

[en][state=start]

Summary

1	Introduction	2
2	Pixel detectors	3
2.1	Hybrid pixels	3
2.2	CMOS MAPS and DMPAS	4
2.2.1	DMAPS: large and small fill factor	4
2.2.2	A modified sensor	5
2.3	Analog front end	6
2.3.1	Preamplifier	7
2.3.2	ALPIDE-like front end	7
2.4	Readout logic	8
3	Use of pixels detector	11
3.1	Tracking in HEP	11
3.1.1	Two HEP experiments who chose CMOS-DMAPS	11
3.2	Dosimetry	11
3.2.1	Applicability to FLASH radiotherapy	11
4	TJ-Monopix1	12
4.1	The sensor	12
4.2	FE flavors	14
4.2.1	FE parameters	15
4.3	Readout logic and Data-packets structure	15
4.3.1	Dead time measurement	16
5	Arcadia-MD1	17
5.1	The sensor	17
5.2	Readout logic and data structure	17
5.2.1	Matrix division and data-packets	17
6	Threshold and noise characterization	21
6.1	Threshold and noise: figure of merit for pixel detectors	21
6.2	TJ-Monopix1 characterization	22
6.3	ARCADIA-MD1 characterization	22
A	Pixels detector: a brief overview	23
A.1	Signal formation	23
A.2	Radiation damages	24
B	FLASH radiotherapy	26
B.1	Cell survival curves	26
B.2	FLASH effect	27
	Bibliography	28

Chapter 1

Introduction

Pixel detectors, members of the semiconductor detector family, have significantly been used since () at the first accelerator experiments for energy and position measurement. Because of their dimension (today $\sim 30 \mu m$ or even better) and their spatial resolution ($\sim 5\text{-}10 \mu m$), with the availability of technology in 1980s they proved to be perfectly suitable for vertex detector in the inner layer of the detector.

Technological development has been constant from then on and today almost every high energy physics (HEP) experiment employs a pixels detector; hybrid pixel currently constitute the state-of-art for large scale pixel detector but experiments began to look at the more innovative monolithic active pixels (MAPS) as perspective for their future upgrades, as BelleII, or they already have installed them, as ALICE.

Requirement imposed by accelerator are stringent and they will be even more with the increase of luminosity/intensity, in terms of radiation hardness, efficiency and occupancy, time resolution, material budget and power consumption.

Qual è invece la richiesta per la dosimetria?

Chapter 2

Pixel detectors

2.1 Hybrid pixels

Hybrid pixels are made of two parts (fig. 2.1a), the sensor and the electronics: for each pixel these two parts are welded together through microconnection (bump bond).

Hybrid pixels provide a practical system where readout and sensor, being independent, can be optimized separately, although the sensors cannot be tested without being connected to the readout and this represents a significant practical issue.

In addition the particular and sophisticated procedure to bond sensor and ASIC (application specific integrated circuit) makes them difficult to produce, delicate, especially for high levels of radiation, and also expensive.

A critical parameter for accelerator experiments is the material budget, which represents the main limit factor for momentum measurement resolution in a magnetic field; since hybrid pixels are thicker (\sim hundreds of μm) than monolithic ones (even less than $100 \mu m$), using the latter the material budget can be down by a third: typical value for hybrid pixels is $1.5 \% X_0$ per layer, while for monolithic $0.5 \% X_0$.

Among other disadvantages of hybrid pixels there is the bigger power consumption that implies, by the way, a bigger cooling system leading in turn to an increase in material too.

DEPFET are the first attempt towards the integration of the front end (FE) on the sensor bulk: they are typically mounted on a hybrid structure but they also integrate the first amplification stage.

Each pixel implements a MOSFET (metal-oxide-semiconductor field-effect transistor) transistor (a p-channel in fig. 2.1b): an hole current flows from source to drain which is controlled by the external gate and the internal gate together. The internal gate is made by a deep $n+$ implant towards which electrons drift after being created in the deplation region (to know how the signal is created in a pixel detector look at appendix A); the accumulation of electrons in the region

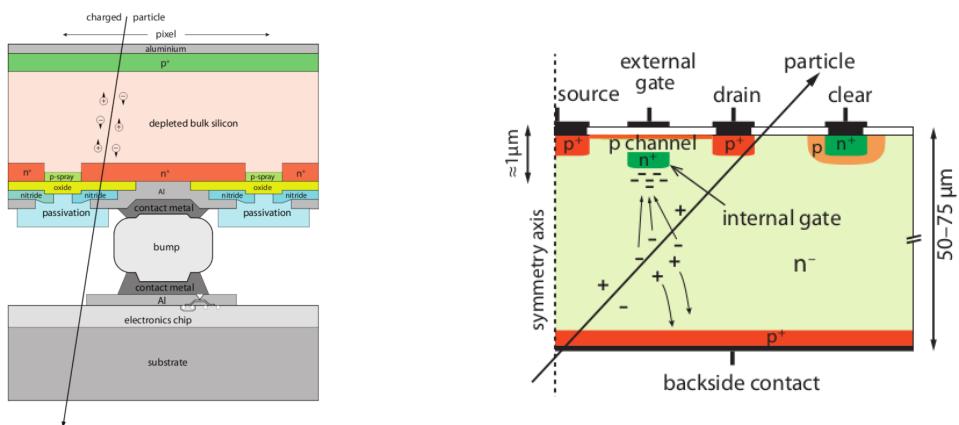


Figure 2.1: Concept cross section of hybrid pixel (a) and of a DEPFET (b)

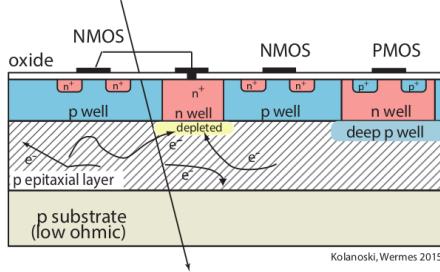


Figure 2.2: Concept cross section of CMOS MPAS pixel

underneath the n implant changes the gate potential and controls the transistor current. DEPFET typically have a good S/N ratio: this is principally due the amplification on-pixel and the large depletion region. But, since they need to be connect with ASIC the limiting factor still is the material budget.

2.2 CMOS MAPS and DMPAS

Monolithic active pixels accomodate on the same wafer both the sensor and the front end electronics, with the second one implanted on top.

MAPS have been first proposed and realized in 1990s and their usage has been enabled by the development of the electronic sector which guarantees the decrease in CMOS transistors dimension at least every two years, as stated by the Moore's law¹.

As a matter of fact the dimension of components, their organization on the pixel area and logic density are important issues for the design and for the layout; typically different decisions are taken for different purposes.

Monolithic active pixel can be distinguish between two main category: MAPS and depleted MAPS (DMAPS).

MAPS (figure 2.2) have typically an epitaxial layer in range 1-20 μm and because they are not depleted, the charge is mainly collected by diffusion rather than by drift. This makes the path of charges created in the bulk longer than usual, therefore they are slow (of order of 100 ns) and the collection could be partial especially after the irradiation of the detector (look at A for radiation damages), when the trapping probability become higher.

In figure 2.2 is shown as example of CMOS MAPS: the sensor in the scheme implements an n well as collection diode; to avoid the others n wells (which contain PMOS transistor) of the electronic circuit would compete in charge collection and to shield the CMOS circuit from the substrate, additionally underlying deep p well are needed. DMAPS are instead MAPS depleted with d typically in $\sim 25\text{-}150 \mu\text{m}$ (eq. A.1) which extends from the diode to the deep p-well, and sometimes also to the backside (in this case if one wants to collect the signal also on this electrode, additional process must be done).

2.2.1 DMAPS: large and small fill factor

There are two different sensor-design approaches (figure 2.3) to DMAPS:

- large fill factor: a large collection electrode that is a large deep n-well and that host the embedded electronics
- small fill factor: a small n-well is used as charge collection node

To implement a uniform and stronger electric field, DMAPS often uses large electrode design that requires multiple wells (typically four including deep n and p wells); this layout adds on to the standard terms of the total capacity of the sensor a new term (fig. 2.4), that contributes to the total amplifier input capacity. In addition to the capacity between pixels (C_{pp}) and between the pixel and the backside (C_b), a non negligible contribution comes from the capacities between wells (C_{SW} and C_{WW}) needed to shield the embedded electronics. These capacities affect the thermal and 1/f noise of the charge amplifier and the τ_{CSA} too:

¹Moore's law states that logic density doubles every two years.

