Table of Contents

Introduction

The advancement of High Energy Physics (HEP) needs more and more powerful equipments. The LHC (Large Hadron Collider), a proton collider of the highest energy level up to now (about 7 TeV for protons-protons), is under contracture and it will be put into operation by the end of April 2008. Corresponding to the on going LHC project, one of the next generation linear collider, called ILC (International Linear Collider), is designed to have a center-of-mass energy up to at least 500 GeV, aiming at measuring the properties of hypothetical Higgs particles around 100 to 300 GeV. Thus the future ILC requires a well designed and more accurate vertex detector, which determines the total performance of the future ILC.

Very strict requirements, such as high readout speed, high spatial resolution, tolerance to radiation hardness and low-power dissipation, have to be achieved simultaneously by the vertex detector. Moreover, very compact devices are needed since room in the vertex detector is limied: located very close from the interaction point, slightly more than a centimeter. Although designs based on different technologies (microstrip, CCD (Charge Coupled Device), hybrid pixel sensors) have been successfully implemented for various applications during the development of particle physics, it is very difficult to find out an appropriate solution that reached all the requirements. However, the advancements in the electronic technology provide us with a new type of device which seems to be a good candidate: the Monolithic Active Pixel Sensors (MAPS), realized by the standard Complementary Metal-Oxide-Semiconductor (CMOS) technology.

Firstly invented in the beginning of 1960's even before the invention of CCD [1], great advances had been done in the development of CMOS APS (Active Pixel Sensor) in the early '90s. Thanks to the improvements in the techniques of microelectronics, pixel size is largely reduced (down to several μm) and in-pixel signal processing circuits realized active elements, which reduce significantly the noise level. As a result, satisfying performances, similar to CCD sensors, are achieved with respect to dynamic range and readout noise but with much faster readout speed, vast on-chip functionalities, lower system power consumption and the reduction of system cost [2].

The studies of the CMOS APS for the vertex detector of the future ILC are firstly carried on by IPHC (Institut Pluridisciplinaire Hubert Curien) at Strasbourg in 1999. Since then, a series of MIMOSA (Minimum Ionizing particle MOS Active pixel sensor) prototypes have been developed. The tests of these prototypes show very encouraging results: low noise of the sensors, high detection efficiency, high spatial resolution and good radiation resistance [3] [4] [5] [6]. All these results indicate that, for MIPs detection, the technique of MAPS works very efficiently and provides excellent tracking qualities.

In the year 2001, the team of CEA-SACLAY (DAPNIA) started research activities on this field, cooperating with the IPHC team with the goal of design of new prototypes which should reach the detection performance demonstrated by the former MIMOSA series but with important new aspects:

- realize pixel level Correlated Double Sampling (CDS) operation;

- reach a high readout speed, about 20∝s/frame, there are 128 lines in one frame;
- integrate column level comparator together with the pixel array and realize 1-bit binary output.

My works focus on this framework: to study the characteristics of the new prototypes and to develop new integrated analog and digital circuits that improve the prototype performance.

In the first chapter, after a brief review of the development of the High Energy Physics, the requirements for the vertex detector of the future ILC are presented in general. Then, after a brief introduction on available technologies, the principle of APS for the visible light is given. Then, the difference and the difficulty of charged particle detection are presented.

In Chapter 2, after a general description of sources of noise in MAPS, the noise for MIPs (Minimum Ionizing Particles) detection is emphasized. The development of MIMOSA series is briefly reviewed. Then, the architecture of the two new prototypes, MIMOSA 8 and MIMOSA 16, is presented. The working principles of two types of pixels (DC-coupling and AC-coupling) and column level discriminator are presented in detail. At the end of this chapter, a study of DC-coupling pixel performance, realized under CADENCE environment, is given, enlightling the important design points of the DC-coupling pixel.

The laboratory tests are carried on with a ^{55}Fe X ray source (energy peaks: 5.9 keV and 6.4 keV). The corresponding tests results and detailed analyses are given in Chapter 3. The important pixel parameters, such as the noise performances, SNR (Signal-to-Noise Ratio), CVF (Charge-to-Voltage conversion Factor) and CCE (Charge Collection Efficiency), are carefully studied as the function of main clock frequency. The results of analog outputs and digital outputs are compared. The algorithms used for analysis are also described in this chapter.

The beam tests results are given in Chapter 4. Two beam tests are carried on. One is made at DESY in 2005 with a 5 GeV electrons beam. The SNR and charge collection performance under real high energy charged particles environment are studied. Detection efficiency is studied for both analog outputs and digital outputs. For the analog outputs, the detection efficiency is studied as the function of the SNR value for seed pixels. For the digital outputs, the detection efficiency is studied as the function of discriminator threshold cuts. The average hit multiplicity and the fake hit rate of the digital outputs are also studied. The influences of the average hit multiplicity and the fake hit rate on the detection efficiency are discussed.

In order to improve the spatial resolution, a column level 5-bit ADC based on successive approximation architecture is developed to replace the actual column level discriminator. The design detail can be found in Chapter 5. Preliminary test results are given and the key points related to its performance are discussed. A new column level discriminator is also described in this chapter. Using a static latch, the architecture is simplified.

At last, the conclusion and the researches prospectives are given in Chapter 6.

Reference

[1] E. R. Fossum, “CMOS Image Sensors: Electronic Camera-On-A-Chip,” *IEEE Trans. Electron Devices*, vol. 44, NO. 10, Oct. 1997, pp. 1689-1698.

[2] M. Bigas, E. Cabruja, J. Forest and J. Salvi, “Review of CMOS image sensors,” *Microelectronics Journal* 37(5), 433-451, May 2006.

[3] G. Deptuch et al., “Design and Testing of Monolithic Active Pixel Sensors for Charged Particle Tracking”, *IEEE Trans. Nucl. Sci.*, vol. 49, April 2002, pp 601-610.

[4] G. Deptuch, “New generation of monolithic active pixel sensors for charged particle detection”, Ph.D. Thesis, Strasbourg, 2001.

[5] G. Deptuch et al., “Development of monolithic active pixel sensors for charged particle tracking”, *Nucl. Instrum. Meth. A 511*, 2003, pp. 240-249.

[6] W. Dulinski et al., “CMOS Monolithic Active Pixel Sensors for minimum ionizing particle tracking using non-epitaxial silicon substrate,” *IEEE Trans. Nucl. Sci.*, vol. 51, Aug. 2004, pp 1305-1310

Chapter I

Vertex Detector and CMOS Active Pixel Sensors

1.1 The development of the particles physic

1.1.1 Early development (before ~ 1930)

Before the year 1930, the only elementary particles known were the proton, the electron and the photon. All of them have an undulatory behaviour (Quantum Mechanics). Integer and half-integer numbers are associated to them (charge, spin). The core, for example, was supposed to be composite but, its inner structure was unknown: it was very difficult to interpret it in terms of protons and electrons, although with the α considered as a possible supplementary component.

During this period, the known interactions are the gravity (negligible on the atomic or subatomic scale) and the electromagnetism, which explain very well the atomic levels, their stability and the properties of the radiation. The relativistic effects are so low that they are generally considered as perturbation to Hamiltonian (development in power of v/c, which is the power of the operator p). The inner cohesion of the core and the nuclear disintegration cannot be theoretically explained in the same way. Thus, conservation laws were used in all the case (without exception):

- Monentum, energy;
- Kinetic momentum (including the spins);
- Electric charge.

At the end of the year 20, a formalism made the synthesis of the known theory: the Relativistic Quantum Mechanics, which introduces the 1/2 spin of the electron (Dirac equation) and gives the ordinary Quantum Mechanics back to the approximation non-relativist. However, the problem of free particles with negative energy states is still unexplained, bringing people to anti-particle notion.

1.1.2 Enlarging and proliferation of the particle family

From the years 30 to 60, numbers of important discoveries enlarged the particle family.

Anderson discovered anti-electron (positron) in 1933 by the production of e+ e- pairs in the cosmic radiation.

In 1932, Chadwick discovered the neutron, describing the composition of the core. The couple of nucleons (p and n) are considered as a doublet of "isospin" (Heisenberg). The strong interaction between nucleons is "isotrope" in the "isoespace" of the nucleons. The core theory began to develop.

In order to explain the obvious non-conversation of the kinetic energy and moment in the β disintegration, the neutrino hypothesis of 1/2 spin was proposed by Pauli in 1930. In 1934, Fermi introduced the weak interaction model, based on the idea that the electron and the neutrino are two states of the charge of the lepton. The β disintegration is then a creation of lepton-antilepton pair with change of the nucleon charge (whereas the electromagnetic interaction does not change the charge).

In 1934, Yukawa gave the hypothesis that meson is quantum of strong field (similarly, the photon is the quantum of the electromagnetic field). The short range of this interaction (about 2 fm) indicates that the mass of the meson is around 100 MeV.

In 1938, the muon was discovered in the cosmic radiation. Because it interacts rarely with the core, it is classified as a new type of lepton.

In 1947, the mesons π+ and π- were also discovered in the cosmic radiation. The π0 was lately discovered and then the "strange" particles (with long life time).

Since 1952, the accelerators with new types of detectors allowed us to observe numerous different interactions. The antiproton was observed firstly in 1955 and then the antineutron was also observed in experiment. The interactions of the neutrinos were detected in 1956. Two distinct variations are shown: the ν_e (deintegration b) and the ν_μ. Therefore, we have two doublets of the leptons. Numbers of hadrons were also found. The hadrons are sensible to strong interaction and lots of them, called "resonances", have a very short lifetime.

The development of the theory for the particle physics is adapted to the numerous experimental discoveries.

The second quantification introduced the operators of creation and destruction, firstly for the field quanta (the photons emission and absorption theory), then for the particles and antipaticles. We completed the Theory Quantum of Field (relativist), which was constructed from Lagrangiens, by using the electromagnetism as the model (generalising the notion of gauge invariance).

The perturbative development for the calculation of energy level and the amplitude of the transition makes the remarkable success of the Quantum ElectroDynamics (QED). The fine structure constant α is very small, allowing precise calculations to the small order. However,

the treatment of the interaction with strong coupling is very complicated, so that there is no precise theory of strong interactions*.

1.1.3 Standard Model

The theory of particle physics continues to complete and numbers of new particles were discovered consecutively since 1960.

The quarks model, proposed by Gell-Mann in 1962, explains all of the known hadrons with three types of flavor: u, d and s. This model previewed the existence of Ω^-, which was discovered lately. The fourth quark (c) and the fifth quark (b) were discovered in 1974 and 1976, respectively. In order to form three doublets, the sixth quark (t) was predicted and was discovered in 1995.

The third charged lepton (τ) was discovered in 1976. It is associated to neutrino ν_τ (supposed because of the lack of energy and the kinetic moment).

Weinberg and Salam unified the electromagnetic and weak interactions in 1971. They also gave the prediction of the quanta of the weak field: W^+, W^- and Z^0, which were observed in 1983. These particles mediate the weak nuclear interactions between particles of different flavors (all quarks and leptons). They are massive. The Z^0 is more massive than the W^+ and W^-, which are equally massive. It is of interest to note that the interactions involving the W^+ and W^- gauge bosons act on exclusively left-handed particles. The right-handed particles are completely neutral to the W bosons. Furthermore, the electric charge carried by W^+ and W^- bosons are +1 and –1 respectively, which make them susceptible to electromagnetic interactions. The electrically neutral Z^0 boson acts on particles of both chiralities, but preferentially on left-handed ones. The weak nuclear interaction is unique in that it is the only one that selectively acts on particles of different chiralities; the photons of electromagnetism and the gluons of the strong force act on particles without such prejudice. These three gauge bosons along with the photons are grouped together which collectively mediate the electroweak interactions.

The strong interaction theory represents the interactions between quarks and gluons as detailed by the theory of Quantum ChromoDynamics (QCD). The massless eight gluons mediate the strong nuclear interactions between color charged particles (the quarks). Each of the eight carry combinations of color and an anticolor charge [1] enabling them to interact among themselves.

The Standard Model of particle physics is a theory which describes three of the four known fundamental interactions between the elementary particles that presented above. It is a quantum field theory developed between 1970 and 1973 which is consistent with both quantum mechanics and special relativity. To date, almost all experimental tests of the three

* In the year 90: things maybe change.

forces described by the Standard Model have agreed with its predictions. The only unknown element in the Standard Model is the Higgs boson, which is predicted by the electroweak unification mechanism. The mass of the Higgs boson is hard to determine. It is the only Standard Model particle not yet observed, but plays a key role in explaining the origins of the mass of other elementary particles, in particular the difference between the massless photon and the very heavy W and Z bosons. Elementary particle masses, and the differences between electromagnetism (caused by the photon) and the weak force (caused by the W and Z bosons), are critical to many aspects of the structure of microscopic (and hence macroscopic) matter; thus, if it exists, the Higgs boson has an important effect on the physics theory.

The following table (Tab. 1.1) describes different components of the Standard Model.

STADARD MODEL				
Family	**Symbol (part./antipart.)**	**Electrical charge**	**Lifetime (s)**	**Interaction type**
Leptons (3 doublets) spin = 1/2	$\nu_e / \overline{\nu_e}$	0	stable?	weak
	e^-/e^+	-1/+1	stable	weak and e.m.
	$\nu_\mu / \overline{\nu_\mu}$	0	stable?	weak
	μ^-/μ^+	-1/+1	2.2×10^{-6}	weak and e.m.
	$\nu_\tau / \overline{\nu_\tau}$	0	stable?	weak
	τ^-/τ^+	-1/+1	0.3×10^{-12}	weak and e.m.
Quarks (3 doublets) spin = 1/2	$d/\overline{d}$	-1/3 and +1/3	0.3	strong, weak and e.m.
	$u/\overline{u}$	+2/3 and -2/3		
	$s/\overline{s}$	-1/3 and +1/3	$\sim 10^{-8}$	strong, weak and e.m.
	$c/\overline{c}$	+2/3 and -2/3	$\sim 10^{-12}$	
	$b/\overline{b}$	-1/3 and +1/3	$\sim 10^{-12}$	strong, weak and e.m.
	$t/\overline{t}$	+2/3 and -2/3	$\sim 10^{-25}$	
Field quanta spin = 1	γ	0	stable	e.m.
	W^+/W^-	+1/-1	$\sim 10^{-25}$	weak
	Z	0	$\sim 10^{-25}$	weak
	g	0		strong
Boson Higgs spin = 0	H^0	0	?	
	$H^{\pm}$?	+1/-1		

Table 1.1: The Standard Model.

1.2 Vertex Detector for the Future ILC

1.2.1 General collider conception

In particle physics, the knowledge of elementary particles is done by the studies of particle interactions, obtained by accelerating particles to very high kinetic energy and letting them impact on other particles. Physical reactions happen for sufficiently high energy and the accelerating particles can be transformed into other particles. Detecting these products gives insight into the physics involved.

There are two possible setups to do such experiments:

A. Fixed target setup

A beam of particles is accelerated with accelerator, and as collision partner, a fixed target is placed into the path of the beam. With the accelerator, the corresponding particles will be firstly accelerated to a high level energy (a few of GeV for example) which means that the beam moves at a very high speed. After acceleration, this high speed beam will be used as the bullet to hit the stable target. By the impact, the original solid and stable nucleus of the target will then be shattered. New particles are generated around the interaction point. Therefore, if we could capture all the particles emitted around the interaction point, the knowledge of the inner structure of atomic nucleus could be obtained by analyzing experimental data. An accelerator of 20 MeV has been realised in the year forties and the accelerator of energy level higher than 6.5 GeV has been achievable in 1955 at Berkeley. The advance of accelerator technology greatly improves the knowledge of particles and the quarks level has been reached in 1968 at SLAC (Standford). Since then, hundreds of new particles have been discovered and the mystery of nucleus' inner structure has been discovered step by step. All of these research results are gradually completing the quantum theory and largely enrich the recognizing of the mysterious micro world for the human beings.

B. Collider

In a collider, two beams of particles are accelerated and the beams are directed against each other, so that the particles collide while flying into opposite directions. Different types of cylindrical detectors are then equipped around the interaction point. The advantage of collider is that the effective impact energy is proportional to that of the accelerated particle (in the case of simple accelerator, it is proportional to the root of the energy of the accelerated particle). Therefore, the impact energy can be efficiently improved and the particle beam with very high energy level could be acquired. For this reason, nowadays the collider is widely used in HEP experiments.

Colliders may either be circular accelerators (ring accelerator) or linear accelerators (linac).

In the circular accelerator, the particle beam is typically bent into a ring using electromagnetic field and particles are accelerated by accelerating structure (HF cavities) until they reach sufficient energy (example: Large Electron Positron Collider at CERN).

But this concept is not suitable for reaching even higher energies, as electrons radiate electromagnetic energy when forced on a circular path. The related energy loss increases by a factor 16 when doubling the particle energy. Therefore the only way to reach electron energies substantially above 100 GeV is by accelerating them on a straight line. This leads directly to the concept of a linear collider, first proposed by M. Tigner in 1965. In this concept, electrons and positrons are accelerated in opposite directions in two linear accelerators and made to collide in the middle of a detector. Each linear accelerator consists mainly of large series of electromagnetic radiofrequency cavities, which efficiently generate the required electric fields to accelerate particles.

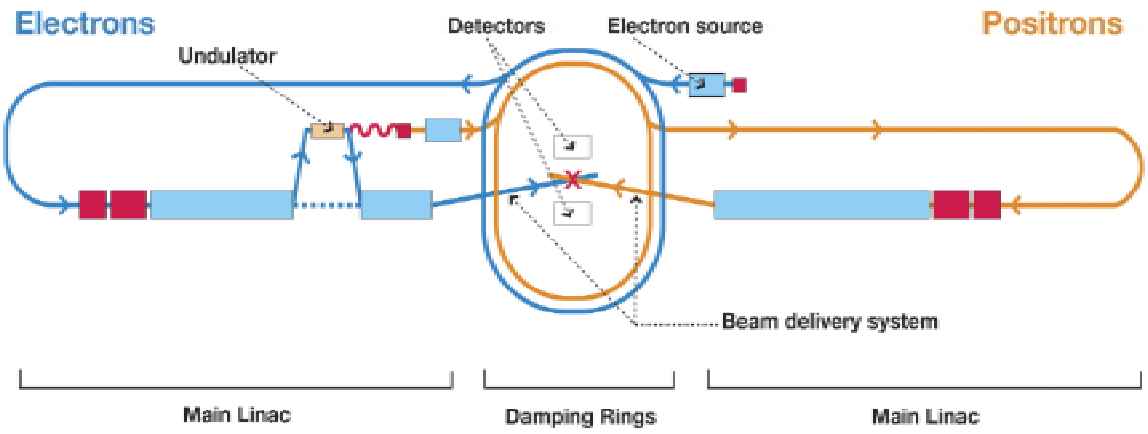
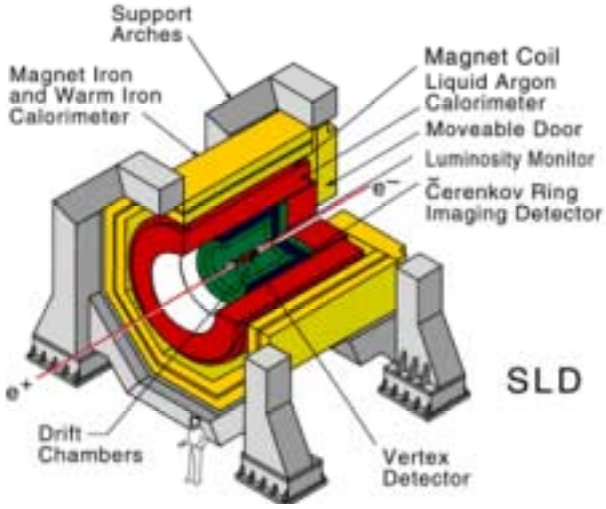


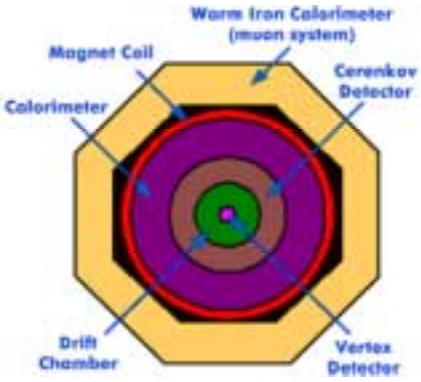
Figure 1.1: General structure of a e- e+ linear collider.

Linacs e+ e- are very widely used now. They are also used to provide an initial low-energy kick to particles before they are injected into circular accelerators. The longest linac in the world is the Stanford Linear Accelerator, SLAC, which is 3 km long. SLAC is a linear electron-positron collider of 100 GeV. Future linear collider should attain the energy level of about 500 GeV. As an example, Fig. 1.1 shows the general structure of e^- e^+ International Linear Collider (ILC project).

In both types of accelerators, during the collision, it is necessary to detect with precision the track of a passing particle, to measure their momentum and/or their energies. For all these measurements, different detectors dedicated to different missions are installed around the interaction point. For example, the complete detector (proposed for ILC in the SLD model), from external layer to inner layer, is made up of Muon chambers, Magnet solenoids, Calorimeter (Electromagnetic and Hadronic), Cerenkov detector, drift chamber and Vertex detector (Fig. 1.2).



(a) Global view of the interaction point and the surrounding detectors.



(b) Cross section of the interaction point.

Figure 1.2: The model of SLD (SLAC LARGE DETECTOR).

As in Fig. 1.2, the vertex detector is always a very small component and its diameter is generally only tens of centimeters in order to approach as close as possible to the interaction point. In spite of its size, the vertex detector is indeed a very accurate measurement tool. Very high spatial resolution (a few μm) and tracking efficiency (more than 99%) are required for the vertex detector in order to obtain better tagging efficiency of the jet flavor. The interaction point is concentrically surrounded by the vertex detector, consisting of several layers of small sensors. By extrapolating tracks measured by different layers of these sensors, vertex reconstruction is performed. In order to get better vertex resolution and better flavor tagging efficiency, the vertex detector should have the following features:

- The length of the extrapolation should be as short as possible, so that the most inner layer should be as close as possible to the interaction point (several centimeters in general).

- To minimize the effect of multiple scattering, the thickness (in radiation length) of the sensors and supporting ladders, particularly that of the most inner layer, should be as thin as possible.

- The sensors should be two-dimensional (pixel type) sensors with high granularity in order to separate near-by tracks in a collimated jet. Of course spatial resolution is a strong requirement.

As a result, a well designed vertex detector is crucial for the full reconstruction of collision events.

1.2.2 Hadron and Lepton Collider: LHC and ILC

The under contracture LHC (Large Hadron Collider) is a proton collider of highest energy level up to now (about 7 TeV for protons-protons) and it will be put into operation in the end of April 2008. As the proton is composed of three quarks and gluons, scattering process of proton continues with energy up to several TeV when two protons hit each other. Hence, the results are very complicated to study and it is very difficult to reconstruct all the events. For this reason, a special trigger system must be equipped in order to get useful information on a relative narrow one depending on the purpose of different experiments. The principle is different in the case of lepton collider.

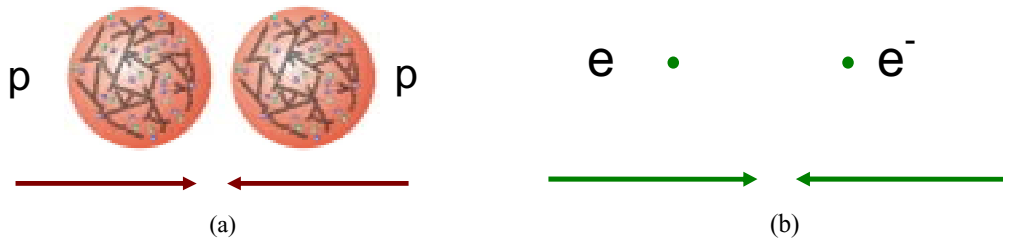


Figure 1.3: Collision of different particles.
(a) Collision of protons. (b) Collision of electrons.

Instead of protons, simple leptonic particle such as electron can be used for collision. For this reason, a well defined centre-of-mass energy could be applied to the collision and the experimental environment is much cleaner than that of hadron collider. Furthermore, the collision events are fully reconstructable, which indicates that signals of unexpected new physics could be found by this measurement. After the above general description, the hadron and the lepton colliders interplay between each other. While the hadron collider produces new heavy states, the lepton collider accomplishes very precision measurements which are very good complementary results to that of hadron collider's. It is desirable that these two types of collider work together for the exploration of the unknown world.

Corresponding to the going LHC project, the next generation of linear collider is ILC (International Linear Collider) with center-of-mass energy up to at least 500 GeV and one of its main objective is to measure the properties of hypothetical Higgs particles around 100 to 300 GeV.

As mentioned above, a well designed vertex detector determines the total performance of a collider. The future ILC requires a more precise vertex detector. My works emphasises on the study of the new type sensors suitable for this application. The sensors constructed by new technologies, the CMOS image sensors, will be carefully considered.

1.2.3 The Vertex Detector of the Future ILC

A. Introduction

As presented at section 1.2.1, the vertex detector, a very compact tracking device about the size of 40 cm × 6 cm × 6 cm, surrounds the interaction region at the heart of the massive ILC detector system. Working like a 3-D camera, the vertex detector measures the tracks of outgoing particles with micron precision. The performances of the new vertex detector for the future ILC will be well beyond what was ever achieved so far.

Thanks to its high luminosity, to well defined interaction parameters and to operating conditions which are much less demanding than for hadron colliders, the future ILC allows for an unprecedented sensitivity to physical processes. The phenomena involving very short-lived particles can be scrutinized by using the future ILC. These particles include taus, cousins of the electrons, charmed particles, made of c-quarks, and beauty particles, made of b-quarks. During a collision, they firstly fly over distances as short as a few hundred micrometers from the interaction point and then decay into secondary - more stable - particles, which can be detected in the vertex detector. An out standing spatial resolution, provided by a high precision vertex detector, is required for reconstructing the decay vertex from these particles. This resolution has to be complemented by a very light and thin material budget, ensuring that particles traversing the detector will only be slightly affected by multiple scattering, which affects the direction of particles thus fooling the reconstruction of their trajectory and their spatial origin.

Mainly coming from electromagnetic interactions between the two incoming beams, the background in the vertex detector results an overwhelming electron and positron cloud, which fires several thousands of pixels every microsecond. As a result, a major technical challenge is then to read out all the information of the vertex detector 200 - 250 times per second, if not more, to maintain the occupancy rate to a reasonable level of a few percent.

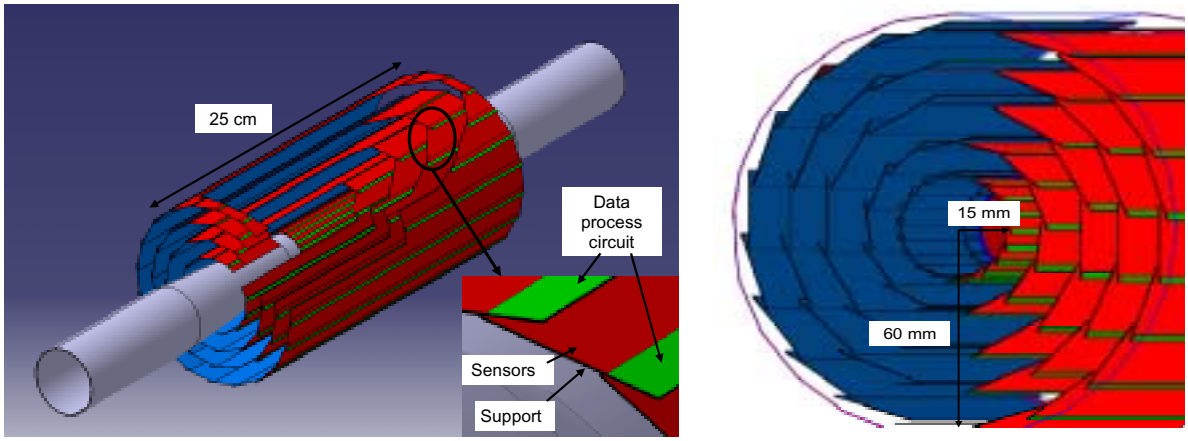
The electrons and positrons also generate radiation damage when traversing the pixels. The detection performance depends on the special care taken to make the pixels radiation tolerant.

B. Main requirements of the vertex detector for the future ILC

The main requirements of the Vertex detector for the future ILC are as following:

a) Multi-layer structure

In order to faithfully reconstruct the decay vertex from the particles, 5 layers will be needed for vertex detector (Fig. 1.4).



(a) (b)

Figure 1.4: The proposed architecture of the vertex detector for future ILC:
(a) Global view with and composition of each layer;
(b) Cross-section and dimensions for the first and last layers.

Layer	Radius (mm)	Pitch (µm)	$T_{r.o}$ (µs)	N_{Ladder}	N_{pix} (10^6)	P^{inst} (W)	P^{moy} (W)
L_0	15	20	25	20	25	< 100	< 5
L_1	25	25	50	26	65	< 130	< 7
L_2	37	30	< 200	24	75	< 100	< 5
L_3	48	35	< 200	32	70	< 110	< 6
L_4	60	40	< 200	40	70	< 125	< 6
Total				142	305	< 565	< 29

Table 1.2: Vertex detector sepcifications.

The length of the vertex detector is about 25 cm. Each layer is designed with ladders consisting of sensors, signal processing circuits and support. The first layer, being very close to the interaction point, should be at a radius of only 15 mm. The most outside layer (the 5th) will be at a radius around 60 mm. The detailed specifications (dimension, pitch of pixel, readout speed $T_{r.o}$, number of ladders, instant power dissipation and average power dissipation) of each layer are given in Tab. 1.2.

b) High Spatial resolution

A very high spatial resolution is mandatory to distinguish clearly the impact parameter of the collision (Fig. 1.5).

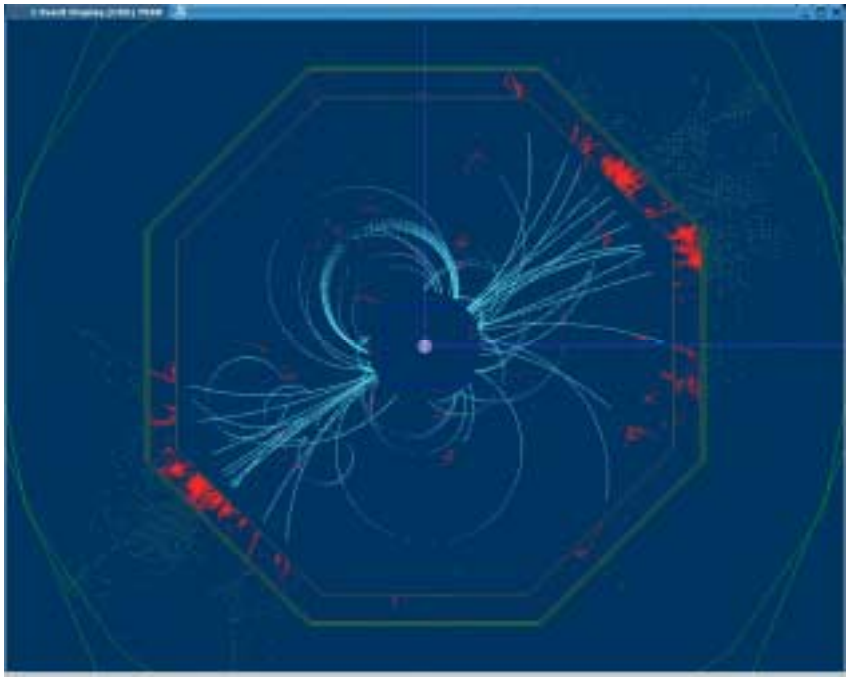


Figure 1.5: Simulation of a Higgs event in a future ILC detector (side view).

According to the physics simulation, the spatial resolution of the vertex detector for the future ILC should be less than 5 µm in order to identify almost every single track. As a result, the corresponding basic detective units have to be enough small.

c) High granularity

A high granularity of the whole vertex detector is needed because a very good spatial resolution is required. Thus, very compact device will be required because they have to be installed inside the very limited space of the vertex detector. The big amount of detection elements needs for on-detector signal processing with almost no external connections.

d) Fast readout

A very high readout speed will be also required in order to avoid the problem of pile up, appearing when two or more particles simultaneously touch the same area of the detector. Obviously, the solution of the problem is to increase the readout speed: to take away the information as fast as possible, so that the probability of pile up reduces. The studies show that a readout time of 25 µs per frame has to be reached in order to insure the separation of the different tracks of particles for the most inner layer. For the most outer layer, the needed time is less than 200 µs.

e) Material budget and power

In order to reduce multiple scattering, the material introduced in the detector must be minimized. As a result, the sensors used have to be thinned to about ~50 μm thickness and no extremely sophisticated cooling system is allowed, so that power consumption should be reduced to the minimum level.

f) Radiation tolerance

Besides all the above conditions, the resistance to radiation damage is also very important. Considering the high center-of-mass energy of the future ILC will be as high as 500 GeV, the radiation hardness for the ionising rays should be higher than 150 kRad/year and the requirement w.r.t the neutron gas should typically be about 10^{10} n_{eq}/cm^2/year. Moreover, it is worth noticing that beamstrahlung electrons may induce bulk damage equivalent to about 10^{11} n_{eq}/cm^2/year.

1.2.4 Different types of semiconductor sensitive detector

From the description of the thorough requirements of the vertex detector for the future ILC, it is a real challenge to design its basic detective units. The detector has to simultaneously achieve all the requirements: the required high granularity, low material budget, fast read-out, satisfactory radiation tolerance and low power consumption. However, designs based on different technologies have been successfully implemented on various applications during the development of particle physics, so that useful experience could be obtained.

A. Silicon Micro-strip:

In a made form of many parallel strips (metal electrode and diodes) placed on a common substrates, microstrip detectors offer a high resolution in one dimension. Since late'70s, successful vertex detectors were mainly based on this kind of sensors. The advantage of this detector is a quite high read-out speed and its good tolerance to radiation. However, the charges sensed by the micro-strip will continuously distribute along the geometrical length of micro-strip which means that ambiguity will be exist in one dimension. Although two planes of micro-strip can be used together to revise this ambiguity and the resolution is down to 10 μm recently, it is still too hard to reach the resolution requirement mentioned above in both directions.

B. Charge-Coupled-Device (CCD):

Successfully implemented on SLD (SLAC Large Detector) applications in the year'90s, the basic element of this detector is a pixel made up of MOS capacitors [2]. A detector is in fact a matrix which consists of thousands of pixels. The read-out mode is realized in a sequential

way. By electrical coupling between two adjacent capacitors, the charge sensed by one pixel passes all the pixels on the same column to the readout circuit at the end of matrix, one after the other and from top to bottom. By this way, all input signals will be collected by a single amplifier and then will be used to compose an image of the whole matrix on output. This image contains all the details of the input information. The resolution of CCD depends on the size of the pixels. Modern fabrication technology makes possible to reduce the pixel size to a rather small level (∝m), then a very high spatial resolution could be reached. Unfortunately, there are two drawbacks for this device which make it difficult to use for the ILC application. The first one is due to its functionality. Generally, the information of a CCD pixel array is successively readout, passing through one pixel to another so that the readout speed is very limited. The increase of the readout speed will make output noise level very important, largely degrading the sensor performance [3]. The second one is the radiation tolerance. An important parameter of the CCD is the charge transfer efficiency describing the ratio of the charge transferred between two adjacent capacitors. For a CCD sensor, this parameter is generally higher than 99% in order to achieve satisfied detecting performance. Unfortunately, the radiation greatly damages this parameter so that the charge lost happens during its transfer for readout. Lowering the working temperature could reduce the charge lost due to radiation, but this will largely increase the cost in turns of cryogenic cooling which is not expected. However, improvements have been made on this type of sensors. The RAL (Rutherford Appleton Laboratory, Chilton, UK) develops full new costum CCD sensor, using parallel readout mode to increase the readout speed and improving the tolerance to radiation of the structure [4].

C. Hybrid Pixel Sensors:

Well performed in LHC experiments (ATLAS and CMS) for vertex detectors, this kind of sensors is also made up of pixels but the read-out access to the pixel matrix is much easier than that of CCD pixel matrix. For hybrid pixel sensors, each pixel can be separately accessed. A very high readout speed can then be reached and sparse data scan, thereby reducing the occupancies due to the dense track environment, is possible. Besides, the radiation hardness of the hybrid pixel sensors is much better than the CCD sensors and can fulfil the radiation requirements needed for future ILC Vertex detector. While the above merits exiting, the need of integration due to the limit space inside the Vertex detector is hard to achieve by the hybrid pixel sensors. In fact, the pixel matrix and their lecture circuits (rows read-out and columns read-out) are realised on different PCB boards. The interconnections between these boards are bump bondings. Due to the technical limits and the mechanical requirements (robustness of connections), the size of pixel is hard to scale down. The minimum pixel size which is available in nowadays is about $50\,\mu m \times 400\,\mu m$ and this dimension does not allow to realize a high granularity pixel matrix. Another problem is that the hybrid pixels bounded on its electronics present a non negligible budget of matter, which makes it very difficult to be compatible with the requirements of the vertex detector for the future ILC. Moreover, the thinning requirement of the vertex detector results in a reduced signal charge compared to the standard silicon detector, so reducing the noise level is an issue.

However, progresses were obtained by a research group working on DEPleted Field Effect Transistor (DEPFET). A DEPFET structure combines detector and amplifier operations and is integrated onto a high ohmic fully depletable detector substrate [5]. Actually, with detectors of about 450 μm active layer, very high charge will be generated and collected by the device, making possible to achieve a very high signal-to-noise ratio. However, because the signal level depends on the thickness of the detector, the detection performance of a thinning detector (about 50 μm) is unknown. Prototypes with small pixel size (25 μm × 35 μm) offering low noise and low power dissipation were reported [6] [7] [8]. But, dedicated process, which is also very complicated, is needed to produce the DEPFET device. The cost of the production is also an issue.

D. Monolithic pixel sensors:

Thanks to the development in the electronic technology, more precisely the advancements in domain of microelectronics, a new type of device seems to be a good candidate: the Monolithic Active Pixel Sensors (MAPS). Two different technologies are available to realize this type of sensor: the standard Complementary Metal-Oxide-Semiconductor (CMOS) technology and the Silicon-On-Insulator (SOI) technology. An advantage of MAPS is that it integrates the signal processing circuits and the charge collecting elements on the same substrate, so that very compact sensor could be fabricated. Moreover, high speed and low-power device with good radiation tolerance could be also realized.

The cross section of standard CMOS and SOI device are shown in the figure below.

(a) CMOS device. (b) SOI device.

Figure 1.6: Cross section of a CMOS device and a SOI device.

As observed in Fig. 1.6, by replacing the deplete region by SiO_2, the SOI device can reduce effectively the capacitance at the source and drain junctions which results in a reduction in the RC delay due to parasitic capacitance, and hence a higher speed performance compared to the standard bulk CMOS device particularly at the downscale power supply voltage. Besides, the device structure also eliminates latch up in bulk CMOS, reduces the short channel effect and improves soft error immunity.

However, despite these advantages of the SOI technology, it faces some key challenges in process and manufacturing availability, devices and circuit design issues. Neither bonded nor separation by implanted oxygen (SIMOX) SOI are mature enough for mass production of

low-cost, low-defect-density substrates [9] and the floating body effect [10] in partially depleted devices poses major challenges for large-scale design.

1.2.5 Choice of the technology for the vertex detector of the future ILC

After the above brief survey of the several technologies already applied in different applications, ILC community has eventually chosen pixel technology as a candidate for the vertex detector. Composed of semi-conductor material, each tiny pixel will be individually read out in order to preserve the high-precision space point it provides, accurate to a few micrometers. The detector will need about 300 millions pixels in total. Located very close from the interaction point, slightly more than a centimeter, the pixels detector will allow reconstructing accurately secondary vertices.

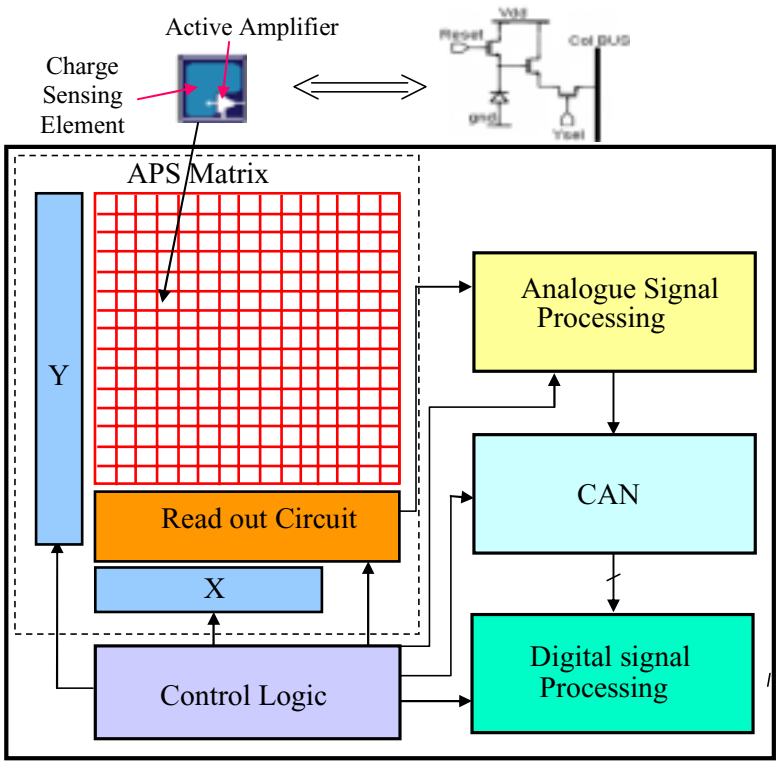


Figure 1.7: General architecture of a CMOS MAPS chip. The pixel matrix and the different processing blocs are integrated on the same chip.

The MAPS fabricated in the standard CMOS processes has been chosen as the research orientation of our group. Thanks to the developments of the standard CMOS processes in recent years, the charge sensing elements, active amplifiers and some basic signal processing circuits (analog memory and noise reduction circuits) can be integrated all together inside the pixel, which is only about 20×20 to $30 \times 30\ \mu m^2$. By this way, the detected signal is amplified very close to the charge sensing elements, avoiding the influence of other noise sources, so that a high Signal-to-Noise Ratio (SNR) is achieved. In addition, all kinds of

signal processing electronics, even sometimes very sophisticated ones (such as analogue-to-digital converter, logical control), can also be integrated on the same substrate as the pixel array (Fig. 1.7). Therefore, a very compact imaging system can be realized by this technique.

The MAPS, highly integrated in a very compact device, can fulfill the dimension limits of the vertex detector for the future ILC. Moreover, other advantages are also available for this technology. Firstly, thanks to the signal processing circuit integrated together with the pixel array, both fast readout and sparse data scan are possible. Secondly, the thickness of the CMOS sensors can be thinned down to about 50 μm, so that very thin sensors can be fabricated. Furthermore, the rapid scaling down of modern CMOS technology will largely reduce not only the dimension of the sensors but also its power consumption. The CMOS sensors require only one hundredth the power of CCDs. In addition, radiation tolerance of the MAPS is adequate for the ILC applications. At the end, the MAPS benefit the steady advancement in standard CMOS technology. As a result, the performance of sensors is expected to be improved while the cost is lowered.

1.3 Monolithic Active Pixel Sensors

1.3.1 General Description

MOS image sensors were firstly invented in the beginning of 1960's even before the invention of CCD [11]. While lots of attempts had been made, it was very difficult to integrate active transistors inside the pixels of the early MOS image sensors due to the size of the transistors. In the early CMOS processes, the pixels, whose size was usually about 20×20 to 30×30 μm^2, was too small to contain the transistors, which was generally larger than 10×10 μm^2. Therefore, the detection element inside the pixels was directly connected to the output bus, suffering from the large capacitive load of the bus. A large readout noise was then introduced and degraded the performances of the early CMOS imager sensors. Thus, the CMOS imager sensors could not offer the same performance compared to the CCD's invented in the 70's when other criteria such as quantum efficiency, fill factor (rate of area used for detection in pixel), charge detection efficiency, dark current, dynamic range, read-out time and noise was considered. However, the continuous progress of CMOS technology and of microelectronics led to resurgence in MOS sensors development. The size of the transistors was significantly scaled down and the market need for highly miniaturized, low-power and low cost imager sensors motivated the research on MOS imager sensors. Great advances had been done in the development of CMOS APS (Active Pixel Sensor) in the early '90s. For recent APS, with the improvements in the techniques of microelectronics, the detection element of pixels is separated from the common bus thanks to the in-pixel signal processing circuits realized by active transistors. Special care has been given to design these in-pixel signal processing circuits, so that the noise level is significantly reduced. As a result, satisfied performances, similar to CCD sensors, are achieved with respect to dynamic range

and readout noise but with much faster readout speed, numerous on-chip functionalities, lower system power consumption and the reduction of system cost [12].

A. General pixel concept

The pixel circuits can be divided into passive pixels and active pixels according to the inside structure. The active pixel contains at least an active element (an amplifier, a source follower etc.) inside it. There are mainly three approaches to realize pixel implementation in standard CMOS processes: 1) photodiode-type passive pixel; 2) photodiode-type active pixel and 3) photogate-type active pixel. The descriptions are as follow.

a) Photodiode-type Passive Pixel

The photodiode-type passive pixel was first suggested by Weckler in 1967 [13], [14]. It consists of a photodiode and a switch (one transistor). When the transistor is activated, the charge of the pixel will be switched to the column level charge integration amplifier (CIA) where the sensing charge will be converted to a voltage. The architecture of passive pixel is shown below (Fig. 1.8).

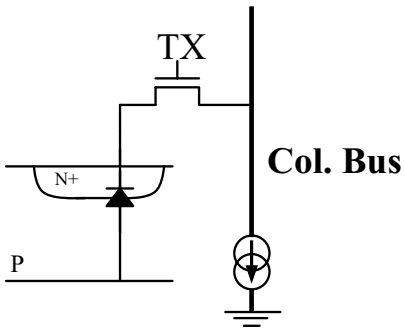


Figure 1.8: Architecture of passive pixel sensor.

Obviously, high fill factor and high quantum efficiency could be achieved by this type of the pixel because there is only one transistor in it. But since the large bus is directly connected to the pixels during the signal readout, the passive pixel has large capacitive load, largely reducing its readout time due to the high RC time constant of the common bus. In addition, the readout noise of passive pixels is typically high, on the order of 250 electrons [15]. Moreover the passive pixel suffers from Fixed Pattern Noise (FPN) which arises from the differences between the column level CIAs. The large FPN value make impossible to realize auto on-chip data processing which is required for particle physics experiments. All these reasons make passive pixel difficult to be used in the application for the charged particle detection [16].

b) Photodiode-type Active Pixel

In 1968, the photodiode-type APS was described by Noble [17]. A typical photodiode-type active pixel is shown in Fig.1.9.

Three transistors are made for the readout circuit within the pixel: a Reset transistor (RST), a Select Row transistor (SR) and a Source Follower transistor (SF). The charge detection diode is separated from the common bus by the SF, avoiding the influence of the common bus capacitance. The readout speed is then largely improved. Moreover, a reference level can be obtained during the reset stage. The reference level is then used to realize a special signal process, called Correlated Double Sampling (CDS), which largely reduces the readout noise of the pixel.

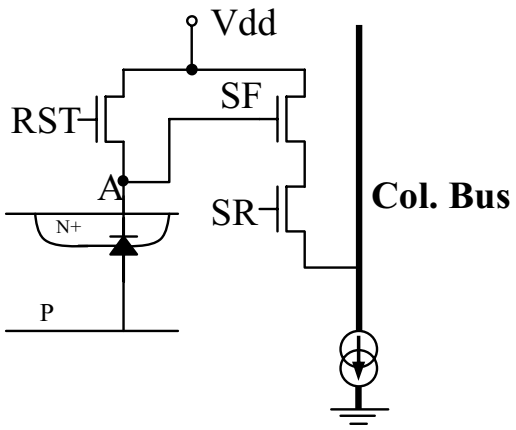


Figure 1.9: Architecture of photo-diode type active pixel sensor.

The information of all the cells of the pixel array is called a frame. A pixel is readout only once when a frame is readout. The readout operation of this classical 3-Transistor pixel can be divided into three steps, as described in Fig. 1.10.

The first step is the reset. The RST is activated during this period. The charge detection diode is initialized to a reference level. By activating the SR, this level is readout and memorized by column level signal treatment circuit.

The second step is the phototransduction. During this stage, the RST and SR are both turned off, leading the photodiode disconnected to its reference voltage. Thus, the charge detection photodiode is equivalent to a floating capacitor (seen into the point A), whose value is equal to the sum of the photodiode equivalent capacitance and the parasitic capacitance around it. Because of the electric field existing in the depletion zone of the P-N junction, the free electrons, generated by incident photons inside the volume of the silicon, are accumulated around the point A (the top plat of the equivalent capacitor). The charge of free electrons integrates on the equivalent capacitor, continuously lowering the potential of the point A (Fig. 1.10).

The third step is the signal readout. At the end of the integration period, the SR is activated, enabling the charge detection diode connects to the common bus through the SF. The voltage of the point A, slightly attenuated by the gain of the SF, is then sampled and stored in the column level signal treatment circuit.

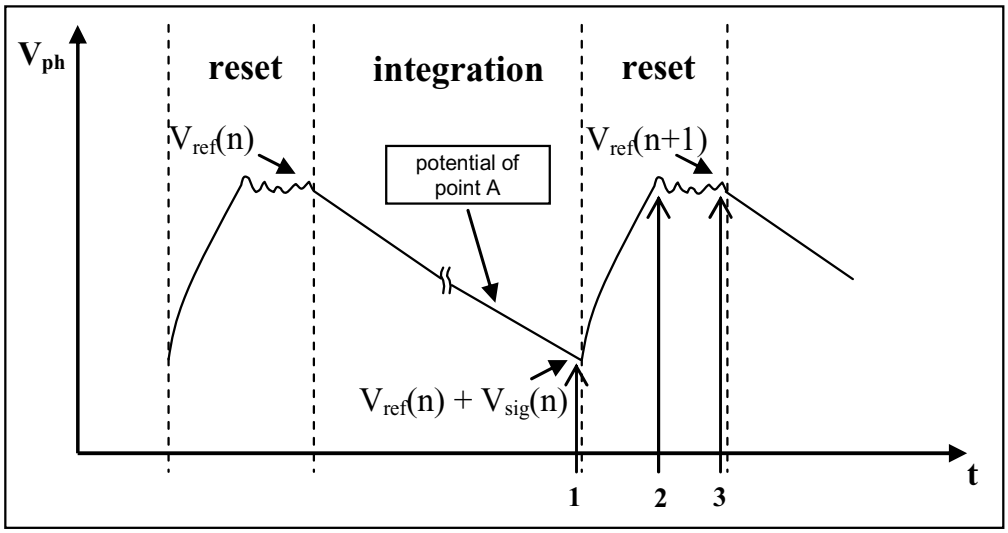


Figure 1.10: Operation of photo-diode type active pixel sensors.

Sequentially, the three steps are:

1) readout and memorize the voltage value (point A) of the frame n, the value is:

$$V_{sign}(n) + V_{ref}(n) + V_{noise}$$

2) reset the pixel;

3) readout and memorize the reference value (point A) of the frame n+1, the value is:

$$V_{ref}(n+1) + V_{noise}$$

In the above expressions, V_{noise} is the noise of the SF.

Obviously, the real signal value is the difference of the voltage value, sampled at step 1, and the referenced value, sampled at step 2. The column level performs the CDS operation, subtracting the two samples, and output the difference. The readout noise and the FPN are largely reduced thanks to this CDS operation.

From the above description, the reference level and the signal level are not taken from the same integration frame, so that these two samples are not truly correlated. As a result, the reset noise can not be removed. Although changing the sequence of the operation steps can fully removed this noise (to memorize the reference value and the signal potential value of the same integration frame), the readout speed is lowered. Fortunately, dedicated circuit

architectures allow to minimize this noise. More detailed discuss is given in the following chapters.

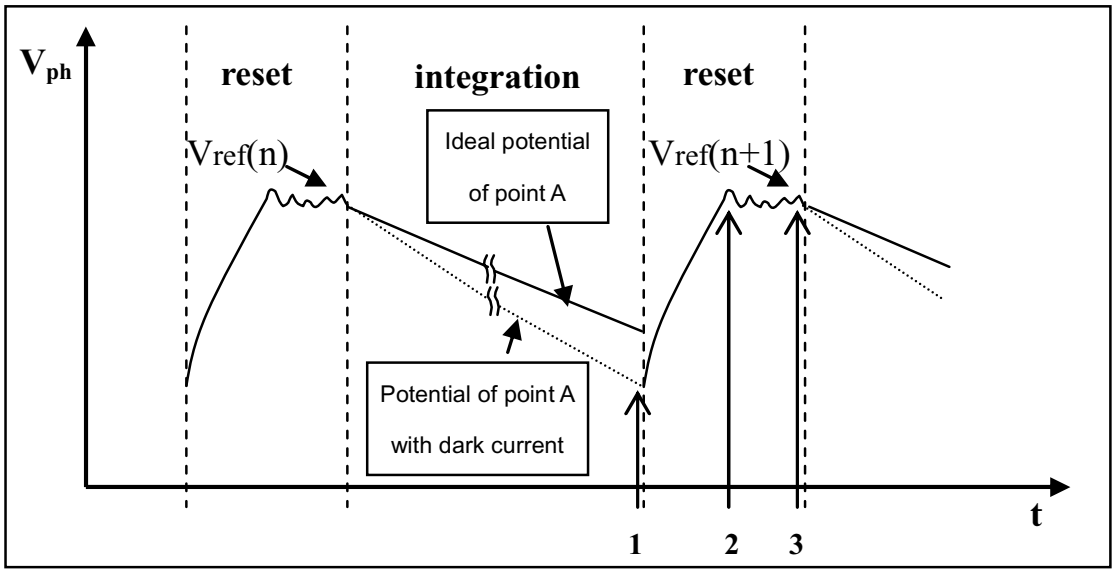


Figure 1.11: Influence of dark current.

Another issue is the dark current (leakage current), causing by electrons due to thermal generation-recombination inside the silicon. As shown in Fig. 1.11, the influence of dark current increases with the growth of the integration time. Although the influence of dark current varies from one pixel to another, for each pixel, this value is stable (changes only with the working temperature). It is therefore considered as a fixed offset. More detailed discussion is presented in the following chapter.

A very important character of active pixel sensors is the ability to convert accumulating charge into voltage. It can be described by Charge-to-Voltage conversion Factor (CVF) which is defined below:

$$Cvf = \frac{qA_1}{C_{fd}} \qquad (1.1)$$

with:

q : charge accumulated on the photodiode;

A_1 : equivalent voltage gain of the Source Follower (SF);

C_{fd}: equals to $C_d + C_p$, where C_d is the equivalent capacitance of the photodiode, being proportional to its size, and C_p represents the total parasitic capacitance around the photodiode.

A direct hint from equation 1.1 is that better conversion factor could be acquired by reducing the size of photodiode. Unfortunately, it is not completely true because there will be less

charge collected when the size of photodiode is reduced. Additionally, a large size of photon sensitive surface is always preferable for the visible light applications. That is why the conversion factor is generally about a few μV/e- for this type of active pixel sensors.

c) Photogate-type Active Pixel

In 1993, the photogate-type APS was introduced by Jet Propulsion Laboratory (JPL) [18] [19] [20] for high-performance scientific imaging and low-light applications. The architecture of photogate-type active pixel is as follow (Figure 1.12).

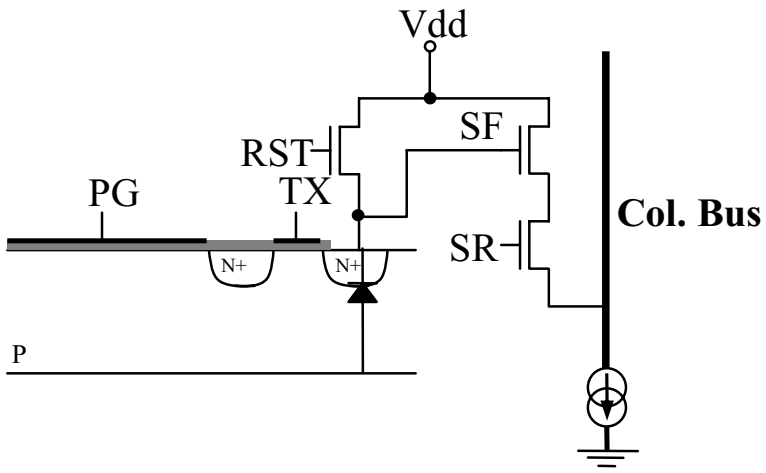


Figure 1.12: Architecture of photo-gate type active pixel sensors.

A positive tension will be applied on the gate PG to form a depletion zone under it. When the photons travel through the gate PG, electrons and holes will be created inside the depletion zone. Due to the existing electrical potential, the electrons are accumulating under the surface beneath PG while the holes push towards inside the substrate.

Greater the size of PG is, more charge will be collected. The continuous biasing transistor TX behaves as a charge transmitter and its source acts as a small reversed biased PN diode which converts the charge to a proportional voltage as output signal. There are three stages for signal readout:

1) Integration stage: positive voltage is applied on the gate PG to form a depletion region beneath it. Negative charge accumulates inside this region.

2) Reset stage: RST is activated and the reset value of the diode is readout and memorized by column level CDS circuit. The voltage applied on the gate PG is still activated during this stage.

3) Signal readout stage: RST is turned off and then the voltage applied on the gate PG is removed. Charge accumulated inside the depletion region transfer to the inversed biased PN

diode where the charge is transformed into voltage thanks to its equivalent capacitor. SR will be activated and the signal level is readout and memorized.

Finally the real output signal is the difference between these two readout values and the subtraction is automatically carried on by the CDS circuit. The signal value and the reset value are truly correlated.

A relative high conversion factor could be reached for this type of active pixel sensors because the charge are collected by the photo gate, whose surface is relative quite large, and then are transferred to the equivalent diode (the source of transistor TX), which could be made very small. The typical value is about 10 ~ 20 μV/e- [11].

B. General Chip concept and readout procedure

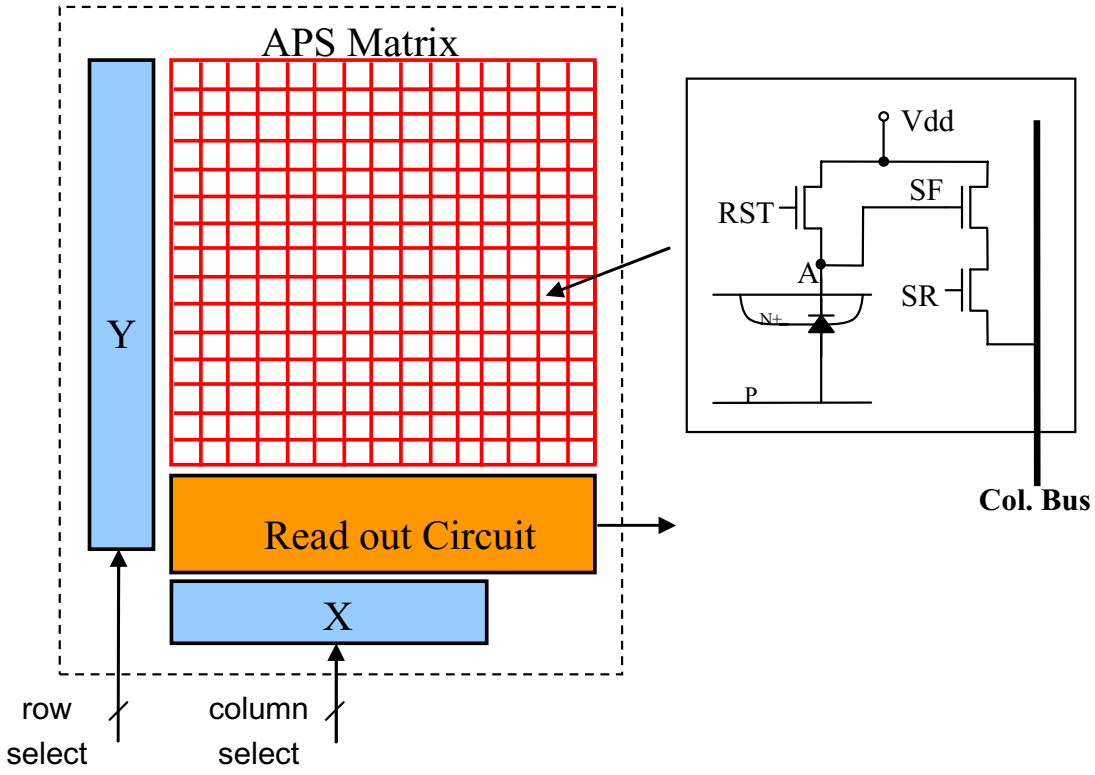
An example of the basic architecture of the CMOS imager sensors is shown in Fig. 1.13 (a), using the photodiode type pixels. Two address decoders are integrated together with the pixel array. Y is used for row selection and X is used for column selection. The pixel information is temporally memorized in the column level readout circuit and then is sequentially sent to the following signal processing circuits. The operation sequence is shown in Fig. 1.13 (b) and is described in detail as follows.

Firstly, the address decoder Y selects the n^{th} row to read the information of the pixels. The Select Row transistors (SR) of the pixels in this row are activated and pixels information is available on the common bus of each column. For the other rows not selected, their SRs are in the state “off”, so that their pixels are in the integration mode.

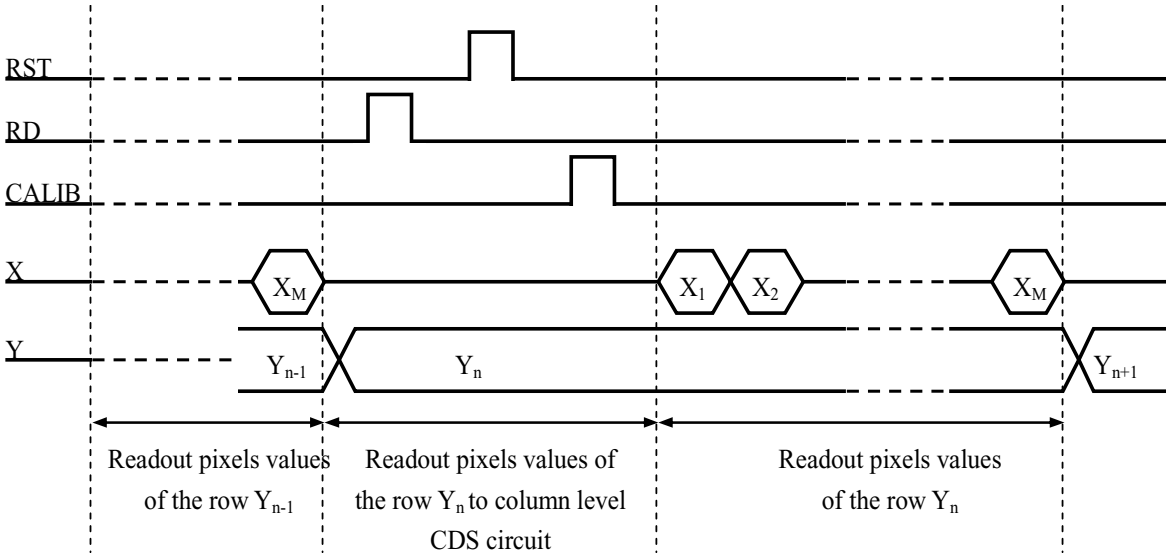
Secondly, the information of pixels is sampled and memorized by the column level signal processing circuit. To do so, two additional column level sample commands, RD and CALIB, are used. RD is firstly activated to sampled and memorized pixel output value ($V_{sign}(n)$ + $V_{ref}(n)$ + Noise). Then, a reset command (RST) initializes the charge detection photodiode and CALIB is used to sample and memorize the reference level after reset.

Thirdly, when the sampling of pixel information has terminated, the address decoder X begins to work. It successively selects the information stored in column level readout circuit and sends them to the following signal processing circuit. The useful signal level is the difference between the pixel output level sampled during RD and the reference level sampled during CALIB. It is important to note that these two values are not truly correlated.

Then, when the information of last pixel has been sent to the following signal processing circuit, the readout process of n^{th} row terminates and the address decoder Y selects the next row to perform the same operation.



(a) Basic architecture of a photodiode type active pixel sensor, the dimension of the pixel array is (row x column): N x M



(b) readout timing diagram of the n^{th} row.

Figure 1.13: Principle of active pixel sensor: (a) basic architecture of photodiode type active pixel sensors and (b) its readout timing diagram.

1.3.2 From Visible Light to Particles Detection

Since 90's, the MAPS earned more and more customer interest thanks to its characteristics like miniaturisation, low-power and low-cost. The growing demand for CMOS image sensors demonstrates its successful applications in the domain of photon detection. Meanwhile, in the domain of particle physics, some future applications and their experimental conditions require to integrate the detecting elements with the front-end electronics on the same silicon substrate. The recent advancements of CMOS image sensors make it a promising candidate.

In the application of visible light domain, the photons will only reach a depth of tens of micrometers inside the MAPS because of their low energy [21]. All the energy of the incident photons is used to create electron-hole pairs. For a classical 3-Transistor photodiode pixel, the charge collection procedure has been already described in section 1.3.1. For the commercial visible light applications, a low CVF does not much influence the performance of MAPS because of the large quantity of photons received. Moreover, in this case, where on-chip sparsification is not required, the FPN can be corrected by subtraction to the previewed FPN value, which is obtain from the mean value of the image of the pixel array taken in darkness.

The charge collection principle of high energy particles is different from the visible photons. The particles at the Ionizing Minimum penetrate through MAPS and therefore electron-hole pairs will be created along particle's track (Fig. 1.14). The free charge carriers thermally diffuse inside the volume of the sensor underneath the readout electronics. A 100% fill factor, as required in physical tracking applications, is then obtained. Because of the electrical potential existing inside the volume of the pixel between the P-N junction and the P-epitaxial layer, the freely diffused electrons can only be collected by the N-Well (Fig. 1.14). For this reason, only NMOS can be used in the pixel design aiming at charged particle detection.

