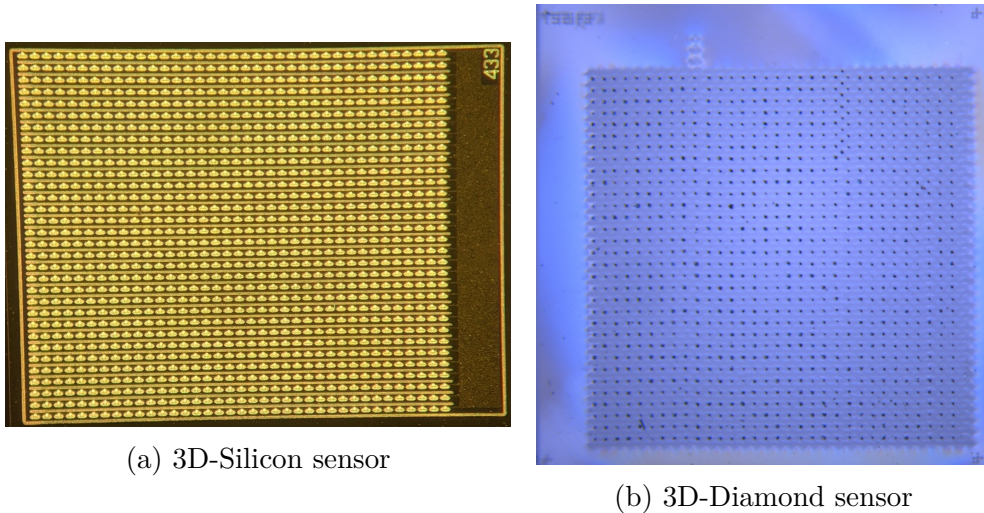
Figure 2.3: Photograph of the TimeSPOT Sensors.

Both the two 3D TimeSPOT sensors share the same pitch of $55\,\mu m \times 55\,\mu m$. A photograph of the two sensors is shown in figure 2.3. These sensors feature totally new topologies that demanded specific advancements in their manufacturing processes. The work on their innovative design was carried out with extensive simulations and characterization through many batches of prototypes. This work was not only necessary for advancing their development, but was also crucial for the front-end design process. This section will not present the history of the development,but rather the characteristics of the latest prototypes, which will be soon coupled with the front-end ASIC.

The 3D Silicon Sensor
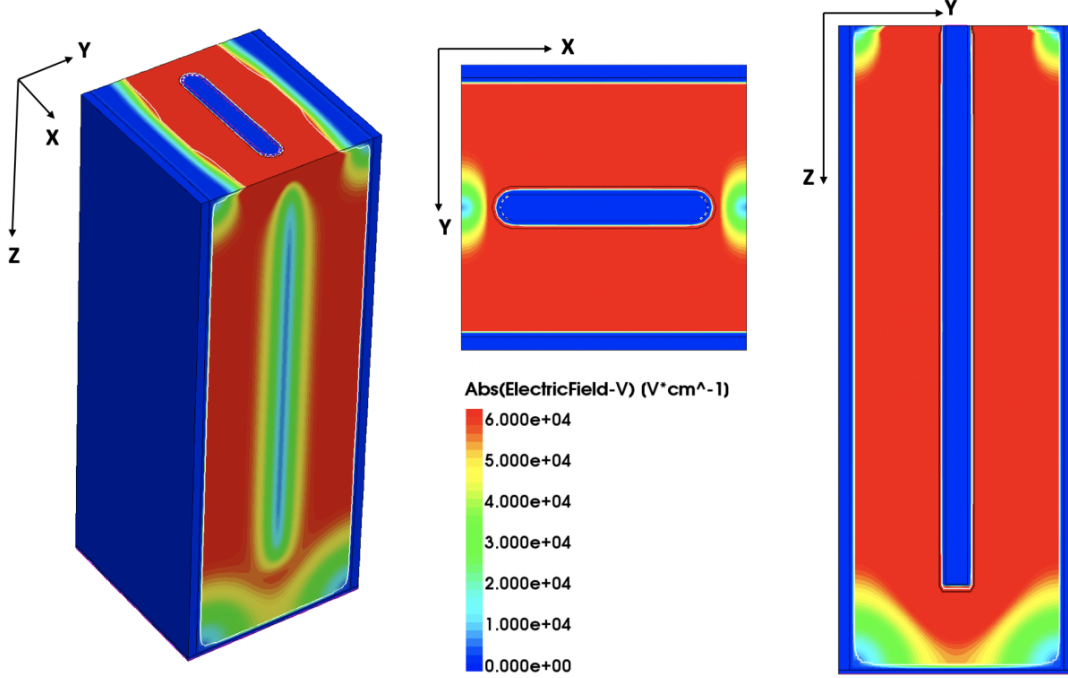
Figure 2.4: Elementary cell of the TimeSPOT 3D-Silicon sensor $(55\,\mu m \times 55\,\mu m \times 150\,\mu m)$. The color code shows the simulated electric field strength and uniformity. Taken from [62].

The TimeSPOT 3D-Silicon sensor is developed by TIFPA[109], manufactured by FBK[31] and simulated by INFN Cagliari.

The pixel features a novel trench-shaped electrode [70]. Each pixel is a parallelepipedal cell measuring $(55 \times 55 \times 150)\mu m^3$. In the center is positioned the collecting trench electrode measuring $(40 \times 5 \times 130)\mu m^3$. Therefore, each pixel row is segmented in many collecting electrodes, which on their top present the ohmic contact for the bonding with the front-end. The bias-electrode is instead continuous in the row dimension. It extends down to the bottom face of the substrate, where the bias voltage is provided. The shape of this cell is presented in figure 2.4.

Simulations show that this electrode shape produces a more uniform field compared to a traditional column electrode, reducing the signal variation based on the interaction position [63]. Figure 2.5 shows the distributions of the TA of the simulated signals. The sensor Minimum Ionizing Particle (MIP) is 2 fC and it develops in under 300 ps. The main drawback of the trench configuration is the relative high capacitance of sensors (120 fC excluding the ohmic contacts) due to its larger parallel capacitive coupling with the bias electrodes.
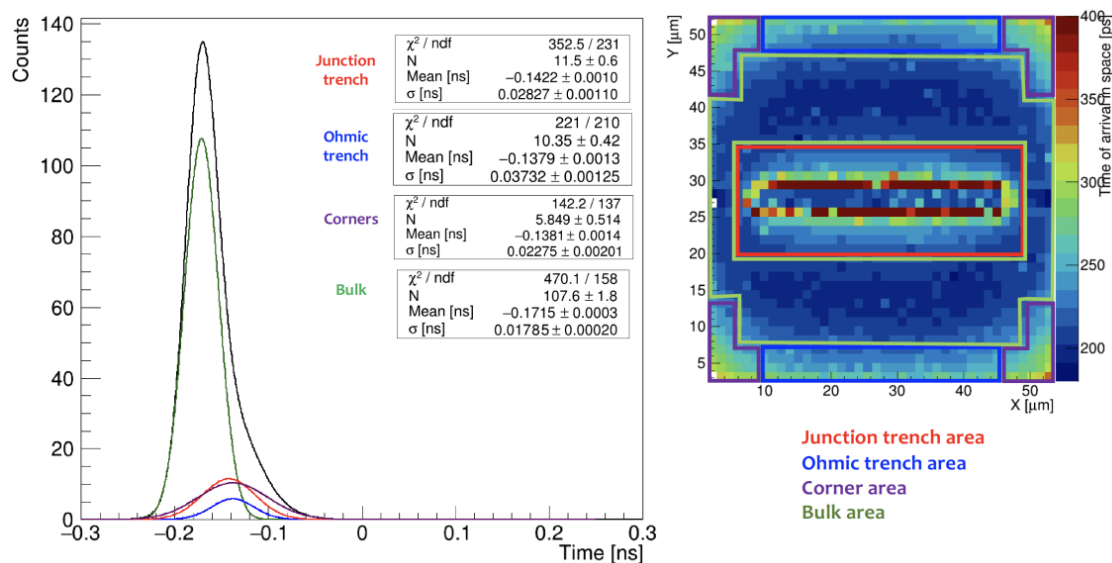
Figure 2.5: Distributions of simulated signals TA for a deposited charge of 2 fC (MIP). The charge is injected in different positions with different field strengths (shown on the right). The σ values of these distribution accounts for the time fluctuation connected to charge generation described in 1.3.1. The total time fractionation due to the free charge generation is the convolution between these distributions and the space distribution on the hit position probability. Therefore, the sensor time resolution is deeply connected to the experiment setup. Taken from [62].

The substrate is p^{--} doped [1], whereas the signal electrodes are n^{++}[2], and the bias ones are p^{++} (highly doped with acceptors) [33]. At the bottom of the cells there is a 3 µm thick p^{++} layer. The bulk silicon is highly doped in order to increase its radiation hardness.

The manufacturing process used consists in etching the silicon bulk with Deep Reactive Ion Etching (DRIE) and then doping the holes with desired atoms[14]. These holes are then filled with polysilicon in order to create the ohmic contacts, this process limits the smallest electrode that can be produced: filling reliably these holes becomes difficult the smaller their section. At the same time, making electrodes excessively large will reduce the active area, reducing the efficiency.

The sensor has been tested both in the laboratory with a laser and at a test-beam facility with a radiation source[57][56]. The sensor shows an intrinsic time resolution of at least 24 ps. Figure 2.6a shows the experimental TA distributions,

---

[1]highly doped with acceptors

[2]highly doped with donors

(a) Time distribution.

(b) Amplitude distribution (related to the charge distribution).
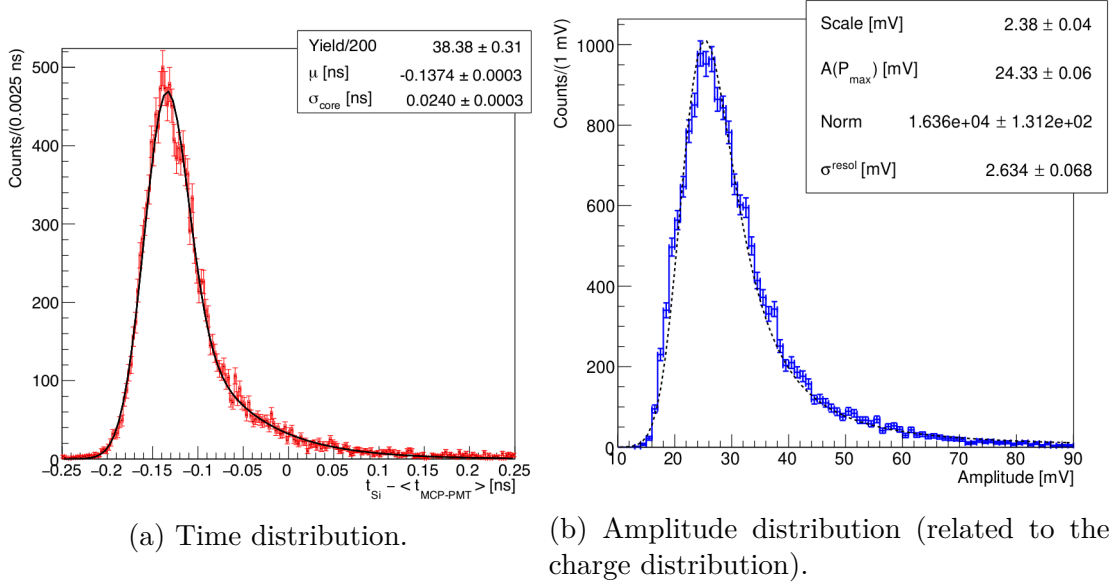
Figure 2.6: Measured signal characteristics of the 3D-Silicon sensor. The amplitude value is shown in millivolt and it refers to the output voltage of the discrete-components integrator used in the experimental setup. Taken from [5].

whereas figure 2.6b shows the shape of its charge distribution.

The 3D Diamond Sensor

The 3D-Diamond sensor is developed by INFN Firenze and simulated by the Perugia University. It is formed by graphite electrodes inside a polycrystalline diamond bulk made by using Chemical Vapor Deposition (CVD) [82]. A sensor made of diamond is naturally more radiation hard compared to silicon ones [10]. Therefore this sensor variant is most suited for zones of the experiment with high radiation fluence.

The elementary cell measures $(55 \times 55 \times 500)\mu m^3$. The utilization of the laser graphitization technique on a transparent mean allows fabricating the electrodes that extend on the whole bulk of the sensor. Each cell includes five electrodes: four bias electrodes on the edge of the cell (which are shared among the neighboring cells) and one collecting electrode. The electrodes shape is a column with a $5\,\mu m$ diameter and a length of $475\,\mu m$. Figure 2.7 shows this elementary cell. Simulations show that with an expected MIP of $2.8\,fC$, the signal develops in under $1.3\,ns$ [74]. Figure 2.8 shows simulated signals for different interaction points.

The processes to manufacture the bulk consist is a heteroepitaxial growth on top of iridium. It represents a low-cost solution enabling the formation of samples with a surface area up to $2\,cm \times 2\,cm$ and $500\,\mu m$ thickness. The column shaped electrodes are made by direct graphitization of the diamond using pulsed laser
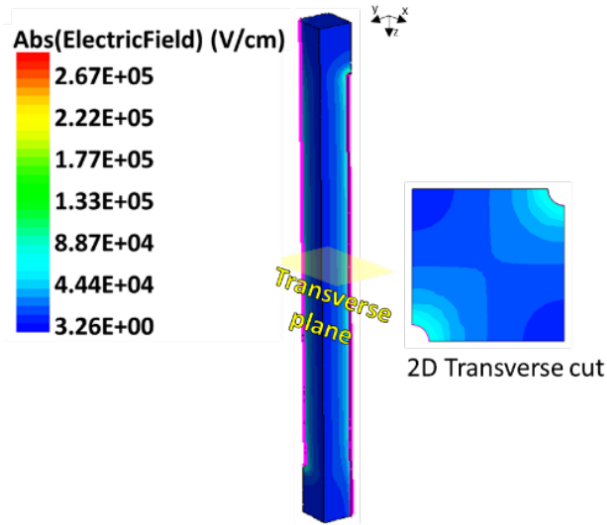
Figure 2.7: Elementary cell of the TimeSPOT 3D-Diamond sensor $(55\,\mu m \times 55\,\mu m \times 500\,\mu m)$. The color code shows the simulated electric field strength and uniformity. Taken from [74].
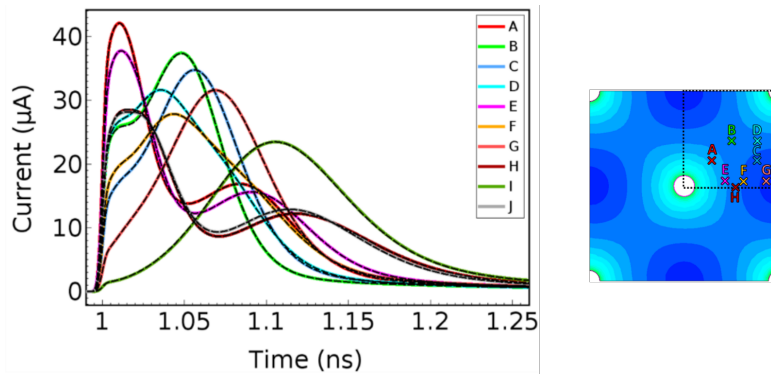
Figure 2.8: Simulated signals of the 3D-Diamond sensor for different injection points. The related injection position is shown on the right. Taken from [74].

along the beam focus. In order to trigger the graphitization process, the laser energy density per pulse must exceed $5\,J/cm^2$. This fabrication technique allows to freely trace the electrodes in three dimensions.

Results from experimental tests of the sensors demonstrate an intrinsic time resolution of 200 ps with a laser induced signal and 250 ps from a test beam. The MIP has a value of 2.8 fC. The plots from the data analysis of these tests are shown in figure 2.9. Due to the column shape, the sensor capacitance has value of 30 fF, very low compared to the silicon sensor.

The major disadvantage of this sensor is the resistivity of the electrodes. In fact,

43
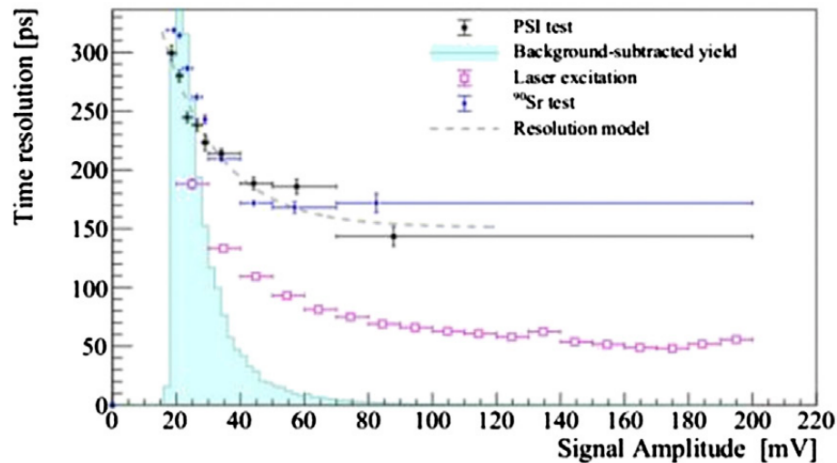
Figure 2.9: Timing characteristics of the 3D-Diamond sensor. The resolution has been tested with different sources. Taken from [6].

only the surface of the column is made of graphite and thus it is the one conducting. The resulting resistance will be in series to the front-end and will determine a loss in jitter performance.

Multiple graphitization steps can be employed to enhance the electrodes resistivity. With multiple steps, it is also theoretically possible to produce trench shaped electrodes. The main drawback of this is the increase in the time required to produce the sensor.

2.1.2 The $28\,\mathrm{nm}$ CMOS Process

This chapter describes the major characteristics of the selected CMOS technology. The technology features standard planar MOS transistors with a minimum feature (the channel length) of $28\,\mathrm{nm}$. Over the course of this thesis, the $28\,\mathrm{nm}$ value is converted to $30\,\mathrm{nm}$ for convenience. Therefore all the sizes of the transistors presented hereafter will be scaled by a factor 1.07 in relation to their actual physical size. The process features eight metal layers with an additional thick redistribution layer. This last layer is used as the base for depositing the bump-bond balls. The technology allows integrating a deep-nwell in order to realize transistors inside triple-N-wells. It is also possible to integrate various flavors of transistors with different threshold voltages.

Although the technology is a planar one, it presents some peculiar characteristics compared to a standard CMOS process. These change will affect the typical design flow of the integrated circuit.

First, some Resolution Enhancement Techniques (RET) have been adopted in order to make possible to realize geometries at this scale. This fact is caused by
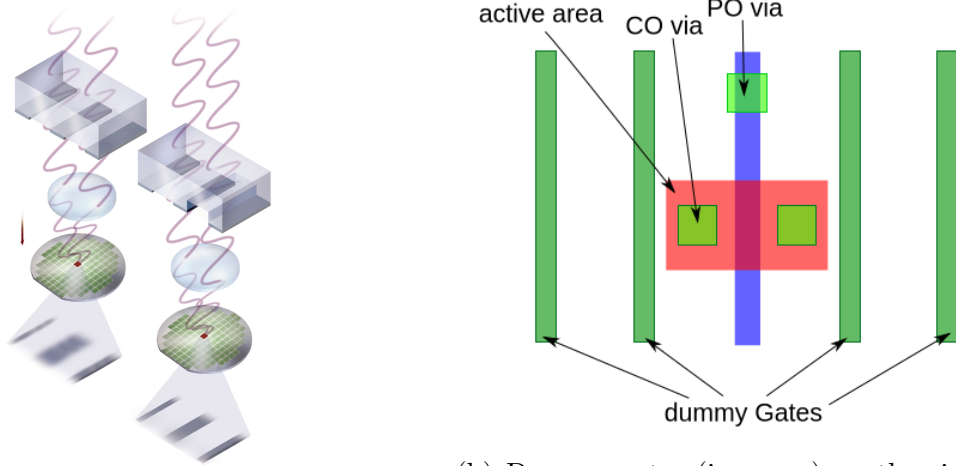
(a) Principle of the PSM technique. Taken from[98].

(b) Dummy gates (in green) on the sided of a true gate (blue). The red area is the transistor diffusion.

Figure 2.10

interference and diffraction effects that occur when the wavelength of the light used for the lithography is approximately of the same order of magnitude as the size of the mask apertures. Usually these effects were avoided by having larger apertures and by increasing the numerical aperture of the shrinking lens positioned before them. This approach, however, has come to its limit, which is dictated by the biggest reliable lens that can be manufactured. RET are used to overcome this limit.

In particular this technology takes advantage of Phase Shifting Mask (PSM) [58]. In PSM the lithography mask couples every aperture with another one which shifts the phase of the light in order to use interference effect to increase or decrease the exposition of the selected wafer areas. This method was initially used to prevent diffractive issues, and it grants a theoretical resolution increment of $\times 2$. However, if the design is regular, adjacent apertures of design can be used to correct their respective interference. In the modern processes masks with etched quartz windows of different thickness are used to make the interference slits. A representation of the PSM principle is illustrated in 2.10a.

PSM requires repetitive patterns in the design's layout in order to operate properly. Transistor gates must therefore be organized in parallel array of identical size and separation. In this technology is no longer possible to produce gates with complex geometries. Moreover all gates inside the chip must be oriented in the same direction. Gate size can be changed from two region of the design, even if the transistors on the boundaries of the two regions will not be manufactured properly. Additionally in order to manufacture a gate two of its replica must be positioned on
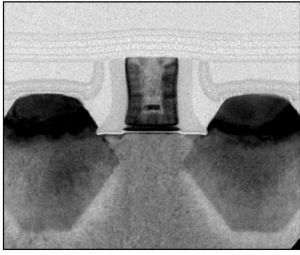
Figure 2.11: Photograph of channel-strained PMOS transistors. The *SiGe* implants can be seen on the sides. Taken from [2].
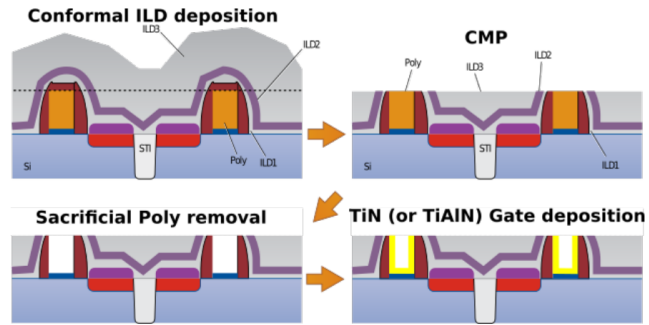
Figure 2.12: Scheme of the gate-last technique. First the material is diced using a Chemical Mechanical Polishing (CMP). Then the sacrificial poly is removed and replaced with the desired metallic material.

each of his sides. At the boundary of a region these replicas can be implemented as dummy transistors, as shown in figure 2.10b. These dummy gates can also be used for transitioning from a size area to another, improving the yield of the transistors on the boundary. In any case, a regular design will improve the overall yield of the circuit and reliability of the simulations.

The usage of the PSM also affects the possibility to realize gate-all-around device. This transistor gate shape was engineered in order to reduce the effect of radiation damage on the transistor threshold voltage. Due to ionization, charges can be trapped below the gate insulator, changing the threshold value required to invert the channel. In any case radiation-damage test on this technology[116] shows that its radiation hardness using linear gates is on par with less scaled technologies with gate-all-around transistors (65 nm and 130 nm)[49][49].

Another characteristic of the process is the presence of a high-k dielectric material used in the gate insulator coupled with a metallic material for the gate. The term high-k means that the dielectric material has a high dielectric constant. Changing the dielectric material is mandatory for technologies scaled below 45 nm in order to limit the tunnel effect through this layer. By using HfO_2 as insulator the gate leakage current can be reduced by a factor 25 for a NMOS transistor and a factor 10^3 for a PMOS one.

The adoption of this material crates two problems: first the HfO_2 can't be deposited by simple oxidation as in the case of SiO_2; second the junction of this material with a standard polysilicon gate will originate an unwanted depletion region. The presence of this depletion region will make it impossible to control the channel inversion by acting on the gate potential. Therefore the high-k dielectric is deposited by first depositing the SiO_2 by oxidation, and then replacing it with HfO_2 via atomic layer deposition.

Similarly, the polysilicon can be replaced with another material using a gate-last technique. A scheme of this technique is shown in figure 2.12. The materials chosen as gate metals in this technology are different for the NMOS and the PMOS. This was done for work function engineering [96] purposes. The chosen metals for the gates are TiN for the NMOS and $TiAlN$ for the PMOS. The main drawbacks of these materials is their higher resistivity.

A consequence of this property of the technology, the gates of NMOS and PMOS transistors can no longer be connected together by joining their gate materials. Moreover all the routing realized with the gate top material is highly discouraged because it will generate a line with high resistance value.

Lastly, mechanical strain on silicon is adopted to enhance the mobility of carriers inside the material. This can be done for two reasons: to counteract the mobility reduction of the scaling (due to the increased doping) and to equalize the difference in mobility of holes and electrons. There are two different kinds of strain: biaxial, which increases the mobility on the entire plane; and uniaxial for an increment in a specified direction (in case of MOSFET the direction of their lenght is the crucial one). A technique for uniaxial strain is the so called Embedded SiGe (e-SiGe)[3]: the strain is caused by SiGe made drain and source electrodes: their material is etched and replaced with SiGe growth epitaxially. These implantations strain uniaxially the material of the channel in between them. With this technique an increase of 45 % is expected for holes mobility. In order to strain the NMOS channel, a silicon nitride cap can be added on top of the gate. In this way, the two types of MOSFETs can be engineered separately. The e-SiGe must be added after the sacrificial poly-gate deposition [2]. The main consequence of these properties is related to an overall equalized response of the NMOS and PMOS, due to their similar mobility terms.

2.2   The ASICs

The plan for the development of the TimeSPOT front-end ASIC was to divide it in two stages with two prototypes. The two ASICs ere named Timespot0 and Timespot1, a photograph of them is shown in figure 2.13. The first prototype was plan to be composed of individual blocks, independent of each other, with each block being a part of the future front-end chain. This decision was taken in order to focus the design efforts on achieving the desired performance from the critical blocks: the analog front-end and the TDC. Two other advantages of this approach are the fact that having separated domains inside the chip will simplify the tests; and that the auxiliary blocks will be ready and functioning for the next prototype.

Consequently, the first ASIC for the TimeSPOT project, named Timespot0, was developed in 2018 and tested in 2019. Both the front-end and the three implemented TDCs obtained satisfying results within the limit of the target 100 ps time resolution
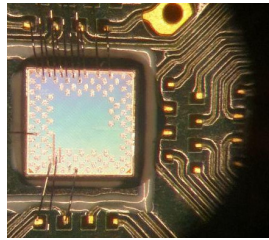
(a) Timespot0 ASIC
(1.4 mm × 1.4 mm)
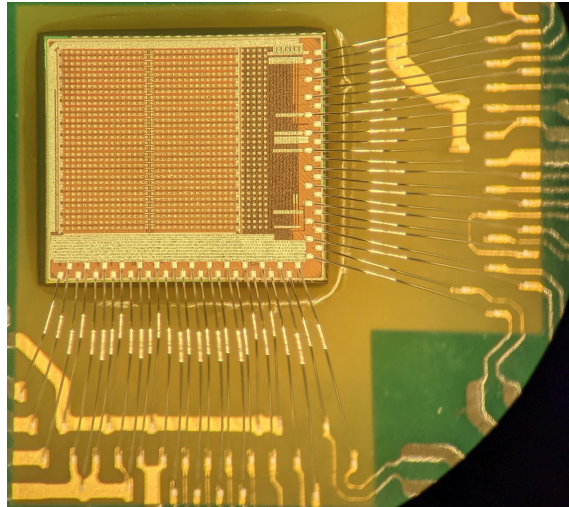
(b) Timespot1 ASIC
(2.6 mm × 2.3 mm)

Figure 2.13: Photograph of the TimeSPOT ASICs in scale.

of the project. The chip was not connectable to the actual sensor since the progress on the two parts was not at a mature stage. The ASIC contained instead a charge injection circuit to emulate the sensor. Meanwhile the research on the sensors progressed with the completion of the first prototypes. The test on the 3D-Silicon sensor showed an intrinsic time resolution of at least 24 ps, well beyond the target initially pursued. This result changed the initial plan for the second prototype ASIC, Timespot1.

The new goal for Timespot1 was to develop the first ASIC organized in a pixel-matrix connectable to a sensor while having a time resolution sufficient to valorize the quality of its sensor. This task was made more easy by the presence of new cooling systems, which opened up new possibilities in terms of power consumption. The suitable goal in terms of performance is to reach a 30 ps time resolution while consuming under 40 μW of power per channel.

Timespot1 is currently the latest front-end ASIC of the project. It was developed in 2020 and it integrates 1024 channels fully connectable to both the TimeSPOT sensors. The ASIC has been developed to be part of the small-scale telescope, therefore a dedicated Printed Circuit Board (PCB), TIMESPOT1, was also developed. The PCB has a form factor designed to be a movable part on the demonstrator rail. The read-out system will be implemented in FPGA, with one system hosting up to 8 TimeSPOT1. The read-out not only has to read the data produced by the ASICs, but it must also set the configurations, provide the timing reference and synchronize the boards. Each TIMESPOT1 mounts only one ASIC and constitutes a single layer of the telescope. Multiple PCBs are connected to the same FPGA
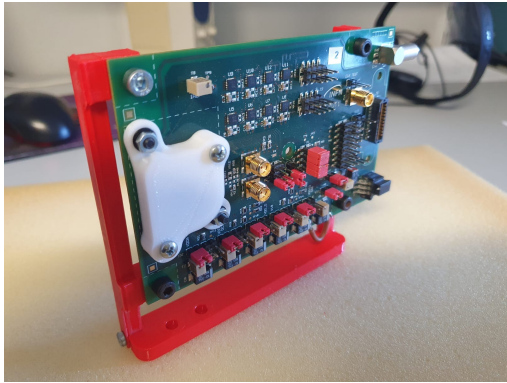
(a) The TIMESPOT1 PCB.
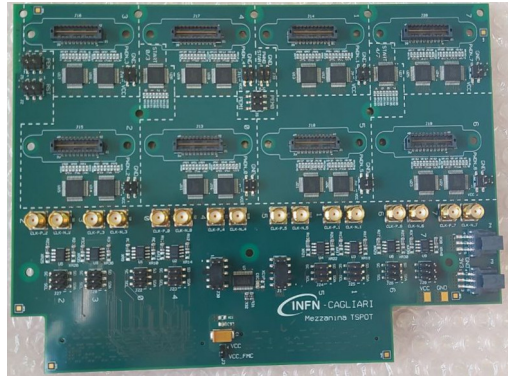
(b) The TimeSPOT Mezzanine Board.

(c) A TIMESPOT1 connected to the mezzanine with a 1 m long cable.

Figure 2.14: Photograph of the components used to interface the up to 8 Timespot1 ASIC to the read-out system.

via the TimeSPOT Mezzanine. This board collects the 8 channel high-speed busses from the Timespot1 and redirects them to the FPGA board. The photograph of the two PCBs and their connection are shown in figure 2.14.

Timespot1 was preliminary tested in 2021. During this test the prototype was not tested with the sensor since the hybrid prototypes were not yet ready. The hybrid has been recently completed, its test will begin in the near future. The results presented in this thesis are taken from electrical only test performed with built-in circuit for self-test and characterization purposes.

2.2.1 The Timespot0 ASIC

Timespot0[89] is the first prototype of the TimeSPOT ASICs and it measures $1.4\,\mathrm{mm} \times 1.4\,\mathrm{mm}$. It contains individual blocks of the final chain, one disconnected from the other for testing purposes. In particular the chip contains the 8 channels of the first prototype of the analog front-end [87], three different TDCs, a 6-bit DAC and a pair of LVDS driver and receiver. The layout of the ASIC with annotations is shown in figure 2.15. The major problem in the integration of this chip was to fit the number of pads required to operate independently the blocks in this small area.
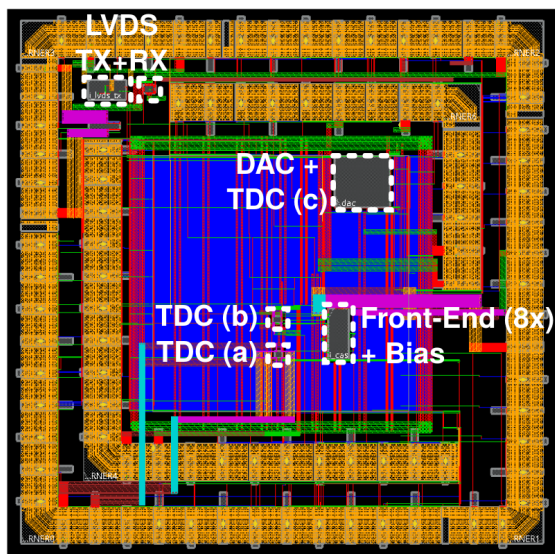
Figure 2.15: Layout of Timespot0 (1.4 mm × 1.4 mm). The implemented blocks are highlighted and annotated. The two double rows of staggered pads can be seen.

The first solution to this problem was to implement the pad frame in a double-row configuration, with the pad staggered. Unfortunately, this remedy was not sufficient, therefore two concentric pad frames like this have been implemented. As a consequence, the required nets can be connected but not simultaneously due to the obstruction of the bonding wires. Therefore the chip can be bonded to the PCB in two different bonding configurations depending on the circuit to be tested. The analog front-end implemented in this prototype is extensively described in chapter 3.

The three TDCs will now be described. The first two designs are totally digital and synthesized, whereas the second is analog and full-custom. The architecture of the first version is shown in figure 2.16a. This TDC is a port of an existing circuit previously implemented in 130 nm [19]. It is based on a DCO, which drives a fast counter performing the phase measurement, with a resolution of about 50 ps. The TDC uses a dithering system to minimize the systematic errors introduced by each delay unit.

The second architecture, shown in figure 2.16b, is based on a tapped delay line looped in a ring oscillator. The clock of the oscillator is then counted with two counter working on opposite edges. Since this architecture is fully digital, the delay of each unit can not be adjusted individually. Therefore the clock frequency is controlled by changing the number of elements in the line.

The third TDC features a two-step conversion using time amplification. The scheme is shown in figure 2.16c. The measurement is activated by the incoming signal that starts the first ring oscillator. The next rising edge of the clock stops

(a) TDC (a): DCO with dithering.

(b) TDC (b): DCO with tapped delay line.

(c) TDC (c): twos stages with amplification.
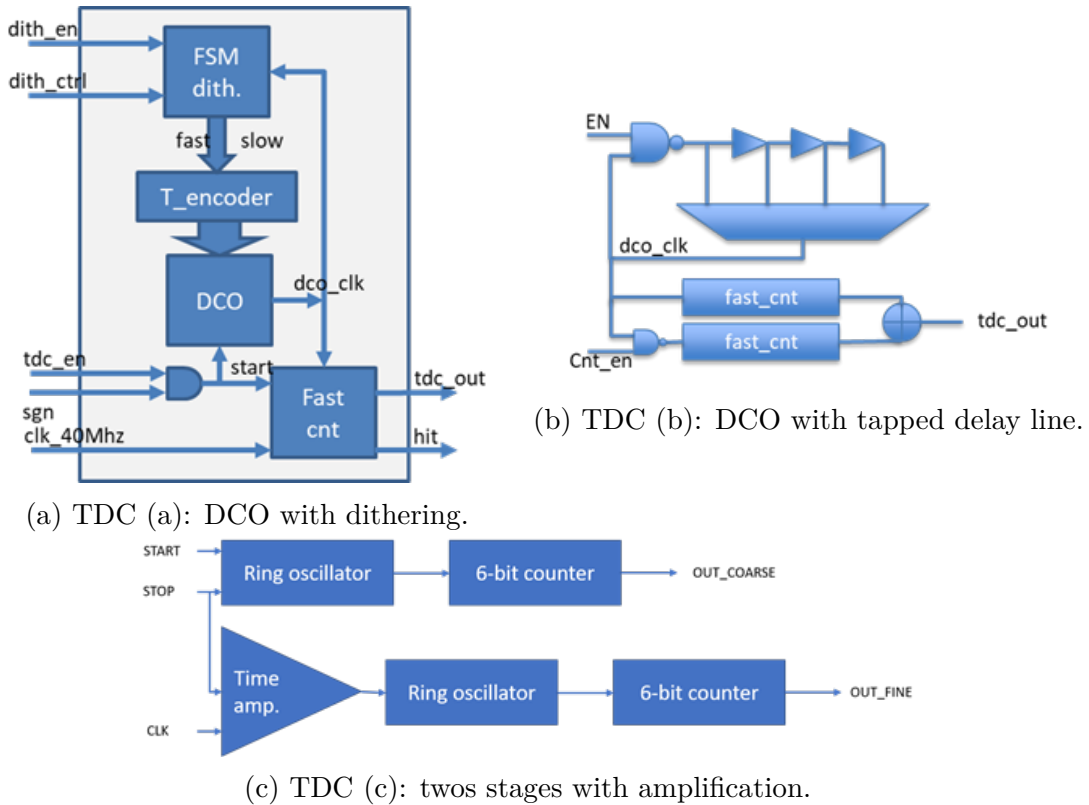
Figure 2.16: Schematic representation of the three TDCs implemented in Timespot0.

the measurement. A 6-bit coarse counter counts the number of oscillations during this period. The residual of the first measurement is fine measured using the second step. The first oscillator is now slowed down, while a second oscillator is started. The number of ticks of the second oscillator elapsed before the next tick of the first oscillator constitutes the fine measurement. This time period is an amplified version of the residual.

The LVDS pair was based on a previously tested design [111]. These LVDS will be used as the IO buffers of the front-end ASIC. Inside Timespot0 they can be looped back in order to test their maximum data rate. Figure 2.17 shows an eye diagram obtained from direct measurements on these LVDS in the loop-back configuration. The LVDS are able to obtain a 1.5 Gbit/s data rate while consuming 2.7 mA of power[11]. The LVDS also constitutes the interface for the input and output signals of the analog front-end. This was done in order to have an interface with less jitter compared to a standard CMOS signal. For this purpose, the total jitter of the combination of the receiver and the transmitter was measured to be 15 ps[88].
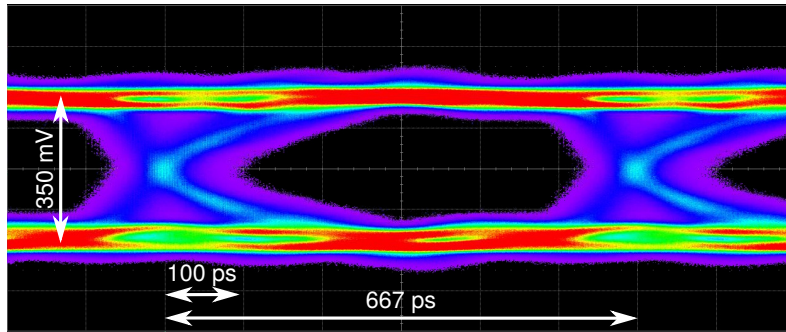
Figure 2.17: Eye diagram obtained from direct measurement of the LVDS receiver and transmitter in loop-back configuration.