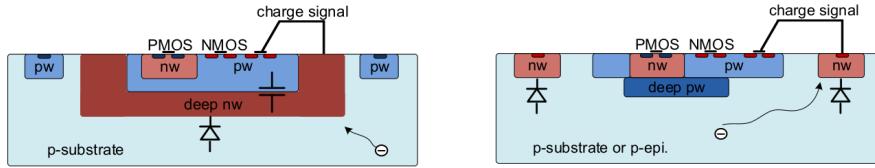


Figure 2.3: Concept cross section with large and small fill factor

	small fill factor	large fill factor
small sensor C	✓ (< 5 fF)	✗ (~ 100-200 fF)
low noise	✓	✗
low cross talk	✓	✗
velocity performances	✓	✗ (~ 100 ns)
short drift paths	✗	✓
radiation hard	✗	✓

Table 2.1: Small and large fill factor DMAPS characteristics

$$ENC_{thermal}^2 \propto \frac{4}{3} \frac{kT}{g_m} \frac{C_D^2}{\tau_{sh}} \quad (2.1)$$

$$\tau_{CSA} \propto \frac{1}{g_m} \frac{C_D}{C_f} \quad (2.2)$$

where g_m is the transconductance, τ_{sh} is the shaping time. Among the disadvantages coming from this large input capacity could be the coupling between the sensor and the electronics resulting in cross talk: noise induced by a signal on neighbouring electrodes; indeed, since digital switching in the FE electronics do a lot of oscillations, this problem is especially connected with the intra wells capacities. So, larger charge collection electrode

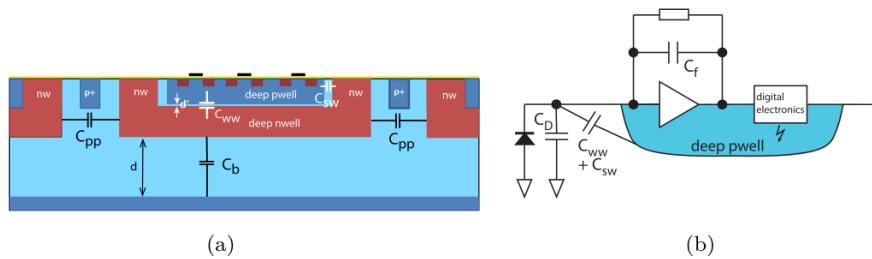


Figure 2.4: C_{pp} , C_b , C_{WW} , C_{SW}

sensors provide a uniform electric field in the bulk that results in short drift path and so in good collection properties, especially after irradiation, when trapping probability can become an issue. The drawback of a large fill-factor is the large capacity (~100 fF): this contributes to the noise and to a speed penalty and to a larger possibility of cross talk.

The small fill-factor variant, instead, benefits from a small capacity (5-20 fF), but suffers from a not uniform electric field and from all the issue related to that. **Ho già detto prima parlando dei MAPS, devo ripetere qui?**

As we'll see these two different types of sensor require different amplifier: the large electrode one is coupled with the charge sensitive amplifier, while the small one with voltage amplifier (sec 2.3.1).

2.2.2 A modified sensor

?? A process modification that has become the standard solution to combine the characteristic of a small fill factor sensor (small input amplifier capacity) and of large fill factor sensor (uniform electric field) is the one carried out for ALICE upgrade about ten years [1].

A compromise between the two sensors could also be making smaller pixels but this solution requires reducing the electronic circuit area, so a completely new pixel layout should be think.

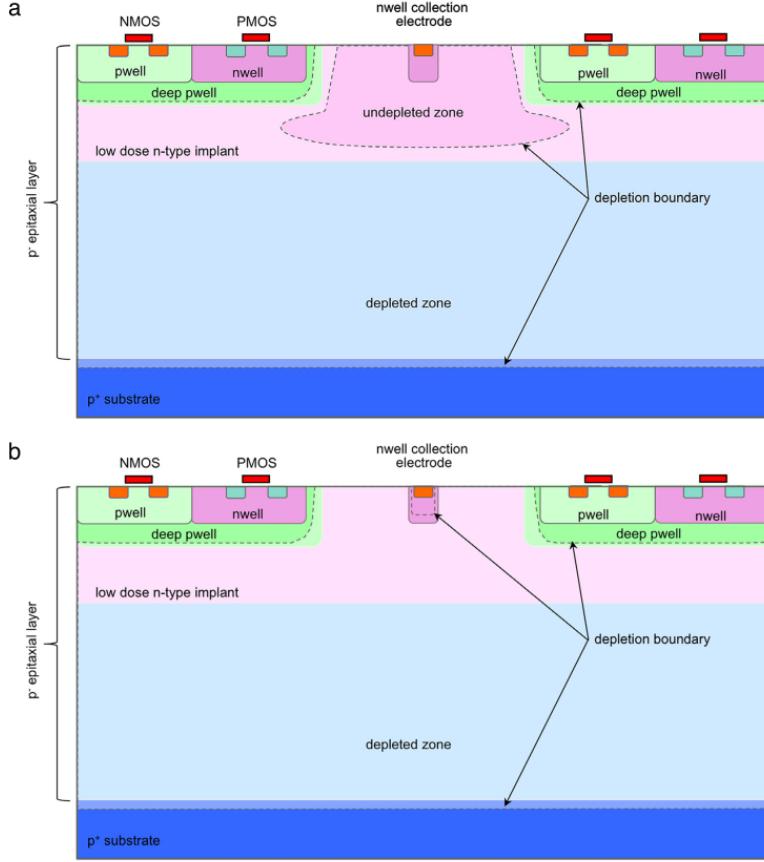


Figure 2.5: A modified process for ALICE tracker detector: a low dose n implant is used to create a planar junction. In (a) the depletion is partial, while in (b) the pixel is fully depleted.

The modification consists in inserting a low dose implant under the electrode and one its advantage lies in its versatility: both standard and modified sensor are often produced for testing in fact.

Before the process modification the depletion region extends below the diode towards the substrate and it doesn't extend laterally so much even if a high bias is applied to the sensor (fig. 2.5). After, two distinct pn junctions are built: one between the deep p well and the n^- layer, and the other between the n^- and the p^- epitaxial layer, extending to the all area of the sensor.

Since deep p well and the p-substrate are separated by the depletion region, the two p electrodes can be biased separately² and this is beneficial to enhance the vertical electric field component.

The doping concentration is a trimmer parameter: it must be high enough to be greater than the epitaxial layer to prevent the punchthrough between p-well and the substrate, but it must also be lower enough to allow the depletion without reaching too high bias.

2.3 Analog front end

After the creation of a signal on the electrode, the signal enters in the front end circuit (fig.2.6), ready to be molded and transmitted out of chip. Low noise amplification, fast hit discrimination and an efficient, high-speed readout architecture, consuming as low power as possible must be provided by the read out integrated electronics (ROIC).

Let's take a look to the main steps of the analog front end chain: the preamplifier (that actually often is the only amplification stage) with a reset to the baseline mechanism and a leakage current compensation, a shaper (a band-pass filter) and finally a discriminator. The whole chain must be optimized and tuned to improve the S/N ratio: it is very important both not to have a large noise before the amplification stage in order to not moltiplicate that noise, and chose a reasonable

²This is true in general, but it can be denied if other doping characteristics are implemented and we'll see that this is the case of TJ-Monopix1

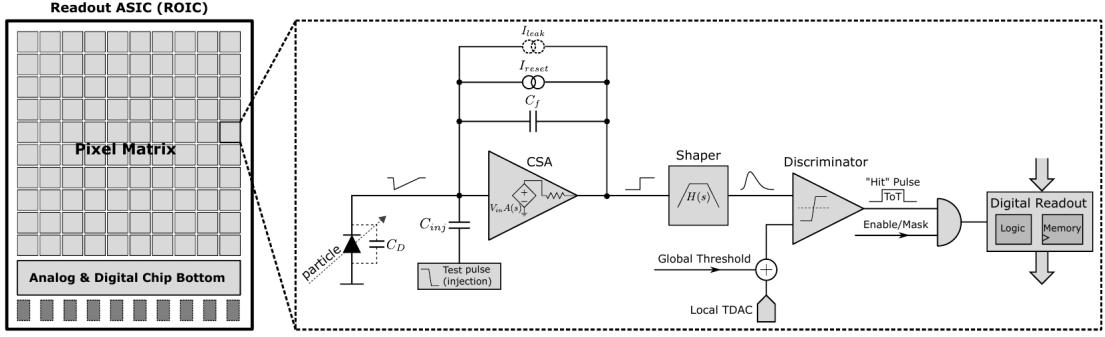


Figure 2.6: Readout FE scheme: in this example the preamplifier is a charge sensitive one (CSA) but changing the capacitive feedback into a resistive one, this can be converted in a voltage or current amplifier.

threshold of the discriminator to cut noise-hits much as possible.

2.3.1 Preamplifier

Even if circuits on the silicon crystal are only constructed by CMOS, a preamplifier can be modelled as an operational amplifier (OpAmp) where the gain is determined by the input and feedback impedance (first step in figure 2.6):

$$G = \frac{v_{out}}{v_{in}} = \frac{Z_f}{Z_{in}} \quad (2.3)$$

Depending on whether a capacity or a resistance is used as feedback, respectively a charge or a voltage amplifier is used: if the voltage input signal is large enough and have a sharp rise time, the voltage sensitive preamplifier is preferred. Consequently this flavor doesn't suit to large fill factor MAPS whose signal is already enough high: $v_{in} = Q/C_D \approx 3fC/100 \text{ pF} = 0.03 \text{ mV}$, but it's fine for the small fill factor ones: $v_{in} = Q/C_D \approx 3fC/3 \text{ pF} = 1 \text{ mV}$.

In the case of a resistor feedback, if the signal duration time is longer than the discharge time ($\tau = R_S C_D$) of the detector the system works as current amplifier, as the signal is immediately transmitted to the amplifier; in the complementary case (signal duration longer than the discharge time) the system integrates the current on the C_D and operates as a voltage amplifier.

2.3.2 ALPIDE-like front end

I've already mentioned ALICE pixel detector talking about the new process modification, now the ALICE name comes up again talking about FE: this is because ALPIDE (ALice PIxel DEtector) is one of the first MAPS detector (TowerJazz 180 nm CMOS) installed ³, therefore it is the current state of art and most of the following chips' FE are inspired by that, making it a standard in the FE design. **ARCADIA MD1 and TJ-Monopix1 are no exception, this is why I'm going to explain some principals characteristics of how it works**

The idea of the amplification stage is to transfer the charge from a bigger capacity[2], C_{source} , to a smaller one, C_{out} : the input transistor M1 with current source IBIAS acts as a source follower and this forces the source of M1 to be equal to the gate input $\Delta V_{PIX_IN} = Q_{IN}/C_{IN}$.

$$Q_{source} = C_{source} \Delta V_{PIX_IN} \quad (2.4)$$

The current in M2 and the charge accumulates on C_{out} is fixed by the one on C_{source} :

$$\Delta V_{OUT_A} = \frac{Q_{source}}{C_{OUT_A}} = \frac{C_{source} \Delta V_{PIX_IN}}{C_{OUT_A}} = \frac{C_{Source}}{C_{OUT_A}} \frac{Q_{IN}}{C_{IN}} \quad (2.5)$$

A second branch (M4, M5) is used to generate a low frequency feedback, where VCASN and ITHR set the baseline value of the signal on C_{OUT_A} and the velocity to goes down to the baseline.