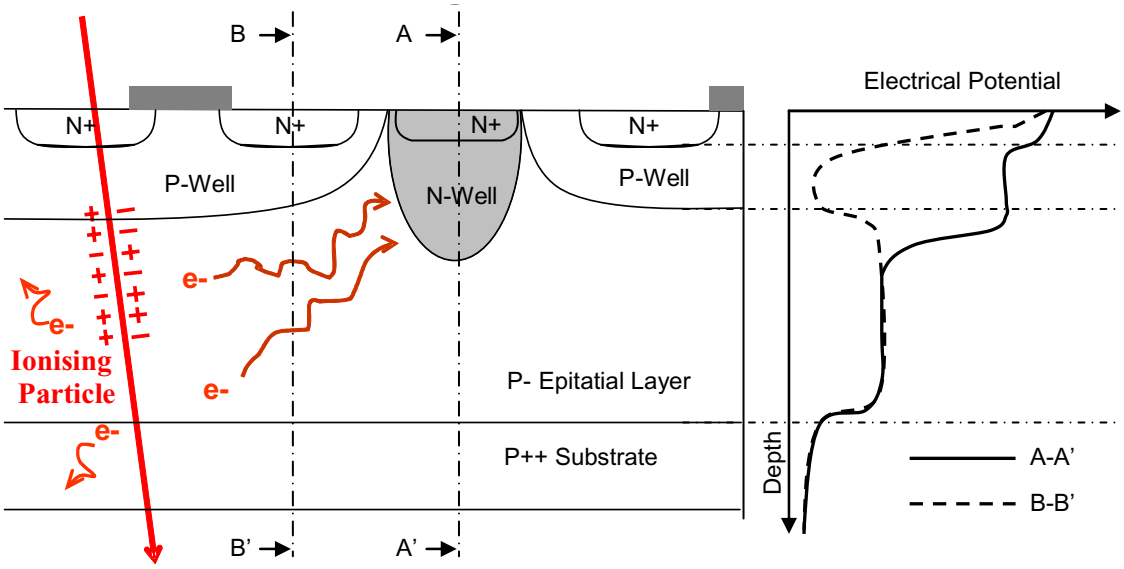


Figure 1.14 Principle of charge collecting in photo-diode type CMOS sensors.

The photogate-type active pixel sensors fabricated in some modern CMOS processes are not appropriate to MIPs detection. The reason is that a P-Well is formed by default for the place where no N-Well is required. As a result, an electrical potential exists, preventing the photo gate to collect efficiently the charge of the free carriers (Fig. 1.15). Therefore, photodiode type pixel is usually used for the application of the charged particle detection.

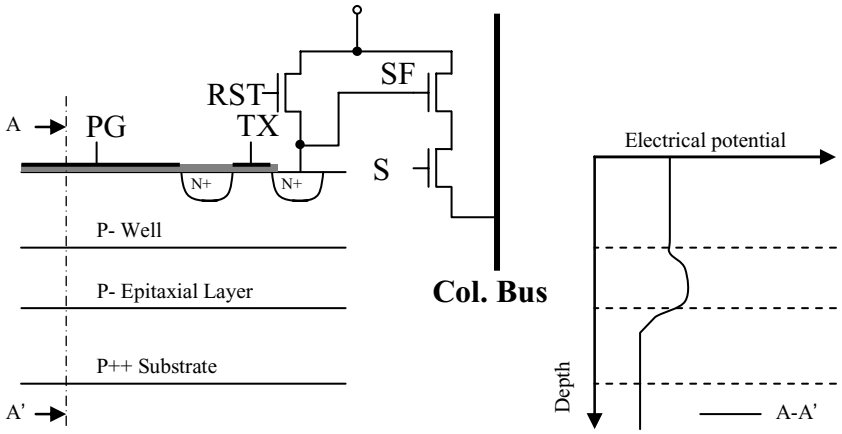


Figure 1.15 Electrical potential inside the photogate type pixel.

The charge collection procedure of the classical 3-Transistor photodiode pixel, different from the visible light application, is described below.

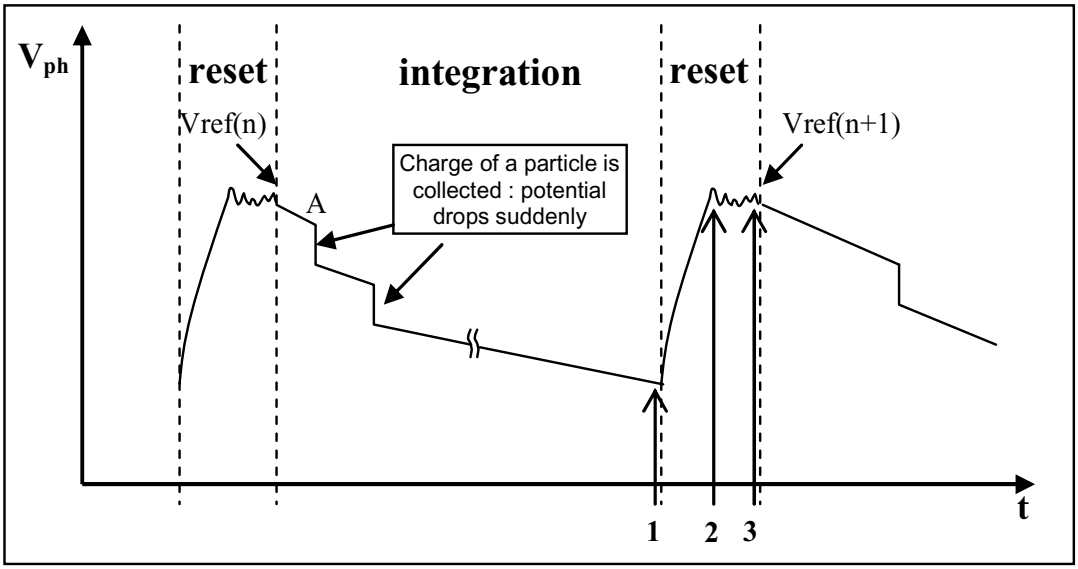


Figure 1.16: Particle charge detection using photo-diode type active pixel sensors.

Similar to the operation in visible light application, three stages are needed to get the real signal level. The difference is in the second stage. The potential of point A drops continuously in the case of intense visible light application, but it drops suddenly with a step in the application of the MIPs detection. Every time the charge of an incident particle is

collected by the photodiode, a quick drop (about 100ns [22]) of the potential on the point A occurs (Fig. 1.16). If the potential valued is not sampled quickly (a long integration time), its value, memorized at the end of the integration period, is the sum of the charge of several particles and the reference level. As the experiments of particle physics need to identify different particles, a very high readout speed is then needed.

In order to avoid that the generated charge carriers diffuse too widely and keep good spatial resolution, an epitaxial layer, grown on a highly p^{++} type doped substrate, is required [23]. For the MIPs, the charge generated inside the epitaxial layer is about 80 e-/µm. Considering the typical thickness of the epitaxial layer (tens of µm), the signal associated to the electrons produced is really limited. Therefore, a very high CVF is strongly required for the MIPs detection.

Although the FPN can be removed by offline signal processing in visible light application, it is another issue for the CMOS sensors designed for MIPs detection where on-chip sparsification is required for most High Energy Particles (HEP) experiments in order to avoid useless data.

In addition, several other differences exist between the application in charged particle detection and in visible light.

- The first one comes from the optical fill factor. For the visible light applications, the number of transistors inside a pixel should be reduced at a minimum level in order to get a large optical fill factor because that the photons of visible light are not able to pass through the component of the circuit covering over the silicon layer. However, the case is different for MIP detection, where the energy is high enough to pass through. Then, a 100% fill factor is reachable.

- The second is the requirement of high spatial resolution for the MIP detection. In visible light applications, the MAPS collect incident photons to form an image without the need to separate the position of each photon. On the contrary, a MIPs detector is asked to accurately identify all the particle tracks emitted during a collision so that a high spatial resolution is needed and extra design consideration should be paid.

- At last, the radiation hardness is another issue that should be taken into account for the applications of MIPs detection. The architectures achieving high radiation tolerance have to be found.

Considering all the above discussions, the CMOS sensors (MAPS type) could be a good potential candidate for the vertex detector of the future ILC. But their characteristics need to be carefully studied and necessary full R&D campaigns have to be carried on.

References

[1] Griffiths, David J. (1987). *Introduction to Elementary Particles*. Wiley, John & Sons, Inc. ISBN 0-471-60386-4

[2] J.E. Brau, "VXD3: the SLD Vertex Detector Upgrade Based an a 307 MPixel CCD System," *IEEE Trans. Electron Devices*, vol. 43, NO. 3, June 1996, pp.1107-1112.

[3] K.Abe et al., "Design and performance of the SLD vertex detector: a 307 Mpixel tracking system," *Nucl. Instrum. Meth. A 400*, 1997, pp. 287-343.

[4] K. Stefanov, "LCFI Status Report: Vertex Detector R&D," *ILC-ECFA Workshop* 2006, Valencia

[5] "A DEPFET based pixel vertex detector for the detector at TESLA," *LC-DET-2002-004*, 5, April, 2002.

[6] R.H. Richter et al., "Design and technology of DEPFET pixel sensors for linearcollider applications,"*Nucl. Instr. and Meth. A* **5**11, 21, Sept. 2003, pp. 250 - 256

[7] M. Trimpl et al., "A DEPFET pixel matrix system for the ILC vertex detector," *Nucl. Instr. and Meth. A* 560, 1st May 2006, pp. 21-25.

[8] C. Sandow, et al., "Clear-performance of linear DEPFET devices," *Nucl. Instr. and Meth. A* 568, 11, Nov. 2006, pp. 176-180.

[9] Z, J. Lemnios, "Manufacturing technology challenges for low power electronics," in Dig. Tech. Papers, Symp. VLSI Technology, 1995, pp. 5-8

[10] R. Howes, W.Redman-White, "A Small-Signal Model for the Frequency- dependent Drain Admittance in Floating Substrate MOSFET's," *IEEE J. Solid-State Circuits*, vol. 27, NO. 8, Aug. 1992, pp. 1186-1193.

[11] E. R. Fossum, "CMOS Image Sensors: Electronic Camera-On-A-Chip," *IEEE Trans. Electron Devices*, vol. 44, NO. 10, Oct. 1997, pp. 1689-1698.

[12] M. Bigas, E. Cabruja, J. Forest and J. Salvi, "Review of CMOS image sensors," *Microelectronics Journal,* vol. 37, No. 5, pp. 433-451, May 2006.

[13] G. P. Weckler, "Operation of p-n junction photodetectors in a photon flux integration mode," *IEEE J. Solid-State Circuits*, vol. SC-2, 1967, pp. 65-73.

[14] R. Dyck and G. Weckler, "Intergrated arrays of silicon photodetectors for image sensing," *IEEE Trans. Electron Devices,*vol. ED-15, Apr. 1967, pp. 196-201.

[15] O. Yadid-Pecht and R. Etienne-Cummings, *CMOS Imagers: From Phototransduction to Image Processing*, Kluwer Academic Publishers, Boston, 2004.

[16] TESLA, Tech. Design Rep. DESY, Hamburg, 2001.

[17] P. Noble, "Self-scanned silicon image detector arrays," *IEEE Trans. Electron Devices,*vol. ED-15, Apr. 1968, pp. 202-209.

[18] S. Mendis, S. Kemeny and E. R. Fossum, "A 128 x 128 CMOS active pixel image sensor for highly integrated imaging systems," in *IEEE IEDM Tech. Dig.*, 1993, pp. 583-586.

[19] S. Mendis et al., "CMOS active pixel image sensor," *IEEE Trans. Electron Devices,* vol. 41, Mar. 1994, pp. 452-453.

[20] S. K. Mendis et al., "CMOS active pixel image sensors for highly integrated imaging systems," *IEEE J. Solid-State Circuits*, vol. 32, Feb, 1997, pp. 187-197.

[21] Y. Degerli, "Etude, modélisation des bruits et conception des circuits de lecture dans les capteurs d'images à pixels actifs CMOS", Ph.D. Thesis, Toulouse, 2000.

[22] G. Deptuch, "New generation of monolithic active pixel sensors for charged particle detection", Ph.D. Thesis, Strasbourg, 2001.

[23] W. Dulinski et al., "CMOS Monolithic Active Pixel Sensors for minimum ionizing particle tracking using non-epitaxial silicon substrate," *IEEE Trans. Nucl. Sci.*, vol. 51, Aug. 2004, pp 1305-1310

Chapter II

The Development of Fast MAPS for Particle Detection

2.1 Introduction

The vertex detector for the future ILC simultaneously requires a high readout speed, high spatial resolution, radiation tolerance and low-power dissipation. Moreover, in order to store only useful information, on-chip auto sparsification is also required. In practice, these parameters trade with each other, making the design work a multi-parameter optimization. These requirements and the trade-offs between them present a big challenge in the design of high performance sensors. Specially, the signal level of MIP (Minimum Ionizing Particles) is very low (only a few mV), which is of the same order of magnitude of typical noise values and mismatches for standard CMOS process. This small signal value makes on-chip auto sparsification a very difficult task. Therefore, for the CMOS sensors aiming at detecting MIPs, an excellent noise performance is strongly required. Dedicated circuit architectures used for reducing sensor noise have to be implemented.

2.2 Sources of noise in MAPS

One of the most important parameters to be considered in MAPS design is its noise level. The power dissipation, readout speed and linearity are all influenced by the system noise level. The minimum signal level that a CMOS sensor can correctly process is determined by its noise level. The small signals, which are in the same range as noise, can not be identified by sensor. Typically, two methods can be used to increase the dynamic range in the entrance of a CMOS sensor. The first consists in decreasing the noise level so that signals with lower amplitude can be distinguished. The second consists in increasing the saturation level of entrance, without increasing the noise level, so that higher signals can be accepted by sensor. As mentioned above, the signal level is very limited in the application of the MIPs detection due to thin epitaxial layer. Then, the first method has therefore been chosen in our design.

2.2.1 General definition

Noise is a random process and has random value. Generally, when the noise of a system is estimated and measured, its rms (root mean square) value is used. The definition of rms value for a signal y(t) is:

$$\sigma_y = \sqrt{\overline{y^2}} = \left[\frac{1}{T} \int_0^T y^2(t) dt \right]^{1/2} \qquad (2.1)$$

where T is the observation time.

In fact, the rms value indicates the normalised noise power of a signal, which is defined as the power dissipated over a $1\,\Omega$ resistor. For example, when a random voltage source $v_n(t)$ is applied, the average power dissipated $P_{v,dis}$ equals to:

$$P_{v,dis} = \overline{v_n^2} \big/ 1\Omega = \sigma_{v_n}^2 \qquad (2.2)$$

Similar result is obtained when a random current source $i_n(t)$ is used:

$$P_{i,dis} = \overline{i_n^2} \cdot 1\Omega = \sigma_{i_n}^2 \qquad (2.3)$$

In frequency domain, the spectral density is used and is defined as the normalised noise power over a 1-Hz bandwidth. The total average noise power is then equal to:

$$\sigma_Y^2 = \int_0^{+\infty} Y(f)df \quad [\text{A}^2/\text{Hz or V}^2/\text{Hz}] \qquad (2.4)$$

where $Y(f)$ is the power spectral density.
Another important notion is the Signal-to-Noise Ratio (SNR), which is defined as:

$$SNR = 10\log\left[\frac{P_{signal}}{P_{noise}}\right] = 20\log\left[\frac{\sigma_{signal}}{\sigma_{noise}}\right] \quad [\text{dB}] \qquad (2.5)$$

2.2.2 Types of Noise

There are two types of noise in integrated circuits: the internal noise and the external noise.

A. Internal noise

The internal noise refers to the electronic noise of the device itself which is due to physical phenomena. For example, the resistance thermal noise is a kind of typical device electronic noise. It is due to the physical phenomena in circuit components under actual biasing conditions. The internal noise determines the lowest level of input signal of a circuit.

a) Thermal noise

The thermal noise is introduced by the random thermal movement of electrons in a conductor. It is associated to the resistance or the resistive part of a component and it is proportional to the absolute temperature. The value of current or tension applied on component has no influence on it. Considering a resistor R, the relationship between the spectral density of its thermal noise and it is given by:

$$V_{th,n}(f) = 4kTR \text{ with } f \geq 0 \ [V^2/Hz] \qquad (2.6)$$

where $k = 1.38\times10^{-23}\,J/K$ is the Boltzmann constant and T is the absolute temperature (in Kelvin).

The formula (2.6) indicates that the thermal noise is a constant when there is no fluctuation of temperature. In fact, up to 100 THz [1], it is quite accurate to consider that the spectrum is flat. That is why it is sometimes called white noise.

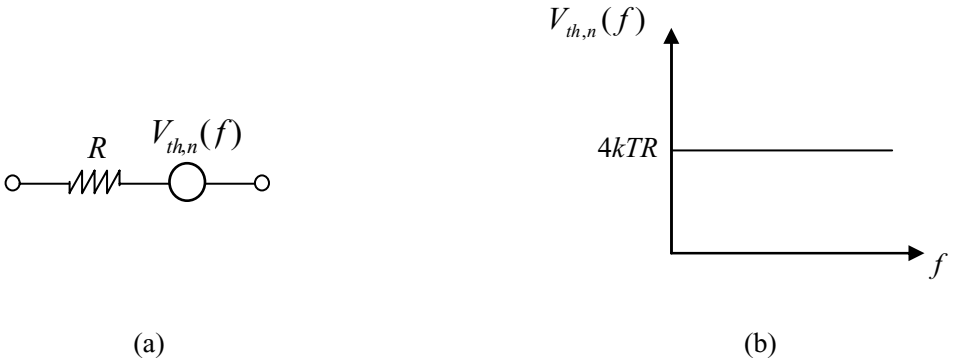


Figure 2.1: Thermal noise of a resistor: (a) equivalent circuit model for noise calculation and (b) its spectrum.

b) Flicker noise (*1/f* noise)

Many lattice defects may exist, they induce extra carrier traps in the interface between the gate oxide and the silicon substrate in a MOSFET, since the silicon crystal reaches its border, many extra carrier traps exist. In this interface, the carriers are captured and released in a random way with different time constants, introducing a noise source associated to the frequency. The spectral density of the flicker noise modelled as a voltage source in series with the gate is given by:

$$V_{1/f,n}(f) = \frac{K}{C_{ox}WL} \cdot \frac{1}{f^{EF}} \qquad (2.7)$$

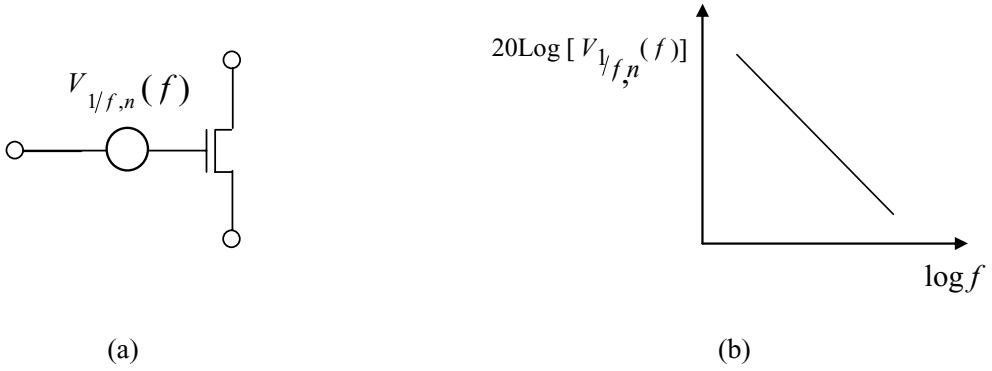
where:
K is a process-dependent constant on the order of $10^{-25}V^2F$,
C_{ox} is the equivalent capacitor of gate oxide,
W is the width of transistor,
L is the channel-length of the transistor,
EF is a process dependant parameter which is close to 1.

From the expression above, when f is a quite low frequency it seems that the flicker noise value is infinite. But it is not really the case. There are two reasons:

- Firstly, the infinite noise simply means that the observation time is infinite which is not true in practice. For this reason, the flicker noise has always a finite value.

- Secondly, there is no interest for the system to function at a significant low-frequency. Actually, the system input signals are much faster and this low frequency is limited by system pass band.

Moreover, because the flicker noise is caused by crystal border in the interface between oxide gate and silicon substrate, it depends on the "cleanliness" of this interface. For this reason, the flicker noise randomly varies from one technology to another one.



(a) (b)

Figure 2.2: Flicker noise: (a) equivalent circuit model and (b) its spectrum.

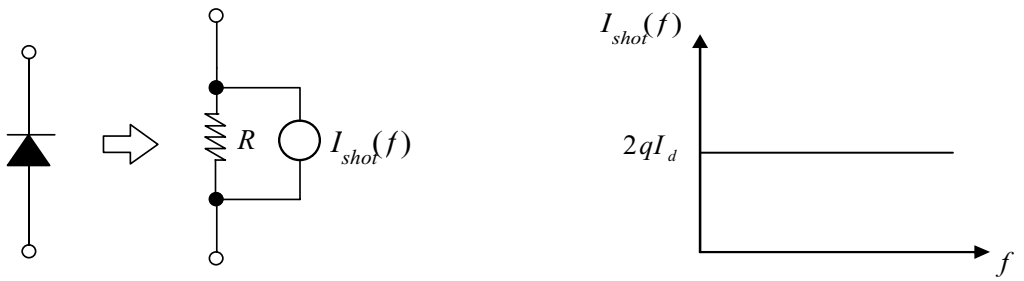
The equation (2.7) is only an approximation. The flicker noise equation in reality is more complicated as mentioned in reference [2].

c) Shot noise

First studied by W. Schottky using vacuum-tube diodes, the shot noise also occurs in PN junctions. The main cause of shot noise is the fluctuation of DC bias current flowing through the diodes. The background current results from ambient background light unrelated to the signal. The dark current generated by thermal movement of carriers inside the depletion region and the variation of photocurrent both contribute to the source of shot noise. The shot noise can be modelled as a current source in parallel with a small-signal resistance of diode and its spectral density given by:

$$I_{shot}(f) = 2qI_d \quad [A^2/Hz] \qquad (2.8)$$

where q is constant and its value is the charge of one electron (1.6 x 10^{-19} C), I_d is the diode DC bias current.



(a) (b)

Figure 2.3: Shot noise spectrum: (a) equivalent circuit model and (b) its spectrum.

B. External noise

The external noise refers to "environment" disturbances. A circuit experiences not only the random disturbances through its power supply or ground lines but also the perturbations due to the couplings of electromagnetic and electrostatic interfaces.

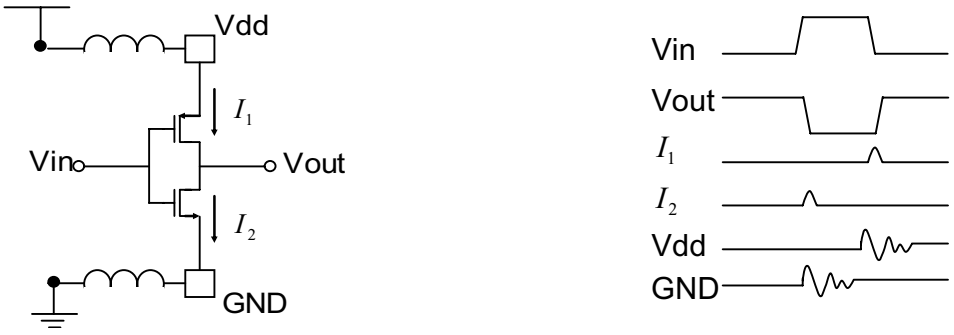


Figure 2.4: Glitches produced by self-inductance.

a) Self-inductance

Self-inductance is introduced by each bond wire and its corresponding package (Fig. 2.4). Nowadays, in mixed-signal circuits, self-inductances of power supply and ground lines impacts the performance. When a charge of state occurs in a digital gate, the glitches (an electronic pulse of short duration) can be produced on the actual power supply and ground lines and the analog signal will be degraded.

b) Parasitic capacitance

Parasitic capacitances exist between the interconnection lines and different layers. When a long interconnection is needed, the transfer speed of signal may be degraded by the parallel-plate (parasitic capacitance due to the surface of different interconnection lines) and fringe capacitances (parasitic capacitance due to the edge of different interconnection lines) of wires. The different kinds of parasitic capacitance are shown in the following figure.

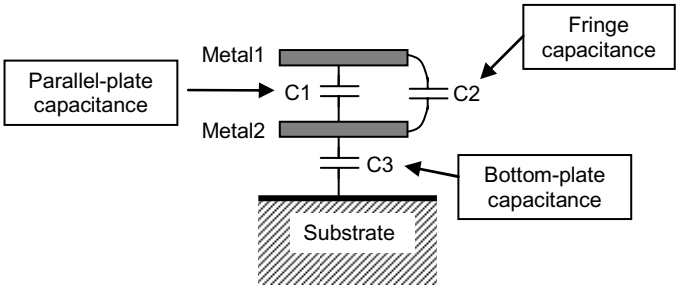


Figure 2.5: Different types of parasitic capacitances:
Parallel-plate capacitance (C_1), fringe capacitance (C_2) and bottom-plate capacitance(C_3).

More importantly, the coupling capacitances between adjacent lines can introduce cross talk of signals. Even if the value of coupling capacitance is very small, the signal may be corrupted.

c) Coupling and mismatch

Another kind of disturbance is the substrate coupling. In order to minimize the phenomenon of latchup, most modern CMOS technologies use a heavy-doped p+ substrate which is low resistive (about $0.1\,\Omega \cdot cm$). But undesired coupling paths between different parts of a circuit are created inside the substrate by this low resistivity. As a result, some sensitive signals may be corrupted.

Moreover, the mismatch due to the fabrication procedure introduces unwanted errors in transistor dimensions and perturbs the performance of the circuit.

However, even though all these external noises affect the circuit performance and dominate the total system noise in certain cases, they can be considerably reduced by applying some conception methods such as [3] – [5]:

- Use of differential pairs in circuit design,
- Realize the layout of circuit with symmetric structures,
- Separate the power supply and ground lines of digital part and analog part,
- Use of the protection structure for some sensitive analog parts,
- Use of decoupling capacitances where it is necessary.

2.2.3 Noise in MAPS for MIPs detection

After reviewing different types of noise, it is necessary to understand the noise sources in MAPS used for MIPs detection. Potential noise sources are present from the pixel through the readout circuit. In fact, the output analog signal of a MAPS circuit mainly contains the following sources of noise:

- Shot noise due to dark current in the pixels,
- Pixel reset noise,
- Column readout noise,
- Fixed Pattern Noise (FPN).

A. Dark current and the shot noise resulting from dark current

The dark current is caused by the thermal generation of electron-hole pairs inside the sensitive region of the pixel and is strongly dependent on exposure time. In addition, the generation rate varies from one pixel to another and this variation results in non-uniformity of dark current between pixels. The dark current closely depends on temperature and its value doubles every 5°C to 8°C [6]. The relationship between the dark current and the temperature is given by [7]:

$$J = AT^{3/2} e^{\frac{Vg \cdot q}{2kT}} \qquad (2.9)$$

where A is a constant and Vg is the width of the band gap. For example, the value of dark current at room ambient temperature is from 200 to 1500 pA/cm^2 in silicon.

According to equation (2.8), the shot noise caused by average dark current can be given by:

$$i_{dark}(noise) = \sqrt{qI_{dark} \cdot \tau} \qquad (2.10)$$

which is equivalent to:

$$v_{dark}(noise) = \frac{q}{C_{pixel}} \sqrt{\frac{I_{dark} \cdot \tau}{q}} \qquad (2.11)$$

where C_{pixel} is the total pixel capacitance and τ is the exposure time.

B. Pixel reset noise

The signal level and the reset level are sampled during a readout period. The difference between these two samples gives the useful signal value. However, the reset level is disturbed by thermal noise and exhibits a random fluctuation. This fluctuation is called reset or kT/C noise [8], where k is the Boltzmann's constant, T is the absolute temperature and C is the equivalent pixel capacitance.

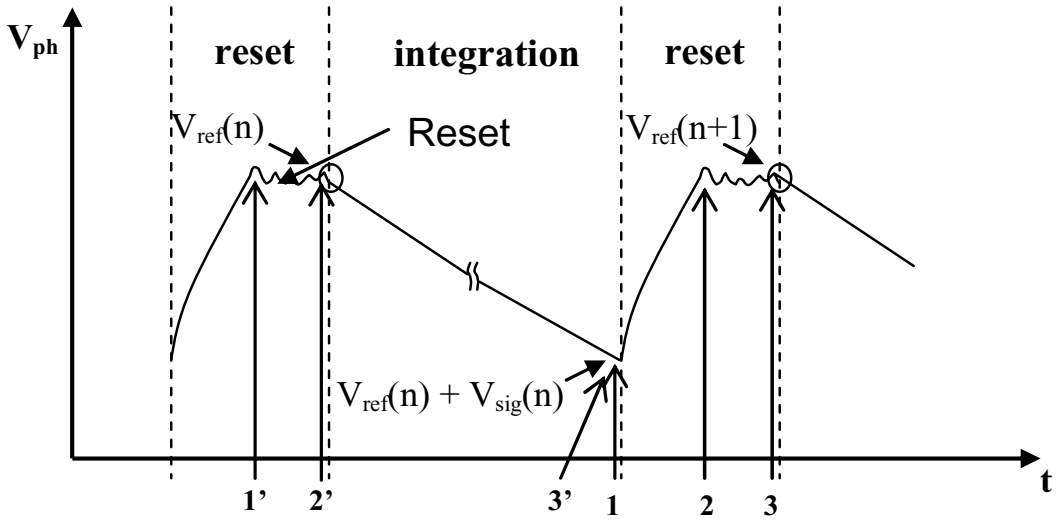


Figure 2.6: A readout cycle for photo-diode type active pixel sensors. The reset level in step n, influenced by thermal noise, is not the same as in step n+1.

The readout step is indicated as 1, 2 and 3 in Fig. 2.6. Step 1 and 3 are used to sample the pixel values. During step 2 the pixel is reset. The useful output signal is the difference between the two level sampled in step 1 and step 3 respectively, which is given by:

$$V_{ref}(n) + V_{sig}(n) - V_{ref}(n+1) \quad (2.12)$$

The reset noise is the difference between $V_{ref}(n)$ and $V_{ref}(n+1)$.

Changing the readout step as 1', 2' and 3' can completely removed the readout noise. But the readout period is too long in this case (equals to the integration time) to use in our application.

C. Noise of readout

The column level readout circuit samples and buffers both the pixel reset and signal levels. The major noise sources are the kT/C noise resulting from the sampling process, the thermal noise and the *1/f* noise associated with the column amplifier MOS devices. Especially, the temporal noise of readout dominates the total system noise level for the applications of low illumination and short exposure time, which is a condition of fast CMOS sensors used for particle detection [9].

The pixel reset noise and the column level readout noise do not change with the integration time and are both dependant to ambient temperature. Thus, we call the sum of these two sources of noise Temporal Noise in our measurements. This type of noise is obtained by statistical method which is presented in next chapter.

D. Fixed Pattern Noise (FPN)

Unlike the Temporal Noise presented above, the FPN is a kind of spatial noise which depends on the different pixels (or different physical channels). It is due to the variations of transistor MOS parameters used in the circuits:

- the dimension mismatch caused by fabrication process;
- the variation of the threshold voltages V_{th} from transistor to transistor;
- the variation of the transconductances g_m from transistor to transistor.

Generally, the difference between parameters of two transistors increases with the distance between them. The variation about tenths of millivolts could be observed in a single chip [10] [11]. The following example gives an image showing the FPN of an APS with 256×256 pixels [9]. The final image is obtained by summing 50 images taken in darkness. The pixel level CDS operation is sufficient to reduce significantly the FPN.

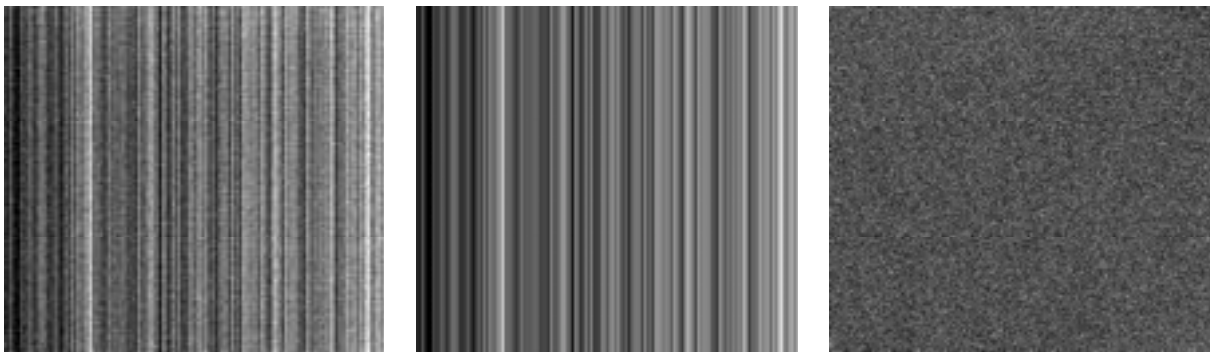


Figure 2.7: FPN level of an APS with 256×256 pixels. The images are obtained by summing 50 times the values taken in darkness [9].

The FPN of one column can be easily eliminated by removing a reference value. But there are usually hundreds of columns in a CMOS sensor matrix and the value of FPN varies from one column to another. Then, it is important to have a small standard variance of FPN for the

overall pixel array, so that only one reference is sufficient to maintain the cleanliness of output signal.

The FPN of a pixel array can be obtained in measurement by calculating the standard variance of the mean values of each pixel, taken without the input signal (in darkness for the visible applications, for example). Detailed analysis is presented in next chapter.

2.3 First MAPS design for particle detection

Although the MIPs detection can be compared to image sensor exposed to low illumination, it is still necessary to test the feasibility of MAPS and its performance to MIPs detection under different strict requirements. For this reason, special care is made in the circuit architecture.

2.3.1 The first MIPs detection MIMOSA series

Since 1999, the series of MIMOSA (Minimum Ionizing particle MOS Active pixel sensor) prototypes have been developed by IPHC (Institut Pluridisciplinaire Hubert Curien) at Strasbourg. The purpose of these prototypes was to test the usefulness of CMOS sensors for application to high energy physics, to verify its ability and efficiency for charged particle tracking.

A. Chip design aspects

The n-well/p-epi diode was chosen as the charge collecting element since nearly 100% fill factor for the visible light detection is allowed with this architecture [12]. Following this idea, four prototype chips (the MIMOSA 1, 2, 3 and 4) were designed and fabricated.

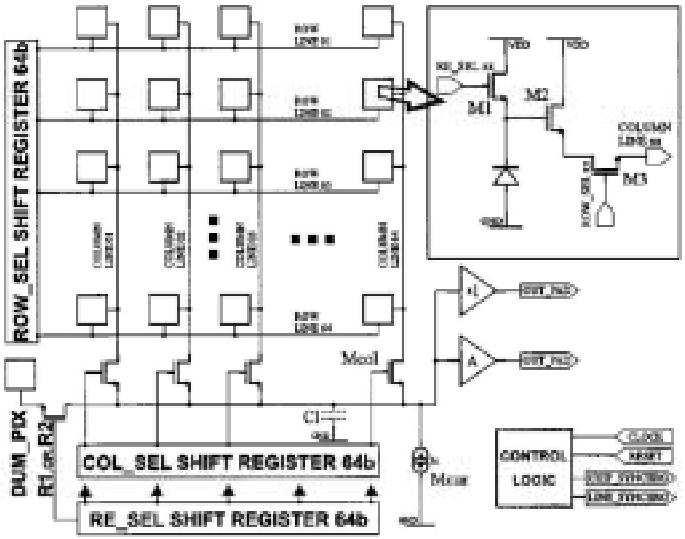


Figure 2.8: Schematic diagram of MIMOSA I.

Each chip was equipped with an analog serial readout circuit. A row shift register and a column shift register were used to serialize the analog signal in each pixel. As an example, the schematic diagram of MIMOSA 1 is shown in Fig. 2.8 [13]. Different CMOS processes have been used to fabricate these chips. As the goal of these chips was only to demonstrate the feasibility of using MAPS for MIPs detection, the basic pixel design has been chosen and only small matrices have been realized. Each pixel only consists of three transistors (Fig. 1.9). The classical 3-T cell operation is carried on for reading out pixel signal.

Although the pixel structure remains similar, different types of charge sensing diodes have been implemented and tested. In addition, in order to test the radiation tolerance of MAPS, some of the pixel matrices were designed by using radiation tolerance layout rules. Other design options, such as several charge collecting diodes per pixel and staggered pixel layout, have been also realized. Moreover, several standard CMOS processes were used and their influences were compared [13] [14]. Tab. 2.1 summarizes the general design features of all these prototypes.

Chip, process, pixel pitch, epitaxial thickness, Frequency	**Chip configuration**	**Sensing element diode size**	**Key design features**
MIMOSA I, 0.6 μm (5V), 20×20 μm^2, 14 ± 2.0 μm, 5MHz	4 arrays 64×64 pixels, square pixel layout	n-well/p-epi diode 3.1×3.1 μm^2	1 and 4 diodes/pixel, analogue serial readout, single output/matrix
MIMOSA II, 0.35 μm (3.3 V), 20×20 μm^2, 4.2 ± 0.2 μm, 25MHz	6 arrays 64×64 pixels, square and staggered pixel layouts	n-well/p-epi diode 1.7×1.7 μm^2	1 and 2 diodes/pixel, analogue serial readout, single output/matrix, use of enclosed transistors
MIMOSA III, 0.25 μm (2.5 V), 8×8 μm^2, 2.0 ± 0.2 μm, 40MHz	2 arrays 128×128 pixels, staggered pixel layout	n-well/p-epi diode 1.0×1.0 μm^2	1 diode/pixel, analogue serial readout, single output/matrix, use of enclosed transistors, varied size of source follow transistor
MIMOSA IV, 0.35 μm (3.3 V), 20×20 μm^2, lightly doped non epitaxial substrate, 40MHz	4 arrays 64×64 pixels, square pixel layout	n-well/psub diode 2.0×2.0 μm^2 and 4.3×4.3 μm^2	1 and 3 diodes/pixel, analogue serial readout, single output/matrix, auto reverse polarisation of diodes, photoFET design

Table2.1: Design features of the fabricated MIMOSA prototypes.

The data process for all these four prototypes had to be performed by software during off-line analysis. Each pixel was sampled twice in order to get the real signal level. One is the reset reference level and the other one is the level after incident charge integration which is called signal level. The useful signal is the difference of these two samples. As the pixel has to be reset after its signal has been read out, the reset noise is present and dominates the total noise level inside the pixel. Fortunately, by implementation of Correlated Double Sampling signal processing [15] [16], this noise can be effectively removed.

As mentioned before, the small prototypes were used to demonstrate the adequacy of the new technique for MIPs detection. That is why different types of CMOS process have been used to fabricate these chips. Although several specific features (the feasibility of using CMOS process for MIPs detection, radiation tolerance, design approaches for low noise operation, etc.) were successfully tested by the small chips, a large scale chip was mandatory because high-energy physics experiments require detector devices with a scale of at least a few square centimeters. In addition, it is also needed to develop the necessary read-out methods matching the requirements for data processing and data transfer in real condition. The MIMOSA V was fabricated for this purpose. The pixel design of MIMOSA V is based on the architecture of MIMOSA I. The basic unit is a full reticule-size device of $19{,}400 \times 17{,}350\ \mu m^2$ [17].

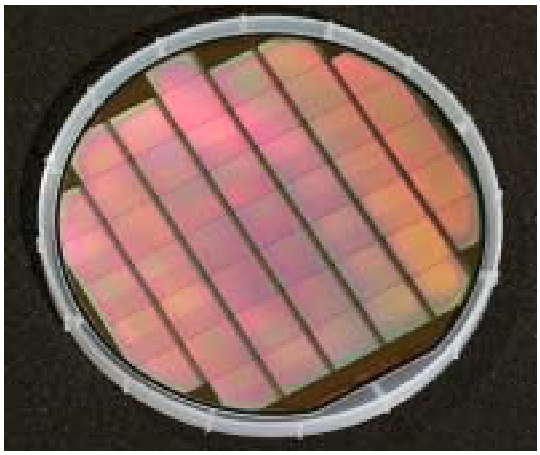


Figure 2.9: MIMOSA V wafer.

B. Performance of these prototypes

All the four chips were calibrated in laboratory by exposing to visible light (by light emitting diode LED) and X-ray radiation (with a radioactive source of ^{55}Fe). High energy beam tests were also performed at DESY and CERN. Very encouraging results have been obtained which demonstrate that, for MIPs detection, the technique of MAPS works very efficiently and provides excellent tracking qualities.

- Noise: Equivalent Noise Charge (ENC) ranges from 6 e^- to 32 e^- much less than signal (80e^- per µm is generated inside the epitaxial layer);
- Particle detection efficiency greater than 99%;
- Spatial resolution: about 1.5 µm has been reached for a 20 µm pixel pitch;
- Radiation tolerance: a resistance to an irradiation of 10^{12} neutrons/cm^2/year and to ionising particles of several hundred kRads has been demonstrated [18].

2.3.2 MIMOSA with on-chip data processing

The first research phase demonstrates that the detection technique of MAPS provides an excellent tracking performance for MIPs. However, the final success of MAPS using for particle physics strongly depends on a successful implementation of on-chip data processing operations, such as hit recognition and sparsification scheme. Moreover, considering the small signal amplitude, special circuit techniques have to be used in order to reduce the noise level. In the year 2001, the team of CEA-SACLAY (DAPNIA) begun to cooperate with the IPHC team and a new prototype, equipped with in-pixel amplifier and column level discriminator, was designed together by them in 2002. The chip is called MIMOSA 6 with the objective of testing the possibility of in-pixel integration of double sampling, which allows reducing FPN and Temporal Noise (kT/C) as well.

The pixel design was based on a principle of switched operation circuits. As the signal level is very small for MIPs detection, an amplifier of a gain around 5~10 [19] was integrated inside the pixel in order to realize column level operation. 15 transistors were used as the switches inside the circuit and 14 transistors were used for signal amplification. The block diagram of pixel concept is shown in Fig 2.10 [20].

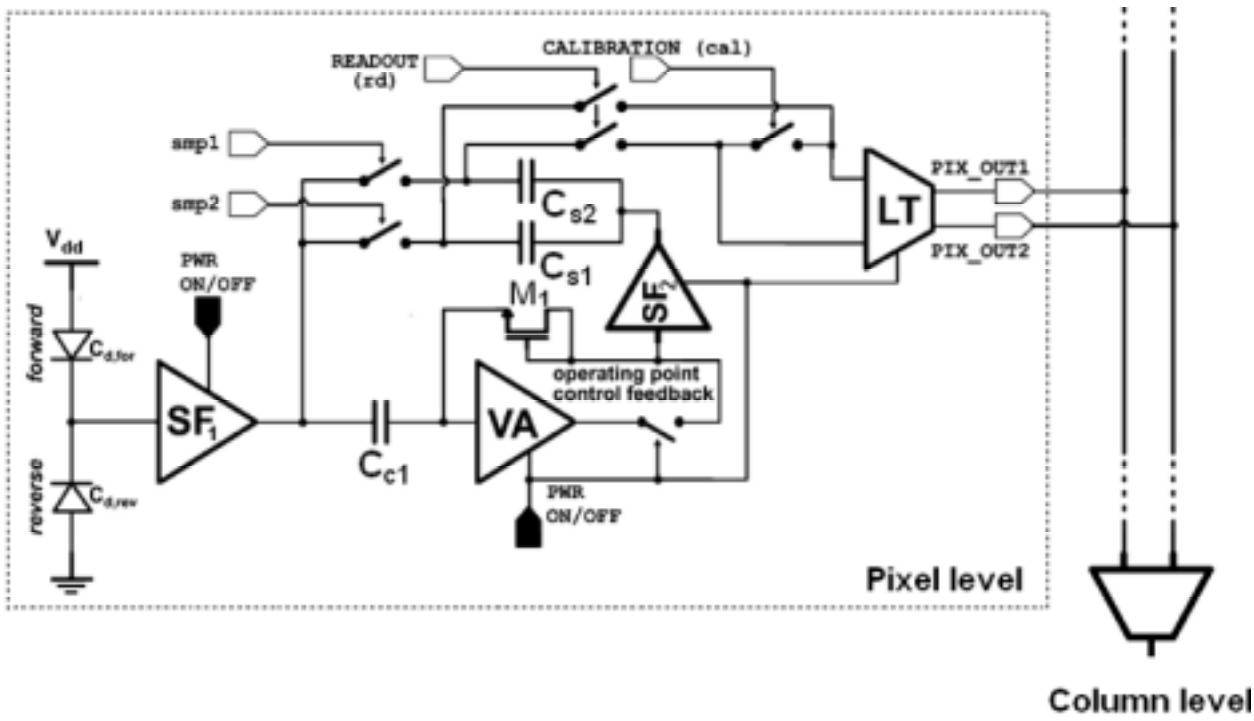


Figure 2.10 Pixel concept of MIMOSA 6.

The on-pixel amplifier comprises three stages:

- The first one is a source follower (SF_1) used to buffer the charge detection diode;
- The second stage is a common source cascode amplifier (VA) AC-coupled by capacitor C_{c1} to SF_1. A feed-back loop of the amplifier is formed by transistor M_1. The bias current of the amplifier is stabilized by this feed-back loop. Two storage capacitors (C_{S1} and C_{S2}), driven by another source follower SF_2, are alternatively connected to the feed-back loop;
- The third stage is the linearized transconductance (LT), providing differential output current.

A column level auto-zero comparator is also realized in MIMOSA 6. One comparator is shared by all the pixels of the same column. The width of the comparator is only 28 µm which matches the pixel pitch. By comparing the input analog signal with its threshold value, it transforms the analog signal into a 1-bit binary code. It is the basic component of more complicated data auto processing circuits (analog-to-digital converter for example). More details concerned the design of comparator are described in the following section.

The main features of MIMOSA 6 chip are summarized in the following table.

MIMOSA VI				**features**		
performances of chip designed				noise ENC: ~20 e^-, conversion gain: 6.5 nA/e^- pixel pedestal variations: ~120 e^-		
discriminator performances						
τ_1 ns	τ_2 ns	τ_3 ns	τ_4 ns	noise inp.ref.	offset inp. ref.	power
90	15	45	30	~85 μV_{rms}	negligible	~200 µW
75	12.5	62.5	25	~100 μV_{rms}	~300 µV	~200 µW
60	15	45	30	~100 μV_{rms}	~800 µV	~200 µW

Table 2.2: Main parameters of the MIMOSA 6 chip [20].

Although the MIMOSA 6 chip validated the possibility of in-pixel sampling and storing signals from different time slots, the FPN value of the chip was too high (more than 120 e^-) due to the complexity of the pixel architecture. The functionality of the column level discriminator was verified. With different configurations (the periods for τ_1, τ_2, τ_3 and τ_4), the input referred noise is around 100 μV_{rms} and the input referred offset is less than 800 μV_{rms}. These results show that the CDS operations realized inside column level comparators work very efficiently. More details can be found in references [20] and [36].

2.4 The fast MAPS with in pixel amplification and reset noise suppression for MIPs detection

The MIMOSA 6 chip, designed with a high gain pixel level amplifier, stepped the first attempt towards the auto on-chip data processing. However, due to the complexity of pixel architecture, the FPN value is too high to achieve the correct column level operation. Since every stage added on the signal path contributes to the total noise level, alternatively to use a complicated high gain amplifier, another solution is to simplify the pixel architecture, using a simple amplifier with moderate gain, in order to reduce the noise sources. Based on this principle, a new prototype chip with a new all-NMOS pixel architecture which integrates in-pixel CDS operation were then designed by CEA-SACLAY and IPHC. Each pixel comprises only 8 transistors. Two CMOS processes with different thickness of epitaxial layer were chosen to realize these prototypes. At the end of 2004, the first prototype was fabricated by TSMC 0.25 µm CMOS digital process with about 8 µm epitaxial layer thickness. This chip is called MIMOSA 8. As the input signal level is proportional to the thickness of the epitaxial layer, a second prototype of the same architecture was designed in AMS 0.35 µm OPTO process, which provides thicker epitaxial layer, in the year 2006. It is called MIMOSA 16. Moreover, this AMS process provides two different thicknesses of the epitaxial layer, so that two versions are realized for MIMOSA 16: M16_14 was with an epitaxial layer of about 14 µm and M16_20 was with an epitaxial layer of about 20 µm.

The main objective of these prototypes is to reach the detection performance which was demonstrated by the MIMOSA 1 to MIMOSA 5 series with some important new aspects:

- realize pixel level CDS operation;
- reach a high readout speed, about 20 μs/frame, there are 128 lines in one frame;
- integrate column level comparator together with the pixel array and realize 1-bit binary output.

Several issues should be carefully considered in order to achieve the above objective.

Firstly, the direct implementation of structures suitable for visible photon detection is not sufficient to overcome the residual pixel-to-pixel and column-to-column FPN due to the low Charge-to-Voltage conversion Factor (CVF) achievable for n-well/p-epi diodes in mainstream CMOS technologies. Therefore, some in-pixel amplification of the signal is mandatory to achieve a higher CVF. Generally, the amplifier has to be placed as close as possible to the charge detection diode in order to obtain an optimum noise performance. But the adding amplifier introduces its offset to the total pixel noise. Special care is needed to cancel the offset due to the in-pixel amplifier.

Secondly, in order to reduce the reset noise of classical 3-T photodiode pixels, some methods were recently presented [21]-[27]. In these methods, the CDS processing, achieved by the implementation of in-pixel capacitors, was used to suppress the reset noise and pixel-to-pixel offset non-uniformities. However, these methods usually need long duration of the reset for a good reduction of noise and practically present high FPN due to the increased number of transistors and capacitors used for the pixels. As a result, several important modifications have been made in the pixel design for the purpose of increasing readout speed and reducing FPN.

2.4.1 Architecture of the prototype

The prototype consists of two main parts: an analog part and a digital part. The global architecture diagram is shown in Fig. 2.11.

The analog part mainly consists of an array of 32×128 pixels and 24 column-level discriminators for signal sparsification. The pixel array is divided into four sub-arrays of 32×32 pixels each. The pixel pitch is 25 μm. Binary readout is realized thanks to the column level discriminators connected to the 24 first columns. The remaining eight columns are connected directly to the buffer of output pad for analog readout.

The digital part, containing a fully programmable digital sequencer block and a serializer block, generates the control logic necessary for the analog part. Thanks to the programmable sequencer, the timing patterns are loaded into the chip during the initial phase of programming. The outputs of 24 column level discriminators are multiplexed by the serializer block.

The chip readout is organized in columns processed in parallel. The first 24 columns are connected to discriminators, multiplexed onto 4 outputs. The last 8 columns are connected

directly to the analog outputs omitting the discriminators. Their analog outputs can be directly observed on the output pads. These direct analogue outputs are very important for evaluating the performance of different pixels. An internal point of the amplifier before the clamping, capacitance used for the CDS processing, was made available for the last column of pixels on one additional output pad of the eights column. This additional line allows to examine the output level of the amplifier.

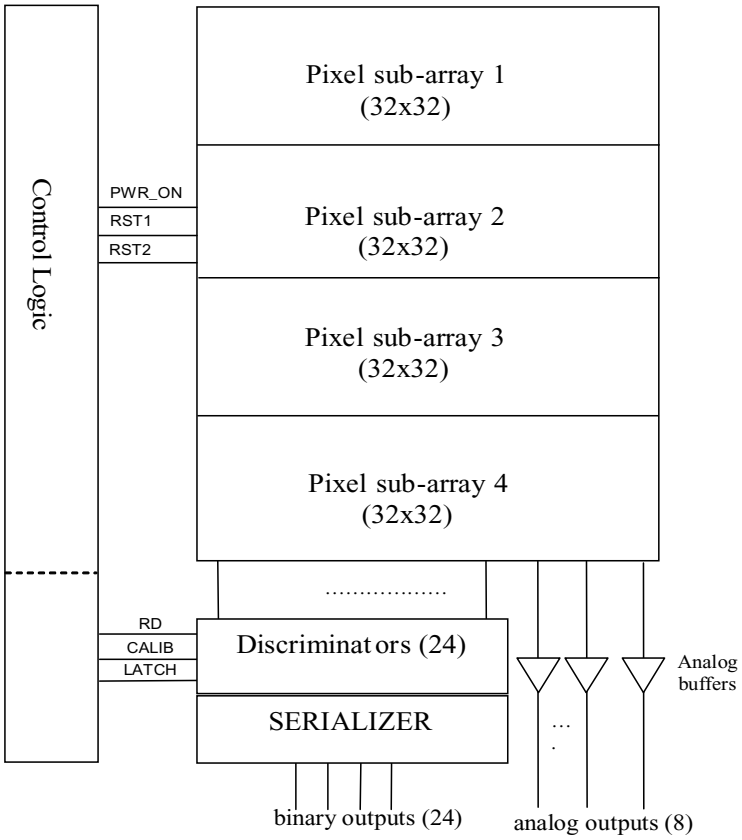
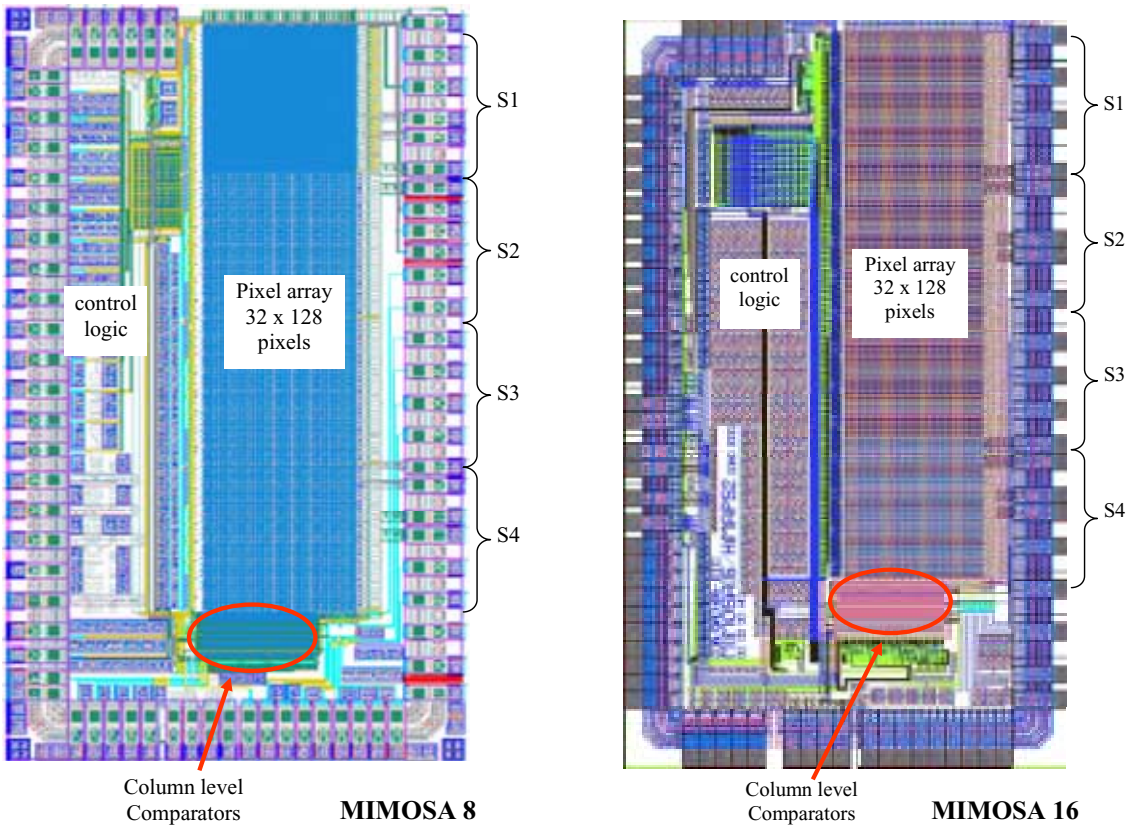


Figure 2.11 Architecture of the prototypes.

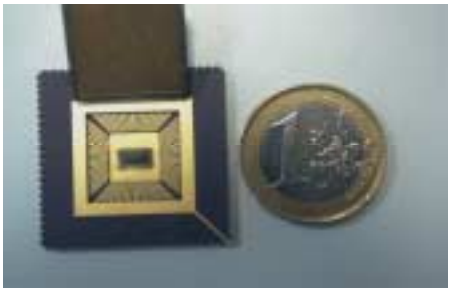
For MIMOSA 8, two types of detection techniques are used for charge detection diode: the AC-coupling diode designed by IPHC-Strasbourg and the DC-coupling diode designed by CEA-SACLAY. The first sub-array (S1) from the top of the chip is made up of pixels with AC-coupling diodes. The bottom three sub-arrays consist of pixels with DC-coupling diodes of different sizes [28]: 1.2×1.2 μm² for sub-array number 2 (S2), 1.7×1.7 μm² for sub-array number 3 (S3), 2.4×2.4 μm² for sub-array number 4 (S4). The diode size influences the equivalent total capacitance which determines the CVF value. From equation (1.1), higher CVF value can be obtained if the equivalent capacitance is small. But the diode with small size collects less charge and therefore degrades the charge collection efficiency. Thus, in order to find out the best compromise between CVF value and efficiency of charge collection, different diode sizes are used in this chip.

For MIMOSA 16, the first sub-array (S1) and the second sub-array (S2) are made up of the DC-coupling pixels with different diode size aiming at different values of CVFs. The diodes size is 1.7×1.7 μm² for sub-array S1 and 2.4×2.4 μm² for sub-array S2. The sub-array S3 use the same pixel architecture and diode size as S2 but radiation tolerance design rules are implemented in order to study the radiation tolerance. The last sub-array S4 is made up of pixels with self-bias diode designed by IPHC of Strasbourg.

The layout of MIMOSA 8 and MIMOSA 16 are shown in the figure below, respectively:



(a) Layout of MIMOSA 8 (TSMC CMOS 0.25 µm digital process) and MIMOSA 16 (AMS CMOS 0.35 µm OPTO process)



(b) Photograph of the real chip MIMOSA 8.

Figure 2.12: (a) The layout of MIMOSA 8 and MIMOSA16 prototypes and (b) photograph of the chip MIMOSA 8.

Thin gate oxide transistors were used for the digital part (V_{DD}=2.5V) and thicker gate oxide transistors for the analog part (V_{DD}=3.3V). 2.5V/3.3V translators were introduced to interface the sequencer and the analog part. Digital I/O pads include also 2.5V/3.3V or 3.3V/2.5V translators. All external digital signals are sent in single ended CMOS 3.3V logic.

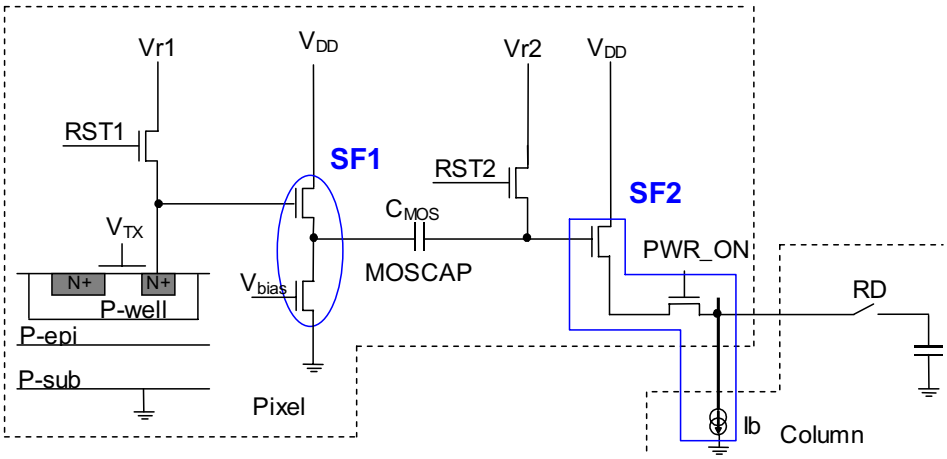
2.4.2 Pixel design

Two types of diode were used in these prototypes: the DC-coupling pixel and the AC-coupling pixel. The DC-coupling pixel uses a transistor as switch to reset the charge detection diode. In the AC-coupling pixel, the reset transistor is replaced by a diode.

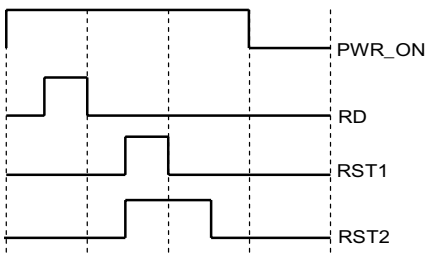
A. DC-Coupling Pixel Design

The success of pixel design strongly depends on the efficiency of noise reduction. While the influence of shot noise is small thanks to high readout speed, the classical 3-T photodiode pixel suffers from the reset noise, temporal readout noise and FPN. The reset noise varies from one pixel to another. It has to be eliminated in order to perform column level discrimination. As been mentioned above, in order to reduce the temporal readout noise, an amplifier has to be added and be placed very close to the detection diode. Moreover, the FPN results from the offset of the mismatches among the transistors and is also crucial for on-chip data processing. Each transistor implemented on the road of signal transfer contributes to the total offset level. Considering the small signal level, special care is needed to remove these offsets while maintaining the desirable readout speed.

The pixel architectures already successfully implemented in the systems for visible photon detection show very good examples of removing or reducing the reset noise [21]-[27]. The reset noise suppression method for the visible light applications proposed in [21] is shown in Fig. 2.13.



(a) Pixel architechture with reset noise cancellation presented in [21].



(b) Timing diagram of the pixel structure presented (not to scale).

Figure 2.13: Pixel structure with reset noise cancellation and its timing diagram.

Compared to the 3-transistor structure, a Source Follower (SF1), a series capacitor and a switch (RT2) were added. The operation is as follows.

- During offset sampling phase, both switches RST1 and RST2 are closed, so that the offset of SF1 is memorized on the series capacitor. The value is $AV_{off,SF1}$, where A is the gain of SF1.

- When entering the integration phase, the RST1 open firstly. Because RST2 remains closed at this moment, the reset noise of the RST1, attenuated by the gain of SF1, is also memorized on the series capacitor. Then RST2 opens and V_{TX} turns on, the integration begins. The value memorized on the capacitor now is equal to: $A(V_{off,SF1} + V_{thm,rst1})$.

- After integration, when the signal is readout, the offset of SF1 and the reset noise of RST1 could be compensated thanks to the information stored on the series capacitor. The following operation is performed:

$$A(V_{sig} + V_{off,SF1} + V_{thm,rst1}) - A(V_{off,SF1} + V_{thm,rst1})$$

The remaining value after readout is then equal to A V_{sig}, free of the offset of SF1 and the reset noise of RST1.

What is not mentioned in the above description is the reset noise of RST2, which is not cancelled at all. But this noise voltage is $\sqrt{kT/C_{MOS}}$, so that it could be made arbitrarily small by choice of a large C_{MOS} value.

Although this pixel structure can efficiently cancel the offset of SF1 and the reset noise of RST1, it is difficult to use it directly for the applications of MIPs detection.

- The charge detection element is a photodiode with readout floating capacitor separated by a transfer gate. Although high CVF can be achieved by reducing the value of the readout floating capacitor, this structure suffers from a charge transfer problem [28] [29] because the transfer transistor (V_{TX}) operates in sub-threshold regime;

- The detected signal is attenuated by the gain of SF1(~0.85);

- The offset of SF2 is not canceled at all;

- The pixel consumes continuously, so that the power dissipation is an issue when millions of pixels is needed.

Derived from the pixel structure presented above, the DC-coupling pixel architecture is proposed in [30] [31]. The pixel structure is shown in below (Fig. 2.14).

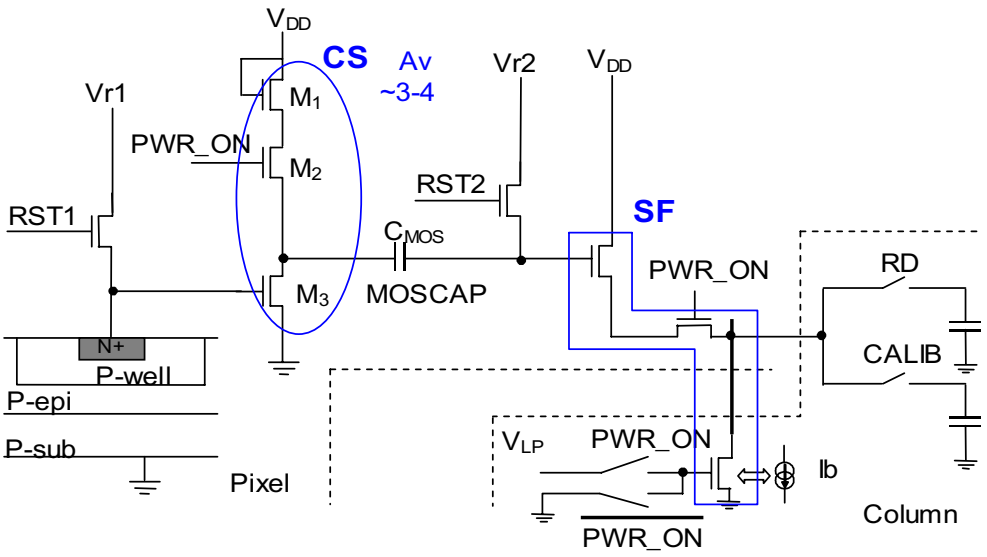


Figure 2.14: Pixel architecture: double sampling and auto-correction circuit are integrated inside pixel.

Compared to the pixel structure of Fig. 2.13 (a), several important changes were made.

- A simple n-well/p-epi diode is used as charge sensing element. The equivalent capacitor is then determined by this diode and its parasitic capacitance;
- SF1 was replaced by a Common Source (CS) preamplifier. The signal would be amplified before readout, so that it is more robust to the noise of column level circuit and a higher CVF could be obtained;
- The CS stage is biased only during the pixel readout phase, which is very short comparing to the time needed for the integration. Therefore, power consumption is largely reduced.
- An additional phase (CALIB) is added in order to memorize the offset due to SF2;
- Moreover, the gate voltage of the bias transistor of output source follower is also switched so that the discharge on the capacitance of output bus is avoided. This limits the signal variations on the column bus and increases the readout speed.