2.2.2   The Timespot1 ASIC
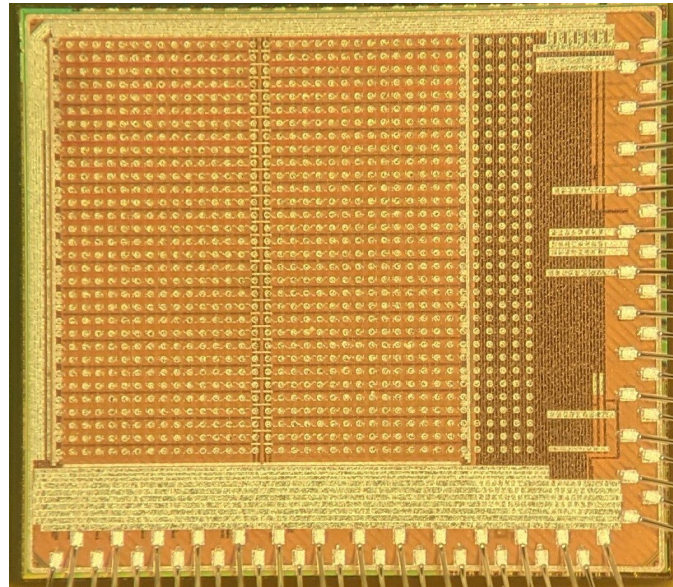
Figure 2.18: Close up photograph of the Timespot1 ASIC ($2.6\,\text{mm} \times 0.24\,\text{mm}$). On the right and bottom edges of the chip the two staggered rows of wire-bond pads are positioned. The $32{\times}32$ bump-bond pads matrix spans from the top-left corner. The last 5 columns of pads are dummy pads inserted for mechanical stability. The two top metal layers used for power and ground delivery can also be seen.

The Timespot1 ASIC [18] is designed to be compatible with the TimeSPOT sensors. In any case it is a front-end pixel ASIC suitable for chip-to-hip bump-bonding with 3D pixel arrays, with a pixel size of $55\,\mu\text{m} \times 55\,\mu\text{m}$. A photograph of the chip is shown figure 2.18. The Timespot1 chip integrates 1024 channels, organized in a $32{\times}32$ matrix, with each channel equipped with its own Analog

Front-End (AFE) and TDC. A schematic block representation of the architecture is shown in figure 2.19. The same figure also shows the full layout of the chip, with the principal structures annotated.

The input channels are organized in two blocks of 512 pixels, each one consisting of 2 groups of 256 channels. Each group is connected to one of four Read-Out Tree (ROT) blocks. Each ROT collects data from the active channels, assigns them a global timestamp and sends the formatted data to one of the two serializers connected to LVDS drivers. The LVDS are used to output data towards the readout system. In total, 8 LVDS drivers are integrated, operating at a data rate of 1.28 Gbit/s each, with an overall nominal data throughput of 10.24 Gbit/s.

From the floor plan point of view, the 32×32 array is arranged in two mirrored 16×32 pixels blocks. The size of each channel is 50 µm × 55 µm. Therefore the pitch is shorter in the horizontal direction compared to the bond-pad matrix. A redistribution layer is used to connect each channel to its own bond-pad. This was done in order to make the design compatible with a future TSV compatible version. In this way, 80 µm are saved every 16 pixels in order to be used for distribution of analog references and power supplies on one side (Analog Column), and for digital power supplies and data transmission lines on the other side (Digital Column). Particular care was used to maintain as separate as possible the digital domain from the analog one. In particular the analog part of channels were integrated in an Analog Row with its own substrate bias and power nets. In the same way the digital part (TDC and control logic) is confined in its dedicated area (Digital Row). The interconnections to the respective service columns were realized only in the opposite sites, making them independent. The two domains creates two matching combs like structures.

The readout scheme is totally data-driven and triggerless, with timestamping from an externally provided signal. Timespot1 also integrates several structures for internal services: two voltage Bandgaps, 8 $\Sigma\Delta$ DACs for voltage references, a PLL and two DCOs generating the needed clock frequencies.

The chip configuration is handled using a slow control interface I^2C like protocol [46]. The top level and each Digital Column and Row features its own independent target interface with its unique address. The AFE requires only one reference current in order to be operated. Dedicated registers are used to configure its power consumption and operating parameters. The AFE shows an updated architecture compared to the one of the previous version with the aim of boosting its time resolution and power consumption. Most of its components have been changed, with only the core discriminator being identical to the previous version. A detailed description of the AFE is presented in chapter 5. The TDC and the data-transmission will now be briefly described.
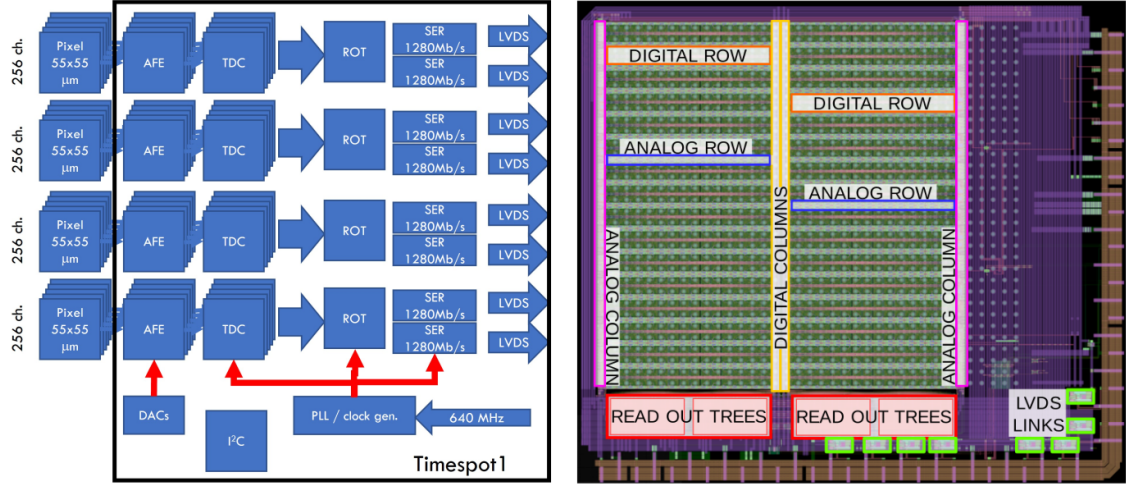
Figure 2.19: The Timespot1 ASIC. On the left: schematic block representation of the architecture. On the right: layout with annotated floor-plan. The two matching comb structures of the analog and digital domains can be seen.

TDC

The TDC measures the TA of the AFE signal, in terms of its phase with respect to a reference clock running at 40 MHz. At the same time the TDC measures the ToT of the signal to correct the time-walk effect. The TDC architecture has been completely changed compared to the previous version.

The new TDC is based on a Vernier architecture, with two identical DCOs working at slightly different frequency (the second is faster than the first one). Figure 2.20 shows the pixel architecture, including a schematic block of the TDC architecture. The DCOs are made with a tapped delay line with tunable length. Each one of the first four DCO step element is made with three different tri-state buffers in parallel, each one with different drive strength. The fine delay is tuned by enabling one of them at a time. The last three stages are fixed delay cells used for coarse adjustment. Each DCO drives a fast counter. When in stand-by, both DCO are inactive, greatly reducing the static power consumption. The signal rising edge starts the first DCO, whereas the second is started by the next reference clock edge (stop signal). A Coincidence detector Circuit (CC) is used to detect when the two positive edges rise at the same time, indicating the end of the measurement. A third counter is driven by the first DCO to calculate the signal ToT.

The TA is computed using:

$$TA = (CNT_{slow} - 1)T_{slow} - (CNT_{fast} - 1)T_{fast} \tag{2.1}$$

where T_{slow} and T_{fast} are the periods of the two DCO while CNT_{slow} and CNT_{fast} are the related counter values. The theoretical resolution of a Vernier TDC is given
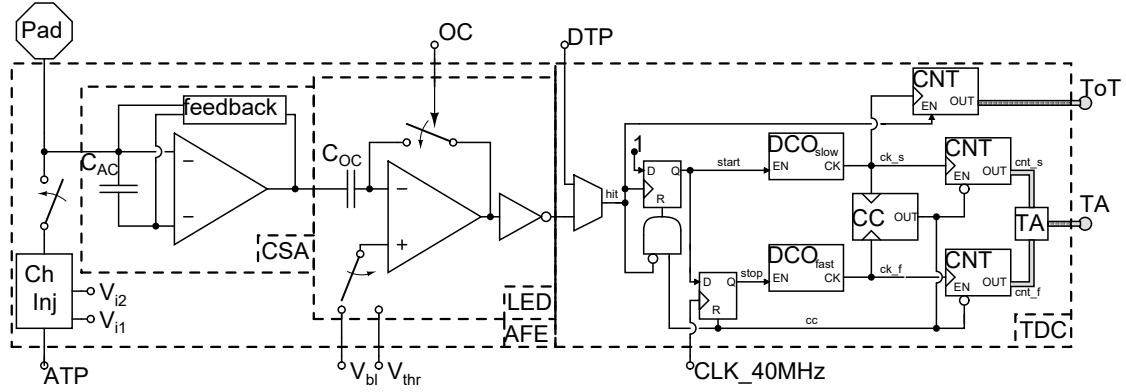
Figure 2.20: Schematic block representation of the Timespot1 pixel architecture. The scheme shows all the signals and voltages required to operate the channel and also the test pulses used in the analog part (ATP) and in the digital part (DTP).

by the periods difference:

$$Res_{th} = T_{slow} - T_{fast} \tag{2.2}$$

The DCOs are calibrated at the beginning to set the resolution and all the parameters needed to calculate the time. The calibration process sets the period difference ($T_{slow} - T_{fast}$) between the two DCOs to be consistent with the required resolution. During calibration, the DCO period is measured and adjusted to obtain the desired value. At the end of the calibration process, the working periods of the two DCOs (T_{slow} and T_{fast}) are stored internally for TA calculation according to (2.1). The full calibration needs less than 4 µs to be completed. The procedure is performed in two steps:

1. The DCO_{slow} is set to generate a clock period around 1.1 ns.

2. The DCO_{fast} is set to have a period slightly smaller than the DCO_{slow}, according to the required resolution. The period is set iteratively while monitoring the period difference $T_{slow} - T_{fast}$ until the desired one is met.

The TA value is calculated internally according to (2.1). The ToT is simply measured by counting the number of DCO_{slow} oscillations occurred while the input pulse is active. The TDC output word is composed of 23 bit where 15 bit are reserved for the TA measurement and 8 for the ToT one.

For debug purposes it is possible to send out the counter values (CNT_{slow} and CNT_{fast}) instead of the TA calculated internally. Furthermore, the TDC is able to self generate test pulses to verify the TA measurement (with 7 different phases) and TOT measurement (with 32 different widths).

The TDC timing performance has been tested in standalone tests[86]. These tests are also instrumental for the AFE characterization since the TDC represents
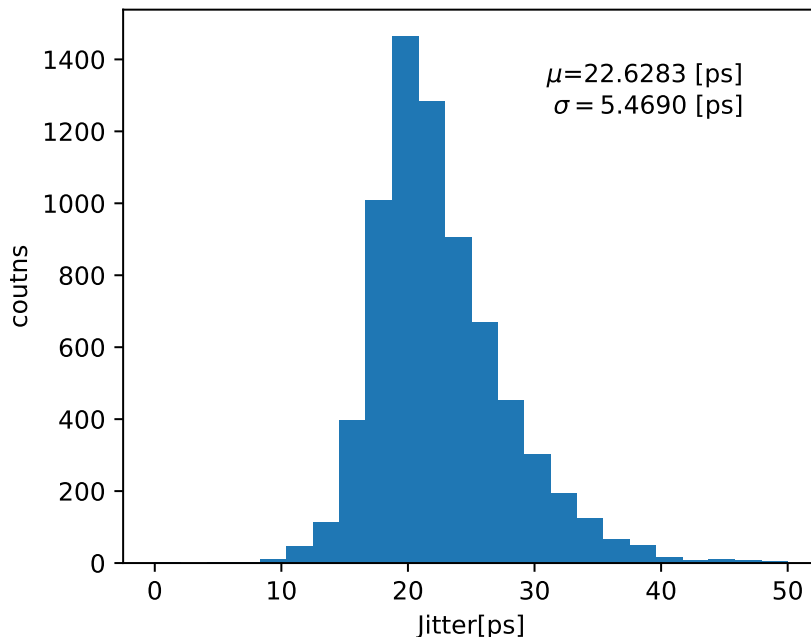
Figure 2.21: Histogram of the measured TA resolution of the Timespot1 TDC. Each entry is the standard deviation of 200 TA measurments. The measurement set includes 1024 channels with 7 different input phases.

its measurement tool. In order to trigger the TDC it is possible to inject test-pulses with 7 different phases and 32 different widths. A histogram collecting multiple measurements of all the possible phases across the whole matrix is shown in figure 2.21. The TDCs have an average resolution of 25 ps with an RMS variation of 5.4 ps from one channel to the others. The variations show a slight dependence to the channel position and input phase.

Data-Transmission

Each TDC is connected throughout the Digital Column via its dedicated 160 MHz serial line to the transmission and processing logic. At every hit, the TDC sends its 23 bit word. Figure 2.22 shows the block diagram of the data transmission circuit that contains the ROT and the data serialization circuit.

Each hit-data coming from the TDC is paired with the corresponding timestamp information (9 bit) that indicates the sequential number of a counter on the 40 MHz clock when the hit data is generated. This counter can be started using a dedicated external signal. This step is required at this point since the successive logic does not have a fix latency. The data is then stored in a register in order to reduce
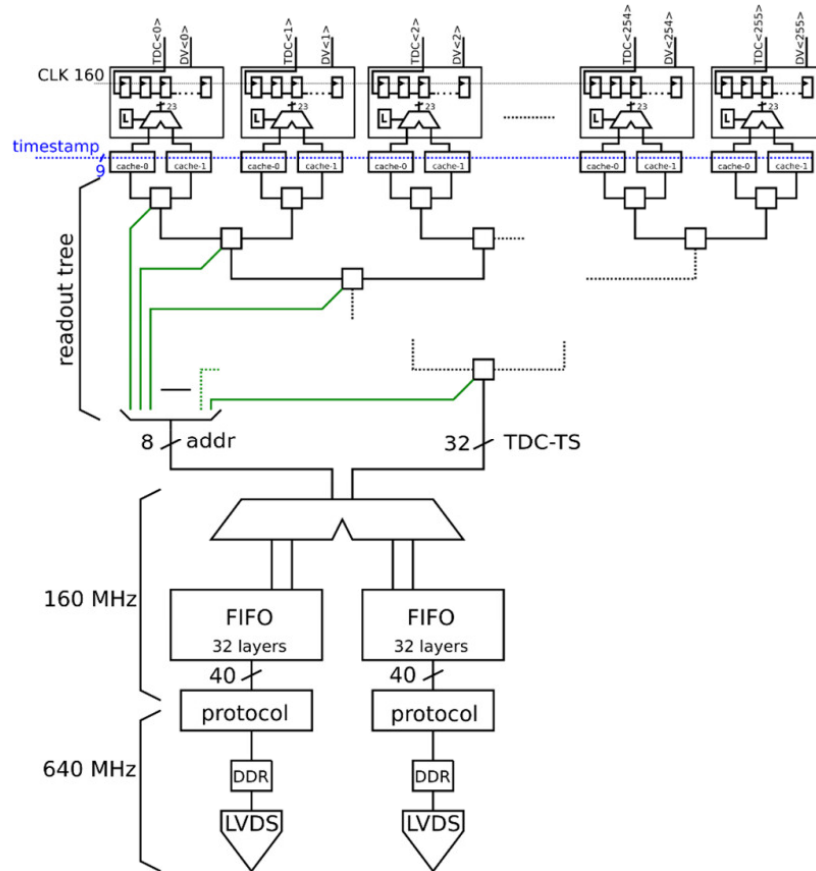
Figure 2.22: Schematic block representation of the data transmission circuit. The data path is organized from the pixels to output drivers from top to bottom.

the risk of data-loss. While the data is stored into the registers, an asynchronous ROT [32] sequentially reads the data from the buffers at a rate of 160 MHz and frees them. The ROT is a fully combinational circuit that generates the geographical coordinates of the pixel (8 bit) and implements a zero-suppression feature.

The output of the ROT is organized as follows: 8 bit address, 23 bit TDC data and 9 bit timestamp information. The resulting 40 bit string is fed to one of two 32 word deep FIFO working at 160 MHz. The FIFOs are used to mitigate activity peaks that otherwise could cause data loss. The 40 bit output of the FIFO is encoded using a custom transmitting protocol. The protocol is organized in 8-bit words. In particular, each communication is constructed as follows: when data is present at the FIFO output, it is divided in five bytes preceded by a header byte, otherwise an idle byte is transmitted. The idle and header words can be configured via I^2C.

The protocol block outputs each byte at a frequency of 160 MHz. The byte is then serialized with a 640 MHz Double Data Rate (DDR) and transmitted to

the output using an LVDS driver working at 1.28 Gbit/s. The maximum output bandwidth of Timespot1 is defined by the 8 LVDS cumulative bandwidth, resulting in 10.24 Gbit/s. The per-channel sustainable hit-rate by the readout processing logic ranges from 3 MHz to 200 kHz depending on the channel occupancy. The first case corresponds to a single active pixel, whereas the second ones corresponds to a totally uniform hit occupancy.

Chapter 3

First Prototype: the Timespot0 Analog-FE

This chapter illustrates the design process of the AFE implemented in Timespot0. A general description of the Timespot0 ASIC can be found in 2.2.1. The chapter begins with a short overview of the architecture in section 3.1, whereas section 3.2 derives analytically the circuit characteristics crucial for timing. Finally, subsection 3.3 will discuss the physical implementation of the circuit. For the sake of convenience, the transistor level schematic, the transistor sizing and the channel layout are shown at the end of chapter.

3.1 Architecture Overview
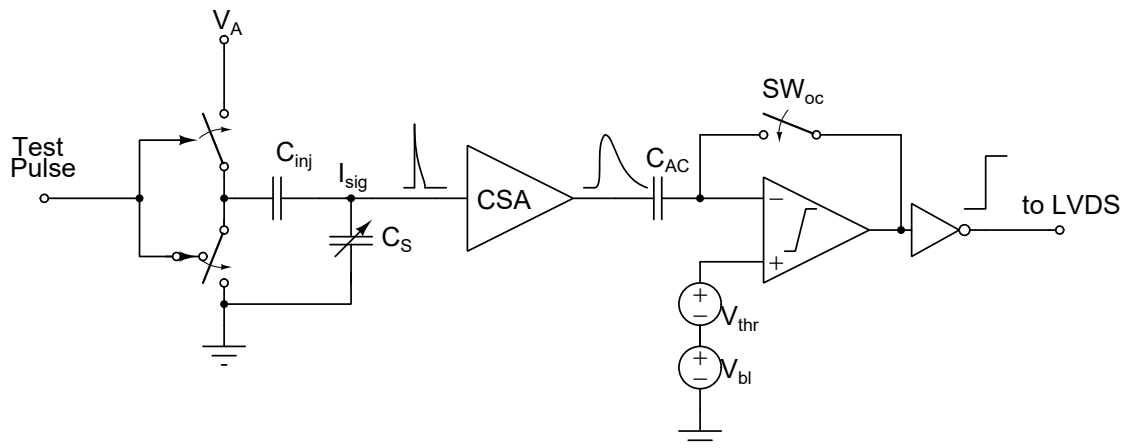
Figure 3.1: Schematic block representation of the implemented analog front-end.

The TimeSPOT AFE must reliably provide the timing pulse to the TDC while

59

minimizing its jitter. The AFE is optimized to be coupled with the TimeSPOT sensors described in subsection 2.1.1. The Timespot0 prototype, as explained in 2.2.1, is not connectable to an actual sensor and its output is directly connected to an LVDS driver. Therefore, the front-end IO interface is digital. All the required bias currents and reference voltages are provided externally. A schematic block representation of the circuit is presented in figure 3.1.

The core of the signal processing chain is formed by a CSA connected to a leading edge discriminator. The latter compares the CSA output with a given threshold in order to produce a digital pulse suitable for the TDC. The leading edge discriminator must provide two critical measurements: the TA and the ToT, which is used to infer the signal charge and perform time-walk corrections. The signal path is designed to minimize the number of stages with the purpose of limiting both the total power consumption for a given bandwidth and the number of noise injectors.

The CSA delivers an output voltage proportional to the integral of the input current. This configuration was chosen because of its lower noise for a given bias current[97] and its excellent frequency stability. The CSA is formed by a core amplifier acting as the gain stage, a feedback capacitor (C_f) for charge integration, a Source Follower and a discharge path implemented with a Krummenacher filter [53]. The circuit is optimized for negative input signals (i.e. sensors collecting electrons). The core amplifier is based on a n-type telescopic cascode amplifier with split bias branches. This topology allows to increase the DC current in the input transistor without reducing the equivalent resistance of the active load. The source follower is used to mask the high-impedance output node of the core amplifier form the capacitive load presented by the following stage. It is implemented with an n-type transistor due to its superior driving capability for the intended signal polarity. The adopted Krummenacher filter has n-type input stage and p-type loads. The circuit provides a constant-current discharge for C_f and filters out the sensor leakage current. The discharge time is crucial in defining the AFE maximum signal rate.

The leading edge discriminator is composed of a core gain stage and an offset compensation circuit. The discriminator core is a three stage voltage amplifier. The first stage is a differential pair which compares the input signal presented to M_{d2} with a given threshold applied to M_{d4}. The second stage is a cascode common source delivering a robust signal with steep edges to the output inverter, that generates full CMOS levels. The discriminator core is AC-coupled to the CSA through capacitor C_{AC}. The DC level at the input node is set via the digitally controlled offset compensation circuit. During the Offset Compensation (OC) procedure the voltage applied to the gate of M_{d4} is saved across C_{AC}. By registering the desired threshold, the circuit is also compensating the channel intrinsic DC variations. Thanks to the auto-zeroing, it is possible to work with a common threshold for multiple channels using two shared reference voltages. The main benefits of this approach compared to

60
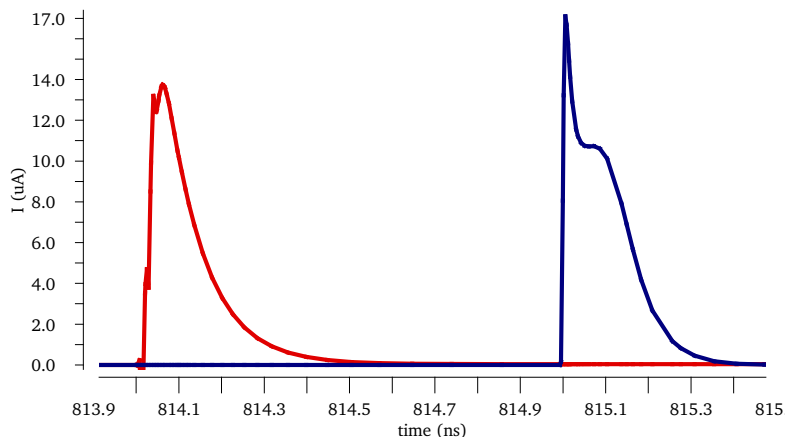
Figure 3.2: Comparison of two simulated input signals: in blue the one extracted from a sensor physics simulation, in red the one produced by the charge injection circuit. The related input charge is 2 fC.

a per-pixel threshold tuning with a DAC are: smaller area and power consumption, overall simpler architecture and less cumbersome calibration procedures.

The charge injection circuit allows to pulse the channel with a current signal with programmable charge provided by a voltage step of amplitude V_A applied to the injection capacitor C_{inj}. This voltage step is produced locally by switching between a reference voltage V_A and ground. The capacitance presented by the sensor, C_S, is emulated by an array of three capacitors that are independently connected to or disconnected from the CSA input through dedicated switches. The expected sensor signal shape can be approximated by a suitable sizing of this network. A simulated signal produced by the charge injection circuit is shown in figure 3.2. This local generation of the input signal enables a stable timing characterization.

A transistor level schematic of the AFE and its sizing are presented in figure 3.7 and table 3.1 respectively, while its layout is shown in figure 3.8.

3.2 Front-End Operation

In this section the timing performance of the AFE will be analytically derived. In order to describe the key aspects of the input stage operation a small signal equivalent circuit is derived from the scheme reported in figure 3.7. The resulting model is shown in figure 3.3.

The relevant small signal parameters of the transistors are the gate transconductance g_m and the output resistance r_0. The bulk-effect induced transconductance g_{mb} has been omitted for the sake of simplicity since its contribution is negligible in the overall estimation of the circuit performance. For g_m the following expression is adopted [27]:
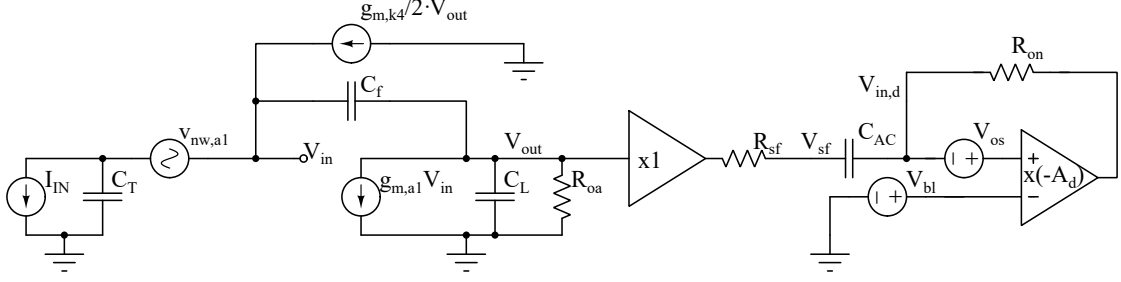
Figure 3.3: Small signal equivalent model of the Analog Front End. This circuit is derived from the one presented in figure 3.7. The portion of the circuit between V_{in} and V_{out} is used to compute the small signal behavior of the CSA. The rest of the circuit is used to analyze the leading edge discriminator operation during the OC procedure.

$$g_m = \frac{I_{ds}}{n\phi_T\sqrt{I_C + \frac{1}{2}\sqrt{I_C} + 1}} \tag{3.1}$$

where I_{ds} represents the drain-source current of the transistor. The inversion coefficient I_C is defined as (3.2)

$$I_C = \frac{I_{ds}}{2n\mu C_{ox}\left(\frac{W}{L}\right)\phi_T} \tag{3.2}$$

where W/L is the transistor aspect ratio and ϕ_T is the thermal voltage. The parameters n, μ and C_{ox} are derived from single transistor simulations for both NMOS and PMOS devices. In the case of the utilized CMOS technology node r_0 can be evaluated[13] as:

$$r_0 = \frac{1}{I_{ds}}\frac{\lambda_C L_C}{LV_E\left(1 + \frac{V_{ds} - V_{ds,sat}}{V_E}\right)} \tag{3.3}$$

in here V_{ds} is the drain-source voltage and $V_{ds,sat}$ it is the minimum value required to keep the transistor in saturation. V_E is a voltage determined experimentally, L_C is a length characteristic of a given technology and depends on oxide thickness and junction depth, while λ_C is a dimensionless fit parameter. In practice, the use of this formula is unfeasible due to presence of various parameters which are not disclosed by the foundry. Therefore, r_0 values for the different conditions have been extracted from the simulator.

In noise calculations the thermal Power Spectral Density (PSD) of a given resistor R is computed as:

$$v_{nt}^2 = 4k_B T R \tag{3.4}$$

where k_B is the Boltzmann constant and T is the absolute temperature. In case of the input-referred thermal PSD voltage of a transistor the following equation is used:

$$v_{nw}^2 = 4k_BT\alpha_w\gamma\frac{1}{g_m} \tag{3.5}$$

where α_w is the excess noise factor[38] and γ is the inversion factor given by:

$$\gamma = \frac{1}{2} + \frac{I_C}{6(I_C + 1)} \tag{3.6}$$

Flicker noise was not considered in first-cut hand calculations, but it was optimized directly with Simulation Program with Integrated Circuit Emphasis (SPICE) simulations.

The equivalent resistance of a transistor in linear region has been quantified as:

$$R_{on} = \frac{1}{\mu C_{ox}\frac{W}{L}(V_{gs} - V_{th})} \tag{3.7}$$

where V_{gs} and V_{th} are respectively the gate-source voltage and the threshold voltage.

As described in more detail in the following, other small signal elements present in figure 3.3 are an approximation of given sub-circuits, such as the Krummenacher feedback and the output buffers. Similarly, the different capacitors inserted in the model are the sum of the dominant capacitors connected to the considered node.

3.2.1 CSA Core Amplifier Gain

The first approximation is about the CSA core amplifier which constitutes the active element of the small signal model. The core amplifier has been approximated with a single pole gain stage made of a transconductance term G_{in} and an output resistance R_{oa}:

$$A_{ol} = -G_{m,in}R_{oa} \tag{3.8}$$

These two terms are computed using (3.9) and (3.10).

$$G_{m,in} = g_{m,a1}\frac{r_{[0,a1//casP2]} \cdot (1 + g_{m,a2} \cdot r_{0,a2})}{r_{0,a2} + r_{[0,a1//casP2]} \cdot (1 + g_{m,a2}r_{0,a2})} \tag{3.9}$$

$$R_{oa} = r_{[casN//casP1]} \tag{3.10}$$

The related transistors used for computation are indicated with the subscript labels. The expression $r_{[A//B]}$ stands for the parallel between resistor A and B. The 'cas' term indicates the equivalent resistance of the cascode configuration formed by the indicated branches. For example the r_{casP2} term is computed as follows:

$$r_{casP2} = g_{m,a5}r_{0,a5}r_{0,a6} + r_{0,a5} + r_{0,a6} \tag{3.11}$$

r_{casP1} is computed in the same way whereas r_{casN} is computed by using (3.11) recursively on transistor M_{a2} and r_{casP2}.

3.2.2 CSA Jitter Computation

The most relevant factor in defining the AFE timing performance is the jitter of the first CSA stage σ_{tn}. In this section this parameter will be computed on the basis of the CSA output noise $v_{nw,out}$ using the relation:

$$\sigma_{tn} = \frac{v_{nw,out}}{\frac{d}{dt}(V_{out})} = \frac{v_{nw,out}}{V'_{out}} \tag{3.12}$$

in which the output signal slew-rate V'_{out} converts the voltage noise into a time fluctuation. In order to compute these two parameters, the CSA transfer function must by firstly derived.

The CSA is modelled as a core transconductance amplifier with a feedback-capacitor, a DC feedback-path and both input and output capacitive loads referenced to ground. The core element has been optimized to maximize G_{in} by leveraging the condition (3.13) obtained from (3.9):

$$r_{0,a2} \ll r_{[0,a1//casP2]} \cdot (1 + g_{m,a2}r_{0,a2}) \tag{3.13}$$

In this condition the transconductance of the input cascode can be approximated to the one of the input transistor $g_{m,a1}$ for sake of simplicity.

The feedback capacitance is quantified as the value of C_f plus the extracted parasitic capacitance between the input and output node.

In the frequency range of the input signal, C_k locks the voltage at the gate of M_{k3}, leaving the differential pair inside the Krummenacher filter as the only feedback path. The equivalent impedance for small signal is $2/g_{m,k4}$. This resistance is the one used to derive the CSA transfer function, however for large signals the feedback path exhibits a non-linear behavior saturating to a constant current discharge equal to $I_k/2$. This simplified behavior of the Krummenacher filter represents the actual operation of the circuit once stability conditions described in 3.3.3 are met.

The input and output capacitance C_T and C_L are computed as:

$$C_T = C_s + C_{gs,a1} + C_{gb,a1} + C_{dg,k2} + C_{dg,k3} + C_{T,par} \tag{3.14}$$

$$C_L = C_{gs,s1} + C_{gb,s1} + C_{gs,k4} + C_{gb,k4} + C_{L,par} \tag{3.15}$$

where C_s is the programmed sensor capacitance and the last terms of both equations are the parasitic capacitance from the specific node to DC nodes.

Solving the nodal equation in the frequency domain for this portion of the circuit, the transimpedance transfer function results in:

$$\begin{aligned} T_A(s) &= \frac{V_{out}(s)}{I_{in}(s)} \\ &= \frac{g_{m,a1} - C_f s}{\zeta s^2 + \left[\frac{C_T}{R_{oa}} + C_f \left(g_{m,a1} + \frac{1}{R_{oa}} - \frac{g_{m,k4}}{2}\right)\right] s + \frac{g_{m,a1} g_{m,k4}}{2}} \end{aligned} \tag{3.16}$$

$$\zeta = C_T C_f + C_T C_L + C_f C_L \tag{3.17}$$

This relation can be simplified by neglecting the zero in the numerator and assuming:

$$g_{m,a1} \gg g_{m,k4} \tag{3.18}$$

$$g_{m,a1} \gg \frac{1}{R_{oa}} \tag{3.19}$$

In this way the following equation is obtained:

$$T_A^*(s) = \frac{2}{g_{m,k4}} \frac{1}{\frac{2\zeta s^2}{g_{m,a1}g_{m,k4}} + \left(\frac{C_T}{R_{oa}g_{m,a1}} + C_f\right)\frac{2s}{g_{m,k4}} + 1} \tag{3.20}$$

This equation can be re-written as:

$$T_A^*(s) = \frac{2}{g_{m,k4}} \frac{1}{(1 + \tau_r s) \cdot (1 + \tau_f s)} \tag{3.21}$$

where τ_r and τ_f are the rising and falling time constants of the output signal. The values of these constants can be estimated assuming that the fall time is much longer than the rise time and using the following approximation:

$$(1 + \tau_r s) \cdot (1 + \tau_f s) \overset{\tau_f \gg \tau_r}{\simeq} \tau_r \tau_f s^2 + \tau_f s + 1 \tag{3.22}$$

which yields:

$$\tau_r = \frac{\zeta}{\frac{C_T}{R_{oa}} + g_{m,a1}C_f} \tag{3.23}$$

$$\tau_f = \frac{2}{g_{m,k4}} \left(\frac{C_T}{R_{oa}g_{m,a1}} + C_f\right) \tag{3.24}$$

With this transfer function it is now possible to convert the CSA local noise sources into the output noise. The major noise source of the CSA is the thermal one of M_{a1}. Its contribution is referred to the CSA input using (3.5). This voltage PSD is converted to a current one by C_T:

$$i_{nw}^2(s) = s^2 \cdot C_T^2 v_{nw}^2 \tag{3.25}$$

In order to calculate the RMS output noise, this component is applied to the transfer function (3.21) and integrated over the whole spectrum using again approximation (3.22):

$$< v_{nw,out} >^2 = \frac{4C_T^2}{g_{m,k4}^2} \int_0^\infty \left| \frac{2\pi f v_{nw}}{(1+\tau_r 2\pi f) \cdot (1+\tau_f 2\pi f)} \right|^2 df$$

$$\overset{\tau_f \gg \tau_r}{\simeq} \frac{C_T^2 v_{nw}^2}{g_{m,k4}^2} \cdot \frac{1}{\tau_f^2 \tau_r} \tag{3.26}$$

$$= \frac{k_B T \alpha_w \gamma C_T^2}{\zeta \left(\frac{C_T}{g_{m,a1} R_{oa}} + C_f \right)}$$

The CSA output noise is strongly related to the input capacitance C_T. The second largest noise source in the circuit is the one generated by the feedback circuit, $i_{n,fb}$. When converted to an output voltage it can be expressed as:

$$v_{n,fb}^2 = \frac{8k_B T \alpha_w \gamma}{g_{m,k4}} \tag{3.27}$$

This component is negligible when compared to $v_{nw,out}$. The value of $v_{nw,out}$ computed using (3.26) is $1.6\,\text{mV}$.

In order to compute V'_{out}, the transfer function (3.21) must be transformed in the time domain and then derived:

$$T_A^*(t) = \mathcal{L}^{-1}[T_A^*(s)]$$

$$= \frac{2}{g_{m,k4}} \cdot \frac{e^{-\frac{t}{\tau_f}} - e^{-\frac{t}{\tau_r}}}{\tau_f - \tau_r} \tag{3.28}$$

Ideally the best value of σ_{tn} is computed for the time value which maximizes the derivative of (3.28). Practically this approach will not lead to an intelligible explicit result while also not representing a valid working point. The threshold computed in this way will be positioned at about half of the signal amplitude. Therefore the optimal threshold in terms of jitter minimization is dependent on the input charge and highly susceptible to the time-walk effect. A real case threshold must be set as close as possible to the baseline value in order to capture a large range of input signals and minimize the time-walk effect. The lower limit is defined by the output noise. Imposing a threshold at $5 \cdot v_{nw,out}$ above the baseline will limit the number of noise induced hits. For the sake of calculation V'_{out} was computed at time zero. This approximation is acceptable considering that $5 \cdot v_{nw,out}$ is small compared to the $300\,\text{mV}$ amplitude of the output signal generated by the most probable input charge ($2\,\text{fC}$). V'_{out} is therefore computed as:

$$V'_{out}(t=0) = \left[\frac{dT_A^*(t)}{dt} \right]_{t=0} \cdot Q_{in}$$

$$= \frac{2Q_{in}}{g_{m,k4} \tau_f \tau_r} \tag{3.29}$$

$$= \frac{g_{m,a1}}{\zeta} Q_{in}$$

In this expression the input signal has been considered to be a delta-like current pulse with a related integrated charge Q_{in}. This approximation is reasonable considering the short time development of the input signal compared to the CSA integration time ($\sim 100\,\mathrm{ps}$ versus $\sim 10\,\mathrm{ns}$). The dependence on Q_{in} will lead to time-walk. The $g_{m,a1}$ term accounts for the dependency on the power consumption via the input transistor I_{ds}. ζ represents the contributions of the sensor and transistor capacitance.

Finally, using (3.29) and (3.26) the CSA output jitter σ_{tn} can be estimated as the ratio between output noise and slew rate:

$$\begin{aligned}
\sigma_{tn}(0) &= \frac{v_{nw,out}}{V'_{out}(0)} \\
&= \frac{v_{nw} C_T \sqrt{\tau_r}}{2Q_{in}}
\end{aligned} \tag{3.30}$$

This equation shows a direct relation between σ_{tn} and v_{nw}. The direct relation between the jitter and C_T/Q_{in} shows the dependence to the sensor technology. The connection with the power consumption is expressed in terms $\sqrt{\tau_r}$ and v_{nw} with both containing the $1/\sqrt{g_{m,a1}}$ term. The σ_{tn} value computed with (3.30) is $64\,\mathrm{ps}$.