³It was installed in the Inner Tracking System during the second long shut down of the LHC in 2019

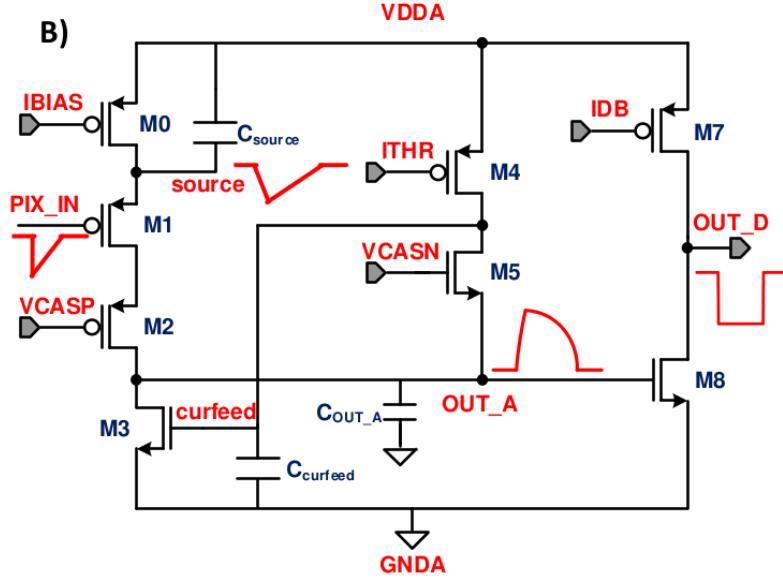


Figure 2.7: ALPIDE like FE

IL RUOLO DI CURVFEED NON L'HO CAPITO, NELL'ARTICOLO C'È SCRITTO: The curfeed net, loaded with $C_{curfeed}$ capacitance and connected to the gate of M3 is adjusted for M3 to absorb IBIAS+ITHR.

Finally IDB defines the charge threshold with which the signal OUT_A must be compared: depending on if the signal is higher than the threshold or not, the OUT_D is high or low respectively.

2.4 Readout logic

Readout logic includes the part of the circuit which takes the FE output singal, processes it and then trsamit it out of pixel and/or out of chip; depending on the situation of usage different readout characteristics must be provided.

To store the analogical informations (i.e. charge collected, evolution of signal in time, ...) big buffers and a large bandwidth are needed; the problem that doesn't occur, or better occur only with really high rate, if one wants record only digital data (if one pixel is hit 1 is recorded, and if not 0 is recorded).

A common compromise often made is to save the time over threshold (ToT) of the pulse in clock cycle counts; this needs of relatively coarse requirement as ToT could be trimmer to be a dozen bits but, being correlated and hopefully being linear with the deposited charge by the impinging particle in the detector, it provides a sufficient information. The ToT digitalization usually takes advantage of the distribution of a clock (namely BCID, bunch crossing identification) on the pixels matrix. The required timing precision is at least around 25 ns, that corresponds to the period of bunch collisions at LHC; for such reason a reasonable BCID-clock frequency for pixels detector is 40 MHz.

Leading and trailing edges' timestamp of the pulse are saved on on pixel within a RAM unitl they have beeen read, and then the ToT is obtained from their difference.

Moreover the readout architecture can be full, if every hit is read, or triggered, if a trigger system decides if the hit must be store or not. On one hand the triggered-readout needs buffers and storage memories, on the other the full readout, because there is no need to store hit data on chip, needs an high enough bandwidth.

A triggered readout is foundamental in accelerator experiments where the quantity of data to store is too large to be hanled and some selection have to be applied by the trigger: to give an order of growth, at LHC more than 100 TBit/s of datas are produced but the storage data limit is about 100 MBit/s [3] (pag. 797).

Typically the trigger signal is processed in a few μs so the pixel gets it only after a hundred clock cycles from the hit arrival time: the buffer dept must than handle the higher trigger latency.

After having taken out the data from the pixel, it have to be transmitted to the end of column

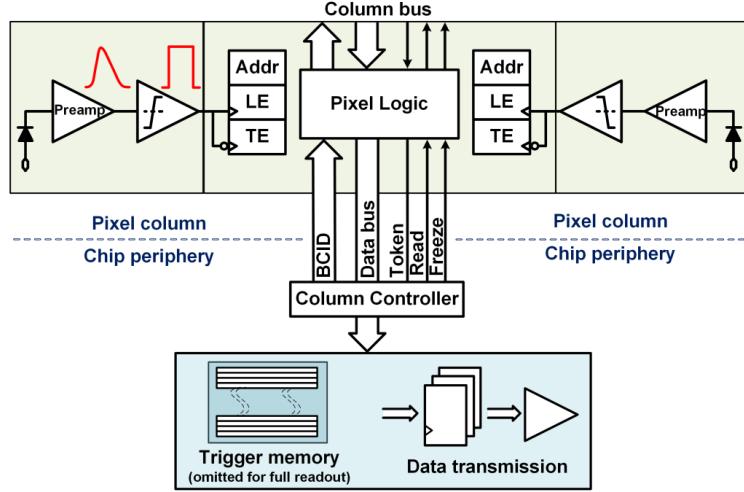


Figure 2.8: Column drain R/O scheme where ToT is saved

(EoC) where a serializer deliver it out of chip, typically to a FPGA.

There are several ways of transmitting data from pixel to the end of column: one of the most famous is the column-drain read out, developed for CMS and ATLAS experiments [4]. All the pixels in a double-column share a data bus and only one pixel at a time, according to a priority chain, can be read. The reading order circuit is implemented by shift register (SR): when a hit arrives, the corresponding data, which can be made of timestamp and ToT, is temporarily stored on a RAM until the SH does not allow the access to memory by data bus.

Even if many readout architectures are based the column-drain one, it doesn't suit for large size matrices. The problem is that increasing the pixels on a column would also raise the number of pixels in the priority chain and that would results in a slowdown of the readout.

If there isn't any storage memory, the double-column behaves as a single server queue and the probability for a pixel of waiting a time T greater than t , with an input hit rate on the column μ and an output bandwidth B_W is [5]:

$$P(T > t) = \frac{\mu}{B_W} e^{-(B_W - \mu)t} \quad (2.6)$$

To avoid hit loss (let's neglect the contribution to the inefficiency of the dead time τ due to the AFE), for example imposing $P(T > t) \sim 0.001$, one obtains $(B_W - \mu) t_t \sim 6$, where t_t is the time needed to transfer the hit; since t_t is small, one must have $B_W \gg \mu$, that means a high bandwidth [5].

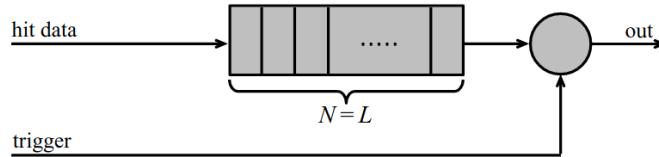


Figure 2.9: Block diagram of a pipeline buffer: N is the dimension of memory buffer and L is the trigger latency expressed in BCID cycles

Actually the previous one is an approximation since each pixel sees a different bandwidth depending on the position on the queue: the first one sees a full bandwidth, but the next sees a smaller one because occasionally it can be blocked by the previous pixel. Then the bandwidth seen by the pixel i is $B_i = B - \sum_j \mu_j$, where μ_j is the hit rate of the j th pixel.

The efficiency requirement on the bandwidth and the hit rate becomes: $B_{W,i} > \mu_i$, where the index i means the constraint is for a single pixel; if all the N pixels on a column have the same rate $\mu = N\mu_i$, the condition reduces to $B_W > \mu$. The bandwidth must be chosen such that the mean time between hits of the last pixel in the readout chain is bigger than that.

In order to reduce the bandwidth a readout with zero suppression on pixel is typically employed;

this means that only informations from channels where the signal exceeds the discriminator threshold are stored. Qualcosa sulla zero suppression? La metto qui questa affermazione?

If instead there is a local storage until a trigger signal arrives, the input rate to column bus μ' is reduced compared to the hit rate μ as: $\mu' = \mu \times r \times t$, where r is the trigger rate and t is the bunch crossing period. In this situation there is a more relaxed constraint on the bandwidth, but the limiting factor is the buffer depth: the amount of memory designed depends both on the expected rate μ and on the trigger latency t as $\propto \mu \times t$, that means that the higher the trigger latency and the lower the hit rate to cope with.

In order to have an efficient usage of memory on pixels area it's convenient grouping pixels into regions with shared storage. Let's compare two different situations: in the first one a buffer is located on each pixel area, while in the second one a core of four pixels share a common buffer (this architecture is commonly called FE-I4).