The CS pre-amplifying stage is placed very close to the charge detection diode. A double sampling circuitry is made up of this CS preamplifier, a serially connected capacitor (C_{MOS}) and two reset switches. The first switch (RST1) is used to reset the detection diode and the second one (RST2) is used to memorize on the capacitor the offset of the preamplifier and the reset level of the diode. A Source Follower (SF) and a row select switch are used to output the signal on the common data bus. The switch RD and CALIB are the column level commands and are used for memorizing the output signal level and the column reference level, respectively.

The timing diagram of the pixel is shown in Fig. 2.15 below.

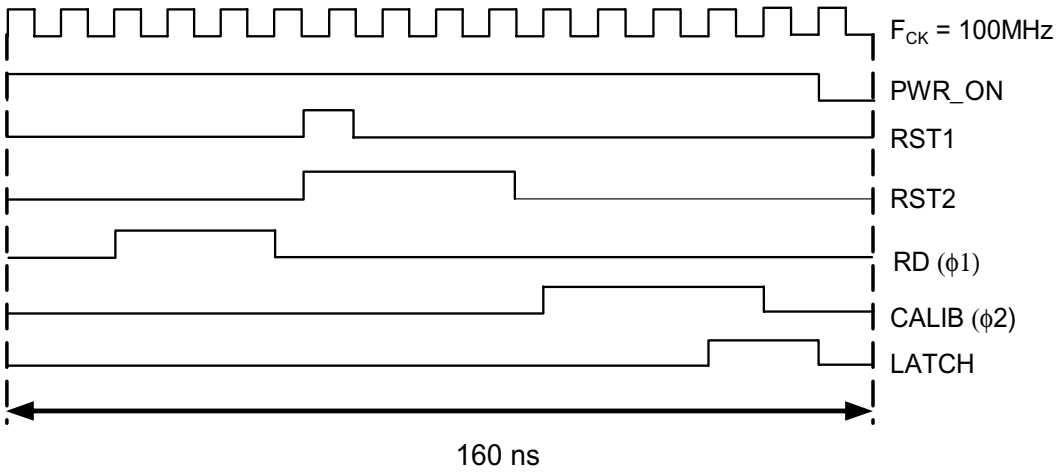


Figure 2.15: Timing diagram: the total time needed for reading a pixel is 160 ns at 100MHz.

The short intervals (tens of μs) between two consecutive readouts in this application give us the opportunity to store easily the reset level of the charge detection diode in the pixel. After RST1 and RST2, this reset level is memorized on C_{MOS}. The reset noise due to RST1 is also memorized because RST2 is still "on" when RST1 is turned off. Moreover, offset mismatches of the amplifiers are memorized too. During the next readout of the pixel, all these noise are automatically eliminated by subtracting from the signal. The threshold voltage mismatches of the SF are successively corrected by a second double sampling process performed at the column level. The voltage V_{RD}, sampled by the readout circuitry during the RD phase (readout of the signal), is the signal which also contains the offset of the SF stage. The voltage V_{CALIB}, sampled during the CALIB phase (readout of the referenced level) just after the reset, is the reference level containing the offset of SF stage. The useful signal is the difference between these two levels, free from the offset mismatches of the SF stages. The noise sources not being removed after the above process are the reset noise of RST2 and the 1/f noise. However, the reset noise of RST2 can be reduced by selecting a large value of C_{MOS} and the 1/f noise is very small because the pixel is designed for high speed operation.

This pixel achieves high CVF (about 50~70 μV/e- in simulation) using only 8 transistors in the pixel. The amplifier is based on the NMOS transistor common source architecture with NMOS transistor diode connected load, both operated in strong inversion. The total voltage gain is the ratio of the transconductances of the current source transistor and the load one. The DC current bias of the amplifier is determined by the voltage across the charge collecting diode. SPICE simulations shown that the pixel can be read in 100 ns. In the timing shown in Fig. 2.15, additional 40 ns for discrimination and 20 ns to simplify the digital part of the whole chip are added. Using this timing diagram, the corresponding readout time of one frame is then given by:

$$T_{r.o} = N_{row} \times 16 \times \frac{1}{F_{ck}} \qquad (2.13)$$

where N_{row} is the number of rows of the pixel array and F_{ck} is the main clock frequency.

The performance of the DC-coupling pixel is studied by using CADENCE microelectronics simulation environment. The detailed simulation results and the key points for its operation are discussed in section 2.5.

B. Optional AC-Coupling Pixel Design*

The pixel design with the direct AC-coupling of the amplifier to the charge sensing diode was used as a test structure for the current chip design. The pixel architecture uses an auto-reverse polarized charge collecting diode [30], in pixel amplifier AC-coupled to the charge sensitive element and circuitry for CDS (Fig. 2.16).

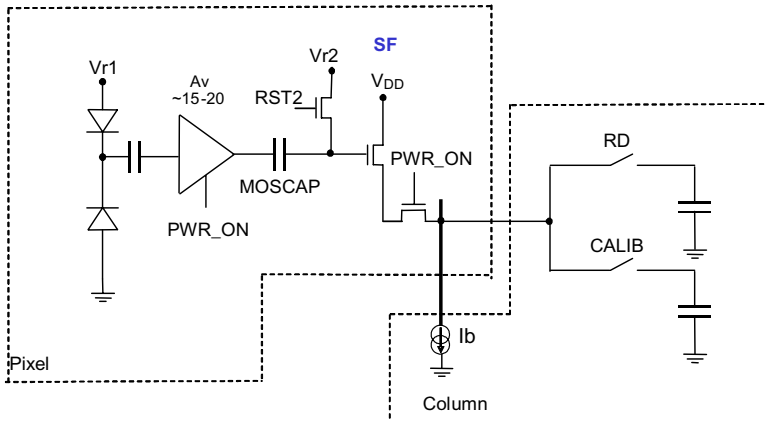


Figure 2.16: Pixel architecture: auto-reverse polarized charge collecting diode with double sampling auto-correction circuit

In order to well understand the principle of the AC-coupling pixel, the principle of the direct AC coupling of the auto-reverse polarized charge sensitive element and the amplifier is shown in Fig. 2.17 below:

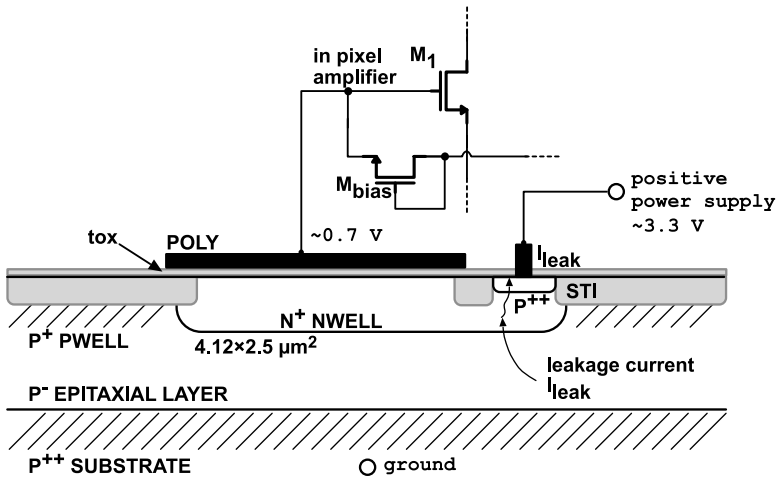


Figure 2.17: principle of the direct AC coupling of the auto-reverse polarized charge sensitive element and the amplifier.

* The AC-coupling diode is not concerned in the following discussion and tests.

The CDS circuitry is similar to that used in the DC-coupled pixel version. The charge sensitive element is a two diode system, with an n-well/p-epi diode, collecting the charge available after particle impact, and a p-plus/n-well diode, providing a constant reverse bias of the first one. The signal of the charge sensitive element is delivered to the amplifier via the series capacitance obtained by placing the polysilicon plate over the n-well area, as it is shown in Fig. 2.17. The gate oxide is used providing a high value of the coupling capacitance. The signal is then amplified with a voltage gain about 15-20, aiming at total Conversion Factor (CVF) about 150 μV/e-. The choice of AC-coupling instead of DC-coupling allows independent bias of the input transistor in the amplifier from the potential settled on the n-well region during the detector operation. At the same time, the n-well diode is polarized with the maximum voltage available in the technology process bearing the optimization of the charge collection process. Since the second pixel does not use reset transistor for the diode, the CDS is used to extract the signal in subtraction from the reference level. The reference level for each new measurement is the state from the previous readout cycle.

C. Radiation tolerant pixel design

A new sub-array with pixels specially designed for radiation tolerance was implemented in MIIMOSA 16. While using the same architecture as DC-coupling pixel, special layout design rules were applied in order to improve the radiation tolerance of pixel.

Two effects of radiation damages are considered here: the dose effects damage that the silicon interacts with charge particles such as electrons and photons and the displacement damage that the silicon interacts with heavy particles such as neutrons. Other radiation phenomenon such as latch-up, Single Event Latchup (SEL), Single Event Upsets (SEU) and Single Event Transient (SET) are not concerned here.

The dose effect damage transfers the energy of radiation to the electrons of material without damages the crystal lattice. Excessive charge will be created inside the material. For a MOS transistor, it is well known that the thick local oxide is the most sensitive part of charge build-up [33]. Positive charge will be created close to the silicon interface. As a result, the threshold voltage of transistor will be changed (Fig. 2.18).

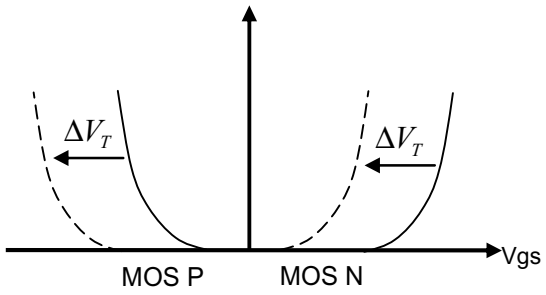


Figure 2.18: The threshold voltage varies due to the charge generated inside oxide.

$$\Delta V_t = -\frac{\Delta Q_{ox}}{C_{ox}} \qquad (2.14)$$

where C_{ox} is the equivalent capacitance of the oxide and is given by:

$$C_{ox} = \frac{\varepsilon_{ox}}{e_{ox}} \qquad (2.15)$$

where e_{ox} is the thickness of oxide.

Moreover, the gate oxide to field oxide bird's beak region forms a non-planar parasitic transistor and leakage current of transistor is caused by the inversion of the silicon in this region due to positive charge build-up in the oxide. Usually, the parasitic structures are located at the edge of the gate where the poly silicon overlaps the bird's beak which limits the active area (Fig. 2.19). For such a structure the parasitic transistors are in parallel with the main transistor. Studies show that the most sensitive part of the bird's beak region is determined by the maximum of the effective thickness of the oxide [34].

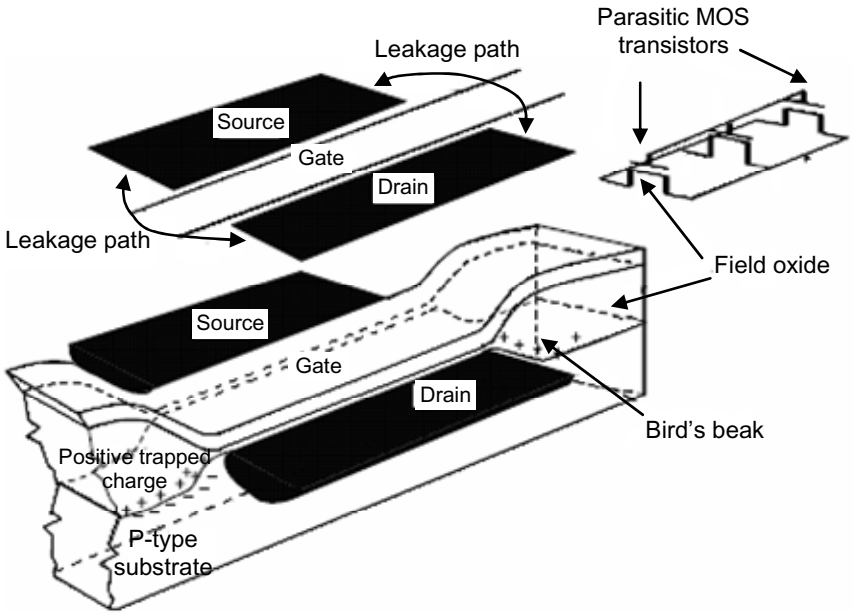


Figure 2.19: Parasitic transistor formed by gate oxide to field oxide bird's beak region.

In displacement damage, the energy of incident heavy particles will be deposited on the crystal lattice and the atom will be kicked away from its original position. It is also called bulk damage. Created by atom displacement, the defect points and the defect clusters reduce the lifetime of the electrons inside the epitaxial layer. Two results due to this phenomenon:

- Firstly, the generation-recombination rate will be increased. It means that the leakage current increases in the depleted region of charge detection diode. This current is directly related to the electron lifetime and normally increases with the density of incident particles.

- Secondly, the free drift length (or the diffusion length) is reduced which means that more charge will be lost before collection.

While the damage of non ionising radiation can not be reduced by applying special layout design rules, the ionising radiation damage can be reduced by using enclosed geometry

transistors and by using guard-rings [35]. The enclosed geometry transistor thins down the field oxide to reduce the charge created by ionization. The guard-ring cuts the path of surface leakage current induced by positive charge build-up on oxide after irradiation. In MIMOSA 16, the guard-ring structure is used in the pixels of sub-array S3 and the layout of this pixel is shown in Fig. 2.20. Here, the size of diode is the same as the diode size used in sub-array S2.

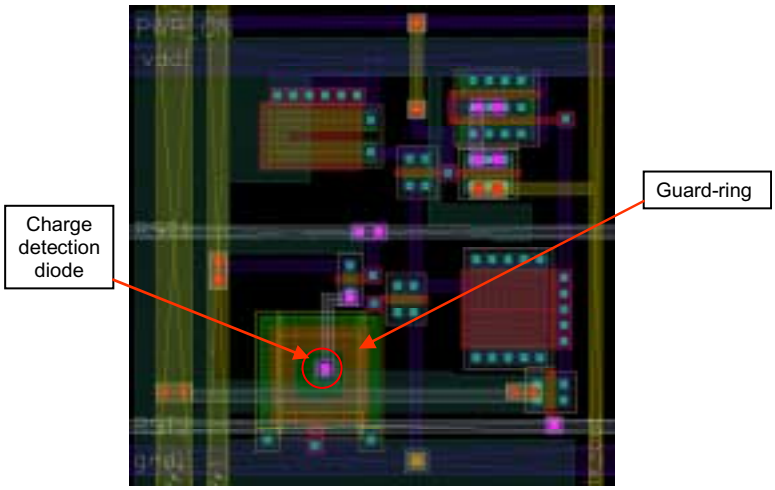


Figure 2.20: Layout of pixels: guard-ring is used around the charge detection diode to improve radiation tolerance.

2.4.3 Column level discriminator

The full digital output data and on-chip real time data processing are required for the final vertex detector of future ILC. The comparator realized on the bottom of each column is the first step towards these goals. It transforms the output analog signal of pixel into a 1-bit digital code. Considering the small value of analog signal (which means that the sensitivity of comparator has to reach a few hundreds of μV), the offset of mismatch resulting from various sources (the pixels reference level, the output buffers of column bus and the charge injected by the necessary switches) and the very high signal processing speed (at 100 MHz, the period is 160 ns which is the same as that of reading a pixel), the design of this comparator is not a trivial task.

For the comparator design, it is mandatory to use an offset compensated amplifying stage, which corrects the residual offset of comparator. Two main techniques are widely used: the IOS (Input Offset Storage) and the OOS (Output Offset Storage).

The architecture of IOS is shown in Fig.2.21 below:

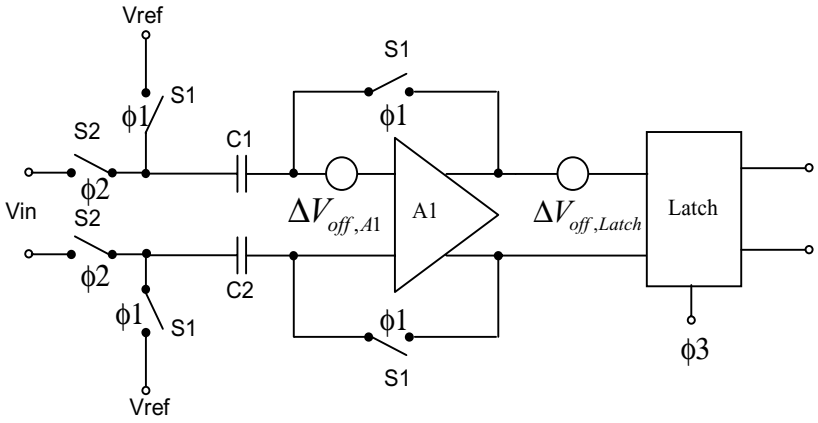


Figure 2.21: IOS architecture.

In this mode, S1 is firstly switched on, S2 is off. A close unit-gain feedback loop is formed and the offset is stored on C1 and C2. Then, S1 is switched off, S2 is on. During this stage, the offset is automatically suppressed while amplifying the input signal. The total residual input referenced offset after an IOS stage is given by [31], [32]:

$$V_{os,total} = \frac{V_{off,A1}}{1+A1} + \frac{\Delta Q}{C} + \frac{V_{off,Latch}}{A1} \qquad (2.16)$$

where A1 is the gain of amplifier, $V_{off,A1}$ is the offset of amplifier, $V_{off,Latch}$ is the offset of latch and ΔQ is the charge injected by S1. Although the IOS achieves a wide input range, it features a large input capacitance and its offset cancellation ability requires a high gain, obviously an issue on high speed operation.

The architecture of OOS is shown in Fig. 2.22.

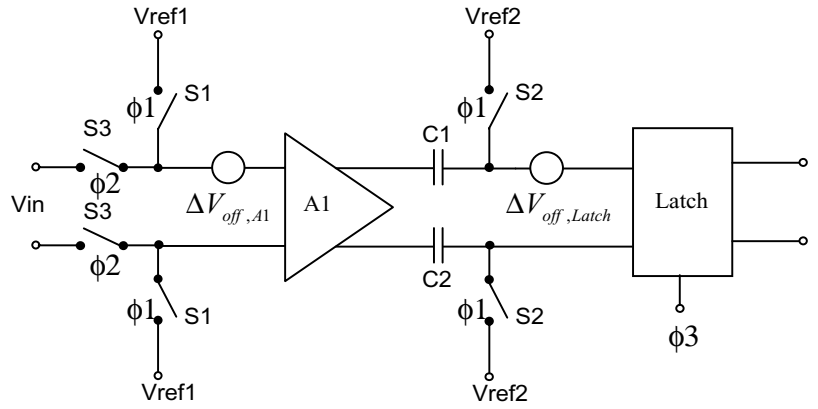


Figure 2.22: IOS architecture.

In order to store the offset, S1 and S2 are firstly on, S3 is off. Similar to the IOS mode, the offset is stored on serious capacitors C1 and C2 during this stage. Then, S3 is on while S1 and S2 are off. The signal is sampled and the offset is automatically corrected. Unlike the IOS mode, the offset of amplifier can be totally cancelled by OOS technique. Hence, the total residual input referenced offset after an OOS stage is [30], [31]:

$$V_{os,total} = \frac{\Delta Q}{A1 \cdot C} + \frac{V_{off,Latch}}{A1} \qquad (2.17)$$

where A1 is the gain of amplifier, $V_{off,Latch}$ is the offset of latch and ΔQ is the charge injected by S3. The advantage of OOS is its low input referenced offset value. There are two main drawbacks of this architecture. Firstly, a high gain can not be used in this architecture in order to avoid saturation of the preamplifier. The typical gain value is less than 20 [32]. Secondly, this structure suffers from input DC-coupling which limits the input common-mode range.

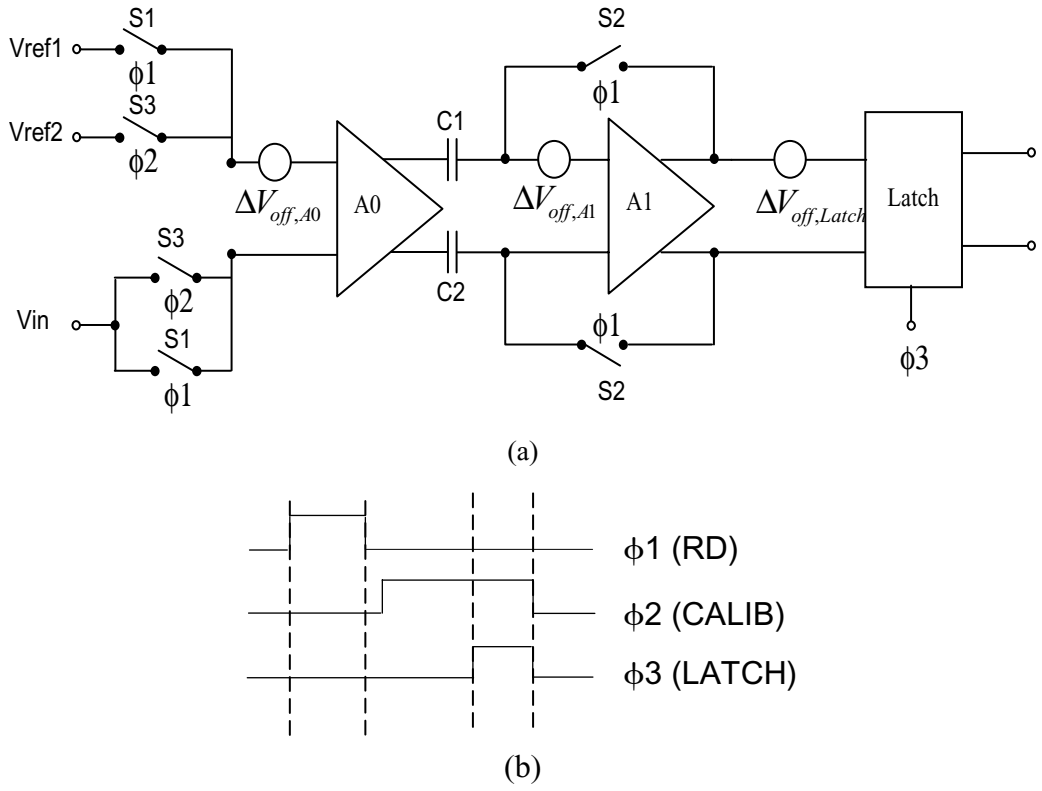


Figure 2.23: Column level comparator: (a) architecture and (b) timing diagram (not to scale).

Considering the advantages and the drawbacks of both IOS and OOS architecture, a combination of them has been chosen to realize the high speed column level comparator for MIMOSA 8 chip. The architecture of this comparator, consisting of an auto-zeroed preamplifying stage and a dynamic latch, is shown in Fig. 2.23 above.

It is an improved version of the previous design MIMOSA 6 presented in [20] [36]. The input switches have been modified to sample voltage signal values. Instead of poly-poly linear capacitors, MOS capacitors were used because the fabrication process does not feature this type of capacitors. To obtain a good linearity, a care had to be taken to bias the capacitors in the deep inversion regime. Level shifters before capacitors were used for this purpose. Thanks to the MOS capacitors, which have small dimensions, the size of the comparators is the same as the ones used in the previous design (220 μm × 25 μm).

Three phases are needed for a complete auto-compensation operation.

- Firstly, during $\phi 1$, S1 and S2 are switched on, S3 is switched off, the residual offset of preamplifier A0 and A1 is memorized in series capacitors C1 and C2. Besides, because $\phi 1$ is equivalent to the RD phase used for reading the pixel output signal of charged particle (Fig. 2.15), this signal is then read out and stored in C1 and C2. Moreover, as the pixel output signal is buffered by source follower (SF), a supplementary value resulting from the mismatch of source follower is also stored in capacitors: the pixel output stage offset, which varies from one pixel to another. This offset is corrected during next phase.

- Secondly, during $\phi 2$, S1 and S2 are switched off, S3 is switched on. The comparator enters its auto-compensation mode. As $\phi 2$ is equivalent to the CALIB phase, the pixel output stage offset value is read out once and automatically used to compensate the offset value stored in series capacitors during $\phi 1$. The threshold value (Vref2) of comparator is also sampled during this phase and is compared with the pixel output signal value.

- At last, at the end of $\phi 2$, $\phi 3$ (LATCH) begins and actives the latch, which rapidly amplifies the difference between the pixel output signal level and the comparator threshold level. A logical signal is given according to the level of difference. If the level is positive, the output of latch is 1; on the contrary, it is 0.

In addition of compensation of the pixel output stage offset and the amplifier residual offset, the charge injected by the switches used in comparators and the parasitic effect related to substrate coupling in the mixed-signal environment are also effectively reduced by the fully differential architecture. The total input referred residual offset of this comparator is given by [36]:

$$V_{OSR} = \frac{\Delta V_{off,A1}}{A_0(1+A_1)} + \frac{\Delta Q}{A_0 C} + \frac{\Delta V_{off,Latch}}{A_0 A_1} \qquad (2.18)$$

where A_0 and A_1 are the gain of preamplifiers, $\Delta V_{off,A1}$ and $\Delta V_{off,Latch}$ are respectively the offsets of preamplifier 2 and latch, ΔQ is the total charge injected mismatch from S2 and C is the equivalent series capacitance.

2.4.4 Digital part of the chips

The digital part generates the patterns necessary for addressing, resetting and double sampling the signals in pixels and discriminators in a column parallel way. The pattern is fully programmable allowing testing the chip under different timing conditions. The corresponding waveforms are shown in Fig. 2.15.

The block diagram of the digital part is shown in Fig. 2.24 below.

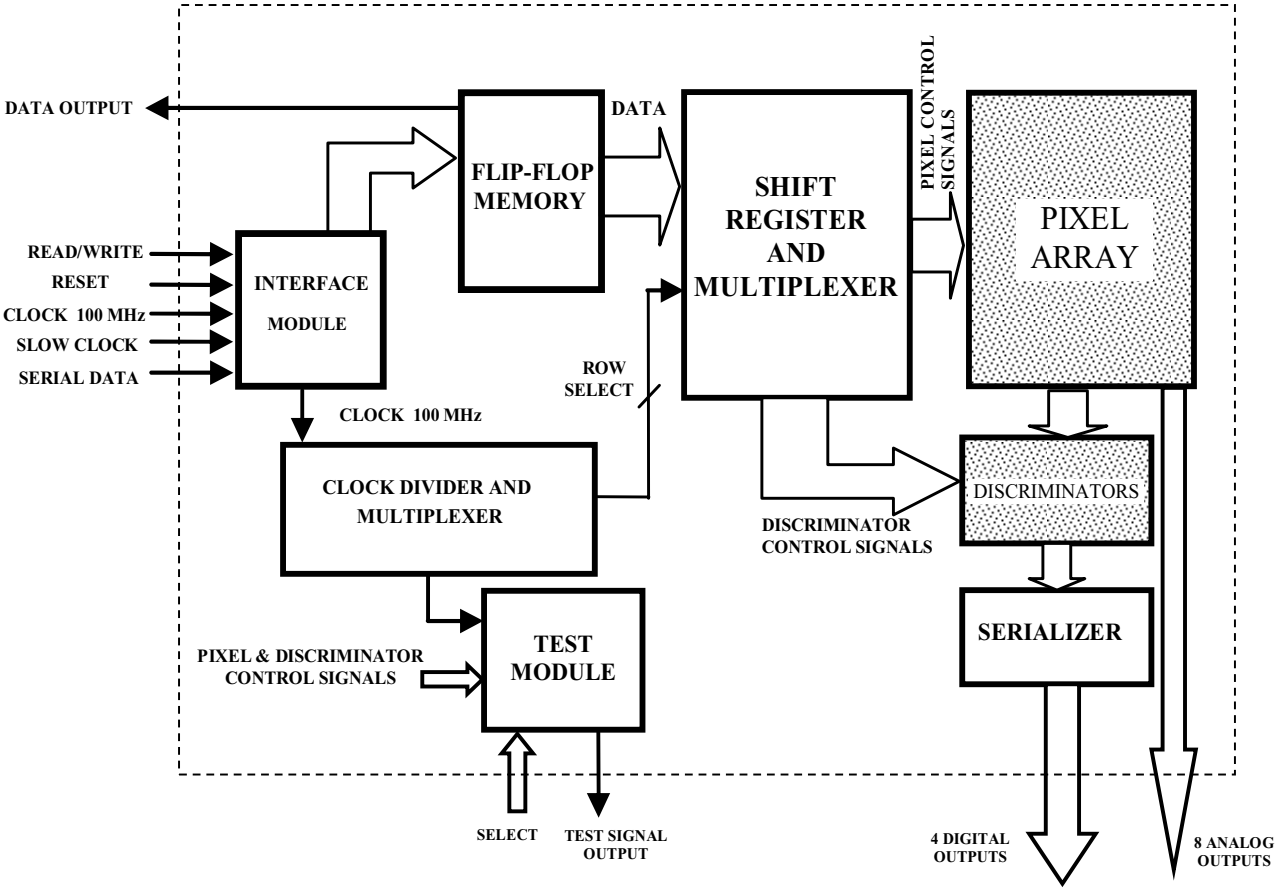


Figure 2.24: The bloc-diagram of digital part of MIMOSA 8. The corresponding analog part is shown in grey (Fig. 2.11).

The working principle of the digital block is as follows. The timing pattern is loaded to the chip during a programming phase. The low speed synchronous serial interface is used for this purpose. The end of the programming phase starts the normal readout operation. The readout starts with the first row. Moreover the resynchronization of the readout can be done anytime sending a short pulse to the chip via the serial interface. An external high speed CMOS signal is used for clocking the chip (max f_{CK} is 100 MHz). The rows are selected sequentially every 16 clock cycles using a multiplexer and the readout pattern is applied to each row selected. The serializer bloc realizes a temporal multiplexing of the binary outputs signals (column discriminators) at a frequency value half that of the main clock frequency.

The test module allows the observation of some internal control signals. One of these signals, generated at the beginning of the readout of the first row, is used to synchronize the chip with the data acquisition board. The chip sends its own clock signal synchronous with the analog data available for the last 8 columns. This clock can be used as an analog-to-digital conversion clock for the data acquisition system.

2.5 DC-coupling pixel performance study

The electrical simulations are performed to study the performance of the DC-coupling pixels. At last, the key points ensuring the working conditions of the DC-coupling pixel are given.

2.5.1 The Common Source (CS) stage

The design of the CS stage is crucial to the pixel performance. A high gain is desirable to achieve a high CVF but it generally requires a high output resistance and increased input capacitance degrading the speed and the noise performance of the circuit.

The CS stage used inside the DC-coupling pixel is represented in figure 2.25:

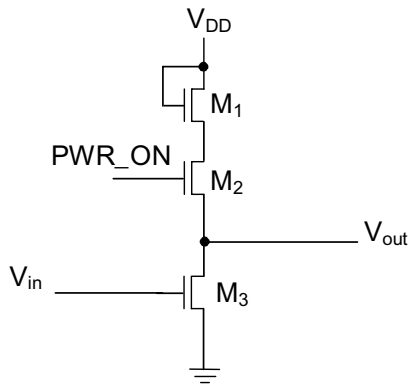


Figure 2.25: Common source amplifier stage used inside the DC-coupling pixel

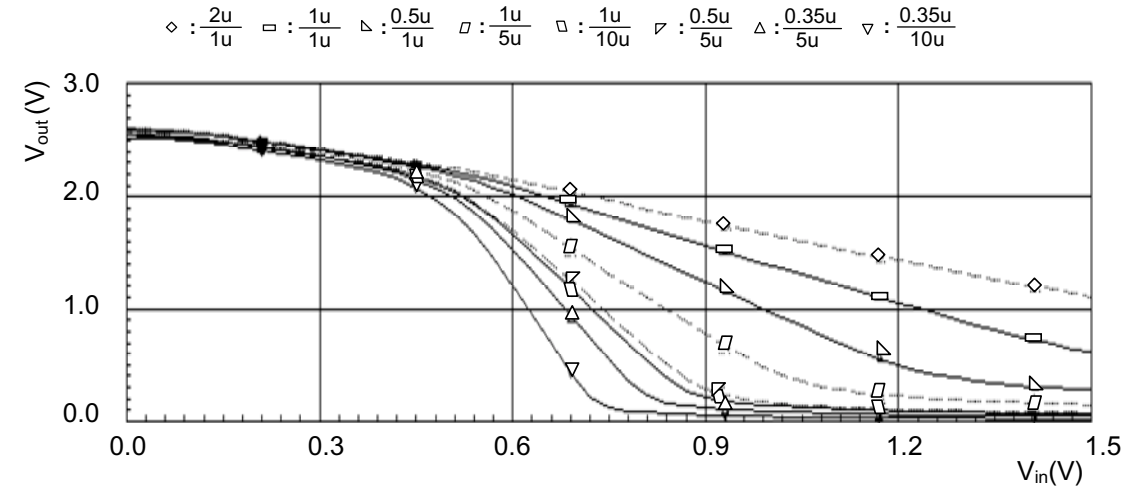
The simplicity of this architecture reduces the offset mismatches of the transistors. One transistor (M_2) is added in the basic architecture. It performs as a switch in order to cut off the power of the CS stage when the pixel is in the integration mode. For this transistor, a rather high W/L (length to width ratio of the transistor channel) is preferable in order to reduce the turn on resistance, so that it has only very slight effect on the performance of the CS stage. The transistors M_1 and M_3 make up the amplifier, where M_1 is the load of M_3. The voltage gain of the amplifier is given by:

$$A_v = -\sqrt{\frac{(W/L)_3}{(W/L)_1}}\frac{1}{1+\eta} \qquad (2.19)$$

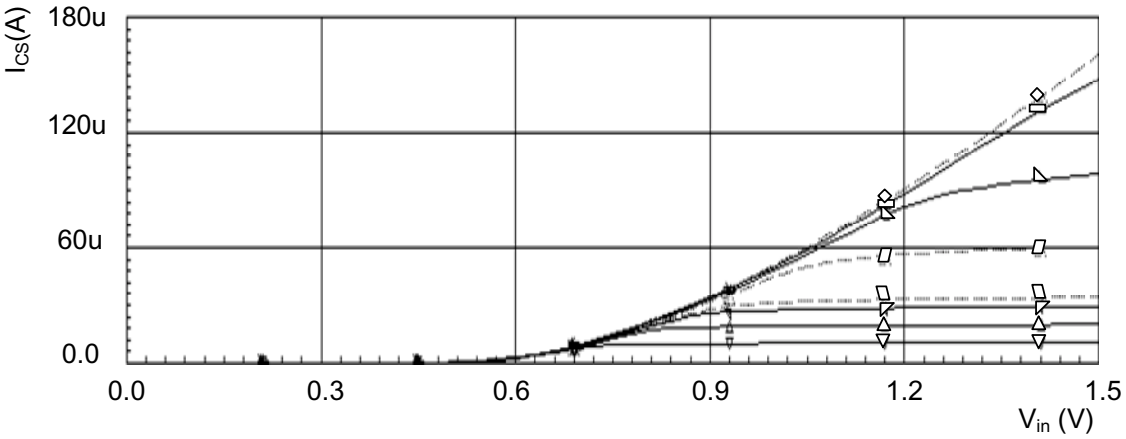
where W and L are the width and the length of the transistor channel, respectively. In the equation 2.19, η is equal to $\frac{g_{mb1}}{g_{m1}}$. Here, g_{m1} is the transconductance of M_1 and g_{mb1} is defined as $\frac{\partial I_D}{\partial V_{BS}}$ of M_1 where I_D is the current passing through the drain of M_1 and V_{BS} is the bulk to source voltage of M_1.

M_3 transfers the input voltage into a current. In order to obtain a high voltage gain and a fast conversion, the minimum L should be used and a large value of W is preferable. However, the charge detection diode is connected with M_3, so increasing the value of W of M_3 means to

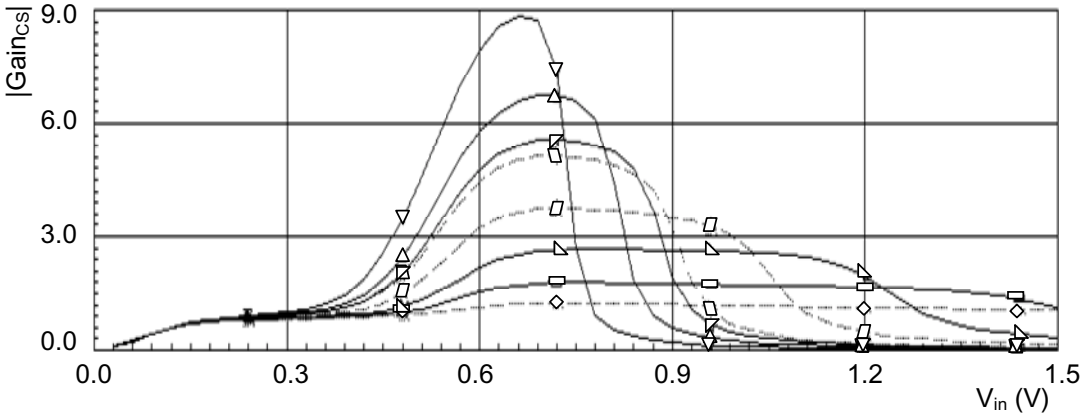
increase also the parasitic capacitance around it, decreasing the corresponding CVF value. For this reason, moderate value of W is chosen to ensure a reasonable current and the voltage gain is mainly obtained by varying W/L of M_1. Several different values of W/L are tested in the simulation. The figure 2.26 gives the simulation results of static characteristics of the CS stage.



(a) Input-output voltage characteristic.



(b) Voltage-current characteristic of M_3.



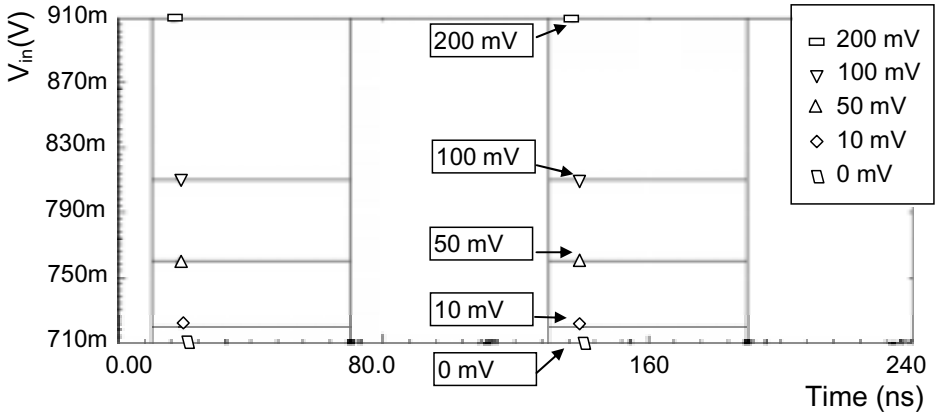
(c) Absolute values of the voltage gain under different conditions.

Figure 2.26: Static characteristics of the CS stage under different W/L values of M_1.
The different values of W/L are as follows:

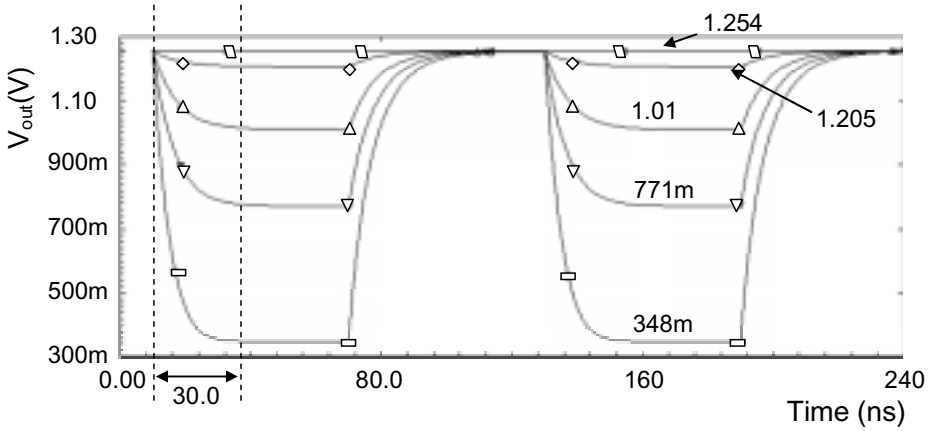
From Fig. 2.26 (c), it can be seen that the absolute value of the voltage gain increases when the W/L ratio of M_1 decreases, as predict by (2.19). The highest value can be achieved with a W/L ratio of 0.35u/10u. But the corresponding dynamic range is rather limited, less than 100 mV. The best compromise between the absolute gain value and the dynamic range can be obtained when W/L is equal to 0.5u/5u or 1u/10u. The absolute gain value is from 5 to 5.6 with a dynamic range of at least 200 mV. The value 1u/10u is then chosen for the better tolerance to mismatch.

From Fig. 2.26 (b), it can be seen that the current passing through the transistor is only about 10 µA if 700 mV is chosen to be the common mode level on the input. As a result, the consumption of the CS stage is about 30 µW when it is powered on.

The transient characteristic of the CS stage is also studied. The responses of the different input signals are as follows (the output charge is set to 100 fF, equivalent to C_{MOS} used in the pixel):



(a)



(b)

Figure 2.27: Transient characteristic of the CS stage: (a) the input signal, (b) the responses of the CS stage for different input signal values.

From Fig. 2.27, the CS stage achieves a dynamic range of 200 mV and the time needed for the output signal to stabilize is about 30 ns. Because the integration time of the whole pixel array is several µs [14], this CS stage is then fast enough for our application.

2.5.2 The SF stage

From the static characteristic of the CS stage Fig. 2.27 (b), its output common level is about 1.25 V. In order to ensure the working condition of the MOS capacitor (strong inversion, $V_{GS} > V_{th}$), the SF stage is realized by N-MOS transistor. The static characteristic of the SF stage is as follows:

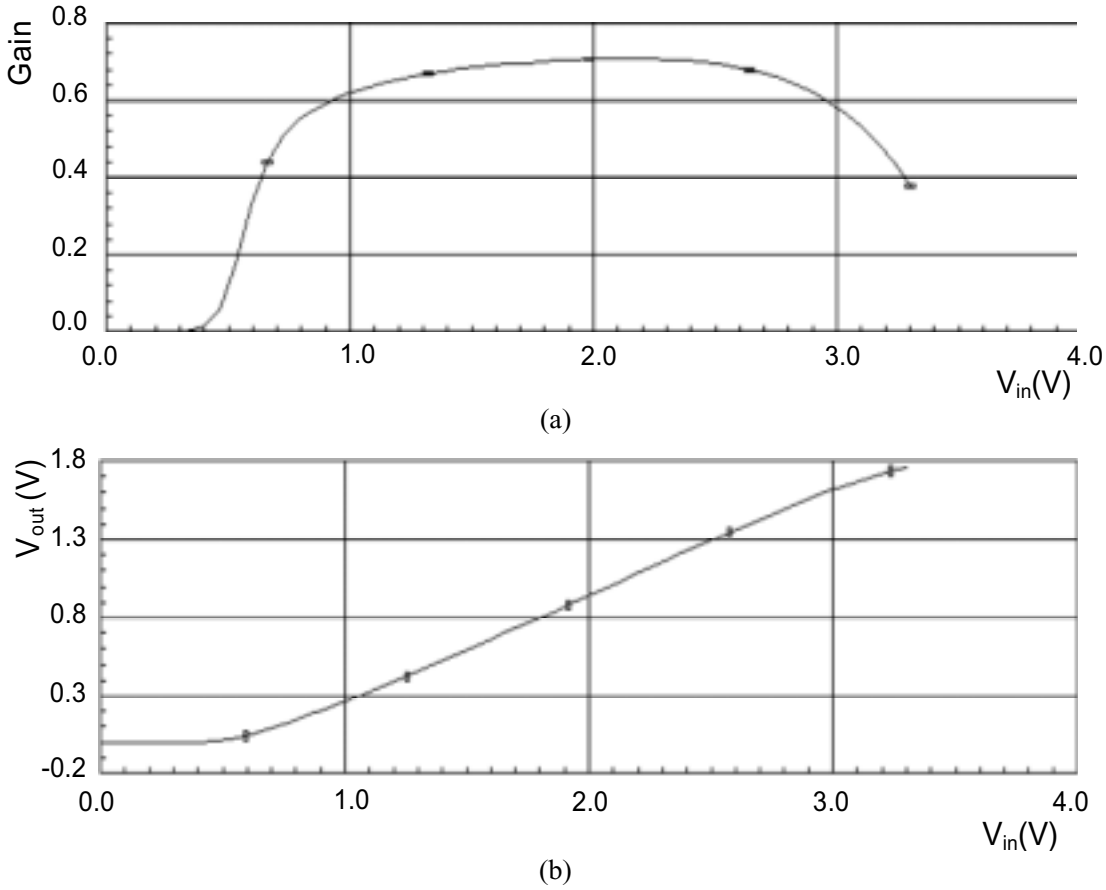


Figure 2.28: Static characteristic of the SF stage: (a) gain, (b) input-output voltage characteristic.

Its input reference level (V_{r2}) is set to 2.1 V by RST2, so that V_{GS} of the MOS capacitor is equal to 0.85 V ensuring the needed working condition.

The bias transistor of the SF stage is also switched by the command signal PWR_ON, so that the discharge on the capacitance of output bus is avoided, limiting the signal variations on the column bus and largely increasing the readout speed. Shown in Fig. 2.29, the reference level remains on the common bus when PWR_ON is turned off. The time needed for the output signal to be stable is less than 15 ns.

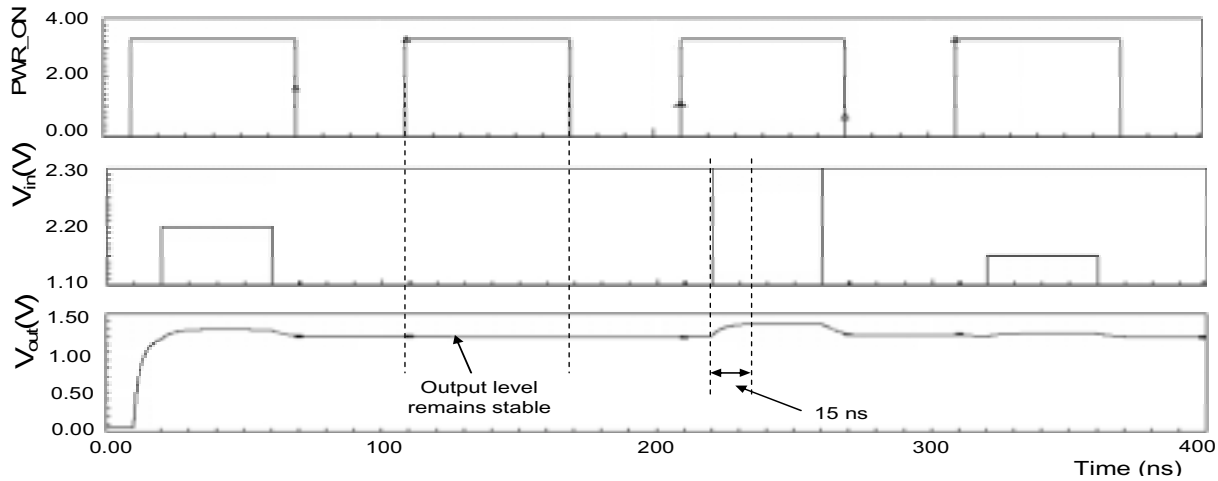


Figure 2.29: Transient characteristic of the SF stage.

2.5.3 The global pixel performance

Taking into account the CS stage and the SF stage together, the pixel's global amplitude and phase responses are as follows:

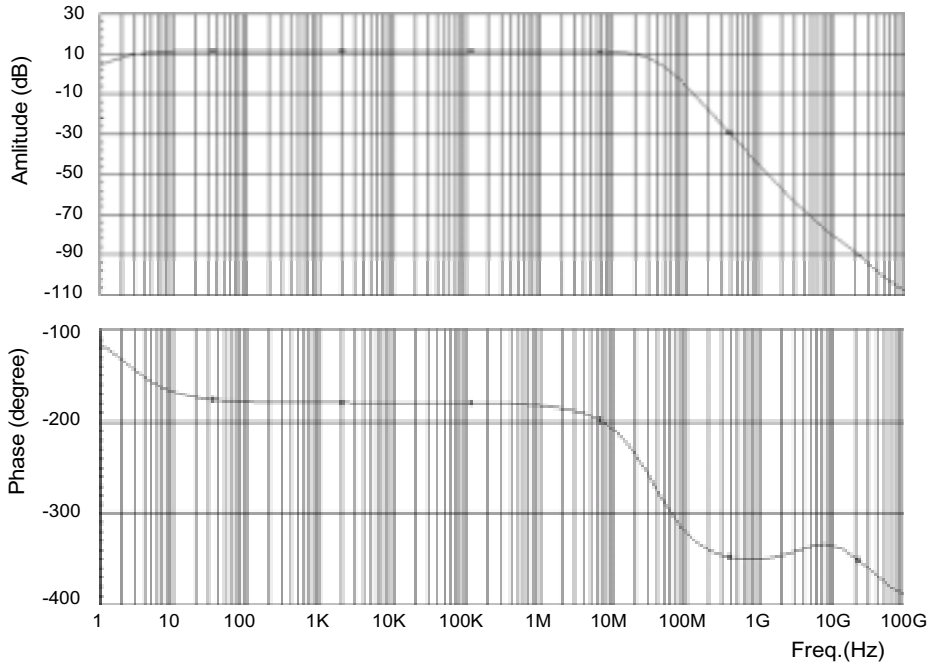


Figure 2.30: Amplitude and phase responses of the pixel.

For the input signals with frequencies from 10 Hz to 10 MHz (the case of MIPs signal detection) the gain of the system is about 13 dB and the phase shift is 180° introduced by the CS stage.

The temporal output noise is estimated by the simulator. The relationship between output noise and the frequency is as follows:

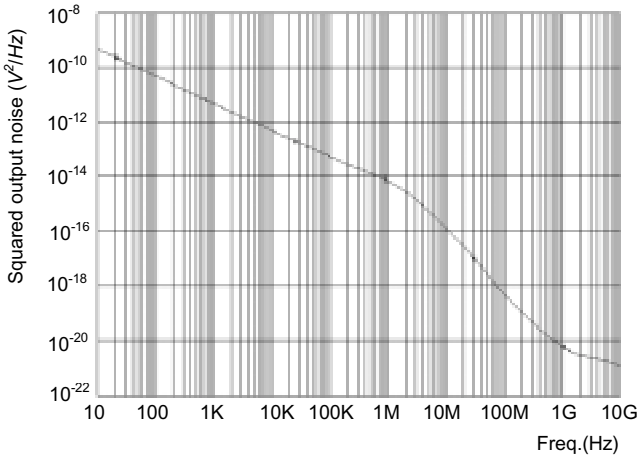


Figure 2.31: Output noise characteristic of the pixel.

It can be seen that the logarithmical characteristic is almost linear below 1 MHz and then it drops rapidly until 1 GHz. As *1/f* noise is not removed according to the pixel's working principle, this simulation result means that the contribution of the 1/f noise to the total output noise level becomes important (compared to the thermal noise) when the observation frequency is below 1 MHz.

The offset cancellation operation has also been simulated. In order to do so, an offset of the range from -50 mV to 50 mV is added on the input of the CS stage. The results are shown below:

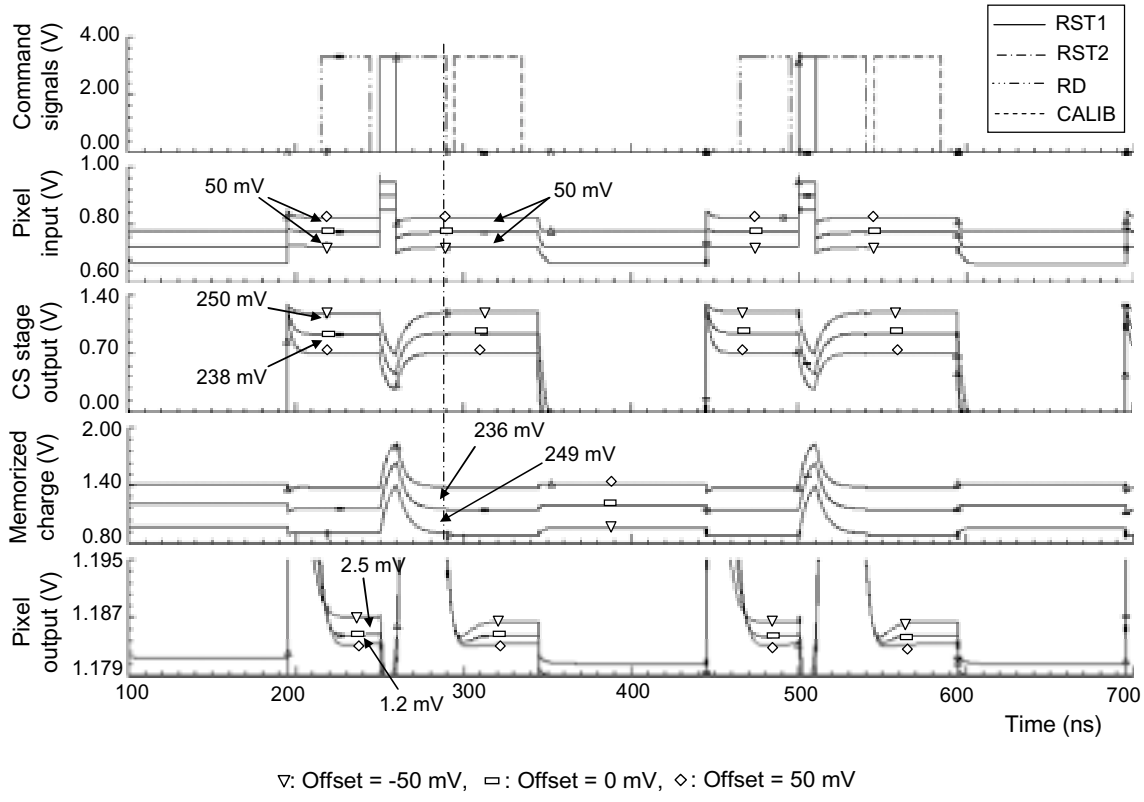


Figure 2.32: Simulation of the offset cancellation operation

It can be seen that the outputs of the CS stage for -50 mV and 50mV offset are 250 mV and 238 mV, respectively. At the end of RST2, these two values are mostly memorized on the series capacitor C_{MOS} (the memorized values are 236 mV for the 50 mV offset input and 249 mV for the -50 mV input). Therefore, the final residual error on the output of the pixel are only 2.5 mV for -50 mV offset input and 1.2 mV for 50 mV offset input. Considering the gain of the system (about 5), these values mean that the input equivalent offset is only about 0.5 mV. Thus, the simulation results prove that this pixel architecture is quite efficient for offset cancellation.

As conclusion, all the key design points are listed:

- The reset transistors should remain in linear region for a fast reset. It requires, for the two reset transistors RST1 and RST2, $V_{GS} - V_{DS} > V_{th}$. Since the reference voltages chosen for RST1 and RST2 are respectively 0.9 V and 2.1 V, this condition can always fulfill when a logic level 3.3 V is applied on their gate as the command (V_{th} is about 0.7V).

- The MOS capacitor should remain in inversion for its linearity. For a transistor used as a capacitor, it requires a relative high gate voltage to ensure the region under the gate is in strong inversion, meaning that V_{GS} must be much greater than V_{th}. As explained above, setting V_{r2} to 2.1 V fulfill this requirement.

- The CS stage should remain in appropriate operation region. In order to benefit the gain, input common mode should choose the value situating in the middle of the dynamic range (700 mV in Fig. 2.26 (c)). However, since the collected charge and the state changing of RST1 both lower the input voltage of the CS stage, choosing a higher input common level of 0.9 V is needed for the design.

- The SF stage should remain in its appropriate operation region. From Fig. 2.28 (a), in order to have a stable gain of the SF stage, the input common level should in the range from 1.0 V to 2.9 V.

2.6 Conclusion

The prototypes of MIMOSA I to MIMOSA V based on the classical 3-transistor pixel architecture demonstrated the feasibility of CMOS sensors for MIPs detection. The encouraging results demonstrate that this detection technique works efficiently and provides excellent tracking performances.

However, faster readout speed and on-chip data processing are required for the final vertex detector. A new pixel structure was studied for this purpose. Two prototypes have been realized: MIMOSA 8 and MIMOSA 16. These chips are optimized to work at a main clock frequency of 100MHz. The corresponding readout speed is 20 μs/frame. These chips contain a 32×128 pixel matrix which is divided into four sub-arrays (S1, S2, S3 and S4).

In the year early 2004, MIMOSA 8 was fabricated in TSMC 0.25 μm digital process with an epitaxial layer of 8 μm. In MIMOSA 8, the sub-array S_1 is made up of AC-coupling pixels and the other sub-arrays consist of DC-coupling pixels. 24 first columns are connected with

column level discriminators. 1-bit binary output is realized. Moreover, in order to reduce the noise level, the circuit realizing CDS operation is integrated inside pixel and column level discriminators, respectively.

In the year 2006, MIMOSA 16 has been fabricated with the same structure as MIMOSA 8 in AMS 0.35 μm OPTO process. Two versions are realized with the different thickness of epitaxial layers: 14 μm and 20 μm. In MIMOSA 16, the sub-arrays S_1, S_2 and S_3 are made up of DC-coupling pixels and the sub-array S_4 contains AC-coupling pixels. The charge detection diode of sub-array S_2 and S_3 has the same size but special layout design rules of radiation tolerance are applied in the design of pixels in sub-array S_3.

Reference

[1] B. Razavi, *Design of Analog CMOS Integrated Circuits*, Boston: McGraw-Hill 2001.

[2] Y. Tsividis, *Operation and Modeling of the MOS Transistor*, Second Ed., Boston: McGraw-Hill 1999.

[3] U. Gatti and F. Maloberti, "Analog and mixed analog-digital layout," in M. Ismail and T. Fiez (eds.), *Analog VLSI: Signal and Information Processing*, McGraw-Hill, New York, 1994, pp. 699-726.

[4] N.K. Verghese, T.J. Schmerbeck, and D.J. Allstot, *Simulation Techniques and Solutions for Mixed-Signal Coupling in Integrated Circuits*, Kluwer, Boston, 1995.

[5] T. Blalack, "Design techniques to reduce substrate noise," in J. H. Huijsing *et al*. (eds.), *Analog Circuit Design: 1-Volt Electronics; Mixed-Mode Systems; Low-Noise and RF Power Amplifiers for Telecommunication*, Kluwer, Boston, 1999, pp. 193-217.

[6] A.J.P. Theuwissen, "Design for manufacturability of CMOS image sensors," IEEE Int. Solid-State Circuits Conference (ISSCC'2000) Tutorial Notes, San Francisco, February 2000.

[7] A.J.P. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices*, Kluwer, Dordrecht, The Nederlands, 1995.

[8] H. Tian, B. Fowler and A. El Gamal, "Analysis of Temporal Noise in CMOS Photodiode Active Pixel Sensor," *IEEE J. Solid State Circuits*, vol.36, NO.1, Jan. 2001, pp. 92-101.

[9] Y. Degerli, "Etude, modélisation des bruits et conception des circuits de lecture dans les capteurs d'images à pixels actifs CMOS", Ph.D. Thesis, Toulouse, 2000.

[10] M.J.M. Pelgrom, H.P. Tuinhout, and M. Vertregt, "Transistor matching in analog CMOS applications", in *1998 IEEE Int. Electron Devices Meeting (IEDM) Tech. Digest*, San Francisco, pp. 915-918.

[11] J. Bastos, M. Steyaert, A. Pergoot, and W. Sansen, "Mismatch characterization of submicron MOS transistors", *Analog Integrated Circuits and Signal Processing*, vol. 12, February 1997, pp. 95-106.

[12] B. Dierikx, G. Meynants, D. Scheffer, "Near 100% Fill Factor CMOS Active Pixels", in *Proc. IEEE CCD&AIS Workshop*, Brugge, Belgium, 1997, p. P1.

[13] G. Deptuch et al., "Design and Testing of Monolithic Active Pixel Sensors for Charged Particle Tracking", *IEEE Trans. Nucl. Sci.*, vol. 49, April 2002, pp 601-610.

[14] G. Deptuch, "New generation of monolithic active pixel sensors for charged particle detection", Ph.D. Thesis, Strasbourg, 2001.

[15]M.H. White, D.R. Lampe, F.C. Blaha and I.A. Mack, "Characterization of Surface Channel CCD Image Arrays at Low Light Levels", *IEEE J. Solid State Circuits*, vol.9, NO.1, 1974, pp.1-13.

[16] J. Hynecek, "Theoretical Analysis and Optimisition of CDS signal Processing Method for CCD Image Sensors", *IEEE Trans. Electron Devices*, vol.39, No.11, 1992, pp.2497-2507.

[17] G. Deptuch et al., "Development of monolithic active pixel sensors for charged particle tracking", *Nucl. Instrum. Meth. A 511*, 2003, pp. 240-249.

[18] M. Deveaux et al., "Neutron radiation hardness of monolithic active pixel sensors for charged particle tracking, " *Nucl. Instrum. Meth. A Vol.512*, Oct. 2003, pp. 71-76.

[19] G. Deptuch, W. Dulinski, Y. Gornushkin, C. Hu-Guo, I. Valin, "Monolithic active pixel sensors with on-pixel amplification and double sampling operation", *Nucl. Instrum. Meth. A 512*, 2003, pp. 299-309.

[20] G. Deptuch et al., "Monolithic Active Pixel Sensors With In-pixel Double Sampling Operation and Column-Level Discrimination", *IEEE Trans. Nucl. Sci.*, vol. 51, No.5, Oct. 2004, pp. 2313-2321.

[21] R. Merrill, "Intra-pixel reset noise cancellation", *Proc.2001 IEEE Workshop on Charge Coupled Devices and Advanced Image Sensors*, Lake Tahoe, Nevada, USA, June 2001, pp.153-156.

[22] I. Takayanagi et al., "A four-transistor capacitive feedback reset active pixel and its reset noise reduction capability", *Proc.2001 IEEE Workshop on Charge Coupled Devices and Advanced Image Sensors*, Lake Tahoe, Nevada, USA, June 2001, pp.118-122.

[23] L. J. Kozlowski and D. L. Standley, "Compact low-noise active pixel sensor with progressive row reset", Patent No. WO9953683, Ocyt. 1999.

[24] B. Fowler et al., "Low noise readout using active reset for CMOS APS", *Proc. of SPIE*, vol.3965, San Jose, CA, Jan. 2000, pp.126-135.

[25] B. Pain et al., "Reset noise suppression in two-dimensional CMOS photodiode pixels through column-based feedback reset", *Proc. 2002 IEEE Int. Electron Device Meeting*, pp. 809-812.

[26] M. Loose et al., " 2/3' CMOS imaging sensor for high definition television", *Proc.2001 IEEE Workshop on Charge Coupled Devices and Advanced Image Sensors*, Lake Tahoe, Nevada, USA, June 2001, pp.44-47.

[27] T. J. Cunningham et al., "A two-dimentional array imager demonstrating active pixel reset suppression of kTC-noise", *Proc. 2003 IEEE Workshop on Charge Coupled Devices and Advanced Image Sensors*, Elmau, Germany, May 2003.

[28] Y. Degerli, G. Deptuch, N. Fourches, A. Himmi, **Y. Li**, P. Lutz, F. Orsini, and M. Szelezniak, "A Fast Monolithic Active Pixel Sensor with Pixel-Level Reset Noise Suppression and Binary Outputs for Charged Particle Detection," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 6, 2005, pp. 3186-3193.

[29] Y. Degerli, “Monolithic active pixels with in-pixel amplification and reset noise suppression for charged particle tracking in a 0.25μm digital CMOS process”, preprint n° DAPNIA-04-184, CEA-Saclay/DAPNIA.

[30] G. Deptuch, "Charge sensitive elements optimised for particle tracking," in *Proc. 2003 IEEE Workshop on Charge Coupled Devices and Advanced Image Sensors*, Schloss Elmau, Elmau, Germany, 2003.

[31] B. Razavi and B. A. Wooley, “Design techniques for high-speed, high-resolution comparators”, *IEEE J. Solid-State Circuits*, vol. 27, Dec. 1992, pp. 1916–1926.

[32] B. Razavi, *Principles of Data Conversion System Design*. Piscataway, NJ: IEEE Press, 1995.

[33] F.B. McLean, H.E. Boesch Jr. and T.R. Oldham, "Electron-hole generation, transport and trapping in SiO_2", in *Ionizing radiation effect in MOS devices and circuits,* edited by T.P. Ma and P.V. Dressendorfer, new-York : Wiley Interscience, 1990, p. 87.

[34] O. Flament et al., “Charge Pumping Analysis of Radiation Effects in Locos Parasitic Transistors,” *IEEE Trans. Nucl. Sci.*, vol. 44, no. 6, Dec. 1997, pp. 1930-1938.

[35] G.Anelli et al, “ Radiation Tolerant VLSI Circuits in Standard Deep Submicron CMOS Technologies for the LHC Experiments: Practical Design Aspects,” *IEEE Trans. Nucl. Sci.*,vol.46, Dec. 1999, pp. 1690-1696.

[36] Y. Degerli, N. Fourches, M. Rouger, and P. Lutz, "Low-power autozeroed high-speed comparator for the readout chain of a CMOS monolithic active pixel sensor based vertex detector," *IEEE Trans. Nucl. Sci.*, vol. 50, no. 5, October 2003, pp. 1709-1717.

Chapter III

Performance of MIMOSA 8 and MIMOSA 16: ^{55}Fe X-ray photons calibration tests

3.1 Introduction

Before test with high energy charged particles, MIMOSA 8 and MIMOSA 16 prototypes are calibrated carefully with a ^{55}Fe X-ray source (energy peaks: 5.9 keV and 6.4 keV) in order to evaluate their performances. The important parameters, such as the Charge-to-Voltage conversion Factor (CVF) and the Charge Collection Efficiency (CCE), are obtained in laboratory tests [1] [2]. We can also deduce the Signal-to-Noise Ratio for X detection as an indication for the following tests.