3.2.3 Leakage Current Compensation

At low frequency the feedback path provided by the Krummenacher filter is not limited by C_k. In this situation the output node drives the gate of M_{k4} controlling the current which flows through M_{k5}. This current is then copied into M_{k3} which injects it in the input node. Therefore the input impedance of the CSA is defined by the current sourced by M_{k3} and can be computed as:

$$Z_{lf} = \frac{V_{out}}{I_{M_{k3}}} = \frac{2C_k s}{g_{m,k3} g_{m,k4}} \tag{3.31}$$

This impedance exhibits an inductive behaviour acting as a high pass filter with a cutoff frequency $f_c = g_{m,k3} g_{m,k4}/2C_L$.

3.2.4 Signal Discrimination

In order to operate the leading edge discriminator in signal discrimination mode, the signals SW_1 and SW_2 are driven low. In this condition M_{o1} and M_{o2} act as open switches disconnecting the compensating capacitor C_o and placing the leading edge discriminator in an open loop configuration. M_{o4} is also switched off driving the gate of M_{o3} in high-impedance which is in turn bypassed with a direct connection between its source and drain terminals.

The leading edge discriminator acts as a high gain differential amplifier with a gain equal to the product of the gain of the two stages of the core discriminator:

$$
\begin{aligned}
A_d &= A_{d1} \cdot A_{d2} \\
&= g_{m,d2} r_{[r_{0,d4}//r_{0,d5}]} \cdot g_{m,d6} r_{[r_{casNd}//r_{0,d8}]}
\end{aligned}
\tag{3.32}
$$

In this way the difference between the input signal in $V_{in,d}$ and the desired threshold V_{thr}^* is amplified to saturation producing a digital pulse which is then inverted and buffered by the inverter formed by M_{d9} and M_{d10}.

3.2.5 Discrete Time Offset Compensation

In the offset compensation phase the voltage on the inverting input of the leading edge discriminator is firstly set to the desired baseline voltage V_{bl}^*.

Then SW_2 is driven high. In this state, M_{o1} connects the discriminator output $V_{out,d}$ to the load capacitor C_o which limits the leading edge discriminator band-width making it more stable when connected in closed loop. M_{o4} connects the gate of M_{o3} to $\overline{SW_1}$.

Subsequently SW_1 is driven high. M_{o2} closes the discriminator core in to an almost unitary loop with a feedback resistance R_{on}. M_{o3} is counter-driven in respect to M_{o2} compensating its charge injection through its gate capacitance. In this state the non-inverting input node is driven to a voltage V_{bl}.
This voltage is saved in C_{AC} which acts as an analog memory and provides AC coupling to the source follower.

After a certain compensation time T_{comp}, SW_1 is driven low in order to re-establish the open loop configuration. Then, the inverting node is placed at the desired threshold value V_{thr}^* in order to avoid noise triggers on the baseline. Lastly, SW_2 is driven low reestablishing the maximum signal band-width.

Both the actual threshold V_{thr} and baseline V_{bl} are affected by core discriminator offset voltage V_{os}. However, since the offset is added with the same polarity, its effect will be compensated by the procedure. A waveform graph of this procedure is shown in figure 3.4.

This effect can be described with an equivalent small signal model. The discriminator offset is modelled as a small signal input voltage V_{os}. Using this model the voltage compensation on the non-inverting terminal ΔV_{comp} can be computed as:

$$
\Delta V_{comp} = \frac{A_d \beta}{1 + A_d \beta} V_{os}
\tag{3.33}
$$

where β is the feedback factor:

$$
\beta = \frac{R_{sf}}{R_{on} + R_{sf}}
\tag{3.34}
$$
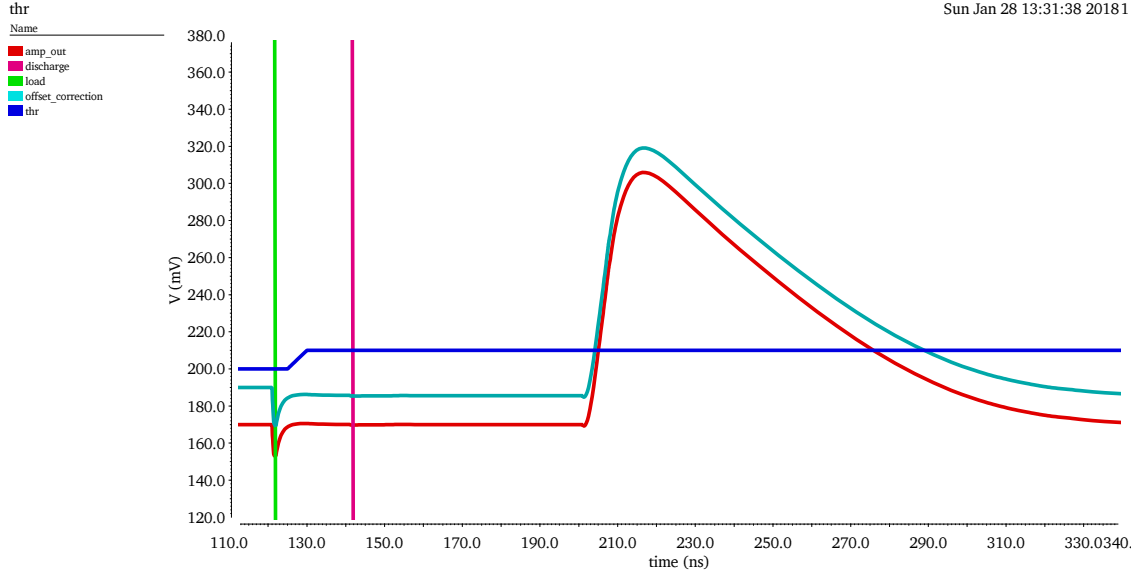
Figure 3.4: Simulated waveform of an offset compensation procedure. First of all an external voltage (blue line) is applied at the desired baseline level. Then the OC procedure is initiated in correspondence of the green line. The discriminator DC level (light blue) is shifted toward the desired level minus its intrinsic offset. The threshold (blue line) is then moved to the desired value relative to the stored baseline. After the purple line, the channel is ready to process incoming signals (in red) with the compensated DC level.

which is defined by R_{on} computed using (3.7), and the source follower output resistance R_{sf} computed using:

$$R_{sf} = \frac{r_{0,s2}}{1 + g_{m,s2} \cdot r_{0,s2}} \tag{3.35}$$

Ideally this procedure will totally compensate V_{os} for an infinite value of $A_d \cdot \beta$. However with a finite A_d value and a non-zero R_{on} it will produce a systematic error V_{err}:

$$V_{err} = V_{os} - \Delta V_{comp} = \frac{1}{1 + A_d\beta}V_{os} \tag{3.36}$$

This computation suggests that V_{os} is reduced by a factor 6×10^{-4}.

3.2.6   Discriminator Jitter

Another source of error on the offset compensation process could arise from the sampling of the noise on the leading edge discriminator input node. This stochastic contribution cannot be eliminated and therefore ads up to the CSA jitter. The
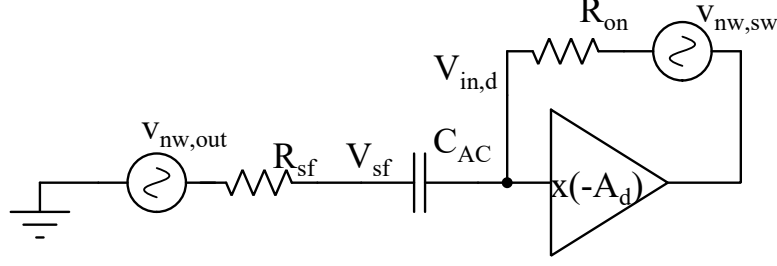
Figure 3.5: Small signal equivalent circuit used to compute the discriminator jitter.

leading edge discriminator contribution to the noise found in $V_{in,d}$ can be computed with the small signal model presented in 3.5. R_{sf} is the output resistance of the source follower, and it is computed using (3.35).
$v_{nw,sw}$ is the series voltage noise produced by R_{on} and has been computed using (3.4). By solving the nodal equation, it is possible to find the voltage noise in $V_{in,d}$ as:

$$V_{in,d} = \frac{(C_{AC}R_{sf}s + 1) \cdot v_{nw,sw}}{[(A_d + 1) R_{sf} + R_{on}] C_{AC}s + A_d + 1} \tag{3.37}$$

This equation can be simplified considering the relative lower resistance R_{on} of the switch and the high gain of CD.

$$A_d \gg 1 \ , \ R_{sf} \gg R_{on} \tag{3.38}$$

which leads to:

$$V_{in,d}^* = \frac{v_{nw,sw}}{A_d} \tag{3.39}$$

The noise generated by M_{o2} is reduced by the discriminator core gain making it negligible when compared to $v_{nw,out}$.

Finally, the OC procedure samples the input noise v_{nw} which affects V_{bl} at each iteration. Therefore the total system jitter for a random signal $\sigma_{t,tot}$ is composed of these two identical independent components. One jitter component affects the signal, the other one the baseline. The jitter $\sigma_{t,tot}$ can be computed as:

$$\sigma_{t,tot} = \sqrt{2}\sigma_{tn} \tag{3.40}$$

using (3.30), $\sigma_{t,tot}$ results to be 90 ps.

3.3 Implementation

This section discusses the issues and design decisions regarding the actual implementation of the front end. The layout of the implemented AFE is presented

in figure 3.8 whereas its transistors sizing is reported in table 3.1. The discussion will start from some general aspects of the implementation; it will then move to the realization of the power, biasing and signals scheme and, finally, it will discuss some topics regarding specific sub-circuits.

3.3.1 General Considerations

Starting from the area utilization, the core channel size is $34.0\,\mu m \times 10.9\,\mu m$. If the charge injection circuit is included, the overall occupied area is $467.3\,\mu m^2$ which will account for just 16% of the total pixel area. It must be noted that a large portion of the charge injection circuit is occupied by the 3 capacitors C_S which emulate the sensor capacitance. In a final version of the circuit, these capacitors will not be present, making the area occupation around 13% of the available one. The rest of the pixel can be used to implement the TDC and the local digital circuitry.

The 28 nm CMOS technology takes advantage of interference-based lithography masks. This technique improves the quality of the patterning of features with a size comparable to the wave length of the employed light. In order to make this technique feasible, the design must feature a certain regularity. In particular all transistors gate metals must have the same orientation across the chip. The gates must also have the same aspect ratios W/L in subsection of the design. In order to transit from a region with a certain W/L to another with a different one, some dummy gates must be placed in between.

For the purpose of increasing the area efficiency and improving the overall transistor yield, only two different gate metal aspect ratios were selected and repeated in clusters across the layout. Each flavor was designed to be combined in a multi-finger configuration in order to build different W/L transistors. The used configurations comprises both the standard in-parallel connection in order to obtain the sum of the fingers W and an in-series connection to sum the L of the fingers. The two aspect ratios were selected in order to better form wide and long transistors. In addition, by combining the same metal finger with different diffusion widths, it was possible to produce long transistors in two sub-variants. Hence, the whole design is built from combinations of transistors with a W/L of 1/0.1, 0.1/0.3 and 0.5/0.3 expressed in µm/µm.

Overall the minimum L adopted was $0.1\,\mu m$, three time larger than the minimum feature allowed by the technology. This decision was taken for several reasons related to the transistor total area. First of all, a large area enables an easier connection to the metal interconnection layers, enabling a wider spread mentalization which will decrease the parasitic capacitance. The flicker noise component is also inversely proportional to transistor area. In order to maximize timing performance this noise component has been made negligible compared to the thermal one. Likewise, process and mismatch variations also decrease with the total area [51]. In

71

particular bias transistors and the Krummenacher filter have been oversized since circuit operation point is defined largely by them.

In order to improve substrate noise insulation and enhance radiation tolerance, each transistor was implemented inside its dedicate well polarized with its own guard ring. Therefore the AFE has been implemented with the triple-N-well technique with a dedicated deep-N-well per channel. This structure also enables to polarize the deep-N-well with a different ground network compared to the substrate one, reducing cross talk with nearby digital circuits.

Capacitors were implemented using both Metal-Oxide-Metal (MOM) capacitors and MOS capacitors. MOM capacitors where selected for sensitive part of the circuit such as C_f, C_{AC} and C_S. Their design values are $C_f = 4\,\text{fF}$, $C_{AC} = 100\,\text{fF}$ and C_S is selected to vary from $25\,\text{fF}$ to $175\,\text{fF}$. C_{AC} was split in two symmetrical capacitors in order to improve the leading edge discriminator symmetry, thus reducing the offset due to the reduction of the mismatch between the two differential branches. The usage of MOM capacitors avoids the voltage non-linearity of MOS capacitors and minimizes the coupling to the substrate. C_f and C_{AC} were implemented using only middle metal layers in order to further reduce coupling with the substrate and the top metal nets. C_k and C_o were implemented with PMOS and NMOS transistors respectively due to their higher capacitance density. Since they act as large damping capacitors, they are less affected by the aforementioned drawbacks.

3.3.2 Power, Biasing and Signals

The AFE was designed in order to work over a power-consumption range of $5\,\mu\text{W}$ to $10\,\mu\text{W}$ per channel. This range was selected based on the power consumption per unit area allowed in the applications [108][107][106] and the performance of state-of-the art cooling systems. [23]. The AFE will thus use from one fourth to one half of the total power budget. The bias current of the various sub-circuits can be regulated independently. In particular, the most relevant branches in terms of power consumption are the two core amplifier branches and the two discriminator core stages. The Krummenacher filter power consumption is negligible when compared to the total one. Other blocks do not consume static power and contribute with negligible dynamic one.

The Krummenacher filter bias current I_k can be modulated from $25\,\text{nA}$ to $100\,\text{nA}$ in order to set the discharge current.

All the bias currents as well as all the cascode voltages (labeled as "cas" in figure 3.7) were provided using a common bias cell. The schematic of this cell is omitted since it is not relevant for circuit description.

Each bias current is directly provided by an external reference throughout its dedicated pad. These currents are then copied using current mirrors to a replica bias cell of the actual AFE. All the cascode voltages are then fed to this replica and
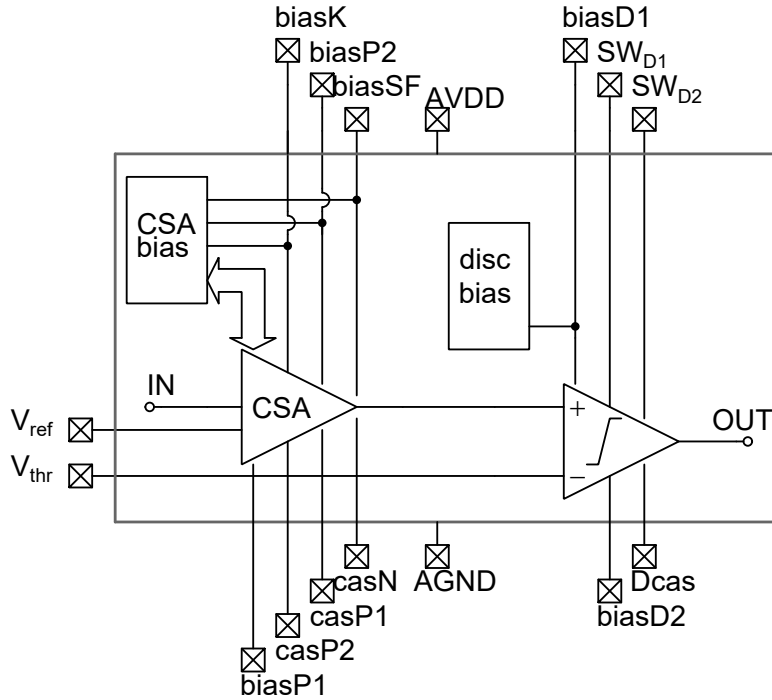
Figure 3.6: Schematic block representation of the Timespot0 AFE biasing scheme. The rectangular frame represents the circuits inside the chip, whereas the crossed boxes represent the analog bond pads.

the actual channels via dedicated pads. The replica bias voltages are again copied to the AFE channels. In order to reduce the process and mismatch variations [83], the bias cell has been replicated two times. This bias scheme is presented in figure 3.6 whereas the layout of the 8 channels plus the bias cell can be found in figure 3.9. Since no current flows from the replica to the channel, this solution can be used to bias hundreds of channels. The operating analog and digital signal SW_1, SW_2, TP, VA, V_{thr} and V_{refK} were supplied via dedicated pads.

3.3.3 Krummenacher Filter Stability

The use of the Krummenacher filter in the CSA feedback introduces four poles which are crucial for the circuit stability. An optimal implementation of the scheme must assure a good separation of these poles. The values of these poles are listed below from the lowest to the highest frequency:

73

$$\omega_{k1} = \frac{g_{m,k3}}{C_k} \tag{3.41}$$

$$\omega_{k2} = \frac{g_{m,k4}}{\frac{C_T}{A_{ol}} + C_f} \simeq \frac{g_{m,k4}}{C_f} \tag{3.42}$$

$$\omega_{k3} = \frac{g_{m,k4}}{C_{k4,gnd}} \tag{3.43}$$

$$\omega_{k4} = \frac{C_f}{C_T} A_{ol}\omega_a \simeq \frac{C_f G_{m,in}}{C_T C_L} \tag{3.44}$$

The term $C_{k4,gnd}$ in (3.43) represents the capacitance between the source of M_{k4} and ground. In (3.44) the core amplifier angular frequency ω_a has been quantified as the reciprocal of the time constant $R_{oa}C_L$. This is the dominant of the two poles of the CSA core amplifier. The separation between ω_{k1} and ω_{k2} can be obtained by increasing C_k. The downside of this solution is the increase in area occupation. ω_{k4} and ω_{k3} can be separated by changing the ratio between the core amplifier current and I_k. Doing so by reducing I_k will not affect the timing performance even though it will increase the signal fall time. Lastly, the separation between ω_{k2} and ω_{k3} is more delicate and relies on transistors sizing. Once C_f size is established M_{k1}, M_{k2} and M_{k4} need to be shrunk in order to reduce their capacitance. However this approach will produce a higher mismatch in the pair generating a trade-off between circuit stability and baseline fluctuation.

3.3.4 Offset Compensation Circuit

The main issue of the OC is the time duration T_{hold} in which the desired voltage is stored inside C_{AC}. Any small conductance is able to discharge C_{AC} and therefore compromise $V_{in,d}$ voltage. The voltage on this node will migrate towards its equilibrium point V_0. From the point of view of timing, a certain difference in V_{bl} will produce an error $t_{err,oc}$ dependent on the elapsed time from the previous OC:

$$t_{err,OC}(t) = \frac{V_{bl}(0) - V_{bl}(t)}{V'_{out}} \tag{3.45}$$

where $V_{bl}(0)$ is the value of V_{bl} immediately after the OC. Since the signal arrival time is unknown, this error is random and therefore it cannot be fully eliminated. The OC needs to be repeated periodically in order to minimize this effect. However a more frequent OC will make the channel blind for a longer time.

A first solution to mitigate the problem can be to increase C_{AC} capacitance since this will produce a smaller voltage drop for the same discharge current. There are however different drawbacks of this approach. First, the increase in the offset compensation setting time T_{comp} due to the higher time constant τ_{comp}.

$$\tau_{comp} \propto R_{on} \cdot C_{AC} \tag{3.46}$$

Second, the area occupation of C_{AC} will also increase with its capacitance. This in turn will also increase the parasitic capacitive coupling to the substrate $C_{par,gnd}[C_{AC}]$. This capacitance will low-pass filter the input signal reducing its slew rate and therefore the timing performance. Its cutoff frequency f_{lp} will depend on R_{sf}:

$$f_{lp} = \frac{1}{2\pi R_{sf} C_{par,gnd}[C_{AC}]} \tag{3.47}$$

C_{AC} sizing was thus selected in order to have the cutoff frequency above the signal band.

Another approach to enlarge T_{hold} is to reduce the discharging currents. The main discharge current comes from the off-current of M_{o2}. This can be limited by increasing the transistor off-resistance R_{off} by an appropriate sizing. Increasing R_{off} will inevitably increase R_{on} with the side effects of affecting τ_{comp} according to (3.46) and increasing V_{err} as shown in (3.36). However M_{o2} can be sized to make its off-current negligible before reaching these trade-offs. At this point the gate leakage currents i_{leak} of the transistors connected to C_{AC} will start to become dominant. i_{leak} is directly proportional to the gate area, for this reason M_{d2} gives the major contribution. In order to limit this component, M_{d2} and M_{d4} were implemented with thick-oxide transistors. The last leakage current is the one of M_{o3}. This component was reduced switching off M_{o4} during the signal discrimination, which will insert a high resistance in series to M_{o4} gate. Generally it is better to set V_{bl}^* close to V_0 because the small voltage difference across the discharging resistors will limit such currents.

Finally M_{o3} was introduced in order to compensate the current injected by M_{o2} throughout its C_{gs}. During OC M_{o3} is counter driven in respect to M_{o2}. M_{o3} is sized to have half of the area of M_{o2}, so that they will inject the same charge into C_{AC} but with opposite sign. The quality of this compensation will depend on the mismatch between these two transistors.
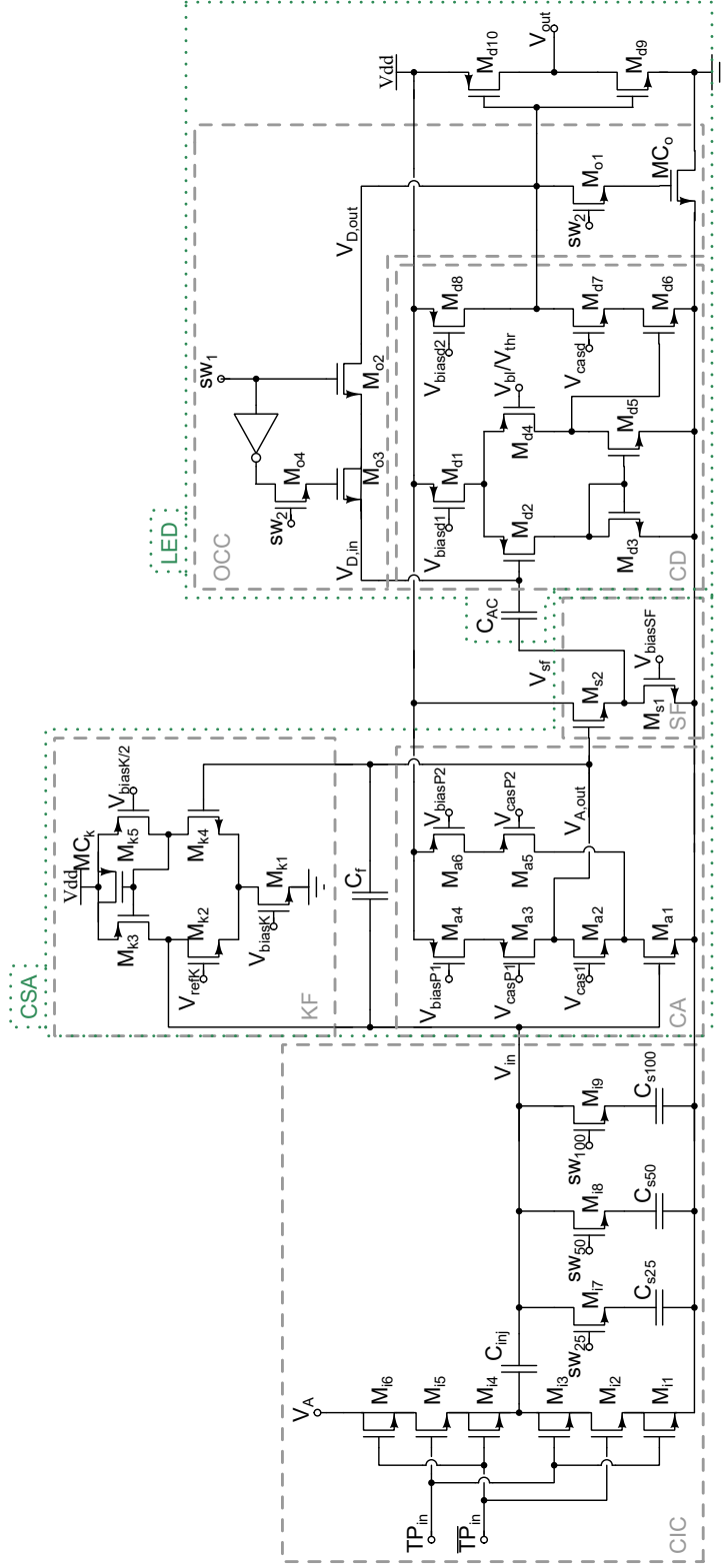
Figure 3.7: Transistor level schematic of the Analog Front End channel. The bias circuit is omitted for sake of clarity. The relevant blocks described in the text are labelled as: Charge Injection Circuit (CIC), Charge Sensitive Amplifier (CSA), Krummenacher Filter (KF), Core amplifier (CA), Source Follower (SF), Leading Edge Discriminator (LED), Core Discriminator (CD) and Offset Compensation Circuit (OCC). M_{d2} and M_{d4} are thick oxide transistors.

Table 3.1: Transistor sizing of the Timespot0 analog front-end. The transistor labels are the same as those used in figure 3.7.

Charge Injection Circuit	Core Amplifier	Krummenacher Filter	Source Follower	Discriminator Core	Offset Correction Circuit
width/length $[\mu m/\mu m]$					
M_{i1} 0.15/0.10	M_{a1} 8.00/0.10	M_{k1} 0.50/4.80	M_{s1} 0.20/3.00	M_{d1} 0.50/3.00	M_{o1} 0.10/0.30
M_{i2} 0.30/0.10	M_{a2} 1.00/0.10	M_{k2} 6.00/0.10	M_{s2} 6.00/0.10	M_{d2} 2.00/0.15	M_{o2} 0.10/0.30
M_{i3} 0.15/0.10	M_{a3} 1.00/0.10	M_{k3} 0.20/7.20		M_{d3} 0.20/1.20	M_{o3} 0.10/0.15
M_{i4} 0.15/0.10	M_{a4} 0.20/2.40	M_{k4} 6.00/0.10		M_{d4} 2.00/0.15	M_{o4} 0.10/0.15
M_{i5} 0.30/0.10	M_{a5} 1.00/0.10	M_{k5} 0.20/7.20		M_{d5} 0.20/1.20	MC_o 15.96/0.30
M_{i6} 0.15/0.10	M_{a6} 0.20/1.80	MC_k 24.32/0.30		M_{d6} 4.00/0.10	
M_{i7} 0.30/0.10				M_{d7} 1.00/0.10	
M_{i8} 0.30/0.10				M_{d8} 0.20/1.80	
M_{i9} 0.30/0.10				M_{d9} 0.50/0.10	
				M_{d10} 0.50/0.10	
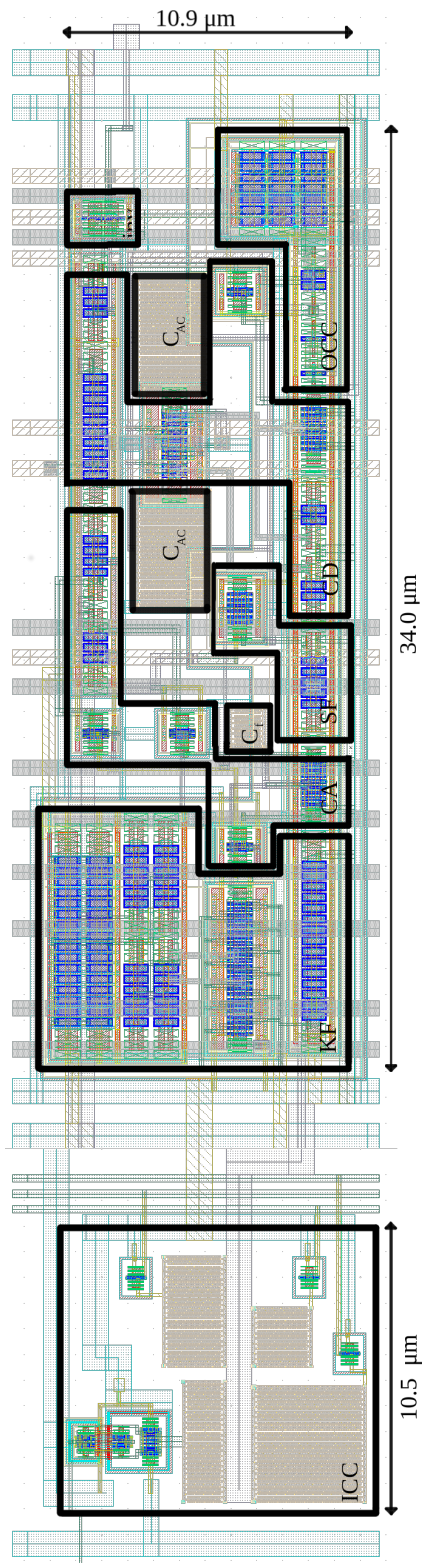
Figure 3.8: Layout of the Analog Front-End presented in figure 3.7. Relevant blocks and MOM capacitors are highlighted. The blocks are labelled as: Charge Injection Circuit (CIC), Krummenacher Filter (KF), Core amplifier (CA), Source Follower (SF), Core Discriminator (CD) and Offset Correction Circuit (OCC). The core channel area occupation is $379.7\,\mu m^2$, the overall area is $467.3\,\mu m^2$.

single channel **core cell** **charge injection** **bias cell**

Figure 3.9: Layout of the Timespot0 AFE channels and bias-cell. Each chip features 8 channels and one bias cell.

Chapter 4

Timespot0: Analog Front-End Characterization

This chapter presents the characterization process and results on the Timespot0 analog front-end, which is described in chapter 3. Firstly, the measurement setup and methodology will be presented in section 4.1, this information is essential to understand the subseque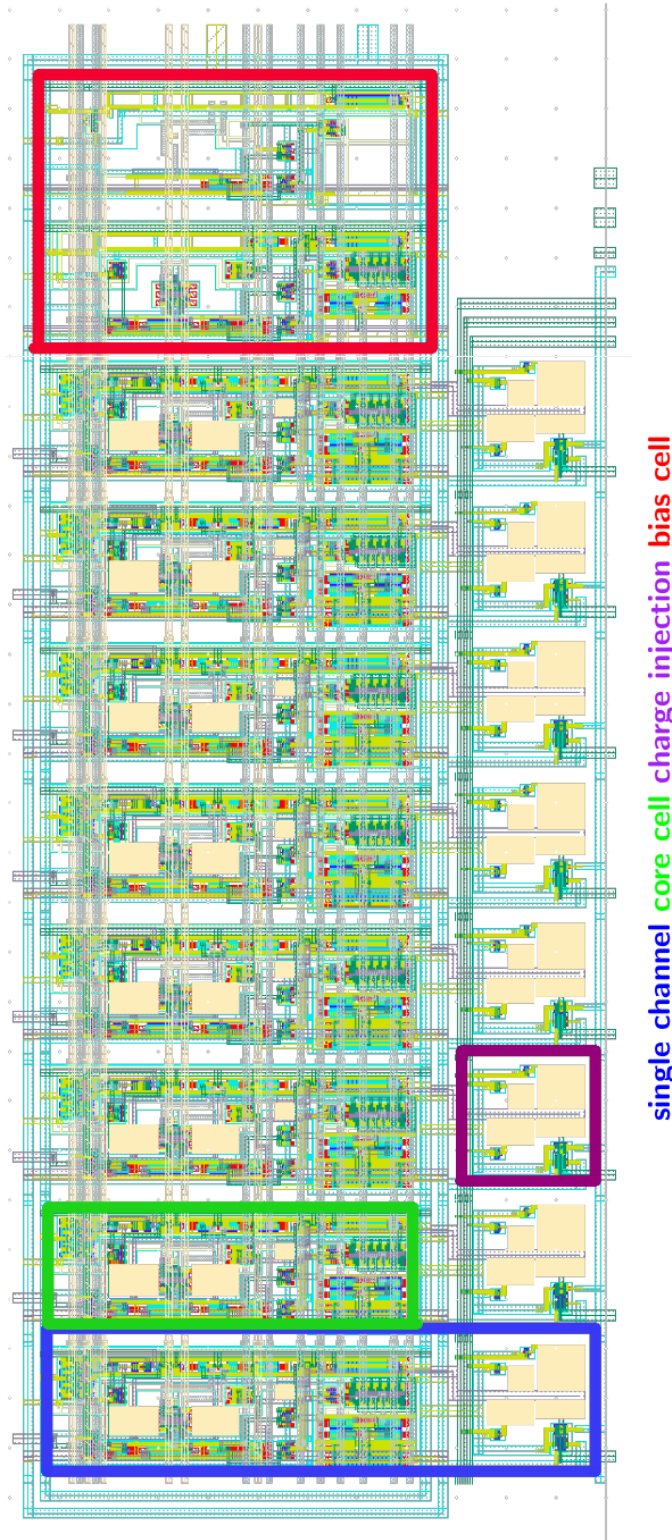nt measurements. Section 4.2 contains all the plots extracted from the experimental data with related discussions. For the sake of convenience, a full schematic of the setup and a flow chart of the Data Acquisition (DAQ) firmware are placed at the end of the chapter. This chapter presents its content assuming that the reader is familiar with the measurement concepts presented in 1.2 and with the AFE architecture presented in 3.

4.1 Setup and Method

This section describes the experimental setup and presents the AFE characterization methodology. The setup description includes the on-chip electronics, the electronics on the tspot0 board and all the measurement instruments. The measurement method was adopted to reliably characterize the circuit from a timing point of view while differentiating the time fluctuation components coming from the Device Under Test (DUT) from the ones of the setup.

4.1.1 Setup

In order to characterize the AFE, the setup must be able to pulse each individual channel and measure the discriminator digital output while performing the offset compensation. A schematic representation of the experimental setup is reported in figure 4.13 at the end of the chapter, whereas a photograph of the setup is shown in figure 4.1. The setup must also provide all the required analog levels. The setup
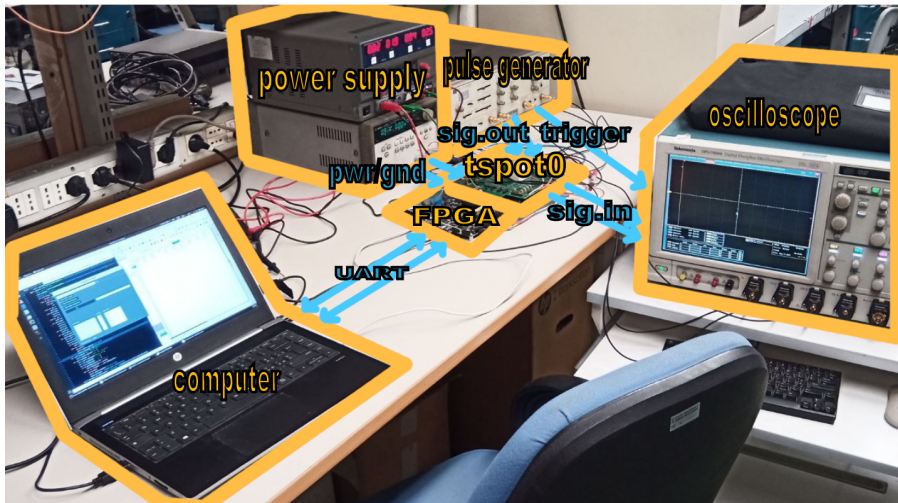
Figure 4.1: Photograph of the setup used for the Timespot0 AFE characterization.

will be described starting from the on-chip electronics.

The Timespot0 chip contains 8 AFE channels. Each channel can be triggered digitally using its dedicated charge injection circuit, whereas the only measurable signal is the digital output of the discriminator. These two signals are routed through the on-chip LVDS receiver and driver respectively, which are directly connected to the respective pair of pads. Having a differential current signal ensures that a digital signal can be transmitted without been degraded by the interconnections impedance. However, only one pair of LVDS is implemented in the ASIC, making a channel selection mechanism mandatory. A combination of one de-multiplexer and a multiplexer is used to select independently the channel to be pulsed and the one to be read. In the same way the channels can be bypassed forming a loop-back of the drivers. In order to program the desired configuration, dedicated registers can be written via a common three wire interface. In addition, the setup must also provide the analog references and the signals used for the offset compensation. All the analog levels are provided externally through their dedicated pads. The two digital signals, SW_1 and SW_2, used in the offset correction circuit, are routed through their dedicated CMOS pads.

The analog references are provided using the dedicated circuit implemented in the tspot0 PCB. The PCB is shown in figure 4.2 in which the components and connections related to the AFE are highlighted. Every analog voltage reference (V_A, V_{refK} and V_{thr}) and the cascode voltages are generated by four LTC2604 DACs [65] mounted on the PCB. The reference currents can be regulated using resistive dividers directly connected to the pads. Each divider is composed of one potentiometer for current regulation and a known resistor for current measurement. An LVDS buffer is also mounted on the PCB in order to strengthen the output signal directed to a SubMiniature version A (SMA) connector.
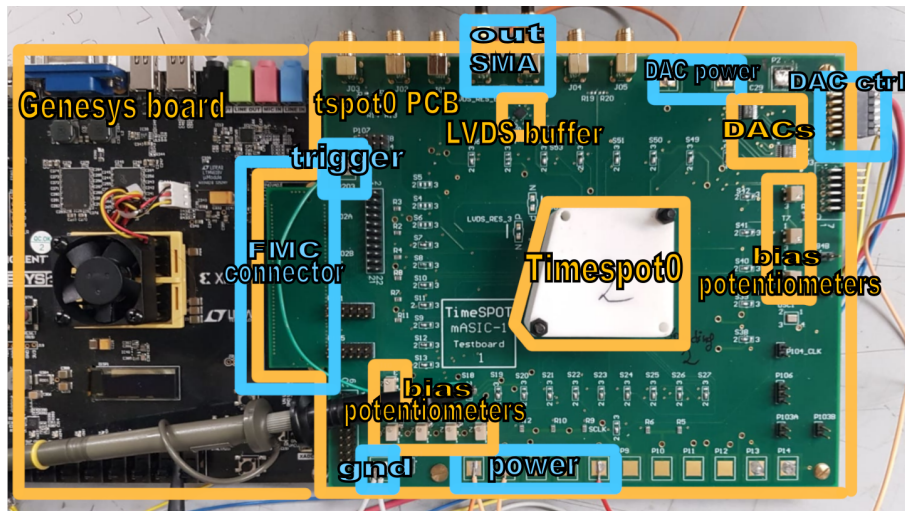
Figure 4.2: Photograph of the tspot0-PCB connected to the Genesys Board via FMC. The Timespot0 chip is under the white cover

An FPGA is used to control the digital signals, write the registers and program the DACs. The adopted FPGA board is a Genesys2 [36] mounting a Kintex7 [115] FPGA. The decision to use this type of digital controller rather than a simpler one comes from the need of executing the measurements routine with the tight timings imposed by the offset correction procedure. The firmware must implement the procedure described in 3.2.5 which involves controlling the OC signals, the threshold DACs and the TP signal. The procedure can be configured by setting various delays between its different phases, as well as the phase between the OC itself and the channel pulsing. The FPGA also writes the configuration registers and programs all the DACs in the PCB. The FPGA is not directly connected to the LVDS receiver, but it triggers the pulse generator which produces the LVDS signal used as the TP. A flow chart of the state machine implemented in this firmware is shown in figure 4.14 at the end of the chapter.