Consider a 50 kHz single pixel hits rate and a trigger latency of 5 μs , the probability of lossing

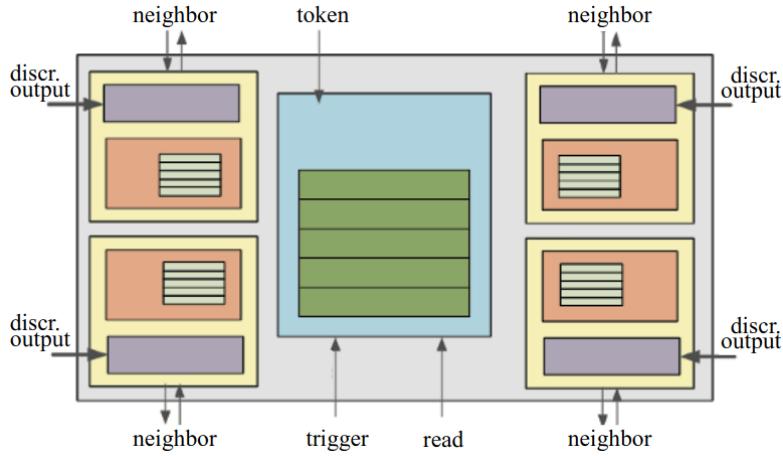


Figure 2.10: Block diagram of the FE-I4 PR organization. Read and memory management section is highlighted in light blue; latency counters and trigger management section are highlighted in green; hit processing blocks are highlighted in purple; ToT counters and ToT management units are highlighted in orange

hits is:

$$P(N > 1|\nu) = 1 - P(N = 0|\nu) - P(N = 1|\nu) = 1 - e^{-\nu}(1 + \nu) \approx 2.6\% \quad (2.7)$$

where I have assumed a Poissonian distribution with mean $\nu = 0.25$ to describe the counts N . To get an efficiency ϵ greater than 99.9 % a 3 hit depth buffer is needed:

$$P(N > 3|\nu) = 1 - \sum_{i=0}^3 P(N = i|\nu) < 0.1\% \quad (2.8)$$

Considering the second situation: if the average single pixel rate is still 50 kHz, grouping four pixels the mean number of hits per trigger latency is $\nu = 0.25 \times 4 = 1$. To get an efficiency of 99.9% (eq. 2.8) a buffer depth of 5 hits in the four-pixels region, instead of 3 per pixels, is needed.

Chapter 3

Use of pixels detector

A partire dal 2017, i sensori CMOS rappresentano l'89% delle vendite globali di sensori di immagine. Ma praticamente dal 2010 in poi solo CMOS e non più CCD.

3.1 Tracking in HEP

Per gli acceleratori la richieste sono molto stringenti e lo saranno sempre di più con l'aumento dell'intensità o della luminosità in termini di radiation hardness (per HL-LHC for example expected in 5 anni 500 Mrad e NIEL di 10 alla 16), efficiency e occupancy (efficienza alta dopo tanta radiazione e noise occupancy bassa), time resolution (bunch crossing 40 Mhz), material budget e power consumption (material budget below 2 per cento e power consumption 500 mW/cm²) Usati come tracciatori per misure di impulso e per misure di energia (per rigettare) ad esempio dati di fondo (topic fondamentale per BELLE-II).

Position measurement resolution

Depending on the type of signal reading the spatial resolution is $\sigma_x = \frac{p}{\sqrt{12}}$ where p is the pitch between pixels, or even better if other analogica information, as the charge, are read and capacitive charge division method is applied.

Momentum measurement resolution

3.1.1 Two HEP experiments who chose CMOS-DMAPS

ALICE

TJ 180 nm CMOS process was firstly used for Alice inner tracker system: ALPIDE (primo ad avere FE sul pixel e sparsified zero suppression readout).

BELLE-II

3.2 Dosimetry

3.2.1 Applicability to FLASH radiotherapy

Chapter 4

TJ-Monopix1

The TJ-Monopix1 is one of the chips fabricated by TowerJazz with 180 CMOS imaging process. From the middle of 2013 a dedicated collaboration, RD 53 ('Development of pixel readout integrated circuits for extreme rate and radiation'), has been established with the specific goal to find a sensor suitable as vertex detector for future upgrade of CMS and ATLAS experiments. Among the main objects of study of the collaboration there are both hybrid pixels and monolithic options as CMOS MAPS: fig 4.1 shows the intermediate MAPS-prototypes made by TowerJazz.

Besides the TowerJazz series, also LFoundry fabricated other similar sensors with 150 CMOS



Figure 4.1: Timeline in TowerJazz productions. In addition to Monopix serie also the small electrode demonstrator TJ-Malta and mini-Malta have been produced and tested[6]. The Malta prototypes differ from TJ-Monopix in the readout: while Monopix implements a column-drain R/O, an asynchronous R/O without any distribution of BCID has been used by TJ-Malta in order to reduce power consumption.

technology [7][8]: LF-Monopix.

The main difference between the LFoundry and TowerJazz's products, summarized in 4.2, is the sensor structure rather than the read out architecture, based on a column drain R/O with ToT capability (LF-Monopix has 8 bits dedicated and TJ-Monopix 6 bits). Concerning the sensor, LFoundry pixels are bigger and have a large fill factor, while TJ-Monopix ones have a small fill-factor electrode.

The performances of both the detectors have been tested before and after irradiation ($\sim 10^{10} n_{eq}/cm^2$) and the result is, as expected since LF-Monopix is a large fill factor electrode (capitolo ??), that LF-Monopix is more radiation hard than TJ-Monopix whereas the main degradation of efficiency in TJ-Monopix chips is due to the low electric field in the pixel corner. On the other hand one more accidental consequence of the large fill factor size in LF-Monopix (the deep p-well covers $\sim 55\%$ of the pixel area) is a significant cross-talk problem.

4.1 The sensor

TJ-Monopix1 adopted the modification I've told about in ?? that allows to achieve a full depletion region within the epitaxial layer: a planar low dose n implant is built on a high resistivity (≥ 1

	LF-Monopix	TJ-Monopix
Bulk Resistivity	p-type substrate maggiore $2\text{k}\Omega\text{cm}$	p-epi. on a low ρ substrate maggiore $1\text{k}\Omega\text{cm}$
Pixel size	$50 \times 250 \mu\text{m}^2$	$26 \times 40 \mu\text{m}^2$
Depth	$100\text{-}750 \mu\text{m}$	$25 \mu\text{m}$
Capacity	$\sim 400 \text{ fF}$	$\sim 3 \text{ fF}$
Preamplifier	CSA	Voltage
Threshold trimming	on pixel (4-bit DAC)	global threshold
Readout mode	Fast column drain	Fast column drain
Consumption	$\sim 300 \text{ mW/cm}^2$	$\sim 120 \text{ mW/cm}^2$
Threshold	$1500 e^-$	$\sim 270 e^-$
ENC	$100 e^-$	$\sim 30 e^-$

Table 4.1: Main characteristics of TJ-Monopix and LF-Monopix [8]

$\text{k}\Omega\text{ cm}$), p-type epitaxial layer.

The modification improves a lot the performances of the detector, especially after radiation, but Technology Computer Aided Design (TCAD) simulations (fig.4.2) have shown that a non uniform electric field is still produced in the substrate after the modification; since the lateral component of the electric field drops at the pixel corner (this point in figure is indicated by a star) the efficiency at the side is reduced.

In some cases, then, a second modification have been made to increase the lateral component of electric field: a portion of low dose implant has been removed creating a discontinuity in the pixel corner. Questo migliora le prestazioni del rivelatore, ma dall'altra parte non rende più possibile fornire tensione separatamente a deep p well e substrate, sennò si ha punchthrough. Moreover, to

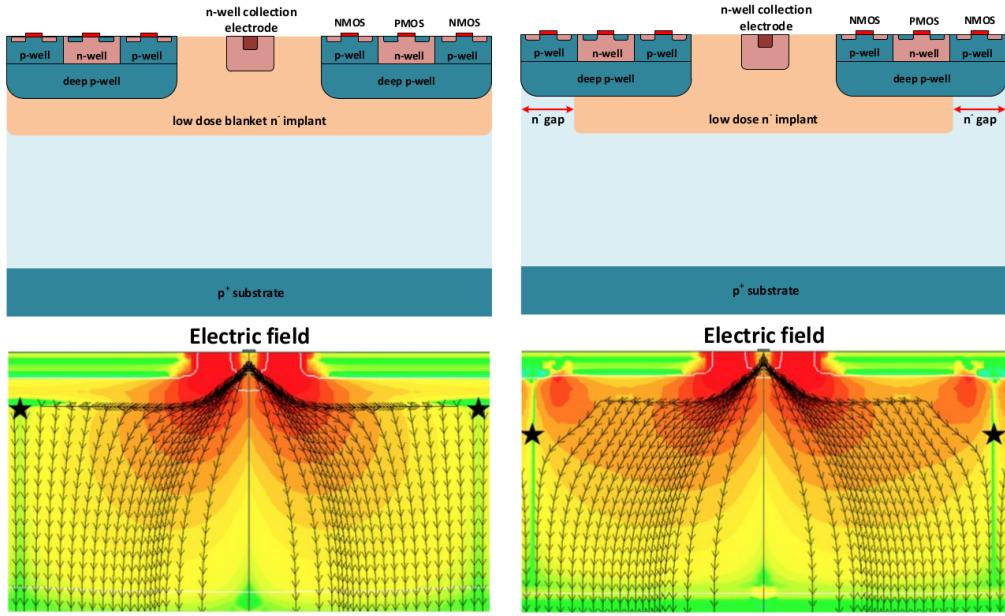


Figure 4.2: (a) The cross-section of a monolithic pixel in the TJ-Monopix 180 nm with modified process; additionally in (b) a gap in the low dose implant is created to improve the collection of charge due to a bigger lateral component of the electric field

investigate the differences in charge collection efficiency, there is a difference between pixels within the matrix: rows from 0 to 111 are fully covered by deep p-well (FDPW) under p-well near the sensor, while rows from 112 till the last 223 have a portion of deep p-well removed (RDPW).

The removing enhance the the lateral electric field component then resulting in a higher efficiency, as we'll see later.