Since the signal level is very small in our applications of Minimum Ionizing Particles (MIPs) detection, the noise level determines the chip's performance. For this reason, Temporal Noise and Fixed Pattern Noise (FPN) evaluated for a whole array of pixels, pixel-to-pixel variation of the conversion gain and properties of the lateral charge spreading onto the neighbouring pixels, are carefully studied.

As the design of MIMOSA 8 and MIMOSA 16 prototypes aim at realizing high readout speed chip with on-pixel data processing (in-pixel double sampling, offset auto-compensation on both pixel and column level), all the parameters are measured as the principle working frequency varies from 1MHz to 150 MHz, corresponding to readout speed from 200 µs/frame to 13 µs/frame respectively.

Both analog output and digital output are acquired. Although the analysis of the analog output give us the CVF and the CCE value, the digital output only show us the sensitivity of the pixels under different threshold value. However, by setting an equivalent threshold value during the analysis of the analog output, the performances of these two types of outputs are compared.

The influence of the thickness of epitaxial layer on detection performance is also studied. The epitaxial layer is about 8 µm for MIMOSA 8, fabricated in TSMC 0.25 µm digital process. MIMOSA 16, using AMS 0.35 µm OPTO process, has two versions with different thicknesses of the epitaxial layer: 14 µm (MIMOSA 16_14) and 20 µm (MIMOSA 16_20).

The spatial resolution and the detection efficiency, obtained using high energy charged particle beams at DESY and CERN, will be discussed in the next chapter.

3.2 Experimental set-up and procedure

In order to record the output data of the pixel array, the test chip is linked to DAQ by two dedicated boards designed to control the chip and to provide signal buffering for transmission. Then, data acquired are analyzed using dedicated off-line software. Firstly, the chip is studied without source, to determine the pedestal of each pixel and their distribution on the array (FPN) and to measure the Temporal Noise level. Secondly, the chip is studied with the source. The CVF and the CCE value of every sub-array are therefore characterized.

3.2.1 Set-up of the experiment

Two specific data acquisition systems, both designed by IPHC, are used to test all the chips. The test system is also under development. For this reason, MIMOSA 8 chips are tested with a data acquisition system based on the VME bus protocol. MIMOSA 16 chips are tested with a system based on based on USB 2.0 bus. The architectures of these two systems are similar to each other, consisting of three cards: the Front-End board, the Interface board and the data acquisition board. The simplified schematic diagram of the experimental set-up for MAPS detectors characterization is shown in Fig. 3.1.

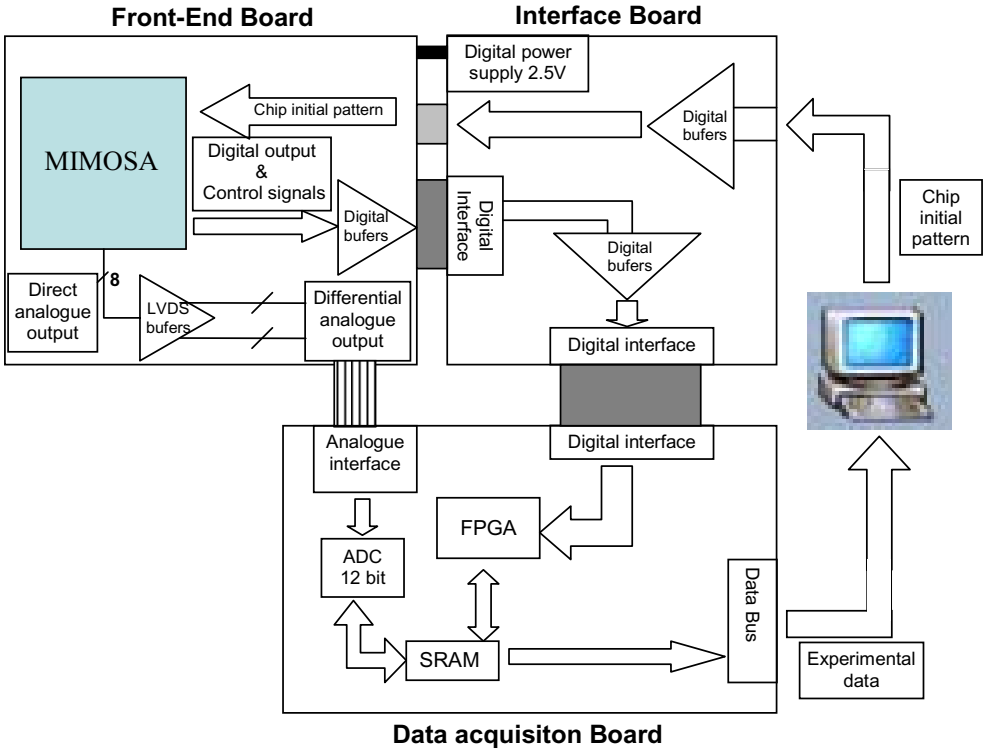


Figure 3.1: Set-up for laboratory test.

A. Front-End board

The Front-End board is a small PCB card, made by CEA Saclay, where the prototype chip is wire-bonded. It comprises the first stage external amplifiers buffering the signals from the chip and tunable sources of bias voltages needed for chip operation. The set of low noise amplifiers is mounted on the card. They convert single-ended analog pixel signals, which directly come from the chip, to LVDS (Low Voltage Differential Signal) signals, more convenient for transmission. In order to reduce the coupling between analog and digital signals on the mixed circuit environment, the power supplies of the analog part and the digital part are separated. The reference voltage (3.3 V), used for on-chip analog parts and for external amplifiers, is generated on this card.

B. Interface Board

The Front-End Board is connected to the Interface Board (made by IPHC-Strasbourg) which provides an interface stage between the chip under test and the VME processor. On one hand, the Interface Board transmits the initial timing program pattern from the computer to the chip. On the other hand, the differential analog data and some digital signals, being necessary for data acquisition, generated by MIMOSA 8 chip are also transmitted to VME card thanks to the Interface Board. In addition, in order to separate analog and digital power supply for avoiding digital-to-analog coupling, the Interface Board is also used for generating and distributing digital power supply voltages needed for on-chip digital parts and external digital components. The generation of power supply is +2.5 V for digital circuits.

C. VME Data acquisition board

The system, elaborated by IPHC-Strasbourg, is composed of:

- two fast analog-to-digital conversion VME boards, installed inside a VME card system under LynxOS;
- The dedicated PC computer under Linux, running the software for data acquisition, is used to control the VME board and to store the acquired data on its disk.

The VME data acquisition board is driven by ELTEC E-16 processor running OS/9. The VME Flash ADC Unit for the Strip Detector Readout (VFAS) modules [3] with 8 and 12 bit versions of precision are installed in this board.

The 12 bit resolution VFAS module features two independent ADC channels with maximum conversion rate of 40Msamp/s, and it was possible to process simultaneously data from arrays of pixels. The full digital control is handled by a XILINX FPGA (Field Programmable Gate Array) chip based, programmable logic unit installed on the VFAS card. It includes frame and line synchronization, timing verification and generation of test patterns. The digital test

patterns, provided by the VFAS module are limited to two channels only, needed to drive the MIMOSA chip i.e. the read-out clock and the reset signals. These signals are transmitted via the digital I/O port.

The individual pixels in the array of the chip are addressed in consecutive clock cycles and the samples, digitized with 12-bit resolution, are stored in the local on-card SRAM (Static Random Access Memory) memory.

D. USB Data acquisition board

The architecture of this new data acquisition board is based on USB 2.0 bus for data transfer and a Virtex II FPGA as board controller. Without the need of dedicated RIO2 CPU board and LynxOs real-time system, the new Imager board can be driven by any simple Windows PCs equipped with a USB 2.0 port. Thus, the system integration is significantly improved. The new Imager board is standard 6U VME size, but they only require power supply from the VME board. The function of USB board is similar to VME version: chip is driven by a sequencer and serial output analog data of pixels is acquired by four differential analog inputs. The data transfer rate has been improved. The full digital control is handled by a XILINX Virtex II FPGA unit installed on the USB board. The main features of USB data acquisition board are summarized as follows [4]:

- 4 analog inputs (ADC) 12-bit 40 MHz
- 2 digital trigger inputs (NIM / TTL)
- 2 digital synchronisation outputs (NIM / TTL)
- 8 LVDS lines for pattern generator (4 in, 4 out)

3.2.2 Experimental procedure

The analog information from the pixel device under test is transmitted to the data acquisition board and is digitized by a 12-bit precision ADC. The sampling clock used for ADCs is made up of two internal digital signals of the chip: the RD signal used for sampling signal level and the CALIB signal used for sampling reference level. Consequently, a pixel is consecutively sampled twice to extract the real signal level of the incident charged particle.

$$V_{signal,real} = V_{RD} - V_{CALIB} \qquad (3.1)$$

In order to avoid the fluctuation of signals, it is important to choose a correct delay so that the signal for sampling is already stable. For this reason, additional delay, comparing to the output analog signals of pixels, can be added to this ADC clock thanks to the Interface Board.

During all the laboratory tests, the pixel sensor is kept at constant temperature (about 20°C). These test conditions are reached using a water circulation in a dedicated cooling system. The major motivation of cooling was to achieve more reliable test conditions. Generally, cooling is used to reduce the leakage current so that the reset cycle could be slowed down to a few Hz. However, during the tests, the device cooling has a negligible effect on the measured noise in that the readout speed is high enough so that the contribution of the shot noise due to the leakage current is negligible compared to other noise sources.

3.3 Test of analog output

The analog output is tested firstly. In the first step, noise performance is studied without the source. The Temporal Noise and the FPN are measured for the main clock frequency varying from 1 MHz to 150 MHz. In the second step, using the source, the important parameters determining the chip performance (CVF, CCE) are also studied as a function of the main clock frequency.

3.3.1 Noise performance

Due to small value of the incident signal, a very low noise level is required for the detectors applied for MIPs detection. As mentioned in chapter 2, the main sources of noise in a CMOS sensor are: shot noise, pixel reset noise, readout noise and FPN. The former three sources of noise occur randomly and can be considered as Temporal Noise. The last one, the FPN, is the dispersion of the average pixel values (without source) for a whole array. The difference of the average values among the pixels results from the mismatches between the transistors used inside each pixel. For every single pixel, this average value remains constant and is called pedestal value. Thus, the FPN value of a pixel array gives the dispersion level of the pedestal values, so that on-chip sparsification operation needs very small FPN value of the pixel array. The main objective of noise measurement is to evaluate the values of the Temporal Noise and the FPN. The measurements are carried out under different conditions and very encouraging results are obtained.

A. Noise calculation algorithm

A dedicated data analysis program has been developed using C++ Builder. The algorithm is described as follows:

According to (3.1), the real signal value is obtained by subtraction of the two consecutive samples of one pixel. This subtraction corresponds to the CDS processing, which largely reduces the disturbance of noise over the signal. The total value of signal of one pixel after CDS subtraction can be given by:

$$S_{pixel} = S_{sig} + \sigma_{tem} + P_{ped} \qquad (3.2)$$

where S_{sig} is the physical charged signal value due to the interaction between the particles and the silicon, σ_{tem} is the residual random Temporal Noise value and P_{ped} is the residual pedestal value. All of these quantities are measured in ADC units, which is equal to analog to digital conversion unit value defined by the data acquisition system.

When there is no any radioactive source, the simplest pedestal estimator can be calculated by averaging measured signals over N events*:

$$S_N^{Ped}(k) = \frac{1}{N}\sum_{n=1}^{N}\left[\sigma_{tem,n}(k) + P_{ped,n}(k)\right] \qquad (3.3)$$

where k represents pixel k in the array of pixel and n represents the n^{th} acquired event in the total N events.

In practice, obviously, it is difficult to isolate the physical signal from the pedestal. A possible estimator for pedestal is given by the mean value of output signal when there is no source:

$$P_{ped}^{est}(k) = \frac{1}{N}\sum_{n=1}^{N} S_{pixel,n}(k) \qquad (3.4)$$

For a total of N events, if we suppose that there are N_p events with no physical signal charge, we have $N_s = N - N_p$ events with physical signal charge. The equation 3.4 can be rewritten as:

$$\begin{aligned} P_{ped}^{est}(k) &= \frac{1}{N}\sum_{n,pedestal}^{N} S_{pixel,n}(k) + \frac{1}{N}\sum_{n,signal}^{N} S_{pixel,n}(k) \\ &= P_{ped,N}(k) + \frac{1}{N}\sum_{n,signal}^{Ns} S_{pixel,n}(k) \end{aligned} \qquad (3.5)$$

After the arrangement equation (3.5), the ratio between this estimator and the real value is:

$$\frac{P_{ped,N}^{est}(k) - P_{ped,N}(k)}{\langle Q_s \rangle} = 1 - \frac{N_p}{N} \qquad (3.6)$$

* In this work, an event means a frame of raw data taken from the pixel matrix.

where $\langle Q_s \rangle = \frac{1}{N_s} \sum_{n,signal}^{N} S_{pixel,n}(k)$, N_p stands for the events without physical signal and N_s stands for the events with physical signal.

By (3.6), it is clear that the estimator, $P_{ped}^{est}(k)$, is equal to the real value when N_p equals to N. For this reason, the noise measurement is always carried out without the radioactive source (equivalent to no input signal) and $P_{ped}^{est}(k)$ is calculated and considered as the real value of pedestal. Obviously, because the pedestal value gives the signal level of a pixel without any incident charged particle, it represents the influence of dark current, injected charge, *1/f* noise and shot noise. Besides, the reset noise and the offset due to mismatch of transistors also influence the pedestal value. Fortunately, since the pedestal value is stable for each pixel, it can be eliminated by applying an external threshold voltage. Thus, the pedestal value is also called offset.

The FPN value, representing the importance of pedestal variation level over all pixels, is the standard variance of pedestal over a whole pixel array. This value is crucial for on-chip sparsification operation. Generally, all the pixels share the same signal processing circuit integrated with the pixel array, which means that an single external threshold valued is settled down to compensate the average offset of pixel array. Thus, an important variation of offset makes this threshold value useless so that signals issue from a large part of the pixels can not be correctly processed. For an array of M pixels, the FPN value is given by the equation below:

$$FPN = \frac{1}{\sqrt{M-1}} \sqrt{\left[\sum_{k=1}^{M} P_k^2 - M\left(\frac{1}{M}\sum_{k=1}^{M} P_k\right)^2 \right]} \qquad (3.7)$$

where P_k is the pedestal value of the k^{th} pixel in the array.

The Temporal Noise value for each pixel is determined for each pixel with the expression below:

$$\sigma_{tem,N}^{est}(k) = \frac{1}{\sqrt{N-1}} \sqrt{\left[\left(\sum_{n=1}^{N} S_{pixel,n}(k)^2 \right) - N \cdot P_{ped,N}^{est}(k)^2 \right]} \qquad (3.8)$$

The equation (3.8) gives the Temporal Noise value of the k^{th} pixel in the array, over total N events. The Temporal Noise of the array is the average value over all the pixels inside it. For an array of M pixels, the Temporal Noise of this array is:

$$\sigma_{tem,N}^{est} = \frac{1}{M} \sum_{k=1}^{M} \sigma_{tem,N}^{est}(k) \qquad (3.9)$$

The array's Temporal Noise represents the influence of random variation over the output of a pixel and determines the minimum input signal that the detector is able to identify.

In the following measurements, Temporal Noise means only the average value for the pixels of sub-array, which is calculated by equation (3.9).

B. Noise performance

At the first step, in order to estimate Temporal Noise and FPN, the analog outputs of MIMOSA 8 are tested with a digital oscilloscope as the basic option aiming validation of the new clamped CDS circuitry. During the tests with the oscilloscope, the clock operating frequency is of the order of 10 MHz, allowing clearly the observation of different phases (the first readout, reset of the charge collecting diode, reset of the clamping capacitance, the second readout—calibration in the pixel access). The direct analog output signal observed for the DC-coupling option of the pixel design is shown in Fig. 3.2.

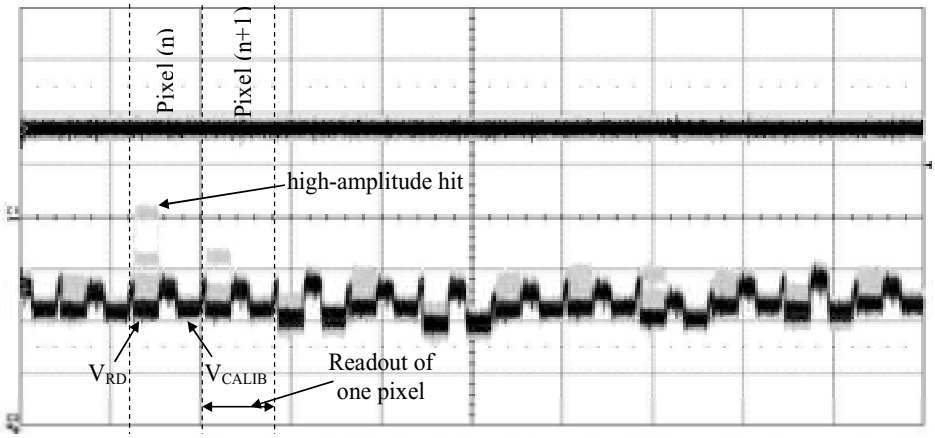


Figure 3.2: Output analog signal from 12 pixels of a column recorded on a scope (50 mV/div, f = 10 MHz).

For each pixel, it can be seen at the oscilloscope that the presence of 2 levels for each pixel (V_{RD} and V_{CALIB}) is visible on the recorded waveform. The output signal is accumulated by the oscilloscope. The black part is the reference level and the grey parts correspond to the signal, appearing when a ^{55}Fe source is put in front of the detector. The useful signal for each pixel is the difference between these two levels, normally extracted by the column readout circuitry. A high-amplitude hit* is clearly detected on pixel (n) during the readout phase, corresponding to full energy deposition of the X photon. Other hits appear also distinctly. In Fig. 3.2, the time needed for the readout of one pixel is 1.6 µs at a main clock frequency of 10 MHz.

Tests without the source showed that the double sampling could eliminate offset dispersions of the in-pixel amplifying stage. The offset dispersions of the source follower output stage are corrected later by the column readout circuitry (discriminators). The two samples are

* An incident photon, whose charge is deposit inside the epitaxial layer, is captured by the detector.

recorded for each pixel and then subtracted offline, resulting in dispersions level less than 1 $mV_{Peak-Peak}$. As only one row of pixels is selected during the readout phase, the power consumption is given by:

$$P = 70\ \mu A \times 3.3\ V \times N_{col} \qquad (3.10)$$

where N_{col} is the number of column. For example, for MIMOSA 8, the power consumption of the analog pixel is about 7.4 mW.

In order to measure more precisely the pixel parameters, the VME-based and the USB-based data acquisition system are used. The experimental data are analyzed by a dedicate off-line software. Formerly, MIMOSA 8 is tested with the VME data acquisition system. Since 2006, the new USB acquisition system has been set up and the two versions of MIMOSA 16 are then tested by using this new system. With these acquisition systems, the noise characteristics are measured as the function of the main clock frequencies.

a) Noise performance of MIMOSA 8

As the first prototype, the noise performance of MIMOSA 8 is studied thoroughly for a frequency varying from 1 MHz to 150 MHz.

As an example, Fig. 3.3 shows the distribution of pixel-to-pixel pedestal dispersions for full readout sequence with on-pixel CDS and subtraction of the calibration level at three clock frequencies (1, 10 and 25 MHz) for the three DC-coupling sub-arrays: S_2, S_3 and S_4. Gaussian fits on the distributions are used to give mean values of 2.47 ADC units for 1 MHz, 0.70 ADC units for 10 MHz and 0.66 ADC units for 25 MHz, respectively. The rms value of pedestal dispersions (FPN) is very small: 1.05, 0.71, and 0.90 ADC units for 1, 10, and 25 MHz. In the acquisition system, 1 ADC unit is corresponding to 0.5 mV. The interpretation of the results will be given later.

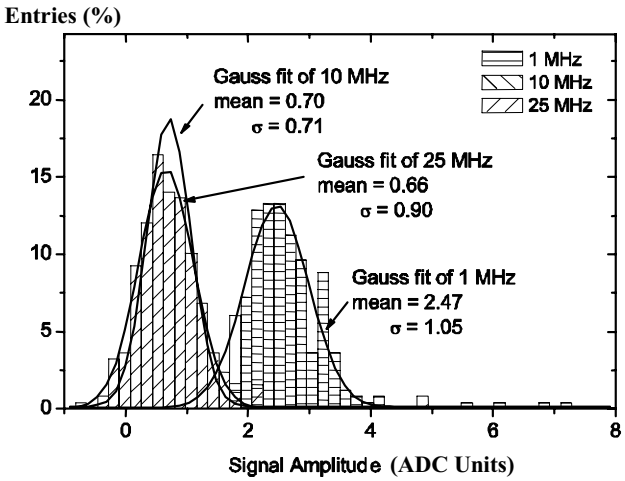


Figure 3.3: FPN for full readout sequence at different readout frequencies.

The distributions of the Temporal Noise at these three frequencies are shown in Figure 3.4, for the three DC-coupling sub-arrays.

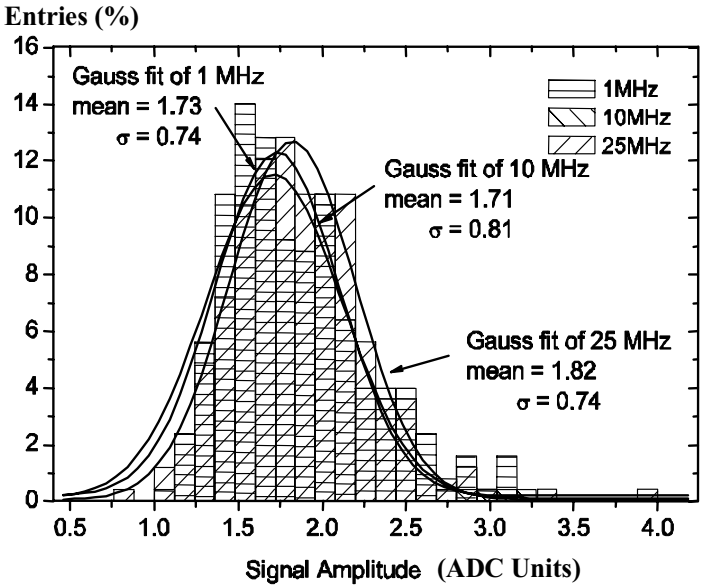


Figure 3.4: Temporal Noise measured for full readout sequence applying on-pixel CDS operation and subtraction of the calibration level for different readout frequencies.

The mean value of these distribution is lower than 2 ADC units, corresponding to 1 mV (Temporal Noise = 1.73, 1.71, and 1.82 ADC units respectively for 1, 10, and 25 MHz). The estimated errors for all these measurements are below 3%. We notice from the results that the Temporal Noise is not dependent on the clock frequency. Thus, the dominant noise source is identified to be the Thermal Noise of the CS stage sampled on the clamping capacitor. The results presented in Figs. 3.3 and 3.4 were obtained putting together distributions corresponding to all sub-arrays with DC-coupling pixels. Independently of the measurements with the VME data acquisition system, the Temporal Noise was measured using the oscilloscope, increasing the readout clock frequency up to 150 MHz. The noise measured was still lower than 1 mV_{rms}.

The above measurements are repeated for different frequencies from 1.25 MHz to 150 MHz and the results are shown in Fig. 3.5 and 3.6. Fig. 3.5 represents the Temporal Noise and Fig 3.6 represents the corresponding FPN, as the function of the main clock frequency.

From Fig. 3.5, the Temporal Noise is below 0.85 mV_{rms} for the frequencies from 5 MHz to 100 MHz (designed frequency of the chip). For the FPN, its value is around 0.25 mV_{rms} below 100 MHz.

For the frequencies below 5 MHz, the Temporal Noise increases slightly. The main reason is that the data processing (double sampling) applied inside pixel is not truly correlated. The value of pixel is not read out right after the reset until the arrival of RD signal for next lecture

cycle. Thus, the 1/f noise can not be cancelled by this operation and exhibits its contribution to the total noise value especially in lower frequency.

For the frequencies below 125 MHz, the Temporal Noise remains more or less stable while the fixed pattern noise increases sharply when the frequency reaches 100 MHz. The reason is that the period for charging the offset storage capacitors is not enough at such frequency and that the pixel signal would be read out before it is stable. In the measurements inside this frequency range, the maximum value of pedestal variation is about 0.75 mV_{rms} at 125 MHz.

For the frequencies of 140 and 150 MHz, the corresponding Temporal Noise is more important than other frequencies. The main reason is that it is very hard to correctly sample an analog signal level whose duration is only tens of ns. The bad position of sampling (influenced by the reset period) disturbs the final results.

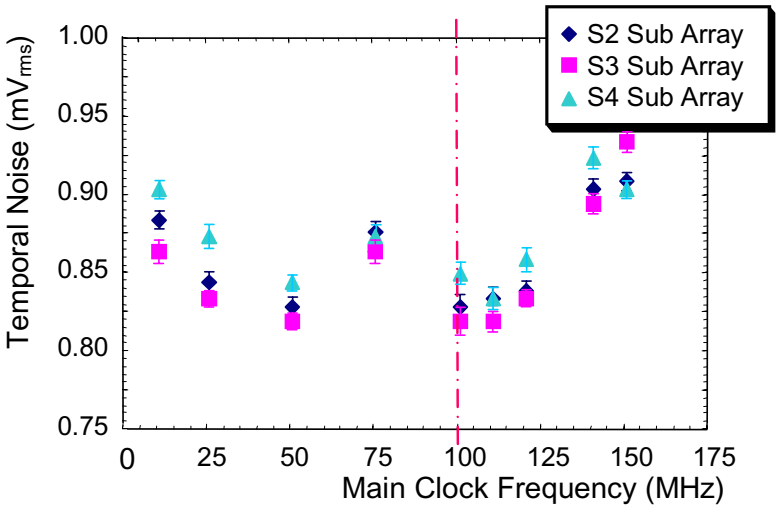


Figure 3.5: Temporal Noise as the function of main clock frequencies.

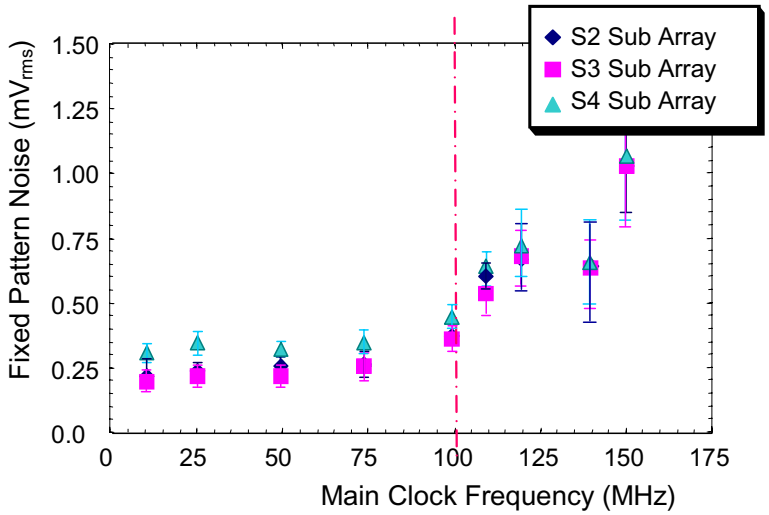


Figure 3.6: FPN as the function of main clock frequencies.

b) Noise performance of MIMOSA 16

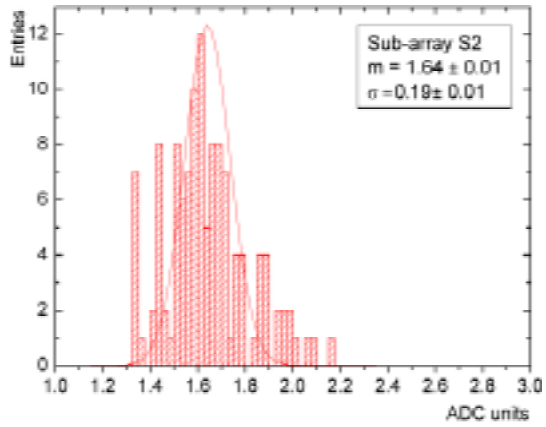
Since the year 2006, a new USB-based data acquisition system has been developed. Compared to the VME DAQ card, which needs a dedicated control system, the control of the USB DAQ card is much simpler: it can be driven by any Windows PCs equipped with a USB 2.0 port. Additionally, the Front-End board was lightly modified. One ground layer was added in order to reduce the coupling between analog and digital signal. The same Interface Board was installed in the system. MIMOSA 16 is tested with this new data acquisition system. The measurements were carefully carried out at different frequencies. During all the tests, a cooling system is used in order to keep the environment at a constant temperature (about 20°C).

Because the results of MIMOSA 8 show that the noise performance remains stable as function of the main clock frequencies, the main objective of the noise measurements of MIMOSA 16 is to verify these results. Thus, the measurements are mainly carried out at three different frequencies (1, 50 and 100 MHz) in order to avoid numerous repeats of test. The three DC-coupling sub-arrays are studied separately. Both of the two versions of MIMOSA 16 (14 µm and 20 µm thickness of the epitaxial layer) are measured. As an example, the distributions of the pedestal are shown in Fig. 3.7.

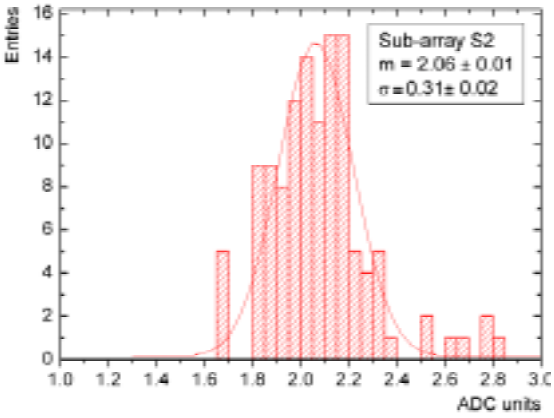
Comparing the results at these three frequencies, the mean value of pedestal is raised slightly at 1MHz. The reason is that the corresponding integration time is about 2 ms at this main frequency and an important number of free electrons, which are created by thermal generation, accumulate inside the p-n junction of the charge detection diode during this period. The accumulation of these free electrons gives rise to dark current. Since the two samples taken from pixel were not truly correlated, the growth of dark current could not be fully suppressed and therefore raises the mean value of pedestal. Fortunately, this phenomenon disappears at higher frequencies. The mean values of pedestal at 50MHz and 100MHz are very small. The corresponding mean value is less than 0.5 ADC units.

The dispersion of pedestal represents the sub-array's FPN value which is crucial for on-chip auto data processing. The FPN values remains very small at all the three frequencies. The value is only about 0.3 ADC units (equivalent to 0.15 mV_{rms}), which indicates that the mismatch between pixels is extremely well controlled.

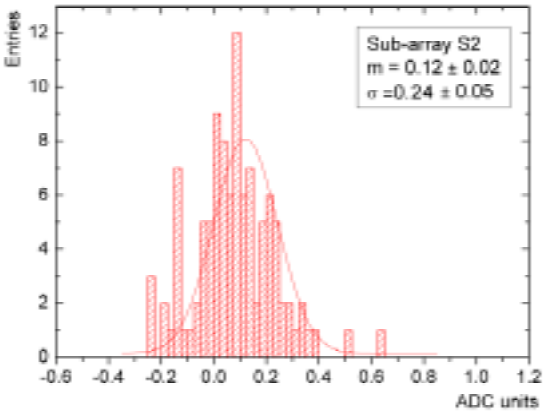
As presented previously, several hundred millions of pixels are required for the vertex detector of the future ILC. Because the on-chip auto sparsification is needed, very low pixel-to-pixel dispersion is mandatory for the success of the design. Thus, the excellent results obtained here, especially for high frequency at 100 MHz, give us the possibility to realize large scale pixel matrix and more complicated on-chip signal treatment circuits.



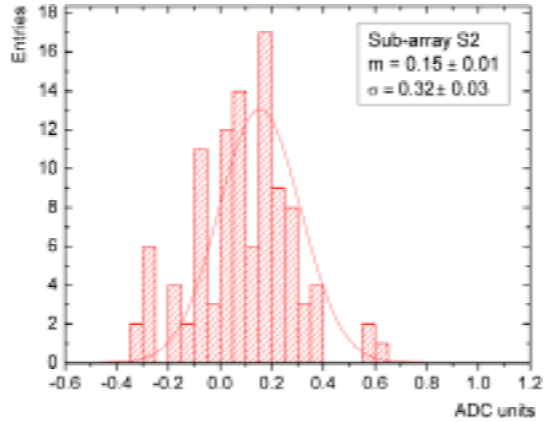
(a) 1 MHz : 14 µm epitaxial layer version



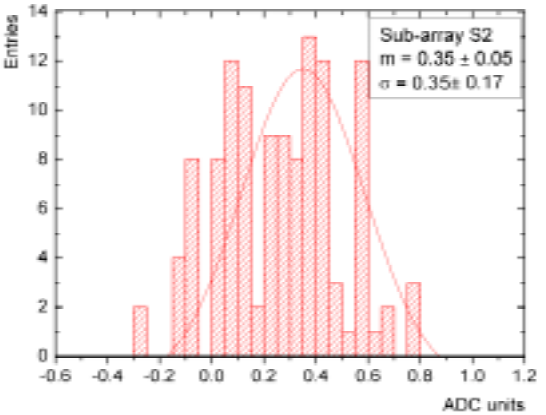
(b) 1 MHz: 20 µm epitaxial layer version



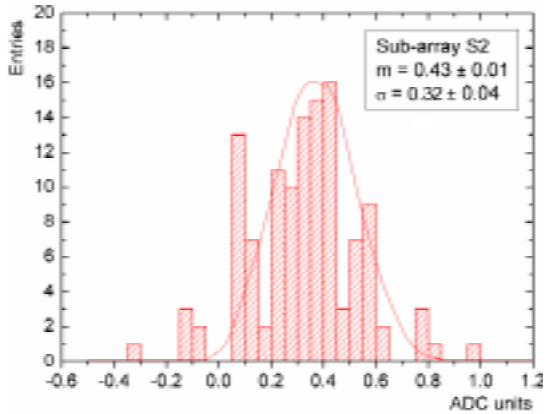
(c) 50 MHz: 14 µm epitaxial layer version



(d) 50 MHz: 20 µm epitaxial layer version



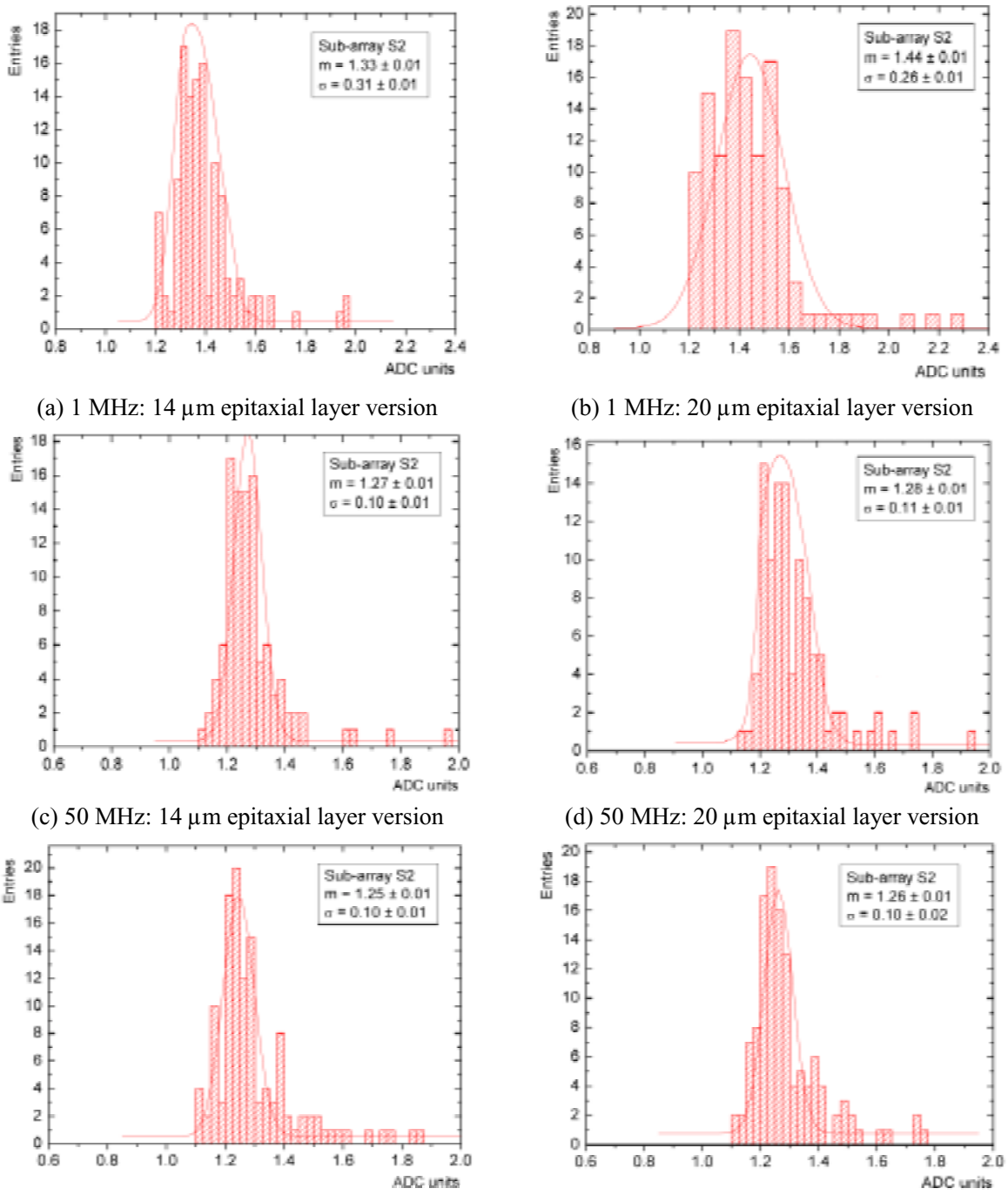
(e) 100 MHz: 14 µm epitaxial layer version



(f) 100 MHz: 20 µm epitaxial layer version

Figure 3.7: Distributions of pedestal values of sub-array S2 for both 14 µm (a) (c) (e) and 20 µm (b) (d) (f) epitaxial layer versions at respectively 1, 50 and 100 MHz.

The distributions of the Temporal Noise for all the three sub-arrays are also obtained. The noise level is very low and the results are very stable as the function of the frequency. The following example gives its distributions for sub-array S2 at 1MHz, 50MHz and 100MHz, respectively (Fig. 3.8).



(a) 1 MHz: 14 µm epitaxial layer version

(b) 1 MHz: 20 µm epitaxial layer version

(c) 50 MHz: 14 µm epitaxial layer version

(d) 50 MHz: 20 µm epitaxial layer version

(e) 100 MHz: 14 µm epitaxial layer version

(f) 100 MHz: 20 µm epitaxial layer version

Figure 3.8: Distributions of Temporal noise of sub-array S2 for both 14 µm (a) (c) (e) and 20 µm (b) (d) (f) epitaxial layer versions at 1, 50 and 100 MHz.

For the main clock frequencies of 50 and 100 MHz, it can be seen that the Temporal Noise is very small with the peak value below 1.3 ADC units, which is equal to only 0.65 mV. No difference is observed between the two versions of the chip. Moreover, the Temporal Noise is quite stable for 50 and 100 MHz. However, the Temporal Noise increases slightly at 1 MHz because of the increasing of 1/f noise. The explanation is the same as MIMOSA 8: the pixel operation is not truly correlated so that the *1/f* noise is not cancelled at all, which contributes a great part to the total noise level at lower frequencies.

The FPN and Temporal Noise values (in mV_{rms}) for all the three sub-arrays are summarized in Tab. 3.1 and Tab. 3.2:

Freq(MHz) / Sub-arrays	1	50	100
S1	0.16±0.01	0.16±0.05	0.16±0.05
S2	0.10±0.01	0.12±0.03	0.18±0.09
S3	1.50±0.01	0.14±0.03	0.13±0.08

(a) FPN for MIMOSA16 of 14 µm epitaxial layer

Freq(MHz) / Sub-arrays	1	50	100
S1	0.11±0.01	0.15±0.03	0.19±0.06
S2	0.16±0.01	0.16±0.02	0.13±0.01
S3	1.52±0.01	0.14±0.08	0.12±0.05

(b) FPN for MIMOSA16 of 20 µm epitaxial layer

Table 3.1: FPN (in mV_{rms}) for all the three sub-arrays of MIMOSA 16 at 1, 50 and 100 MHz: (a) 14 µm version and (b) 20 µm version.

Freq(MHz) / Sub-arrays	1	50	100
S1	0.65±0.01	0.65±0.01	0.66±0.01
S2	0.69±0.01	0.64±0.01	0.63±0.01
S3	0.80±0.01	0.67±0.01	0.68±0.01

(a) Temporal Noise for MIMOSA16 of 14 µm epitaxial layer

Freq(MHz) / Sub-arrays	1	50	100
S1	0.67±0.01	0.64±0.01	0.63±0.01
S2	0.73±0.01	0.65±0.01	0.63±0.01
S3	0.84±0.01	0.69±0.01	0.67±0.01

(b) Temporal Noise for MIMOSA16 of 20 µm epitaxial layer

Table 3.2: Temporal Noise (in mV_{rms}) for all the three sub-arrays of MIMOSA 16 at 1, 50 and 100 MHz: (a) 14 µm version and (b) 20 µm version.

In Tab 3.1, we notice that the FPN is very small (around 0.15 mV_{rms}) for all sub-arrays at 50 and 100 MHz. At 1 MHz, the FPN for the pixels of sub-array S3 is much higher than the pixels in the other sub-arrays. The pixels of sub-array S3 are designed with radiation tolerance rules with a guard-ring added around the charge detection diode, increasing both its equivalent capacitance and FPN. As a result, the FPN for the pixels in S3 is much more important than the pixels in S1 and S2 at 1 MHz. No significant difference can be observed between the 14 and 20 µm versions.

From Tab. 3.2, the Temporal Noise behaviors among the three sub-arrays are similar. More Temporal Noise is found at 1 MHz due to the increasing *1/f* noise. Stable values are obtained

at 50 and 100 MHz for each sub-array. The sub-array S3 has Temporal Noise slightly higher than the others due to its increasing equivalent capacitance.

c) Comparison of the noise performance of MIMOSA 8 and MIMOSA 16

In order to evaluate the noise performances of MIMOSA 8 and MIMOSA 16 (both 14 µm and 20 µm versions), the sub-array S_4 of MIMOSA 8 and the S_2 of MIMOSA 16 are chosen because of the same charge detection diode size: 2.4 × 2.4 µm^2. Their Temporal Noise and FPN values are summarized in the table below.

Freq. (MHz)	Temporal Noise (mV_{rms})		
	S_4 of MIMOSA 8	**S_2 of MIMOSA 16 14 µm epitaxial layer**	**S_2 of MIMOSA 16 20 µm epitaxial layer**
1	0.90±0.01	0.69±0.01	0.73±0.01
50	0.84±0.02	0.64±0.01	0.65±0.01
100	0.85±0.02	0.63±0.01	0.63±0.01

Table 3.3: Temporal Noise for S_4 of MIMOSA 8 and S_2 MIMOSA 16 at three different main clock frequencies: 1, 50 and 100 MHz.

From the results above, for all the three versions of the chip, the Temporal Noise increases slightly at 1 MHz due to the un-cancelled 1/f noise. For the frequencies 50 and 100 MHz, the Temporal Noise is rather stable. The value of the Temporal Noise of MIMOSA 8 is more important than that of MIMOSA 16. For example: for the sub-array S_4 of MIMOSA 8, its value is 0.85 mV_{rms} and for the sub-array S_2 of MIMOSA 16, the value is less than 0.65 mV_{rms}. In order to understand the reason of this improvement, a simple test has been performed with MIMOSA 8 by using the USB data acquisition system at 10 MHz. The Temporal Noise found by this test is also 0.85 mV_{rms}, the same as the former results. For this reason, the improvement of the Temporal Noise performance is probably due to the different CMOS processes. The AMS 0.35 µm OPTO process is optimized for the mixed circuit design. The circuit elements (such as the MOS capacitor used inside pixel for memorizing the reference level) in the simulator are generally well modelled. However, the TSMC 0.25 µm digital process is optimized for the digital circuit design. We choose it for the only reason that the mixed-signal version of this process has no epi-layer.

The FPNs are summarized in Tab. 3.4. We notice that the FPN values are stable as the function of the main clock frequency except that it increases slightly at 100 MHz for MIMOSA 8. As explained before, one possible reason is that the period for charging the offset storage capacitors is not sufficient for the signal to stabilize. Another important reason is the difficulty of the measurement at such a high readout speed. The additional delay of the sampling clock is needed in order to sample the analog output signal when it is already stable, avoiding the glitches and the fluctuation of the output signal. But it is very difficult to choose a correct delay value when the readout period is only tens of nano-seconds.

Freq. (MHz)	Fixed Pattern Noise (mV_{rms})		
	S_4 of MIMOSA 8	**S_2 of MIMOSA 16 14 µm epitaxial layer**	**S_2 of MIMOSA 16 20 µm epitaxial layer**
1	0.30±0.02	0.10±0.01	0.16±0.01
50	0.32±0.01	0.12±0.03	0.16±0.02
100	0.45±0.05	0.18±0.09	0.16±0.02

Table 3.4: FPN for S_4 of MIMOSA 8 and S_2 MIMOSA 16 at three different main clock frequencies: 1, 50 and 100 MHz.

In general, from Tab. 3.3 and 3.4, the pixels fabricated in two different CMOS processes exhibit very low level Temporal Noise (less than 1 mV) and very small dispersion (the FPN level is about 0.3 mV, largely below Temporal Noise level). This excellent noise performance indicates that the in-pixel auto-offset cancellation mechanism works efficiently.

3.3.2 Charged particle detection performance

Having acquired the noise parameters of these two kinds of chips, other important parameters determining the performance of chip, such as Charge-to-Voltage conversion Factor (CVF) and Charge-Collection Efficiency (CCE), are calibrated by using a ^{55}Fe radioactive source delivering low energy X-ray photons with well-known energy peaks. This source delivers photons with two emission energies: K_α = 5.9 keV (24.4% emission probability) and K_β = 6.4 keV (2.8% emission probability). In silicon, it is known that to create an electron-hole pair needs a deposit energy of 3.6 eV. Therefore, about 1640 electron-hole pairs can be created in silicon by the X-ray photons of 5.9 keV. This value gives us a reference during our test.

A. The charge collection mechanisms

Fig. 3.9 illustrates three mostly possible charge collection mechanisms inside a CMOS sensor volume, where three different locations for charge collection are indicated.

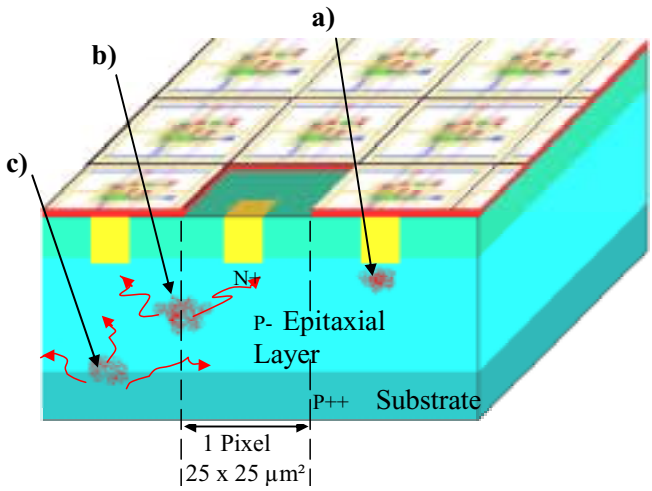


Figure 3.9: The charge collection mechanisms: (a) Single diode total charge collection; (b) Cluster collection and (c) Diffusion of charge inside substrate.

The three cases of interaction of photons with CMOS sensors are described below:

a) Single diode total charge collection

An X photon interacts in the depletion zone of the n-well in pixel and all the charge is collected by the single diode of the pixel, which is a quite rare case due to the small diode size.

A CCE of 100% is achieved in this case and the Calibration Peak, corresponding to the 5.9 keV energy level, is visible in signal height distribution spectrum for photons (Fig. 3.11). All the 1640 electrons produced by a photon are collected by a single diode so that the CVF value can be calculated with the position of the Calibration Peak. The relationship is given by the following expression:

$$CVF\ (\mu\mathrm{V/e\text{-}}) = \frac{P_{Calib\_Peak} \times 500\,\mu\mathrm{V}}{1640 e^{-}} \qquad (3.11)$$

where P_{Calib_Peak} is the position of the Calibration Peak expressed in ADC units. 1 ADC unit is equal to 500 µV in our data acquisition system.

Since the case where all the energy is absorbed by one single diode occurs rarely, a large amount of experimental data is required in order to get a visible Calibration Peak.

b) Cluster collection

An X photon interacts in the epitaxial layer and the generated charge is shared by the diodes of adjacent pixels.

It is the most general case which appears during the measurement. The charge is naturally distributed among several pixels and the charge collected by one pixel is thus only a part of the total charge. By summing the charge collected by a cluster of neighboring pixels, the total charge is more or less reconstructed. But 100% of the charge induced in the epitaxial layer is practically impossible to build due to loss during diffusion from the interaction point and due to thermal recombination of electrons before collection. During our tests, clusters of 3×3 and 5×5 are used to calculate the total charge collected by sensors and the corresponding dominant broad peak which appears in photon spectrum is called Charge Collection Peak (Fig. 3.20). The ratio between the position of the Charge Collection Peak and that of Calibration Peak is defined as CCE (Charge Collection Efficiency):

$$CCE \quad (\%) = \frac{P_{cluster}}{P_{calib}} \times 100\% \qquad (3.12)$$

c) Diffusion of charge inside substrate

An X photon interacts in the heavily doped substrate and most of the generated charge is lost. In this case, there is only a very small charge collected, which corresponds to the lower part of photon spectrum (see Fig. 3.20).

B. Measurement of the analog output

The ^{55}Fe source is placed in a dark box at approximately 1 cm of the chip with no material in between and the analog outputs are firstly observed on an oscilloscope. An example of the sub-array S2 of MIMOSA 8 is shown in Fig. 3.10 below. After 15 min of acquisition time, the hits could be seen on the scope as jumps on the output signal during the RD phase. The detailed studies are carried out by data acquisition system.

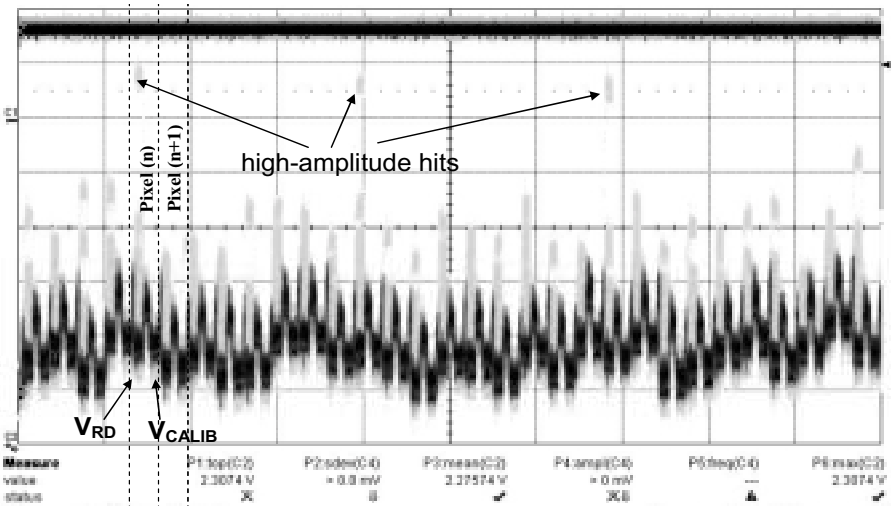
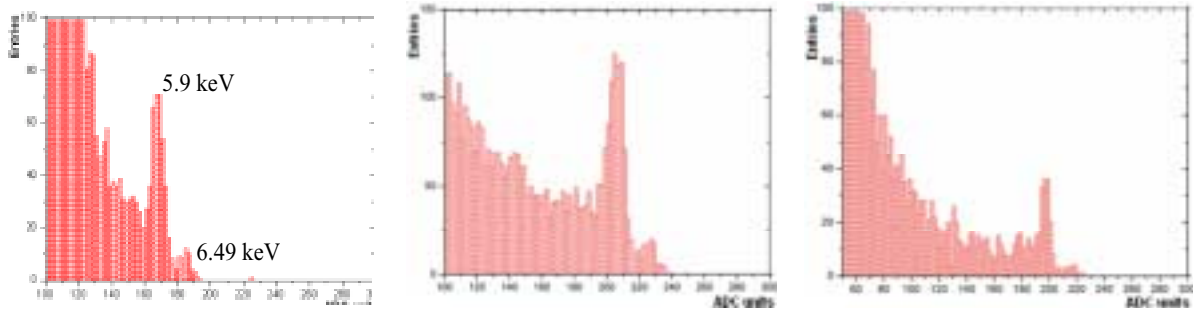


Figure 3.10: Output analog signal from 32 pixels of a column recorded on a scope (20 mV/div, f = 10 MHz) with multiple high-amplitude hits for the DC-coupling version of the pixel design.

a) The measurements of the Calibration Peak

The Calibration Peak is obtained by the spectrum of the collected charge of one pixel, corresponding to the case of total charge collection. The following example (Fig. 3.11) gives the Calibration Peaks for the three prototypes at 10 MHz. The corresponding sub-arrays are S4 for MIMOSA 8 and S2 for MIMOSA 16 (both 14 µm and 20 µm versions) with the diode size of 2.4 × 2.4 µm².

It can be seen that the Calibration Peak situates almost in the same position for the two versions of MIMOSA 16 but the case is different for MIMOSA 8. This difference can be explained by the different equivalent capacitance of the charge collection diode in the different CMOS processes.

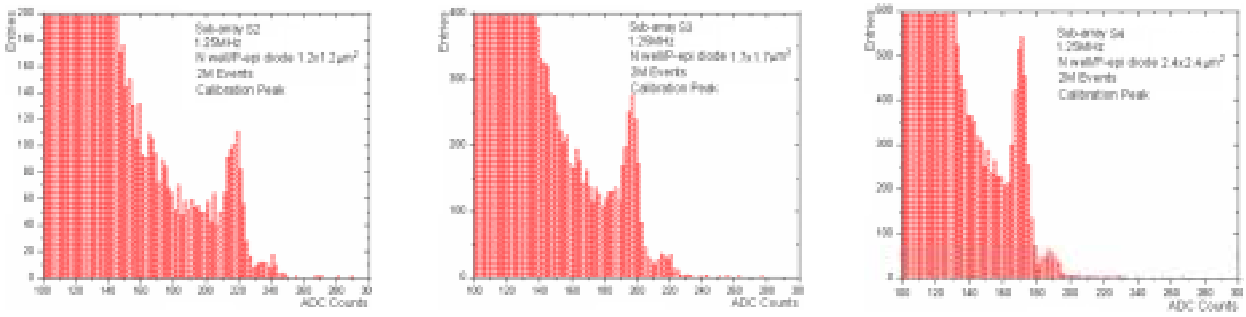


(a) Sub-array S4 of MIMOSA 8 (b) Sub-array S2 of MIMOSA 16: 14 µm version (c) Sub-array S2 of MIMOSA 16: 20 µm version

Figure 3.11: Calibration Peaks of the sub-arrays that the diode size of the pixel is 2.4 x 2.4 µm². The measurements are performed at 10 MHz for all the three prototypes.

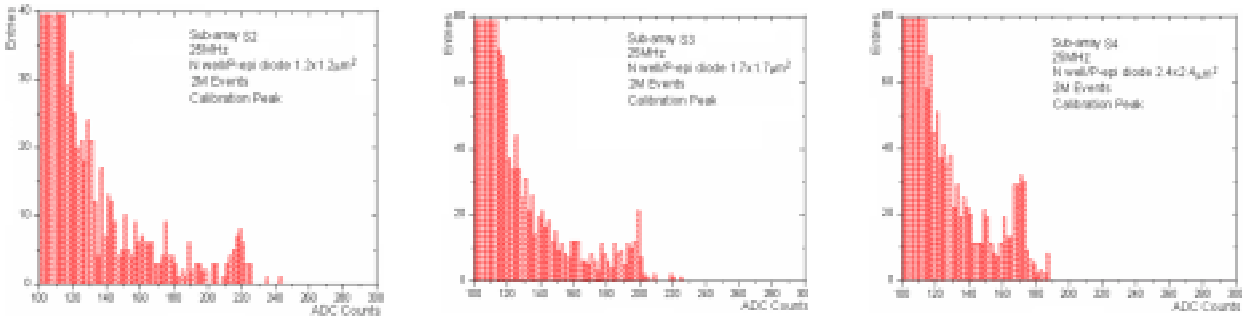
The Calibration Peaks have been measured as the function of the main clock frequencies. The measurements have been performed up to 75 MHz for MIMOSA 8 and 170 MHz for MIMOSA 16.

The following examples (Fig. 3.12 and Fig. 3.13) give the Calibration Peaks of all the three sub-arrays for MIMOSA 8 at main clock frequency of 1.25 MHz and 25 MHz, respectively.



(a) Sub-array S2 (b) Sub-array S3 (c) Sub-array S4

Figure 3.12: Calibration Peaks for the three sub-arrays of MIMOSA 8 at 1.25 MHz.



(a) Sub-array S2 (b) Sub-array S3 (c) Sub-array S4

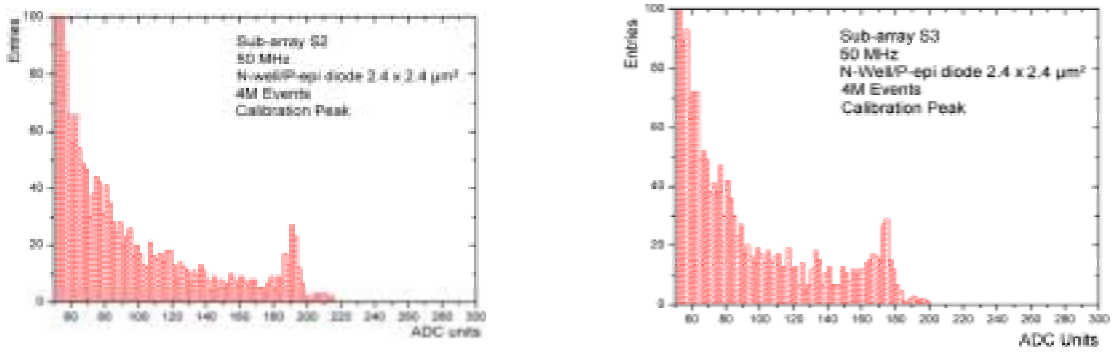
Figure 3.13: Calibration Peaks for the three sub-arrays of MIMOSA 8 at 25 MHz.

Two millions events are taken in both frequencies in order to obtain a clearly visible Calibration Peak. Comparing the results in Fig. 3.12 and Fig. 3.13, the amplitude of the Calibration Peak found at 1.25 MHz is more important than at 25 MHz. The reason is that readout period is longer at 1.25 MHz and in this case more photons are emitted as the activity of source is constant, increasing the number of hits found in the measurement during one readout cycle.

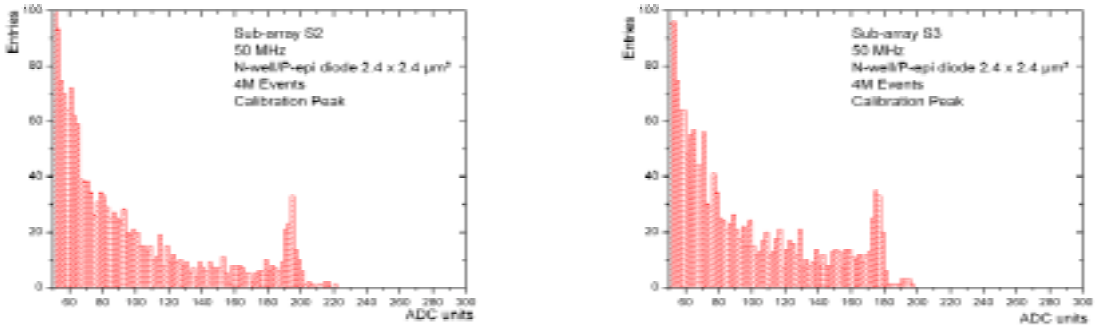
The Calibration Peak for the sub-arrays of MIMOSA 16 are obtained at much higher main clock frequencies thanks to the more active radioactive source used since the year 2006. However, the Calibration Peak of sub-array S1 has not been obtained due to its small diode size ($1.7 \times 1.7\ \mu m^2$). In MIMOSA 8 (TSMC 0.25 µm), this diode size (sub-array S3) gives interesting performances but in MIMOSA 16 (AMS 0.35 µm OPTO) the same diode size (sub-array S1) gives very poor performances for charge collection efficiency. This illustrates perfectly the complexity to keep the same performances with the same circuit architecture for different CMOS processes.

For this reason, the sub-array S1 will not be concerned in the following tests. For the sub-arrays S2 and S3 of MIMOSA 16 (both 14 µm and 20 µm versions), the examples below give their Calibration Peaks for the main clock frequencies of 50 MHz and 100 MHz.

As an example, the Calibration Peaks of sub-array S2 for these two frequencies are given in Fig. 3.14 and Fig. 3.15

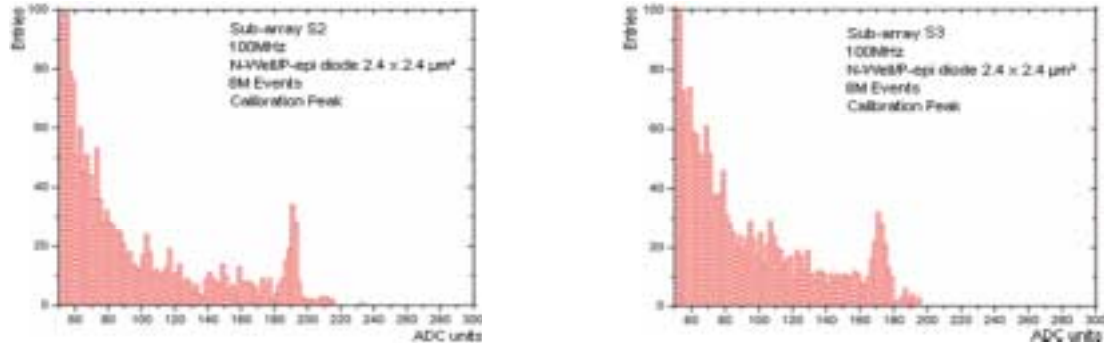


(a) Chip with 14 µm epitaxial layer

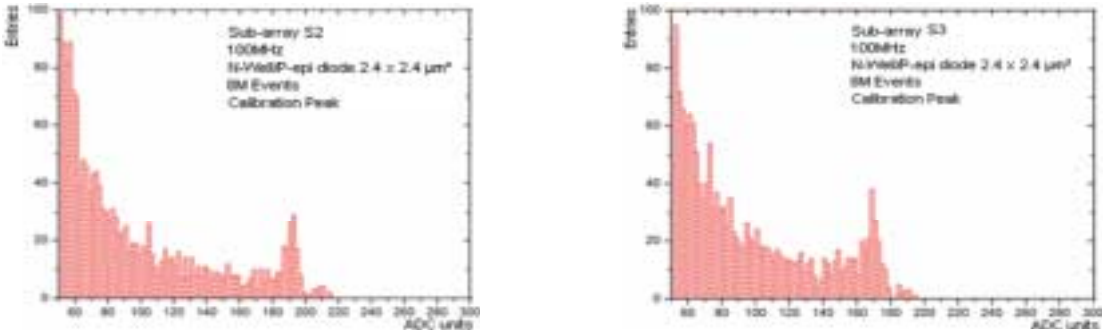


(b) Chip with 20 µm epitaxial layer

Figure 3.14: Calibration Peaks for the sub-arrays S2 and S3 of MIMOSA 16 at 50 MHz.



(a) Chip with 14 µm epitaxial layer



(b) Chip with 20 µm epitaxial layer

Figure 3.15: Calibration Peaks for the sub-arrays S2 and S3 of MIMOSA 16 at 100 MHz.

From the above examples, we notice that the position of the Calibration Peak for each prototype is almost unchanged when increasing the clock frequency. It is a very important result. As explained in the former section, the Calibration Peak is obtained when all the generated charge is collected by a single pixel. Theoretically, its position does not depend on the working frequency and varies only with the diode size when a ^{55}Fe source is used (the total charge generated inside the silicon for this source is constant: about 1640 e-). However, it is not the case in practical. The position of the Calibration Peak is usually affected by the temporal readout noise and the offset, which means that the performance of the chip is deteriorated. The deterioration appears especially when the chip works at high readout frequency. Fortunately these sources of disturbance are well controlled by using the DC-coupling pixel structure and the results obtained here are quite stable even at 100 MHz (corresponding to a readout speed of about 20 µs/frame). Therefore, the pixel level double sampling data processing works efficiently and is stable in function of the main clock frequency.

No difference has been observed between the two versions of MIMOSA 16.

Using (3.11), the CVF values are calculated. The results of the different sub-arrays for all the prototypes are summarized in the following table:

	Charge-to-Voltage Conversion Factor (CVF)
MIMOSA 8	S2: CVF = 66 µV/e-; S3: CVF = 60 µV/e-; S4: CVF = 52 µV/e-
MIMOSA 16: 14 µm	S2: CVF = 59 µV/e-; S3: CVF = 54 µV/e-
MIMOSA 16: 20 µm	S2: CVF = 60 µV/e-; S3: CVF = 53 µV/e-

Table 3.5: CVF values for the different sub-arrays of all the prototypes. The error of measurement is negligible.