The pulse generator is adopted since it produces a stable and reliable succession of pulses. The output of the pulse generator is split in two different paths, one is connected to the PCB and the other one to the oscilloscope. The two paths use the same coaxial cables in order to have the same propagation delay. The output of the LVDS transmitter is also connected to the same oscilloscope. In this way, each time measurement consists in the difference between the input signal phase and the output one. In this way, external stochastic contribution to the absolute phase of the input signal are eliminated. FPGA and pulse generator jitters can be the sources of these variations.

Finally, a software was developed in Pyhton3 [93] in order to control the FPGA via USB using an Universal Asynchronous Receiver-Transmitter (UART) interface.
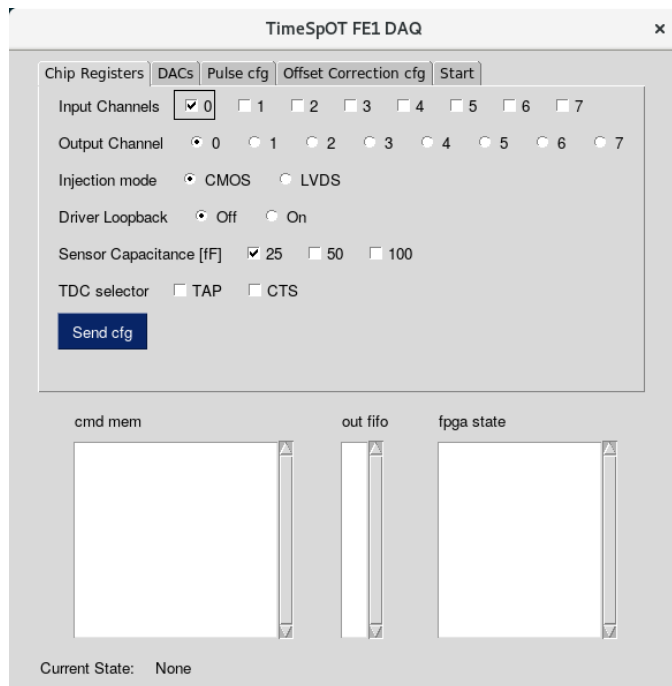
Figure 4.3: Screen of the GUI developed for the Timespot0 DAQ.

The software uses the pySerial library [92] to implement the UART interface and the pyThread multi-threading library [94] to enable continuous write and read procedures without dead times. A Graphical User Interface (GUI) was also developed in order to consistently monitor the DUT state. A screen of this GUI is presented in 4.3.

4.1.2 Method

The AFE characterization process must begin by identifying the operative conditions. First, both the input signal and the AFE state can be changed to test different aspects of the circuit. As described in 3.1 both the input signal charge Q_{in} and the input capacitance C_S can be set via V_A and their respective capacitors. In terms of the main circuit, the power can be set for the different stages as described in 3.3.2. Moreover, the leading edge discriminator can be configured with different V_{bl} and V_{thr} when using the OC, as well as operating without OC. Over the course of the section, certain conditions will be labeled as nominal. This is the condition that better represents the target system, both in terms of sensor characteristics and power consumption. The values for this condition are: $C_S =150$ fF, $Q_{in} =2$ fC, total power $P_{tot} =6.51$ µW and $V_{thr} =20$ mV with OC enabled.

For a given condition, the circuit response can be extracted from repeated measurements on many signals. Figure 4.4 shows an example of a procedure with
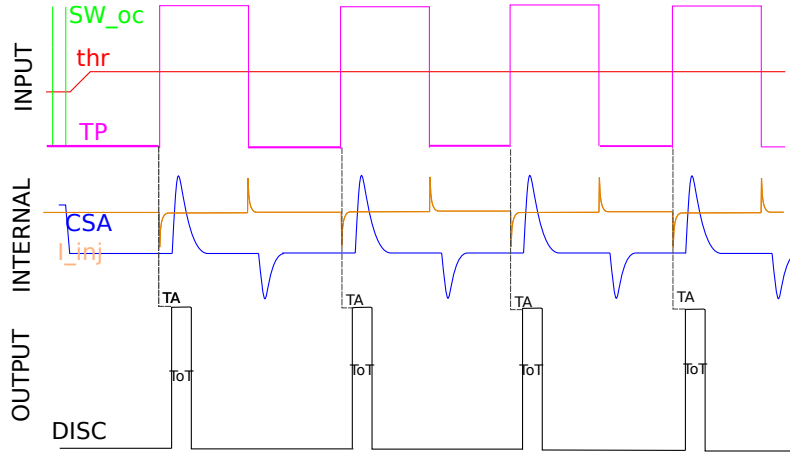
Figure 4.4: Pulsing procedure used to test the Timespot0 AFE with offset compensation (OC) enabled. The waveform graph differentiates the signals observable with the setup from the ones internal of the chip.

repeated pulses after an OC procedure. Since this process involves repeatedly pulsing the AFE, the measurement will be susceptible to different sources of stochastic variations. The variations could be intrinsic to the DUT or due to the setup. However, the average of these measures will be independent on both type of variations, depicting the actual mean behavior of the circuit. The standard deviation of the set of measurements σ_{mes}, on the other hand, will quantify the entity of the variations. For the purpose of characterizing the AFE, only intrinsic variations σ_{AFE} must be measured since they provide its final time resolution limit. By leveraging the LVDS drivers loop-back, most of the setup related variations σ_{setup} can be quantified using the standard deviation of a repeated set and then square removed from the DUT measurements. The only non-removable component of external variation is the one introduced by the charge injection circuit σ_{CIC}. Overall, the different sources of variation can be decomposed as:

$$\sigma_{mes}^2 = \sigma_{AFE}^2 + \sigma_{setup}^2 + \sigma_{CIC}^2 \tag{4.1}$$

Only the combined contribution of the AFE and the charge injection circuit can be measured: this measurement overestimates the AFE intrinsic time-variation. Sources of systematic variation cannot be pointed out by direct measurement, therefore they must be assumed and verified a posteriori.

Experimental conditions are also replicated in simulation in order to be used as a reference. These simulations consist in post-layout Montecarlo simulations, in which parasitic elements are extracted from the circuit layout and inserted into the circuit netlist. In all the plots presented in this chapter (figures 4.5-4.12) simulation results are also presented in the form of a trend line representing the average behavior. The three different filled areas represent the 1σ, 3σ and 5σ variation

bounds, respectively.

4.2 Measurements

In this section, the experimental measurements are presented from the point of view of the CSA and the leading edge discriminator. The focus of the test was to measure the overall system jitter and the proper operation of the OC circuit.

4.2.1 CSA Characterization

The first characterization of the CSA is carried out by acquiring the s-curves of the channels in different conditions. This analysis quantifies the channel noise, its baseline spread and its charge to voltage gain. The measurement is carried out by pulsing the channel 200 times and counting the hit ratio for different threshold values. All the other parameters remain unchanged. At the start of the acquisition, the leading edge discriminator is under threshold. By rising the threshold, the hit ratio will increase. The data are therefore fitted using a reversed Gaussian CDF:

$$f_S(x, \mu, \sigma) = 1 - \frac{1}{2} \left[1 + erf \left(\frac{x - \mu}{\sigma \sqrt{2}} \right) \right] \tag{4.2}$$

where $erf(x, \mu, \sigma)$ is the error function. This process was repeated across the 8 channels in nominal condition. A first set of s-curves were acquired in the absence of signal, with the inflection point corresponding to the channel DC baseline. The s-curves have also been acquired with the 2 fC signal; this time the inflection point corresponds to the signal peak voltage. It is possible to evaluate the signal amplitude by subtracting the μ parameters obtained in the two conditions. In order to check for consistency, these operations were repeated while using the OC procedure and without it. Figure 4.5 shows these measurements, their fitted curves and the values expected in the simulation. In the case of s-curves derived from the channel without signal and without OC, (4.2) proved to be inadequate as fit function. By assuming that the observed behavior is caused by the convolution of two normal distributed fluctuation terms with different mean values μ, the curve can be fitted using the product of two Gaussian Cumulative Density Function (CDF):

$$f_S, prod(x, \mu_1, \sigma_1, \mu_2, \sigma_2) = f_S(x, \mu_1, \sigma_1) \cdot f_S(x, \mu_2, \sigma_2) \tag{4.3}$$

One of the two contributions is assumed to be generated by the set-up; the one which best matches the other evaluations is considered to be originated by the AFE.

From the point of view of the signal amplitude, a noticeable discrepancy from the expected value can be observed. The origin of this discrepancy can be related
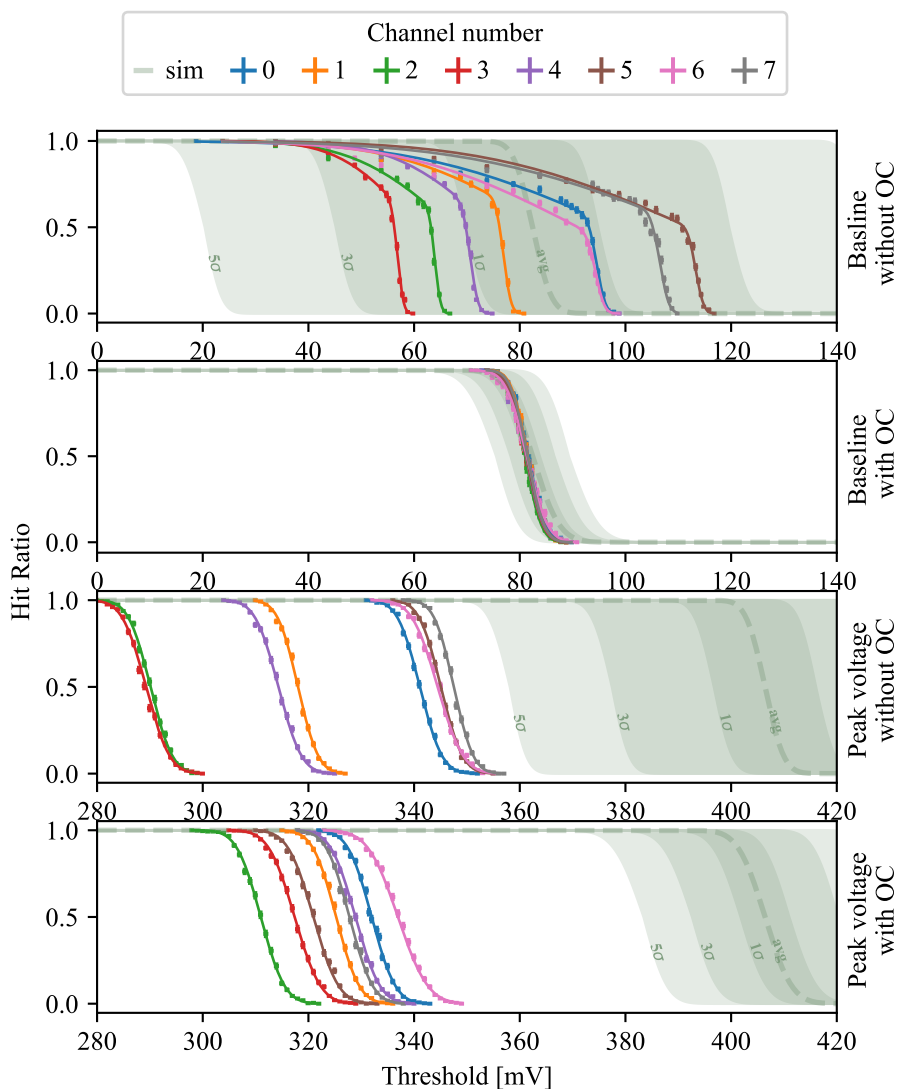
Figure 4.5: S-Curves extracted from 8 different AFE channels compared with the simulated values. The curves were measured in four different combinations: with and without an input signal (baseline and peak voltage) of $2\,\text{fC}$ and with and without offset compensation (OC). Every point is obtained from 200 repeated measures in nominal condition. The channel mismatch in terms of peak position is imputable to variations on the charge to voltage gain. Each channel gain is consistent regardless of the usage of the OC.

to two effects: a lower than expected CSA gain or a lower than expected charge injection. The first hypothesis may not be attributed to process variations, since the measured values are significantly outside the 5σ expectancy range. This fact
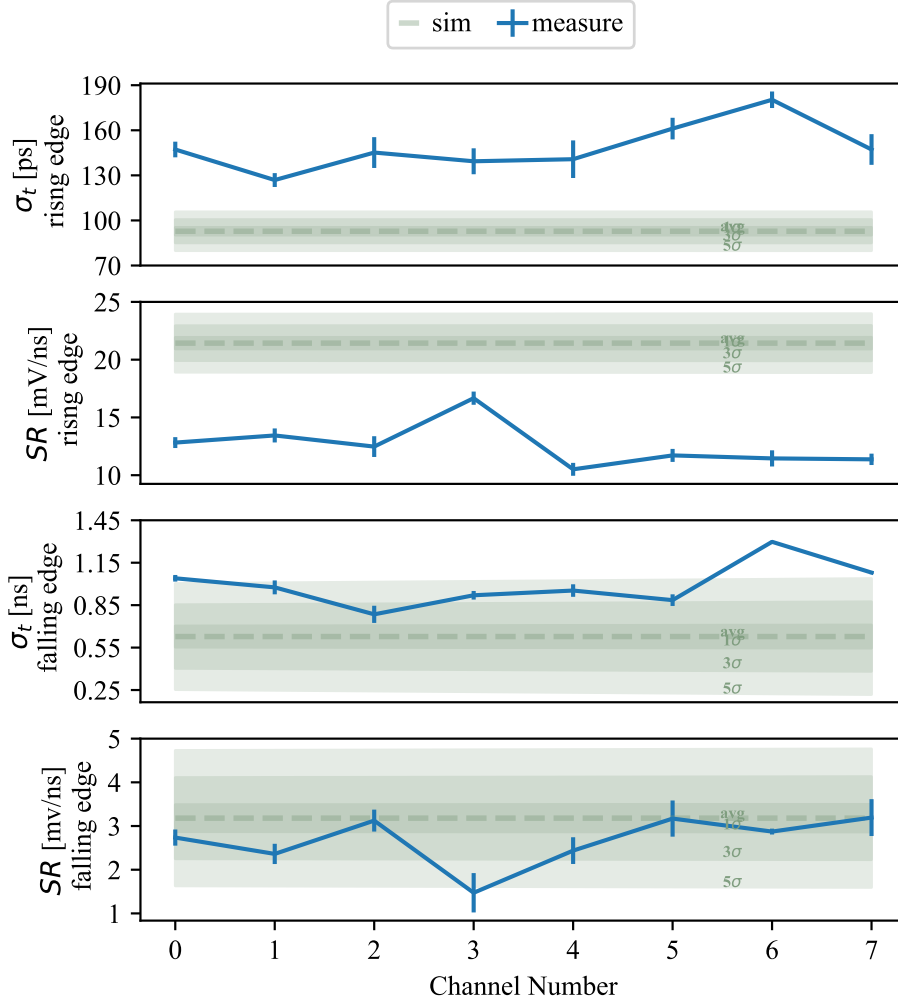
Figure 4.6: Jitter (σ_t) and Slew-Rate (SR) computed on both the rising edge and falling edge of the signals across 8 channels. The values were extracted from 4 independent sets of 200 measures taken with V_{thr} values separated by $1\,\mathrm{mV}$ around the nominal condition. σ is computed as the average of the 4 sets, while SR is the slope of the fit line.

can also be observed in the signal threshold scan presented in figure 4.7. In order to validate the hypothesis of an input charge deficit, some other effects on the signal must be observed. First, the slew-rate will directly relate to input charge as shown in (3.29). As a consequence, the jitter will follow the same behavior, as shown in (3.12). Moreover, the ToT will also be reduced by the same factor. If the constant discharge current is the one expected, the ToT will indicate the input charge. Using the s-curves of figure 4.5 the deficit factor results to be around 0.73. The same factor of 0.73 can be observed in the threshold scan of figure 4.7. A
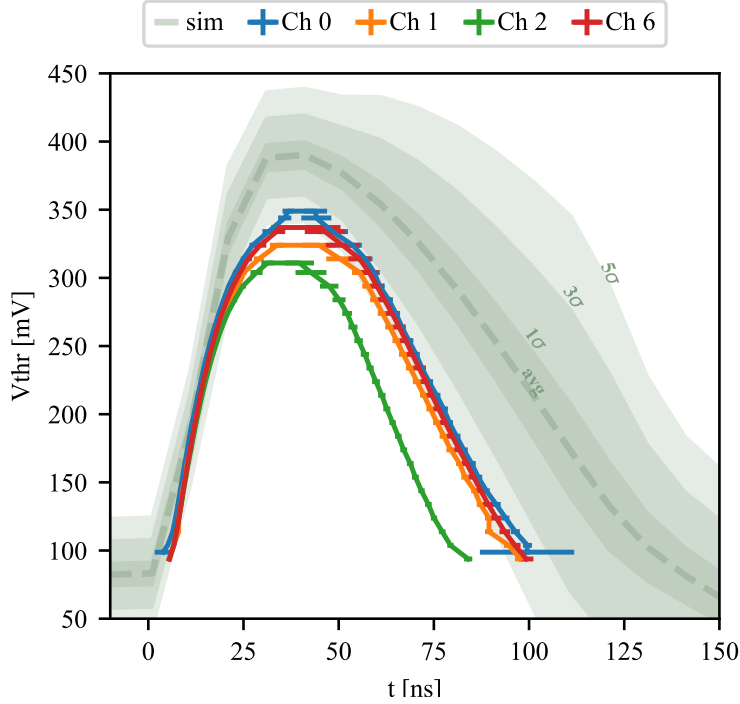
Figure 4.7: Threshold scan reconstruction of signals in nominal condition and expected simulated signal. Every point is obtained from 200 repeated measures. The four shown channels are the ones with highest and lowest gain (Ch 0 and Ch 2) and the ones with better and worst timing performance (Ch 1 and Ch 6).

value of 0.71 can also be found when comparing the ToT of the same measurement. In figure 4.6 the result of the direct measurements of the slew-rate and jitter are presented. Also, in this case, the values related to the rising edges show similar deficits of 0.68 for the slew-rate and 0.67 for the jitter. However, the falling edge slew-rate measurements of figure 4.6 are in line with what expected in simulation. This suggests that the discharge current is the desired one, and therefore the ToT can be trusted when used to quantify the input charge. The same deficit can also be observed in figure 4.9.

The fit parameters σ from figure 4.5 will quantify the CSA noise. This noise can also be computed from the ratio of the measurement of the jitter and slew-rate presented in figure 4.6 according to (3.12). The combined results from these analyses can be found in figure (4.8). First, it can be observed that, overall, the noise is lower than expected. This can be imputed to an underestimation of the noise due to a filtering effect by the setup.

Lastly, figure 4.9 presents the direct measurement of the jitter as a function of the input capacitance. This curve represents a key factor during sensor engineering in order to predict the front-end jitter. The measured trend is in line with what
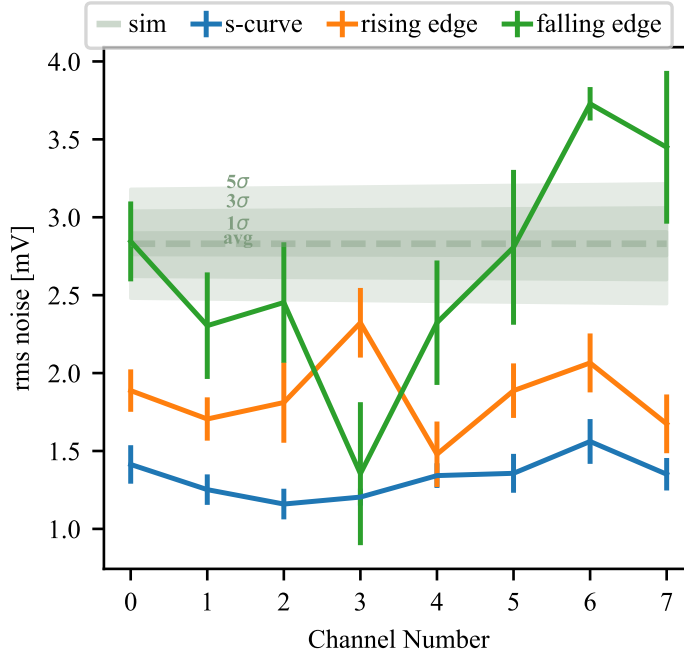
Figure 4.8: Evaluated noise across 8 channels computed with two independent methods. Data labeled as s-curve has been extracted from the average of the σ from (4.2) parameter of the fits presented in figure 4.5. Rising edge and falling edge were computed as SR/σ_t from the data presented in figure 4.6.

predicted with (3.30).

4.2.2 Discriminator Characterization

The leading edge discriminator characterization focuses on the verifying the correct operation of the OC. Figure 4.5 shows that the OC is working properly. The baseline position is in line with what expected from simulation in terms of systematic error $t_{err,oc}$ and spread.

In figure 4.11 the measurement of the actual baseline V_{bl} as a function of the T_{comp} is presented. The offset starts to become constant above 100 ns, which defines the optimal T_{comp} value. Below 60 ns a residual instability can be observed. This oscillation is damped afterward.

Lastly, the OC coherence time τ_{hold} is evaluated. In order to perform it, the baseline position was measured after increasing time periods from the last OC. The results are presented in figure 4.12. In particular, these measurements were taken around the channel asymptotic equilibrium point. This point is evaluated to be around 73.5 mV. The two baselines V_{bl}^* were set to be 20 mV above and below the
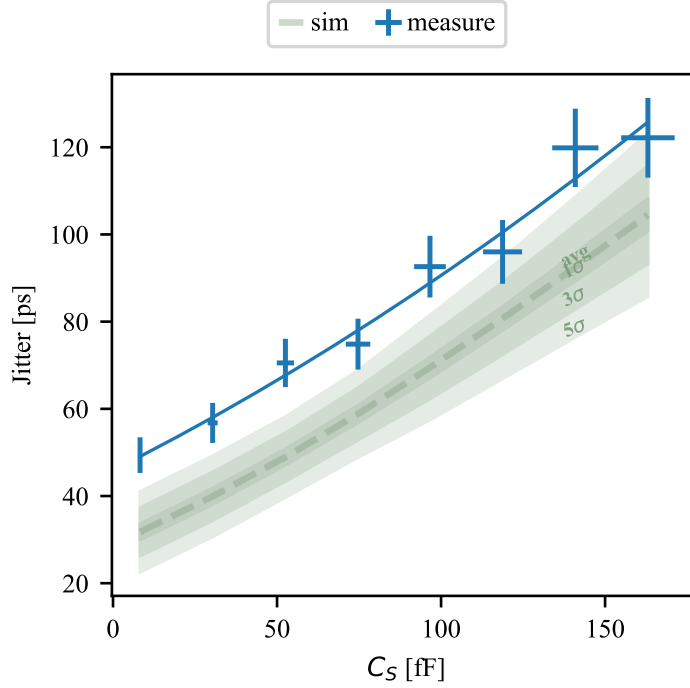
Figure 4.9: Jitter trend measured in standard condition versus sensor capacitance. The test was performed on channel 1. Capacitance values have been evaluated from post-layout parasitic extraction.

equilibrium. The crossing time is then measured. The characteristic times τ_{hold} can be quantified using an exponential fit. Their values result to be of (31 ± 3) ms when the threshold is reached from below, and (46 ± 5) ms when it is reached from above. The same values extracted from the average simulation were $46\,\text{ms}$ and $37\,\text{ms}$.

These characteristic times are critical to choose the appropriate T_{hold} which defines the OC frequency. A possible value for T_{hold} can be computed by imposing an average time error due to the OC.

$$\sigma_{t,err,OC} = \frac{t_{err,OC}}{\sqrt{12}} \tag{4.4}$$

This value can be imposed relative to the channel resolution σ_{TA} with a factor err_{rel}. From now on $\sigma_{t,err,OC}$ is set to be the $10\,\%$ of σ_{TA}. Using (3.45), $t_{err,oc}$ can be converted to a voltage error and thus connected to τ_{hold} by inverting its exponential relation. Using this method T_{hold} can be quantified as:

$$\begin{aligned} T_{hold} &= \ln(\Delta V_{bl} + 1)\tau_{hold} \\ &= \ln\left(\frac{\sqrt{12} \cdot \sigma_{TA} V'_{out}}{err_{rel}} + 1\right)\tau_{hold} \end{aligned} \tag{4.5}$$

91
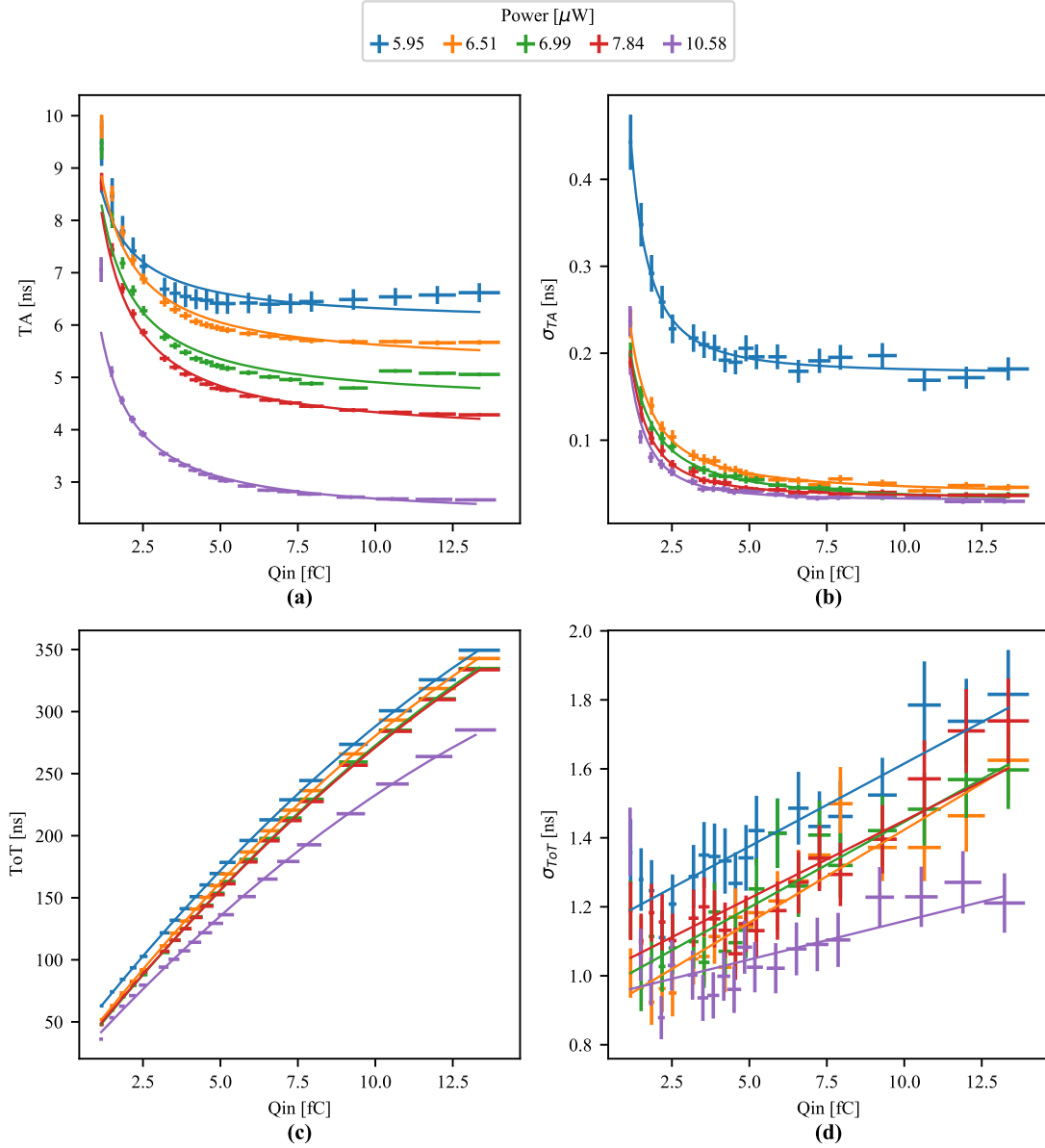
Figure 4.10: Timing characteristics of the signal for different input charge Q_{in} and power conditions. Every other parameter is set according to the nominal condition. The point with 2 fC and 6.51 μW represents the nominal condition. These curves are used to estimate the timing resolution of the system.

Using the σ_{TA} and V'_{out} of the most probable charge taken from figure 4.6, T_{hold} results to be 1.24 ms. The fraction of time χ_{oc} in which the channel is inactive due to the OC can be computed based on the compensation time T_{comp}:

Figure 4.11: Base-line value after OC for different OC times T_{OC} for Ch 0. The desired base-line V_{bl}^* is 90 mV. Baseline values have been estimated using the μ parameter of (4.2).
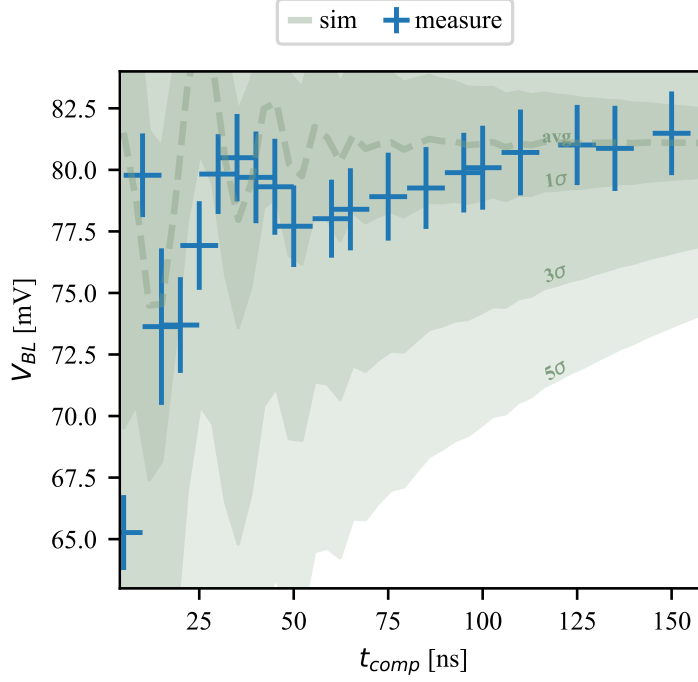
$$\chi_{oc} = \frac{T_{comp}}{T_{comp} + T_{hold}} \tag{4.6}$$

Assuming $T_{comp} =$ 100 ns based on the result showed in figure 4.11, a χ_{oc} of 0.08 ‰ is obtained. The average fraction of lost signals χ_{loss} due to this procedure can be computed as:

$$\chi_{loss} = f_{hit} \cdot \overline{ToT} \cdot \chi_{oc} \tag{4.7}$$

where f_{hit} is the expected average hit-rate and $\overline{ToT}$ is the average ToT of the signals. $\overline{ToT}$ is assumed to be 75 ns according to the measured MIP signal showed in figure 4.10.c. This will correspond to a 0.018 ‰ signal loss at 3 MHz.

It must be noted that T_{hold} is particularly susceptible to radiation damage. As explained in 3.3.4, one of the contributors to the current discharging C_{AC} is M_{d2} off current I_{off}. This current increases with the Total Ionizing Dose (TID) by the mean of two effects. The first one is the degradation of the transistor threshold voltage. The second effect is related to the formation of parallel parasitic paths along STI sides. Since the STI can trap ionized charges, increasing the TID will tend to locally invert the transistor channel. This effect is dependent on the number of STI near the transistor and therefore the number of its fingers. While it is difficult to
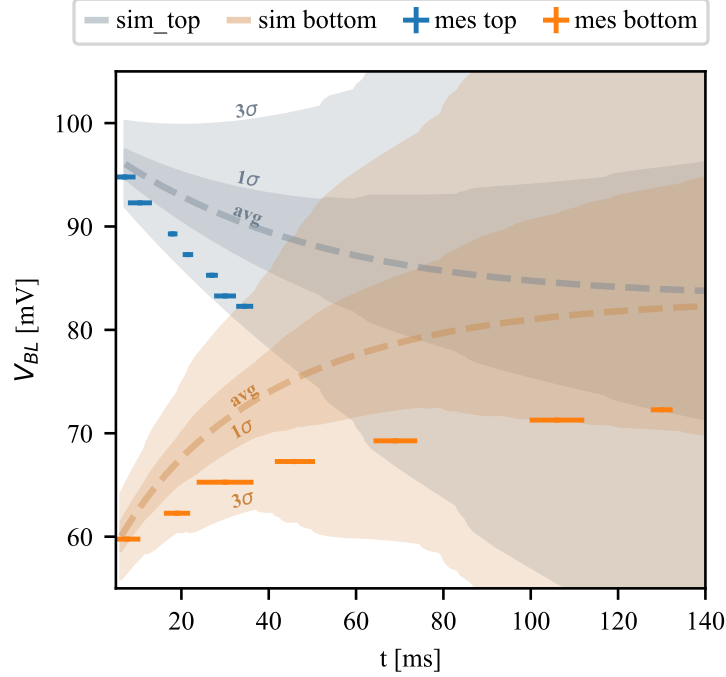
93

Figure 4.12: Measurement of the drifting of the baseline level performed using the discriminator to detect threshold crossing in absence of signal. The measure was performed in two conditions: in top conditions the threshold was set at $103.5\,\mathrm{mV}$ whereas in bottom it was set at $63.5\,\mathrm{mV}$

accurately evaluate these contributions, considering the device sizing and number of fingers, it is safe to assume that the off current will increase by no more than two orders of magnitude at $1\,\mathrm{Grad}$[117]. By assuming that the discharge is mainly caused by this current, τ_{hold} can be linked to I_{off} using:

$$\tau_{hold} = 2\pi R_{off} C_{AC} = \frac{2\pi V_{DS} C_{AC}}{I_{off}} \tag{4.8}$$

Here C_{AC} is considered to be the main capacitance connected to the node. This determines an inverse relation between I_{off} and τ_{hold}. It should be noted that these assumptions identify the worst case scenario in terms of TID dependence. T_{hold} can be computed as above, consequently its value will also decrease proportionally. The new value for T_{hold} results to be $12.4\,\mathrm{\mu s}$. In this case χ_{oc} is $8.0\,\text{‰}$ with an average signal loss χ_{loss} of $1.80\,\text{‰}$ at $3\,\mathrm{MHz}$.
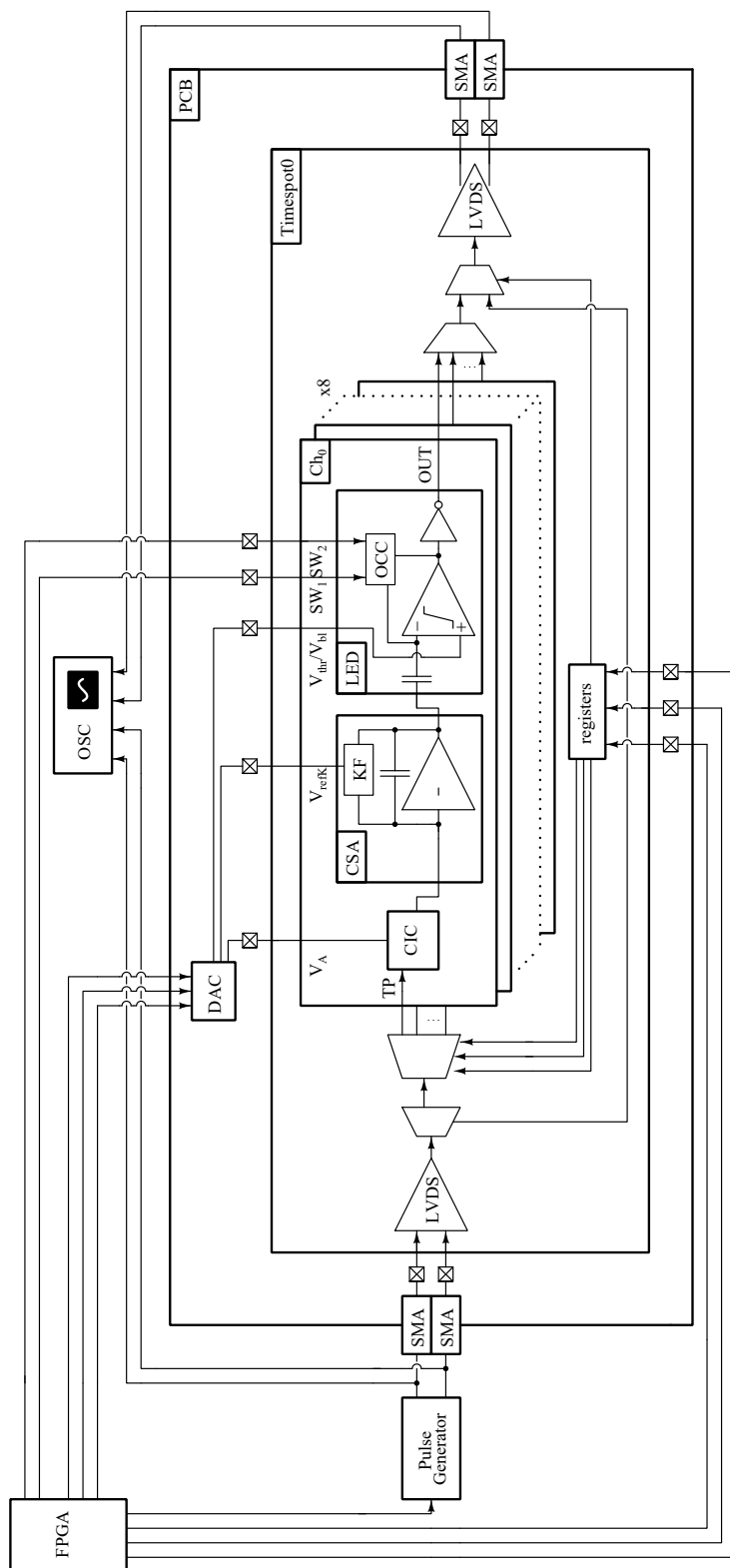
Figure 4.13: Schematic representation of the experimental setup. Bias currents and reference voltages are omitted for sake of clarity. They are provided using dedicated resistive dividers and DACs respectively. The FPGA board acts as master for the measure triggering and circuit setting. Differential trigger and output signals are used to avoid the degradation of the signals slew rate. The trigger signal is split and sent to the oscilloscope in order to be used as timing reference for the output signal. Internal registers are used to both select the DUT channel or bypass it.