Parameter	Value
Matrix size	
Pixel size	$26 \times 40 \mu\text{m}^2$
Depth	$25 \mu\text{m}$
BCID	40 MHz
ToT-bit	6
Power consumption	$\sim 120 \text{ mW/cm}^2$

Table 4.2

4.2 FE flavors

TJ-Monopix1 has been implemented in four different flavors, each one corresponding to a different sector on the matrix (fig. 4.3) and thus having a separate readout and data transmission, in order to explore different variations of the FE. The four flavors mainly differ in the reset input circuit. R resistenza di reset deve essere abbastanza grande da far sì che il ritorno allo zero

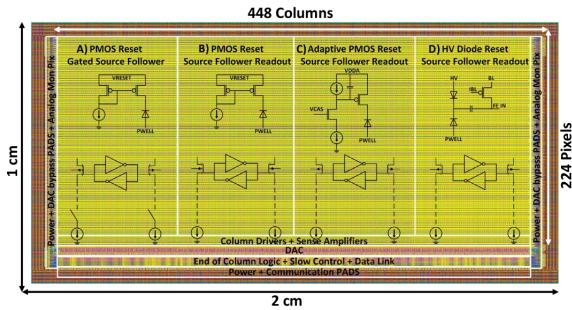


Figure 4.3

è abbastanza lento (non devi "interferire" con la tot slope e non devi più corto del tempo del preamplificatore, sennò hai perdita di segnale).

Baseline reset: all'input solitamente hai un PMOSS o un diodo;

The FE circuit 4.4 is ALPIDE-like, so it is similar to the one described in ??; a quanto già detto voglio però aggiungere due parole: come viene implementato il mascheramento dei pixels e il reset.

Prevedere un modo di mascherare gli screaming pixel, tipicamente pixels con manufacturing defects, è fondamentale per poter ridurre il rate molto alto di dati e non saturare la banda.

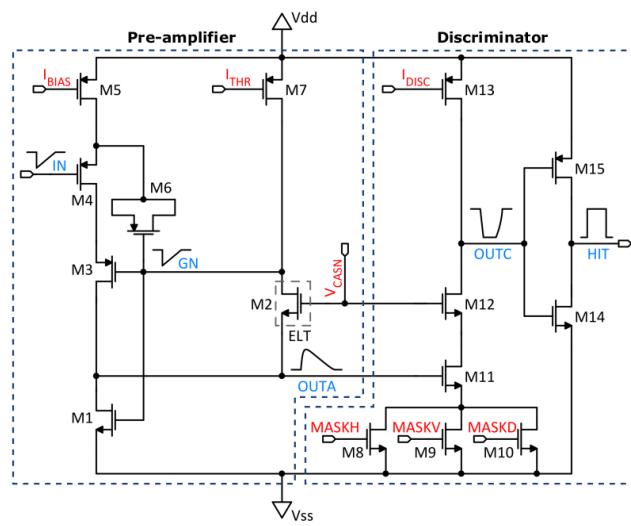


Figure 4.4

Parameter	Meaning
-----------	---------

Table 4.3

In the circuit in fig. 4.4 transistors M8, M9 and M10 implement are used to disable pixels-readout, where MASKH, MASKV and MASKD represent respectively the vertical, orizontal and diagonal coordinates of the pixel that one want to mask.
If all three transistors-signals are low, the discriminator is disabled and the pixel is masked. The masking is implemented in this way (with three cordinates instead of one) in order to avoid masking too many ghost pixels (fig. 4.5). Un modo standard che si usa di solito è allocare un registro su

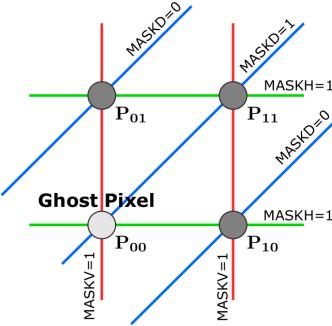


Figure 4.5

ogni pixel periphery: il vantaggio di questo modo è che si può disabilitare ogni pixel individualmente. Questo metodo pur essendo più comodo richiede less amount of area ha però come drawback che il registro può essere soggetto a SEU ¹ problema non trascurabile in acceleratori come HL-LHC adronici

The implemented approach of masking in Monopix-1 funziona però solo se il numero di pixel da mascherare non è troppo alto dato che il numero di pixel unintentionally masked ("ghost pixels") increase with the number of pixels masked.

Nel caso in cui solo due coordinate vengono utilizzate il numero di pixel unintentionally masked scales with N^2 , where N is the number of the intentionally masked; if instead three coordinates are given the ghost pixels are N^α where $\alpha \min 2$.

4.2.1 FE parameters

Descrivo un po' le misure fatte sul fe e sul significato dei vari parametri.

4.3 Readout logic and Data-packets structure

TJ monopix ha un colum drain readout proven by the ATLAS FEI3 front end chip (I. Peric et al., The FEI3 readout chip for the ATLAS pixel detector, Nucl. Instrum. Meth. A 565 (2006) 178, ed. by J. Grosse-Knetter, H. Krueger, and N. Wermes (cit. on pp. 42, 50, 60))

TJ-Monopix is a triggerless. It sends data whenever it gets hits. Only thing we can do is to record timestamp of the external triggers and correlate with the hits.

La lettura si può descrivere come una macchiana a stati in cui gli stati possibili sono READ, WAIT, DATA TRASMISSION

¹SEU = Single Event Upset, in sostanza è quando un bit ti cambia valore (da 0 a 1 o viceversa) perché una particella deposita carica nell'elettronica che fa da memoria (registro/RAM/...). Questo tipo di elettronica ha bisogno di un sacco di carica prima che il bit si "flippi" (cambi valore), infatti tipicamente per avere un SEU non basta una MIP che attraversa esattamente quel pezzo di chip in cui è implementata la memoria, ma un adrone che faccia interazione nucleare producendo più carica di quanto farebbe una MIP.

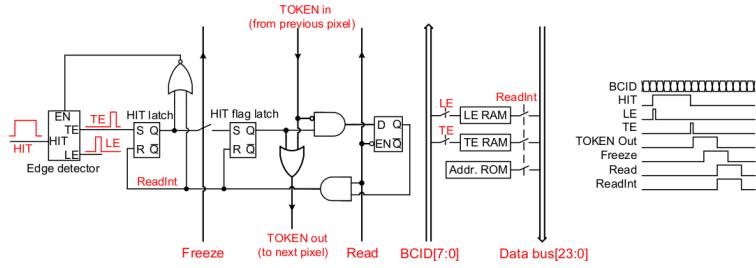


Figure 4.6

4.3.1 Dead time measurement

Descrivo come ho fatto il test. Fattore limi

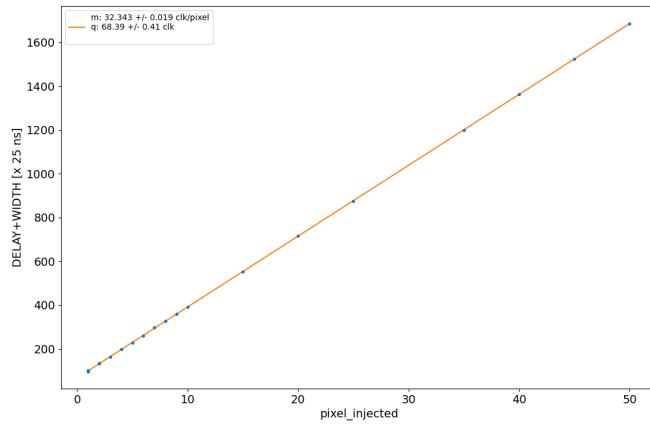


Figure 4.7

Chapter 5

Arcadia-MD1

[9] [10]

Breve introduzione analoga a Monopix1 in cui descrivo brevemente la "timeline" da SEED Matisse a Md1 e Md2

5.1 The sensor

ARCADIA-MD1 is an LFoundry chip which implements the technology 110 nm CMOSS node with six metal layer ???. The standard p-type substrate was replaced with an n-type floating zone material, that is a tecnique to produce purified silicon crystal. (pag 299 K.W.).

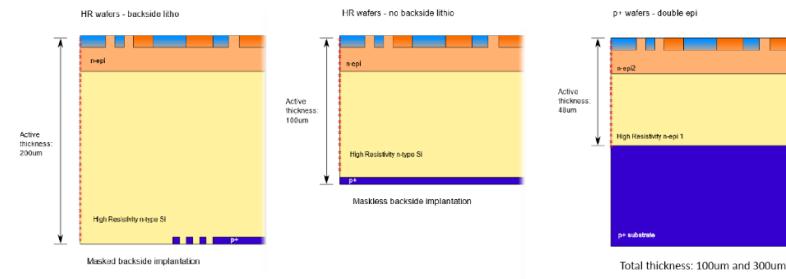


Figure 5.1

Wafer thinning and backside lithography were necessary to introduce a junction at the bottom surface, used to bias the substrate to full depletion while maintaining a low voltage at the front side. C'è un deep pwell per - priority chainseparare l'elettronica dal sensore; per controllare il punchthrough è stato aggiunto un n doped epitaxial layer having a resistivity lower than the substrate.

RILEGGI SUL KOLANOSKY COS'È IL PUNCHTHROUGHT, FLOAT ZONE MATERIAL, COME VENGONO FATTI I MAPS COME FAI LE GIUNZIONI

It is part of the cathegory of DMAPS Small electrode to enhance the signal to noise ratio.

It is operated in full depletion with fast charge collection by drift.

Prima SEED si occupa di studiare le prestazioni: concept study with small-scale test structure (SEED), dopo arcadia: technology demonstration with large area sensors Small scale demo SEED(sensor with embedded electronic developement) Quanto spazio dato all'elettronica sopra il pwell e quanto al diodo. ..