The smaller CVF is obtained when the size of the charge detection diode is bigger. It is a reasonable result because from equation (1.1) in Chapter 1, we have:

$$\mathrm{CVF} \propto 1/C_{eq}, \qquad (3.13)$$

where C_{eq} is the equivalent capacitance of the charge sensing diode and is proportional to the diode size. We notice that the CVF values are different for sub-array S4 of MIMOSA 8 and sub-array S2 of MIMOSA 16 while the diode size of the pixels inside these two sub-arrays is the same. This is due to the different equivalent capacitance introduced by the different CMOS processes used for these two prototypes.

From equation 3.13, although a relative higher CVF can be achieved by pixel of smaller diode size, it collects less charge. It means that the pixel with small diode size is lack of charge collection ability compared to the pixel where the diode size is bigger. On the contrary, for the pixel with big diode size, although they collect more charge, its CVF is relatively low and it suffers from more disturbance from the noise. Therefore, compromise has to be found between the CVF value and the charge collection ability.

We can compare the real CVF values measured with experimental data with the theoretical values obtained by simulation. For this purpose, it is necessary to learn the parasitic capacitances that contribute to the total value of equivalent capacitance. The circuit of DC-coupling pixel is represented in Fig 3.16, with all the possible parasitic capacitance. Here, the circuit of CS stage with a select switch is shown in detail.

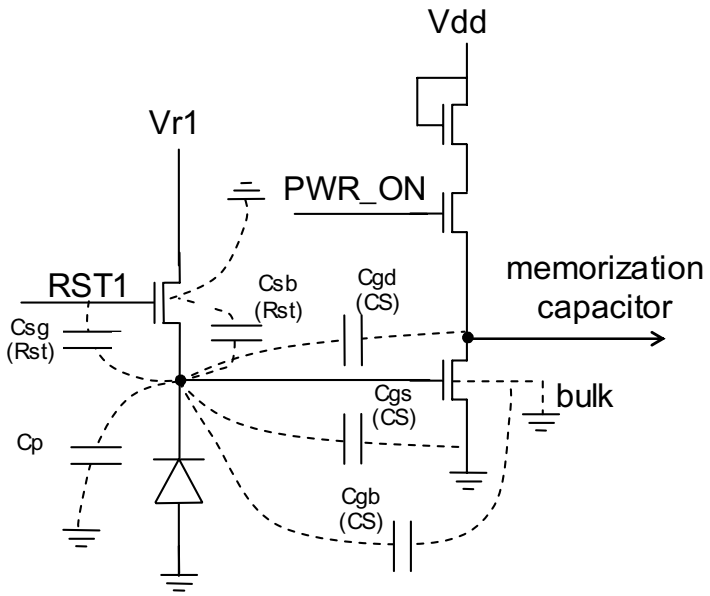


Figure 3.16: The interconnection of charge detection diode, reset switch and CS amplifier stage with all parasitic capacitances of the diode.

C_{gd}(CS), C_{gs}(CS) and C_{gb}(CS) are gate-drain parasitic capacitance, gate-source parasitic capacitance and gate-bulk capacitance of the CS stage, respectively. C_{sb}(Rst) and C_{sg}(Rst) represent source-bulk and source-gate parasitic capacitance of the reset transistor (RST1). Cp is the equivalent capacitance of the line used for the interconnection between the diode and the transistors, which is the sum of C_{m1} and C_{poly}. C_{m1} and C_{poly} are the equivalent capacitances of metal 1 and polysilicon used for realizing the layout of the pixel. The values of all these capacitances can be found in the technical document of TSMC 0.25 µm digital process.

The total equivalent capacitance is the sum of all parasitic capacitances and the equivalent capacitance of the charge detection diode.

$$C_{total} = C_d + C_p + C_{gs}(CS) + (1 + A_{CS})C_{gd}(CS) + C_{gb}(CS) + C_{sg}(Rst) + C_{sb}(Rst) \quad (3.14)$$

where C_d is the equivalent diode capacitance and A_{cs} is the gain of the CS stage.

The theoretical values of capacitors for MIMOSA 8 are calculated. Using the value of area junction capacitance and side wall junction capacitance given by TSMC 0.25 µm digital process, C_d is calculated for pixels of the three different sub-arrays. The results are as follows:

S2: C_d = 2.2 fF

S3: C_d = 3.2 fF

S4: C_d = 4.7 fF

Using the TSMC design guide, the value of equivalent parasitic capacitor is 5.04 fF. Then, the different CVF values obtained by simulation are:

S2: CVF = 71 µV/e-

S3: CVF = 63 µV/e-

S4: CVF = 53 µV/e-

Compared to the real CVF values of MIMOSA 8 (Tab. 3.5), the similar results indicate that the objective of design is achieved. The bigger charge detection diodes have more accurate previewed CVF value only because that their high capacitance is more robust to the parasitic parameters.

b) The measurements of charge collection peak

As explained before, the most common case of the charge collection is, when the generated charge distributes inside the epitaxial layer is collected by neighboring pixels thanks to the thermal diffusion. Therefore, in order to study the charge distribution on the pixel array, clusters of pixels are analyzed. The collected charge was studied as the function of the cluster size (the number of the pixels contained inside the cluster) [5] [10]. The results of beam test with MIPs (measured in 20°C) are shown in Fig. 3.17. In the analysis procedure, the pixels are ordered w.r.t. the collected charged and the last pixels contribute with noise only. Due to the ordering procedure, these noise contributions are positive at first and then negative. It can be seen that the charge spread is limited to the closest 25 pixels (a 5×5 cluster built around the central pixel who collects the maximum part of charge). Within this area, almost all the available charge is collected using a lower limit threshold of 5σ on measured signals (the value of σ represents the Temporal Noise) [5] [6] [10]. For this reason, the maximum cluster size studied here is 5×5 pixels.

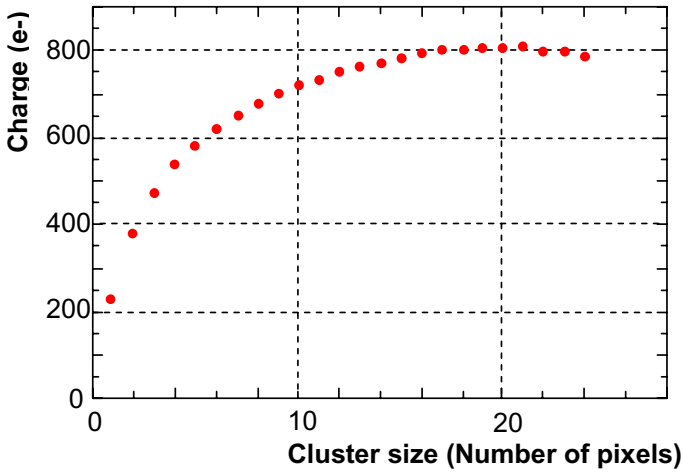


Figure 3.17: Collected charge as a function of the number of pixels contributing to the cluster at 20°C for MIPs. MIMOSA 9 fabricated in AMS 0.35 µm OPTO [10].

The algorithm of the cluster construction is described in the following.

Because the pixels of different sub-arrays have different parameters, each sub-array is considered separately in analysis. The procedure repeats event per event.

For the nth event of the experimental data:

- The first step is to calculate the SNR value for the whole sub-arrays of this event.

  As the Temporal Noise value of each pixel has been already obtained through the noise measurements, we have the SNR value for the kth pixel in the sub-array:

$$SNR_n(k) = \frac{S_{signal,n}(k)}{\sigma_{tem}(k)} \qquad (3.15)$$

  where $\sigma_{tem}(k)$ is the Temporal Noise of the kth pixel, $S_{signal,n}(k)$ is the signal value of this pixel in the nth event.

  The value of $S_{signal,n}(k)$ is given by:

$$S_{signal,n}(k) = S_{pixel,n}(k) - P_{ped}(k), \qquad (3.16)$$

  where $S_{pixel,n}(k)$ is the CDS value (V_{RD} - V_{CALIB}) of the kth pixel in the event n and $P_{ped}(k)$ is the pedestal value of the kth pixel obtained by noise measurements.

- The second step is the hit search.

The former SNR value calculated for each pixel is used to determine the central pixel (or seed pixel) around which the cluster is constructed. The detail algorithm is described in the diagram below:

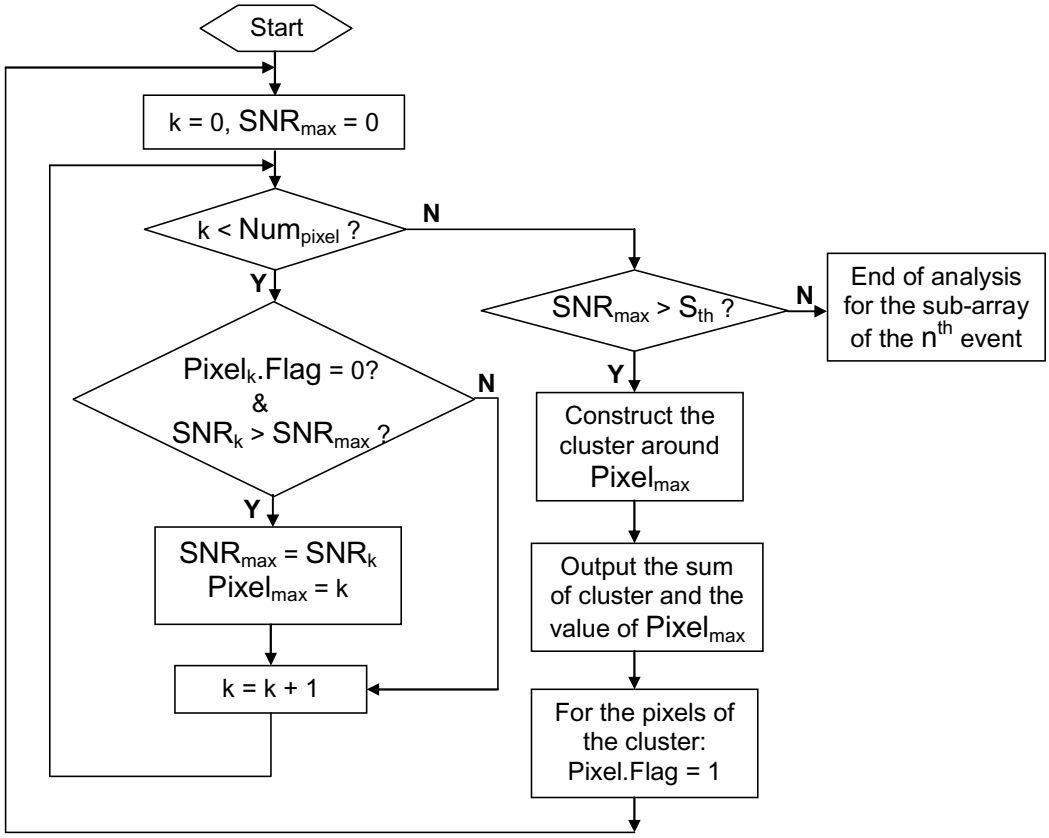


Figure 3.18: Diagram of cluster construction algorithm.

In Fig. 3.18, Num_{pixel} is the total pixel number of the sub-array and S_{th} is a predetermined threshold value used for analysis.

The algorithm searches firstly for the pixel with maximum SNR value. When it has been found, it is used as central pixel for the construction of the cluster. As an example, Fig. 3.19 shows how a 5 × 5 pixels cluster is constructed around the central pixel. In fact, the cluster represents a hit generated by the diffused charge of an incident photon. Therefore, the charge collection peak can be obtained by the sum of the charge of the pixels in the cluster. The value of the central pixel and the sum of the charge collected by all pixels in this cluster are stored in result file.

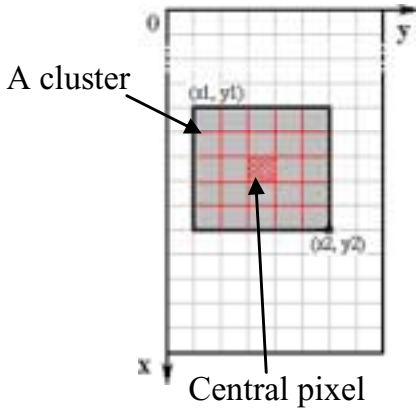
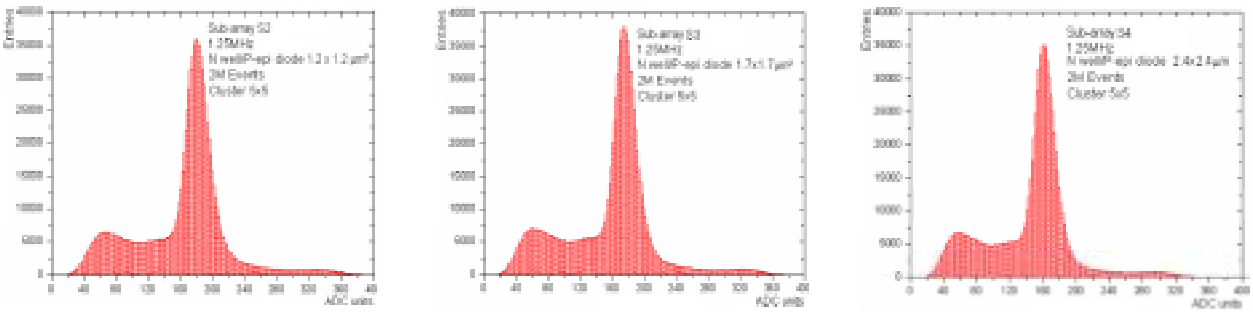


Figure 3.19: Construction of 5 × 5 cluster around the central pixel.

The pixels containing inside the cluster found are then flaged and a new search of central pixel, excluding the former flaged pixels, begins again for the remaining pixels of the sub-array. The search stops when there is no pixel with its SNR value more important than S_{th} (5σ for example). The analysis of one event terminates when the search of all sub-arrays is finished.

Sometime it is possible that two photons hit simultaneously adjacent area. The result is that two hits found in the same event are very close, making two corresponding clusters overlap. In this case, it is impossible to correctly identify the charge collected of each cluster so that the overlapping clusters are removed in analysis.

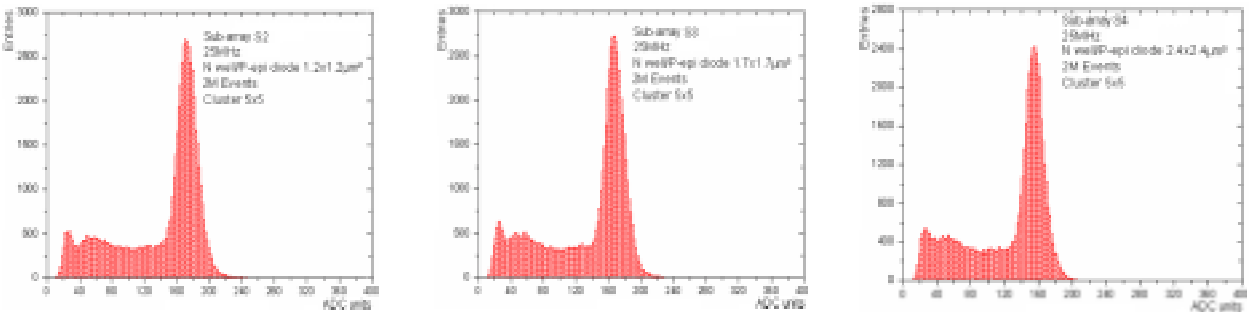
Using this algorithm, the charge collection peak of 5 × 5 pixels cluster is obtained from 1.25 MHz to 150 MHz for MIMOSA 8 and the charge collection peak of 3 × 3 pixels cluster is obtained from 1 MHz to 170 MHz for MIMOSA 16*. The examples of the Charge Collection Peaks for MIMOSA 8 at frequencies of 1.25 MHz, 25 MHz are shown in Fig 3.20 and Fig. 3.21, respectively.



(a) Sub-array S2 (b) Sub-array S3 (c) Sub-array S4

Figure 3.20: Charge Collection Peaks for the three sub-arrays of MIMOSA 8 at 1.25 MHz.

* MIMOSA 16 is tested with USB data acquisition system which allows only 4 columns of analog data to be stored simultaneously. Consequently, we are limited by cluster with maximum size of 3 x 3.

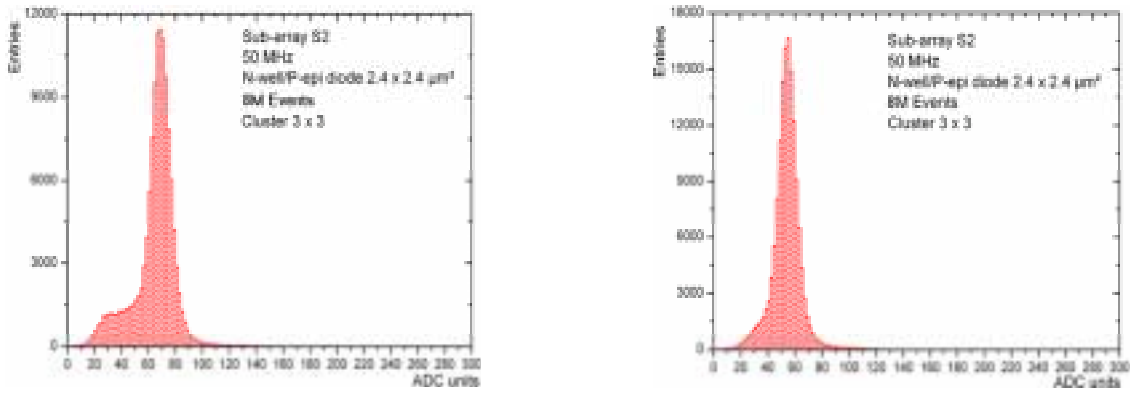


(a) Sub-array S2 (b) Sub-array S3 (c) Sub-array S4

Figure 3.21: Charge collection Peaks for the three sub-arrays of MIMOSA 8 at 25 MHz.

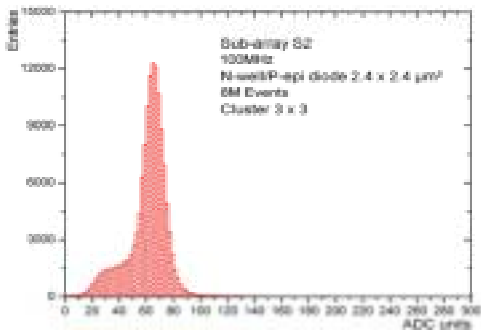
At 1.25 MHz, it can be seen that the charge distribution spectrum takes a much larger range. For example, shown on Fig. 3.20 the range of the charge distribution spectrum for the sub-array S2 at 1.25 MHz reaches at about 380 ADC units, largely beyond the position of corresponding calibration peak (Fig. 3.12). This phenomenon is due to the long integration time. It is about 200 µs when the main clock frequency is 1.25 MHz. Since the ^{55}Fe source used in measurement is very active, one pixel is probably touched more than once by incident photon during one integration period (about 200 µs, which is calculated by Eq. 2.13: $128\times16\times1/1.25\times10^{-6}$). As explained before, the charge detection diode is not reset during the integration, the charge issued from several photons accumulates and is collected at the end of integration period. The most frequent case is that two photons hits successively in one cluster so that a small peak is visible where the position is twice of the Charge Collection Peak. But at higher main clock frequency, such as 25 MHz, the integration period is relatively short and only one photon can interact with the same pixel.

The Charge Collection Peaks of sub-array S2 for MIMOSA 16 (14 µm and 20 µm version) at main clock frequencies of 50 MHz and 100 MHz are shown in Fig. 3.22 and Fig. 3.23, respectively.

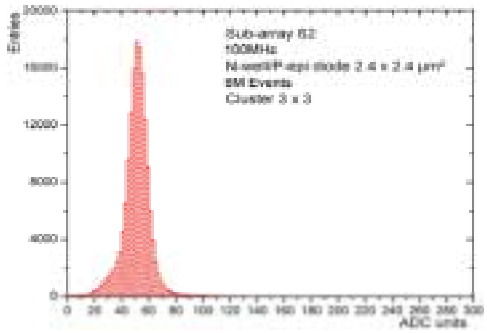


(a) Sub-array S2 for 14 µm epitaxial layer version (b) Sub-array S2 for 20 µm epitaxial layer version

Figure 3.22: Charge collection Peaks for sub-array S2 of MIMOSA 16 at 50 MHz.



(a) Sub-array S2 for 14 µm epitaxial layer version



(b) Sub-array S2 for 20 µm epitaxial layer version

Figure 3.23: Charge collection Peaks for sub-array S2 of MIMOSA 16 at 100 MHz.

Charge collection distributions of sub-array S3 are very close to those obtained for sub-array S2.

Comparing the results of MIMOSA 16 with 14 and 20 µm epitaxial layers, we notice that more important CCE value is obtained for the version having 14 µm epitaxial layer. One possible explanation is as follows:

With the ^{55}Fe source, the total charge of a free electron generated by a photon inside epitaxial layer is constant. In the most probable case, the free carriers of photon are generated inside the epitaxial layer and move towards to the charge sensing diode by thermal diffusion. Thus, the chip of 20 µm epitaxial layer increases the diffusion length before the collection of charge. Two results issue from this assumption. Firstly, during the travel of free electrons, more charge will be lost by generation-recombination due to the increasing diffusion path length. Secondly, the electrons can diffuse more widely so that its charge could be shared by more pixels. Therefore, in order to have the same rate of charge collection, a cluster with more pixels is needed when the analysis of data for the 20 µm epitaxial version is carried on.

Further simulations are needed in order to verify the above assumption.

According to (3.12), the CCE values of different sub-arrays for MIMOSA 8 and MIMOSA 16 can be then calculated. The results are summarized in Tab. 3.6 and 3.7.

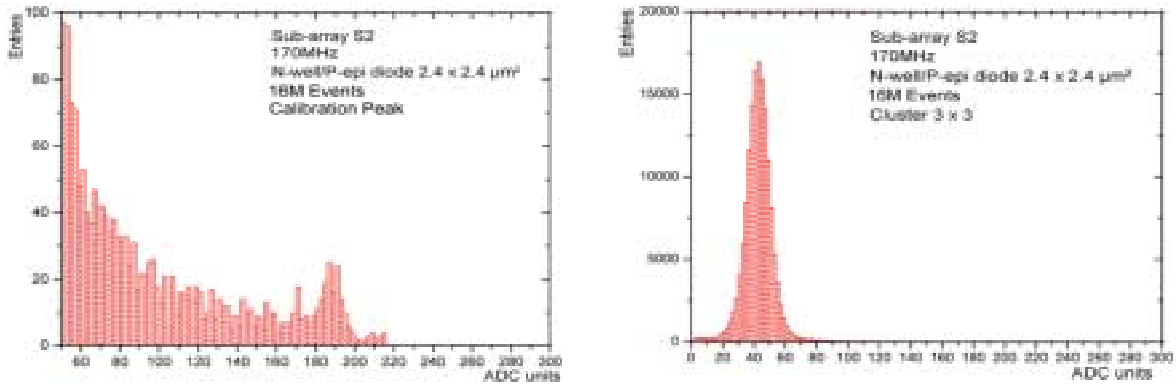
MIMOSA 8	**Charge Collection Efficiency (CCE) (5×5 cluster)**	
	1.25 MHz	**25 MHz**
S2	81.8%	79.5%
S3	90.4%	85.9%
S4	94.1%	92.9%

Table 3.6: CCE of sub-arrays S2, S3 and S4 for MIMOSA 8 at 1.25 MHz and 25 MHz. The measurement error is negligible.

MIMOSA 16	Charge Collection Efficiency (CCE) (3×3 cluster)	
	50 MHz	100 MHz
14 µm	36.5%	35.8%
20 µm	26.9%	25.9%

Table 3.7: CCE of sub-arrays S2 for 14 and 20 µm epitaxial layer version of MIMOSA 16 at 50 MHz and 100 MHz. The measurement error is negligible.

In the measurement, the error is negligible. The maximum main clock frequencies used for testing MIMOSA 8 and MIMOSA 16 are 150 MHz and 170 MHz. Using Eq. 2.13, we can get the equivalent readout speeds are 13 µs/frame and 12 µs/frame, respectively. The positions of the Charge Collection Peak for all these prototypes remain stable at these frequencies. Considering that the stable position of the Calibration Peak, it can be said that the CCE for the DC-coupling pixel is unaffected by the main clock frequency. Although the DC-coupling pixel is originally made to ensure functionality at 100 MHz of the main clock frequency, the measured results indicates that the pixel chip works efficiently at main clock frequency up to 170 MHz (Fig. 3.24). A very fast readout speed of about 12 µs/frame, being one of the design objectives, is almost reached. Slight decrease of the CCE value can be observed at 170 MHz. It is because the corresponding readout speed reaches the limits of the chip. Readout speed of 12 µs/frame used for charging the series capacitor inside the pixel is not long enough for signal to be fully stored on it before it is sampled. So that only part of the detected signal is sampled and transferred to the output buffer. For this reason, the measurement at higher frequency (more than 170 MHz) is not carried out.



(a) Calibration Peak (b) Charge Collection Peak

Figure 3.24: The Calibration Peak and (a) the Charge Collection Peak (b) of sub-array S2 for MIMOSA 16 of 20 µm epitaxial layer at main clock frequency of 170 MHz.

To summarize, the evaluations of different sub-arrays' CCE as the function of the main clock frequency are given in the following. Fig. 3.25 is for MIMOSA 8's, from 1.25 MHz to 150 MHz. Fig. 3.26 is for different versions of MIMOSA 16's.

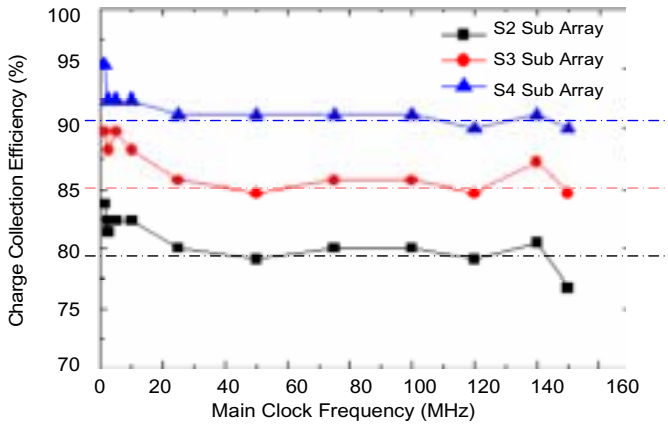
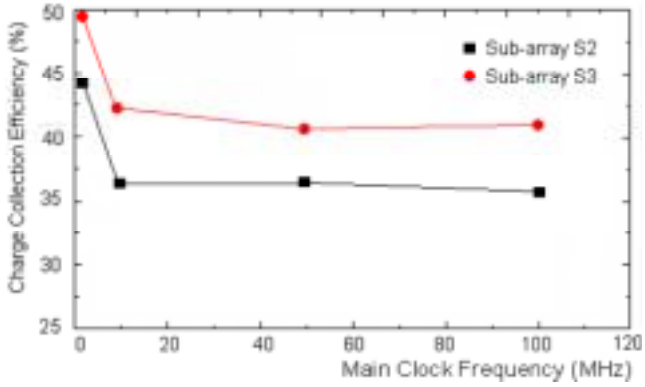
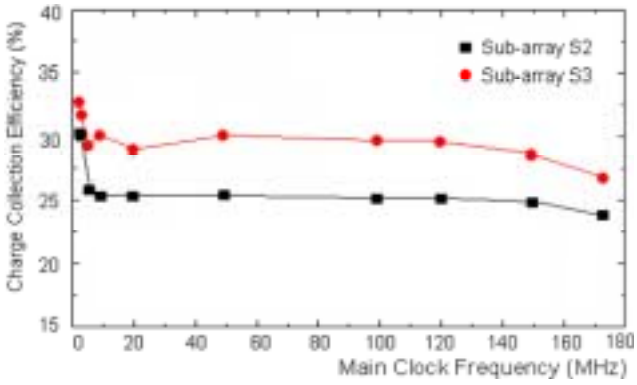


Figure 3.25: Evaluation of CCE of S2, S3 and S4 for MIMOSA 8 as the function of main clock frequency.



(a) 14 µm epitaxial version



(b) 20 µm epitaxial version

Figure 3.26: Evaluation of CCE of S2, S3 for MIMOSA 16 as the function of main clock frequency.

From these results, the measurement error is negligible. We notice that remarkably stable performances are obtained for both MIMOSA 8 and MIMOSA 16 on a rather large frequency range (from 24 MHz to 140 MHz for MIMOSA 8 and from 10 MHz to 150 MHz for MIMOSA16). Higher CCEs are obtained when the prototypes work at very low frequency (1 MHz). As explained before, the reason is that the source used in test is so active that the corresponding integration time is long enough for more than one photon interacting with the

same pixels. Therefore, the charge collected at very low frequency is probably deposit by several photons.

The fastest readout speed, 12 µs/frame, has been obtained with MIMOSA 16 of 20 µm epitaxial layer when the prototype is tested with a main clock frequency of 170 MHz. In the application of CMOS sensor for MIPs detection, it is the first chip that achieves such a high readout speed.

At one frequency, the CCE depends only on the equivalent capacitance of the charge detection diode determined by its size and the parasitic capacitance around it. For example, in MIMOSA 8, the CCE varies from 80% for S2 sub-array up to 90% for S4 sub-array. The sub-array with the biggest diode size achieves the most important CCE value. However, a main drawback of the bigger diode size is that its noise increases with the area. Moreover, the corresponding CVF value decreases also. Thus, a compromise has to be found among noise level, CVF value and CCE value.

3.3.3 Comparison between MIMOSA 8 and MIMOSA 16

Using USB data acquisition system, measurements have been made at 10 MHz for both MIMOSA 8 and MIMOSA 16. The results are compared in this section.

A. Noise performance

With the knowledge of the CVF for different sub-arrays, the equivalent input referred noise can be calculated. The results measured at 10 MHz are summarized below in Tab. 3.8 and Tab. 3.9 for MIMOSA 8 and MIMOSA 16, respectively.

	S2	**S3**	**S4**
Temporal Noise (e-)	11	12	14
Fixed Pattern Noise (e-)	4	5	6

Table 3.8: Input referred noise of MIMOSA 8. Measured at 10 MHz with USB data acquisition system.

	S2	**S3**
Temporal Noise (e-)	11	12
Fixed Pattern Noise (e-)	2	2

(a) 14 µm epitaxial version

	S2	**S3**
Temporal Noise (e-)	11	13
Fixed Pattern Noise (e-)	2	2

(b) 20 µm epitaxial version

Table 3.9: Input referred noise performance of MIMOSA 16. Measured at 10 MHz with USB data acquisition system.

The noise (Temporal Noise and FPN) is very small for both of the prototypes and the estimated error is less than 2% as mentioned before. The noise of MIMOSA 8 is slightly higher than MIMOSA 16 for equivalent diode size. One of the probable reasons is the different data acquisition system used. But tests have been performed with the USB data acquisition system for measuring the noise of MIMOSA 8 and similar results have been obtained. Thus, the difference of the noise performance between these two prototypes is most probably due to the different CMOS processes. Nevertheless, the excellent noise performance shows that the DC-coupling pixel structure with the integrated inside-pixel CDS operation works efficiently for all of these CMOS processes despite their different offsets and mismatches.

B. Charge collection performance

Take into account the pixels with the same charge detection diode size for MIMOSA 8 (sub-array S4) and MIMOSA 16 (sub-array S2), the CCE decreases with the increasing of the thickness of the epitaxial layer (Tab. 3.10). One of the probable explanations is that the increasing epitaxial layer increases also the diffusion path length which means that there would be more loss due to the thermal recombination and that the charge diffuses more largely. Further studies are needed to verify this hypothesis.

Chip	Thickness of epitaxial layer (µm)	3 x 3 cluster CCE (%)
MIMOSA 8 (S4, 2.4×2.4 µm²)	8	~82%
MIMOSA 16 (S2, 2.4×2.4 µm²)	14	~36%
	20	~25%

Table 3.10: The thickness of epitaxial layer and the corresponding CCE values. Measured at 10 MHz with USB data acquisition system. The measurement error is negligible.

3.4 Test of binary output

As mentioned in Chapter 2, there are two main design objectives of DC-coupling pixels:

- to realize high readout speed pixels with integrated offset-cancellation architecture;
- to realize 1-bit binary outputs by the integrated column level comparator.

The tests of the analog outputs demonstrate the excellent performance of the DC-coupling pixels. Very small FPN has been obtained for the pixels. However, the digital outputs suffer also from the FPN of the comparators due to the numerous parallel columns. Moreover, although the external threshold compensates the influence of mean offset level for the

comparators, it decreases on the other hand the sensibility of the sensor to small input signals. Thus, a compromise has to be found.

3.4.1 Noise performance of the digital outputs

Apart from the pixels, the Temporal Noise and FPN of the comparators are studied thanks to the test comparator separated from the pixel matrix. All the test comparators are studied here for the main clock frequency up to 100 MHz (equivalent to a readout speed of about 20 µs/frame). The test system, including a precision threshold voltage generator and a digitizing scope, is shown in Fig. 3.27.

Once the chip is programmed, the control signal needed for comparator is also available. Generated by a precision voltage generator, the threshold voltage is applied to two reference voltage inputs available on the Frond-End Board. A digitizing scope is used to visualize the output of the 1-bit comparator and also to compute the average number of "1" over a significant number of cycles.

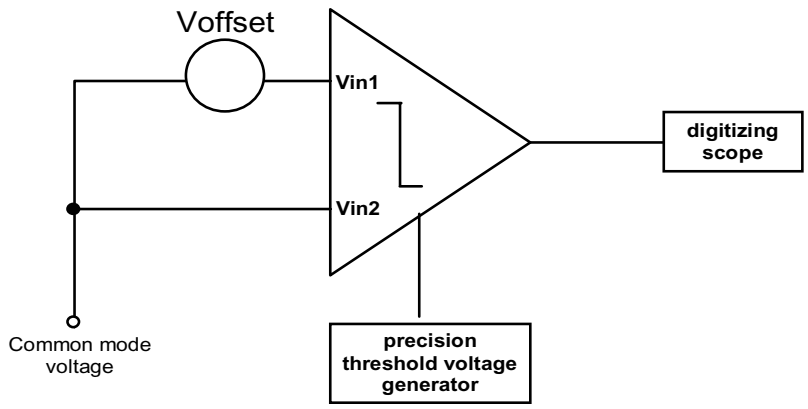


Figure 3.27: Test system with its precision threshold voltage generator, a digitizing scope and the comparator.

The test is performed with MIMOSA 8 chips.

Firstly, the mean residual offset of the test comparators for 9 chips have been measured. For this purpose, the threshold voltage is set to a value which makes the number of "1" on the output is the half of the total statistical output number computed by the scope. This threshold value is the proper offset of the current comparator because the probability that the output is "1" is equal to "0". The results are shown in Fig. 3.28. We notice that the fluctuation of offset between different comparators is very small. The maximum value is about 0.5 mV and the dispersion is only about 0.22 mV_{rms}.

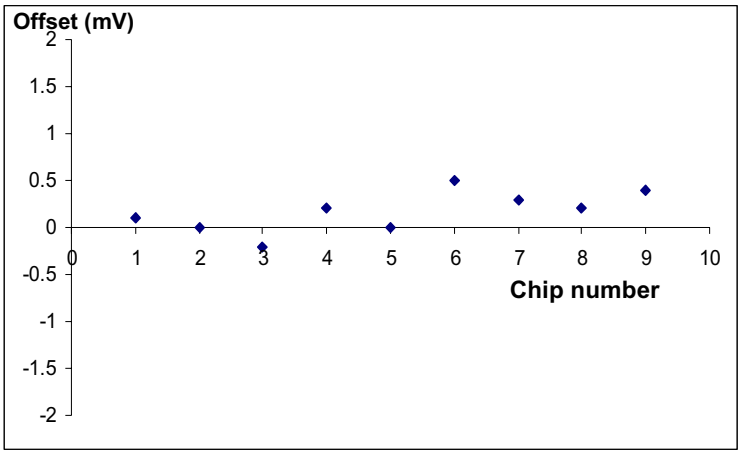


Figure 3.28: Offset dispersion of comparators on different MIMOSA 8 chips (f_{clock} = 100MHz).

Secondly, the most important character of the comparator, the transfer curve, is obtained by applying different threshold voltages. As described in Fig. 3.27, the two inputs of the tested comparator and the threshold voltage are shorted to a same common mode value. The threshold voltage varies from a very low value, which fixes the output to "1", and increases step by step up to a few mV. When the threshold voltage is close to the offset, the comparator triggers on its own Temporal Noise and changes randomly its state between "1" and "0". For one threshold voltage given, after normalization of the number of "1"s over the total number of cycles, the statistical value showing the probability of "1" for the threshold value used can be acquired. For the voltage range from -2mV to 2mV, these points of normalization compose the characteristic transfer curves of the discriminators.

Using the above method, test comparators of three different chips are measured and the results are shown in Fig. 3.29.

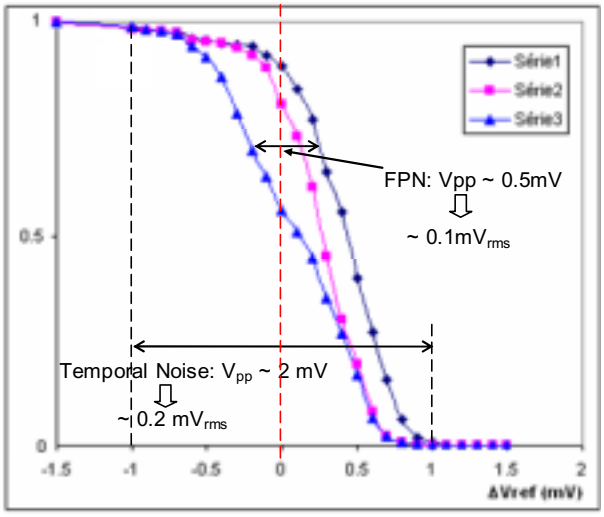


Figure 3.29: Normalised responses of test comparators on three different MIMOSA 8 chips (f_{clock} = 100MHz).

From Fig. 3.29, the FPN of the test comparators has a value around 0.1 mV_{rms} and the Temporal Noise is around 0.2 mV_{rms}.

Considering the results shown in Fig. 3.28 and Fig. 3.29, we can say that the Temporal Noise of the column level comparators is less than 0.5 mV_{rms} and its FPN is about 0.22 mV_{rms} (for 9 chips). Compared to the corresponding values of the analog output of the pixels, the comparators have small Temporal Noise but suffer more from the FPN. Therefore, for the global performance (pixel with discriminator), the Temporal Noise is dominated by the pixel level and the FPN is dominated by the column level comparators.

The transfer curves for the global performance are obtained by the data acquisition system. Since every pixel in the matrix has its digital output, one transfer curve is related to one pixel. The transfer curves give the noise performance for both pixels and column level comparators. Limited by the speed of the data acquisition system, the measurements are performed at 40 MHz. Similar to the method used for characterization of the test comparators in different chips, there is no input signal and the value of threshold voltage varied in a certain range. For each threshold value, a run of 1000 events is statistically enough to estimate the corresponding point of the transfer curve. The following diagram (Fig. 3.30) described the algorithm.

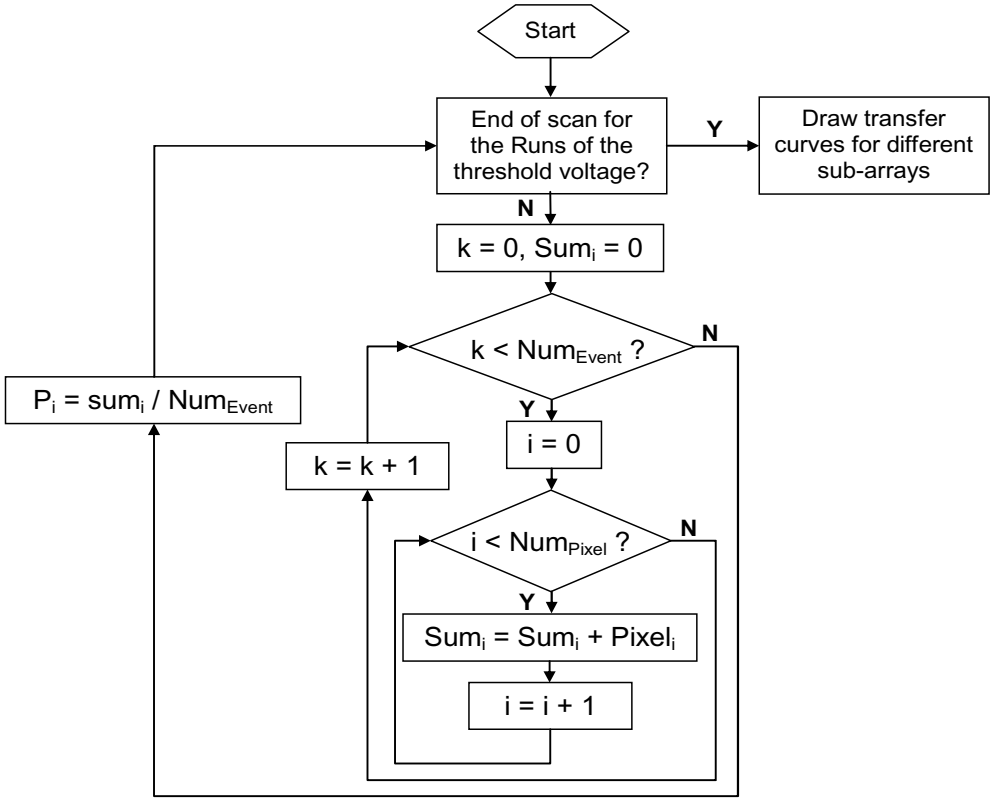


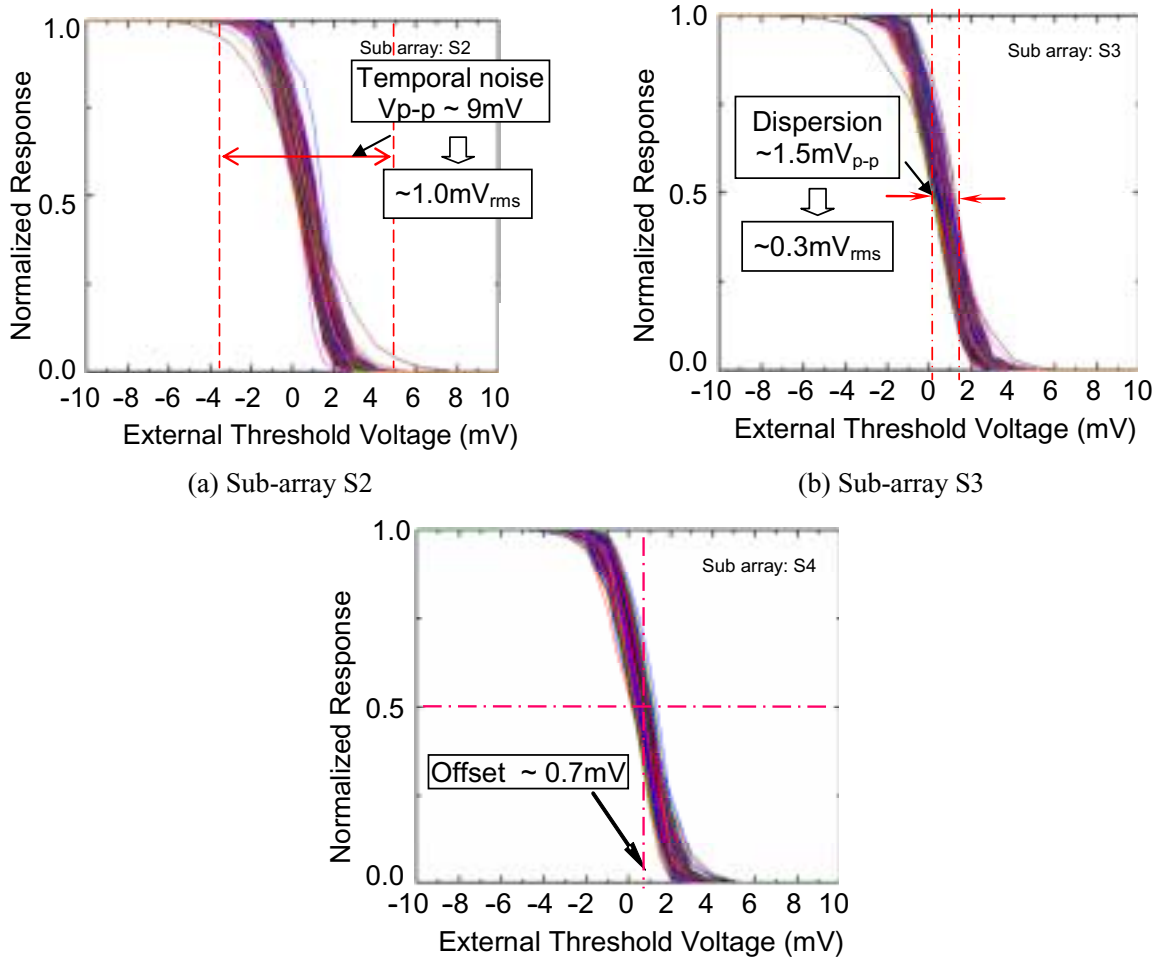
Figure 3.30: Diagram of the algorithm for obtaining transfer curve. Num_{Event} is the number of total events in the Run, Num_{pixel} is the number of the pixels in the whole matrix (all sub-arrays).

Two steps are needed in order to obtain the transfer curves for each sub-array:

- The first step is to calculate the normalized position of each pixel under a certain threshold voltage (P_i in Fig. 3.30). Since the pixel output value is "1" or "0", the value Sum_i is the total times that "1" appears on the output for i^{th} pixel (under the analyzed Run of the specified threshold voltage). Thus, the ratio between Sum_i and Num_{Event} gives the normalized position value of i^{th} pixel. For different threshold voltages, the values of P_i for all the pixels are obtained respectively.

- The next step is only to specify the pixels belonging to the same sub-array. Using the P_i values of each pixel, the transfer curves for the pixels of the sub-array are then obtained for the threshold voltage range applied.

Using the above algorithm, the transfer curves are obtained for both MIMOSA 8 and MIMOSA 16. The results are shown in Fig. 3.31 and Fig 3.32 respectively.

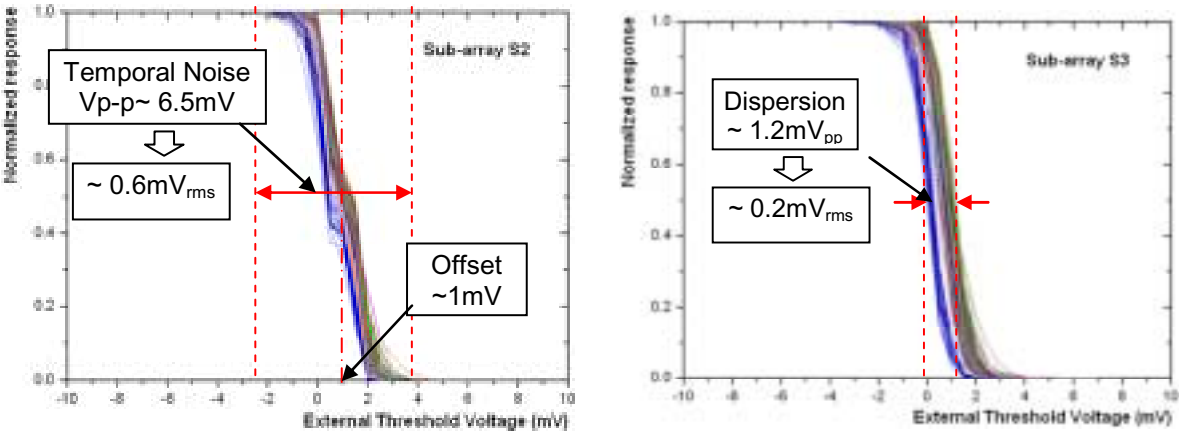


(a) Sub-array S2

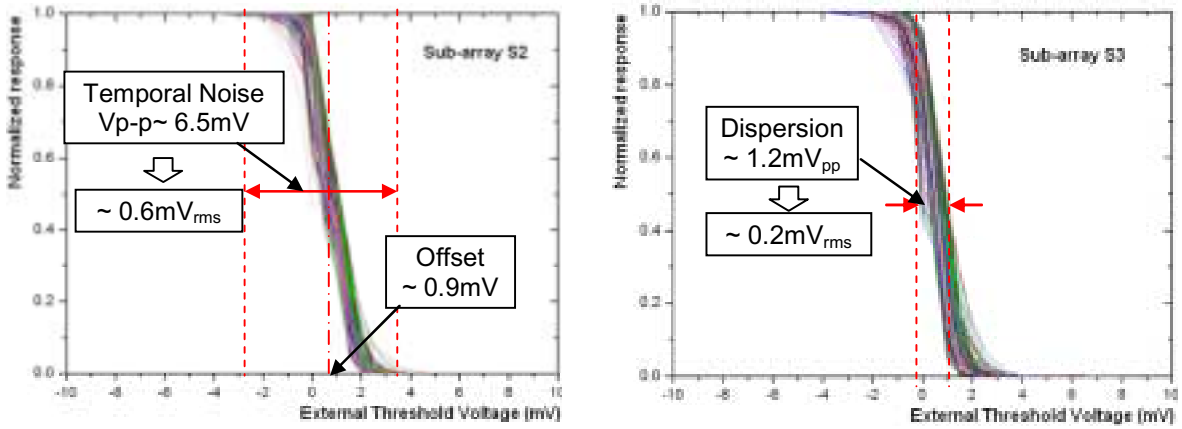
(b) Sub-array S3

(c) Sub-array S4

Figure 3.31: The transfer curves for the pixels of S2, S3 and S4 sub-arrays for MIMOSA 8.



(a) 14 µm epitaxial layer



(b) 20 µm epitaxial layer

Figure 3.32: The transfer curves for the pixels of S2 and S3 sub-arrays for MIMOSA 16.

Very similar results are obtained from the three sub-arrays of MIMOSA 8 and from the two sub-arrays of MIMOSA 16. For MIMOSA 8, the Temporal Noise is about 1.0 mV_{rms} and the FPN is about 0.3 mV_{rms}. These results are quite compatible with the results obtained on the analog outputs. The results for MIMOSA 16 are better: the Temporal Noise is 0.6 mV_{rms} and the FPN is only 0.2 mV_{rms}, also compatible with the results of the analog outputs. Another very important information indicated by the transfer curves is the minimum threshold voltage which prevents all the pixels to give a response from the noise. Its value determines the minimum signal level that the comparator can distinguish with very little influence of the noise sources. This value is about 6 mV for MIMOSA 8 and is only 4 mV for MIMOSA 16.

Moreover, the transfer curves of 5 different MIMOSA 16 chips (both 14 µm and 20 µm versions) are plotted together, as shown in Fig. 3.33 for sub-array S2.

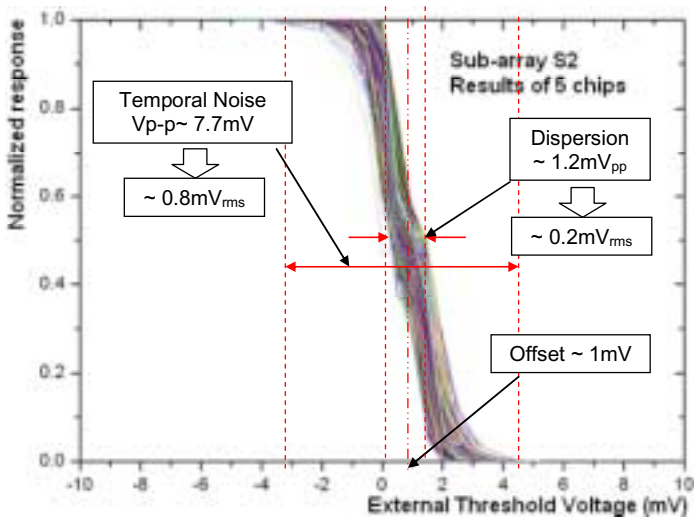
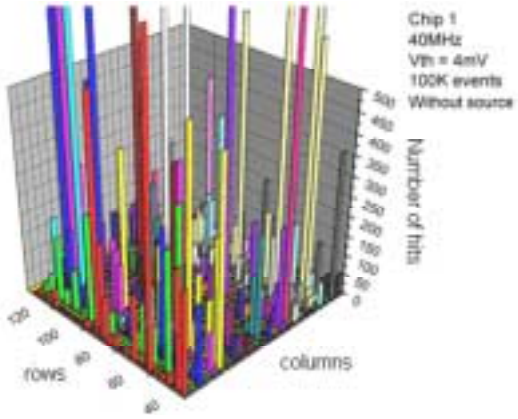


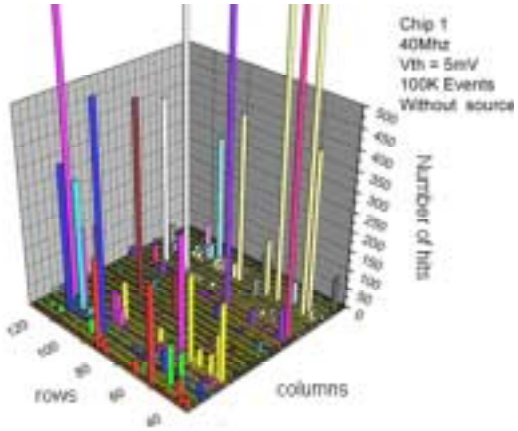
Figure 3.33: Transfer curves of sub-array S2 for 5 MIMOSA 16 chips

For 5 chips, there are totally 3840 pixels and 120 column level comparators. The results presented in Fig. 3.33 exhibit the remarkable noise performance of the column level comparators. The very small value of the Temporal Noise and the FPN indicates that the offset-cancellation structure applied inside the comparator works efficiently. Future design can be based on this architecture.

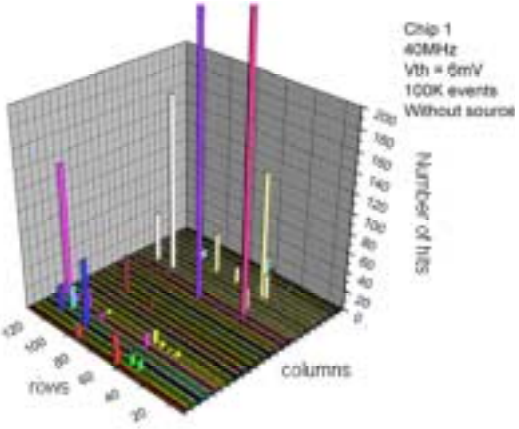
During measurements, the threshold value applied is very important because it influences the detection efficiency and the fake hit rate (the probability of hits caused by noise) [7]. When there is no input signal, the random responses of comparator could be either “1” or “0” and the “1” is obviously caused by noise in this moment. By raising the threshold voltage, the probability of “1” caused by noise can be reduced but the minimum input signal level will be greater on the contrast, reducing the sensibility of the comparator. Thus, the detection efficiency will be drawn down. On the other hand, if the threshold value is two small, the unwanted “1” caused by noise increases and leads to an increase of fake hit rate. So that compromise has to be found between the noise performance and the sensibility of the comparator. A preliminary study of different threshold values in function of the pixels responses is carried out. Three different values are tested, 4 mV, 5 mV and 6 mV, 100K events are taken for each threshold value. As an example, the responses of the whole matrix (Sub-arrays S2, S3 and S4) for MIMOSA 8 are shown in Fig. 3.34. We notice that with a threshold of 6mV, very few pixels are activated by the noise.



(a) Whole matrix (Sub-array S2, S3 and S4) with a threshold voltage applied on comparators: V_{th} = 4mV



(b) Whole matrix (Sub-array S2, S3 and S4) with a threshold voltage applied on comparators: V_{th} = 5mV



(c) Whole matrix (Sub-array S2, S3 and S4) with a threshold voltage applied on comparators: V_{th} = 6mV

Figure 3.34: MIMOSA 8: The influences of comparator threshold value in function of the responses of all the pixels.

3.4.2 Measurements with ^{55}Fe source

A. Measurement of the digital output

The analysis of the digital outputs (24 columns) with X-ray exposure leads to an array of 1s (hits) and 0s (no hits) as represented in Fig. 3.35, for a single frame. We notice that noise can induce non-physical isolated hits. For example, the cluster containing only one pixel in Fig. 3.35 is caused probably by noise. Increasing the comparator common threshold can limit the influence of noise. Cluster of 1s are present and may be used for hit position determination. In this case discrimination is equivalent to one-bit analog to digital conversion.

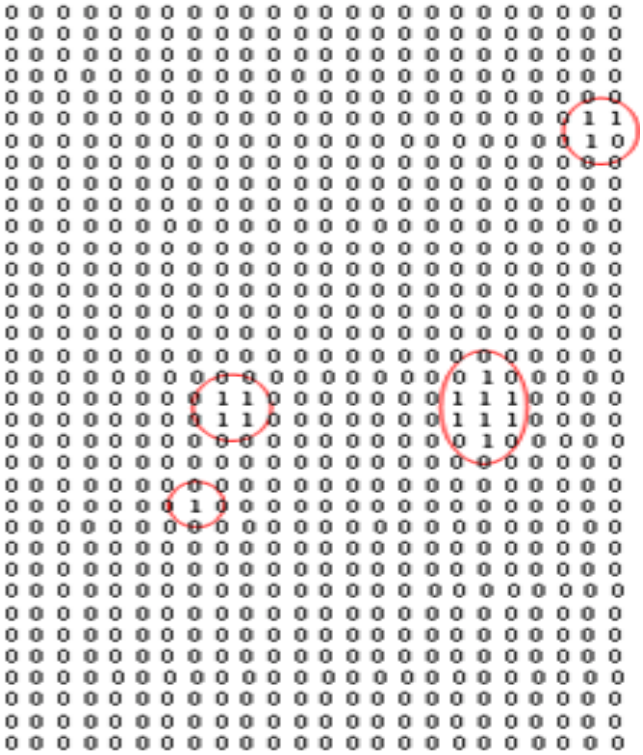
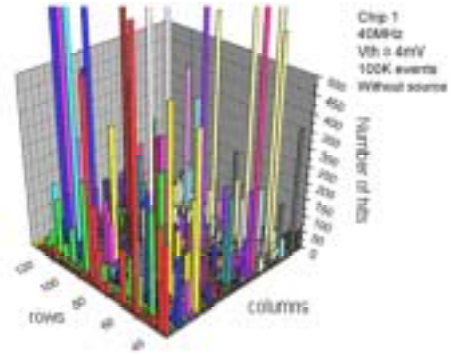


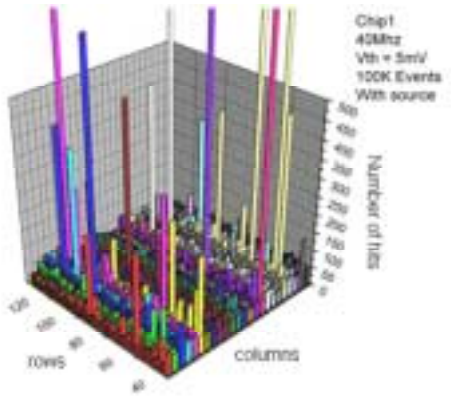
Figure 3.35: Example for a sub-array of pixels with the digital values of their outputs (1: pixel hit, 0: no hit). Clusters of hits distinctly appear. V_{th} = 4 mV and the main clock frequency is 40 MHz.

Furthermore, in order to carefully study the sensitivity of comparators under different conditions, we vary the threshold voltage of the comparators from 4 mV to 6 mV.

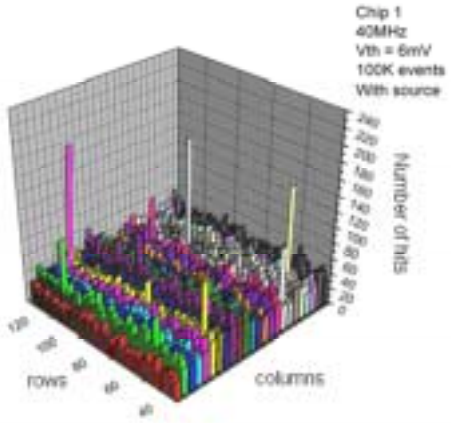
Fig. 3.36 gives the results obtained with MIMOSA 8 at 40 MHz. 100K events are taken for each threshold value. The binary response of each pixel is accumulated. We see that for V_{th} = 4mV, lots of pixels suffer from its noise and the real hit is merged inside the fake responses. However, when V_{th} = 6mV, the hits due to noise are efficiently cut off thanks to this high enough value. There is only a few numbers of pixels which are still high.



(a) V_{th} = 4mV



(b) V_{th} = 5mV



(c) V_{th} = 6mV

Figure 3.36: 3-D histogram (number of hits for each pixel) for the whole pixel array of MIMOSA 8. With the source, different external threshold voltages are applied.

B. Photon detection efficiency: comparison of analog and digital outputs

After the main test for both analog part and digital part, the photon detection efficiency is compared between them. The different external threshold voltages (Th_{real}) are applied to the comparators and the number of "1"s (a hit) is counted for each threshold voltage. For the analog data, an equivalent threshold values (Th_{eq}), corresponding to Th_{real} of comparator, is used in the analysis software. The number of hits having a value more important than Th_{eq} is also taken into account. The result of MIMOSA 8 is presented in Fig. 3.37. Very similar behaviors are obtained. Here, the effective threshold is the difference between the value of external voltage applied on the discriminator and the offset value (about 1 mV).

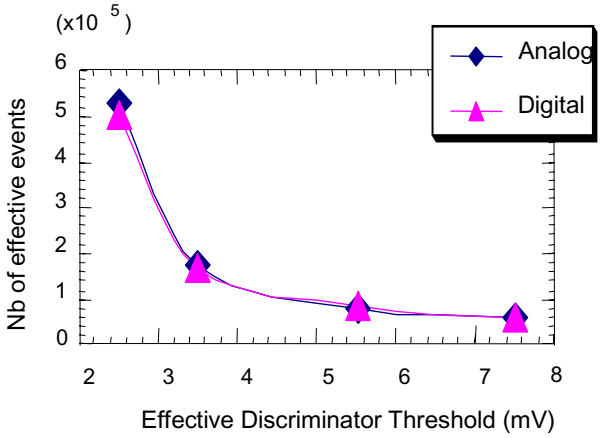


Figure 3.37: MIMOSA 8: Similar behavior between analog and digital part.

3.5 Fast neutron irradiation

In this work the fast neutron irradiations of chip were made at CERI (Orléans) in 2005. The neutron flux has a spectrum which extends up to 20 MeV. Integrated fluxes vary from 10^{11} up to 10^{13} neutrons/cm^2. Preliminary results of analog outputs were obtained at a low clocking frequency (1MHz).

In epitaxial processes, such as the one used for this MIMOSA 8 prototype (TSMC 0.25 µm), the degradation due to neutron irradiation is mostly due to atomic displacements in the thin epitaxial layer [8]. A number of defect points and defect clusters are created which reduces the lifetime of the minority carriers (the electrons in this case) and induces deep traps that affect the transport properties in the silicon. There are clearly two kinds of effects:

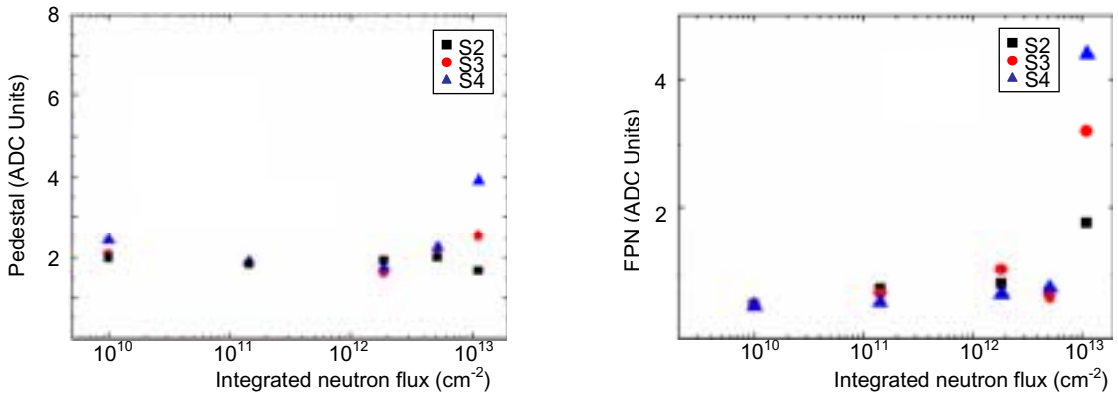
- Firstly, a continuous effect is the increase of the generation-recombination current in the depleted region of the p-n junction of charge detection diode. This current is

directly related to the electron lifetime and normally increases when the neutron exposure is enhanced.

- Secondly, the free drift length (or the diffusion length) and the lifetime of free electrons are reduced.

3.5.1 The effect on noise performance

The neutron irradiation effect on pedestal is shown in Fig. 3.38. The points at 10^{10} neutrons/cm^2 represent the reference values of a non-irradiated chip. Two effects can be observed. Firstly the pedestals mean value and the FPN remain remarkably stable for the integrated neutron flux from 10^{10} up to a few 10^{12} neutrons/cm^2. Secondly, the pedestal and FPN increase suddenly when the total integrated flux equals to 10^{13} neutrons/cm^2. The reason is that the integrated neutron flux of 10^{13} neutrons/cm^2 forces supplementary generation-recombination current, increasing the leakage current. Then, the growth of leakage current introduces a pedestal more important. In addition, since the defect clusters non-homogeneously distributed in the pixel induce a dispersion of leakage currents, the FPN increased with integrated neutron flux [9].



(a) (b)

Figure 3.38: (a) The pedestal and (b) its dispersion (FPN) as the function of integrated neutron flux for MIMOSA 8. Measured at 20°C with main clock frequency of 1 MHz.

The Temporal Noise of the three sub-arrays is weakly dependent on the total integrated flux (Fig. 3.39). There is no perceptible noise increase. Most of the noise is due to the electronic noise.