95

Figure 4.14: Flow chart representing the Finite State Machine of the firmware implemented inside the Kintex7 FPGA. The firmware must implement both the procedure described in figure 4.4, program the onboard DACs and write the chip internal registers. The timing characteristics of the pulse, the OC timings, the analog levels can be programmed via an UART interface.
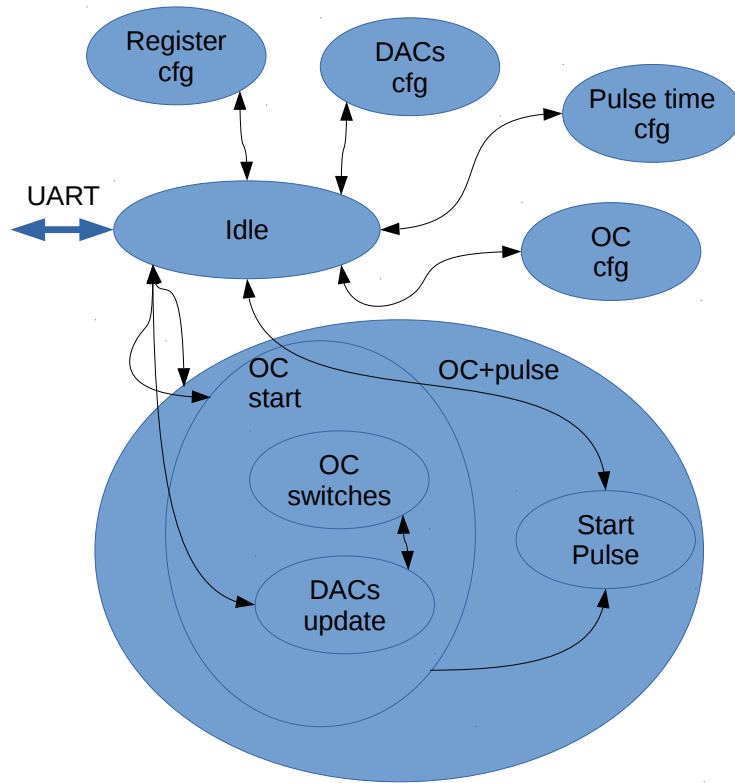
Chapter 5

Second Prototype: the Timespot1 Analog-FE

This chapter presents the Timespot1 analog front-end architecture and its implementation. Compared to the Timespot0 AFE (presented in 3) this front-end is built to be part of a 1024-channels matrix which implies additional problems from an architectural and an implementation point of view. This circuit must in fact be reproducible in many channels connectable to the respective TDC, and it must be organized in matrix configuration.

In this ASIC it is not possible to supply all the required reference, biases and digital controls via dedicated pads since they are reserved mainly for IO communication and power delivery (the ASIC architecture is explained in 2.2.2). The circuit must therefore include smart bias and reference cells which are able to generate the needed levels and built-in digital automation for the channel control.

Moreover, the pixel architecture has been updated to enhance its jitter performance versus power consumption and to make the scheme more radiation tolerant.

Section 5.1 presents the AFE architecture starting from the updated pixel electronics, moving then to the analog periphery and ending with the digital controls and TDC interface. Section 5.2 presents the circuit implementation of the pixel and service electronics, as well as their integration inside the matrix. The core channel operation is not described in this chapter since it is conceptually similar to the one of the first prototype which has been already discussed in 3.1. The only exception is the operation of the new core amplifier, which is discussed in 5.1.1. For the sake of convenience, the transistor level schematic, the transistor sizing and the channel layout are shown at the end of the chapter. . In order to better understand the design concepts and decisions presented in this chapter, it is suggested to have first read sections 1.2 and 1.3 of chapter 1.
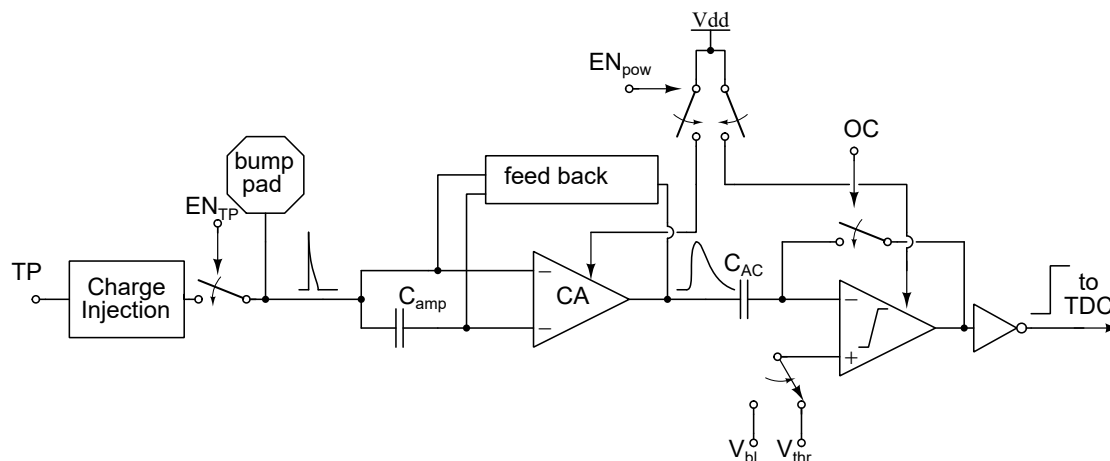
Figure 5.1: Analog pixel architecture of the Timespot1 ASIC. The channel features a CSA pre-amplifier connected to a discrete-time leading edge discriminator with offset compensation. The CSA capacitive feedback path (not drawn in the picture) is formed by the input transistors gate-drain capacitance, whereas the active feedback path is based on a Krummenacher filter. The whole channel can be powered down digitally. It also presents a charge injection circuit for test purposes, which can be disconnected during the normal operation.

5.1   Architecture

The AFE will be described by dividing it in three main blocks: the pixel electronics, the periphery electronics and the digital controls. From the point of view of the pixel electronics, the argument will focus on the updates made in relation to the first prototype.

5.1.1   Core Channel Updates

A schematic representation of the analog front-end is shown in figure 5.1 whereas its transistor level representation is presented in figure 5.15 at the end of the chapter. The scheme is largely based on one of the first prototypes (described in 3.1), therefore this section will outline the difference between the two architectures. In any case, a brief description of the electronics chain will follow.

The AFE channel comprises a CSA as first stage and a leading edge discriminator as second stage. The CSA is used to convert the input current signal into a voltage one with an amplitude proportional to the input integrated charge. Since this CSA implementation features a constant current discharge, the output signal exhibits a ToT proportional to its amplitude and therefore the input charge. The discriminator features a discrete time offset compensation circuit used to cut down the per-channel variations. Each CSA can be both connected to the respective

sensor bond-pad and to its own charge injection circuit. Additionally, this version, offers the possibility of shooting down the whole channel by acting on a switch positioned between the circuit and its power rail. This function can be used to eliminate the power consumption of inactive channels as well as preventing faulty channels from interfering with neighboring ones.

The updates on the previous scheme focus on three aspects: timing resolution improvements, radiation hardness improvements and integration of new auxiliary features. A general measure used in the whole scheme was to use only PMOS transistors to realize the switches. This was done because, pre-irradiation, this realization does not present any downside compared to an N-type one; but it presents an important advantage after irradiation. As shown in [116], the PMOS transistor features a higher resistance to TID in terms of its off-current, making it more suitable for realizing switches. The changes applied to the individual blocks will be described in the following paragraphs. The only block that has not changed is the discriminator core.

Charge Injection Circuit

First, the switches inside the circuit were changed to p-type ones for the aforementioned reason. The cell was also made removable in order to avoid interference during the normal operation. This can be performed by acting on the switch formed by M_{i7} using the digital signal $\overline{EN}_{inj}$. This switch is connected after the injection capacitor C_{inj} in order to mask its capacitance during normal operation. C_{inj} itself was reduced from $35\,\mathrm{fF}$ to $10\,\mathrm{fF}$ in order to produce a larger range of input charges. The injection levels were also changed: in the previous case the analog voltage step swings between ground and level V_A, this time around the voltage is formed providing two levels V_{A1} and V_{A2}. This approach offers a more flexible setup which allows producing current pulses of both polarities and decides the absolute position of the pulse inside the range. Lastly, the sensor capacitance emulation portion of the circuit has been removed due to the fact that the real sensor will be connected.

Core Amplifier

The core amplifier architecture has been changed in comparison to its predecessor in order to maximize its jitter performance versus power consumption. The function of the core amplifier is to charge the feedback capacitance. In order to produce signals with high time accuracy, this block must feature high gain, high bandwidth and low noise. These characteristics are all tied to the available power; it is therefore critical to maximize the architecture performance for a power consumption under $15\,\mathrm{\mu W}$.

In this version, the block has been implemented using a single inverter-like stage. A challenging aspect of a linear inverter-based amplifier is the proper biasing of its
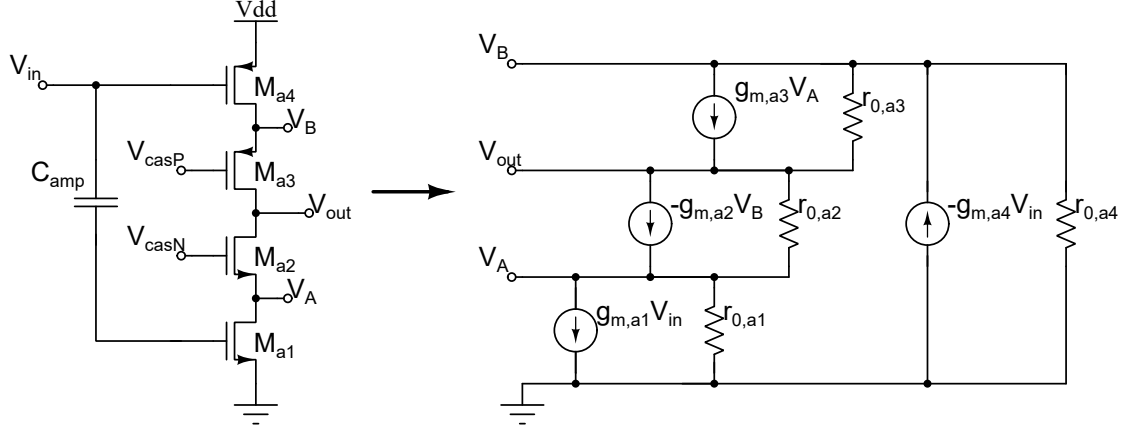
Figure 5.2: Small signal equivalent circuit of the core amplifier of the Timespot1 CSA. The transistor level circuit is shown on the left as a reference. The approximation is derived at the band pass frequency of the AC coupling.

DC operating point. In particular, it is difficult to force a reliable DC input voltage while also forcing the desired bias current on the stage. To achieve this, the NMOS input has been AC coupled to the PMOS one, splitting the DC voltages of the two transistors. These two nodes are then biased using two separate feedback paths.

The capacitance value of the AC coupling capacitor C_{amp} has been optimized using the simulation in order to obtain the optimal band-pass frequency. In particular C_{amp} size influences both its intended value and the parasitic capacitance coupling its terminal to the substrate. The first capacitance determines the high-pass component of the equivalent filter, whereas the second one is a low-pass component.

Compared to the single input transistor stage of the previous prototype, this architecture provides double the transconductance with the same bias current. In order to boost the open loop gain of the architecture, the two transistors have been cascoded. These two parameters are critical to minimize the jitter of the CSA as shown in equation (3.30) derived in 3.2.

These advantages can be verified considering the small signal model of the core amplifier presented in figure 5.2. The nodal equation system extracted from this model is:

$$
\begin{cases}
g_{m,a1}V_{in} + \frac{V_B}{r_{0,a1}} & = -g_{m,a2}V_B + \frac{V_{out}-V_B}{r_{0,a2}} \\
-g_{m,a2}V_B + \frac{V_{out}-V_B}{r_{0,a2}} & = g_{m,a3}V_A + \frac{V_A-V_{out}}{r_{0,a3}} \\
g_{m,a3}V_A + \frac{V_A-V_{out}}{r_{0,a3}} & = -g_{m,a4}V_{in} - \frac{V_A}{r_{0,a4}}
\end{cases}
\tag{5.1}
$$

Even though the system can be easily solved, it risks to be expressed in a confusing form. Therefore, it can be useful to introduce the two following terms in order to simplify the notation:
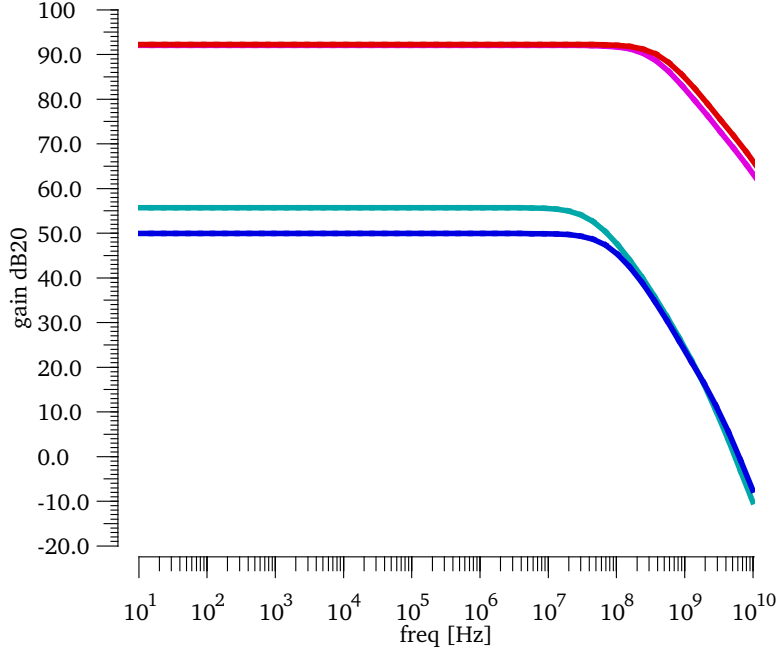
100

Figure 5.3: Simulated frequency responses of the core amplifiers found in Timespot0 and Timespot1. The light blue line represents the split-branches telescopic cascode amplifier, whereas the purple one represents the cascoded inverter. The cascoded inverter exhibits an overall higher gain and bandwidth. The blue and red lines represent a higher power condition. The cascoded inverter is able to better retain its gain even while being biased with a larger current. Both the core amplifiers were simulated by forcing their input DC points to the optimal voltage. For this reason the cascoded inverter simulation does not include the AC coulping.

$$\chi_{12} = \frac{g_{m,a2}r_{0,a1}r_{0,a2} + r_{0,a1}}{g_{m,a2}r_{0,a1}r_{0,a2} + r_{0,a1} + r_{0,a2}} \tag{5.2}$$

$$\chi_{43} = \frac{g_{m,a3}r_{0,a4}r_{0,a3} + r_{0,a4}}{g_{m,a3}r_{0,a4}r_{0,a3} + r_{0,a4} + r_{0,a3}} \tag{5.3}$$

Using χ_{12} and χ_{43} it is straightforward to write the equations which relates V_A and V_B to V_{in} and V_{out}, these can be used to derive the input-output equation:

$$-V_{in}\left(g_{m,a1}\chi_{12} + g_{m,a4}\chi_{43}\right) = V_{out}\left(\frac{1 - \chi_{12}}{r_{0,a2}} + \frac{1 - \chi_{43}}{r_{0,a3}}\right) \tag{5.4}$$

At this point, it is possible to compare the single input cascode to this architecture, making explicit their common terms. First of all the equation (3.9) expresses the cascode equivalent transconductance G_m which in this case it can be expressed as:

$$G_{m,a1} = g_{m,a1}\chi_{12} \tag{5.5}$$

$$G_{m,a4} = g_{m,a4}\chi_{43} \tag{5.6}$$

101

In the same way, the cascode equivalent output resistance can be found using (3.11):

$$r_{cas12} = \frac{1 - \chi_{12}}{r_{0,a2}} \tag{5.7}$$

$$r_{cas43} = \frac{1 - \chi_{43}}{r_{0,a3}} \tag{5.8}$$

The low-frequency transfer function for the cascoded inverter can thus be written as:

$$\frac{V_{out}}{V_{in}} = -(G_{m,a1} + G_{m,a4}) \cdot r_{[r_{cas12}//r_{cas43}]} \tag{5.9}$$

This relation makes it clear that the total stage transconductance is the sum of the ones of the two inputs making larger than the one found in (3.8) for the same bias current. The overall open-loop gain is also larger, since both the transconductance and the output resistance are larger, with comparable sizing, than the previous case . In fact, in this the output resistance is the parallel of two cascodes equivalent resistances whereas in the previous case (equation (3.10)) it was the parallel between a cascode resistance and a transistor output resistance. Figure 5.3 shows this aspect by comparing the transfer functions on the frequency domain of the two architectures with similar sizing.

CSA

The first update was done in order to maximize the charge to voltage gain by minimizing the feedback capacitance. The solution was to remove the capacitor inserted in feedback. The capacitive feedback path is then formed by the input transistors gate-drain and parasitic capacitance. The capacitance value is estimated to be around 3 fF.

The active feedback path has been updated in order to accommodate the cascoded inverter architecture. The two feedback circuits are implemented using two complementary realizations of the Krummenacher Filter in order to have the best DC matching. The NMOS variant is used to bias the PMOS input, while the PMOS one is used for the NMOS input. The p-type branch is also used to compensate the sensor leakage current. A detailed description on how this biasing is implemented is discussed in 5.1.2. In order to obtain the same discharge time of a single Krummenacher feedback, the individual bias currents needs to be halved. The discharge current value is programmable and can be set between 25 nA and 100 nA, this results in a ToT around 100 ns to 20 ns for a 2 fC input charge.

Using the feedback path to set the proper DC input level was not the only explored solution. It has been tried to bias the inputs using two external AC coupled voltage levels. This approach brings two downsides: the input capacitance will be larger and the circuit will suffer more mismatch variation. The first point is
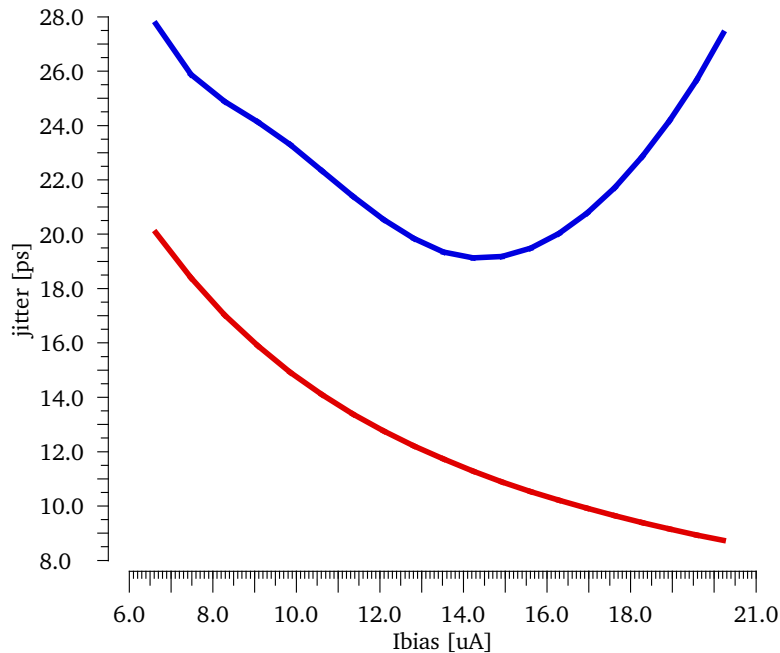
Figure 5.4: Simulated comparison of the output CSA jitter as function of the bias currents between Timespot1 (in red) and Timespot0 (in blue). The two circuit were sized similarly. The new approach offers a general jitter improvement over the old one. The old architecture offers a trend inversion for high current values related to the lack of biasing in the range, and thus an output resistance degradation. This can be prevented with a larger sizing, which will in turn spoil the performance at low bias current. The input capacitance C_S has been set to 150 fF. Both the CSA were optimized to operate with a nominal bias current I_{bias} of 15 µA.

related to the sizing of the coupling capacitor: in order to avoid signal loss from the channel to the reference, it needs to adequately large to low-pass filter the signal band. It will however increase the input capacitance spoiling the jitter performance of the architecture. The second point is related to the necessity to distribute a single voltage level across many channels, making the bias conditions susceptible to the per-channel variations of the input transistors.

Another solution could have been to implement an AC coupled auto-zeroing scheme similar to the one of the discriminator. In this way, each channel will auto compensate itself, eliminating the mismatch dependency. This approach, however, will come with the already described complications and it will risk interfering with the OC procedure. Hence, the adopted solution was the optimal one for this application.

Overall, the larger conductance and open loop gain of this configuration will result in better jitter performance in relation to the power consumption (shown in figure 5.4) and input capacitance (figure 5.5). Overall, the updated CSA has been
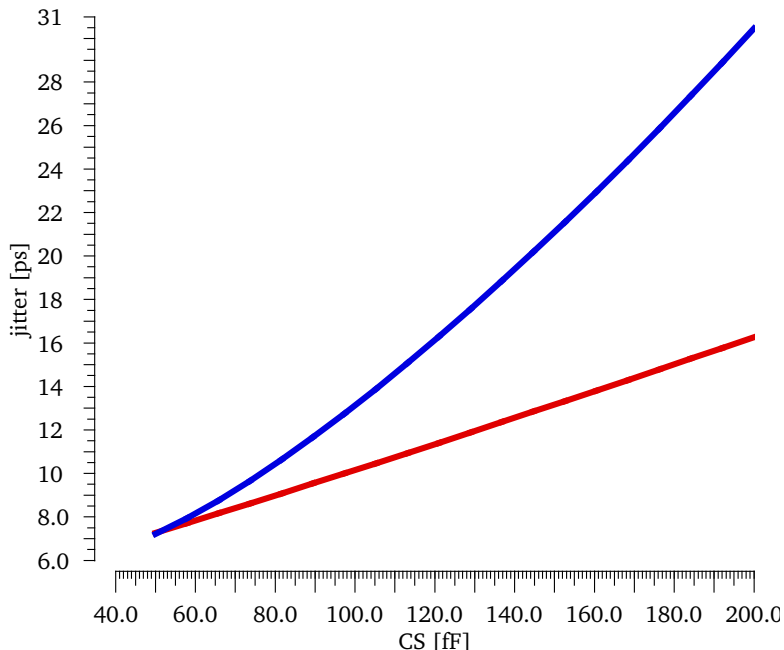
Figure 5.5: Simulated comparison of the output CSA jitter as function of the input capacitance between Timespot0-like architecture (in blue) and Timespot1 (in red). The higher transconductance and open-loop gain ensures a lower jitter to capacitance slope. The two circuit have been simulated with the same bias current I_{bias} of 15 µA.

designed to absorb more current, corresponding to a maximum power consumption of 15 µW. This value corresponds to approximately one third of the pixel power budget, leaving the rest to the TDC. In comparison, the CSA consumed at most 6 µW of power in the previous version.

Offset Compensation Circuit

The first modification applied to the offset correction circuit is to change every switch to PMOS switches. The second difference is the change of the way in which the threshold terminal is connected. In the previous case, the baseline and threshold values were provided with a single wire, in which the signal is dynamically changed when required by the OC procedure. In this implementation, there are lines dedicated to the baseline and threshold levels which are connected to the discriminator through a switch selector. This selector is composed of two P-type switches (M_{o5} and M_{o6}) and driven with a digital signal SEL_{thr}. The signal is inverted at the gate of one of the two transistors in order to avoid undesired states such as: driving the discriminator in high impedance or shorting the two lines. The first reason to employ this method is that the integrated sigma-delta DAC would

not allow such a fast change in its output. Additionally, this method also allows performing the OC in one channel independently from the others. Since the OC consumes a noticeably higher power compared to a channel at rest, this method allows avoiding the interference between the channels and distributes more evenly the consumption spikes across the matrix.

5.1.2 Analog Periphery

This part of the circuit is common to many channels and is located in the analog-services column as shown in figure 5.14. Its function is to provide all the analog levels necessary to operate the pixel electronics. Each pixel requires 13 analog references: 4 levels, 2 references, 4 biases and 3 cascode voltages. Obviously, cascode and bias voltages are naturally the same for every channel; on the contrary the Krummenacher voltage reference can be fine-tuned per-channel. Regulating V_{ref} per channel is however unfeasible: it has been chosen, instead, to equalize the channel via the OC and per-channel TA versus ToT curves. The four voltage levels are used in pairs for the charge injection and the OC. Due to the architecture explained above, these levels can be shared among all the pixels leaving only the digital controls tied to the single channel. In summary, all the 13 levels can be simultaneously shared between all the pixels.

The actual composition of the periphery is shown in figure 5.6. See the figure caption, for a detailed scheme description. The bias-cell is programmable via a thermometric code in order to set precisely the channel power consumption, whereas the reference-cell is designed to automatically adapt to these settings. These two blocks will be detailed in the next paragraphs.

Bias Cell

The bias-cell can be conceived as the input stage of the analog-matrix setting circuit. This block allows setting independently the bias currents of: the core amplifier, the core discriminator stages and the Krummenacher filters. The first two settings are crucial to the definition of the channel power consumption: by lowering these currents, it is possible to lower the power consumption at the expense of the system jitter and the total propagation delay. On the other hand, the Krummenacher controls are used to program the discharge time, which in turns changes the slope of the charge versus ToT curve.

A schematic representation of the cell is shown in figure 5.7 which includes a detailed description. In short, the bias-cell allows to obtain the desired currents from a reference current, and setting their value via a digital control. Therefore, the cell requires only one analog reference to produce its outputs. The outputs of the cell are the bias voltages of the related current mirrors. Distributing these static voltages rather than the currents reduces the IR drops along the lines, since
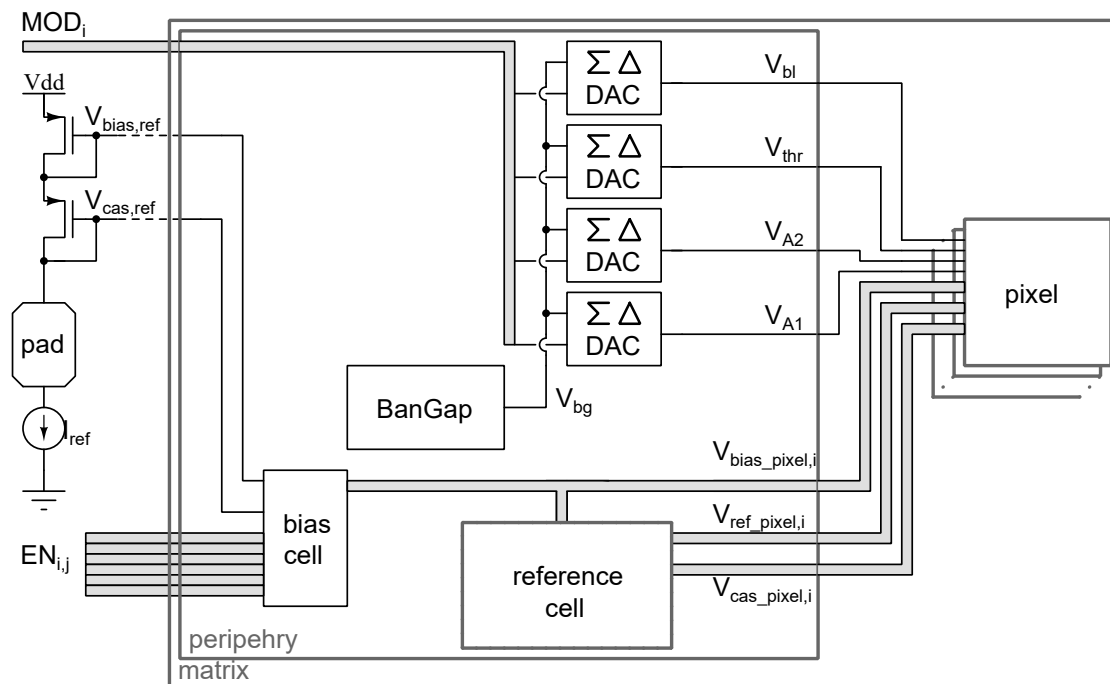
Figure 5.6: Schematic representation of the Timespot1 analog periphery. The periphery includes a programmable bias-cell (detailed in figure 5.7), a reference-cell (detailed in figure 5.9) and four $\Sigma\Delta$ DACs connected to a Bandgap. This block provides all the 13 voltage levels required to operate the pixel electronics. The DACs are used to produce the service levels V_{thr}, V_{bl}, V_{A2} and V_{A1}; whereas the rest is used for references ($V_{ref\_pixels,i}$), biases($V_{bias\_pixels,i}$) and cascodes($V_{cas\_pixels,i}$). The analog inputs of this block are the bias and cascode voltages ($V_{bias,ref}$ and $V_{cas,ref}$) generated by the current mirror of the external reference current I_{ref}, whereas the digital ones are the four modulators of the DACs (MOD_i) and the thermometric code used to program the bias-cell (EN_{ij}).

they will only drive transistor gates. It is noted that the reference current is fed to the bias cell with the same mechanism with the purpose of both limiting the IR drop and in order to use this reference in other blocks of the ASIC.

Reference Cell

The purpose of this block is to produce the cascode voltages for the cascoded transistors inside the channel and to produce the two voltage references required by the Krummenacher filters.

This operation is trickier than it seems since the core amplifier bias current is not directly set by acting on a bias transistor using a current mirror configuration. Apparently, the only external controls of the core amplifier are constituted by the
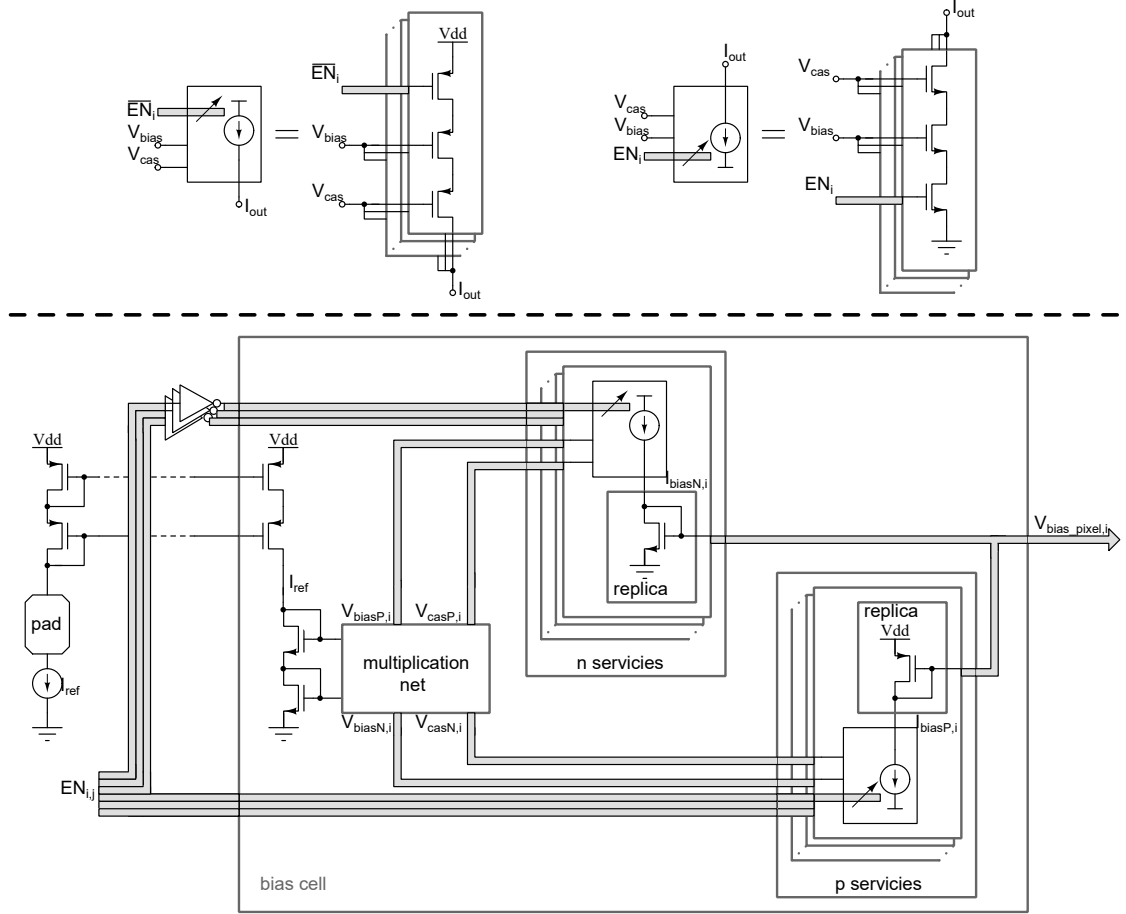
Figure 5.7: Schematic representation of the Timespot1 programmable bias-cell. The external reference current I_{ref} is copied inside the cell and connected to two cascoded current mirror chains: one for multiplication and the other for de-multiplication. The desired value for each cell is obtained by programming with a set of thermometric codes (EN_{ij}) a parallel configuration of switched current mirrors (shown on top) copying the desired unit currents. These currents are locally fed to a replica transistor of the target circuit in order to produce the voltage biases ($V_{bias\_pixel,i}$) to be distributed to the reference-cell and to the pixel matrix.

two cascode voltages V_{casP} and V_{casN} driving the gates of transistor M_{a2} and M_{a3}. When biased properly, this configuration makes these transistors act as active loads rather than current sources, making them poor candidates to control the branch current.

A starting point to design this bias-circuit is to treat the cascoded inverter as a single-ended telescopic cascode amplifier. These two circuits present the same topology since they have the same number of transistors connected in the same way. The only difference between the two configurations are the functions of their
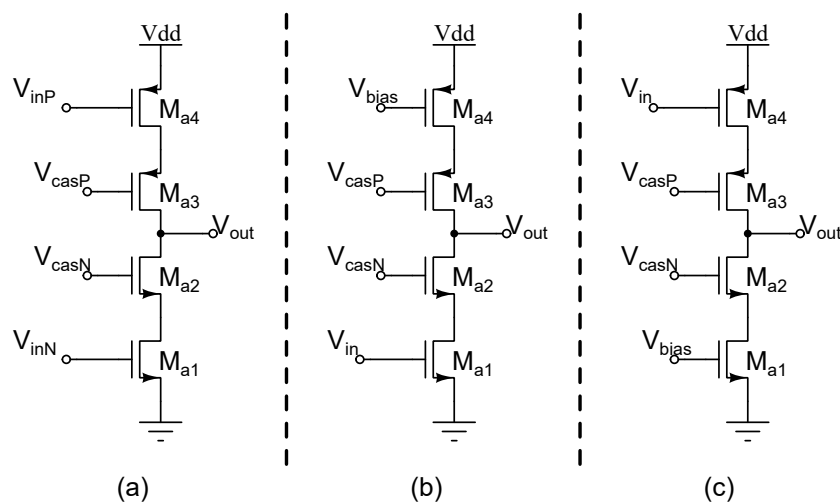
Figure 5.8: Analogy between the cascoded inverter (a) and the telescopic cascode amplifier in the N-type (b) and P-type (c) implementations.

terminals. This concept is shown in figure 5.8. Taking for example a N-type telescopic cascode amplifier, the only difference between it and the cascoded inverter is the function of the signal at the gate of M_{a4}. In the first case it controls the bias current whereas in the second case it acts as one of the two input terminals. Using this analogy, it is clear that the DC levels of M_{a1} and M_{a4} can be used to set their branch currents.

As explained in 5.1.1, these DC voltages are set by the feedback paths, in this way the related mismatch dependency is moved from the highly sensitive M_{a1} and M_{a4} to the one of the two Krummenacher filters. In particular, this is done by acting on the two reference voltages V_{refN} and V_{refP} since they control the DC levels at the Krummenacher filter terminals.

Therefore, the reference-cell scheme must reactively find the correct values for these voltages. The cell scheme is presented in figure 5.9 The description of the mechanism is placed on the caption in order to facilitate the discussion.