5.2 Readout logic and data structure

5.2.1 Matrix division and data-packets

The matrix is divided into an internal physical and logical hierarchy: The 512 columns are divided in 16 section: each section has different voltage-bias + serializzatori. Each section is devided in

cores () in modo che in ogni doppia colonna ci siano 1Pacchetto dei dati 6 cores. ricordati dei serializzatori: sono 16 ma possono essere ridotti ad uno in modalità spazio

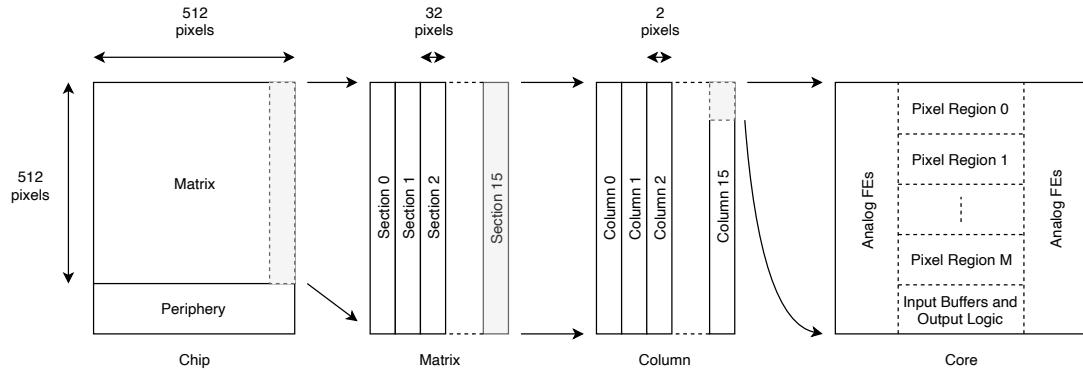


Figure 5.2

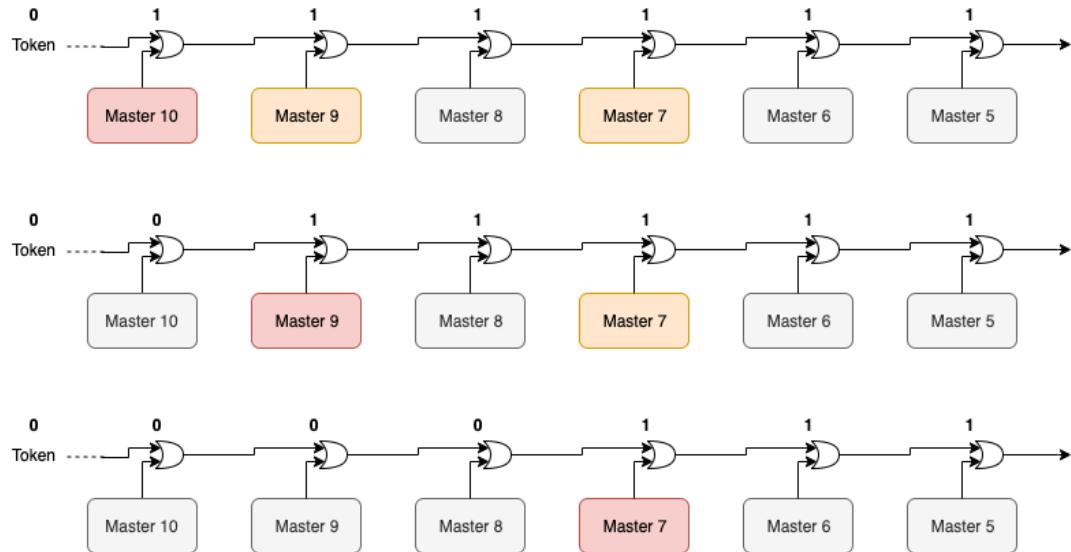


Figure 5.3

Questa divisione si rispecchia in come sono fatti i dati: scrivi da quanti bit un dato è fatto e le varie coordinate che ci si trovano dentro; devi dire che c'è un pixel hot e spieghi dopo a cosa serve, e devi accennare al timestamp

"A core is simply the smallest stepped and repeated instance of digital circuitry. A relatively large core allows one to take full advantage of digital synthesis tools to implement complex functionality in the pixel matrix, sharing resources among many pixels as needed.". pagina 28 della review.

TABELLA: con la gerarchia del chip Matrix (512x512 pixels) Section (512x32 pixels) Column (512x2) Core (32x2) Region (4x2)

Nel chip trovi diverse padframe: cosa c'è nelle padframe e End of section.

"DC-balance avoids low frequencies by guaranteeing at least one transition every n bits; for example 8b10b encoding n =5"