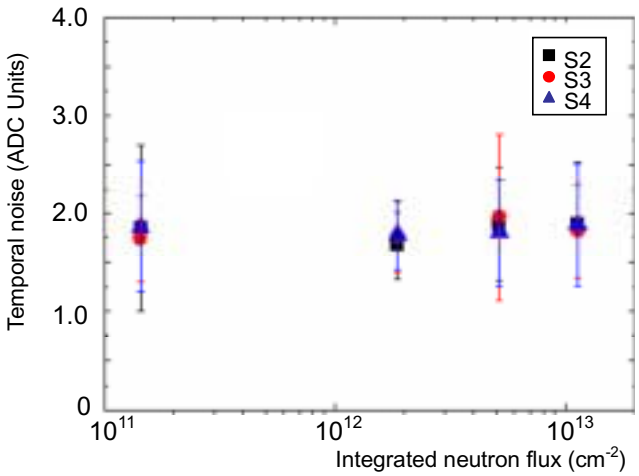


Figure 3.39: Temporal Noise as the function of integrated neutron flux for MIMOSA 8. Measured at 20°C with main clock frequency of 1 MHz.

3.5.2 The effects on charge collection

The effects on charge collection were tested by ^{55}Fe source. The calibration peaks of three sub-arrays remain stable with integrated flux. It is because charge generated by incident photon was in the depletion region that very close to the detection diode. Thus, the charge was collected very quickly and no charge loss was happened. On the contrary the position of charge collection peak continuously drops as the growth of integrated flux. As a result, the values of CCE decrease as function of integrated neutron flux (Fig. 3.40). The explanation is as follows: generally, the charge was generated inside epitaxial layer and was collected by thermal diffusion, changed by neutron irradiation. The increasing rate of generation-recombination current and the reduction of free electron's drift length introduce more energy loss.

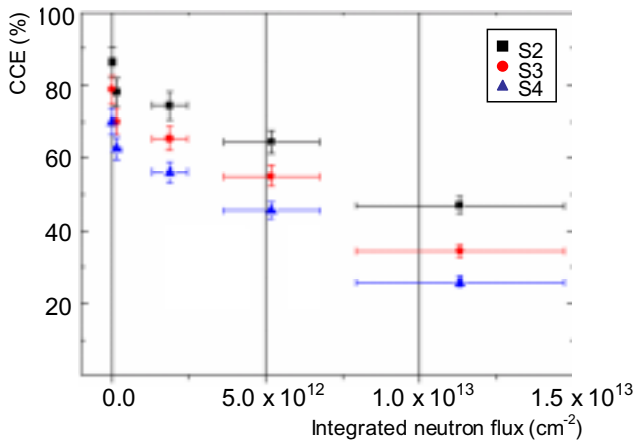


Figure 3.40: Charge collection efficiency in function of integrated neutron flux for the three sub-arrays for MIMOSA 8. Measured at 20°C with main clock frequency of 1 MHz. As a reference, the CCE of an un-irradiated chip is shown at the origin point.

3.6 Conclusion

The DC-coupling pixel architecture is realized in two different standard CMOS processes: MIMOSA 8 in TSMC 0.25 µm Digital and MIMOSA 16 in AMS 0.35 µm OPTO processes. For both prototypes, very encouraging results are obtained in laboratory tests with a ^{55}Fe X-ray source. Their main characters (at 100 MHz) of analog output are summarized in Tab. 3.11, Tab. 3.12 and Tab. 3.13 respectively. The measurement error for the noise is less than 2% and the error for the CCE values is negligible.

Sub array	CVF (µV/e-)	Input Referred Noise (ENC)	S/N (^{55}Fe)	Input Referred FPN (ENC)	Charge Collection Efficiency (5×5 cluster)
S2	66	11 e-	130	4 e-	80%
S3	60	12 e-	110	5 e-	85%
S4	52	14 e-	90	6 e-	92%

Table 3.11: Characteristics of MIMOSA 8.

Sub array	CVF (µV/e-)	Input Referred Noise (ENC)	S/N (^{55}Fe)	Input Referred FPN (ENC)	Charge Collection Efficiency (3×3 cluster)
S2	59	11 e-	152	2 e-	36%
S3	54	12 e-	127	2 e-	37%

Table 3.12: Characteristics of MIMOSA 16 with 14 µm epitaxial layer.

Sub array	CVF (µV/e-)	Input Referred Noise (ENC)	S/N (^{55}Fe)	Input Referred FPN (ENC)	Charge Collection Efficiency (3×3 cluster)
S2	59	11 e-	151	2 e-	25%
S3	52	13 e-	132	2 e-	29%

Table 3.13: Characteristics of MIMOSA 16 with 20 µm epitaxial layer.

For the results shown in Table 3.12 and 3.13, the error of noise measurements is less than 2%. Remarkable low noise is obtained in the measurements, showing the successfully implementation of the offset-cancellation architecture inside the pixel. The future pixel design will be based on this structure.

Tested with ^{55}Fe source, the CCE value of MIMOSA 16 is relatively low compared the results of MIMOSA 8. For the sub-array S2 with the charge detection diode of $2.4 \times 2.4\ \mu m^2$, the value is about 36% for 14 µm epitaxial version and about 27% for 20 µm epitaxial

version. The CCE is 92% for the pixels of sub-array S4 of MIMOSA 8. One probable reason is that the increasing thickness of epitaxial layer increases the lateral diffusion length of them. Therefore, more charge is lost and more widely the charge is distributed. The solution is to increase the size of charge detection diode. However, the noise level increases with the diode size. A compromise has to be found among the diode size, the noise performance and the CCE value.

Very small Temporal Noise and FPN have been observed for the digital outputs (the error of measurements is less than 2%), compatible with the results obtained by analog outputs. The results are summarized in Tab. 3.14.

	Temporal Noise (mV_{rms})	**FPN (mV_{rms})**
MIMOSA 8	1.0	0.3
MIMOSA 16	0.6	0.2

Table 3.14: Noise performance of the digital outputs for both MIMOSA 8 and MIMOSA 16 at 40 MHz.

The very small values of the FPN indicate that the offset-cancellation structure used inside comparator works efficiently. Forward design can be based on this architecture.

Tolerance to fast neutrons is also under study. The MIMOSA 8 is irradiated by integrated neutron flux from 10^{10} up to 10^{13} neutrons/cm². The noise performance remains stable with the neutron flux of 5×10^{12} neutrons/cm². These are only preliminary results and thorough radiation tests are needed to be carried on in the future.

A test with high energy particle beam is necessary to study the characteristic of the DC-coupling pixel in a real working environment. Very important parameters, such as MIPs detection efficiency and spatial resolution, are measured and presented in the next chapter.

Reference

[1] **Y. Li**, et al., "CMOS Sensors for the Vertex Detector of the future International Linear Collider," *Nucl. Instrum. Meth. A 572*, 2007, pp. 300-304.

[2] Y. Degerli, M. Besançon, A. Besson, G.Claus, G. Deptuch, W. Dulinski, N. Fourches, M. Goffe, A. Himmi, **Y. Li**, P. Lutz, F. Orsini and M. Szelezniak, "Performance of a fast binary readout CMOS active pixel sensor chip designed for charged particle detection," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 6, 2006, pp. 3949-3955.

[3] W. Dulinski, "VME Flash ADC Unit for the Strip Detector Readout (VFAS) – User's guide," LEPSI, 1999.

[4] G. Claus, et al., "A portable system for monolithic active pixel sensors characterization," *IEEE NSS Conference Record*, October 2004, Rome, Italy, pp. 1512.

[5] Y. Gornushkin, "Test results of monolithic active pixel sensors for charged particle tracking," *Nucl. Instrum. Meth. A 478*, Feb. 2002, pp. 311-315.

[6] G. Deptuch, et al., "Design and testing of monolithic active pixel sensors for charged particle tracking," *IEEE Trans. Nucl. Sci.*, vol. 49, April 2002, pp 601-610.

[7] A. Besson, G. Claus, C. Colledani, Y. Degerli, G. Deptuch, M. Deveaux, W. Dulinski, N. Fourches, M. Goffe, D. Grandjean, F. Guilloux, S. Heini, A. Himmi, C. Hu, K. Jaaskelainen, **Y. Li**, P. Lutz, F. Orsini, M. Pellicioli, E. Scopelliti, A. Shabetai, M. Szelezniak, I. Valin and Marc Winter, "A vertex detector for the international linear collider based on CMOS sensors," *Nucl. Instrum. Meth. A 568*, Nov. 2006, pp. 233-239.

[8] M. Deveaux, et al., "Neutron radiation hardness of monolithic active pixel sensors for charged particle tracking," *Nucl. Instrum. Meth. A 512*, Oct. 2003, pp. 71 -76.

[9] N. T Fourches, M. Besançon, **Y. Li**, P. Lutz, F. Orsini, "Fast neutron irradiation of Monolithic Active Pixel Sensors dedicated to particle detection," *Nucl. Instrum. Meth. A 576*, Jun. 2007, pp. 173 - 177.

[10] W. Dulinski, "Development of CMOS Sensors for future Hige Precision Position Sensitive Detectors", talk given at Vienna Conference on Instrumentation, Feb. 2007.

Chapter IV

MIMOSA 8 performances with high energy charged particles

4.1 Introduction

The calibration tests with a ^{55}Fe X-ray source show that the prototypes (MIMOSA 8 and MIMOSA 16) achieve very good performances [1] [2]:

- the noise is very small for both prototypes: Temporal Noise is only tens of electrons and the FPN, crucial to sparsification, is only several electrons;
- stable Charge Collection Efficiency (CCE) has been achieved in function of the main clock frequency;
- auto offset-cancellation column level comparator works efficiently: very similar behaviors comparing to the analog outputs are obtained;
- very high readout speed is achieved: about 12 μs/frame, approaching the requirement for the vertex detector of the future ILC.

Moreover, the performances of both prototypes are remarkably stable on a rather large main clock frequency range from about 20 MHz to 150 MHz. As an example, the input referred Temporal Noise of the pixel using the charge detection diode of 2.4×2.4 μm² is only 14 e- MIMOSA 8 and 11 e- for MIMOSA 16, respectively. The corresponding input referred FPN is 6 e- for MIMOSA 8 and 2 e- for MIMOSA 16. The CCE for MIMOSA 8 is about 92% (obtained with a 5×5 cluster) and the corresponding value for MIMOSA 16 is 36% (the 14 μm epitaxial layer version, obtained with a 3×3 cluster).

However, other important properties such as charged particles detection efficiency at the minimum ionization and spatial resolution can not be calibrated using the ^{55}Fe X-ray source. Tests with high energy beams are needed to study the responses of the prototypes and to calibrate their properties. Different from the photons generated by X-ray ^{55}Fe source, the high energy charged particles cross through the sensor and create the charge inside the epitaxial layer along its track. The charge created is then proportional to the thickness of the epitaxial layer. Obviously, the total charge created inside the epitaxial layer determines the input signal level to the sensors, influencing directly the performance of the prototypes. For the two prototypes, as more charge can be created inside the epitaxial layer of MIMOSA 16 than for MIMOSA 8, the input signal level of MIMOSA 16 is probably higher despite its lower CCE.

Two beam test campaigns have been made for MIMOSA 8. The first one was carried on at DESY (Hamburg, Germany) with a 5 GeV/c electrons beam in September 2005 [1] [2] and the second was carried on at CERN with a 180 GeV/c pions beam in September 2006. For MIMOSA 16, the beam tests are programmed in September 2007.

4.2 Experimental set-up of the beam tests

A high precision beam reference telescope developed by IPHC, Strasbourg [3] is used for the beam tests. The beam telescope is based on eight individual planes of high precision silicon microstrip detectors, which measure the trajectories of the impinging particles. The size of all detectors is the same: 12.8mm × 12.8mm × 300 µm (length × width × thickness). The simplified schematic diagram of test bench is shown in Fig. 4.1.

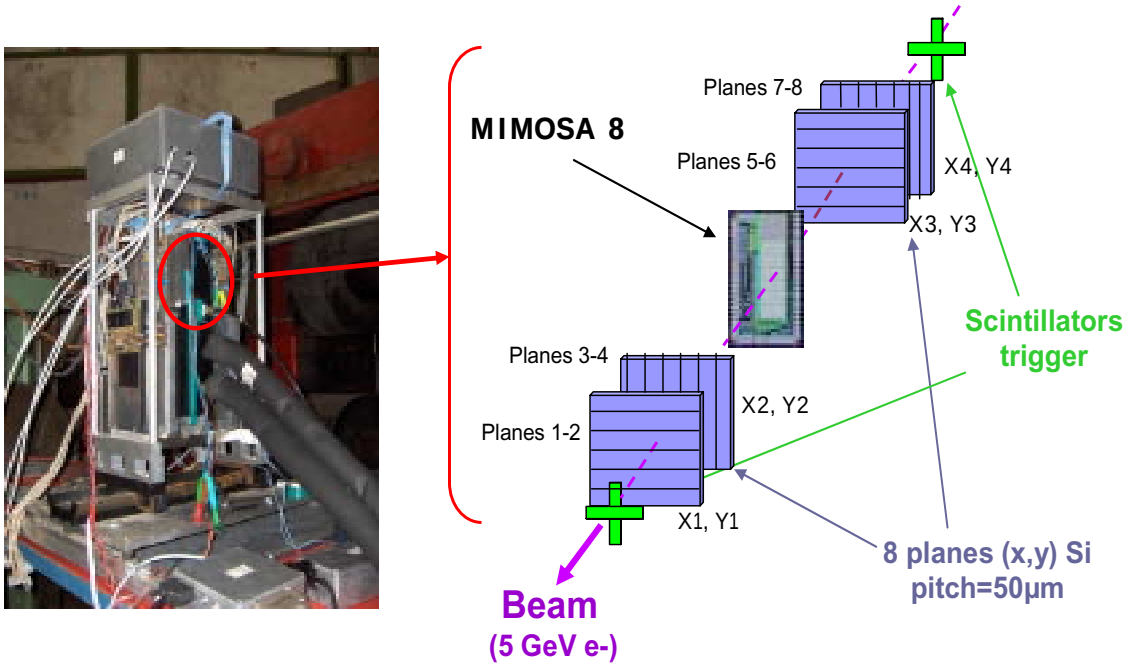


Figure 4.1: High precision telescope with eight reference planes.

As shown in Fig. 4.1, the reference detectors are arranged in four pairs allowing to measure particle positions in two orthogonal coordinates X-Y. For both sides, two of them are installed upstream and the other downstream according to the sensor under test. There are a total of 256 strips per plane and the strips are oriented in the order of XYYX. With a low noise VA2 electronics and digitized by 8-bit resolution VFAS card installed in the system, the analog information from the micro-strips is read out. Two scintillator pairs, 2×4 mm^2 for the first and 7×7 mm^2 for the second, are used as a fast trigger of the data acquisition system. When the prototypes are under test, the scintillation counters are mechanically aligned with the sensitive area of a chosen array of pixels. The typical trigger rate is within the range from a few hundred Hz to 1 KHz, but the acquisition system, based on the slow VME OS/9 processor, could accept only part of these triggers.

The analog signals are digitized in 12-bit ADC module and the raw experimental data are on the hard disk of the computer under LINUX connected to the OS/9 processor in the VME card by an Ethernet cable link. The data acquisition rate is limited by the time needed for Ethernet cable connection and for writing data on the disk. Each data acquisition is followed by a single reset cycle and the new data acquisition starts immediately after the acceptation of a trigger signal.

During the tests, the ambient temperature is about 20°C. The testing prototypes work at 40 MHz for all the tests.

Stand alone programs based on LabVIEW are implemented in the processor used for the data acquisition which allow on-line monitoring of the behaviors for the prototype under test and get a fast estimation of the test chip performances such as noise level, hit rate and signal amplitude.

An analysis C++ software using ROOT interface under LINUX environment, called "MAF" (Mimosa Analysis Framework) which is developed by IPHC since 2003, is used to analyze the experimental data [4].

4.3 The alignment process

As shown in Fig. 4.1, eight telescope planes with the chip under test consist of the test set-up. In order to simplifying the analysis procedure, the prototype is divided into two planes: the 9^{th} plane for digital outputs and the 10^{th} plane for analog outputs. Hence, the analyses for the digital and analog outputs are separated. For the analysis concerning the digital outputs, the 9^{th} plane and the eight telescope planes are used. On the contrary, the 10^{th} plane and the telescope planes are used for the analysis concerning the analog outputs.

For the analysis of both analog and digital outputs, the alignment of the different planes is mandatory. Although they are already aligned mechanically, more sophisticated alignments using statistical methods are needed in order to carry on very precise measurements. Two types of statistical alignments precede the analysis software in turn.

- The first one is tracker alignment, which align the tracks of the eight telescope planes and the two scintillator pairs. The tracker alignment proceeds automatically by the software after simply set down the bound distance in µm for the fit. Generally, the bound distance is set to a relatively large value in order to cover the area of the two scintillator pairs. Only the tracks found inside these areas are considered as effective tracks, which indicate that a particle has crossed through all the telescope planes. Thanks to the tracker alignment procedure, only the events containing good tracks (found in both scintillator pairs) are kept. Then, the raw experimental data are reconstructed. It is a pre-selection of hits and tracks based on the required signal to noise ratio for the telescope planes and the sensor plane. The events containing good data are prepared for the following procedure.

- The second one is the chip alignment itself which aligns the sensor plane under analysis (the 9^{th} or the 10^{th}, respectively) to the telescope planes (the plane (x, y) in Fig. 4.1).

For the second alignment, a coordinate system proposed in reference [4] is used. It is shown in Fig. 4.2 below.

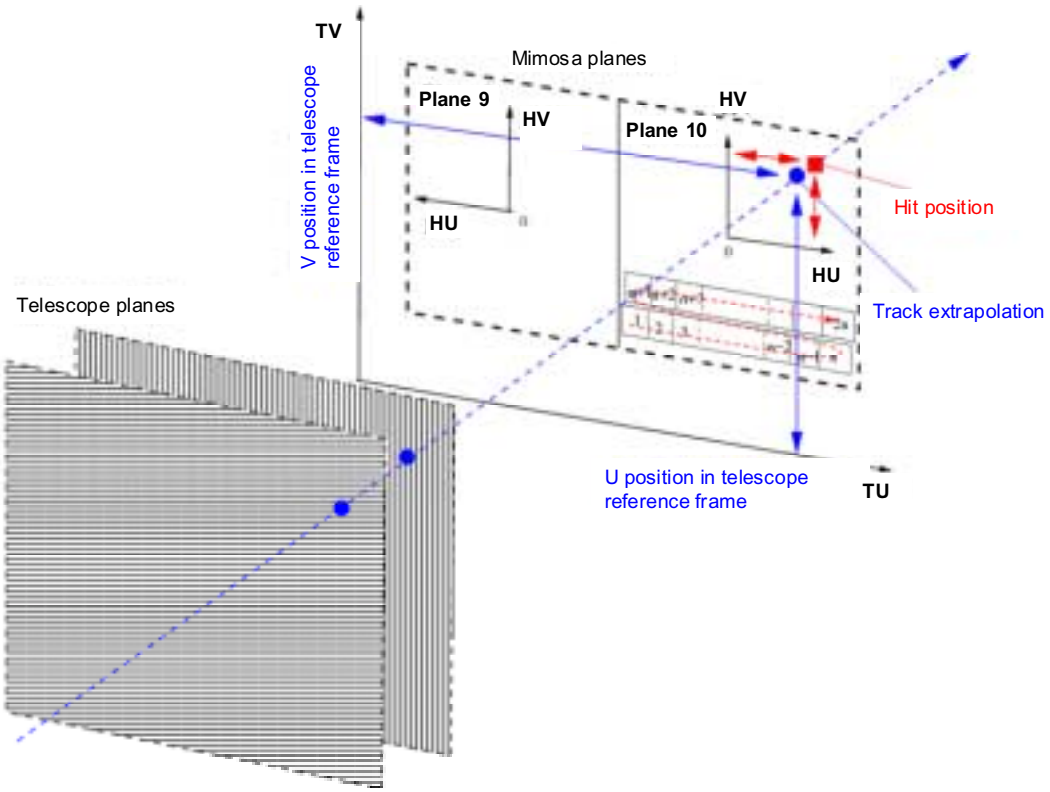


Figure 4.2: MIMOSA alignment: readout example and corresponding reference frames.

There are two reference frames in this system:

- The (HU, HV) sensor reference frame is related to the sensor plane under study (the MIMOSA prototype). The U and V direction in the frame are defined by the way that pixels are read and stored in the raw data. Data are stored row by row. These rows define the U direction and the columns define the V direction (Fig. 4.2). This reference frame is fixed. The original point (0, 0) of the frame is the center of the matrix.

- The (TU, TV) telescope reference frame is related to the track extrapolation.

The goal of the second alignment is to superimpose these two reference frames. It requires determining the shifts (U, V) and the possible rotation angles (th1, th2 and th3).

Several parameters are used for the sensor plane alignment in the off-line data analysis:

- TrackHitdist: cut on the distance between Hit position and track extrapolation position (µm). When the distance between a Hit and a track is smaller than this value, the Hit found by the sensor is considered to be associated to this track.

- GeoMatrix: windows for tracks. It defines the area inside which the tracks are compared to the Hits. Four available are listed as follows. 0: the area of the total telescope plane. 1: the area of the total pixel array. 2: the area of the sub-array under study. 3: the area of the sub-array under study without pixels on its boundary. The

former two values (0 and 1) are used to make alignment and the last two ones (2 and 3) are used for sensor performance analysis.

- S2N_seed: cut on the Signal-to-Noise Ratio (SNR) of the seed (central) pixel.

- S2N_neighbour: cut on the SNR of the 8 neighbour pixels: $\sum_{i=1}^{8} signal \Big/ \sqrt{\sum_{i=1}^{8} noise^2}$

The cuts S2N_seed and S2N_neighbour are used together in order to ensure to find out a good Hit[*]. Only the Hits fulfilling the following conditions (at the same time) are considered as good Hits and are kept in the analysis: the SNR of its central pixel should be more important than S2N_seed and the SNR for the neighbouring pixels should be more important than S2N_neighbour. The first condition guarantees that the Hit collects an important amount of charge (comparing to the noise level) and the second condition makes sure that the pixel with a high output value is not due to its offset.

The chip alignment is a recursive procedure. Several steps are needed. To begin, a large TrackHitdist value is firstly set down with the largest window size (GeoMatrix = 0). In our analysis, we usually start with a TrackHitdist value of 800 µm. The following example (Fig. 4.3) gives the distribution of the distance between the Hit position and the track position (for sub-array S2). Here, it is the 10th plane (plane for analog outputs) under study.

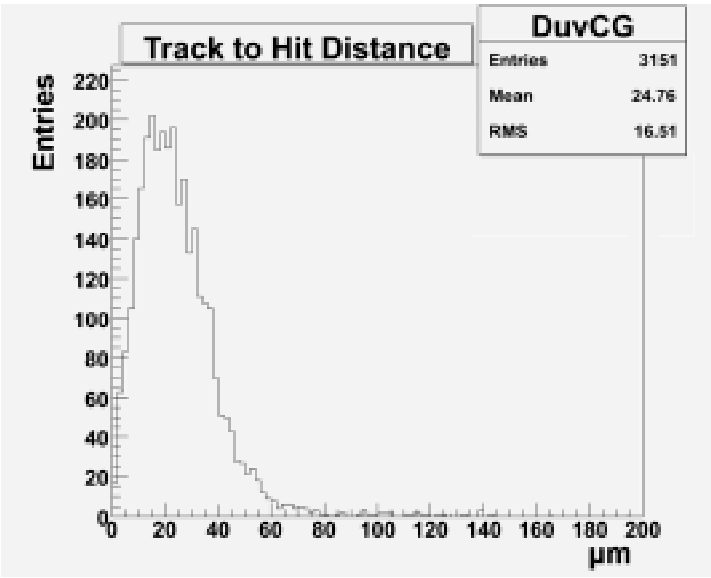
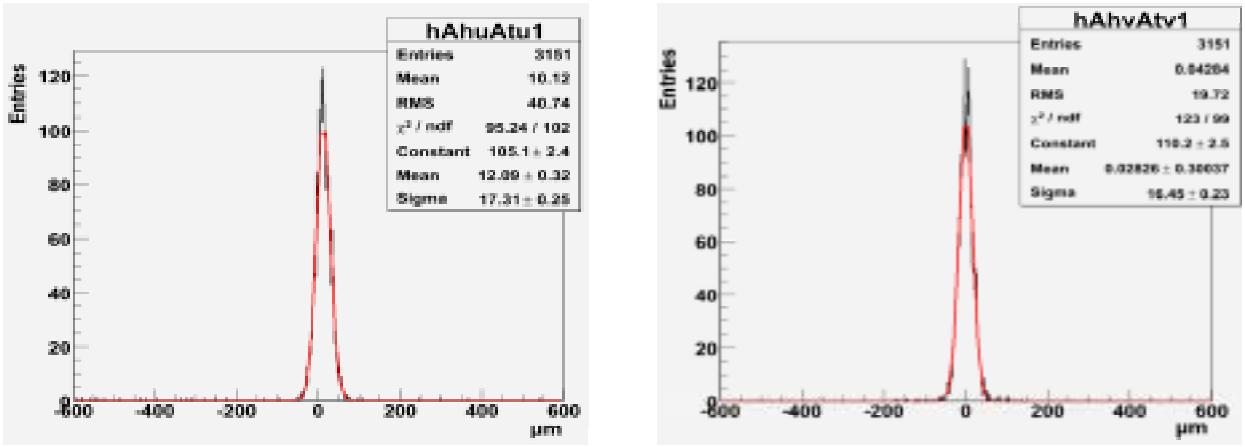


Figure 4.3: Distributions of track to Hit distance when TrackHitdist is set to 800 µm. Measured at DESY, 2005. 140 000 events are analyzed.

From Fig. 4.3, despite of using a large TrackHitdist value (800 µm) in the alignment procedure, the mean value of the distance between Hit and track position is about 25 µm. Thus, small value of TrackHitdist can be used.

The residuals (track position – Hit position, TU-HU and TV-HV) are also obtained, as shown in Fig. 4.4.

[*] For the digital outputs of MIMOSA 8, as the value is either "1" or "0", S2N_seed is set to 1 and S2N_neighbour is set to negative.



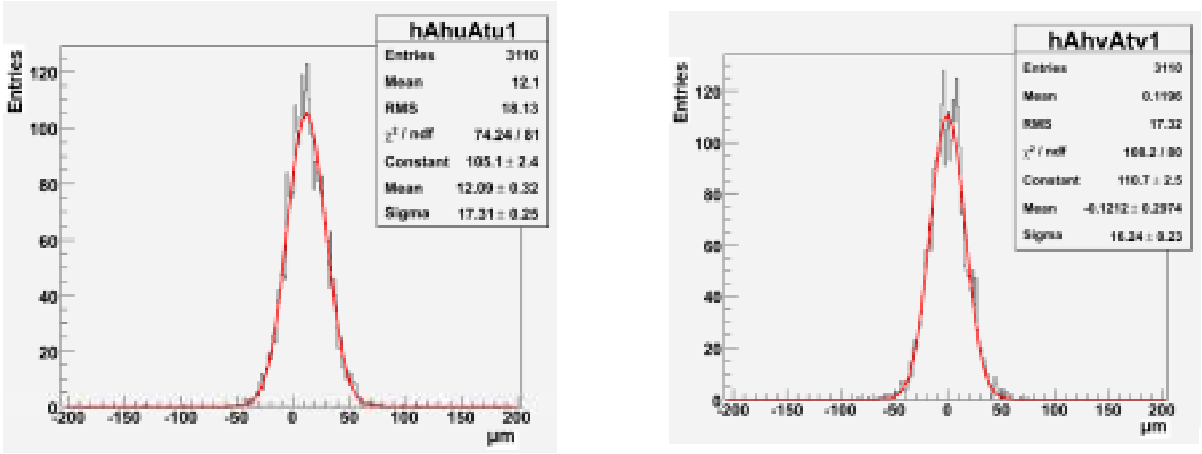
(a) Residuals of U direction from track position to Hit position.

(b) Residuals of V direction from track position to Hit position.

Figure 4.4: For the analog outputs, distributions of the residuals: difference between track and Hit position for U and V direction when TrackHitdist is set to 800 µm. Measured at DESY, 2005. 140 000 events are analyzed.

The residuals of the difference between track and Hit position are different for U and V direction. The mean value is 10.1 µm with a rms value of 40.7 µm for TU-HU and is –about 0.04 µm with a rms value of 19.7 for TV-HV. This difference comes from the lever arm which is different for U and V directions. While there are 128 rows for V direction, there are only 8 columns available for U direction, rather limit to obtain a good residual.

Decreasing the TrackHitdist value can improve the precision of the alignment. The residuals of the difference between track and Hit position when TrackHitdist is equal to 200 µm are given in Fig. 4.5 below.



(a) Residuals of U direction from track position to Hit position.

(b) Residuals of V direction from track position to Hit position.

Figure 4.5: For analog outputs, distributions of the residuals: difference between track and Hit position for U and V direction when TrackHitdist is set to 200 µm. Measured at DESY, 2005. 140 000 events are analyzed.

Comparing Fig. 4.4 and 4.5, in decreasing TrackHitdist value from 800 µm to 200 µm, we find that the residuals of U directions are improved (rms values): the mean value is 12.1µm

with a residual of 18.1 µm. For the V direction, only slight improvement is observed. The same distribution of the distance between track and Hit position (Fig. 4.3) is found with the TrackHitdist of 200 µm. Thus, the residuals are smaller than the mean value of the distance between the track and Hit positions.

Another useful information is the number of good Hits per event that we can find under the actual windows for tracks (GeoMatrix is set to 0 for the former examples). The result is shown in Fig. 4.6 below.

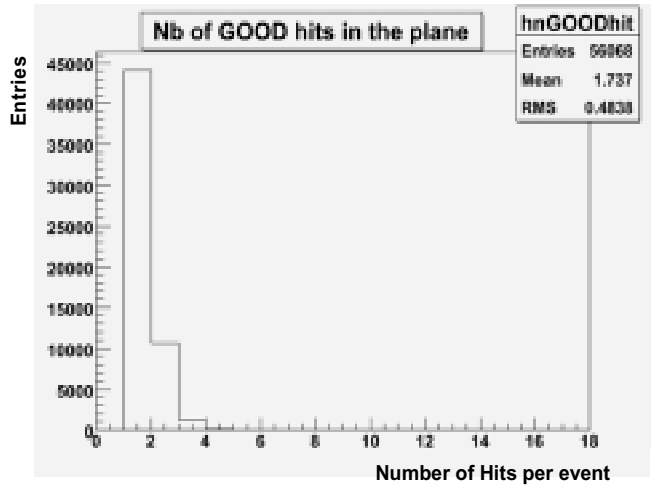


Figure 4.6: Distributions of number of good Hits for the 10th plane (analog outputs) under study. The total event number is 140000. Measured at DESY in 2005.

For the above example, there are 56068 good Hits are found for the total 140000 events analyzed. The most probable case is one Hit per event (about 78%). The probability of two Hits per event is about 20%. The case that one event contains more than two Hits is rather rare.

It seems that a large TrackHitdist value is more appropriate for the following analysis because there is that only one Hit in the event for about 80% of the case. However, it is not useful to have a too large value since the area of the pixel array for analog outputs is rather small (only 8 columns with a width of 25 µm per column). The total area under study is:

$$(\text{Number of pixel per sub-array}) \times (\text{Pixel area}) \times (\text{Number of sub-arrays})$$

Thus, for one sub-array of MIMOSA 8, the total area is:

$$32 \times 8 \times 25\ \mu m \times 25\ \mu m \times 3 = 480\,000\ \mu m^2$$

When TrackHitdist cut is equal to 800 µm, the surface cover by one Hit is:

$$800\ \mu m \times 2 \times 8 \times 25\ \mu m = 320\,000\ \mu m^2$$

Fig. 4.7 indicates the corresponding area inside the whole pixel matrix.

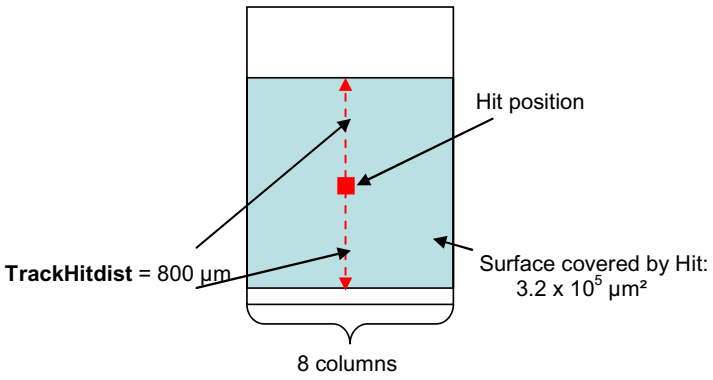
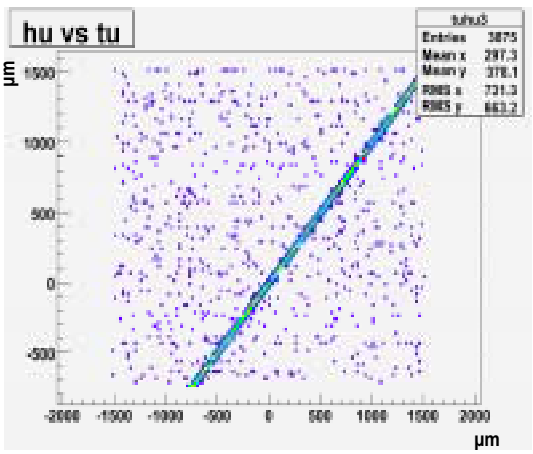


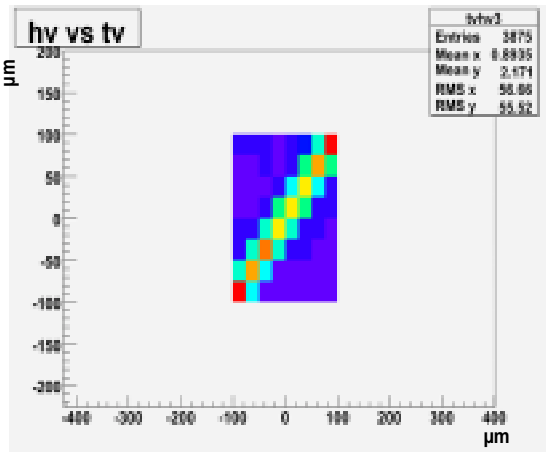
Figure 4.7 The area occupied by one Hit when TrackHitdist is set to 800 µm.

Compared to the total sensor surface, it is about 67% is covered under this condition. Taking into account that there are two Hits per event in about 20% of the case, the probability that the track is associated to another Hit becomes non-negligible, about 13%. When TrackHitdist cut is set to 200 µm, a Hit covers about 16% of the total surface. It is an acceptable value considering the multiple scattering effect which makes the reconstructed track far away from its real trajectory.

Since the errors coming from bad associations and multiple scattering are not taken into account, the χ^2 of the alignment procedure is usually far above 1 so that a value below 100 is good enough. It is possible to check the alignment thanks to the option of the analysis software. It shows correlations between Hit positions in the (HU, HV) sensor reference frame and track position in the (TU, TV) telescope reference frame. If the alignment is correct, the correlation is close to a line like y = x. Fig. 4.8 gives an example of the correlations between track and Hit position when the alignment is correct. More details concerned the chip alignment are available in [4]. When the alignment procedure is finished with favorable results, the analyses of the prototype performance can begin.



(a) Correlation between Hit and track positions in U direction.



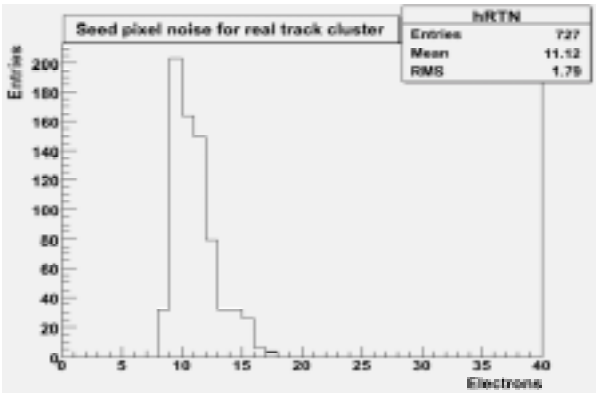
(b) Correlation between Hit and track positions in V direction.

Figure 4.8: Correlation between Hit and track positions. Total event number is 140 000. TrackHitdist and GeoMatrix are set to 200 µm and 1, respectively. Measured at DESY in 2005. 140 000 events are analyzed.

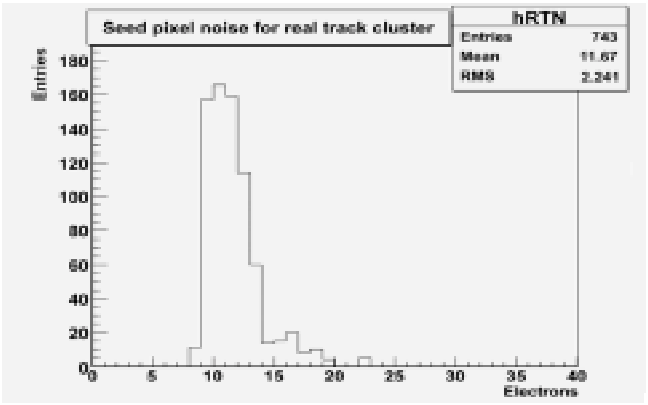
4.4 Tests of analog outputs

4.4.1 Noise measure

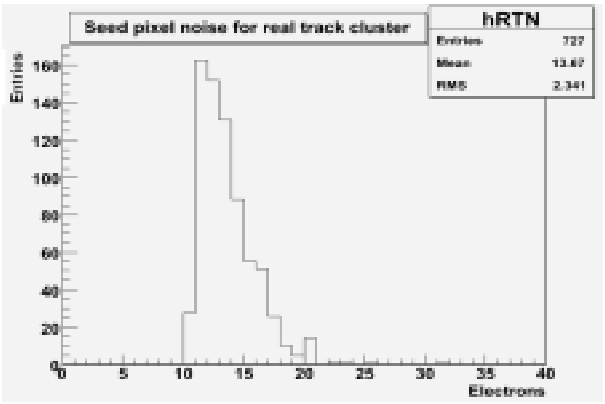
The Temporal Noise is measured firstly under the test beam environment. With the knowledge of CVF obtained in laboratory's measurement, the equivalent input referred noise of the seed pixel (central pixel) measured at DESY in 2005 is given in Fig. 4.9 below.



(a) Sub-array S2.



(b) Sub-array S3.



(c) Sub-array S4.

Figure 4.9: The distribution of the Temporal Noise for central pixel of MIMOSA 8. Measured at DESY, 2005. 140 000 events are analyzed.

The mean values of the Temporal Noise for the different sub-arrays are: 11 electrons for sub-array S2, 11 electrons for sub-array S3 and 13 electrons for sub-array S4. These values are compatible with laboratory's measurements (Tab. 3.8 in chapter 3). Similar results are found for the beam test at CERN in 2006. The results are summarized in Tab. 4.1 below.

Temporal Noise of analog outputs		
Sub Array	**Input referred Noise (ENC) in e-**	
	Results at DESY	**Results at CERN**
S2	11±2	10±2
S3	12±2	11±2
S4	14±2	12±2

Table 4.1: Input referred noise of the analog outputs for MIMOSA 8 for the beam tests at DESY (2005) and at CERN (2006).

The small values of the Temporal Noise show that the CDS structure integrated inside pixels works efficiently under the environment of real high energy charged particle.

4.4.2 Charge collection performance and detection efficiency

A. Charge collection performance

The average charge of MIPs generated inside epitaxial layer is proportional to the thickness of the layer: 80 e-h/μm. Therefore, the total average charge available for the sensor to collect is given by:

$$C_{Tol} = L_{epi} \times 80e^{-} h / \mu m, \qquad (4.1)$$

where L_{epi} is the thickness of epitaxial layer.

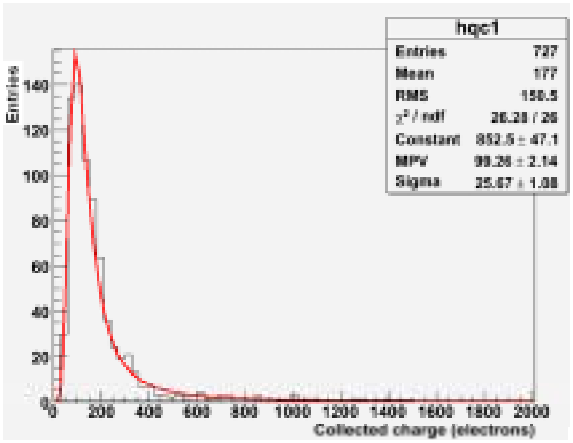
MIMOSA 8 is fabricated in TSMC 0.25 μm CMOS digital process with about 8 μm epitaxial layer thickness. Considering the different CCE (Charge Collection Efficiency) of the three sub-arrays, in principle the total charge values for the different sub-arrays are (the estimated error is about 3%):

Sub-array S2: $C_{Tot,S2} = 8 \times 80 \times 80\% = 512e^{-}$

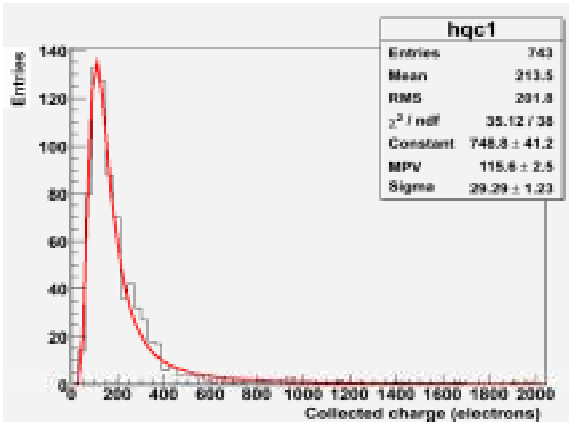
Sub-array S3: $C_{Tot,S3} = 8 \times 80 \times 85\% = 544e^{-}$

Sub-array S4: $C_{Tot,S4} = 8 \times 80 \times 92\% = 588e^{-}$

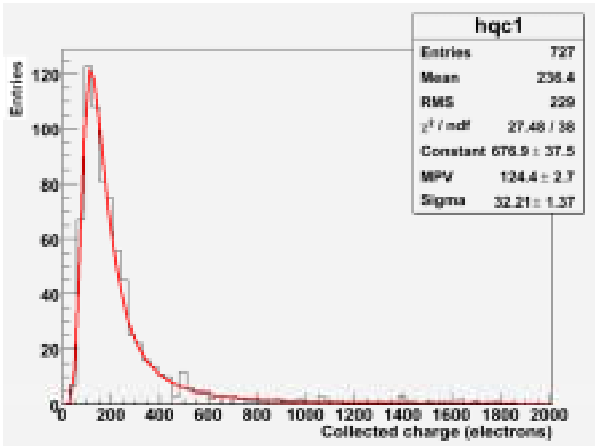
By setting the analysis cuts S2N_seed to 4 and S2N_neighbour to 2, the charge collected by the central pixel for good Hits are obtained. The results of the beam test at DESY are given in Fig. 4.10 below, for sub-array S2, S3 and S4 respectively.



(a) Charge collected by seed pixels for sub-array S2.



(b) Charge collected by seed pixels for sub-array S3.

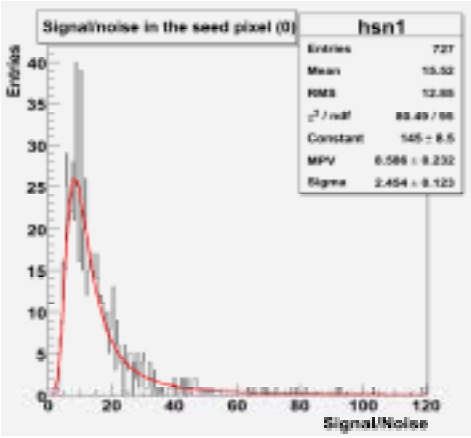


(c) Charge collected by seed pixels for sub-array S4.

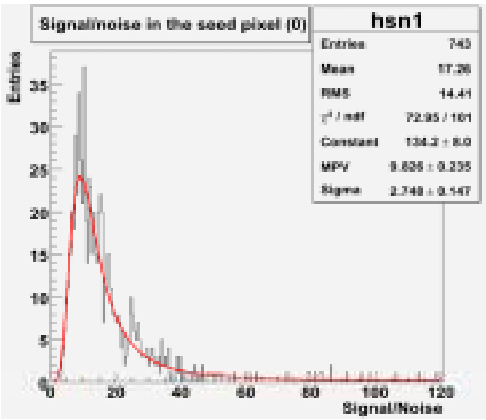
Figure 4.10: The charge collected by seed pixel (central pixel) of a cluster for the good Hits under the environment of 5 GeV electron beams at DESY. 140 000 events are analyzed.

The MPVs (Most Probable Value) of these results give us the charge collected by the seed pixel in the most general case. It is about 99 electrons for the pixels in sub-array S2, which is equivalent to 19.3% of the total available charge. For the sub-arrays S3 and S4, the corresponding values are 115 electrons (21.1% of the total charge) and 124 electrons (21.1% of the total charge), respectively. The charge collected on the seed pixel is rather small (hundreds of electrons in average, which is equivalent to only tens of μV) due to the limited

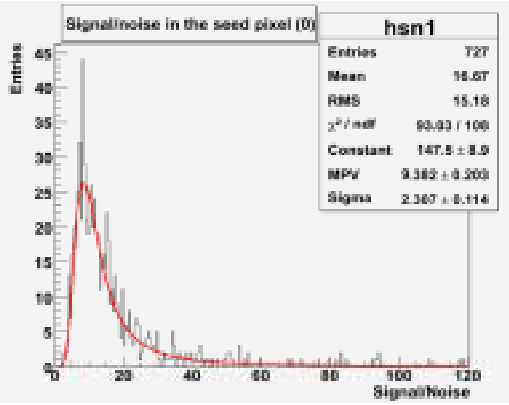
free charge created by the MIPs. Thus, the input signal level to the pixel and to the column level readout circuit is also limited. However, thanks to the CDS structure successfully integrated inside the pixels, the equivalent input noise level is very small (tens of electrons, Fig. 4.9), which makes the pixel and the column level readout circuit correctly working despite the small input signal. The SNR (Signal-to-Noise Ratio) of the seed pixels for different sub-arrays are also obtained. The results for the tests made at DESY are given in Fig. 4.11 below.



(a) SNR of seed pixel for sub-array S2.



(b) SNR of seed pixel for sub-array S3.



(c) SNR of seed pixel for sub-array S4.

Figure 4.11: The SNR of the seed pixels for the three sub-arrays. Tested at DESY (2005) with 5 GeV electron beams. 140 000 events are analyzed.

Despite very low level of noise (tens of electrons, Tab. 4.1), the SNR values are limited due to the small input signal level. It is about 8.6 for sub-array S2, 9.8 for sub-array S3 and 9.4 for sub-array S4. Moreover, the SNR of the seed pixels are measured as a function of the column level discriminator threshold value. Remarkable stable results are obtained, as shown in Fig. 4.12 below.

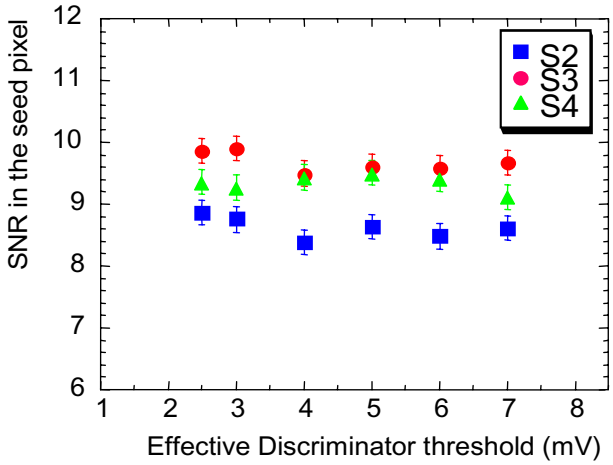
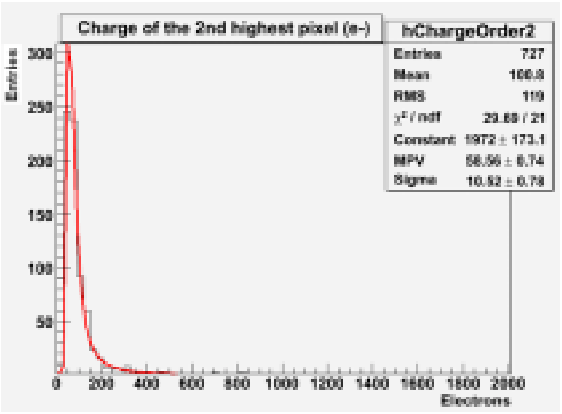


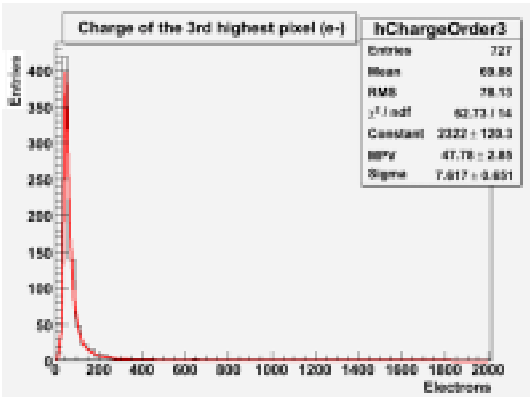
Figure 4.12: The SNR of the seed pixels for different sub-arrays. Tested at DESY (2005) with 5 GeV electron beams. 140 000 events are analyzed.

An important issue of high speed mixed signal circuit is the digital-analog coupling. The performance of analog part in a mixed circuit is generally degraded by digital signals. In MIMOSA 8, the decrease of the discriminator threshold value will increase the activities of the digital components. In Fig. 4.12, the stable SNR values under different threshold values of the column level discriminator indicates that the analog-digital coupling is well controlled for the prototype MIMOSA 8.

The charge distribution in the pixels of a cluster for the good Hits is also study. For sub-array S2, the distribution of the collected charge for the second and third highest pixel is given below.



(a) The charge collected on the second highest pixel.

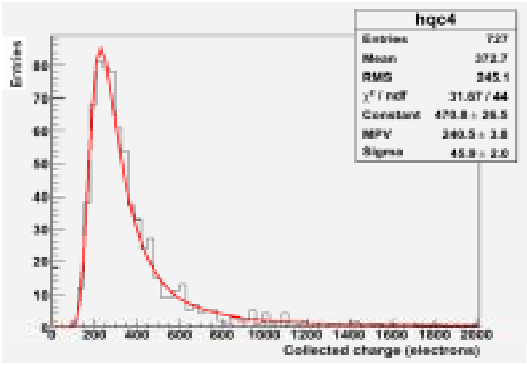


(b) The charge collected on the third highest pixel.

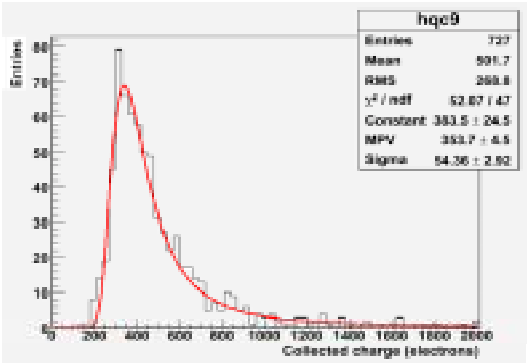
Figure 4.13: The charge collected on the pixels around the seed pixel.
Tested at DESY (2005) with 5 GeV electron beams. 140 000 events are analyzed.

From Fig. 4.13, the charge collected on the second and the third highest pixels around the seed pixel are 58 and 47 electrons respectively. This result shows that these pixels collect almost equal charges. By the same way of analysis, the charge collected on the fourth and the fifth highest pixel are 23 and 17 electrons respectively, less than 25% of the charge collected on the seed pixel. The above results indicates that the charge collected is concentrated in several pixels, corresponding to a rather small area.

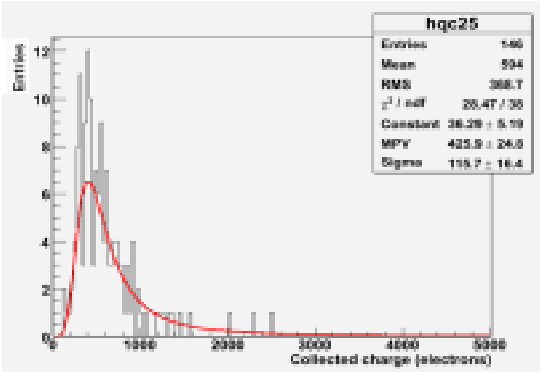
The variation of the collected charge as a function of the cluster size is also study. As an example, the results of the charge collected by the cluster of 4 (2×2), 9 (3×3) and 25 (5×5) pixels for sub-array S2 are given below:



(a) 4 pixels of 2×2 cluster.



(b) 9 pixels of 3×3 cluster.



(c) 25 pixels of 5×5 cluster.

Figure 4.14: For sub-array S2, distribution of the collected charge as a function of the number of pixels in the cluster. Tested at DESY (2005) with 5 GeV electron beams. 140 000 events are analyzed.

From Fig. 4.14, a 2×2 cluster collects 240 electrons and a 3×3 cluster collect 353 electrons. The charge collected by a 5×5 cluster is about 425 electrons, 14% higher than the charge

collected with 3×3 cluster. With 16 pixels in addition in a 5×5 cluster, no significant gain in the collected charge is made. Moreover, the small number of entries (146) indicates that only a few of good Hits of 5×5 cluster is found (the clusters meet the conditions that S2N_seed is no less than 4 and S2N_neighbour is no less than 2). The reason is that there are only 8 columns for the analogue outputs and a cluster bigger than 3×3 has really no sense in this condition. Therefore, we use 3×3 clusters in our following studies.

The results of charge collection performance for the beam tests at DESY are summarized in Tab. 4.2 below:

Sub Array	Charge in 3×3 cluster (e-)	SNR of seed pixel
S2	353±4	8.6±0.2
S3	392±4	9.8±0.1
S4	433±5	9.4±0.2

Table 4.2: Charge collection performance of the analog outputs for MIMOSA 8 measured at DESY, 2005.

The corresponding results obtained at CERN in 2006 are given in Tab. 4.3.

Sub Array	Charge in 3×3 cluster (e-)	SNR of seed pixel
S2	306±4	6.2±0.5
S3	328±4	6.3±0.2
S4	381±5	6.6±0.6

Table 4.3: Charge collection performance of the analog outputs for MIMOSA 8 measured at CERN, 2006.

Although the noise levels measured at DESY and at CERN are very similar (Tab. 4.1), the SNR of the seed pixels measured at CERN is slightly lower than the results obtained at DESY. One of the probable reason is that one column of analog output is used for the synchronization signal of the data acquisition system at CERN. The remaining seven columns are too limited to obtain a good alignment, which influences all the following analyses.

Moreover, the charge obtained after the analysis is lower than the charge expected. For example, only a charge of 425 electrons (using 5×5 cluster) is found for sub-array instead of 512 electrons, which is predicted by calculation. About 17% of difference is presented. Two reasons are possible.

- The first hypothesis is that 5×5 clusters do not collect all the charge. But, as shown in Chapter 3 (Fig. 3.17), we have seen that almost all the available charge is collected within a 5×5 cluster [5][6]. So that the low charge collected level is not due to this reason.

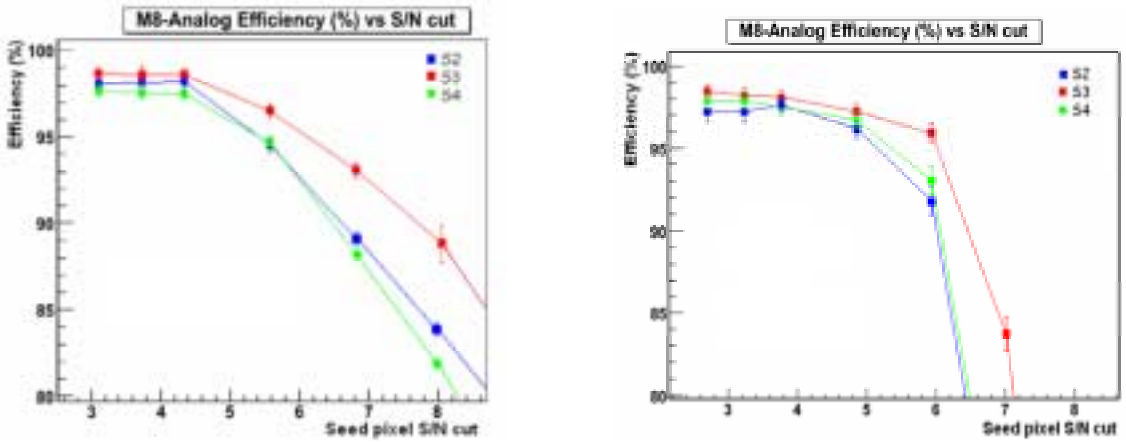
- The second is that the thickness of epitaxial layer is lower than the value indicated by the foundry. The real thickness is:

$$\frac{425}{80\times 80\%} \approx 6.6 \pm 0.4\ \mu m$$

Therefore, a CMOS process with higher thickness of epitaxial layer is needed.

B. Detection efficiency

The detection efficiency is defined as the proportion of events which have a reconstructed hit in MIMOSA8 close to the track extrapolation of the telescope. Obviously, two analysis cuts, TrackHitdist and S2N_seed, influence the final result. The TrackHitdist is set to 200 µm as explained in the former sections. For the value of S2N_seed, while a too small cut value suffers from the pixel noise, a too high cut value barricades pixels with small signal amplitude and therefore makes the detection efficiency decrease. Several cut values of seed pixel are tried during the analysis in order to find out the best compromise between noise level and detection efficiency. The results are given in Fig. 4.15 below. Encouraging results have been obtained. For the cut value on seed pixel varying from 3 to 4, the detection efficiencies are quite stable for all the three sub-arrays. For small cuts (around 3), the values are 98.0% for sub-array S2, 98.6% for sub-array S3 and 98.2% for sub-array S4 respectively. The detection efficiency begins to decrease when the seed pixel cut value reaches 4.



(a) Measurements of the beam tests at DESY, 2005.

(b) Measurements of the beam tests at CERN, 2006.

Figure 4.15: Analog outputs detection efficiency as the function of the Signal to Noise ratio cut value on the seed pixel.

Comparing the two measurements (at DESY and at CERN), we notice that for the cut values below 6, very similar results are obtained. The detection efficiency is around 95% (cut is equal to 5.6) for the measurements at DESY and is also about 95% for the measurements made at CERN. Moreover, the detection efficiencies remain remarkable stable when the cut is below 4 for both beam tests. The values are around 98% for all the three sub-arrays. Despite the cut values, the results of sub-array S3 with the pixel size of $1.7\times 1.7\ \mu m^2$ are better than the other sub-arrays. The best results are found with a cut equal to 3. The values are $98.6 \pm 0.3\%$ for the tests of DESY and $98.1 \pm 0.5\%$ for the tests of CERN. Although the

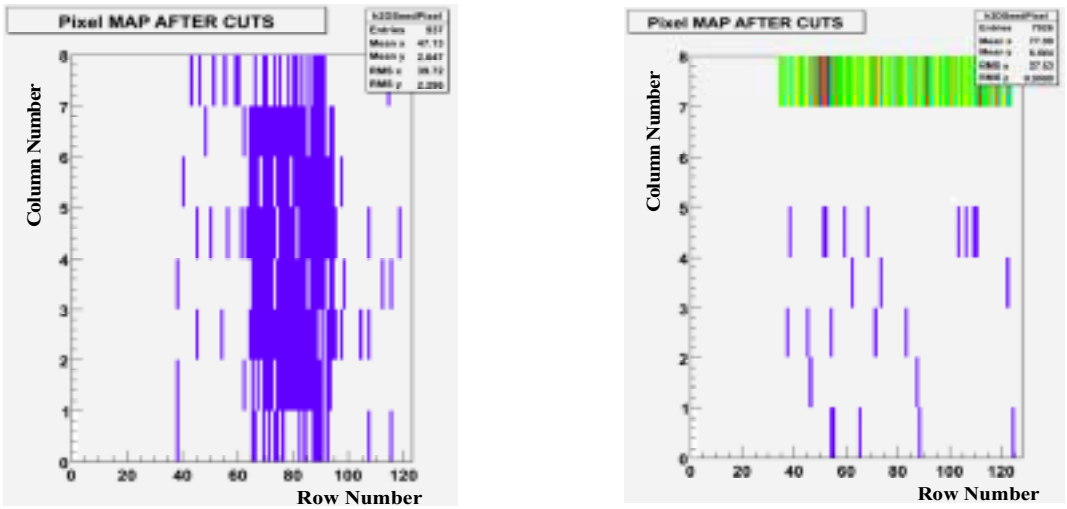
pixels of sub-array S4 collect more charge (Tab. 4.2 and 4.3) due to the largest size ($2.4\times2.4\ \mu m^2$), they suffer from more noise degrading the detection efficiency of S4.

For the beam tests made at DESY, the detection efficiency continuously decreases for the cut values more important than 4. With a cut value of 8, the detection efficiencies are about 84% for sub-array S2, 89% for sub-array S3 and 82% for sub-array S4.

For the beam tests made at CERN, the detection efficiency drops rapidly after the cut value reaches 6. The reason is the low SNR of the seed pixels. As the SNR of the seed pixels is equal to ~6.3, it is hard to find a good Hit which fulfils the condition that the S2N_seed is equal to 7. For the tests made at DESY, because the SNR of the seed pixel is bigger than 7, the sharp drop does not happen.

The above explanation can be easily understood by observing the pixel maps (Fig. 4.16: all reconstructed hits) and the track-without-Hits maps (Fig. 4.17) obtained at DESY and at CERN with the same cut value.

For the whole pixel matrix (sub-array S2, S3 and S4), the pixel map records the pixels which have ever been taken as the seed pixel of a good Hit during the analysis process. At the end of an analysis, the pixel map indicates the situation of good Hits found in the whole pixel matrix. With a cut value of 7, the pixel maps of the tests made at DESY and at CERN are given in Fig. 4.16.



(a) Tests made at DESY.

(b) Tests made at CERN*.

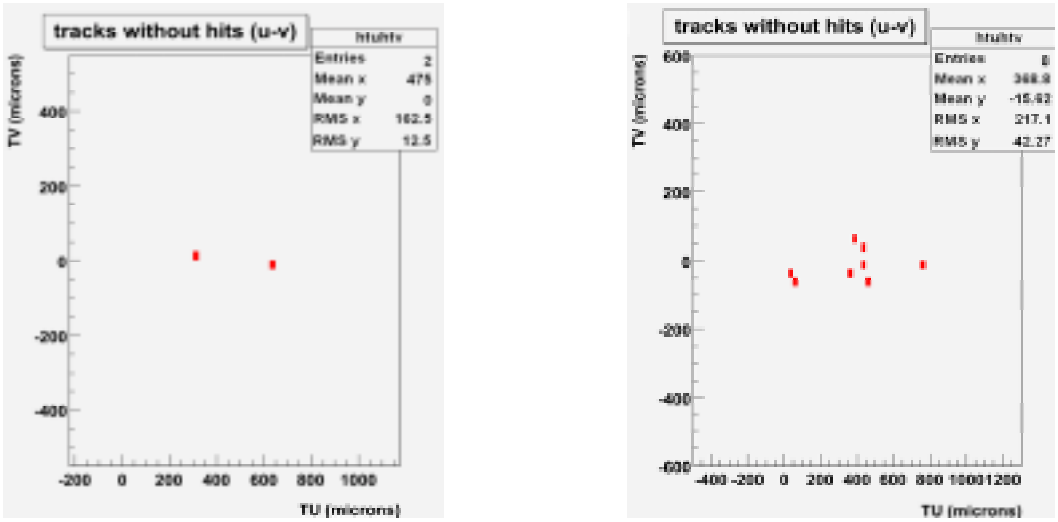
Figure 4.16: Pixel map with the cut of 7 obtained with the tests made at (a) DESY and (b) CERN.

Comparing the two pixel maps, only a few hits are found for the tests at CERN when a cut of 7 is applied. Hence, there are not enough statistics to analyze properly.

The track-without-Hits map records the tracks found in the sensor area which is not associated with a good Hit. These tracks are not detected by the sensor. The following examples

* For the tests made at CERN, the 8th column is used for synchronization purpose. It is not concerned in the analysis

(Fig. 4.17) give the track-without-Hits maps obtained at DESY and CERN when a cut of 7 is applied.



(a) Tests made at DESY. (b) Tests made at CERN.

Figure 4.17: Track-without-Hits map with the cut of 7 obtained with the tests made at (a) DESY and (b) CERN.

Only two tracks have been found for the result of DESY, while eight tracks have been found for the results of CERN. Considering the results given by Fig. 4.16 and 4.17, for the tests made at CERN, the small number of good Hits and the increasing undetected track number explain why the detection efficiency drops sharply when the cut value reaches 7 (Fig. 4.15 (b)).

The detection efficiencies are also obtained with the variation of the column level discriminator threshold value. The seed pixel SNR cut (S2N_seed) is set to 4 for this analysis. As an example, Fig. 4.18 gives the results obtained with the tests made at DESY.

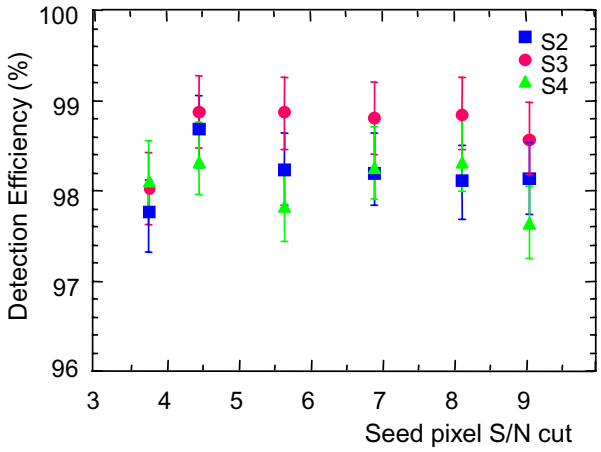


Figure 4.18: Detection efficiency of analog output under different column level discriminator threshold values. Tested at DESY (2005) with 5 GeV electron beams.