The principle used to create this cell is to use combinations of replicas of the core channel components with the purpose of improving the overall matching. The circuit in the cell automatically finds its equilibrium point in relation to the currents imposed from the bias-cell. The two advantages of this architecture are connected to its ease of use and adaptability.
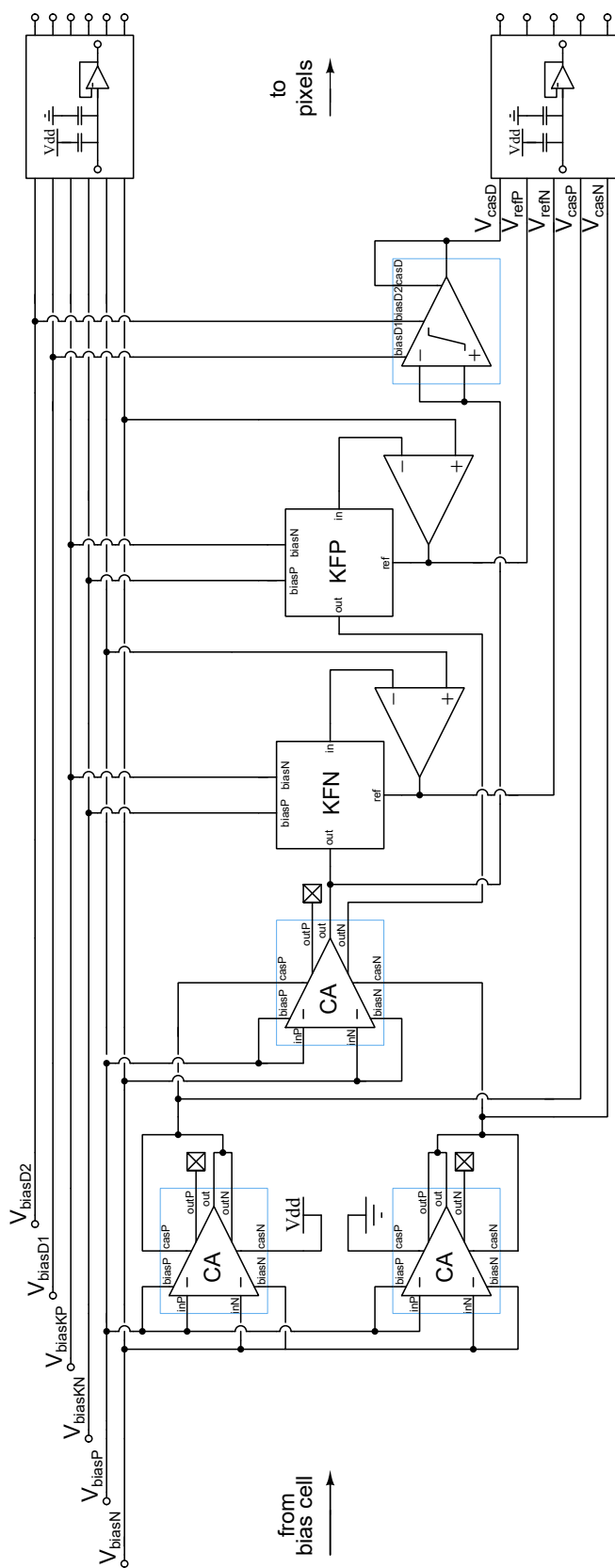
Figure 5.9: Schematic representation of the Timespot1 reference-cell. This cell receives the bias voltages generated in the bias-cell and automatically produces the 3 cascode voltages and the two reference voltages needed to operate the pixel. First, the core amplifier cascode voltages are generated separately in two replicas: the input transistor is biased using the level generated with a current mirror while the cascode transistor is diode-connected using the rest of the circuit as a load. The obtained levels are fed to another replica, which is used to obtain the core amplifier output levels. This block is then connected to the replicas of the feedback paths, as in the main channel. Two negative feedback loops are used to obtain the V_{ref} values that would bias the CSA to the desire DC points. Finally, the core amplifier output is also used as the common mode of the core discriminator replica in order to set the proper cascode voltage. All the signals are decoupled and buffered with a voltage follower towards the pixel matrix. Particular attention has been paid to ensure the frequency stability of these low-frequency loops.
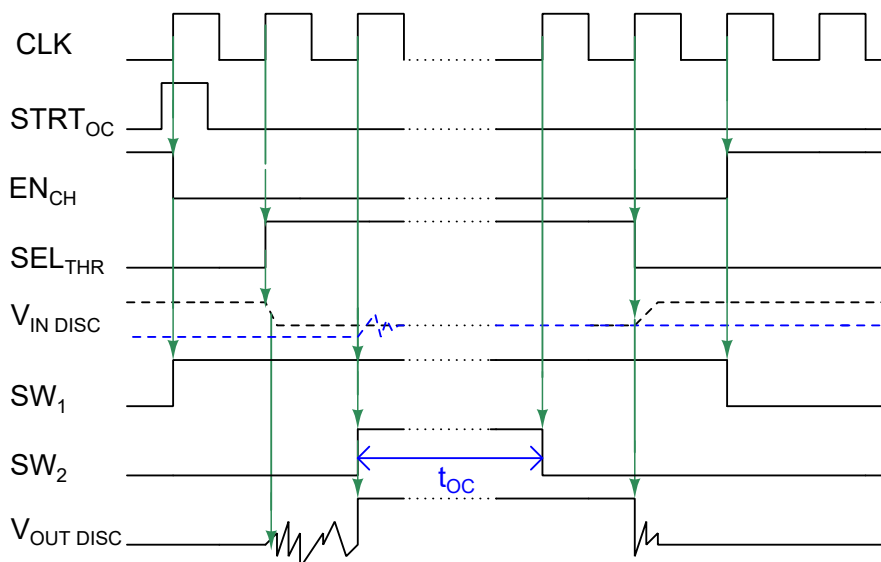
5.1.3 Digital Controls and TDC interface

Figure 5.10: Waveform representation of the OC control logic. The arrival of the $STRT_{OC}$ signal issues the start of the OC procedure at the next clock (CLK) pulse. The procedure will start by dis-enabling (acting on EN_{CH}) the channel and acting on the SW_1 signal in order to stabilize the discriminator. After this, the discriminator input is switched between the baseline voltage to the threshold one acting on SEL_{THR}. At this point, the effective compensation will take place by rising the SW_2 signal for a time period t_{OC}. Finally, the channel is re-activated by rising EN_{CH} and by lowering SW_1. Disabling the channel during the OC protects the TDC from spurious signals generated by the discriminator (indicated by the jagged line in $V_{OUTDISC}$).

The AFE digital controls are implemented in the digital-row that is the same in which the TDC is integrated. The controls signals must be dedicated to each channel, whereas their circuitry can be shared among many channels. In this way the controls are applied to one channel at a time. The digital controls must be able to:

- Switch on and off the channel power.

- Switch between the test mode and the measure mode.

- Perform the offset compensation procedure (explained in figure 5.10).

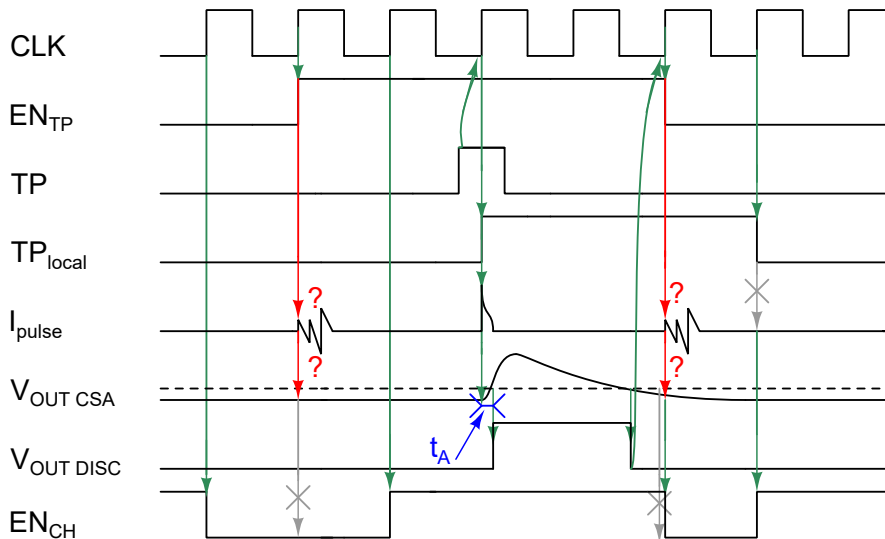- Perform the test pulse procedure (explained in figure 5.11).

Figure 5.11: Waveform representation of the test pulse control logic. The channel is enabled by acting on the EN_{TP} signal. Whenever this signal changes, it causes an undefined behavior in the channel, therefore the channel must be disconnected from the TDC using EN_{CH}. Before injecting a signal it is necessary to wait the end of the transient of the EN_{TP} perturbation. The test pulse (TP) is always injected with a fixed phase in relation to the 40 MHz clock (CLK). In order to avoid the issues related to EN_{TP}, it is better to activate the test-pulse mode and to pulse the channel in two different moments.

The two procedures involve a larger power consumption compared to a static channel, which risks to overload the power net. Therefore, these must be performed simultaneously only on a small subset of channels. In particular, the OC procedure must be repeated periodically: this has been implemented by daisy-chaining together the 32 channels of a row so that the OC will be performed in only one channel each time. The period of the OC procedure is programmable between 125 µs and 1 ms; and it is common to the whole matrix.

During these two procedures the channel will behave in an unwonted way generating spurious events that risk to produce spurious data or block the TDC. For this reason, the TDC is able to veto the channel via an enable signal EN_{CH}. The veto is performed with a multiplexer inserted between the AFE and the TDC. This veto is also used during the TDC self-calibration in order to prevent incoming signals to spoil the procedure.

A schematic view of the control logic with a detailed explanation is presented in figure 5.12 in the next page.
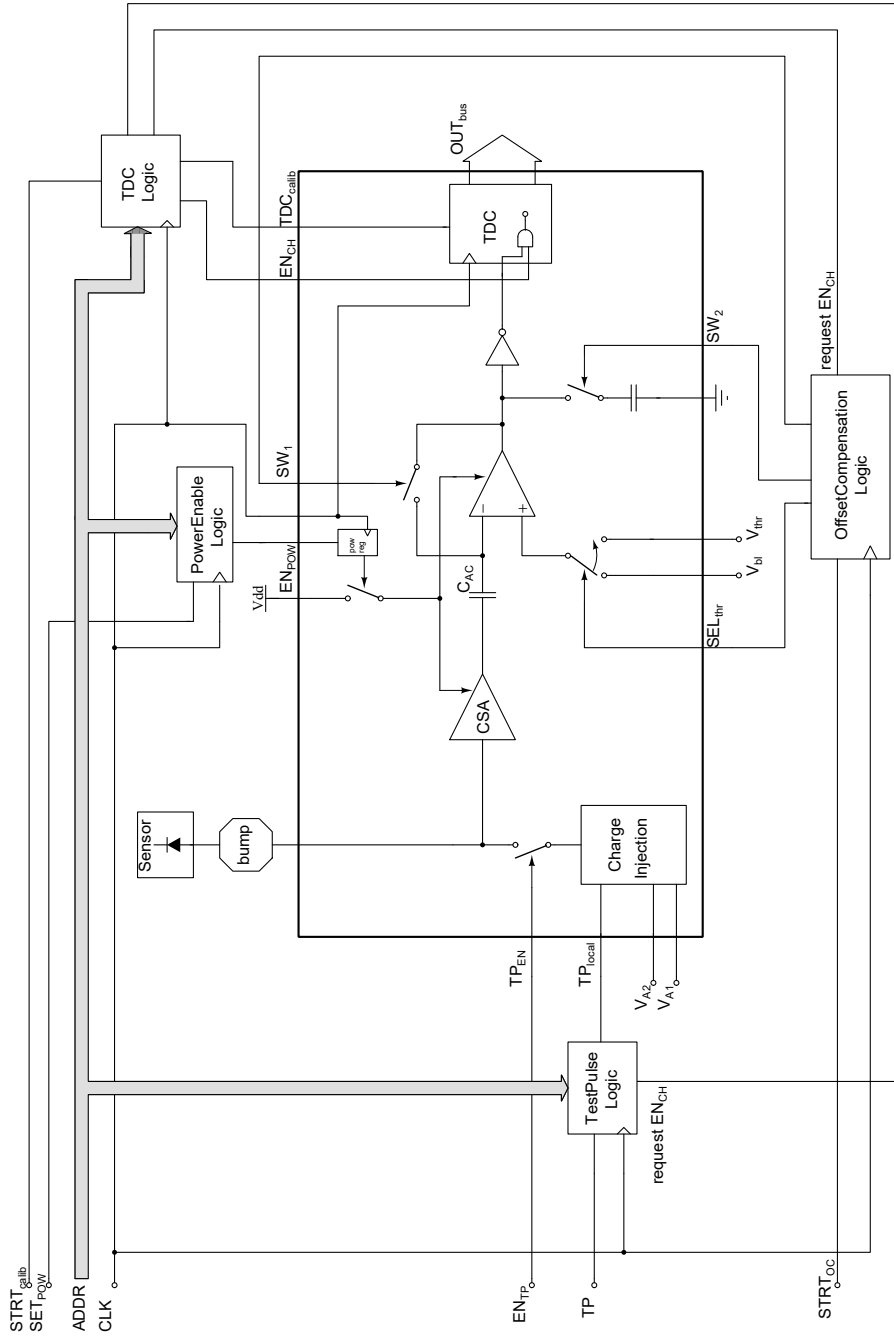
Figure 5.12: Schematic representation of the Timespo1 AFE control logic. The logic comprises some blocks which act on the individual channels, whereas others act on a line basis. In any case, the logic is shared among all the 32 pixels of a double digital-row. Power enabling and test pulsing are per-channel controls and can be addressed to the desired pixel using its address number (*ADDR*). The OC procedure is instead issued using *STRT_{OC}* in a daisy-chain like way from the first channel to the last. All procedures must be actuated in tandem with the TDC.
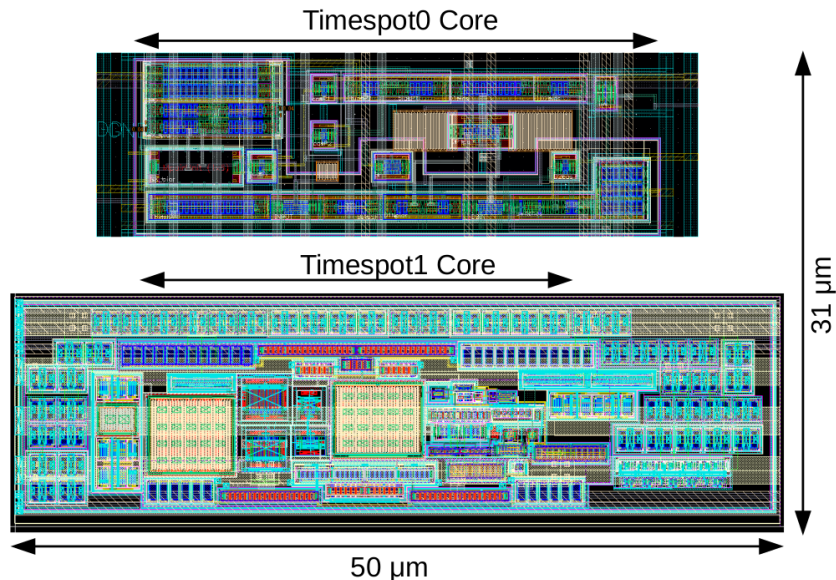
5.2 Implementation

Figure 5.13: Comparison between the Timespot0 and Timespo1 pixel electronics. Despite integrating more features, the core area of the two pixels is the same. All the free area around the core pixel is used to integrate decoupling capacitors.

This section discusses the issues and design decisions regarding the actual implementation of the AFE both in the pixel circuits and in the periphery. For general implementation concepts regarding the technology node, please read section 3.3.1. The section will start by describing the implementation of the core channel inside the pixel area. It will then move to describe the implementation of the periphery. Lastly, the overall floor plan will be discussed.

5.2.1 Analog-Pixel

The total pixel area is $50\,\mu\text{m} \times 55\,\mu\text{m}$: it is recalled that the reduced pitch in the horizontal direction was chosen in order to gain some area to implement the periphery in the analog-column. Of the total pixel area, the AFE core occupies $50\,\mu\text{m} \times 16\,\mu\text{m}$, which also includes the required digital buffers and decoupling capacitors. This corresponds to $29\,\%$ of the pixel area: the rest of the area is occupied by the TDC, the digital controls and the data transmission lines. Figure 5.13 shows a scale comparison between the old pixel electronics and the new one.

The layout of the pixel is shown in figure 5.16 at the end of the chapter. Of the nine available metal layers, the first four are used for internal interconnection. In this way, it was possible to tightly interconnect the transistors with the thinnest

metal lines. The first metal was used exclusively to connect the substrate contacts on the guard-rings, since this metal is the most coupled to the substrate ground. These guard-rings polarize the deep-nwell and surround every block including the whole channel.

All the devices are positioned on the substrate in a configuration which minimizes the length of the signal path, reducing its parasitics. The transistor sizing of the core channel is shown in table 5.1. Particular care was used to over-size the CSA elements in order to contain its process and mismatch related variations, given that this block is the most sensitive one. Additionally, the P and N-type Krummenachers were sized in order to have their bias voltages compatible: the current source of one has half the width of the other one. In this way it was possible to spare two signals from the periphery.

The other connections were traced afterwards. All the auxiliary signals are routed using metal five in order to be connected to the common lines that are found from metal six to eight and spans from the left to the right of the cell. Metal six and seven are used for all the analog levels driving gates, whereas thicker metal eight is used for the OC and injections levels, since they experience the largest current spikes. All these analog DC signals are decoupled using MOS capacitors.

Finally, the power and ground nets are found on metal nine, which is drastically thicker than the other ones and thus will sustain the dynamic IR drop of the active channels. The lines are built in two sets of interleaved parallel lines in order to make them self-decoupling. The set of switches for the power-enable of the amplifier and the discriminator were oversized in order to keep the resistance to VDD as low as possible. Two independent switches were implemented for the two blocks with the aim of preventing as much as possible the formation of cross talk between the two circuits. All the power and ground nets were also decoupled with MOS capacitors.

The digital signals including the TDC connection are derived on the bottom of the cells where the TDC will be found. The top priority was reserved to the TDC input since it is a critical signal for timing. All the input signals were buffered with two inverters in series for each one. These inverters are powered with analog power and ground: in this way the analog circuit will be insulated from the disturbances on the digital and ground.

The cell is built to be tileable in the horizontal direction, the analog connections are simply realized by joining two cells on the short side. Likewise, the TDC connections are simply formed by joining the TDC top side to the AFE bottom one.

The connection to the sensor is realized with hexagonal bond-pads about $30\,\mu\mathrm{m}$ wide. The bump-pads are built using the top metal, therefore the power rails cannot be made in this area. The matrix of the bond pads is skewed in the horizontal direction in relation to the one of the channels, therefore each input net was specifically routed to its channel inside the analog-row. The vertical position of the pad is centered on the pixel area, therefore it is found mostly above the TDC with a minor
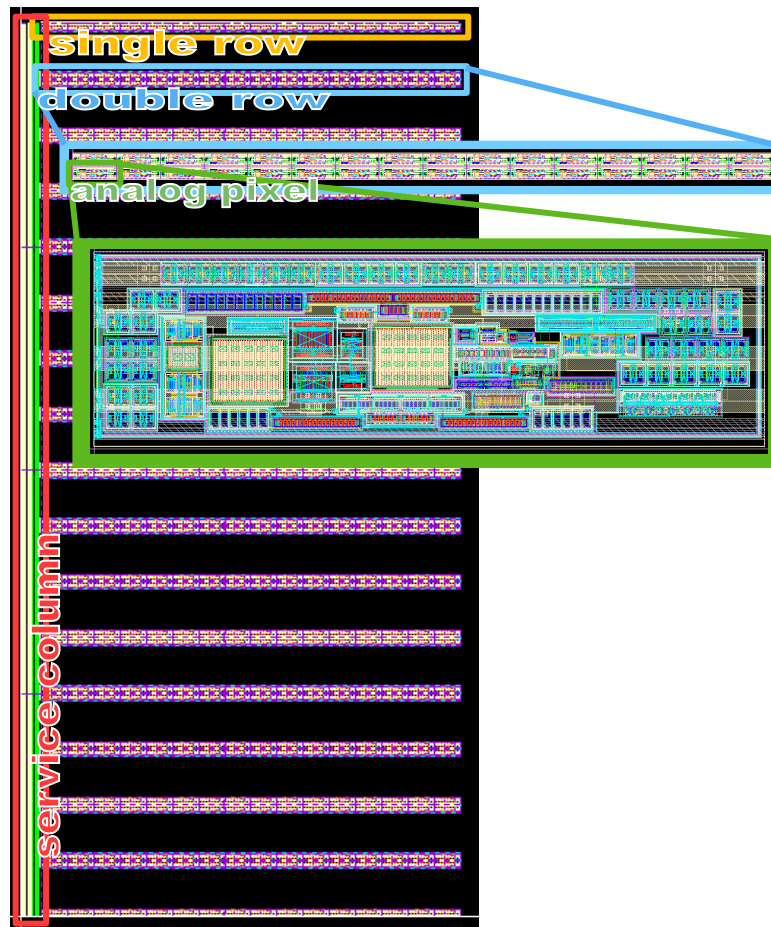
Figure 5.14: Floor-plan of the Timespot1 analog matrix. The most relevant blocks are highlighted. The two single-rows can be seen at the top and bottom of the matrix. The digital-matrix is complementary to the analog one and occupies the block space.

superposition with the AFE. In any case, no net passes directly below the pad or around it in order to reduce capacitive coupling which can disturb this sensitive node.

5.2.2   Analog-Column

The periphery is implemented in the analog-column which is obtained from the sum of the pixels residual areas. The total width of this residue is $16\times5\,\mu\text{m}$ for a total of $80\,\mu\text{m}$, but only half of it is dedicated to the analog-column whereas the other half is dedicated to the digital-column. Therefore it is shaped like an elongated rectangle in the vertical direction with an area of $40\,\mu\text{m}\times1.769\,\text{mm}$.

The power distribution and voltage levels distribution are implemented with the

same approach as in the pixel using the highest metal layers, with the difference that they run in the vertical direction. The connection to the power and ground pads is created at both the bottom and top of the column. These nets are joined together with the ones of the rows at the geometrical junction points.

The floor plan of the analog-matrix is shaped like a comb, with the analog-column being the shaft and the double rows being the teeth. This structure is shown in figure 5.14. 30 of the 32 rows are organized in a double-row configuration, with a bottom row adjacent to a mirrored top-row. The other 2 rows are the terminal ones, and therefore they are unpaired. This configuration allows connecting the TDC matrix to the AFE matrix by fitting together the two comb like structures. The deep-N-well in which all the analog circuits are built is shaped like the analog matrix with guard-rings insulating each sub-block.

The devices are implemented homogeneously along the column using only metal one to three in order to give more headroom to the redistribution nets. However, the usage of the bottom layers would not affect the operation of the periphery since it is composed of only low-frequency blocks. For the redistribution of the nets used inside the periphery, a vertical lain is obtained on the external edge of the analog-column. In this way, the signals can be routed using the first three metal layers.

The bias-cell is positioned on the bottom of the analog-column since it has the highest number of input signals (the thermometric codes bus) which are fed from the bottom. The current mirrors inside this block are oversized in order to reduce the mismatch between them. For the same reason, the replica of the transistors to be biased are repeated four times and connected in parallel. As a result, the bias currents programmed inside the cell are four times larger than the ones in the channels. In the same way, the reference-cell was repeated five times: this is the number of cells that would fill the analog-column empty area. The resulting signals are all shorted together, creating an averaging effect. The four $\Sigma\Delta$ DACs are positioned alternately with the reference cells.

All the nets distribute along the column and to the rows were carefully buffered and decoupled. The voltage followers used as buffers have both the function of protecting the source circuit from the destination one and to provide enough current to sustain the pixels demand. The gate leakage current of a typical transistor inside the core channel is in the order of tens of picoampere. However, when multiplied per 512 channels, the total current would reach tens of nanoampere. This could deviate a substantial amount of current from the current mirrors. The current is dynamic and even higher for the voltage level provided by the DACs. These last nets were especially carefully decoupled in order to damp the absorption spikes and filter the modulator residuals.
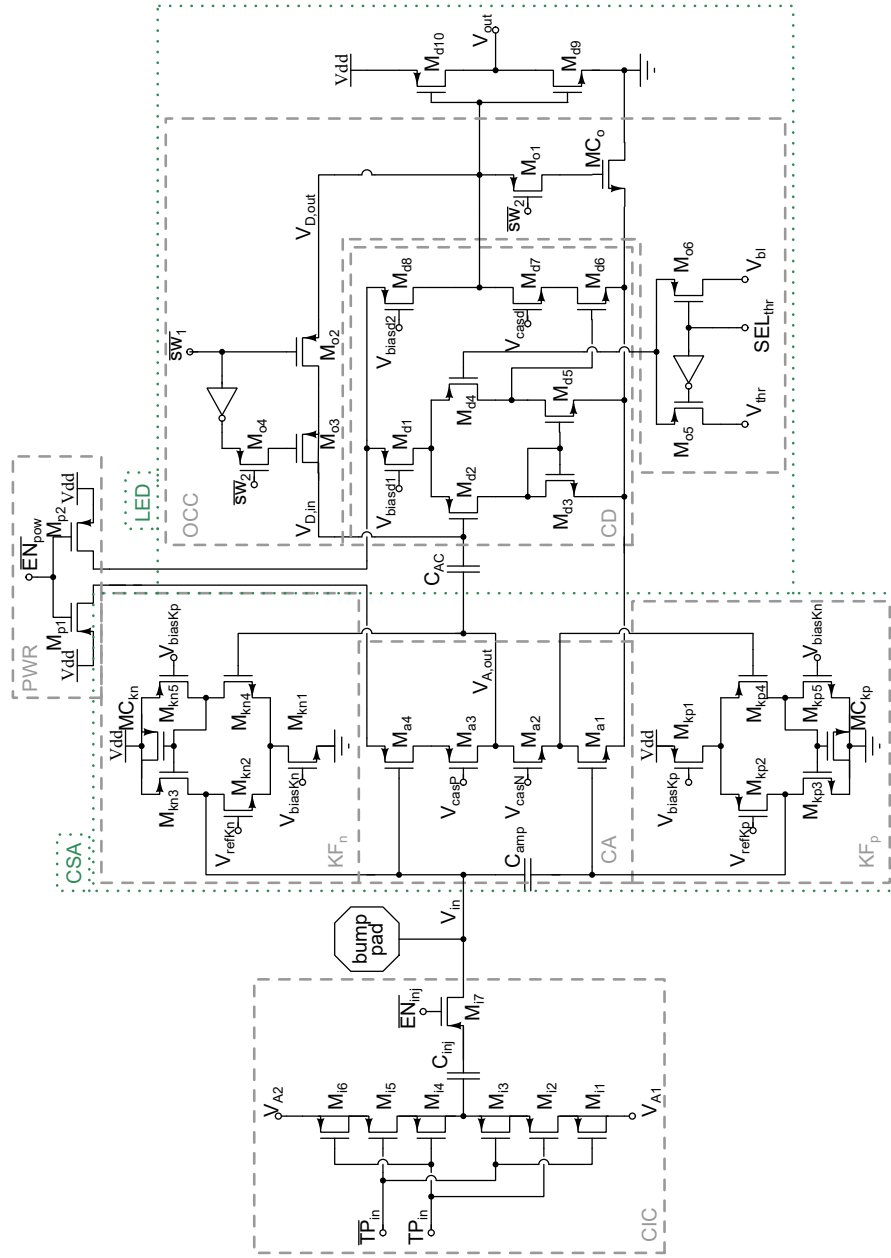
Figure 5.15: Transistor level schematic of the pixel electronics of an AFE channel. The relevant blocks described in the text are labelled as: Charge Injection Circuit (CIC), Charge Sensitive Amplifier (CSA), N-type Krummenacher Filter (KFN), P-type Krummenacher Filter (KFP), Core amplifier (CA), Leading Edge Discriminator (LED), Core Discriminator (CD) and Offset Compensation Circuit (OCC) and Power enable switches (PWR). M_{d2} and M_{d4} are thick oxide transistors.

Table 5.1: Transistor sizing of the Timespot1 analog front-end. The transistor labels are the same as those used in figure 5.15.

width/length $[\mu m/\mu m]$

| Charge Injection Circuit		Core Amplifier		Krummenacher Filter N		Krummenacher Filter P		Discriminator Core		Offset Correction Circuit		Power Enable Switches	
M_{i1}	4.00/0.03	M_{a1}	24.00/0.03	M_{kn1}	1.00/4.32	M_{kp1}	1.00/1.44	M_{d1}	0.50/3.00	M_{o1}	0.10/0.30	M_{p1}	16.00/0.30
M_{i2}	8.00/0.03	M_{a2}	2.40/0.03	M_{kn2}	0.60/0.60	M_{kp2}	3.00/0.48	M_{d2}	2.00/0.15	M_{o2}	0.10/0.30	M_{p2}	16.00/0.30
M_{i3}	4.00/0.03	M_{a3}	2.40/0.03	M_{kn3}	0.50/1.44	M_{kp3}	0.50/4.32	M_{d3}	0.20/1.20	M_{o3}	0.10/0.15		
M_{i4}	4.00/0.03	M_{a4}	24.00/0.03	M_{kn4}	0.60/0.60	M_{kp4}	3.00/0.48	M_{d4}	2.00/0.15	M_{o4}	0.10/0.15		
M_{i5}	8.00/0.03			M_{kn5}	0.50/1.44	M_{kp5}	0.50/4.32	M_{d5}	0.20/1.20	M_{o5}	8.00/0.03		
M_{i6}	4/0.03			MC_{kn}	28.8/0.51	MC_{kp}	28.8/0.51	M_{d6}	4.00/0.10	M_{o6}	8.00/0.03		
M_{i7}	8/0.03							M_{d7}	1.00/0.10	MC_{o}	15.96/0.30		
								M_{d8}	0.20/1.80				
								M_{d9}	0.50/0.10				
								M_{d10}	0.50/0.10				
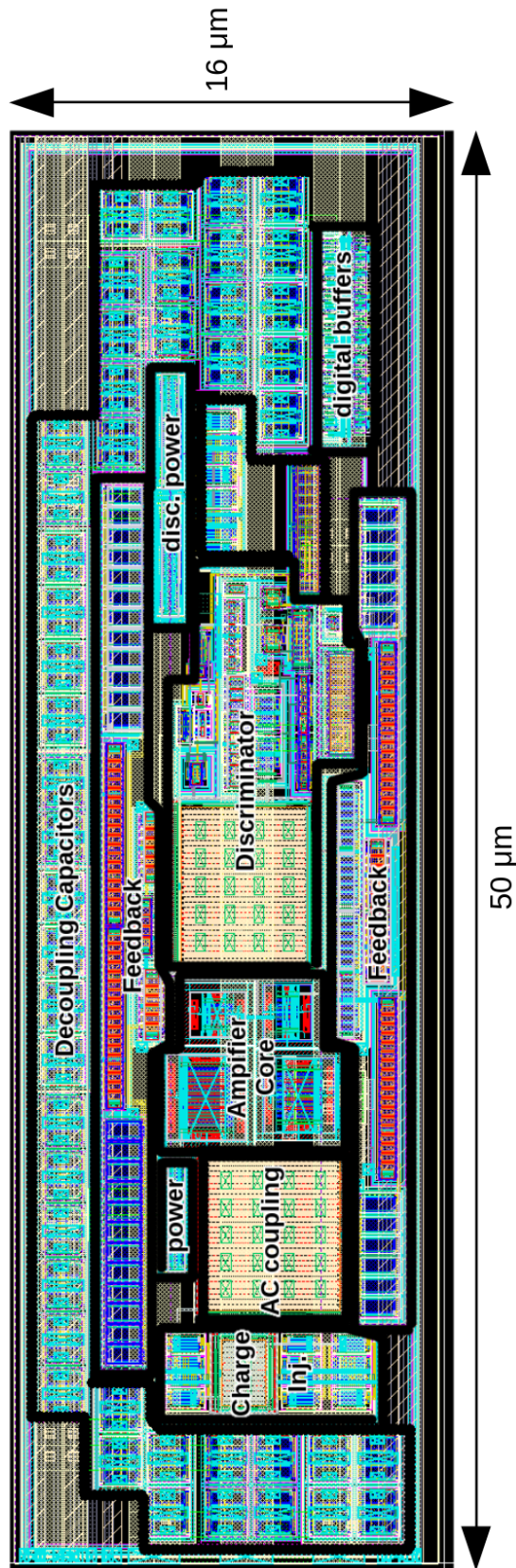
Figure 5.16: Layout of the Timespot1 AFE core channel. The relevant blocks are highlighted. The transistors level schematic is presented in figure 5.15. The connection with the TDC and the digital controls are on the bottom side of the cell. The total area is $800\,\mu m^2$.

Chapter 6

Timespot1: Analog Front-End Characterization

This chapter presents the characterization process and results on the Timespot1 analog front-end, which is described in chapter 5. Firstly, the measurement setup and methodology will be presented in section 6.1, this information is essential to understand the subsequent measurements.

Despite the possibility to couple this ASIC with its sensor, the tests were performed on the front-end chip alone, without hybridization with the sensor. This was due to the fact that the hybrid prototype is not ready for the time being. The hybrid measurements will be performed in the near future.

The characterization effort focused on the verification of the correct operation and performance of the AFE. Since there are 1024 channels inside a single ASIC and they are only accessible through TDC measurements, the AFE will be characterized statistically on many channels as part of a complex system. Therefore, the results are presented in section 6.2 in the form of one-dimensional or two-dimensional histograms.

This chapter presents its content assuming that the reader is familiar with the measurement concepts presented in 1.2.3 and with the AFE architecture presented in 5.

6.1 Setup and Method

This sections both describes the experimental setup and presents the AFE characterization methodology. Since the DUT can be any AFE channel, the experimental setup is constituted by the TDC, as the direct measurement instrument, and all the components DAQ chain used to read it out. The measurement method was adopted to reliably characterize the circuit from a timing point of view while differentiating the time fluctuations inherit in the TDC.
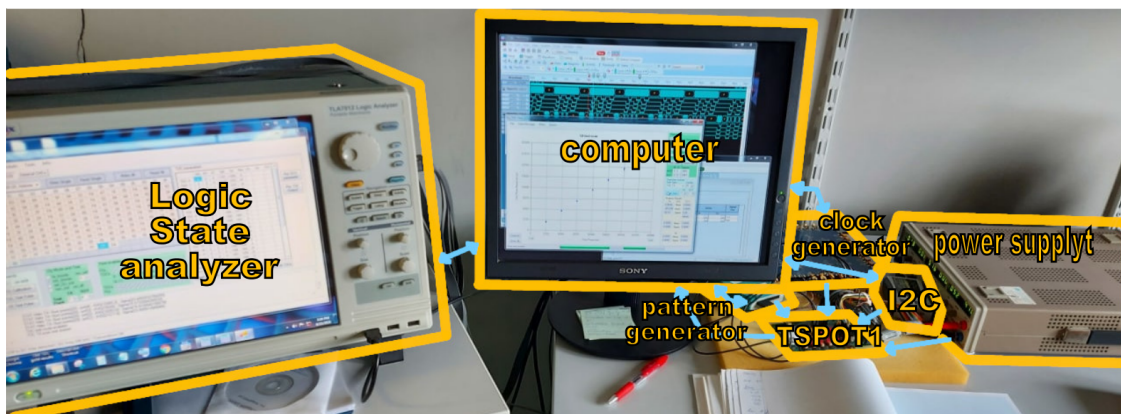
6.1.1 Setup

Figure 6.1: Photograph of the experimental setup used to test Timespot1. One of the three TSPOT1 board is connected to the DAQ system. The DAQ central device is a computer. Since these tests are all electrical tests, no optical instruments are required. The chip is mainly controlled and tested via a slow I^2C interface.

The Timespot1 measurement setup can be divided in 3 major layers: the on-chip electronics, the TSPOT1 PCB and the DAQ equipment. A photograph of the setup is shown in figure 6.1, whereas its schematic representation is shown in figure 6.15 at the end of the chapter.

The main ASIC blocks used to characterize the AFE are the analog periphery, the AFE digital controls and the TDC. The analog periphery contains the bias and reference cells which are used to set all the voltage levels required to regulate the channel power. This block also contains the DACs used to set the input charge and the OC baseline and threshold voltages. This part of the circuit is controlled via an I^2C interface [46], a dedicated I^2C peripheral is implemented at this level.

Another I^2C peripheral is implemented at the row level, controlling both the AFE digital controls and the TDCs. The AFE digital controls are responsible for the channel pulsing and for the OC operation and configuration. The peripheral is also used to issue the TDC calibration and self test. Although it is possible to read the data generated by the TDC via its main data transmission line, the I^2C was used instead. This decision was taken due to the fact that the FPGA-based readout chain was not at a sufficiently mature state during the time of this data taking. The main disadvantage of using the I^2C over the normal path is its maximum sustainable event-rate limitation of about 10 kHz per chip. In normal operating condition this will determine a huge data-loss, whereas in this case it does not represent a problem since each channel is pulsed and read on demand. Anyhow, the overall characterization process is drastically slowed down by the adoption of this interface.
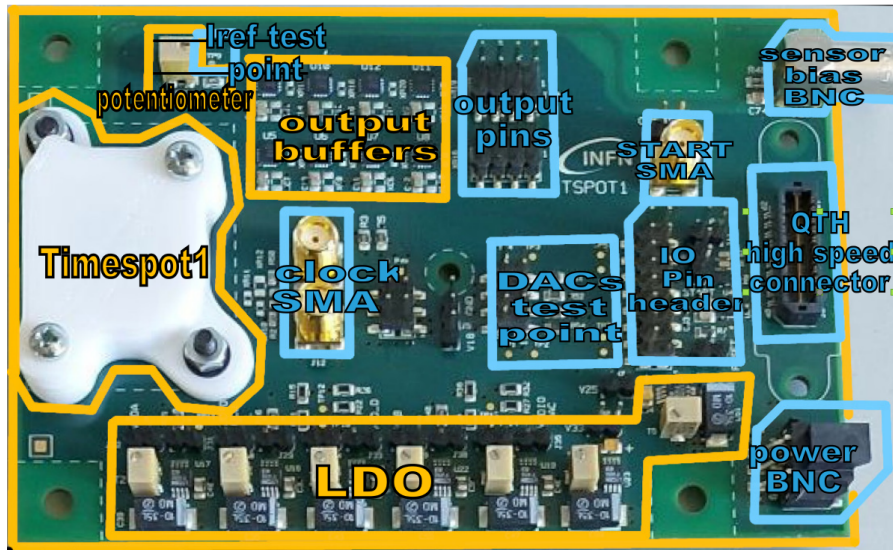
Figure 6.2: Photograph of the TSPOT1 PCB. The board size is $12\,\text{cm} \times 8\,\text{cm}$, and it is designed to fit inside the telescope demonstrator. Relevant blocks are highlighted in gold, whereas the connections are highlighted in blue. The Timespot1 ASIC is inside the white protection box. The IO pin header is the mostly used in this characterization.

The ASIC is mounted on the TSPOT1 PCB, which has been designed specifically as part of a small scale particle telescope. A photograph of this PCB is shown in figure 6.2. The TSPOT1 is conceived to have the maximum possible access to the ASIC I/O via an FPGA high-performance read-out board. In the target application, one FPGA can read out up to four TSPOT1 boards using a dedicated mezzanine and the on-board high-density high-speed QTH[95] connector. The PCB also allows controlling the ASIC and visualizing the signals by means of a pattern generator and a high-performance logic analyzer or oscilloscope using dedicated pin headers for connection. This was done mainly to verify the correct operation of the data transmission circuit. Finally, the last mode of operation is carried out via I^2C, which is the mode used in this characterization. The PCB mounts all the SN65 LVDS buffers [102] and headers used to rout these signals. From an analog point of view, four LTC2604 DACs [65] are mounted and can be used as backups in case of a malfunction of the internal DACs, and a potentiometer with a series known resistor is mounted for regulating the reference current I_{ref}. The PCB also mounts the LT3080 Low-DropOut regulator (LDO) [64] used to derive the required power domains from the system $5\,\text{V}$ one. Particular attention was paid to insulate the PCB for external electromagnetic interferences[26] with a Faraday cage.