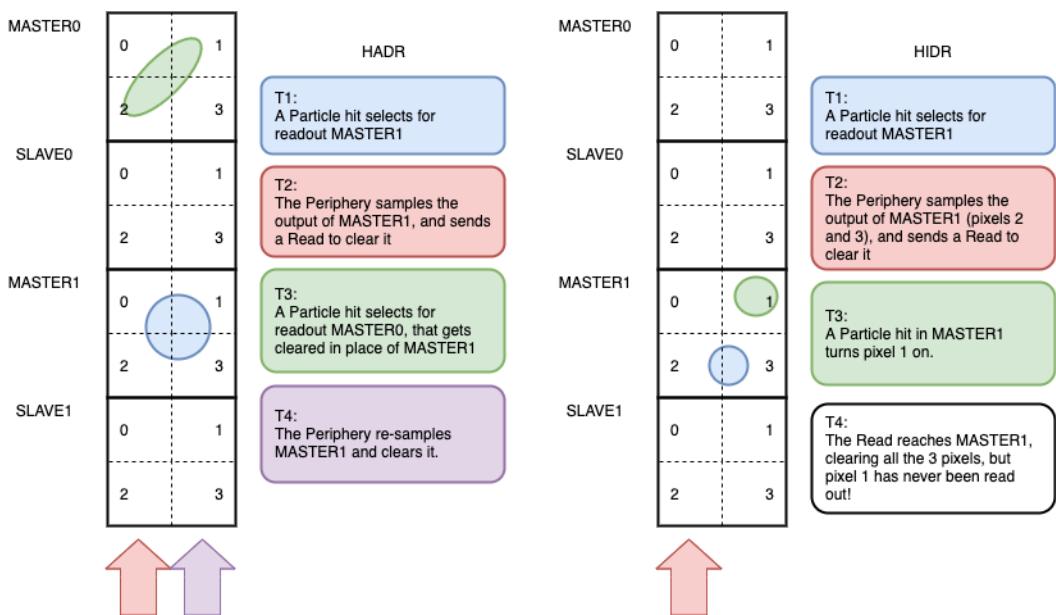


Figure 5.4

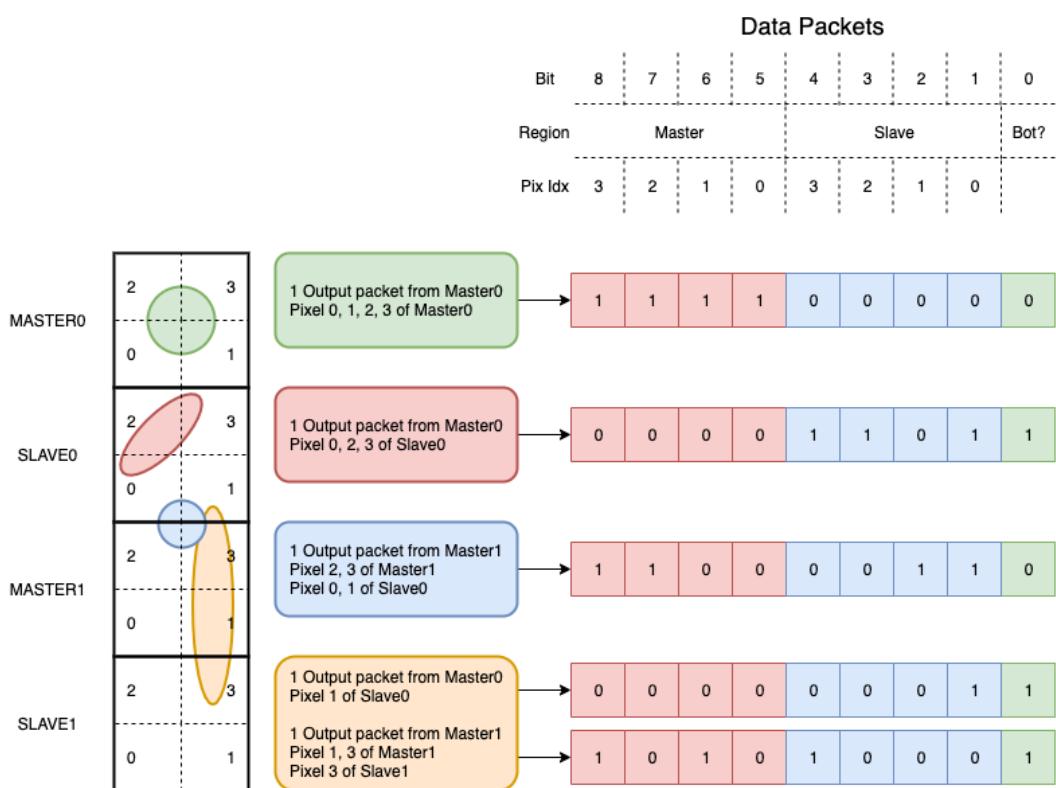


Figure 5.5

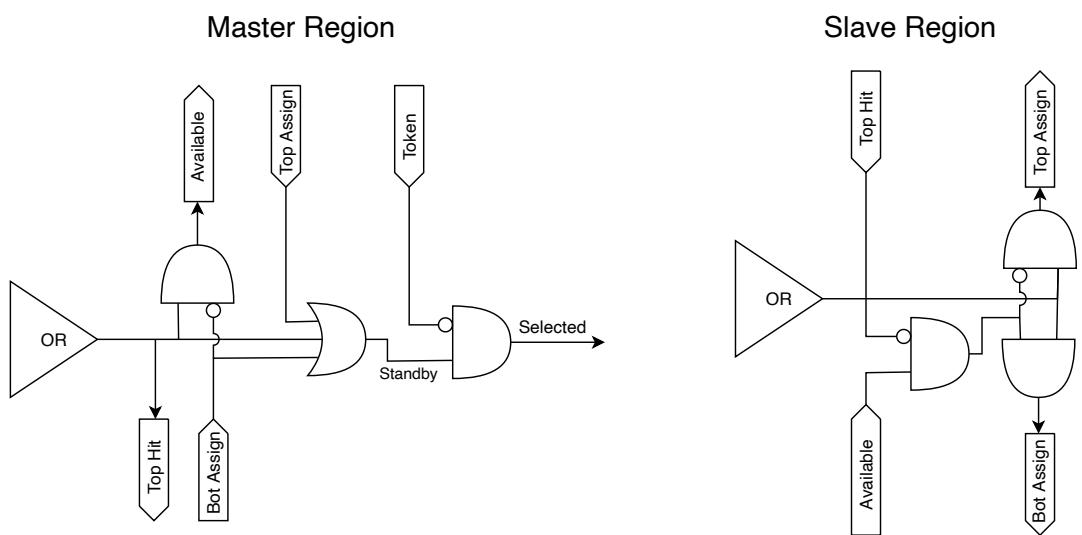


Figure 5.6

Chapter 6

Threshold and noise characterization

6.1 Threshold and noise: figure of merit for pixel detectors

The signal to threshold ratio is the figure of merit for pixel detectors.

la soglia deve essere abb alta da tagliare il rumore ma abb bassa da non perdere efficienza.
Invece di prendere il rapporto segnale rumore prendi il rapporto segnale soglia. Perchè? la soglia è collegato al rumore, nel senso che: supponiamo di volere un occupancy di 10-4 allora sceglierò la soglia in base a questo. (plot su quaderno) Da questo conto trovo la minima soglia mettibile
In realtà quello che faccio è mettere una soglia un po' più grande perchè il rate di rumore dipende da molti fattori quali la temperatura, l annealing ecc, e non voglio che cambiando leggermente uno di questi parametri vedo alzarsi molto il rate di rumore. In realtà non è solo il rumore sensibile a diversi fattori, ma anche la soglia: ad esempio la cosa classica è la variabilità della soglia da pixel a pixel.

In questo modo rumore e soglia diventano parenti.

Review pag 26.

The noise requirement can be expressed as:

Questo implica tra le altre cose che voglio poter assegnare delle soglie diverse a diversi pixel:
Drawback è dare spazio per registri e quantaltro.

Questo lascia però ancora aperto il problema temporale delle variazioni del rumore: problema per cui diventano necessarie le misure dei sensori dopo l'irraggiamento.

Per arcadia i registri (c'è un DAC) per la soglia (VCASN) si trovano in periferia. Non fare trimming sulla soglia è uno dei problemi che si sono sempre incontrati: a casusa dei mismatch dei transistor le soglie efficaci pixel per pixel cambiano tanto. La larghezza della s curve è il noise se assumi che il noise è gaussiano

Il trimming della soglia avviene con dei DAC: la dispersione della soglia dopo al tuning e dovuta al dac è:

$$\sigma_{THR,tuned} = \frac{\sigma_{THR}}{2^{nbit}} \quad (6.1)$$

dove il numero di bit cambia varia tra 3-7 tipicamente. Monopix è 7 Arcadia 6

Each ROIC is different in this respect, but in general the minimum stable threshold was around 2500 electrons (e) in 1st generation ROICs, whereas it will be around 500 e for the 3rd generation. This reduction has been deliberate: required by decreasing input signal values. Large pixels (2 104 um²), thick sensors (maggiore di 200 um), and moderate sensor radiation damage for 1st generation detectors translated into expected signals of order 10 ke, while small pixels (0.25 104 um²), thinner sensors (100 um), and heavier sensor radiation damage will lead to signals as low as 2 ke at the HL-LHC

The ENC can be directly calculated by the Cumulative Distribution Function (CDF) (scurve) obtained from the discriminator "hit" pulse response to multiple charge injections

6.2 TJ-Monopix1 characterization

Com'è fatto il set up per le misure.
FPGA BB, Chip con FE board, qualche foto

6.3 ARCADIA-MD1 characterization

Com'è fatto il set up per le misure.
FPGA BB, Chip con FE board, qualche foto

Appendix A

Pixels detector: a brief overview

A.1 Signal formation

When a charge particle passes through a pixel and loses energy by ionization a part of that energy is used to generate electron-hole pairs (an other part is used for other processes, as the lattice excitation) which are then separated by the electric field and collected at their respectively electrodes (p for holes and n for electrons)¹; by the drift of these charges, a signal i_e is generated on the electrode e as stated by the Shockley–Ramo's theorem:

$$i_e(t) = -q v(t) E_{WF,e} \quad (\text{A.1})$$

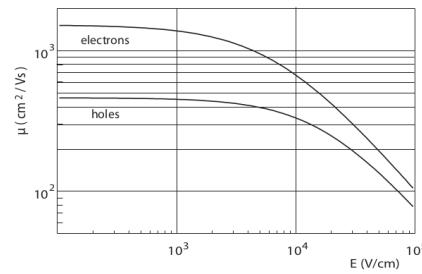
where $v(t)$ is the instantaneous velocity of the charge q and E_{WF} is the weighting field, that is the field obtained biasing the electrode e with 1V and all the others with 0V.

The drift velocity of the charge depends on the electric field and on the mobility of the particle:

$$v = \mu(E) E \quad (\text{A.2})$$

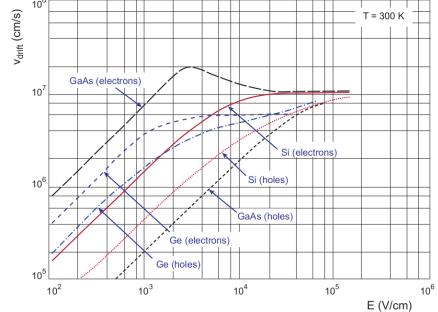
where $\mu(E)$ is a function of the electric field and is linear with E only for small E : at higher values the probability of interactions with optical phonons increases and the mobility drops and this leads to an indipendence of the velocity from the electric field (fig. A.1b).

SECONDO ME MANCA ANCORA UNA FRASE DI CONNESSIONE



(a) Typical values for electrons and holes mobility in

silicon at room temperature are $\mu_n \sim 1450 \text{ cm}^2/\text{Vs}$, $\mu_h = 500$

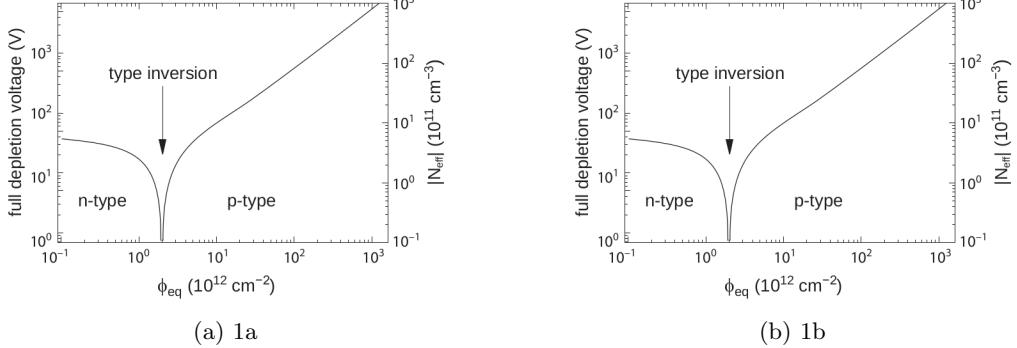


(b) Drift velocity at room temperature in different semiconductors

The average energy needed to create a pair at 300 K in silicon is $w_i = 3.65 \text{ eV}$, that is more than the mean ionization energy because of the interactions with phonon, since for a minimum ionizing particle (MIP) the most probable value (MPV) of charge released in the semiconductor is 0.28 keV/ μ , hence the number of e/h pairs is:

$$\langle \frac{dE}{dx} \rangle \frac{1}{w_i} \sim 80 \text{ e}/\text{h} \sim \frac{1.28 \cdot 10^{-2} fC}{\mu m} \quad (\text{A.3})$$

¹Even if in principle both the electrode can be used to read a signal, for pixel detectors, where the number of channel and the complexity of readout are high, only one is actually used. In strip and pad detectors, instead, is more common a dual-side readout



CON UN'INCERTEZZA CHE È RADICE DI N; ED EVENTUALEMTE SI AGGIUNGE IL FATTORE DI FANO NEL CASO DI ASSORBIMENTO TOTALE. IL FATTORE DI FANO È 0.115 NEL SILICIO

It is fundamental that pairs e/h are produced in the depleted region of the semiconductor where the probability of recombination with charge carriers is low to avoid loss of signals.

Pixel detectors are then commonly reverse biased: a positive bias is given to the *n* electrode and a negative to the *p* to grow the depletion zone in the epitaxial layer below the electrode. The width of the depletion region is related with the extern bias V_{ext} , the resistivity ρ and also with the dopant:

$$d_n \sim 0.55 \sqrt{\frac{\rho}{\Omega cm} \frac{V_{ext}}{V}} \mu m \quad (\text{A.4})$$

$$d_p \sim 0.32 \sqrt{\frac{\rho}{\Omega cm} \frac{V_{ext}}{V}} \mu m \quad (\text{A.5})$$

DA DOVE VENIVA QUESTA DIFFERENZA? DALLA MOBILITÀ MA NON MI RICORDO COME CI SI ARRIVAVA.