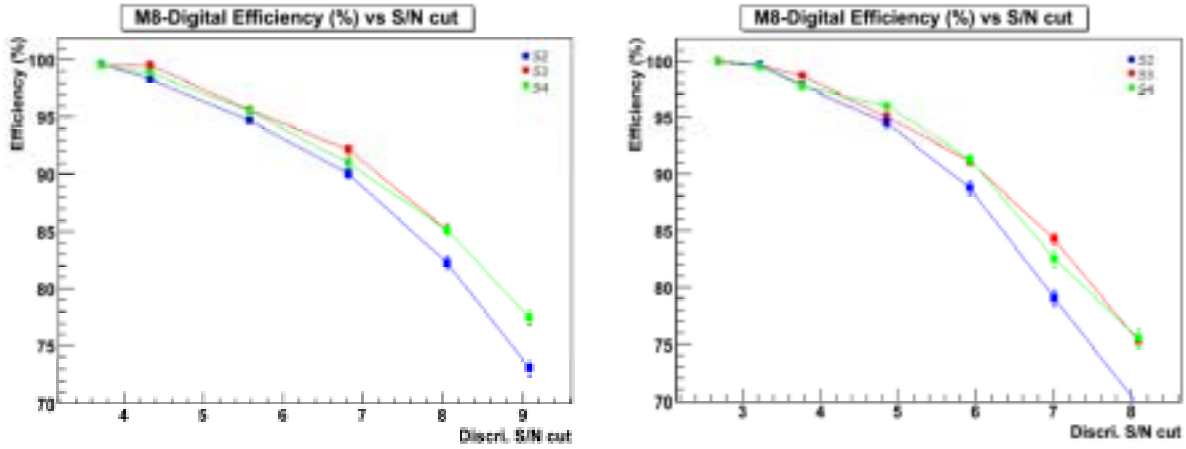
Although decreasing the threshold level of the column level discriminator activates the digital components inside MIMOSA 8, the detection efficiencies of the analog outputs are not disturbed. For the tests made at CERN, similar results are obtained. The increasing activities of the digital components do not affect the performance of the analog outputs. This result shows once more that the digital-to-analog coupling is well controlled for MIMOSA 8.

4.5 Tests of digital outputs

4.5.1 Detection efficiency

The performance of the digital outputs is studied by applying different threshold values to the column level discriminators. As studied in Chapter 3, the applied threshold value influences the sensitivity of the column level discriminator. An important threshold value makes it more robust to noise but the small signals are not detectable by the comparator under this condition. We need to find an appropriate threshold value which reaches high detection efficiency while maintaining small level of noise disturbance.

Fig. 4.19 gives the detection efficiencies as a function of the applied threshold value on the column level discriminator. We notice that very similar results are obtained for the tests made at DESY and CERN.



(a) Detection efficiency vs. threshold cut on discriminator. Tests made at DESY.

(b) Detection efficiency vs. threshold cut on discriminator. Tests made at CERN.

Figure 4.19: Digital outputs detection efficiencies as a function of the SNR threshold value of the discriminators.

In both tests, very high detection efficiencies (about 99%) are achieved when a small threshold cut is applied (less than 4). The detection efficiency decreases with the increase of the threshold cut value since the more hits are cut off because the charges generated is not sufficient to reach the threshold cut level.

Comparing the results of DESY and CERN (Fig. 4.19 (a) and (b)), we notice that these results are compatible with the results obtained with the analog outputs (Fig. 4.15). Very high detection efficiency, above 98%, is achieved when the discriminator threshold value is set to 4. For the threshold cut values below 5, the detection efficiencies are quite stable, varying

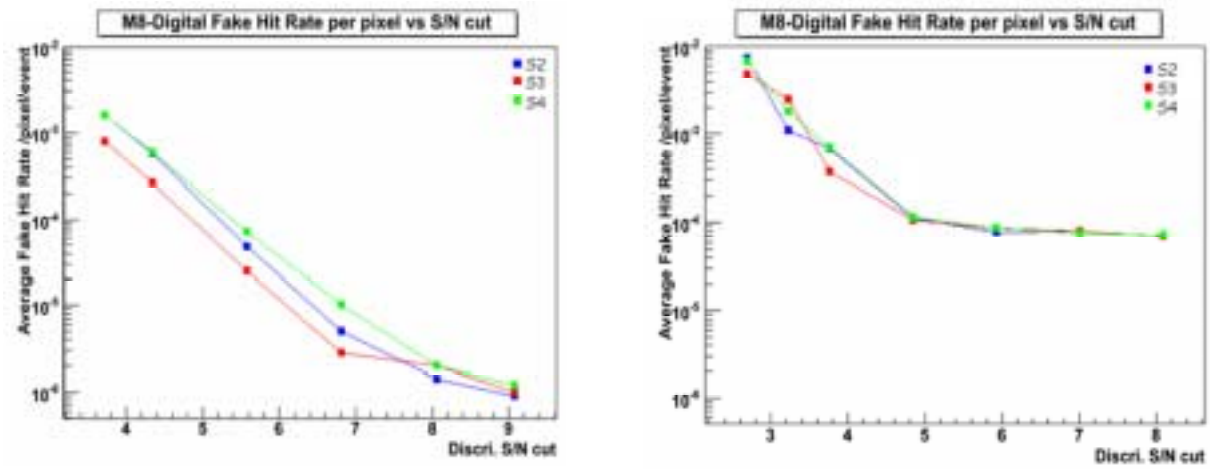
from 99% to 95%. When the threshold cut value reaches 6, the detection efficiencies of CERN drop more quickly than the results measured at DESY. The reason is the same as formerly explained for the analog outputs: the SNR of the seed pixel is lower for the tests made at CERN due to the low charge collected so that the influence of noise increases rapidly when the threshold cut reaches this value.

4.5.2 Average fake hit rate

From the results of Fig. 4.19, it seems that high detection efficiency (~99%) can be simply achieved by lowering the discriminator threshold value. In fact, this is not completely true because lots of Hits, caused by noise, appears when the threshold cut is too small. In order to study the relationship between the detection efficiency and the Hits caused by noise, we use the notion of fake hit rate. The fake hit rate is defined as the probability of Hits caused by noise. It increases with the reduction of threshold level. From the studies of the laboratory tests, the hits are most probably caused by noise instead of by a real signal when the threshold level reaches the limits (Fig. 3.33, Chapter 3). A quantitative study of the fake hit rate as function of external threshold value is carried out. The following method is used to calculate the fake hit rate for each sub-array:

- selection of the events when the track extrapolation is outside the sensor area (the total number of events is noted as M);
- for the selected events, calculation of the total number of hits (N) for each pixel (these hits are obviously caused by the noise);
- calculate the fake hit rate for each pixel which is equal to N/M;
- average the fake hit of the pixels inside the same sub-array.

The fake hit rates for the three sub-arrays are obtained for both tests made at DESY and CERN. The results are shown in Fig. 4.20.



(a) Tests made at DESY. (b) Tests made at CERN.

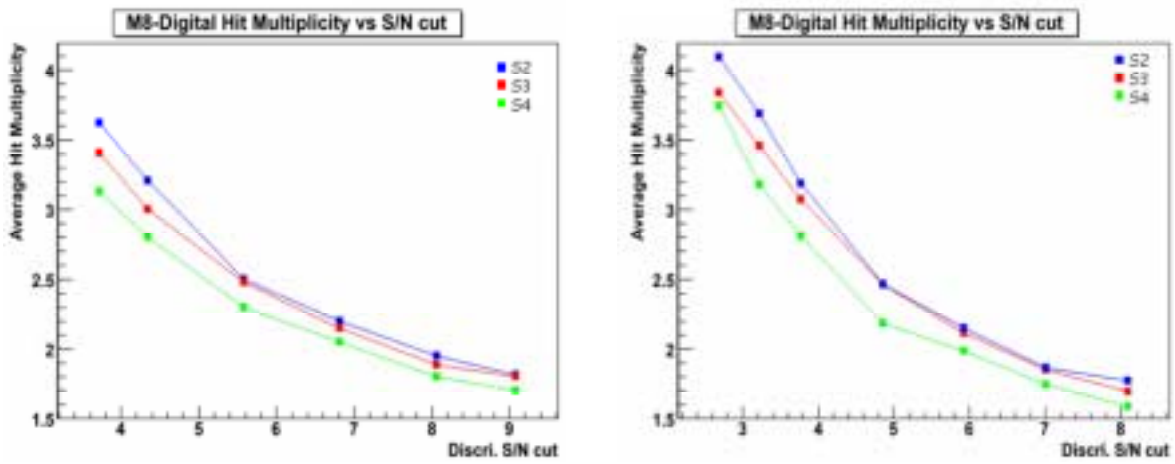
Figure 4.20: Average pixel fake hit rates as a function of the SNR threshold value of the discriminators.

For the results of DESY, the average fake hit rate varies from 10^{-3} to 10^{-6} according to the applied threshold cut values. For the threshold cuts below 7, the sub-array S3 has the lowest fake hit rate. With a threshold cut of about 3.7, the fake hit rate of sub-array S3 is around 8×10^{-4}, which is not negligible. As a result, the highest detection efficiency (about 99%) achieved under this condition is not completely reliable. On the contrary, the fake hit rate drops to 7×10^{-5} when the threshold value is set to about 5.5. Therefore, the detection efficiency (about 96%) found under this condition is quite reliable. For the cut values more important than 6, although the fake hit rate decreases to a very small value (less than 5×10^{-6}), the corresponding detection efficiency is too low comparing the requirements of the vertex detector for the future ILC. More detailed analysis concerned fake hit rate and hit multiplicity can be found in [7].

For the results obtained from the tests made at CERN, the fake hit rate is compatible with the results of DESY when a threshold cut of 3.7 is applied. The value is around 6×10^{-4} under this cut value. For high threshold values (bigger than 5), the fake hit rate at CERN reaches the lowest limit of about 10^{-4}. This value is more important than the fake hit rate obtained at DESY. This is probably due to real random hits created by a more intense beam flux at CERN than at DESY.

4.5.2 Hit multiplicity

Unlike the analog outputs results, the cluster found for the digital contains only the pixels with the value of "1". The average hit multiplicity of the clusters is defined as the average number of "1" inside a cluster for each sub-array. It is very important in that it gives the information on the average size of the clusters, which is very useful for the estimation of the spatial resolution. The results are obtained with different threshold cut values, as shown in Fig. 4.21.



(a) Tests made at DESY. (b) Tests made at CERN.

Figure 4.21: Average hit multiplicity of clusters as a function of the SNR threshold value of the discriminators.

Very similar results are obtained for both the tests made at DESY and CERN. In Fig. 4.21, we notice that the multiplicity of sub-array S2, which contains the pixels of the smallest size,

attains the maximum value. It is a reasonable result. Because of its smallest pixel size, there are more pixels for sub-array S2 inside the diffusion area of generated charge, which is the same for all the three sub-arrays.

The multiplicity decreases when increasing the threshold cut values. Before the threshold cut value reaches 4, the average multiplicity is above 3 for sub-array S2 and S3. For sub-array S4, the values are slightly lower: about 2.7 when the threshold cut is around 4. When an important threshold is applied (greater than 6 for example), the average multiplicity is lower than 2, indicating that the cluster size is generally equal to one under such conditions. Additionally, more important average multiplicity is observed for the tests made at DESY compared to the tests made at CERN. The reason is the different SNR for the seed pixels, which is compatible with the results of the detection efficiencies and the average fake hit rates.

While the average multiplicity for each sub-array gives us a global view of the average cluster size under the corresponding threshold cut, more detailed information is needed because the average value is generally influenced strongly by the exceptional maximum (or minimum) value. Thus, it is also very useful to study the distribution of the multiplicity for the different threshold cuts applied. The distribution of the multiplicity gives us the information on the probability (or the percentage) that a certain size of the cluster can be found when a certain threshold cut value is applied. In other words, we get the information on most probable multiplicity value under a certain threshold cut thanks to the distribution of the multiplicity for the different threshold cuts. This information is useful to understand the analysis results of the spatial resolution that will be presented in the following section of this chapter.

In Fig. 4.22 below, as an example, the distributions of the multiplicity under the different threshold cuts for the tests made at DESY are given.

The behaviors for the three sub-arrays are similar, as described in the following:

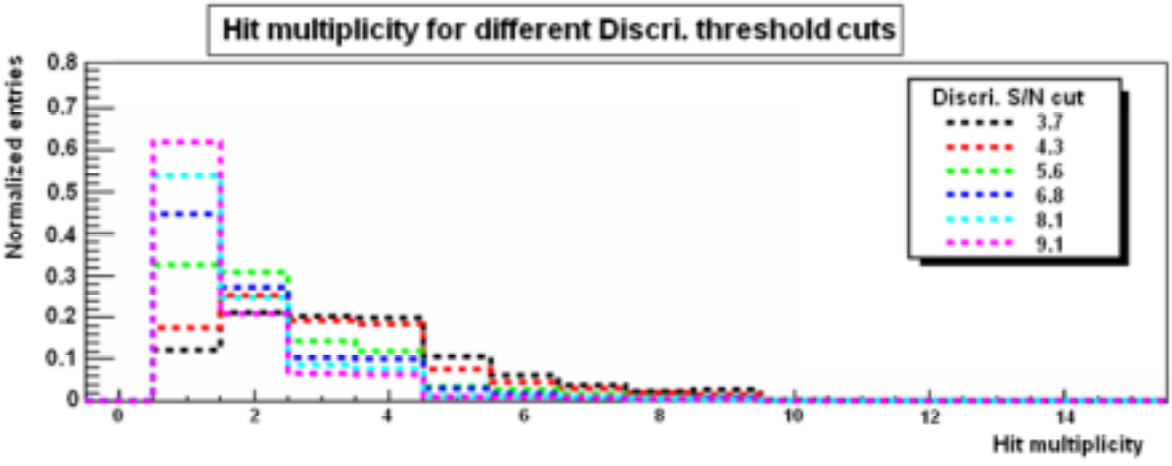
- When the applied threshold cut is low (3.7 or 4.3 for example), there is no significant difference among the probabilities for the multiplicity of 1, 2, 3 and 4. The values vary from 12% to 28% according to the different sub-arrays. Taking sub-array S3 as an example, the probabilities for the hit multiplicities of 1, 2, 3, 4 are 19%, 28%, 20% and 19% respectively when the threshold value is equal to 3.7 (Fig. 4.22 (b)). Almost 86% of the total cluster found in analysis is covered.

- When a high threshold cut is applied (6.8 to 9.1 for example), almost half of the clusters found contains only one pixel with the value of "1". This means that it is impossible to get a spatial resolution better than what is predicted by uniform distribution: $25\,\mu m / \sqrt{12} \approx 7.2\,\mu m$ for half of the case.

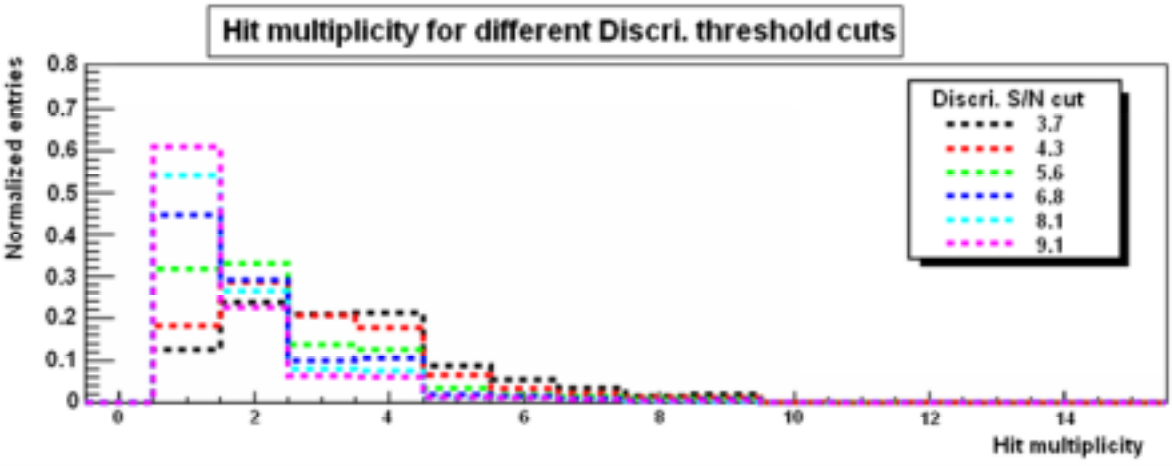
- The multiplicity of important values (more than 5) is rare case despite the different threshold cut applied. If we calculate the total probability for all the multiplicity values more important than 5, the sum is less than 15% for all the three sub-arrays.

Considering the results of the detection efficiency, the average fake hit rate and the hit multiplicity, the most appropriate threshold cut should be set between 3.5 and 5. For the threshold values among this range, a high detection efficiency (more than 95%) is achievable with very small fake hit rate (around 10^{-4}) and an average hit multiplicity above 2.

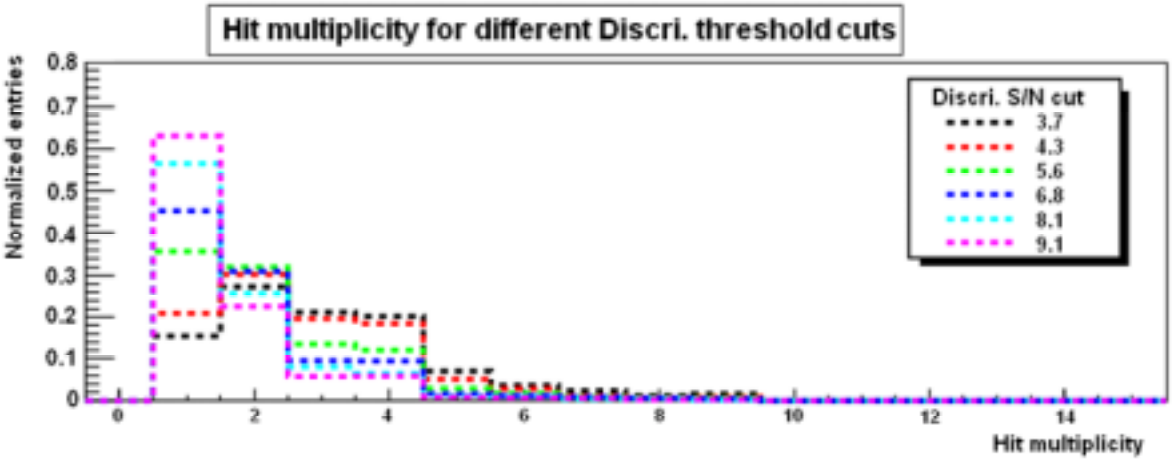
Similar results are obtained for the tests made at CERN.



(a) Distributions of the multiplicity under different discriminator threshold cuts for sub-array S2.



(b) Distributions of the multiplicity under different discriminator threshold cuts for sub-array S3.



(c) Distributions of the multiplicity under different discriminator threshold cuts for sub-array S4.

Figure 4.22: Distributions of the multiplicity under different discriminator threshold cuts for the three sub-arrays. Tested at DESY (2005) with 5 GeV electron beams.

4.6 Spatial resolution of the analog and digital outputs

For the 5 GeV electron beams used for the tests at DESY, the multiple scattering is so important that the spatial resolution could not be measured properly. The pion beams of 180 GeV were used for the tests at CERN in order to study the spatial resolution of MIMOSA 8 without multiple scattering.

For a position sensitive detector, the spatial resolution is in fact equivalent to the error on the real hit position. Thus, measurement of the spatial resolution consists to find out the error between the track extrapolation determined by the 8 reference planes and the hit position found in the sensor. Several methods are available to find the position of the hit by the charge collected, such as: digital hit position, center of gravity hit position and non linear "eta" corrected hit position. In our analysis, we used only the Center Of Gravity (COG) hit position finding method.

On the sensor plane (HU, HV), the COG method gives the hit position in terms of the charge fractions of the pixels in the cluster. For example, for HU, the hit position U_h is given by:

$$U_h = \frac{\sum_{i}^{N} Q_i U_i}{Q_{total}} \qquad (4.2)$$

where N is the total number of pixels inside cluster, U_i is the u position of the i^{th} pixel and Q_i is the corresponding charge collected and Q_{total} is the total charge collected by cluster.

After having obtained the hit position, the residual (σ) of the difference between hit position (U_h, V_h) and the track extrapolation (U_t, V_t) is the spatial resolution. It is given for both U and V coordinates. Spatial resolutions are obtained for both analog and digital outputs.

For the analog outputs, the results are very bad as it is very difficult to correctly align the reference planes and the sensor plane due to the small dimension of the pixel array (only 7 columns* of outputs are available).

The results of sub-array S2 are shown in Fig. 4.23. The threshold voltage is set to 7 mV.

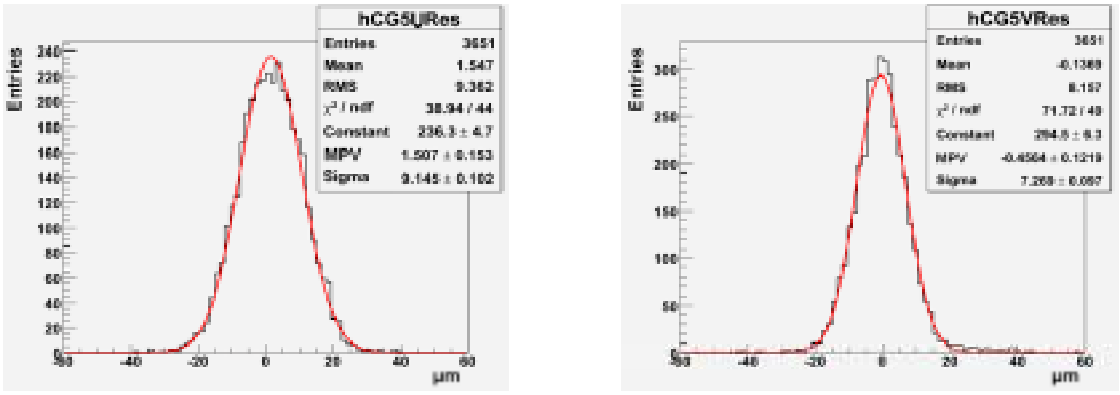
For digital outputs, the spatial resolution is $9.14 \pm 0.10\,\mu m$ for U coordinate and $7.27 \pm 0.10\,\mu m$ for V coordinate. The spatial resolution of V coordinate is better than U coordinate simply because that there are 128 rows for the alignment of V coordinate while there are only 24 columns for the alignment of U coordinate.

The results are approached to the prediction of uniform distribution (the best spatial resolution for one pixel detector):

$$\text{Resolution} = P/\sqrt{12} = \frac{25}{3.46} \approx 7.23\,\mu m$$

* The last column of outputs is used for synchronisation purpose for the data acquisition system.

The reason is that the value of the "charge" for the digital outputs is always equal to 1, so that the contribution of each pixel to the total "charge" of cluster is equal. It is very difficult to identify the seed pixel. Additionally, the hit multiplicity under a threshold cut of 7 mV is 1 for almost 45% of the cases and is 2 for 29% of the case (Fig. 4.22 (a)). When the hit multiplicity is equal to 2, it is impossible to identify the seed pixel by the COG method. Thus, for about 74% of the case, we can consider that there is only one pixel which contributes to the measurements of the spatial resolution.



(a) Residual distribution in U direction.

(b) Residual distribution in V direction.

Figure 4.23: Residuals of the difference between the hit position and the track extrapolation for both U and V directions. Tests made at CERN in 2006.

The spatial resolutions are also obtained as a function of the discriminator threshold cut values. The results of V direction are given in Fig. 4.24 below.

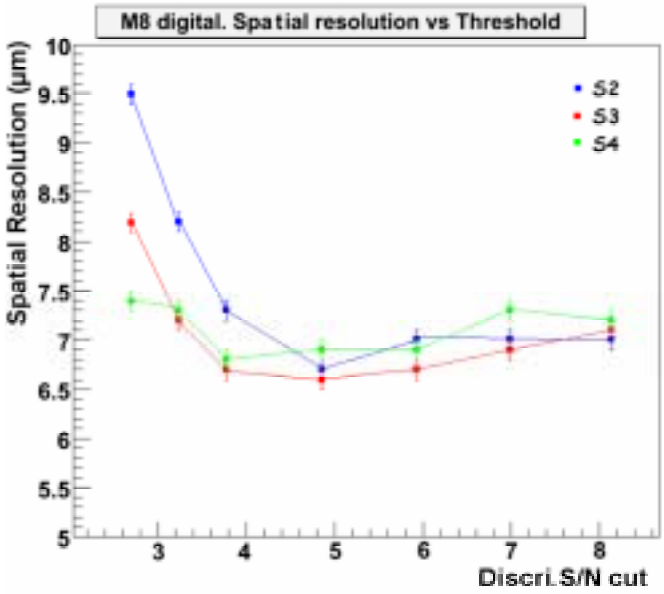


Figure 4.24: Spatial resolution in V direction of the digital outputs as the function of the column level discriminator threshold cut values. Tests made at CERN in 2006.

For very low threshold cut values (below 3), spatial resolutions are a bit worse. The reason is probably the increased number of hit multiplicity for this threshold cut value. This point has to be studied in more details in the future. As shown in Fig. 4.20 (b), the fake hit rate is on the

order of 10^{-2} when the threshold cut is below 3. The increasing false hits increase the size of the clusters so that the precision of the hit position is degraded.

We notice that the spatial resolution slightly increases when an important threshold cut value is applied (around 7 or above). It is reasonable because most of the clusters contain only one or two pixels according to the hit multiplicity under such a condition (Fig. 4.21 and Fig. 4.22).

For the threshold cut values from 3.7 to 6, the spatial resolutions for all the sub-arrays remain quite stable. The best spatial resolution is found with sub-array S3, which has the lower fake hit rate with the best detection efficiency.

4.7 Conclusion

The prototype MIMOSA 8 has been tested with high energy particle beams at DESY with 5 GeV electrons in 2005 and at CERN with 180 GeV pions in 2006, respectively. The results are summarized in the following.

A. Results of analog output:

- The results of the tests made at DESY in 2005:

Sub Array	Temporal Noise (e-)	Charge in 3×3 cluster (e-)	SNR of seed pixel	Detection efficiency to MIPs
S2	11±2	353±4	8.6±0.2	98.2±0.4 %
S3	12±2	392±4	9.8±0.1	98.6±0.3 %
S4	14±2	433±5	9.4±0.2	98.0±0.4 %

Table 4.4: The performance of the analog outputs for MIMOSA 8.
Measured at DESY in 2005.

- The results of the tests made at CERN in 2006:

Sub Array	Temporal Noise (e-)	Charge in 3×3 cluster (e-)	S/N of seed pixel	Detection efficiency to MIPs
S2	10±2	306±4	6.2±0.5	97.6±0.6 %
S3	11±2	328±4	6.3±0.2	98.1±0.5 %
S4	12±2	381±5	6.6±0.6	97.5±0.5 %

Table 4.5: The performance of the analog outputs for MIMOSA 8.
Measured at CERN in 2006.

Very good results are obtained in both tests. The charge collected in the tests made at CERN is less than at DESY because the alignment for the tests made at CERN is very difficult to

optimize. Only 7 columns are available. However, very high detection efficiencies are also achieved for the tests made at CERN.

The charge collected is not very high due to the limit thickness of the epitaxial layer. Indicated by the constructor that the thickness is 8 µm, we found that the real thickness is probably only around 6.5 µm, which means that the total average charge generated by the MIPs is only 528 e-. Considered a CCE of about 90%, the maximum charge that can be collected is around 475 e-. It is a very limit value. Thus, a CMOS process with thicker epitaxial layer is preferred.

B. Results of digital outputs

- The results of the tests made at DESY in 2005:

Sub Array	Average Hit multiplicity	Fake hit rate	Detection efficiency to MIPs
S2	3.7	1.6×10^{-3}	99.6±0.1%
S3	3.6	0.8×10^{-3}	99.5±0.1%
S4	3.1	1.6×10^{-3}	99.5±0.1%

Table 4.6: The performance of the digital outputs for MIMOSA 8.
Measured at DESY in 2006.

- The results of the tests made at CERN in 2006:

Sub Array	Average Hit multiplicity	Fake hit rate	Detection efficiency to MIPs
S2	3.7	1.1×10^{-3}	99.7±0.1%
S3	3.4	2.5×10^{-3}	99.6±0.1%
S4	3.2	1.8×10^{-3}	99.5±0.2%

Table 4.7: The performance of the digital outputs for MIMOSA 8.
Measured at CERN in 2006.

The above results are obtained with a discriminator cut around 3.5. Very high detection efficiencies are achieved with moderate fake hit rate.

Moreover, the spatial resolutions for all the three sub-arrays are obtained with the tests made at CERN. The value is around 7 µm, close to the minimum value predicted by uniform distribution. The reason is that only two values are available for the column level discriminator: “1” or “0”. A column level ADC is needed in order to improve the spatial resolution.

Sub-array S3 gives the best performance for both analog and digital outputs. It indicates that the diode size used inside ($1.7 \times 1.7\ \mu m^2$) achieves the optimal compromise between charge to voltage conversion gain and noise for the TSMC 0.25 µm digital process.

All the tests have been carried on for a main clock frequency of 40 MHz. Tests at higher readout frequency (the readout speed is more fast) are needed.

Reference

[1] **Y. Li**, et al., "CMOS Sensors for the Vertex Detector of the future International Linear Collider," *Nucl. Instrum. Meth. A 572*, 2007, pp. 300-304.

[2] Y. Degerli, M. Besançon, A. Besson, G.Claus, G. Deptuch, W. Dulinski, N. Fourches, M. Goffe, A. Himmi, **Y. Li**, P. Lutz, F. Orsini and M. Szelezniak, "Performance of a fast binary readout CMOS active pixel sensor chip designed for charged particle detection," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 6, 2006, pp. 3949-3955.

[3] C. Colledani et al., "A Submicron Precision Silicon Telescope for Beam Test Purposes," *Nucl. Instrum. Meth. A 372*, 1997, pp. 379-384.

[4] A. Besson, D. Grandjean and A. Shabetai, "MAF – Mimosa Analysis Framework Documentation," IPHC, Strasbourg, 2006.

[4] Y. Gornushkin, "Test results of monolithic active pixel sensors for charged particle tracking," *Nucl. Instrum. Meth. A 478*, Feb. 2002, pp. 311-315.

[5] G. Deptuch, et al., "Design and testing of monolithic active pixel sensors for charged particle tracking," *IEEE Trans. Nucl. Sci.*, vol. 49, April 2002, pp 601-610.

[7] A. Besson, G. Claus, C. Colledani, Y. Degerli, G. Deptuch, M. Deveaux, W. Dulinski, N. Fourches, M. Goffe, D. Grandjean, F. Guilloux, S. Heini, A. Himmi, C. Hu, K. Jaaskelainen, **Y. Li**, P. Lutz, F. Orsini, M. Pellicioli, E. Scopelliti, A. Shabetai, M. Szelezniak, I. Valin and Marc Winter, "A vertex detector for the international linear collider based on CMOS sensors," *Nucl. Instrum. Meth. A 568*, Nov. 2006, pp. 233-239.

Chapter V

Prospect: Column level ADC and new column level comparator

5.1 Column level ADC

5.1.1 Introduction

The first fast CMOS active sensor chip designed for charged particle detection, MIMOSA 8, exhibits very encouraging performances. Thanks to the double sampling architecture integrated inside the pixel, the noise of the circuit has been largely reduced. The input referenced noise is only about tens of electrons. The first step towards final on-chip data processing has been realized by a column level comparator which digitalizes analog pixel signals into a 1-bit digital code. The readout speed reaches up to 12 µs/frame (corresponding to 170 MHz main clock frequency) for the analog outputs and 20 µs/frame (corresponding to a 100 MHz clocking frequency) for the digital outputs. With the test of 5 GeV electron beams at DESY in 2005, detection efficiency of Minimum Ionizing Particles reaches 98% for both analog outputs and digital outputs with a SNR threshold cut small than 5.

However, the spatial resolution of the digital output of the MIMOSA 8, tested with 180 GeV/c pion beam at CERN in 2006, was only around 7 µm. The reason is that it was impossible to identify the center of the cluster because each digital output contained only one bit. Then, center of gravity reconstruction in a series of "1" is very difficult. The spatial resolution is close to what is only given by uniform distribution: $Pitch\_pixel / \sqrt{12}$ [1]. It is around 7 µm when the size of pixel is 25×25 µm² with an epitaxial layer of about 8 µm. In order to reach a good spatial resolution (around 3 µm) required by the vertex detector of the future ILC, the comparator has to be replaced by a column level ADC (Analog-to-Digital Converter) which gives more accurate digital information on every analog pixel output.

Since the signal level is very small (tens of mV), the major difficulty of this column level ADC is that it has to reach a resolution of a few hundreds of µV with the auto compensation of offsets arising from various sources: the pixels reference level, the output buffers of column bus and the residual offset introduced by the ADC itself. Moreover, a high signal processing speed is also required (about 160 ns per conversion, which is equivalent to reading a pixel at 100MHz). In addition, this ADC should be very compact because the width of each column is only around 25 µm. Power dissipation is another important issue because large

pixel array will be realized for the final sensors. Considering all the requirements above, the design of this ADC is not a trivial task and was never achieved so far.

5.1.2 General description of ADC

Being the link between the analog world and the digital processing system, the ADCs provide discrete outputs that approximate the input continuous signal. Nowadays, in mixed signal circuits, the performance of the ADC can be a limit to the overall speed and precision. The fundamental aspects of the ADCs are presented in this section.

A. Ideal ADC model

The block diagram of an N-bit ADC is shown in Fig. 5.1 below.

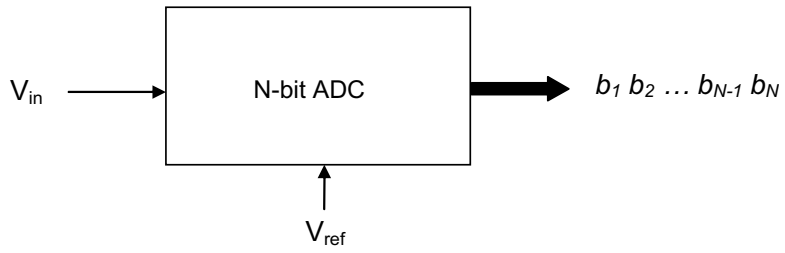


Figure: 5.1 Block diagram of an N-bit ADC.

In the diagram, the relationship between the input analog signal V_{in} and the output code B_{out}, which is composed of $b_1\ b_2 \ldots\ b_{N-1}\ b_N$, is given by:

$$V_{ref}(b_1 2^{-1} + b_2 2^{-2} + \cdots + b_{N-1} 2^{-(N-1)} + b_N 2^{-N}) = V_{in} \pm V_{err} \quad (5.1)$$

where V_{err} represents the difference between the converted value (reconstructed by the output digital code) and the real input value. Its range is:

$$-\frac{1}{2}V_{LSB} < V_{err} < \frac{1}{2}V_{LSB}$$

Here, V_{LSB} is the value of the Least Significant Bit, which is defined as:

$$V_{LSB} = \frac{1}{2^N}V_{ref} \quad (5.2)$$

V_{ref} is the analog reference voltage which the analog input is compared.

A typical transfer curve of the N-bit ADC is shown in Fig. 5.2. From this curve, it can be seen that a range of the input analog values is represented by the same digital output code. For example, for the analog input in the range from $\frac{1}{2}V_{LSB}$ to $\frac{3}{2}V_{LSB}$, the output code is always 00…01. This effect, resulting from the finite resolution of the ADC, introduces an ambiguity between the original analog input and its digitalized value, which is called "quantization error" as presented by V_{err} in the equation (5.1). The quantization error of an ideal N-bit ADC is also shown in Fig. 5.2. When the input analog signal begins from 0, the corresponding output digital code is also 00…00 and no error is produced. As the input signal increases towards $\frac{1}{2}V_{LSB}$ level, the error increases because the input is no longer zero. When the input reaches $\frac{1}{2}V_{LSB}$, the output code changes from 00…00 to 00…01. But the real input is only at $\frac{1}{2}V_{LSB}$ (instead of 1 V_{LSB}), so the error is now $-\frac{1}{2}V_{LSB}$. As the input increases past $\frac{1}{2}V_{LSB}$, the error becomes less negative, until the input reaches 1 V_{LSB}, where the error is zero. This process continues through the entire input range. The range of quantization error is therefore from $-\frac{1}{2}V_{LSB}$ to $\frac{1}{2}V_{LSB}$.

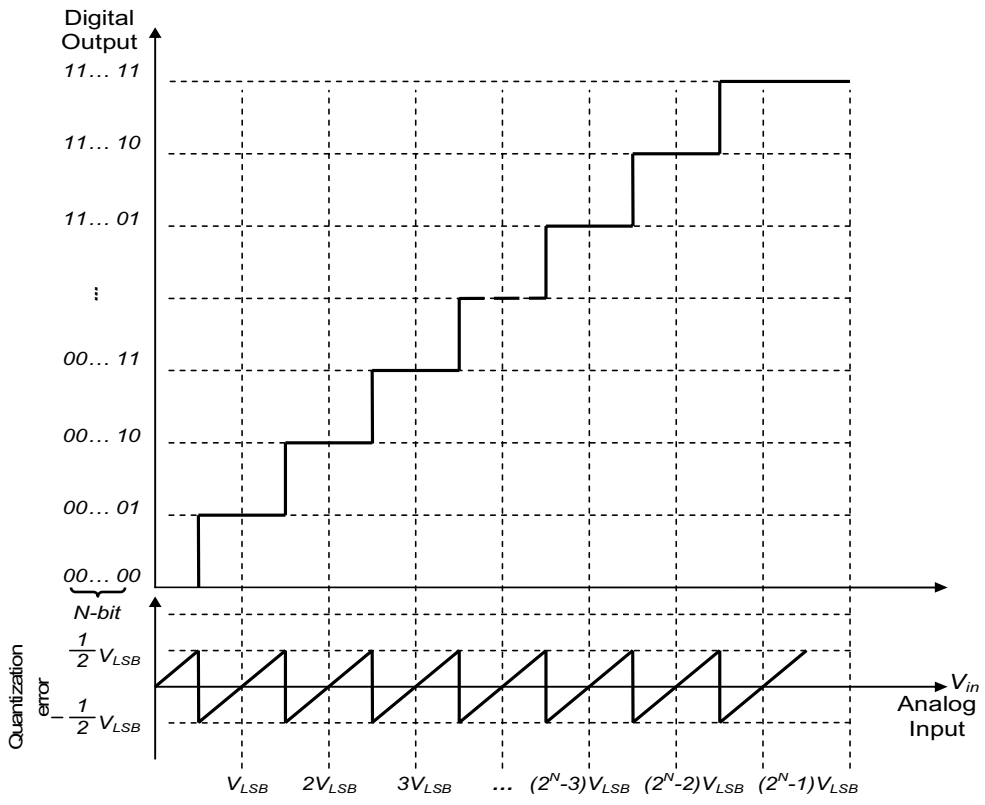


Figure: 5.2 Input/output characteristic of an N-bit ADC.

B. Specifications of ADC

A number of important terms describing the performances of the ADC are defined as follows.

Resolution

The resolution of an ADC is the smallest analog change that can be resolved by the ADC. It is defined as the number of distinct analog levels corresponding to the different digital words. Therefore, an ADC of N-bit resolution implies that the converter can resolve 2^N distinct analog levels. Referred to the definition of the LSB, the resolution of an ADC is generally therefore by its LSB.

Offset and Gain error

In an ideal ADC, an input signal of $\frac{1}{2} V_{LSB}$ will cause the output code passes from zero to one. Any deviation from this point is called the Offset Error (called also Zero scale error or Zero scale offset error). It is defined to be the input analog signal that should produce zero output. It can be expressed as the deviation of $V_{00\ldots01}$ from $\frac{1}{2} V_{LSB}$, given by the following equation:

$$E_{off} = \frac{V_{00...01}}{V_{LSB}} - \frac{1}{2} LSB \qquad (5.3)$$

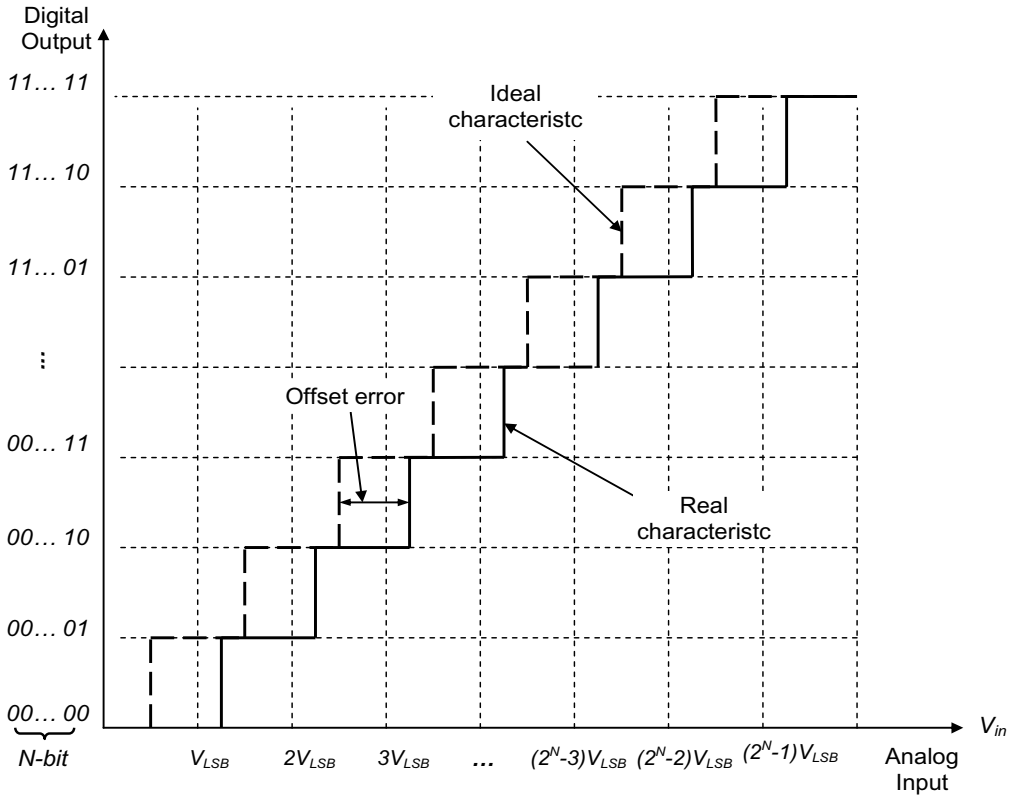
where $V_{00...01}$ represents the transition voltage of the code 00…01, which is $\frac{1}{2} V_{LSB}$ in the ideal case. The Offset error is a constant and can be calibrated out.

The Gain Error is defined to be the difference at the full-scale value between the ideal and actual curves when the Offset Error has been reduced to zero. For an ADC, it is given by:

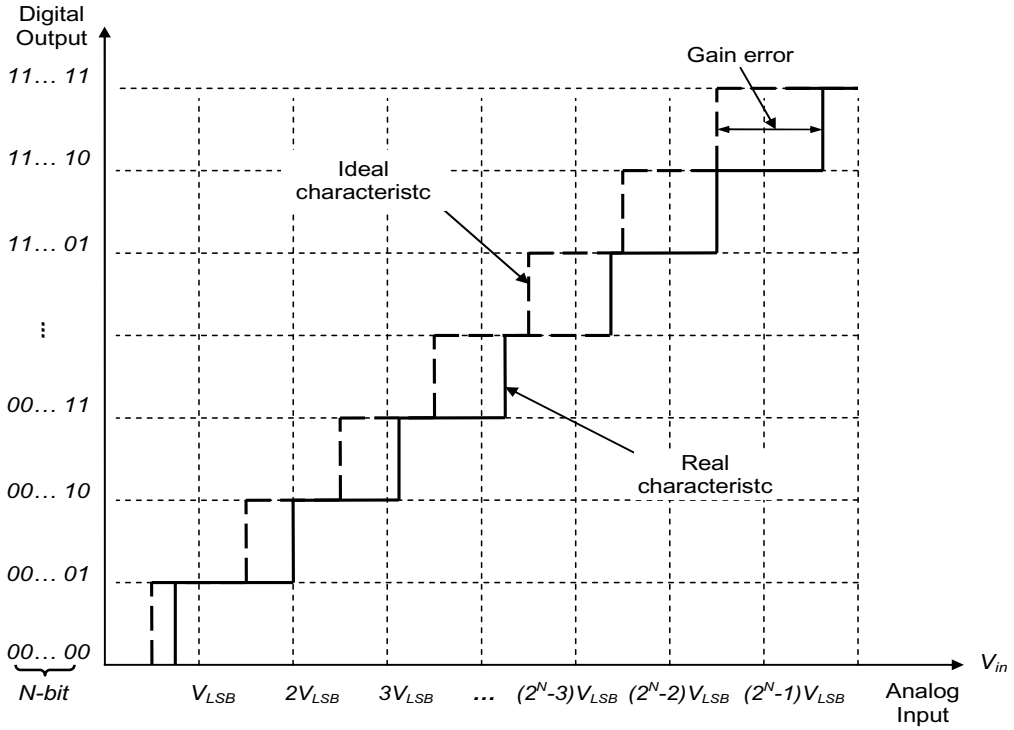
$$E_{gain} = \left(\frac{V_{11...11}}{V_{LSB}} - \frac{V_{00...01}}{V_{LSB}}\right) - (2^N - 2) \qquad (5.4)$$

where $V_{00...01}$ and $V_{11...11}$ represent respectively the transition voltages for the codes 00…01 and 11…11, which are $\frac{1}{2} V_{LSB}$ and $(((2^N - 2) + \frac{1}{2}))V_{LSB}$ in the ideal case.

Graphically, the Offset and Gain Error are illustrated in Fig. 5.3.



(a) Offset error of an N-bit ADC



(b) Gain error of an ADC. It is assumed that all other errors are not presented.

Figure: 5.3 Offset and gain error of an N-bit ADC.

Integral Nonlinearity (INL) error

After removing the Offset error, the INL of the ADC measures the straightness of the transfer function. It is defined as the maximum difference between the best linear fit of the actual transfer function and the ideal curve. It is the "large scale" overall transfer function error. Practically, it can be calculated by the differences between the real output code and the ideal output code. The INL is usually measured in LSB.

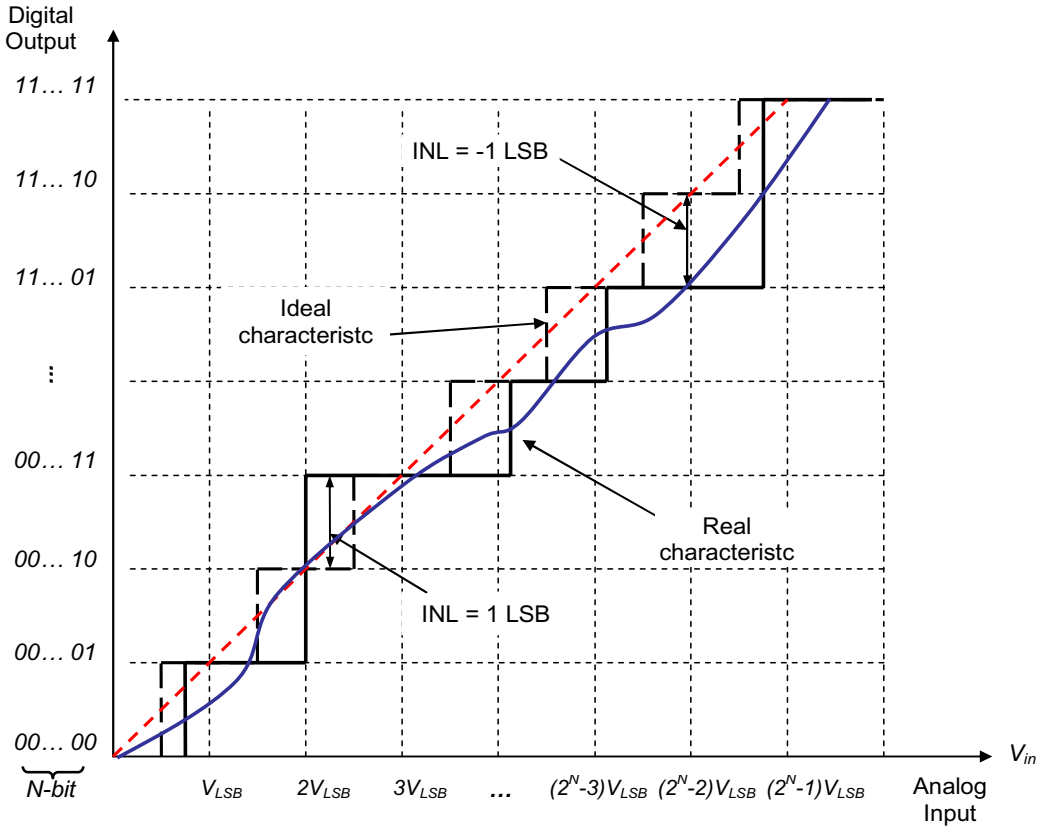


Figure: 5.4 INL error of an N-bit ADC.

Differential Nonlinearity (DNL) error

The DNL of the ADC describes the error in each step size. It is the "small scale" code to code errors. The DNL is defined as the maximum difference between two consecutive output codes. It is given by:

$$DNL = (D_{step} - 1)LSB \qquad (5.5)$$

where D_{step} is the maximum difference between two consecutive output codes.

If there were no missing code, the D_{step} is equal to 1 LSB and the corresponding DNL is equal to 0.

Fig. 5.5 illustrates the DNL errors.

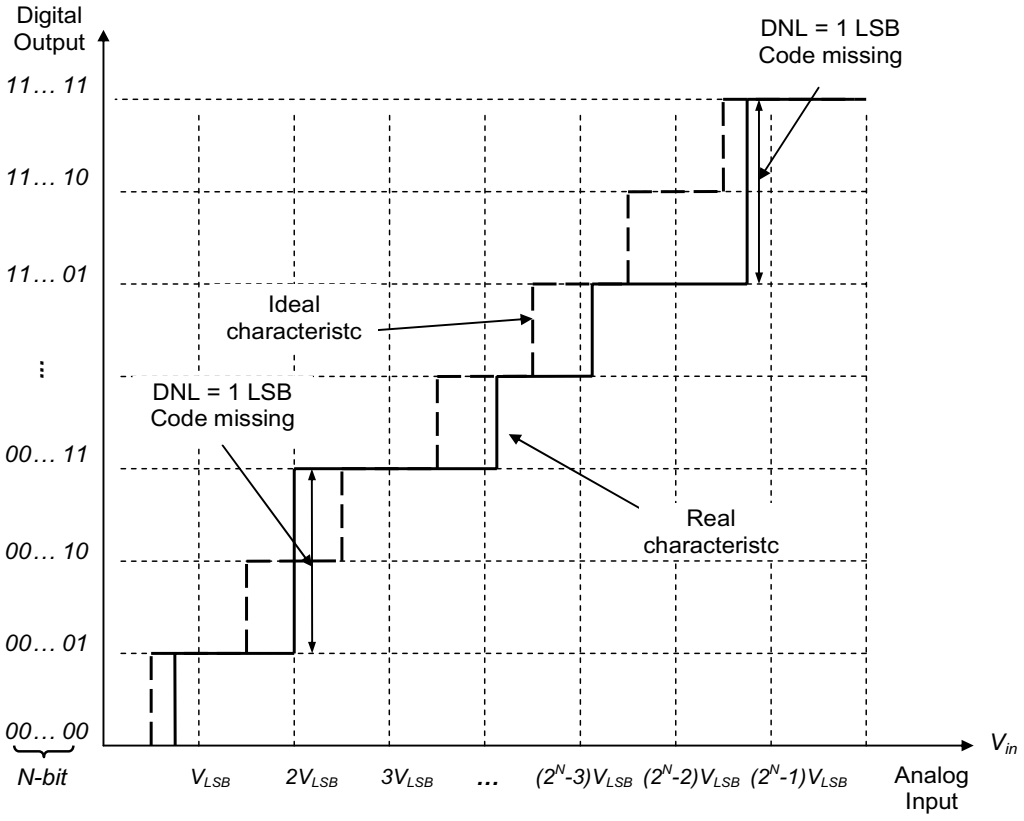


Figure: 5.5 DNL errors of an N-bit ADC.

AD conversion time and Sampling rate

In an ADC, the conversion time is the time needed for the converter to complete a full single measurement including sampling of the input signal and output the digital codes. The maximum sampling rate is defined by the speed at which samples can be continuously converted and is generally the inverse of the conversion time.

Signal-to-Noise ratio (SNR)

When a sinusoidal signal is used at input, the SNR is defined as the ratio of the output signal power to the total output noise power, not including harmonic content.

For an ideal N-bit ADC, the maximum output signal is the full scale signal and the output noise is due to the quantization noise. As $-V_{LSB} < \epsilon < V_{LSB}$, The error due to the quantization noise can be calculated:

$$\sigma = \sqrt{\frac{1}{V_{LSB}} \int_{-\frac{1}{2}V_{LSB}}^{\frac{1}{2}V_{LSB}} \varepsilon^2 d\varepsilon} = \frac{V_{LSB}}{\sqrt{12}} \qquad (5.6)$$

For a sinusoidal signal between 0 and V_{FS}, its average power is $V_{FS}/2\sqrt{2}$. The full scale value V_{FS} is generally equal to V_{ref}. As a result, considering the equation (5.2), the SNR of an ideal N-bit ADC for an input sinusoidal signal is given by:

$$SNR = \frac{V_{FS}/2\sqrt{2}}{\sigma} = \frac{V_{ref}/2\sqrt{2}}{V_{LSB}/\sqrt{12}} = \frac{2^N\sqrt{6}}{2} \qquad (5.7)$$

and its value in decibel is:

$$SNR(dB) = 6.02N + 1.76 \qquad (5.8)$$

where N is the total number of bits.

Effective-Number-Of-Bits (ENOB)

The ENOB of an N-bit ADC means that this ADC performs like an ideal ADC whose resolution is ENOB.

Because the ideal ADC doesn't suffer from distortion, the ENOB can be calculated by equation (5.8):

$$ENOB = \frac{SNR - 1.76}{6.02} \qquad (5.9)$$

Dynamic range

The dynamic range is defined as the ratio between the maximum and the minimum of input signals that can be resolved by the converter. As the minimum input signal is equal to V_{LSB}, the dynamic range of an ideal N-bit ADC equals to its maximum SNR which is given by the equation (5.8).

For High Energy Physics applications, the dynamic range of an ADC can be also expressed by its number of bits. For example, for an 8-bit ADC, it has a 256 dynamic range.

5.1.3 Standard ADC types

Four main types of ADC are introduced in this section. The different architectures are: the integration ADC, the successive approximation ADC, the flash ADC and the delta-sigma ADC. The last type, the delta-sigma ADC which is not appropriate to our application, is not

discussed here. Following the brief introduction of their working principles, the advantages and the limits of each ADC are presented.

A. Integration ADC

Two basic architectures of integration ADC are: single slope ADC and dual slope ADC.

The single slope ADC is derived from digital ramp ADC, which compares the input signal with the analog output of a DAC (Digital-to-Analog Converter) connected to a binary counter and uses the comparator's output to conrol the counter when to stop counting and reset. The following schematic diagram shows the basic idea of a digital ramp ADC.

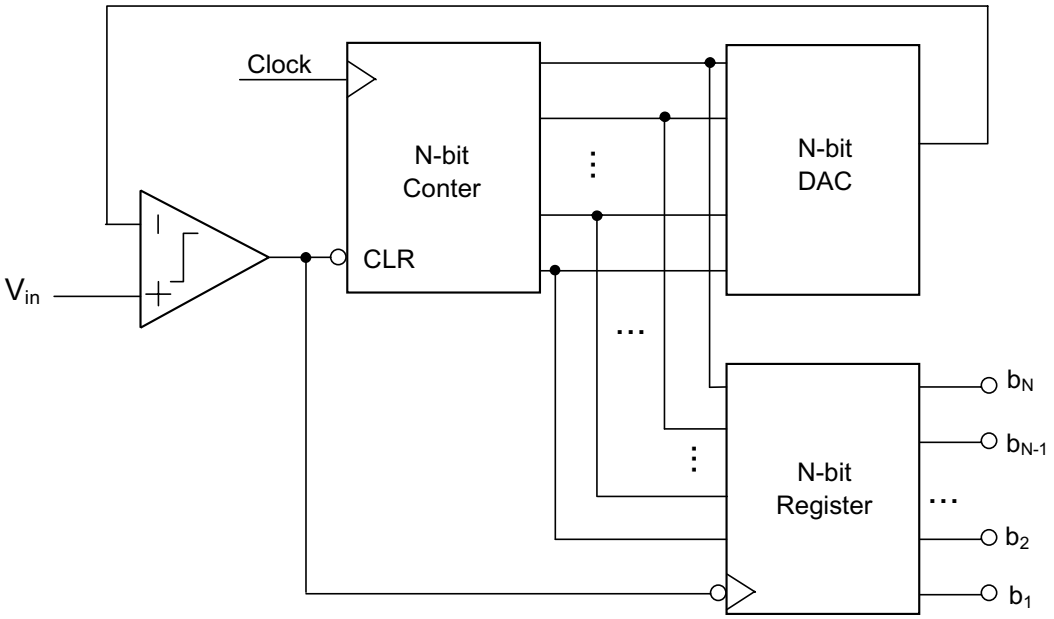


Figure 5.6: Basic architecture of a digital ramp ADC

In the single slope ADC, instead of using a DAC to generate a step-by-step ramped analog output, an operational amplifier (op-amp) circuit called integrator is used to generate a smooth analog signal. The amplitude of this ramp signal is equal to:

$$V_{ramp} = (-V_{ref}) \times t \qquad (5.10)$$

where V_{ref} is the reference voltage and t is the conversion time.

The basic schematic of a single slope integration ADC is shown as follows.

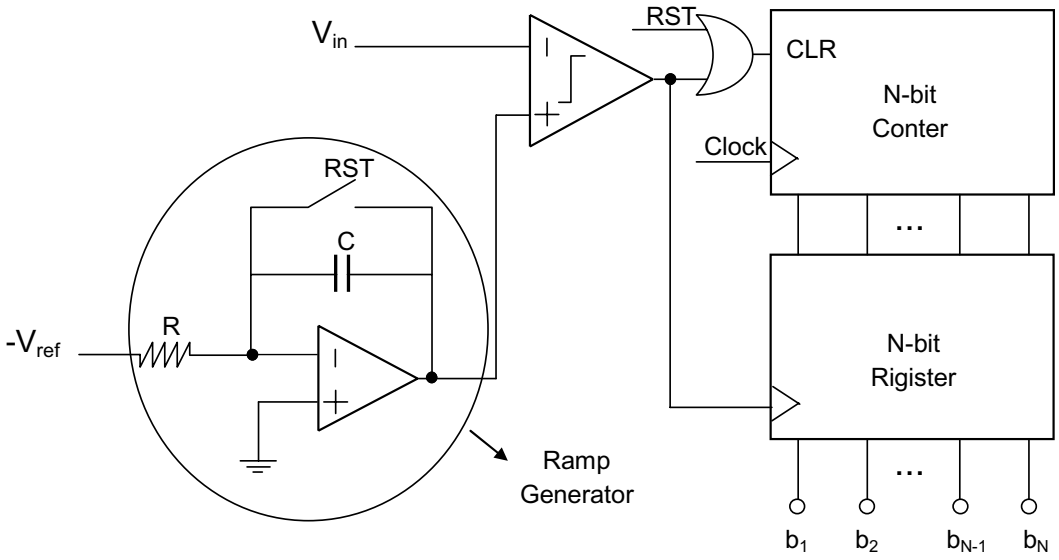


Figure 5.7: Single slope ADC

The analog input is compared with the smooth ramp by a comparator. A digital counter, controlled with a precise timing, is used to measure the time needed for the smooth ramp to exceed the input signal. When V_{ramp} exceeds V_{in}, the comparator changes its output state from "0" to "1" memorizing the current output of the counter, which is the corresponding digital code of the input analog signal, in the N-bit Register. Simultaneously, the counter is reset.

The conversion time is equal to:

$$t = RC\frac{V_{in}}{V_{ref}} \qquad (5.11)$$

The digital output is then equal to:

$$B_{digital} = t \times f_{clk} \qquad (5.12)$$

where f_{clk} is the clock frequency.

With the above equations, the single slope ADC is equivalent to a precious timing measurement system. However, since the rate of integration and the rate of count are independent of each other, the linearity of the integrator is very important. Any variation of it comparing to the rate of count will result in a loss of accuracy. For overcoming this problem, the dual-slope ADC is proposed. A system diagram is shown in the figure below.

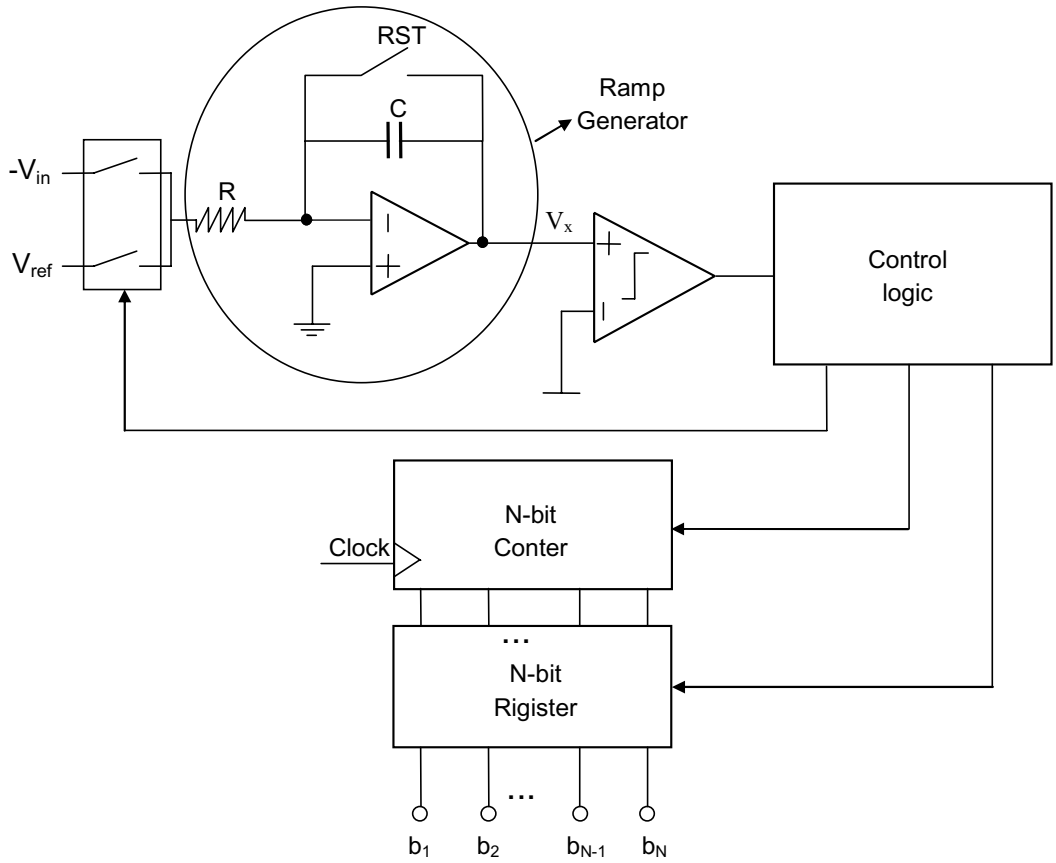


Figure 5.8 : Dual slope ADC

The dual slope ADC performs its conversion in two steps.

The first step fixes the time integration. Starting from a reset integrator, the input signal is integrated during a fixed amount of time T_1 which corresponds with a full count of the counter.

The second step is signal measurement. The reference voltage is integrated during this step. Because the sign of the reference voltage is opposite to the input signal, the integrator is discharged. The counter is counting during the discharge time. When V_x is smaller than zero, the comparator changes its output state and makes the control logic to stop the counter. The discharge time T_2 is then measured by the counter and the current output of counter is the digital code corresponding to the input signal. The following figure gives the voltage of V_x for three different inputs.

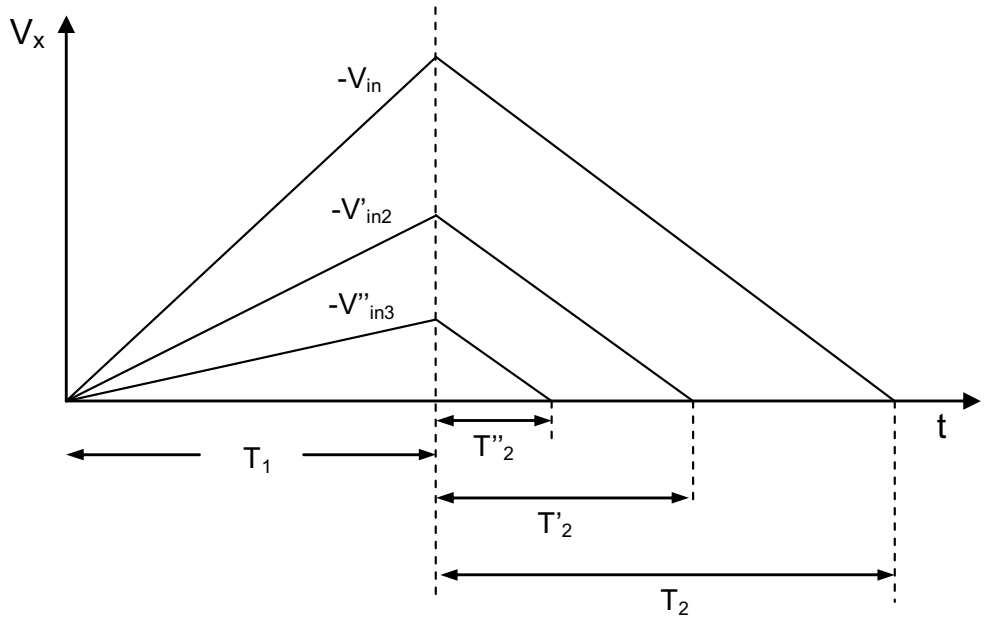


Figure 5.9: Operation of dual slope ADC for three different input voltages.

Comparing the voltage of V_x for the both two steps, it can be given:

$$V_{in} = V_{ref} \frac{T_2}{T_1} \qquad (5.13)$$

Since T_1 is a fixed value, the digital output is then determined by T_2. The result is not dependent on the linearity of the integrator, so that very high accuracy can be achieved.