The board is powered with a HAMEG-HM7942 tabletop low-noise power supply [40]. The $40\,\text{MHz}$ master clock is provided using a si5341-d-evb evaluation board
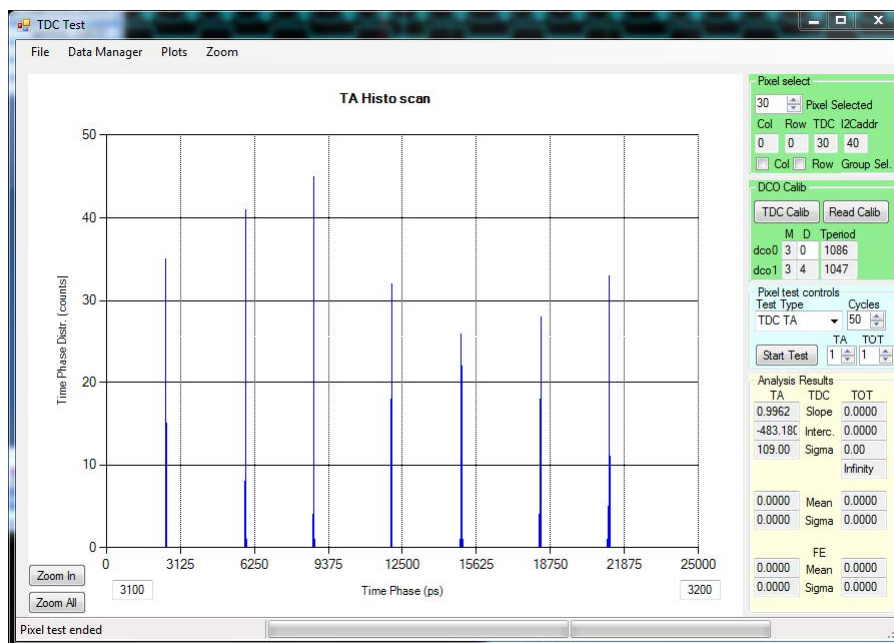
Figure 6.3: Example of a screen of the DAQ software used to test the Timespot1 front-end developed by INFN Cagliari. The screen shows a TA histogram for 7 different injection phases. The measurements were performed on the TDC.

[100] mounting a Si5341 clock generator[99]. This board is able to generate stable low-jitter clocks and phase locking them to an external source clock. The DAQ equipments is composed of an USB-TO-I2C ELITE I^2C interface module [112], a Moving Pixel PG3A pattern generator[85] and a Tektronix TLA7012 logic state analyzer[110]. The pattern generator is used to produce the START signal and the DACs controls. During this characterization, the logic state analyzer is only used to verify the correct operation of the transmission lines.

The entire DAQ is controlled using a computer via a dedicated GUI developed by INFN Cagliari using the C# programming language [17]. This tool allows to easily program and verify the content of the I^2C registers, including the TDC measurements. It is also possible to issue specific routines for characterization purposes, which are described in the next section. An example screen of this GUI is presented in figure 6.3.

6.1.2 Method

The method used to characterize the AFE is tightly connected to the fact that the DUT is part of a complex system. Since each AFE output is connected to its dedicated TDC, the only means of measuring the channel is using the TDC as an instrument. Each AFE channel can be individually pulsed using its charge injection
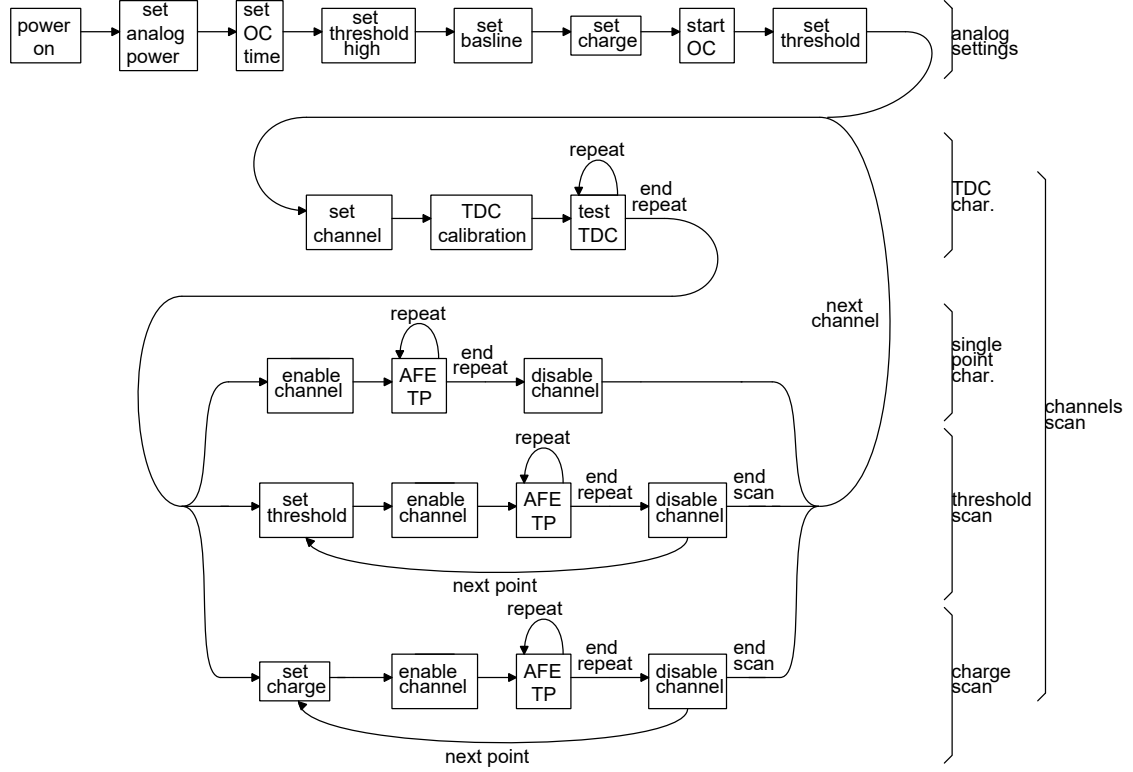
Figure 6.4: Flow chart representing the characterization procedure adopted to test the Timespot1 AFE. The procedure guaranties a reliable and reproducible analog setting while avoiding the injection of spurious signals on the TDC. The TDC is calibrated and characterized before use in order to evaluate its contribution to the time measurements fluctuations.

circuit. In response to this signal, the AFE will produce a digital pulse which is then measured by the TDC. Consequently, with a single measurement, it is only possible to measure the TA and ToT of a signal. In order to extract more information, the channel can be statistically investigated by repeatedly pulsing the same channel. In this way, the time resolution for a given condition can be evaluated as the standard deviation of TA measurement set ($\sigma_{\text{AFE+TDC}}$). However, this contribution accounts for both the AFE and TDC time fluctuations. The AFE component σ_{AFE} can be evaluated by removing the previously measured TDC contribution σ_{TDC}. This can be computed under the hypotheses that the two sources of variation are independent using:

$$\sigma_{\text{AFE}} = \begin{cases} \sqrt{\sigma_{\text{AFE+TDC}}^2 - \sigma_{\text{TDC}}^2} & \text{if } \sigma_{\text{AFE+TDC}} \geq \sigma_{\text{TDC}} \\ 0 & \text{if } \sigma_{\text{AFE+TDC}} < \sigma_{\text{TDC}} \end{cases} \tag{6.1}$$
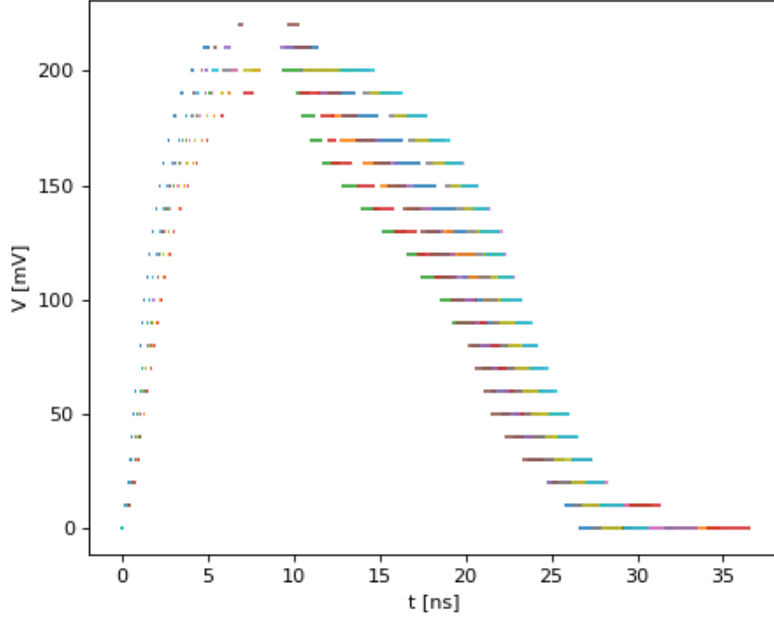
Figure 6.5: Example of a threshold scan reconstruction of signals of 32 channels of a Timespot1 double-row.

Contrary to what is expected, values of $\sigma_{\mathrm{AFE+TDC}}$ smaller than σ_{TDC} can be measured. This case can be explained as an effect of a too small statistic: the measurement sets were too small to distinguish this difference. This case was foreseen, and therefore the resulting value is set to zero for all the presented measurements. The same method is used for the ToT measurements. It must be noted that the TDCs differs from one other, this forces to characterize each TDC of the matrix and to couple this measurement to the related channel.

In order to characterize the channels, the system must be first reliably set to a well known state. It is remained that the parameter which can be varied in test phase are: the individual stages power consumption, the OC repetition frequency, the injected charge and the discriminator baseline and threshold. The change of these parameters may activate the AFE in unexpected ways and inject spurious signals into the TDC. This fact will inevitably produce spurious data and may also risk blocking the TDC. Luckily, it is possible to insulate the TDC from the AFE by acting on the EN_{ch} signal. Moreover, a low initial threshold may activate simultaneously the whole matrix, producing a high and unsustainable power demand that will make the system unstable. This situation can be avoided by setting the threshold as high as possible at the start-up, and by lowering its value only after the OC has been activated.

The characterization procedures must account for these aspects as well as for the TDCs characterization. In particular it is important to characterize the TDC
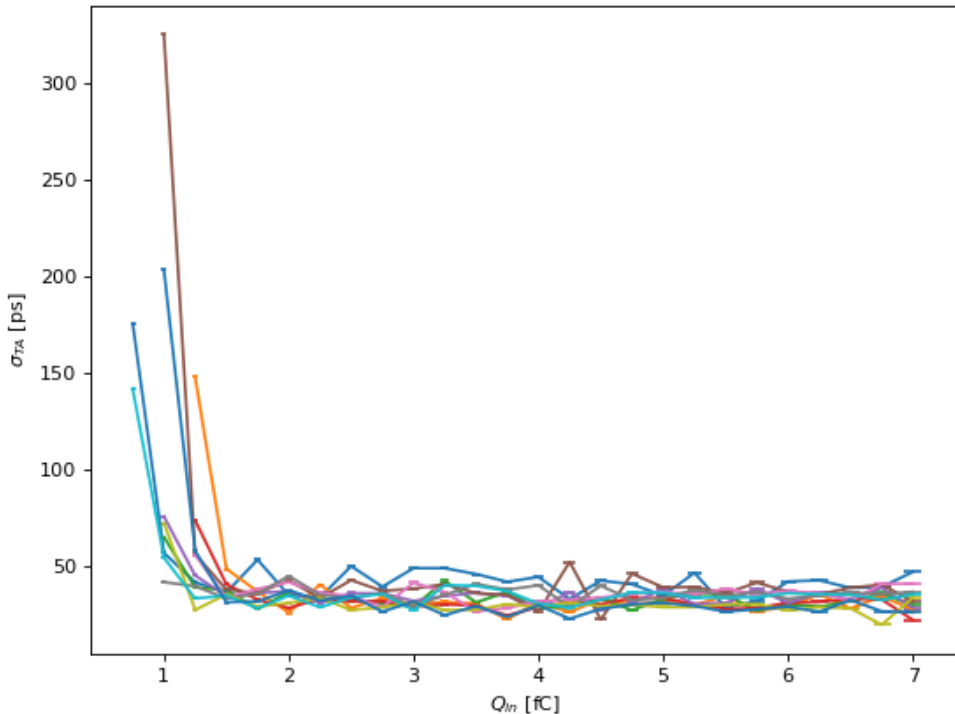
Figure 6.6: Example of a σ_{TA} charge scan of 32 channels of a Timespot1 double-row.

immediately before measuring the AFE in order to avoid the risks of a change in the test conditions (such as the TDC calibration point). Figure 6.4 shows a flow-chart of the measurement procedures adopted to test various aspects of the circuit. These procedures are implemented in software and can be performed automatically on the full matrix. In particular it is important to perform: a signal reconstruction via threshold scan (an example is shown in figure 6.5), and charge scan (an example is shown in figure 6.6) to characterize the timing performance in the input range. Each charge reconstruction measures the four key parameters of the front-end: the TA and ToT average values and standard deviations.

The results presented in the next section are extracted from a data-analysis performed on measurement sets extracted using these procedures. The data analysis software has been implemented in Python3 using the *pandas* [80] library to produce a per-channel detailed database. The database is organized on a per channel bases. The pixel information contained in each entry is: the TDC characterization, the AFE threshold scan reconstruction and the charge reconstruction. Every data set is also paired with the set of all the programmable conditions.

When not explicitly said, the measurement has been taken in the normal condition defined as: average power consumption per channel of $13\,\mu W$, $2\,fC$ of input charge, a baseline value of $450\,mV$, a threshold value of $30\,mV$ above the baseline and an OC period of $125\,\mu s$. The average power consumption per channel is derived from a direct measurement of the power consumption of the analog matrix divided by the number of channels. This value will therefore take into account the contribution of the analog periphery. The input capacitance is extracted from simulations and its value is around $45\,fF$. This capacitance is the sum of: the bump capacitance, the charge injection capacitance and all the parasitic capacitance from the input nodes to ground.

6.2   Measurements

This section focuses on the timing characterization of the Timespot1 AFE core channel. The correct operation of the functionalities integrated on the analog periphery has been verified to work properly. These tests include: the power settings, the DACs configuration and the channel digital controls. A detailed characterization of the AFE in relation to these parameters has not been performed yet. Regarding these settings, this section presents measurements in standard condition.

During this analysis an issue with the OC was found which prevents the channels to behave consistently for certain baseline conditions. For this reason, this OC issue is the first presented in the next subsection, whereas the timing characterization is presented afterward.

6.2.1   Offset Correction Operation

The main consequence of the OC issue is that the procedure has proven to be inadequate to compensate baseline values to low voltages. Figure 6.7 shows the baseline positions spread across half matrix. The analysis was performed with two desired baseline values V_{bl}^*. A low-baseline was set to $100\,mV$ whereas a high one was set to $450\,mV$. It can be seen that most channels are not able to reach the low-baseline, whereas all of them are able to reach the high one. It must be reminded that the OC circuit correct operation was verified in the previous prototype (see 4.2.2).

This problem is probably caused by the fact that the voltage to be compensated is higher than expected. Moreover, the OC time assigned by design is insufficient to discharge the capacitor to the desired level. Unfortunately this value is not programmable in the current iteration of the Timespot ASIC. A possible explanation for this behavior must be investigated on the basis of the differences on the discriminator implementation between the two prototypes. As explained in 5.1.1
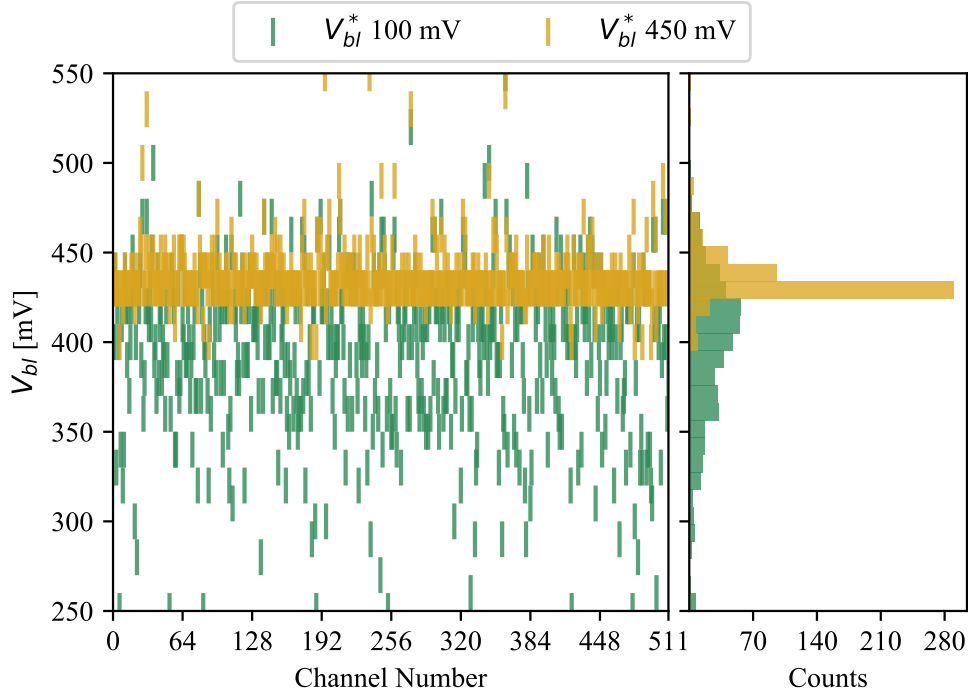
Figure 6.7: Measured baseline distribution V_{bl} across 512 channels. The measurement was repeated for two desired baselines V_{bl}^*: 100 mV and 450 mV. It can be observed that the OC fails to adequately compensate the channels for low baseline values. Rising V_{bl}^* to 450 mV will move the channel population to the desired baseline.

the only change to the discriminator design is in the OC switching circuit. In the current implementation, the circuit has been changed to use p-type switches since they have proven to be more radiation resistant. This change may have risen the discriminator pre-OC DC voltage point, making it more difficult to compensate. The magnitude of the problem varies widely from channel to channel, but it is always possible to perform the OC properly by setting high baseline values. This behavior of the discriminator DC input operating point was not found in simulation. This can be related to the fact that the SPICE simulations hardly simulates high impedance conditions such as this one.

However, setting a high baseline comes with its shortcomings. The side effect of imposing a high-baseline is related to the fact that the threshold value V_{thr} must be set to an even higher value in order to measure signals with positive polarity. In this range, the p-type input discriminator will not operate at its optimal common mode voltage, thus limiting its bandwidth. This behavior is showed in figure 6.8 in which σ_{AFE} is correlated to the measured baseline V_{bl}.

A good indicator that this hypothesis describes the observed situation is the fact
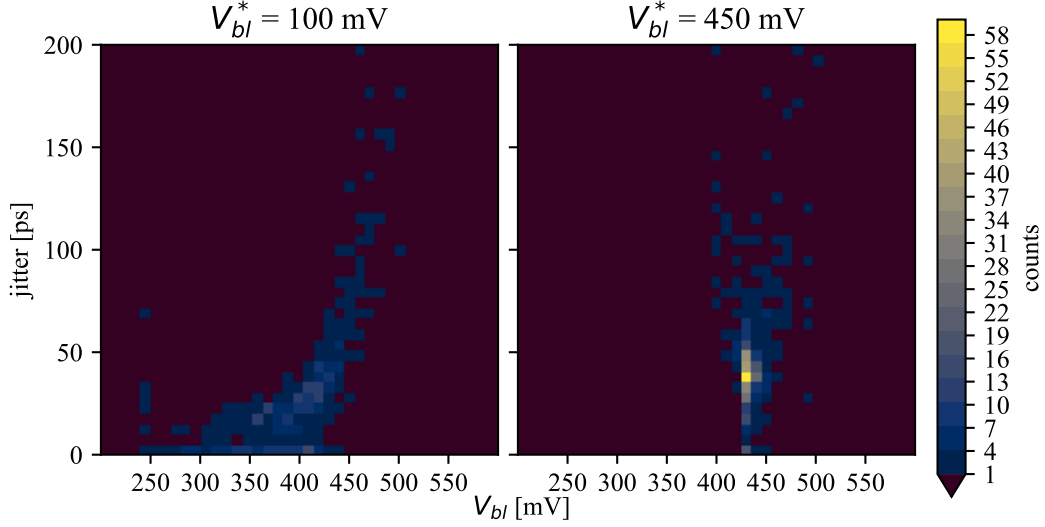
Figure 6.8: Correlation of σ_{AFE} to V_{bl} for two V_{bl}^* values. In the left plot the OC is not working properly. It can be observed a correlation between the AFE resolution and the baseline position. Measurements at low baseline values (on the left) suggest that the CSA intrinsic resolution is below the one of the TDC (20 ps). When the OC is working (right plot) the performances are in line with what is observed at the corresponding baseline in the left plot.

that when the OC is not working properly, a correlation between V_{bl} and σ_{AFE} can be observed. When the OC is set to high values, in fact, the compensated channels will feature the same σ_{AFE} which was found with the corresponding baselines in the previous case. These two facts suggest that the discriminator core is bandwidth limited for high input voltages. It is stressed that the two cases represent the same channels in two different discriminator conditions and, therefore, a dependence of this behavior on the CSA operation must be excluded.

In any case, the AFE time resolution has been evaluated for a 2 fC signal when the OC is working properly. This resulted in an average of 43 ps for σ_{AFE} with a power consumption around 13 μW per channel. Figure 6.9 shows the histogram derived from this measurement. This evaluation represents a pessimistic view of the CSA intrinsic jitter performance, but it is however a good indicator of the matrix performance in this prototype.

However, if the hypothesis of the discriminator malfunction holds true, the CSA performance can be evaluated separately. By considering the low-baseline channels, it can be observed that the CSA is able to reach jitter levels under 20 ps. This limit is imposed by the TDC resolution: the CSA jitter performance is therefore masked by the TDC. Additionally, this value is the result of the own CSA jitter plus the OC contribution, which can be decomposed using (3.40). In this way, the CSA

performance can be really close to what is observed in simulation.
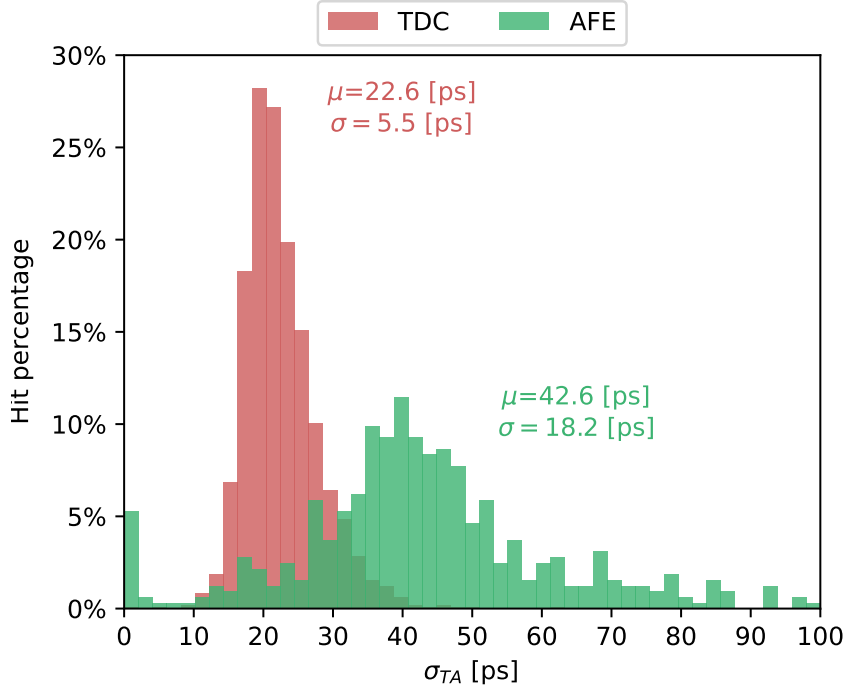
6.2.2 Timing Performance

Figure 6.9: Red histograms: $\sigma_{TA,TDC}$ on 100 repeated DTP across 1024 channels and 7 phases. Green histogram: $\sigma_{TA,AFE}$ on 100 repeated ATP across 512 channels for a 2 fC input signal (MIP), the TDC contribution has been square subtracted according to (6.1).

The data collected for this characterization come from 512 channels of one half-matrix connected to one pair of analog and digital services columns. The collective single point $\sigma_{TA,AFE}$ measurements on these channels in standard condition (please note: $V_{bl}^* = 450\,\mathrm{mV}$) are shown in figure 6.9. This figure also presents the histogram of the respective $\sigma_{TA,TDC}$, these are the values used to obtain $\sigma_{TA,AFE}$ according to (6.1). This histogram presents a real case scenario of the intrinsic average resolution of the Timespot1 AFE.

The most important characterization of the AFE is its performance for different input charges. The parameter which can be measured are the TA and the ToT. Therefore the next plots present their average values and their standard deviations. Each one of the next plots is a 2D-histogram collecting the data from the 512 channels.

The plots shown in the next pages presents:

- Figure 6.10: TA_{AFE} versus Q_{in}. This relation is important to quantify the time-walk effect and to calibrate the system propagation delay (which represents a systematic in the timing measurement).

- Figure 6.11: $\sigma_{TA,\text{AFE}}$ versus Q_{in}. This relation indicates the core resolution of the AFE. This value is the upper limit of what is achievable in a system featuring this architecture.

- Figure 6.12: ToT_{AFE} versus Q_{in}. This relation shows the quality of the charge-to-time conversion capability of the CSA.

- Figure 6.13: $\sigma_{ToT,\text{AFE}}$ versus Q_{in}. This parameter, along with the previous one, is useful to quantify the resolution of the charge-to-time conversion. This resolution, will also define the binning of the ToT-correction, and therefore its contribution to the total resolution.

- Figure 6.14: TA_{AFE} versus ToT_{AFE}. This is the relation used to apply the ToT-correction.

All the data were acquired with both the previously described OC conditions. The specific information and discussion on each plot are presented on the respective caption. Overall, with the OC working properly, less than $2\,\%$ of the channel does not operate adequately.
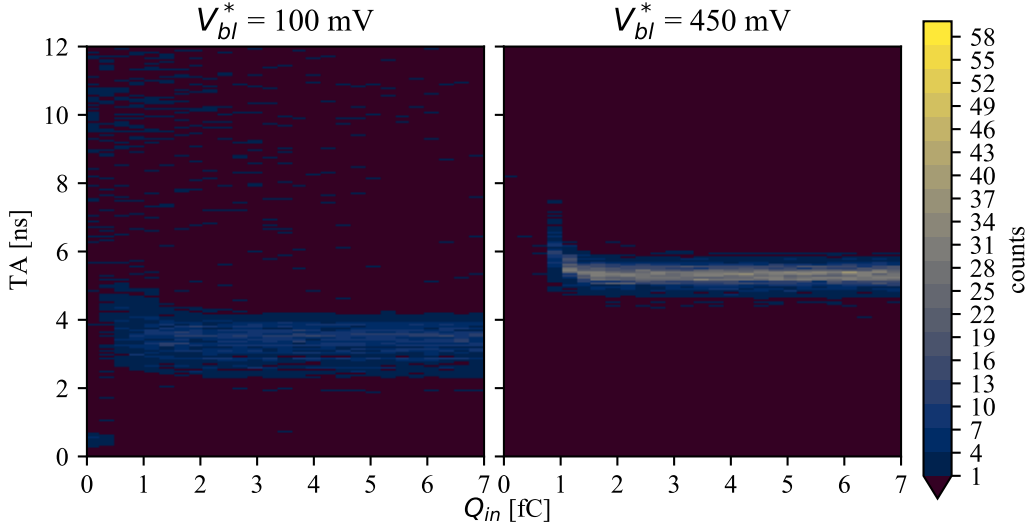
Figure 6.10: TA_{AFE} values as function of the input charge. The time-walk effect can be observed in this plot. As expected, the channels behave more consistently when the OC is working properly. Input charges lower than $1\,\text{fC}$ can be detected, which are well below what the sensor is expected to produce.
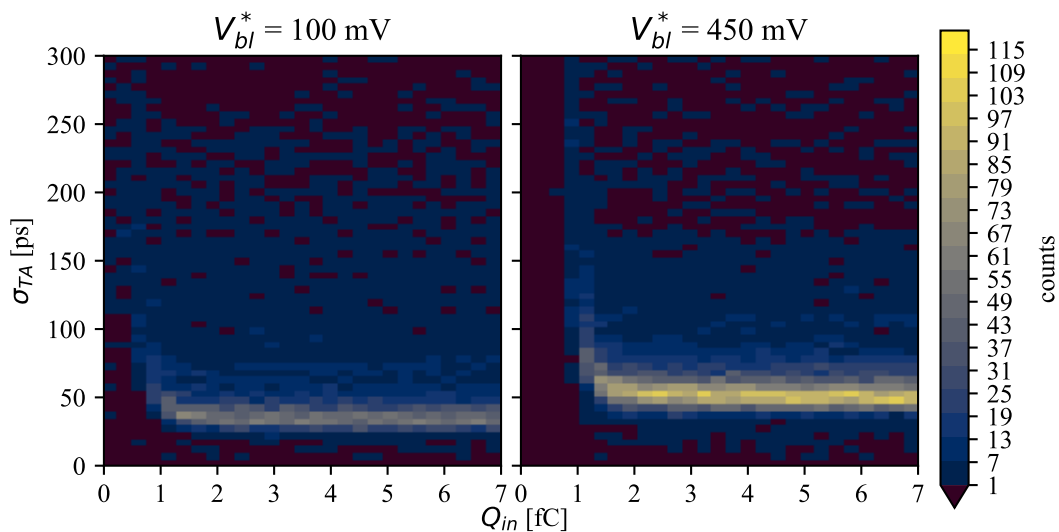
Figure 6.11: $\sigma_{\mathrm{TA,AFE}}$ values as function of the input charge. As expected, the time resolution improves with larger input charges. Consistently with what is shown in figure6.8, the jitter performance is improved for low baselines, but the per-channel variation is improved when the OC is working properly. Even in the second case, the jitter performance is more than double compared to Timespot0.
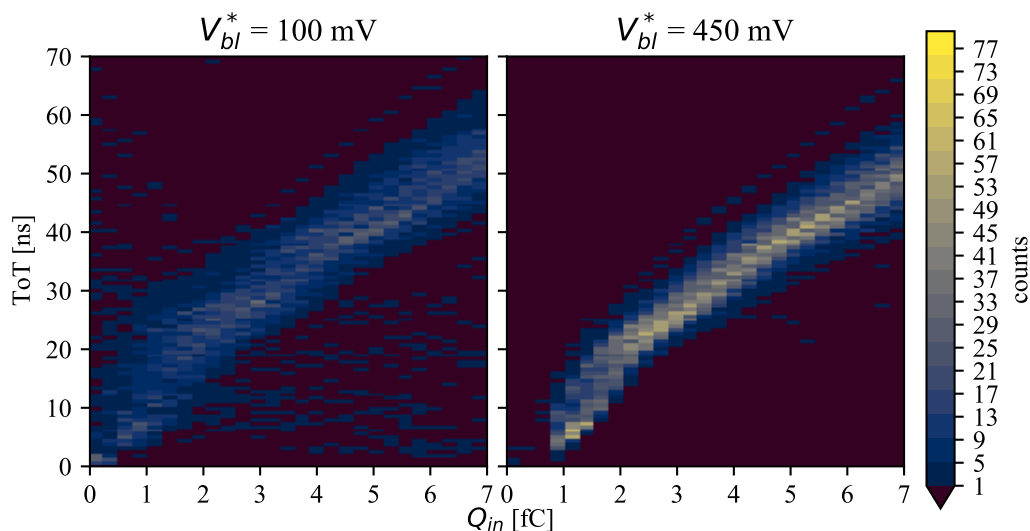

Figure 6.12: ToT_{AFE} as function of the input charge. The ToT shows a good linearity. Consistently with what it is shown in figure 6.8, channel variation improves when the OC is working properly. The non-linearity at low charge values is due to the relatively low threshold.

133
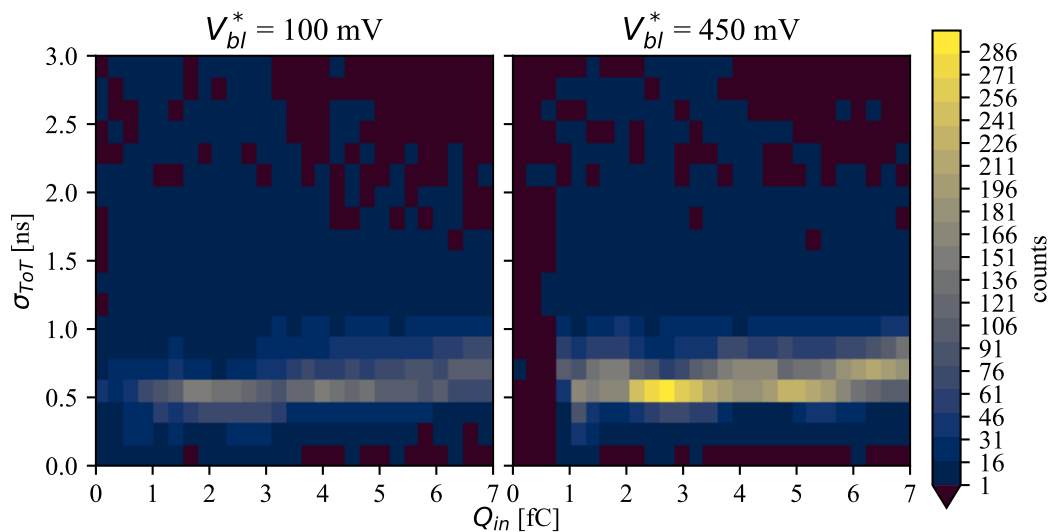
Figure 6.13: σ_{ToT} values of the Timespot1 AFE as function of the input charge. The trend shows a small dependency on the input charge. Overall the values are draftily improved compared to the previous version of about a factor two.
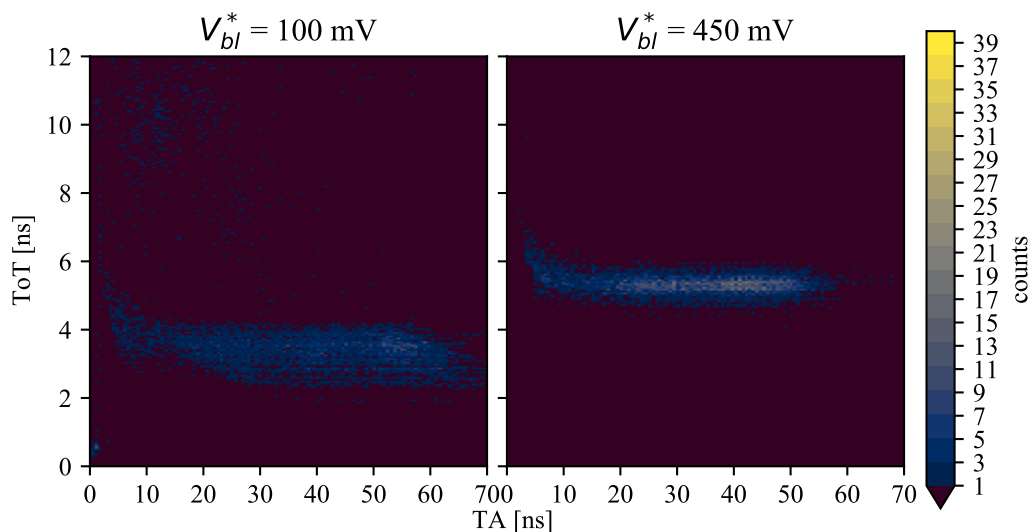

Figure 6.14: TA versus ToT correlation. This relation is used to perform a ToT-correction of the time-walk effect. Consistently with what is shown in figure 6.8, channel variation improves when the OC is working properly. In any case, the correction must be applied on a per-channel basis, making both cases correctable. However, it must be said that the low-baseline condition features more dead channels due to the fact that their baseline is above threshold or to low.
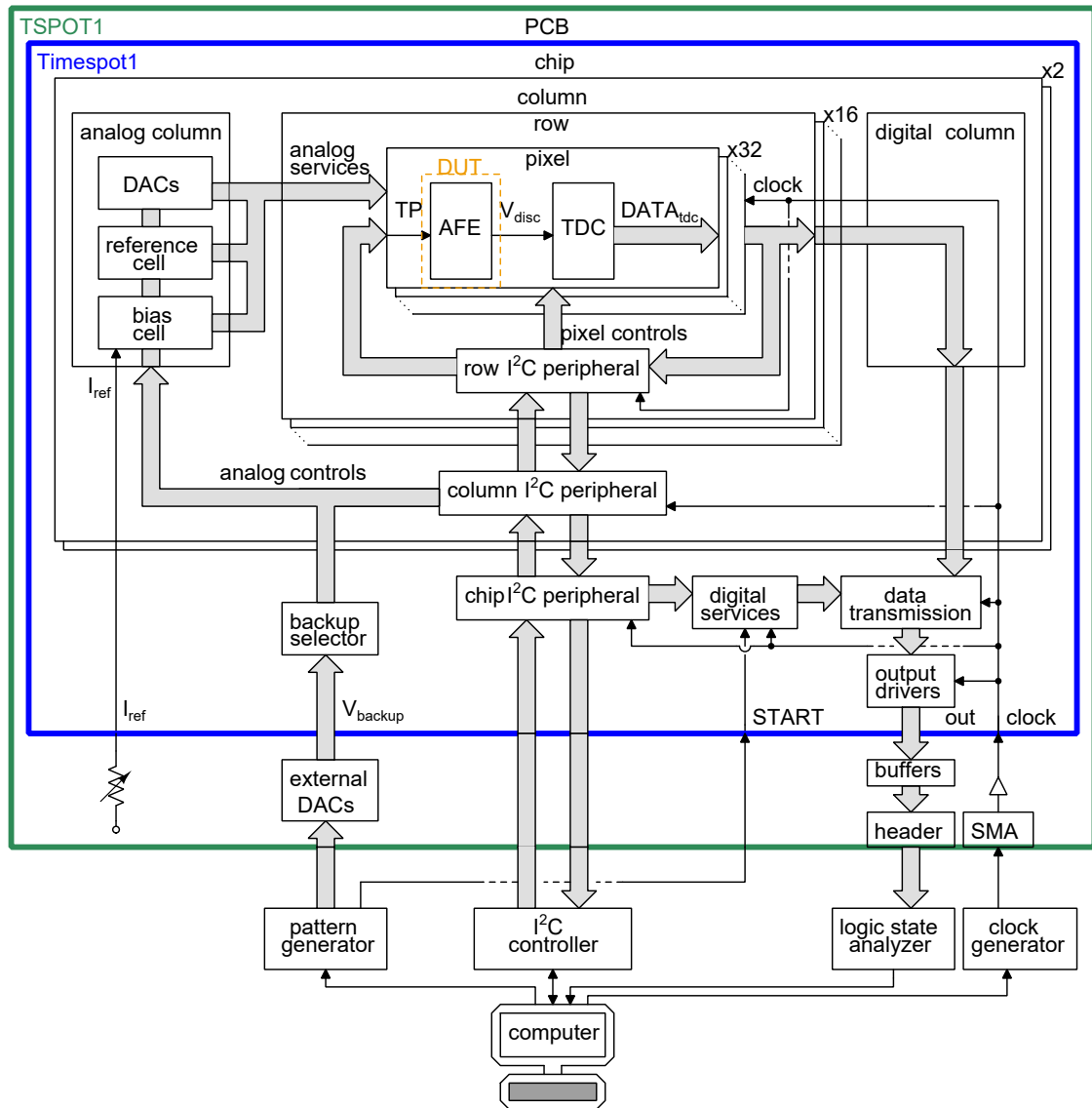
Figure 6.15: Schematic representation of the experimental setup. Power and ground nets are omitted for sake of clarity. The scheme distinguishes the major three layers of the setup: the on-chip electronics, the TSPOT1 PCB and the external DAQ. The main mean of interaction with a single AFE channel is the I^2C interface: the scheme illustrates its structure. The analog setups are controlled via the column peripheral: they set the values for all the 512 channels of one half-matrix. The row peripheral is used to pulse the individual channels and control the OC procedure. Although it is possible to read the output data through the main LVDS outputs, for this test the I^2C was used for the read-out.
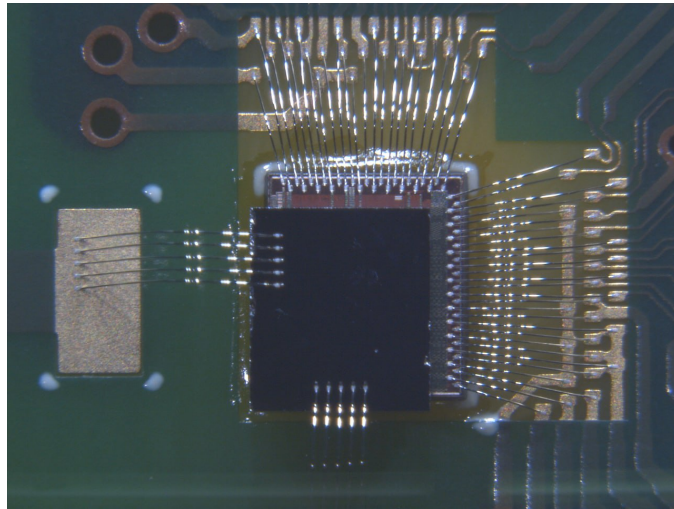
Conclusions

Figure 6.16: Timespot1 ASIC hybridized with the TimeSPOT 3D-Silicon sensor. The sensor chip can be seen flipped on top. Its bias is provided though the wire-bond on the bottom left corner. Timespot1 can be seen in the top right corner, with its wire-bond connections.