For that reason high resistivities wafer ($100 \Omega cm - k\Omega cm$) are typically prefered because they allow bigger deplation zone with smaller voltage bias.

A.2 Radiation damages

Radiation hardeness is a fundamental requirement for pixels detector especially in HEP since they are almost always installed near the interaction point where there is a high energy level of radiation. At LHC the ϕ_{eq} per year in the innermost pixels detector is $10^{14} n_{eq}/cm^2$; this number reduces by an order passing to the outer tracker layer. (referenza: pag 341 Wermes)

Here the high fluence of particles can cause a damage both in the substrate of the detector and in the superficial electronics.

The first one has a principal non ionising nature (non ionizing energy loss, NIEL) since it is related with the dislocation of the lattice caused by the collision with nuclei; by this fact the NIEL hypothesis states that the substrate damage is normalized to the damage caused by 1 MeV neutrons. Differently, surface damages are principally due to ionising energy loss.

DUE PAROLE IN PIÙ SUL SURFACE DAMAGE

a charge accumulation in oxide (SiO_2) can cause the generation of parasitic current with an obvious increase of the 1/f noise.

Anyway surface damages are less relevant then the previous one since with the development of microelectronics and with the miniaturization of components (in electronic industry 6-7 nm transistors are already used, while for MAPS the dimensions of components is around 180 nm) the quantity of oxide in circuit is reduced.

Let's spend instead two more other words on the more-relevant substrate damages: the general result of high radiation level is the creation of new energy levels within the silicon band gap and depending on their energy-location their effect can be different, as described in the Shockely-Read-Hall (SRH) statistical model. The three main consequence of radiation damages are the changing of the effect doping concentration, the leakage current and the increasing of trapping probability.

Changing of the effective doping concentration: is associated with the creation/removal of donors and acceptors center which trap respectively electrons/holes from the conduction band and cause a change in effective space charge density. Even an inversion (p-type becomes n-type²) can happen: indeed it is quite common and happens at not too high fluences ($\phi_{eq} 10^{12-13} n_{eq} cm^{-2}$). A changing in the doping concentration requires an adjust of the biasing of the sensor in time (eq.A.1) and sometimes can be difficult keeping to fully deplete the bulk.

Leakage current: is associated with the generation-recombination centres. It has a strong dependence with the temperature ($I_{leak} \propto T^2$), whose solution is therefore to operate at lower temperature.

Increase of trapping probability: È ASSOCIATA CON QUALE TIPO DI CREAZIONE DI LIVELLO ENERGETICO? since the trapping probability is costant in the depleted region, the collected charge decreases exponentially with the drift path. The exponential coefficient, that is the mean trapping path, decreases after irradiation and typical values are 125-250 μm and must be compared with the thickness of the depleted region which () corresponds to the mean drift path.

Different choises for substate resistivity, for junctions type and for detector design are typically made to fight radiation issues. Some material with high oxygen concentration (as crystall produced using Czochralki (Cz) or float-zone (Fz) process (CONTROLLA SE SONO LORO QUELLI GIUSTI)) for example, show a compensation effect for radiation damage; an other example is the usage of n+ -in-p/n sensors (even if p+ -in-n sensors are easier and chieper to obtain) to get advantage of inversion/to have not the inversion (since they are already p-type). After inversion the n+p boudary, coming from n+ in-n, but to keep using the sensor the depletion zone still must be placed near the diode³.

Radiation damage in CMOS circuits is entirely due to charge carriers generated by ionization in the dielectric layers of the process

²L'INVERSIONE OPPSTA NON CE L'HAI PERCHÈ L'INVERSIONE È ASSOCIATA AD UN CAMBIO DELLA CONCENTRAZIONE DA ... A ...
E COME MAI L'ALTRO NON È FAVORITO?

³With inversion some isolation process of electrodes can become important and p-spray/p-stop tecnique can eventually be applied. PERCHÈ CON L'INVERSIONE TI POTREBBE SERVIRE UNA TECNICA DI ISOLAMENTO?

Appendix B

FLASH radiotherapy

La radioterapia si usa nel 60 per cento dei pazienti, sia come cura che come trattamento palliativo. Si associa spesso ad altre cure e si può fare prima/durante/dopo un intervento.

Si può fare in modi diversi: da dentro (brachytherapy) oppure da fuori (quella standard). Un requisito importante è la delinazione del target (non vuoi rischiare di beccare i tessuti sani), per cui prima tipicamente si fanno esami di imaginig del tumore. Tipicamente anche gli acceleratori stessi per la terapia sono provvisti di radiografia.

Un problema dei fotoni ad esempio è che il loro rilascio di dose è lineare, per cui danneggia anche i tessuti sani. Il problema dei protoni invece è che hanno un picco troppo stretto per cui non puoi coprire grosse zone e soprattutto se sbagli rischi davvero di danneggiare molto i tessuti sani.

B.1 Cell survival curves

Curva di efficacia del trattamento in funzione della dose:

$$\frac{S(D)}{S(0)} = e^{-F(D)} \quad (\text{B.1})$$

dove $F(D)$

$$F(D) = \alpha D + \beta D^2 \quad (\text{B.2})$$

dove α e β rappresentano due tipi di danno diversi: coefficients, experimentally determined, characterizing the radiation response of cells. In particular, alpha represents the rate of cell killing by single ionizing events, while beta indicates the maximal rate of cell killing by double hits observed when the repair mechanisms do not activate during the radiation exposure. Si ottiene una curva di sopravvivenza dove si vede la possibilità delle cellule di autoripararsi. A basse dosi infatti le cellule possono ripararsi.

Per introdurre l'effetto FLASH introduco prima la therapeutic window.

TCP è la tumor control Probability che indica la probabilità delle cellule del tumore di essere uccise dopo una certa dose (con in riferimento a dose in acqua)

Se una media di $\mu(D)$ di cellule di tumore are killed con una dose D , la probabilità che n cellule sopravvivono è data da $P(n|\mu)$ poisson:

$$P(n|\mu) = \frac{\mu(D)^n e^{-\mu(D)}}{n!} \quad (\text{B.3})$$

$$TCP(D) = P(n=0|\mu(D)) = e^{-\mu(D)} \quad (\text{B.4})$$

D'altra parte hai una probabilità di fare danno su normal tissue NTCP Normal Tissue Complication Probability, che rappresenta il problema principale e che limita la massima radiazione erogabile. Una scelta bilanciata si applica guardando a questi due fattori; si usa il therapeutic index definito come TCP/NTCP.

La cosa ottimale è ampliare la finestra del therapeutic ratio.

CONV-RT 0.01-5 Gy/min. A typical RT regime today consists of daily fractions of 1.5 to 3 Gy given over several weeks.

Nell Intra operative radiation therapy (IORT), where they reach values respectively about 20 and 100 times greater than those of conventional radiation therapy.

FLASH vuole ultrahigh mean dose-rate (maggione di 40 Gy/s) in modo da ridurere anche il trattamento a meno di un secondo.

B.2 FLASH effect

Ci sono due effetti che affect the fsh effect and la sua applicabilità: Dose rate effect e oxygen

Cellule che esibiscono hypoxia (cioè cellule che non hanno ossigeno sono radioresistenti); al contrario normoxia e physoxia non lo sono. la presenza di ossigeno rende la curva steeper indicando che lo stesso danno si raggiunge a livelli di dose più bassi rispetto al caso senza ossigeno.

FIGURA con una curva a confronto con e senza ossigeno.

Typically, the OER is in the order of 2.5–3.5 for most cellular systems

Quindi si vogliono sfruttare questi effetti per diminuire la tossicità sui tessuti sani

Bibliography

- [1] W. Snoeys et al. “A process modification for CMOS monolithic active pixel sensors for enhanced depletion, timing performance and radiation tolerance”. In: (2017). DOI: <https://doi.org/10.1016/j.nima.2017.07.046>.
- [2] D. Kim et al. “Front end optimization for the monolithic active pixel sensor of the ALICE Inner Tracking System upgrade”. In: *JINST* (2016). DOI: doi:10.1088/1748-0221/11/02/C02042.
- [3] H. Kolanoski and N. Wermes. *Particle Detectors: Fundamentals and Applications*. OXFORD University Press, 2020. ISBN: 9780198520115.
- [4] E. Mandelli. “Digital Column Readout Architecture for the ATLAS Pixel 0.25 um Front End IC”. In: (2002).
- [5] M. Garcia-Sciveres and N. Wermes. “A review of advances in pixel detectors for experiments with high rate and radiation”. In: (2018). DOI: <https://doi.org/10.1088/1361-6633/aab064>.
- [6] M. Dyndal et al. “Mini-MALTA: Radiation hard pixel designs for small-electrode monolithic CMOS sensors for the High Luminosity LHC”. In: (2019). DOI: <https://doi.org/10.1088/1748-0221/15/02/p02005>.
- [7] M. Barbero. “Radiation hard DMAPS pixel sensors in 150 nm CMOS technology for operation at LHC”. In: (2020). DOI: <https://doi.org/10.1088/1748-0221/15/05/p05013>.
- [8] K. Moustakas et al. “CMOS Monolithic Pixel Sensors based on the Column-Drain Architecture for the HL-LHC Upgrade”. In: (2018). DOI: <https://doi.org/10.1016/j.nima.2018.09.100>.
- [9] L. Pancheri et al. “A 110 nm CMOS process for fully-depleted pixel sensors”. In: (2019). DOI: <https://doi.org/10.1088/1748-0221/14/06/c06016>.
- [10] L. Pancheri et al. “Fully Depleted MAPS in 110-nm CMOS Process With 100–300-um Active Substrate”. In: (2020). DOI: 10.1109/TED.2020.2985639.