The work described in this thesis has made possible to develop the Timespot1 ASIC. This ASIC has been developed in 2020 in a commercial 28 nm CMOS technology. Timespot1 integrates 1024 channels in a 32×32 matrix of 55 µm × 55 µm pixels. Each channel is able to perform a time measurement with a resolution of 43 ps, at a signal rate of 3 MHz, with a power consumption per unit area under 1.3 W/cm².

This achievement is in line with the requirements dictated by the foreseen future upgrades of HEP experiments and it represents a new time resolution record for pixel ASICs with comparable spatial resolution. It also surpasses the initial time resolution goals of 100 ps of the TimeSPOT project. However, the TimeSPOT 3D-Silicon sensor has proven to achieve an intrinsic resolution of 20 ps or better. Therefore, more progress can be made from the point of view of front-end ASIC. The TDC has already reached a 24 ps resolution, while the front-end currently
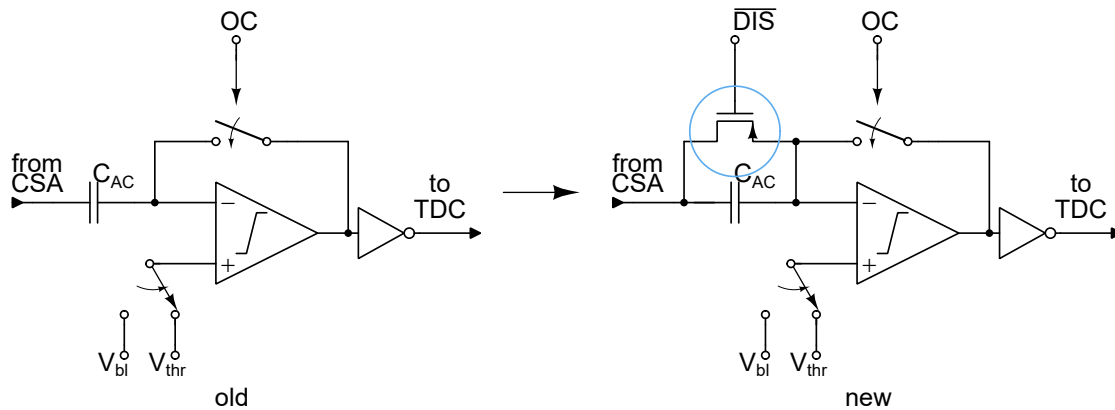
Figure 6.17: Purposed correction for the offset compensation circuit. The offset compensation issue can be corrected by inserting a single transistor which acts as a discharging switch. The compensation procedure must be changed to include a discharge phase before the actual compensation. Having the baseline set close to the CSA output level (around half of the power supply voltage) will improve the compensation quality (as described in 3.2.5). With the insertion of this transistor it is also possible to DC couple the CSA with the discriminator for characterization purposes.

represents the limiting factor on the measurement chain.

However, tests have made it clear that the current architecture of the analog front-end is able to reach resolutions well beyond 43 ps. During the tests, an issue with the discrete time offset compensation was found. This issue prevents the usage of the offset compensation while achieving the optimal resolution. During the characterization process, it was possible to quantify the channel resolution in the optimal condition, resulting in a time resolution of 20 ps or better. The measurement of this resolution was ultimately limited by the TDC sensitivity. It can be said with confidence that this resolution can be achieved with only a small tweak on the current architecture, without spoiling the overall performance. Figure 6.17 presents a simple correction to the scheme that will likely prevent this problem.

From the specific point of view of the analog front-end, both the two prototypes were overall successful. The first one was able to achieve a 60 ps time resolution with a power consumption of 11 µW per channel. The discrete-time offset corrected discriminator has proven to be operating properly with a 0.08 ‰ channel dead time. The second prototype was able to reach an average of time resolution of 43 ps across 1024 channels, with a comparable power consumption of 13 µW per channel. However, when the channel is operated in optimal condition, its resolution is better than 20 ps. This will determine a performance increase in respect to the previous prototype of a factor 3. This aspect also demonstrates the quality of the new very front-end architecture. In this condition, the analog-front end is able to adequately

measure the performance of the 3D-Silicon sensors.

In this regard, all the results mentioned above are in fact extracted from ASIC self tests. The ASIC has been recently hybridized with 3D-Sensor in order to be tested with real signals generated from radiation interactions. This hybrid prototype is shown in figure 6.16. This test is not the only future development for the project. A radiation hardness test on Timespot1 is also foreseen in the near future. Additionally, by the end of 2022, the full demonstrator will be assembled and tested in a test-beam.

Finally, a new version of the Timespot ASIC family is also in development. This new chip will be in principle an updated version of Timespot1 with errors corrected and an improved clock distribution network. This prototype will be the ideal opportunity to implement solutions for the offset correction issue. There is also the possibility to increase the ASIC size in the next prototype, integrating a larger matrix. A last possible update in the next version is the usage of TSV to integrate multiple ASICs on the same surface without dead areas.

Acronyms

ADC Analog to Digital Converter

AFE Analog Front-End

APD Avalanche Photo Diode

ASIC Application Specific Integrated Circuit

CCD Charge-Coupled Device

CDF Cumulative Density Function

CERN Conseil Européen pour la Recherche Nucléaire

CFD Constant Fraction Discriminator

CML Current Mode Logic

CMOS Complementary Metal–Oxide–Semiconductor

CSA Charge Sensitive Amplifier

DAC Digital to Analog Converter

DAQ Data Acquisition

DCO Digital Controlled Oscillator

DDR Double Data Rate

DLL Delay Locked Loop

DNL Differential Non-Linearity

DSP Digital Signal Processing

DUT Device Under Test

EoC End of Column

FIFO First In First Out

FPGA Field Programmable Gate Array

GUI Graphical User Interface

HEP High Energy Physics

LDO Low-DropOut regulator

LGAD Low Gain Avalanche photo-Diode

LHC Large Hadron Collider

LSB Least Significant Bit

LVDS Low-Voltage Differential Signaling

MIP Minimum Ionizing Particle

PCB Printed Circuit Board

PDF Probabilty Density Function

PLL Phased Locked Loop

PSD Power Spectral Density

PSM Phase Shifting Mask

RET Resolution Enhancement Techniques

RMS Root Measn Square

ROT Read-Out Tree

S/H Sample and Hold

SEM Scanning Electron Microscope

SLVS Scalable Low-Voltage Signaling

SMA SubMiniature version A

SNR Signal to Noise Ratio

SPICE Simulation Program with Integrated Circuit Emphasis

TA Time of Arrival

TDC Time to Digital Converter

TIA Trans Impedance Amplifier

TimeSPOT Time and SPace real-time Operating Tracker

TMR Triple Modular Redundancy

ToT Time over Threshold

TSV Through Silicon Via

TVC Time to Voltage Converter

UART Universal Asynchronous Receiver-Transmitter

VCO Voltage Controlled Oscillator

Bibliography

[1] C. Agapopoulou et al. "ALTIROC 1, a 25 ps time resolution ASIC for the ATLAS High Granularity Timing Detector". In: *2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. 2020, pp. 1–4. DOI: 10.1109/NSS/MIC42677.2020.9507972.

[2] K. Mistry et al. "45nm Logic Technology with High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging". In: *IEEE* (Dec. 2007).

[3] T. Ghani et al. "A 90nm High Volume Manufacturing Logic Technology Featuring Novel 45nm Gate Length Strained Silicon CMOS Transistors". In: *IEEE IEDM* (Dec. 2003).

[4] S.R. Amendolia et al. "A multi-electrode silicon detector for high energy experiments". In: *Nuclear Instruments and Methods* 176.3 (1980), pp. 457–460. ISSN: 0029-554X. DOI: https://doi.org/10.1016/0029-554X(80)90368-7. URL: https://www.sciencedirect.com/science/article/pii/0029554X80903687.

[5] L. Anderlini et al. "Intrinsic time resolution of 3D-trench silicon pixels for charged particle detection". In: *Journal of Instrumentation* 15.09 (Sept. 2020), P09029–P09029. DOI: 10.1088/1748-0221/15/09/p09029. URL: https://doi.org/10.1088%2F1748-0221%2F15%2F09%2Fp09029.

[6] Lucio Anderlini et al. "Fabrication and Characterisation of 3D Diamond Pixel Detectors With Timing Capabilities". In: *Frontiers in Physics* 8 (Nov. 2020). DOI: 10.3389/fphy.2020.589844.

[7] G. Apollinari et al. *High Luminosity Large Hadron Collider HL-LHC*. Yellow Report. CERN, May 2015.

[8] R. Ballabriga, M. Campbell, and X. Llopart. "An introduction to the Medipix family ASICs". In: *Radiation Measurements* 136 (2020), p. 106271. ISSN: 1350-4487. DOI: https://doi.org/10.1016/j.radmeas.2020.106271. URL: https://www.sciencedirect.com/science/article/pii/S1350448720300354.

[9] Rafael Ballabriga et al. "Imaging by single quantum processing: large pixels with brains or attopixels without?" In: July 2019.

[10] L Bäni et al. "A study of the radiation tolerance of poly-crystalline and single-crystalline CVD diamond to 800 MeV and 24 GeV protons". In: *Journal of Physics D: Applied Physics* 52.46 (Aug. 2019), p. 465103. DOI: 10.1088/1361-6463/ab37c6. URL: https://doi.org/10.1088/1361-6463/ab37c6.

[11] Massimo Barbaro et al. "A Pixel Read-Out Front-End in 28 nm CMOS with Time and Space Resolution". In: *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. 2019, pp. 1–4. DOI: 10.1109/NSS/MIC42101.2019.9059838.

[12] Werner Beusch et al. *R&D proposal: development of hybrid and monolithic silicon micropattern detectors*. Tech. rep. Geneva: CERN, 1990. URL: https://cds.cern.ch/record/292598.

[13] David M Binkley. "Tradeoffs and optimization in analog CMOS design". In: *2007 14th International Conference on Mixed Design of Integrated Circuits and Systems*. IEEE. 2007, pp. 47–60.

[14] Maurizio Boscardin et al. "Advances in 3D Sensor Technology by Using Stepper Lithography". In: *Frontiers in Physics* 8 (Jan. 2021). DOI: 10.3389/fphy.2020.625275.

[15] W. S. Boyle and G. E. Smith. "Charge coupled semiconductor devices". In: *The Bell System Technical Journal* 49.4 (1970), pp. 587–593. DOI: 10.1002/j.1538-7305.1970.tb01790.x.

[16] Justus Braach et al. "Performance of the FASTPIX Sub-Nanosecond CMOS Pixel Sensor Demonstrator". In: *Instruments* 6.1 (2022). ISSN: 2410-390X. DOI: 10.3390/instruments6010013. URL: https://www.mdpi.com/2410-390X/6/1/13.

[17] *C# programming guide*. Microsoft Corporation. Mar. 2022. URL: %5Curl%7Bhttps://docs.microsoft.com/en-us/dotnet/csharp/programming-guide/%7D.

[18] Sandro Cadeddu et al. "A 28-nm CMOS pixel read-out ASIC for real-time tracking with time resolution below 20 ps". In: *2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. 2020, pp. 1–5. DOI: 10.1109/NSS/MIC42677.2020.9507912.

[19] Sandro Cadeddu et al. "A Time-to-Digital Converter Based on a Digitally Controlled Oscillator". In: *IEEE Transactions on Nuclear Science* 64.8 (2017), pp. 2441–2448. DOI: 10.1109/TNS.2017.2726822.

[20] Chun-Chi Chen, Shih-Hao Lin, and Chorng-Sii Hwang. "An Area-Efficient CMOS Time-to-Digital Converter Based on a Pulse-Shrinking Scheme". In: *IEEE Transactions on Circuits and Systems II: Express Briefs* 61.3 (2014), pp. 163–167. DOI: 10.1109/TCSII.2013.2296192.

[21] Zeng Cheng, M. Jamal Deen, and Hao Peng. "A Low-Power Gateable Vernier Ring Oscillator Time-to-Digital Converter for Biomedical Imaging Applications". In: *IEEE Transactions on Biomedical Circuits and Systems* 10.2 (2016), pp. 445–454. DOI: `10.1109/TBCAS.2015.2434957`.

[22] Cern Collaboration et al. "Development of a pixel readout chip compatible with large area coverage". In: *Nuclear Instruments and Methods in Physics Research A* 342 (Feb. 1994), pp. 52–58. DOI: `10.1016/0168-9002(94)91410-9`.

[23] Paula Collins et al. "Microchannel Cooling for the LHCb VELO Upgrade I". In: (Dec. 2021). arXiv: `2112.12763 [physics.ins-det]`.

[24] *Collision events recorded by CMS in 2016*. `https://cds.cern.ch/record/2241144`.

[25] Alexander Dierlamm. "The CMS Outer Tracker Upgrade for the HL-LHC". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 924 (2019). 11th International Hiroshima Symposium on Development and Application of Semiconductor Tracking Detectors, pp. 256–261. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/j.nima.2018.09.144`. URL: `https://www.sciencedirect.com/science/article/pii/S0168900218313019`.

[26] D. C. Smith Consultants Douglas C. Smith. "A New Type of Furniture ESD and Its Implications". In: *EOS/ESD Symposium Proceedings*. ESD Association. 1993, pp. 3–7.

[27] C. C. Enz, F. Krummenacher, and E. A. Vittoz. "An Analytical MOS Transistor Model Valid in All Regions of Operation and Dedicated to Low-Voltage and Low-Current Applications". In: *Analog Integrated Circuits and Signal Processing Journal* 8 (1995). special issue of the Analog Integrated Circuits and Signal Processing Journal on Low-Voltage and Low-Power Design, pp. 83–114. DOI: `10.1007/BF01239381`. URL: `http://infoscience.epfl.ch/record/149574`.

[28] Christian Fabjan and Thomas Ludlam. "Calorimetry in High-Energy Physics". In: *Annual Review of Nuclear and Particle Science* 32 (Nov. 2003). DOI: `10.1146/annurev.ns.32.120182.002003`.

[29] *FAST2: a new family of front-end ASICs to read out thin Ultra-Fast Silicon detectors achieving picosecond time resolution*. 2021. URL: `https://indico.cern.ch/event/1019078/contributions/4443951/`.

[30] *FastIC: A Fast Integrated Circuit for the Readout of High Performance Detectors*. 2021. URL: `https://indico.cern.ch/event/1019078/contributions/4443966/`.

[31] *FBK Fondazione Bruno Kessler official website*. `https://www.fbk.eu/en/`.

[32] P. Fischer. "First implementation of the MEPHISTO binary readout architecture for strip detectors". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 461.1 (2001). 8th Pisa Meeting on Advanced Detectors, pp. 499–504. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(00)01283-3.

[33] G.T. Forcolin et al. "Development of 3D trenched-electrode pixel sensors with improved timing performance". In: *Journal of Instrumentation* 14.07 (July 2019), pp. C07011–C07011. DOI: 10.1088/1748-0221/14/07/c07011. URL: https://doi.org/10.1088/1748-0221/14/07/c07011.

[34] Steve Gaalema. "Low Noise Random-Access Readout Technique for Large Pin Detector Arrays". In: *IEEE Transactions on Nuclear Science* 32.1 (1985), pp. 417–418. DOI: 10.1109/TNS.1985.4336866.

[35] S. Garbolino, S. Martoiu, and A. Rivetti. "Implementation of Constant-Fraction-Discriminators (CFD) in sub-micron CMOS technologies". In: *2011 IEEE Nuclear Science Symposium Conference Record*. 2011, pp. 1530–1535. DOI: 10.1109/NSSMIC.2011.6154364.

[36] *Genesys 2 Reference Manual*. Digilent. 2015. URL: https://digilent.com/reference/programmable-logic/genesys-2/reference-manual.

[37] Abderrahmane Ghimouz, Fatah Ellah Rarbi, and Olivier Rossetto. *DIAMA-SIC: A multichannel front-end electronics for high-accuracy time measurements for diamond detectors*. 2021. arXiv: 2110.12440 [physics.ins-det].

[38] Jung-Suk Goo et al. "Physical Origin of the Excess Thermal Noise in Short Channel MOSFETs". In: *Electron Device Letters, IEEE* 22 (Mar. 2001), pp. 101–103. DOI: 10.1109/55.902845.

[39] F Hahn et al. *NA62: Technical Design Document*. Tech. rep. Geneva: CERN, Dec. 2010. URL: https://cds.cern.ch/record/1404985.

[40] *HAMEG Triple Power Supply HM7042-3*. ROHDE & SCHWARTZ GmbH & Co. Feb. 2020. URL: %5Curl%7Bhttps://scdn.rohde-schwarz.com/ur/pws/dl_downloads/dl_common_library/dl_manuals/gb_1/h/hm7042_5/HM7042-5_UserManual_de_en_06.pdf%7D.

[41] Frank Hartmann. "Silicon tracking detectors in high-energy physics". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 666 (2012). Advanced Instrumentation, pp. 25–46. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2011.11.005. URL: https://www.sciencedirect.com/science/article/pii/S0168900211020389.

[42] K. Heijhoff et al. *Timing performance of the Timepix4 front-end.* 2022. DOI: 10.48550/ARXIV.2203.15912. URL: https://arxiv.org/abs/2203.15912.

[43] Erik H.M. Heijne. "Semiconductor micropattern pixel detectors: a review of the beginnings". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 465.1 (2001). SPD2000, pp. 1–26. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(01)00340-0. URL: https://www.sciencedirect.com/science/article/pii/S0168900201003400.

[44] M. R. Hoeferkamp et al. "Novel Sensors for Particle Tracking: a Contribution to the Snowmass Community Planning Exercise of 2021". In: (Feb. 2022). arXiv: 2202.11828 [physics.ins-det].

[45] Gerald C. Huth. "Recent Results Obtained with High Field, Internally Amplifying Semiconductor Radiation Detectors". In: *IEEE Transactions on Nuclear Science* 13.1 (1966), pp. 36–42. DOI: 10.1109/TNS.1966.4323942.

[46] *I2C-bus specification and user manual.* UM10204. Rev. 7.0. NXP Semiconductors. Oct. 2021. URL: %5Curl%7Bhttps://www.nxp.com/docs/en/user-guide/UM10204.pdf%7D.

[47] Tetsuya Iizuka et al. "A fine-resolution pulse-shrinking time-to-digital converter with completion detection utilizing built-in offset pulse". In: *2016 IEEE Asian Solid-State Circuits Conference (A-SSCC).* 2016, pp. 313–316. DOI: 10.1109/ASSCC.2016.7844198.

[48] *INFN Istituto Nazioanle di Fisica Nucleare official website.* https://home.infn.it/it/.

[49] Jize Jiang et al. "Total Ionizing Dose (TID) effects on finger transistors in a 65nm CMOS process". In: *2016 IEEE International Symposium on Circuits and Systems (ISCAS).* 2016, pp. 5–8. DOI: 10.1109/ISCAS.2016.7527156.

[50] J. B. Johnson. "Thermal Agitation of Electricity in Conductors". In: *Phys. Rev.* 32 (1 July 1928), pp. 97–109. DOI: 10.1103/PhysRev.32.97. URL: https://link.aps.org/doi/10.1103/PhysRev.32.97.

[51] P.R. Kinget. "Device mismatch and tradeoffs in the design of analog circuits". In: *IEEE Journal of Solid-State Circuits* 40.6 (2005), pp. 1212–1224. DOI: 10.1109/JSSC.2005.848021.

[52] Junjie Kong et al. "A 9-bit, 1.08ps resolution two-step time-to-digital converter in 65 nm CMOS for time-mode ADC". In: *2016 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS).* 2016, pp. 348–351. DOI: 10.1109/APCCAS.2016.7803972.

[53] F. Krummenacher. "Pixel detectors with local intelligence: an IC designer point of view". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 305.3 (1991), pp. 527–532.

[54] Thanushan Kugathasan et al. "Monolithic CMOS sensors for sub-nanosecond timing". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 979 (2020), p. 164461. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2020.164461. URL: https://www.sciencedirect.com/science/article/pii/S0168900220308585.

[55] A. Lai. "Sensors, electronics and algorithms for tracking at the next generation of colliders". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 927 (2019), pp. 306–312. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2019.02.050. URL: https://www.sciencedirect.com/science/article/pii/S0168900219302384.

[56] A. Lai. "Timing characterisation of 3D-trench silicon sensors". In: *Journal of Instrumentation* 15.09 (Sept. 2020), pp. C09054–C09054. DOI: 10.1088/1748-0221/15/09/c09054. URL: https://doi.org/10.1088/1748-0221/15/09/c09054.

[57] A. Lai et al. "First results of the TIMESPOT project on developments on fast sensors for future vertex detectors". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 981 (2020), p. 164491. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2020.164491. URL: https://www.sciencedirect.com/science/article/pii/S0168900220308883.

[58] Marc D. Levenson, N.S. Viswanathaan, and Rovert A. Simpson. "Improving Resolution in Photolithography with a Phase-Shifting Mask". In: *IEEE Transactions on Electron Devices* 29.12 (Dec. 1982).

[59] Ted Liu. "The ETROC1: the first full chain precision timing prototype for CMS Endcap Timing Layer upgrade". TWEPP 2021. Sept. 2021. URL: https://indico.cern.ch/event/1019078/contributions/4443946/.

[60] X. Llopart et al. "Timepix, a 65k programmable pixel readout chip for arrival time, energy and/or photon counting measurements". In: *Nucl. Instrum. Meth. A* 581 (2007). Ed. by Josef Hrubec et al. [Erratum: Nucl.Instrum.Meth.A 585, 106–108 (2008)], pp. 485–494. DOI: 10.1016/j.nima.2007.08.079.

[61] X. Llopart et al. "Timepix4, a large area pixel detector readout chip which can be tiled on 4 sides providing sub-200 ps timestamp binning". In: *Journal of Instrumentation* 17.01 (Jan. 2022), p. C01044. DOI: 10.1088/1748-0221/17/01/c01044. URL: https://doi.org/10.1088/1748-0221/17/01/c01044.

[62] A. Loi, A. Contu, and A. Lai. "Timing optimisation and analysis in the design of 3D silicon sensors: the TCoDe simulator". In: *Journal of Instrumentation* 16.02 (Feb. 2021), P02011–P02011. DOI: 10.1088/1748-0221/16/02/p02011. URL: https://doi.org/10.1088/1748-0221/16/02/p02011.

[63] A. Loi et al. "Simulation of 3D-Silicon sensors for the TIMESPOT project". In: *Nucl. Instrum. Meth. A* 936 (2019). Ed. by Giovanni Batignani et al., pp. 701–702. DOI: 10.1016/j.nima.2018.10.134.

[64] *LT3080 Adjustable1.1A Single Resistor Low Dropout Regulator*. Rev. C. Linear Technology Corporation. 2017. URL: %5Curl%7Bhttps://www.analog.com/media/en/technical-documentation/data-sheets/3080fc.pdf%7D.

[65] *LTC2604-LTC2614-LTC2624 Quad 16-Bit Rail-to-Rail DACs in 16-Lead SSOP*. Rev. D. Linear Technology Corporation. 2004. URL: %5Curl%7Bhttps://www.analog.com/media/en/technical-documentation/data-sheets/2604fd.pdf%7D.

[66] Gerhard Lutz. *Semiconductor Radiation Detectors*. Berlin Heidelberg: Springer Berlin, Heidelberg, 2007.

[67] B. Markovic et al. "ALTIROC1, a 20 ps time-resolution ASIC prototype for the ATLAS High Granularity Timing Detector (HGTD)". In: *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*. 2018, pp. 1–3. DOI: 10.1109/NSSMIC.2018.8824723.

[68] Sorin Martoiu et al. "A low power front-end prototype for silicon pixel detectors with 100ps time resolution". In: *2008 IEEE Nuclear Science Symposium Conference Record*. 2008, pp. 2958–2961. DOI: 10.1109/NSSMIC.2008.4774984.

[69] *Medipix Collaboration*. URL: https://medipix.web.cern.ch/.

[70] Roberto Mendicino et al. "3D trenched-electrode sensors for charged particle tracking and timing". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 927 (2019), pp. 24–30. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2019.02.015. URL: https://www.sciencedirect.com/science/article/pii/S0168900219301901.

[71] N. Moffat et al. "Low Gain Avalanche Detectors (LGAD) for particle physics and synchrotron applications". In: *Journal of Instrumentation* 13.03 (Mar. 2018), pp. C03014–C03014. DOI: 10.1088/1748-0221/13/03/c03014. URL: https://doi.org/10.1088/1748-0221/13/03/c03014.

[72] Hasan Molaei and Khosrow Hajsadeghi. "A low power high resolution time to digital converter for ADPLL application". In: *2016 IEEE 59th International Midwest Symposium on Circuits and Systems (MWSCAS)*. 2016, pp. 1–4. DOI: 10.1109/MWSCAS.2016.7870107.

[73] Hasan Molaei, Ata Khorami, and Khosrow Hajsadeghi. "A wide dynamic range low power $2\times$ time amplifier using current subtraction scheme". In: *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*. 2016, pp. 462–465. DOI: 10.1109/ISCAS.2016.7527277.

[74] A. Morozzi et al. "3D Diamond Tracking Detectors: numerical analysis for Timing applications with TCAD tools". In: *Journal of Instrumentation* 15.01 (Jan. 2020), pp. C01048–C01048. DOI: 10.1088/1748-0221/15/01/c01048. URL: https://doi.org/10.1088/1748-0221/15/01/c01048.

[75] Makoto Motoyoshi. "Through-Silicon Via (TSV)". In: *Proceedings of the IEEE* 97.1 (2009), pp. 43–48. DOI: 10.1109/JPROC.2008.2007462.

[76] N. Neri et al. "4D fast tracking for experiments at high luminosity LHC". In: *Journal of Instrumentation* 11.11 (Nov. 2016), pp. C11040–C11040. DOI: 10.1088/1748-0221/11/11/c11040. URL: https://doi.org/10.1088/1748-0221/11/11/c11040.

[77] Tahereh Sadat Niknejad. *Results with the TOFHIR2X revision of the front-end ASIC of the CMS MTD Barrel Timing Layer*. Tech. rep. Geneva: CERN, Dec. 2021. URL: https://cds.cern.ch/record/2799504.

[78] Matthew Noy et al. "The TDCPix ASIC: Tracking for the NA62 Giga-Tracker". In: *PoS* TIPP2014 (2014), 183. 9 p. DOI: 10.22323/1.213.0183. URL: https://cds.cern.ch/record/2025938.

[79] Peter Osheroff, George S. La Rue, and Subhanshu Gupta. "A highly linear 4GS/s uncalibrated voltage-to-time converter with wide input range". In: *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*. 2016, pp. 89–92. DOI: 10.1109/ISCAS.2016.7527177.

[80] *pandas: powerful Python data analysis toolkit*. Rel. 1.4.2. Wes McKinney and the Pandas Development Team. Apr. 2022. URL: %5Curl%7Bhttps://pandas.pydata.org/docs/pandas.pdf%7D.

[81] S.I. Parker, C.J. Kenney, and J. Segal. "3D — A proposed new architecture for solid-state radiation detectors". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 395.3 (1997). Proceedings of the Third International Workshop on Semiconductor Pixel Detectors for Particles and X-rays, pp. 328–343. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(97)00694-3. URL: https://www.sciencedirect.com/science/article/pii/S0168900297006943.

[82] Giuliano Parrini. "Laser graphitization for polarization of diamond sensors". In: *PoS* RD11 (2012), p. 017. DOI: 10.22323/1.143.0017.

[83] M.J.M. Pelgrom, A.C.J. Duinmaijer, and A.P.G. Welbers. "Matching properties of MOS transistors". In: *IEEE Journal of Solid-State Circuits* 24.5 (1989), pp. 1433–1439. DOI: 10.1109/JSSC.1989.572629.

[84] Matteo Perenzoni, Leonardo Gasparini, and David Stoppa. "Design and Characterization of a 43.2-ps and PVT-Resilient TDC for Single-Photon Imaging Arrays". In: *IEEE Transactions on Circuits and Systems II: Express Briefs* 65.4 (2018), pp. 411–415. DOI: 10.1109/TCSII.2017.2694482.

[85] *PG3A User's Manual*. Ver. 1.13. The Moving Pixel Company. Jan. 2018. URL: %5Curl%7Bhttp://www.movingpixel.com/PG3AUsersManual1_13.pdf%7D.

[86] L. Piccolo et al. "First measurements on the Timespot1 ASIC: a fast-timing, high-rate pixel-matrix front-end". In: *Journal of Instrumentation* 17.03 (Mar. 2022), p. C03022. DOI: 10.1088/1748-0221/17/03/c03022. URL: https://doi.org/10.1088%2F1748-0221%2F17%2F03%2Fc03022.

[87] Lorenzo Piccolo. "A Timing Pixel Front-End Design for HEP Experiments in 28 nm CMOS Technology". In: *2019 15th Conference on Ph.D Research in Microelectronics and Electronics (PRIME)*. 2019, pp. 205–208. DOI: 10.1109/PRIME.2019.8787759.

[88] Lorenzo Piccolo. "First measurements on a discrete-time front-end in 28-nm CMOS technology for timing pixel detectors". In: *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. 2019, pp. 1–4. DOI: 10.1109/NSS/MIC42101.2019.9059904.

[89] Lorenzo Piccolo et al. "The first ASIC prototype of a 28 nm time-space front-end electronics for real-time tracking". In: *PoS* TWEPP2019 (2020), p. 022. DOI: 10.22323/1.370.0022.

[90] T Poikela et al. "Timepix3: a 65K channel hybrid pixel readout chip with simultaneous ToA/ToT and sparse readout". In: *Journal of Instrumentation* 9.05 (May 2014), pp. C05013–C05013. DOI: 10.1088/1748-0221/9/05/c05013. URL: https://doi.org/10.1088/1748-0221/9/05/c05013.

[91]   *Precision timing ASIC for LGAD sensors based on a Constant Fraction Discriminator – FCFD0.* 2021. URL: https://indico.cern.ch/event/1019078/contributions/4443948/.

[92]   *pySerial's documentation.* 2015. URL: https://pythonhosted.org/pyserial/.

[93]   *Python 3.10.4 documentation.* Python organization. 2022. URL: https://docs.python.org/3/.

[94]   *pythread github page.* DedInc, 2022. URL: https://github.com/DedInc/pythread.

[95]   *QSH and QTH product specification.* Rev. F. Samtec, Inc. Mar. 2021. URL: %5Curl%7Bhttps://suddendocs.samtec.com/productspecs/qsh-qth.pdf%7D.

[96]   John Robertson. "Band offsets and work function control in field effect transistors". In: *J. Vac. Sci. Technol. B* 27.1 (Feb. 2009).

[97]   Eduard Säckinger. "The Transimpedance Limit". In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 57.8 (2010), pp. 1848–1856. DOI: 10.1109/TCSI.2009.2037847.

[98]   Frank Schellenberg. "A little light magic". In: *IEEE Spectrum* (Sept. 2003).

[99]   *Si5341-40 Rev D Data Sheet.* Rev. 1.1. Skyworks Solutions, Inc. Nov. 2021. URL: %5Curl%7Bhttps://www.skyworksinc.com/-/media/SkyWorks/SL/documents/public/data-sheets/si5341-40-d-datasheet.pdf%7D.

[100] *si5341-d-evb - evaluation board's guide.* Rev. 1.0. Skyworks Solutions, Inc. Dec. 2021. URL: %5Curl%7Bhttps://www.skyworksinc.com/-/media/Skyworks/SL/documents/public/user-guides/Si5341-D-EVB.pdf%7D.

[101] Kyle Siegrist. *Probability, Mathematical, Statistics, and Stochastic Processes.* University of Alabama in huntsville: LibreTexts.

[102] *SN65LVDS4 1.8-V High-Speed Differential Line Receiver.* Texas Instruments Incorporated. Nov. 2017. URL: %5Curl%7Bhttps://www.ti.com/lit/ds/symlink/sn65lvds4.pdf?HQS=dis-dk-null-digikeymode-dsf-pf-null-wwe&ts=1654191093295&ref_url=https%253A%252F%252Fwww.ichunter.com%252F%7D.

[103] W. Snoeys. "Monolithic pixel detectors for high energy physics". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 731 (2013). PIXEL 2012, pp. 125–130. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2013.05.073. URL: https://www.sciencedirect.com/science/article/pii/S0168900213006840.

[104] Valentina Sola. "Precision Timing with the CMS MTD Endcap Timing Layer for HL-LHC". In: *JPS Conf. Proc.* 34 (2021), p. 010013. DOI: 10.7566/JPSCP.34.010013.

[105] H. Sun et al. "Characterization of the CMS Endcap Timing Layer readout chip prototype with charge injection". In: *Journal of Instrumentation* 16.06 (June 2021), P06038. DOI: 10.1088/1748-0221/16/06/p06038. URL: https://doi.org/10.1088/1748-0221/16/06/p06038.

[106] The ATLAS Collaborationa. *High-Granularity Timing Detector for the ATLAS Phase-II Upgrade.* Technical Proposal. CERN, July 2018.

[107] The CMS Collaboration. *Technical Proposal for a MIP Timing Detector in the CMS Experiment Phase 2 Upgrade.* Tech. rep. CERN, Nov. 2017.

[108] The LHCb Collaboration. *LHCb Upgrades and operation at 10e34cm-2s-1 luminosity, A first study.* CERN-ACC-note. CERN, Aug. 2018.

[109] *TIFPA - Trento Institute for Fundamental Physics and Applications official website.* https://www.tifpa.infn.it/.

[110] *TLA7000 Logic Analyzers TLA7000 Series Data Sheet.* Tektronix, Inc. A. Mar. 2017. URL: %5Curl%7Bhttps://download.tek.com/datasheet/TLA7000-Logic-Analyzer-Datasheet-52W1505322_0.pdf%7D.

[111] Gianluca Traversi et al. "Design of LVDS driver and receiver in 28 nm CMOS technology for Associative Memories". In: *2017 6th International Conference on Modern Circuits and Systems Technologies (MOCAST).* 2017, pp. 1–4. DOI: 10.1109/MOCAST.2017.7937618.

[112] *USB-to-I2C Elite I2C-SMBus and SPI control DLL Users'Manual.* SB Solutions, Inc. Feb. 2013. URL: %5Curl%7Bhttps://i2ctools.com/Downloads/USBtoI2Celite/USB-to-I2C_Elite_DLL_Users_Manual.pdf%7D.

[113] I. Vornicu, R. Carmona-Galán, and Á. Rodríguez-Vázquez. "In-pixel voltage-controlled ring-oscillator for phase interpolation in ToF image sensors". In: *2016 IEEE International Symposium on Circuits and Systems (ISCAS).* 2016, pp. 1906–1909. DOI: 10.1109/ISCAS.2016.7538945.

[114] S. White. "R&D for a Dedicated Fast Timing Layer in the CMS Endcap Upgrade". In: *Acta Physica Polonica B Proceedings Supplement* 7.4 (2014), p. 743. DOI: 10.5506/aphyspolbsupp.7.743. URL: https://doi.org/10.5506%2Faphyspolbsupp.7.743.

[115] *Xilinx Kintex-7 FPGA KC705 User Manual.* Xilinx. 2014. URL: https://www.manualslib.com/manual/1436729/Xilinx-Kintex-7-Fpga-Kc724.html?page=2#manual.

155

[116] Chun-Min Zhang et al. "Bias Dependence of Total Ionizing Dose Effects on 28-nm Bulk MOSFETs". In: *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*. 2018, pp. 1–3. DOI: 10.1109/NSSMIC.2018.8824379.

[117] Chun-Min Zhang et al. "Characterization and Modeling of Gigarad-TID-Induced Drain Leakage Current of 28-nm Bulk MOSFETs". In: *IEEE Transactions on Nuclear Science* 66.1 (2019), pp. 38–47. DOI: 10.1109/TNS.2018.2878105.

[118] Wei Zhang et al. "A Low-Power Time-to-Digital Converter for the CMS Endcap Timing Layer (ETL) Upgrade". In: *IEEE Transactions on Nuclear Science* PP (June 2021), pp. 1–1. DOI: 10.1109/TNS.2021.3085564.