The disadvantage of this system is the conversion speed (proportional to 2^N). It takes a long time if a high resolution is required. For example, for a 12-bit ADC, the full conversion period is about 41 μs even if the conversion time for each bit is only 10 ns.

B. Successive approximation ADC

As shown in Fig. 5.10, The N-bit successive approximation ADC consists in general a sample-and-hold (S/H) circuit, a comparator, a Successive-Approximation-Register (SAR) and a Digital-to-Analog Converter (DAC).

The S/H circuit is used to convert the continuous time input signal into a discrete time signal. The SAR and the DAC work together. The DAC converts the corresponding digital code, which is dependent on the response of the comparator, into a certain analog signal. This signal will be compared with the input signal and the difference between these two signal decides the following operation of the ADC.

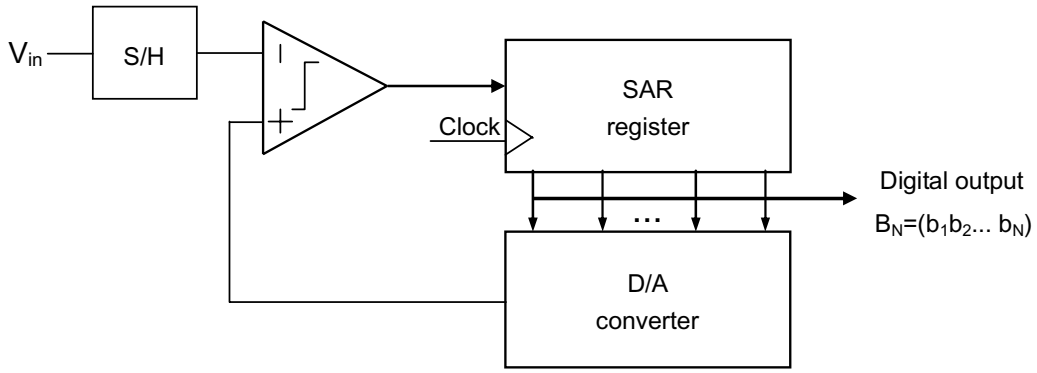


Figure 5.10: Basic architecture of an N-bit successive approximation ADC.

The SAR applies a binary search algorithm to determine the closest digital code to match the input signal. Instead of counting up in binary sequence, the SAR counts by trying all values of bits starting with the Most-Significant Bit (MSB) and finishing at the Least-Significant Bit (LSB). In the first period, after the initial reset, the MSB is "1" and the others are "0". The corresponding output of the DAC is compared with the analog input. If the DAC's value is bigger, the MSB remains at "1", otherwise it is set to "0". The MSB is then determined. The same process is repeated bit by bit until the LSB. Throughout the count process, the register monitors the comparator's output to see if the binary count is less than or greater than the analog signal input, adjusting the bit values accordingly. An example of the searching process is shown in the following figure.

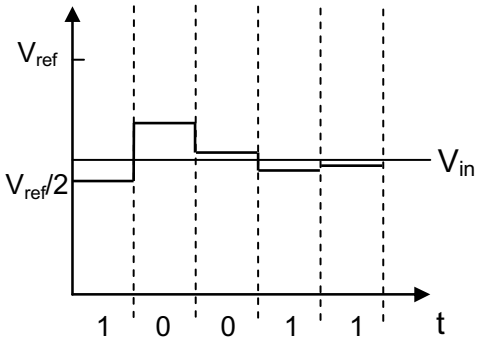


Figure 5.11: Example of the searching process of a 5-bit successive approximation ADC. The digital output (from MSB to LSB) is 10011.

One advantage to this counting strategy is to obtain much faster results: the DAC output converges on the analog signal input in much larger steps than with the 0-to-full count sequence of a regular counter. Another advantage is the reduced power consumption since this structure needs only one comparator.

A key design point of this type of ADC is that it requires a high precision and fast response DAC. Generally realized using charge redistribution capacitor array, the nonlinearity and the size of the array are the issues for achieving a high resolution component.

C. Flash ADC

The flash ADC is the fastest ADC. It converts the analog input signal in a signal pulse. The general architecture is as follows.

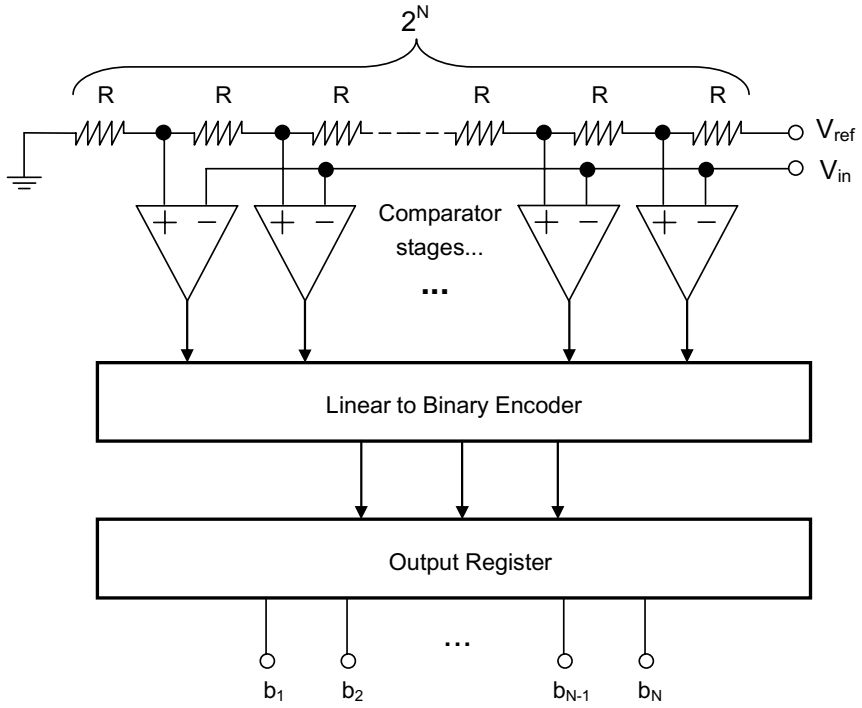


Figure 5.12: Architecture of an N-bit flash ADC.

For an N-bit flash ADC, 2^N resistors are required in order to generate 2^N reference values. The input signal compares directly with these reference value and the outputs of comparators are then coded by a binary encoder to generate N-bit digital code corresponding to the input signal.

The flash ADC is the fastest ADC architecture as its speed is limited only in comparator and gate propagation delay. But it is not practical to use this structure to realize high resolution system because the number of comparators and resistors it needs is 2^N. The modified architecture is proposed, such the pipeline ADC, which realize the full conversion in several steps. Each step is in fact an ADC of low resolution (2-bit for example). An important issue is that the dispersion is very hard to control. In addition, the power dissipation is another issue as the large number of comparators used in this type of ADC.

5.1.4 Design of the column level ADC

In order to increase the spatial resolution of the sensors, an ADC is needed to replace the actual column level discriminator. Evaluating the ADC needed for the vertex detector

application and considering the requirements (size, speed and power dissipation), the ADC of the successive approximation architecture is chosen and prototypes have been realized and tested.

A. Choice of the ADC architecture

The number of the bits needed for this ADC was studied by the IPHC group. The studies were based on experimental analog data taken from two chips: MIMOSA 1 and MIMOSA 2.

Firstly, analysis software was used to simulate the work of an N-bit ADC: the analog data was digitalized and encoded using N bits. It could be considered as a virtual ADC with N bits resolution. This process is repeated for the different N values, so that the digital data of the different virtual ADCs was obtained.

Then, for each virtual ADC, the corresponding digital data was used to calculate the spatial resolution. A curve representing the relationship between the spatial resolution and the number of ADC bits was obtained. It is shown in Fig. 5.13 [2].

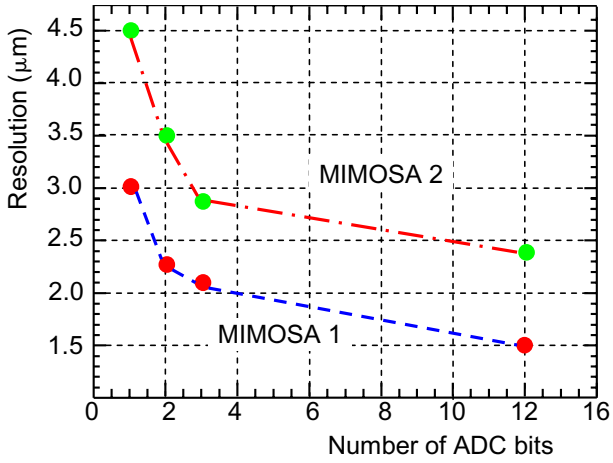


Figure: 5.13 Spatial resolutions vs. number of ADC bits [2].

Similar behaviours were obtained for the data taken from the two chips. With the growth of the number of ADC bits (from 1-bit to 3-bit), the spatial resolution decreases rapidly at the beginning. For example, for MIMOSA 2, the spatial resolution was 4.5 μm for a 1-bit ADC and was improved to 2.9 μm for a 3-bit ADC. But the improvement of the spatial resolution slows down for the ADCs with higher resolution. It can be seen that although the number of ADC bits triples (from 4-bit to 12-bit), the spatial resolution decreases only of 0.5 μm. As a result, when the pitch of pixel is around 20 μm (the size of the pixels used in MIMOSA 1 and MIMOSA 2), an ADC with 5 or 6 bits of resolution should be enough to fulfil the spatial resolution requirement needed for the vertex detector (1.5 μm).

After determination of the number of ADC bits, the next step is to choose an appropriate architecture. The specifications of the sensors which should be used in the ILC vertex detector require a very compact and rapid design with minimum power dissipation. Moreover,

the need for auto sparsification requires that the column level ADC should be the type of auto-offset compensation, making the design task a real challenge.

The flash ADC, which theoretically completes the conversion in one clock duration, can fulfil the speed requirement. The problem is its dimension and the power dissipation. For a full flash structure, the number of the comparators needed is 2^N-1. It means that 31 comparators are needed in order to realize a 5-bit ADC. It is very difficult to integrate all these comparators inside an only 20 µm width area and the consumption is obviously very large. Instead of one step of conversion, some modified structures use multistage conversion and reduce the number of comparators needed. But the auto-offset cancellation becomes a big issue because each single stage contributes with its own offset to the whole system. Special care is needed to cancel all these offsets during the operation of each stage.

The slope ADC has generally very good linearity. As the slope generation circuit and the counter are common to all the columns, only one comparator is needed for each column. Thus, it does not suffer from the area and the power issue. The conversion time is proportional to 2^N. For a 5-bit ADC, a full conversion needs 32 comparisons and the conversion time is $32 \times t_{clk}$ (t_{clk} is the duration of one clock). It seems that this duration is rather short if a small t_{clk} is used. However, due to the need of the auto-offset cancellation, the comparator needs tens of ns for completing the signal comparison and making the decision. Therefore, it is difficult to achieve a high conversion speed.

The structure of successive approximation is finally chosen for our design. Only one comparator is needed for this structure and five comparisons are already enough for the complete conversion of a 5-bit ADC. As a result, the power dissipation and the speed requirements can be fulfilled. According to the size of this architecture, the capacitor array needed can be successfully implemented in a space of 25 µm width because only moderate resolution (5 bits) is required.

B. ADC design detail

According to the requirements of the vertex detector for the future ILC, the column level ADC should be a very compact component (the width is only about 20 µm) operating at high speed (tens of µs for a full conversion of one frame) with small power dissipation (µW level). Moreover, the residual offset and noise of this ADC should be reduced to a very little level because of the small input signal, so that extra offset-cancellation and noise-reduction operation are mandatory. Considering all the above requirements, the architecture of successive approximation is chosen based on the following reasons:

- The conversion time is moderate: N x T_{clk}, where *N* is the number of the bits and T_{clk} is the duration of one working clock. Thus, for a 5-bit ADC working at 100 MHz, the conversion time is theoretically 50 ns, largely sufficient for the ILC application.

- The SAR uses a chip level main control logic which generates the necessary timing chronogram. It is common to all the columns. The column level part circuit consists of one comparator, a capacitor array and simple control logic. The capacitor array could be realized by poly-poly type capacitor in order to obtain better linearity. The size of the capacitor array is not an issue in that the resolution required is 5 bits.

- The column level power consumption of this type of architecture is perhaps the lowest because there is only one comparator per column and the logic components consume only when their states change.

- The offset-cancellation mechanism is feasible thanks to its simple architecture. The mismatch of the comparator could be sampled and cancelled by proper operation.

A prototype based on the successive approximation architecture has been designed in CEA/Saclay in 2006 [3]. The proposed global ADC structure is shown in Fig. 5.14:

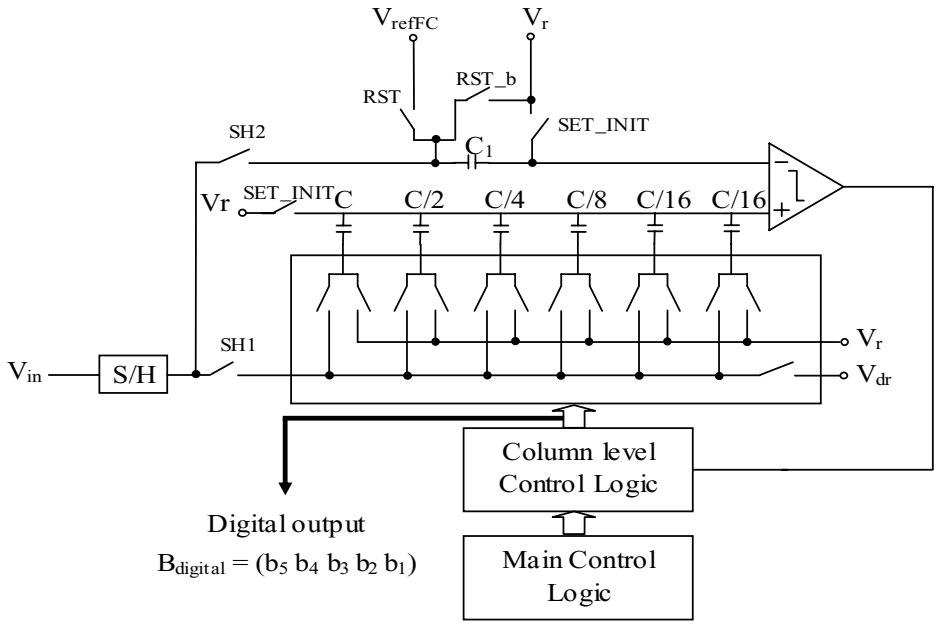


Figure 5.14: Architecture of the column level 5-bit ADC realized in the successive approximation architecture.

A sample-and-hold block (S/H) is used to maintain the input signal level for every comparison. A capacitor DAC based on the "charge redistribution" principle [4] is implemented. It consists of binary weight capacitors which correspond to the different bits of the output digital code. The total capacitance of the array is 2C. A switch array is connected to the capacitor DAC to select the different terminals to the capacitors during the conversion. A main control logic is common to all the columns, sending the global commands (reset, initial, serial switch select signal). A column level control logic is integrated inside each column, transferring the serial switch select signal and making the decision (maintain the "on" signal for the switch on or turn it off) according to the response of the column level

comparator. The column level comparator presented in Chapter 2 is used here. It compares the input signal with the reference value generated by the capacitor DAC and gives the logic response to the logic control part for making the decision. The different types of offset, stemming from the input signal, the reference value and the comparator, are compensated thanks to the offset-cancellation structure used inside the comparator [5] [6]. In the diagram of Fig. 5.14, V_r is the common mode of the comparator's input, V_{dr} defines the full dynamic range (the resolution is $(V_{dr} - V_r) / 32$) and V_{refFC} is an external threshold value used for triggering the conversion. An adjustable voltage is applied on V_{dr}, so that different dynamic range is obtained.

The timing diagram is shown below.

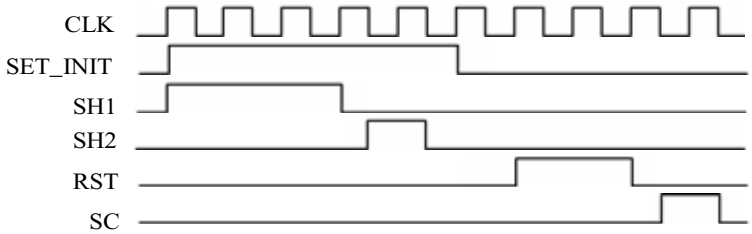


Figure 5.15: Timing diagram (not scaled) for the column level 5-bit successive approximation ADC.

The operation proceeds as follows:

- At first, during the initial phase (SET_INIT), the variable input signal value ($V_{RD} - V_{CALIB}$, as described in Chapter 2) and the offset value of the S/H block will be sampled and stored respectively. The offset value is memorized on the capacitor C_1 and the signal value is memorized on the capacitor array. All the elements of the capacitors array are switched to V_r.

- Then, during the RST phase, the decision comparison will be done. While C_1 is switched to V_{refFC}, the capacitor array is switched to V_r, so that the signal sampled could be compared with the external threshold value given by V_{refFC}. This works as a trigger: only a signal with its value bigger than V_{refFC} could turn on the column ADC. By this way, the ADC is inactive for the small input signals, which will be the most common cases during the working period of ILC, and the power consumption is in turn largely reduced.

- For the signals whose values exceed V_{refFC}, the ADC process turns on and C_1 is switched back to V_r. The conversion is based on charge redistribution principle. Beginning with SC, the MSB is firstly switched to V_{dr} while the others are still connected to V_r (simplified equivalent circuit is shown in Fig. 5.16). Thus, the difference of the two inputs of the comparator is:

$$\frac{(V_{dr} - V_r)}{2} - V_{sig} - V_{off} \qquad (5.14)$$

where V_{sig} is the sampled signal value and V_{off} is the sampled offset value of the S/H block.

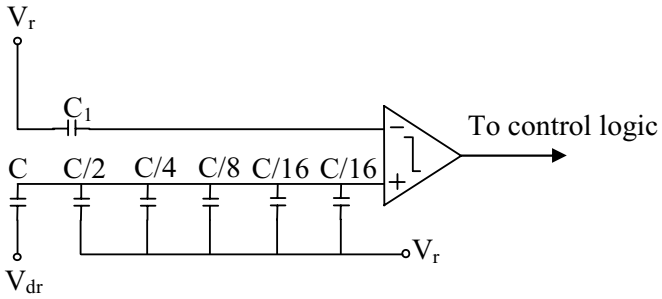


Figure 5.16: Equivalent circuit for the conversion of the MSB.

- The final 5 bits digital value will be ready after five conversions. For the first version, it takes 30 ns for one bit conversion and the time needed for a whole successive approximation process, in addition of initial phase and rest phase, is about 270 ns. For the future, we hope to reach at least the readout speed of the prototypes (MIMOSA 8 and MIMOSA 16) presented in the former chapters (160 ns/row).

a) Design of the sample-and-hold circuit

The column level ADC requires a very low noise and high speed sample-and-hold circuit with low power dissipation. Moreover, the offset-cancellation operation is necessary because of the small input signal. Generally, the sampling techniques are classified in two methods: parallel sampling and series sampling, as presented in Fig. 5.17 below.

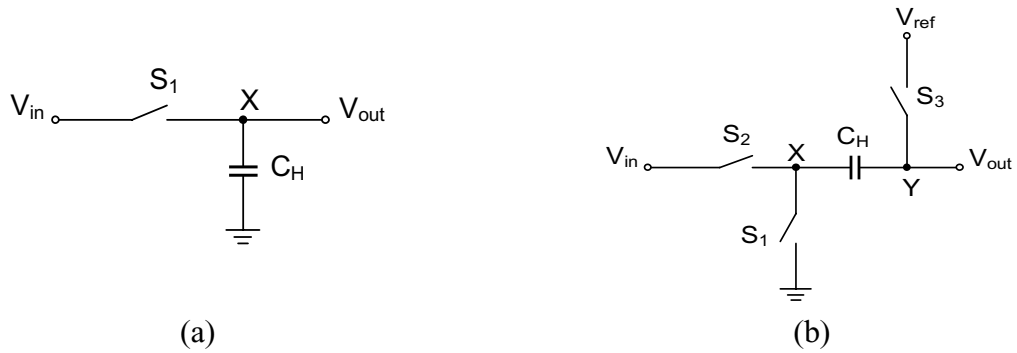


Figure 5.17: Two general sampling structure: (a) parallel sampling, (b) series sampling

In the parallel sampling structure, S_1 is turned on in the acquisition mode for the voltage of the sampling capacitor C_H (point X) to track the level of the input signal and it is turned off in the transition to the hold mode. The input and the output are DC-coupled.

In the series sampling structure, S_2 and S_3 are on and S_1 is off during the acquisition mode, sampling the input signal on C_H. When entering the hold mode, S_3 is firstly turned off and then S_2 is subsequently turned off while S_1 is turned on at the same moment, producing a voltage change equal to the input voltage at the output.

The series sampling structure offers several advantages over the parallel sampling structure. The input is separated from the output by the series sampling capacitor C_H, so that the common mode levels between the input and the output node is isolated. Another advantage is the charge injection error stems from changing the states of the switches. While in parallel structure, this error is dependent on the input signal, it is a fixed value in the series structure because S_3 turns off before S_2, injecting the same value of charge due to the constant gate voltage applied on S_3. Additionally, C_H has little effect of the clock feed through behaviour in the series structure as in the parallel structure [7] [8].

Comparing to the parallel structure, the disadvantages of the series structure is that it suffers from the nonlinearity due to the parasitic capacitance at point Y and that the hold settling time is longer than in parallel structure because the output node is reset to a reference level for every sample [7] [8].

Both of two sampling techniques are implemented in our design. The parallel sampling capacitor C_{H2} is used for fast sampling the pixel signal and memorizing the pixel output reference level. The series capacitor C_{H1} makes the common mode of the following circuit independent from the pixel output level. As presented in previous chapter, the pixel has to be sampled twice: one is the RD signal, representing the signal level of the detected charged particle ($V_{RD} = V_{signal} + V_{ref}$), and the other is the CALIB signal, which is the reference level of the output SF of the pixel ($V_{CALIB} = V_{ref}$). Thus, two capacitors are used inside the S/H circuit in order to memorize these two samples. The circuit schematic is as follows.

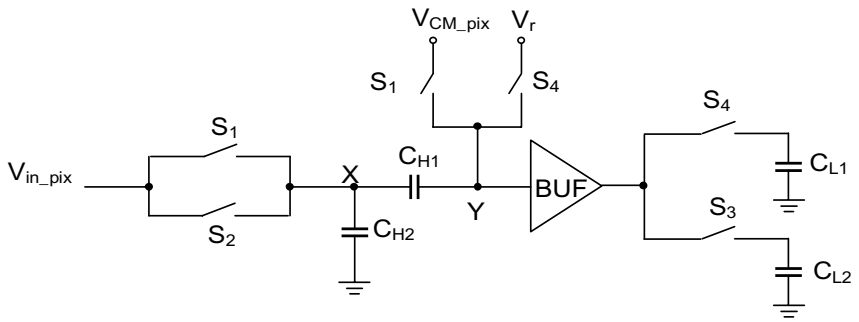


Figure 5.18: Schematic of the S/H circuit.

The circuit consists of four capacitors, an output buffer (BUF) and six switches. V_{CM_pix} is the output common level of the pixel and V_r is the input reference level of BUF. The capacitors C_{H1} and C_{H2} are used to sample and store the pixel output signal, where the offset of the pixel output reference level is suppressed. The buffer provides a current of 20 µA to charge the output capacitors. The current value determines the speed of the signal transfer. The

capacitors C_{L1} and C_{L2} are equivalent to the capacitor array and the C1 in Fig.5.19. The timing diagram is shown below:

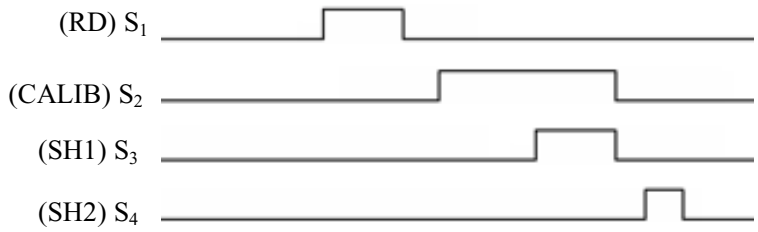


Figure 5.19: Timing diagram (not scaled) of the S/H circuit.

The switches S_1 and S_2 use the signals RD and CALIB which are used inside the discriminators (Ch. 2). The switch S_3 and S_4 are controlled by signal SH1 and SH2 of the ADC, respectively. While S_3 stores output pixel signal on C_{L2}, S_4 memorizes the offset level of the BUF on C_{L1}. The detailed procedure is following:

- The first step is pixel output signal (V_{RD}: signal of charged particle with offset) sampling. S_1 is turned on and the pixel output signal is sampled both on C_{H1} and C_{H2}. After turning off S_1, C_{H1} and C_{H2} are disconnected from the pixel and the signal level is memorized on both of the capacitors.

- The second step is pixel output reference level (V_{CALIB}) sampling. S_2 is firstly turned on, making the voltage of point X equal to the pixel output value. C_{H2} tracks the voltage change of the point X but the signal stored previously on C_{H1} is not damaged because the switch S_1 is turned off, making C_{H1} unconnected. Therefore, at this moment, the voltage of point Y is the difference of the previous sampled value and the pixel output reference level: V_{CALIB} - V_{RD}, which is the useful signal free of offset due to the pixel output stage. S_3 is activated and the useful signal, slightly attenuated by the gain of the voltage buffer (BUF), is stored on C_{L2}. S_3 is turned on after S_2 in order to wait for the end of the above auto-offset cancellation operation.

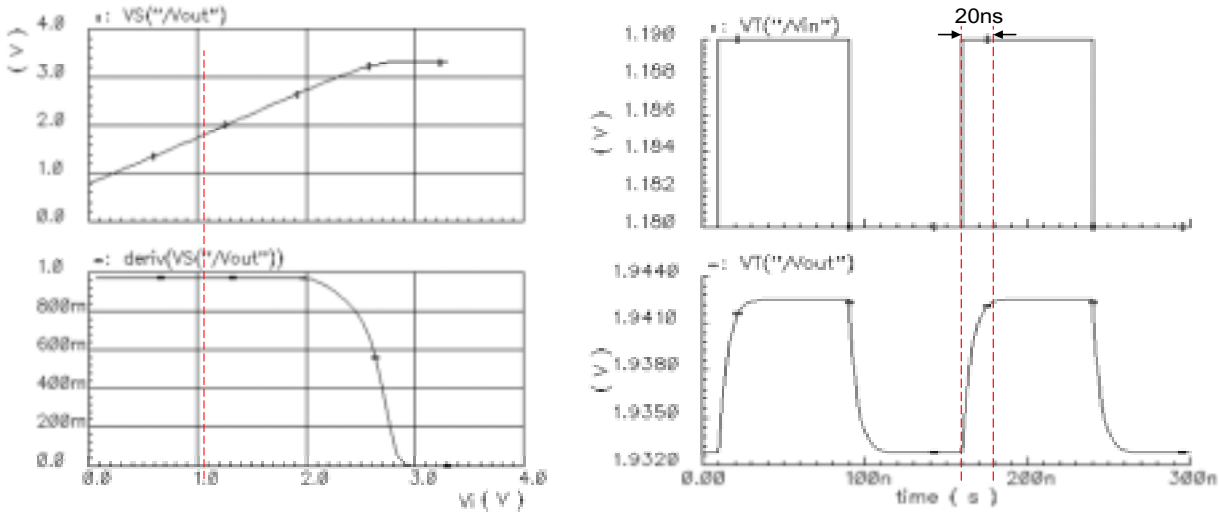
- At the third step, S_4 is turned on, sampling and storing the offset value of the BUF on C_{L1}.

After all the above three step, the difference between the samples stored on C_{L1} and C_{L2} is the real signal value of the charged particle, V_{CALIB} - V_{RD}, free from all the offsets (both the pixel output stage and the S/H circuit).

A PMOS buffer is chosen in order to achieve a unit gain. Simulated in CADENCE environment, the characteristics of the BUF is given in Fig. 5.20.

According to Fig. 5.20 (a), the static gain is about 0.97 when the input common mode is around 1.1 V. Using this common mode value, its transient response with an input signal equalling to 10 mV is shown in Fig. 5.20 (b). The output load capacitor is set to a value of 1 pF, equal to the total capacitance of the following 5-bit capacitor DAC. The settling time is

about 20 ns for this 10 mV input signal, compatible with the time of RD and CALIB signals used for reading out the pixel information (30 ns and 40 ns respectively, Chapter2).



(a) Static response (b) Transient response

Figure 5.20: Static and dynamic characteristics of the BUF in sample-and-hold circuit. Obtained with an output load capacitor equals to 1 pF.

The column level readout offset can be cancelled thanks to the series capacitor C_{H1}. In order to verify the offset cancellation efficiency, an offset with the range from -50 mV to 50 mV has been added on the input signal in the simulation. The results are shown in Fig. 5.21 below:

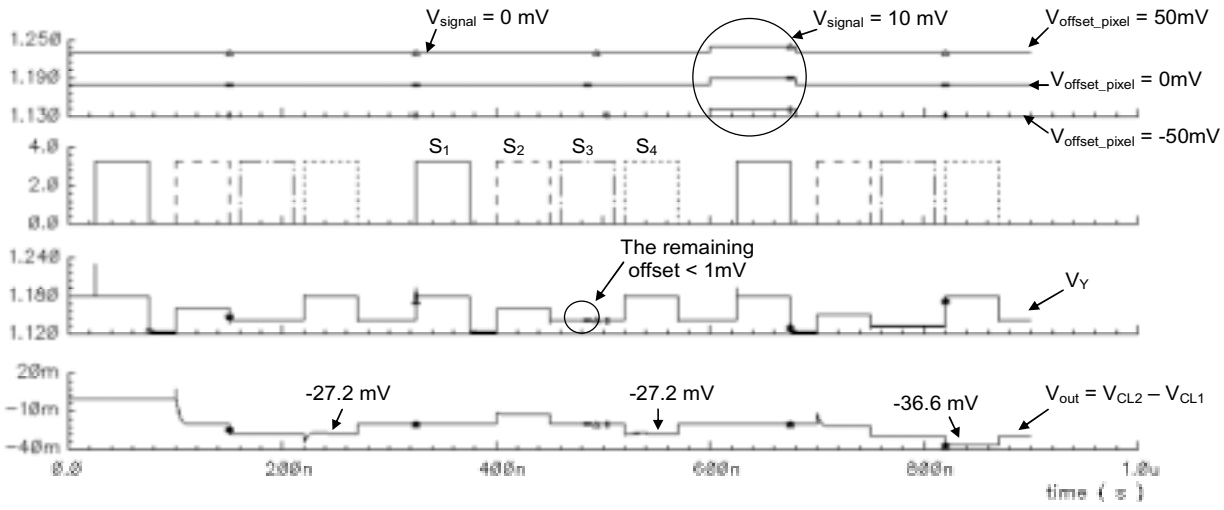


Figure 5.21: Cancellation of the offset of the column level readout circuit.

We notice from Fig. 5.21 that the remaining offset level on the point Y is less than 1 mV during S_3 is turned on. For a signal level of 10 mV, the output signal level after the offset cancellation is:

$$-36.6 - (-27.2) = -9.4 \text{ mV}$$

The negative sign appears since the signal is inversed on the series capacitor C_{H1} during the sampling. This simulation results show that the column level offset is suppressed efficiently.

Another source of offset stems from the source follower BUF. However, this offset is sampled and stored on CL1 when S4 is switched on. As the final output is VCL2 – VCL1, this offset is also cancelled. The following simulation results show that the offsets of the range from -50 mV to 50 mV are successfully cancelled.

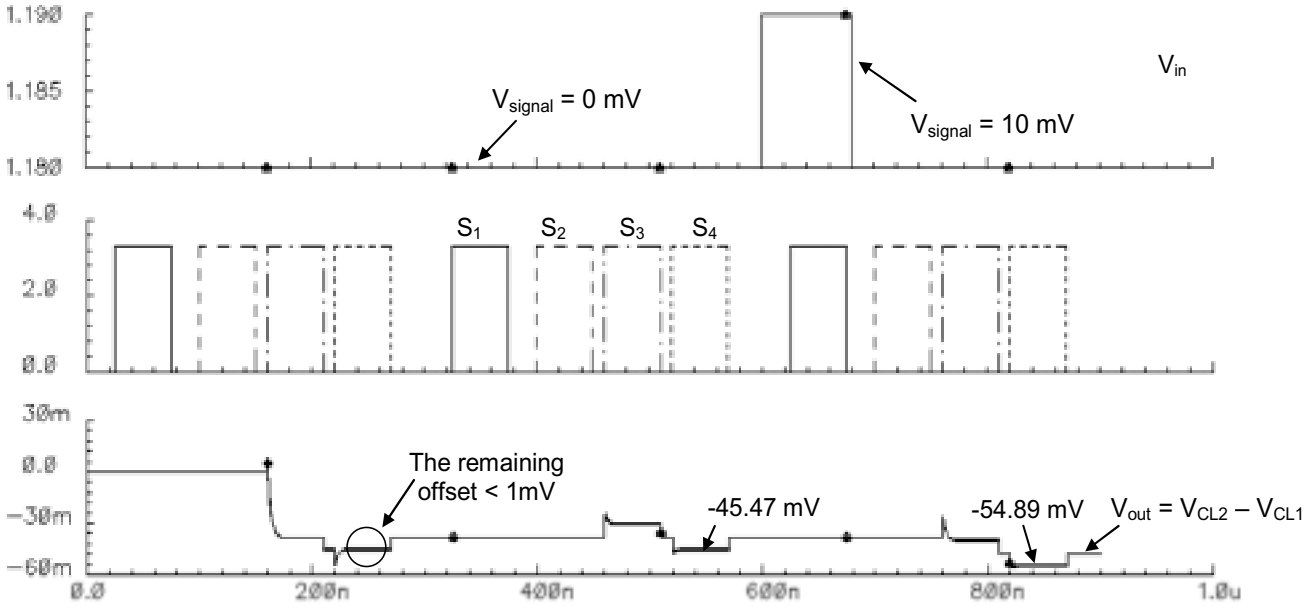


Figure 5.22: Cancellation of the offset of the source follower.

From Fig. 5.22, the final output signal level is:

$$-54.89 - (-45.47) = -9.42 \text{ mV}$$

Similar results as in Fig. 5.21 is obtained here.

The duration of the control signals for S_1, S_2, S_3 and S_4 is 50 ns in simulation. But noticing that the sampling signal and the final output signal is stable within 20 ns, the design requirement of speed is also fulfilled.

b) Design of the control logic

The schematic of the main control logic and the column level control logic are shown in Fig. 5.23 and Fig. 5.24 below, respectively.

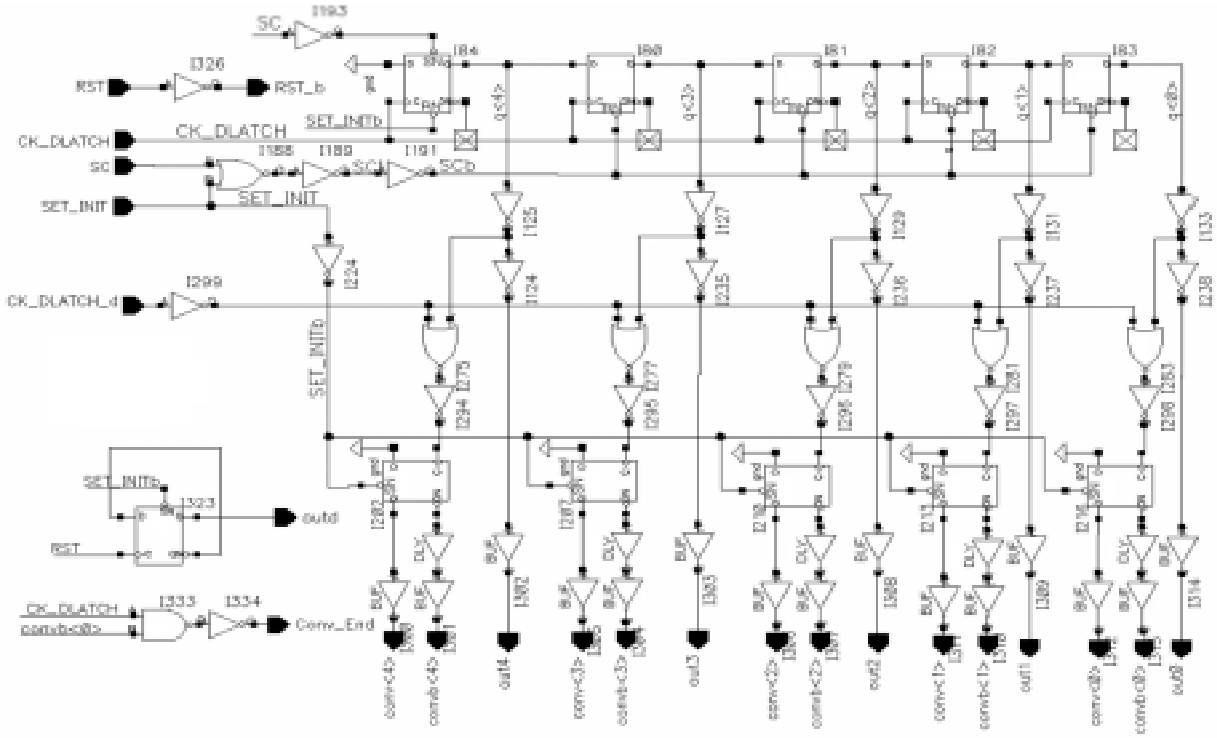


Figure 5.23: Main control logic of the column level 5-bit successive approximation ADC. It is common to all the column ADCs implemented in one chip.

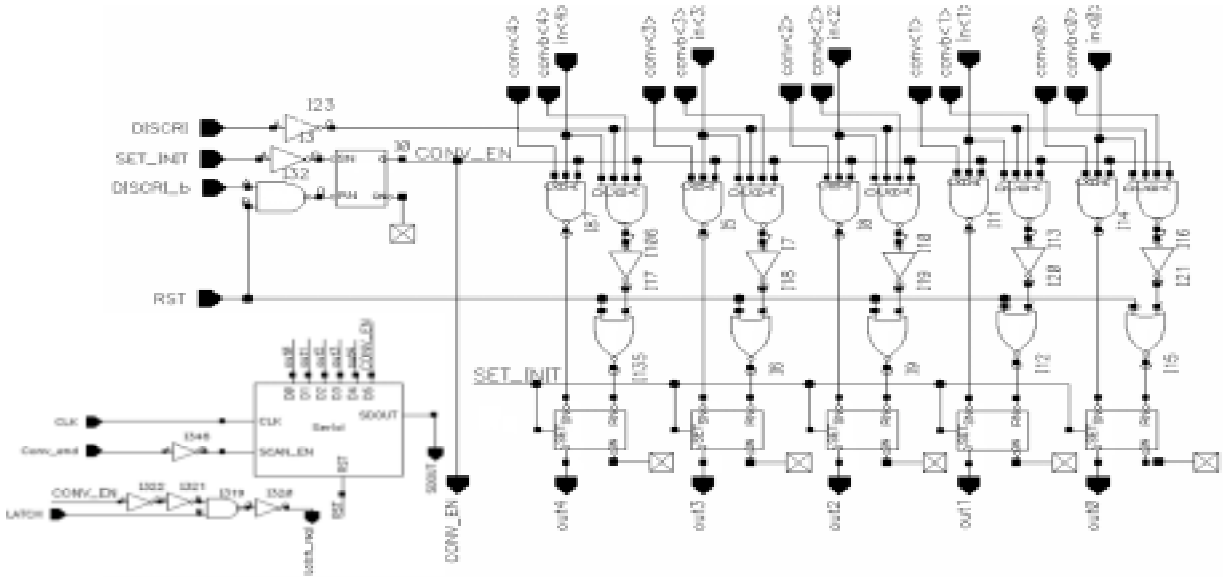


Figure 5.24: Column level control logic of the column level 5-bit successive approximation ADC. It consists of the simple logic gates, consuming only when their output states are change.

The bit selection signals, from out4 to out0 of the main control logic, are connected with the inputs from in<4> to in<0> of the column level control logic, respectively. The signals from out4 to out0 of the column level control logic are connected to the corresponding switches of the different bits inside the capacitor array (see Fig. 5.14). Based on the structure proposed in

[9], the main design features of the control logic for our successive approximation ADC are the following:

- Instead of using one signal, CLOCK, to control the selection and the determination for the conversion, as proposed in [9], the control signals used for selecting the current converting bit and for determining its state are separated. CK_DLATCH is used to drive the shift register in the main control logic to generate the necessary signal to select the current conversion bit. Another input signal, CK_DLATCH_d, is used for generating the signals necessary for setting initially the conversion bit to "1" and for determining its final state according to the comparison result. For example, the output signal "out4" in the main control logic (Fig. 5.23) is used to select the MSB (the bit number is 4). The signal "conv<4>" is used to set the column level output to "1" and the signal "convb<4>" is used to determine the final state of the column level output. "conv<4>" works together with the column level discriminator's output (DISCRI) which is dependent on the result of the comparison. A possible example of the working chronogram is shown in Fig. 5.25.

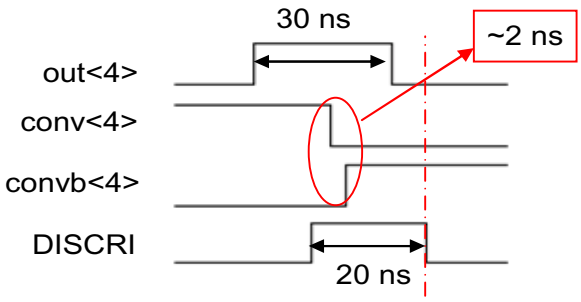


Figure 5.25: An example of the timing diagram for out<4>, conv<4>, convb<4> and DISCRI.

Two important working conditions are demonstrated in Fig. 5.25. The one is that conv<4> and convb<4> are not overlapped otherwise logic conflict will happen. The other is that the output of the comparator (DISCRI) must maintain until the signal out<4> pass to "0". The reason is that any "0" input of the NAND gate (Fig. 5.24) during it is selected will force its output to "1", which will reset the corresponding column level output although the current comparator's output is "1". There are several advantages of using CK_DLATCH and CK_DLATCH_d. The first is that the period during which the current bit is selected ("out4" for bit number 4 in Fig. 5.23) is control by CK_DLATCH, which is fully adjustable by initial programming (the step is 5 ns, half of the clock period when the main clock frequency is 100 MHz). The second, using two different signals make the test conditions less strict. As these two signals are both programmable, we can adjust them easily by changing the initial pattern of the ADC. For example, the signal conflict due to the undesirable transfer delay can be avoided by this way.

- A fixed delay unit, DLY, is added in order to avoid the overlap of the signals for initial setting and for final determination. For example, the signal "convb<4>" is always behind the signal "conv<4>" because "convb<4>" has to pass through the DLY unit. It is very

delicate to choose an appropriate delay value. For our application, a well controlled delay is desirable in order to achieve fastest conversion speed without the logic conflict. In fact, in high speed integrated mixed circuits, the delay is influenced by numerous source of disturbance: the difference of the clock caused by the transfer bus, the difference between the interconnection of the components and the delay between the columns caused by the signal transfer. For this reason, the delay value is very difficult to be preciously modelled in the simulator. Thus, great effort has been made in order to find out a compromise value. At last, we set the delay added by DLY unit to about 2 ns.

- A conversion enable option is added in order to make the column control logic inactive when the input signal of the ADC is less than V_{refFC}. The compensate output of the comparator (DISCRI_b, Fig. 5.24) is used for this purpose. If the input signal level is less than V_{refFC}, DISCRI_b is "1". The signal CONV_EN is firstly set to "1" by SET_INIT at the beginning of conversion. Then, using an NAND gate, DISCRI_b is allow to reset the signal CONV_EN when RST is activated. A CONV_EN of "0" state blocks the logic inputs of all the bits (Fig. 5.24).

- A serialization block is used inside the column level logic (Fig. 5.24). When the conversion is done, a signal called CONV_END is generated by main control logic (Fig. 5.23). After receiving this signal, the CONV_EN signal and the conversion results (the states of the bits) are serialized using the main clock frequency (100 MHz in our design). The output code begins with the CONV_EN signal, which is applied for synchronization with the data acquisition system.

- The input signals of the control logic (SET_INIT, RST, SC, CK_DLATCH and CK_DLATCH_d) are loaded during the initial programming phase of the ADC. A sequencer is designed for this purpose. An external high speed CMOS signal is used for clocking the chip (max f_{CK} is 100 MHz).

The simulation results are shown in Fig. 5.26.

In the simulation shown in Fig. 5.26, the full dynamic range V_{dr} of the ADC is set to 32 mV and V_{refFC} is set to 0 mV. Three conversions with different input signals are simulated. The signal SDOUT is the serialized output code. For the first conversion, an input signal of 22.5 mV is applied. The input signal is set to 0.5 mV for the second one and is set to -0.5 mV for the last one. The first output code is 10110 (in Fig. 5.26, the first pulse is the CONV_EN signal), corresponding to 22 mV within a dynamic range of 32 mV. The second output code is 00000. Only the synchronisation signal is visible. For the last input signal, DISCRI_b is activated during RST period (the first conversion) and the CONV_EN is set to "0" by this signal. As a result, no output is produced for this input signal. The simulation results show the good working behaviour of the ADC.

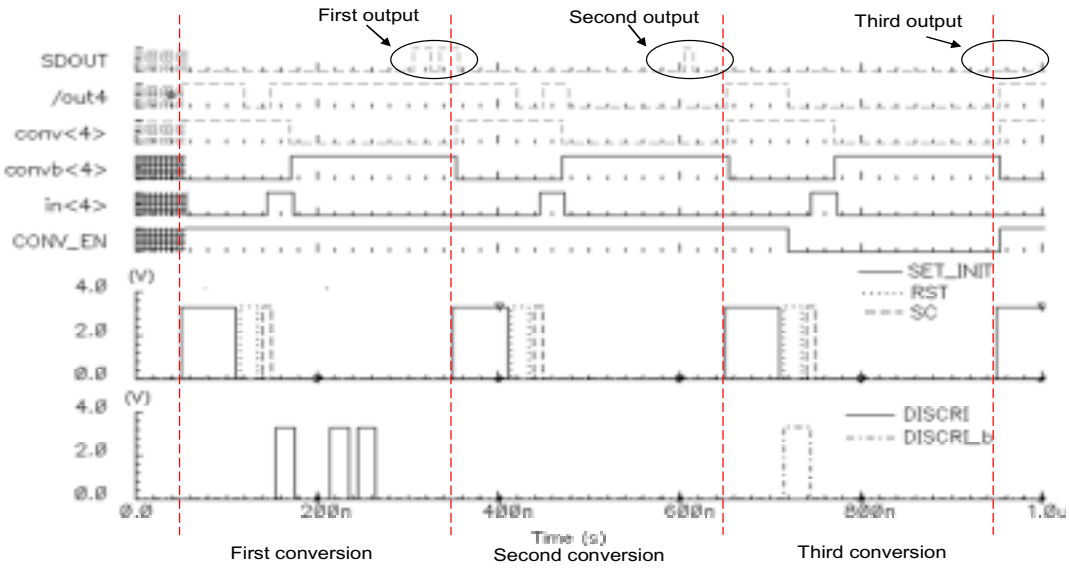


Figure 5.26: Simulation of the ADC conversion.

Fabricated in AMS 0.35 µm OPTO CMOS process, the layout of the prototype is shown in Fig. 5.27 below:

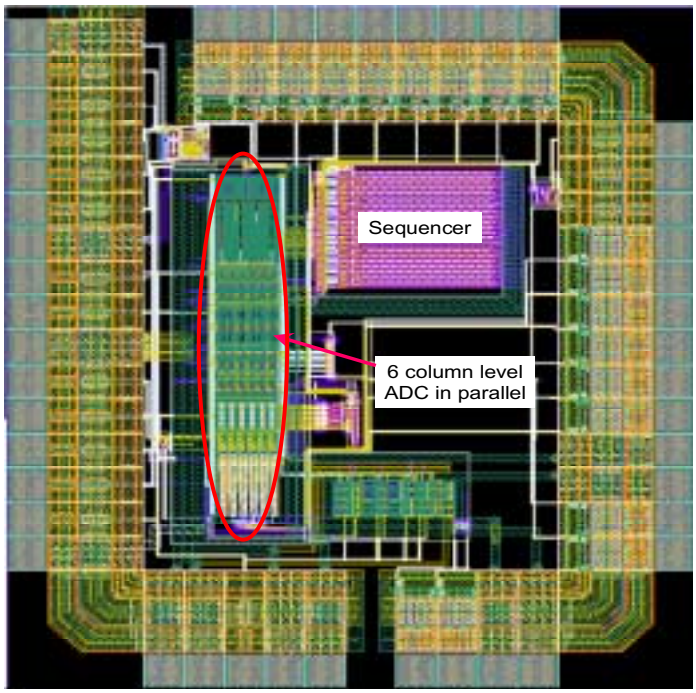


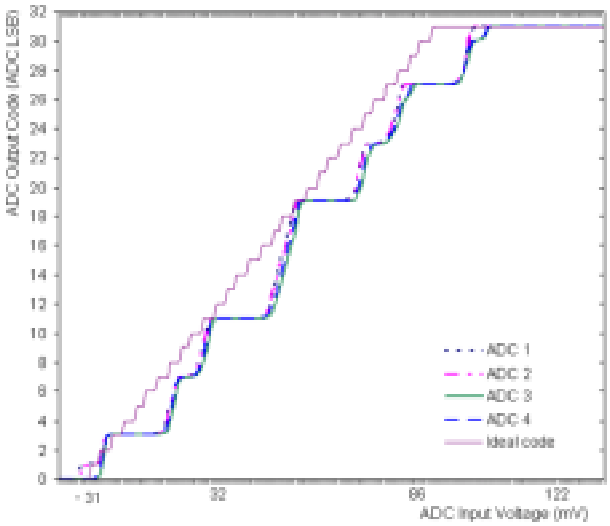
Figure 5.27: Layout of the prototype of the ADC.

Six column level ADCs have been fabricated in parallel. On each side, two of them are used as the guard structure. The four ADCs in the middle are tested. The dimension of the ADC is 25 µm x 1 mm. The conversion period is about 250 ns at 100 MHz and the static power consumption is about 300 µW.

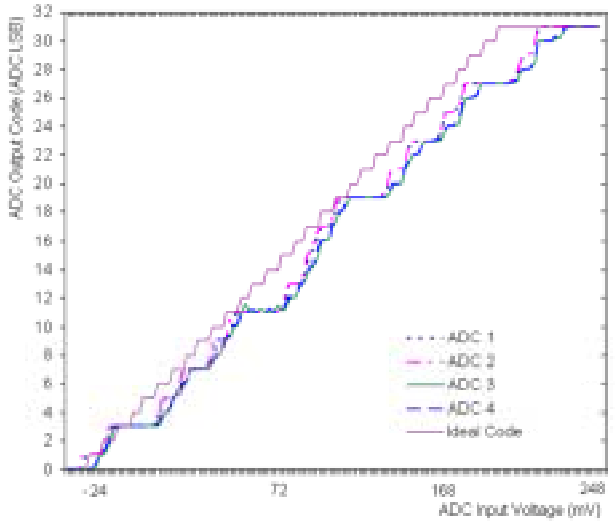
5.1.5 Test of the column level ADC

The ADC prototype has been tested using a USB test bench based on a 16-bit 4-channel DAC with a 75 µV LSB. The DAC channels are used as quasi-static inputs for the ADC prototype.

The transfer function of the ADC is tested firstly. The frequency is set to 100 MHz. Four different ADCs from different chips are tested. Thanks to the adjustable V_{dr}, the results for two different dynamic ranges are shown in Fig. 5.28.



(a) Transfer function for the resolution of 3.75 mV



(b) Transfer function for the resolution of 7.5 mV

Figure 5.28: Transfer functions of the four different ADCs for the resolution of: (a) 3.75 mV and (b) 7.5 mV.

The examples given in Fig. 5.28 display the transfer function for the resolutions of 3.75 mV and 7.5 mV, respectively. The corresponding dynamic range is 120 mV and 240 mV. During the test, we notice that it is hard to achieve a resolution of 1 mV (the corresponding dynamic

range is 32 mV) as we simulated in CADENCE environment. Thus, we applied much larger resolution value in the tests.

After removing the Offset Error, the Gain Error is about 4 LSB and 3 LSB for the resolution of 3.75 mV and 7.5 mV, respectively. The dispersion for the outputs of the four ADCs from the different chip is smaller than 1 LSB for the two different resolutions. However, the transfer function is very poor (Fig. 5.28(a)), when the resolution is 3.75 mV, almost half of the output codes are lost.

Doubling the resolution value improves the quality of the transfer function (Fig. 5.28(b)). Observing the behaviour of ADC 4 (Fig. 5.29), we see that almost all the output codes are visible. But the linearity is still not good. Continuing improving the resolution values changes slightly the transfer function's quality.

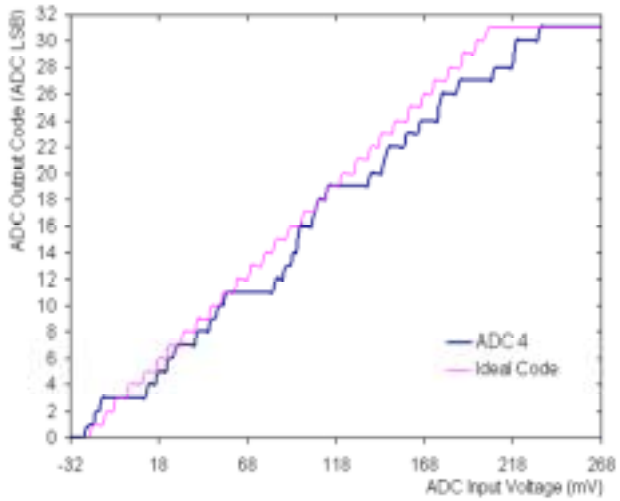


Figure 5.29: Transfer functions of ADC 4 for the resolution of 7.5 mV.

If we carefully observe the transfer function, we notice that the linearity disturbs mostly when following codes appear: 3, 7, 11, 19, 23 and 27. A common point for all these codes is that they all end with "11". For the three last bits, the step from "011" to "100" is much longer than the others. Several possibilities can cause these phenomena. Considering the nonlinearity of the capacitor array, one of the possibilities is that this error influences the capacitor (C/8) which makes the decision of the third bit. Therefore, a larger input signal is needed to compensate this error. Another possibility is that the injected charge on the capacitors accumulates during the charge redistribution process. Thus, it becomes an important value for the last two bits. In fact, the capacitance of the last two bits is rather small: only 25 fF for LSB and 50 fF for the 4^{th} bit. As a result, the offset due to the injected charge becomes quite important and it is big enough to affect the decision of the last two bits, which is rather a small value: 1/8 of the dynamic range value V_{dr}. Then, a large input signal level is also needed to compensate this offset. Moreover, the parasitic capacitance around the last two small capacitors perturbs also their linearity, affecting the charge redistribution process. There is also a possibility due to the very compact structure of the ADC. The bus connected

to this bit suffers from the coupling of the other signals. In this case, the layout of the circuit needs to be carefully modified which is not an easy task because we have only 4 levels of metal which are available for interconnection in AMS 0.35 µm OPTO process.

A new prototype with modified structure of the capacitor array is fabricated at the end of the year 2006. The new capacitor array is given in Fig. 5.30.

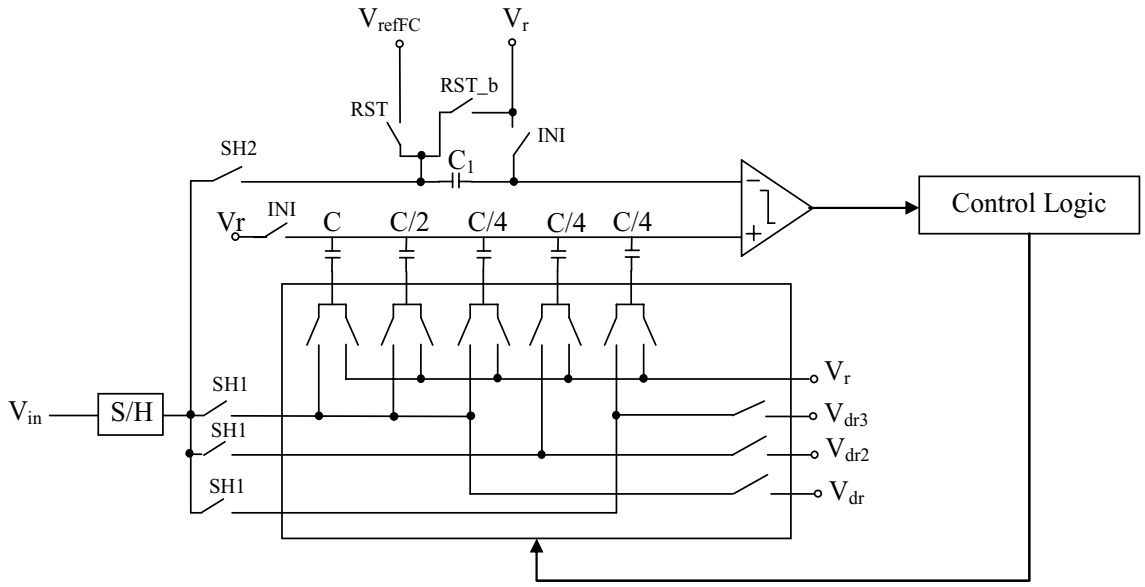


Figure 5.30: Architecture of the column level 5-bit ADC with modified capacitor array.

In Fig. 5.30, instead of using one dynamic reference voltage V_{dr}, three reference voltages with proportional values are applied. While the first three bits use V_{dr}, the last two bits use V_{dr2} and V_{dr3} separately. The capacitors for the last two bits have the same value as the third bit: C/4. Two versions are designed with capacitors of different values. For the first one, C/4 is equal to 100 fF, so that the total capacitance of the capacitor array is 900 fF. For the second, C/4 is set to 50 fF so that the total capacitance is 450 fF. The proportional charge redistribution realised by the different capacitor values in the former architecture (5.14) can be achieve here by applying different dynamic reference voltage with the same values of the capacitors. All the three dynamic reference voltages are independent and are adjustable externally. This structure gives us the access to compensate the undesirable errors bit by bit.

Only the version that the total capacitance equals to 900 fF is measured. The tests begin with a resolution of 7.5 mV, corresponding to a dynamic range of 240 mV, at 100 MHz. In Fig. 5.31, the example gives results when V_{dr2} is equal to 70 mV and V_{dr3} is equal to 30 mV. We notice that the linearity is better than the results shown in Fig. 5.29.

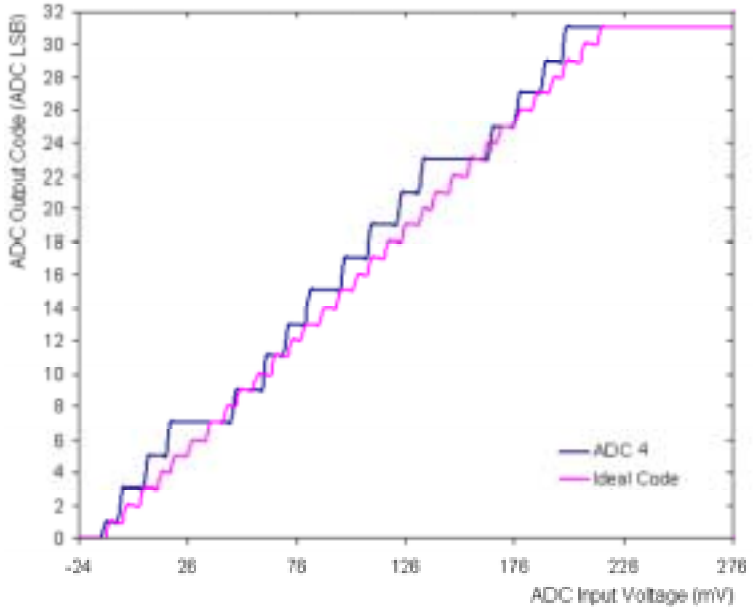


Figure 5.31: Transfer functions of ADC 4 for the resolution of 7.5 mV. Tested at 100 MHz. The corresponding dynamic range is 240 mV with V_{dr2} = 70 mV, V_{dr3} = 30 mV.

While the new structure improves the linearity, we observed that half of the output codes are lost in the transfer function shown in Fig. 5.31. The state "10" of the last two bits is missed. Varying the values of V_{dr2} and V_{dr3}, the problem remains, what indicates that the main reason is not the offset.

It is mainly due to the time needed for the charge redistribution. The high frequency time constant for each capacitor can be estimated by the open-circuit time constant approach [10]. The time constant for each capacitor is given as [11]:

$$T = (R_s + R)C \qquad (5.15)$$

where R_s is the switch-on resistance of all the switches connected with the capacitor, R is the switch-on resistance of the bit line and C is its capacitance.

From the equation 5.15, the capacitor's time constant is proportional to its capacitance, so that the time constant of the new architecture is four time of the original architecture. We repeat the test of 7.5 mV resolution at 50 MHz, maintaining all the other conditions, and every output code is visible. The result is given in Fig. 5.32.

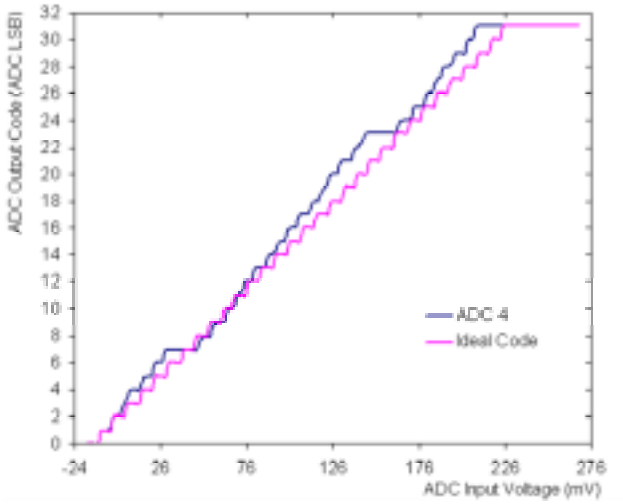


Figure 5.32: Transfer functions of ADC 4 for the resolution of 7.5 mV. Tested at 50 MHz. The corresponding dynamic range is 240 mV with V_{dr2} = 70 mV, V_{dr3} = 30 mV.

From Fig. 5.32, the linearity disturbs mostly when the output code is 7 or 23. Compared to Fig. 5.29, the nonlinearity of the last two bits has been removed thanks to the modifications effectuated in the new structure: the capacitors with large values and the independent external reference voltage applied on them.

5.1.6 Conclusion

A column level ADC has been designed and fabricated in AMS 0.35 µm OPTO process. Considering the requirements of speed, consumption and simplicity, the successive approximation ADC architecture is chosen. Auto-offset cancellation operation is realized inside this column level ADC. Very compact structure is achieved. The size of the ADC is only 25 µm×1 mm.

Preliminary results are obtained for this prototype. The auto-offset cancellation structure integrated inside the ADC works efficiently. With different ADCs from different chips, very small dispersion is observed. The first conversion option works correctly. No conversion will be carried on if the input signal level is less than the preset threshold value. Thanks to the S/H circuit, the ADC works continuously in parallel with the pixel. No interruption is needed for sampling the signal and preparing the conversion. A full conversion period (sampling, conversion and output the 5-bit digital result) is 270 ns for this first prototype. The static consumption is about 300µW. However, because no conversion will perform if the first comparison with a preset threshold value gives a negative result, the power consumption will be largely reduced in the ILC working environment.

However, the resolution observed during the test is largely above the simulation results. We noticed that the original structure with only one reference dynamic range suffers from offsets due to charge injection of the switches and analog-digital coupling. The linearity disturbs due

to the perturbation on the last two bits because of their small capacitance value (25 fF and 50 fF). Thus, a new structure with additional independent reference voltage applied on the last two bits is realized. The value of the capacitors connected with the last two bits is also increased. Two versions with different capacitance are realized. From the tests of the version with a total capacitance of 900 fF, the results show that this new structure improves the linearity of the ADC. But the speed problem appears in that the increasing capacitance increases the time constant of the capacitor for the charge redistribution. The tests for the second version with small capacitance (450 fF) are needed to evaluate profoundly the performance of this new architecture.

For both original and modified structure, the resolution is hard to reach 1 mV. So that an amplifier, replacing current source follower used in S/H circuits, is mandatory for improving the resolution. According to the test results, the gain of the amplifier should be not less than 7 as a resolution of 1 mV is required. With a resolution of 7 mV, the ADC works correctly.

A fully differential structure is efficiently for suppress the offset due to the charge injection. Because the dimension is an issue when two capacitor arrays have to be used in a differential structure, it is worth to study a CMOS process providing more compact devices.

5.2 New column level comparator

The column level comparator previously presented uses a dynamic latch [12]. Although it offers good speed and no-static power dissipation, it absorbs a large amount of current while switching from one state to another, causing a significant and prompt change in the supply voltage which affects the operation of all the front-end stages. Moreover, the latches generally exhibit large input referred offset voltages of several tens of millivolts [13]. The offset of the dynamic latch used in our column level comparator can not be memorized so that an amplifying stage with a gain of a few hundreds is mandatory in order to reduce the input referred residual offset of the comparator (equation 2.14).

A new column level comparator using an offset corrected latch has been designed. Thus, the gain stage in the closed-loop bloc is not necessary for this new comparator. A static latch is used for this purpose, so that the current absorbed is fixed by a current source. The preliminary test results show that very small Temporal Noise and FPN have been achieved by this new comparator.

5.2.1 Design of the new column level comparator

Static and dynamic latches suffer from large input referred offset voltages in standard CMOS technologies. For the application of the vertex detector of the future ILC, the offset of the individual component is crucial in that millions of the sensors are to be integrated together and be processed by the same circuit. For this reason, the latches are often preceded by an

offset compensated amplifying stage. Two stages are used in the structure of the column level comparator presented in Chapter 2 in order to supply a gain of a few hundreds to reduce the input referred offset of the dynamic latch. While the dynamic latch works only when the LATCH command signal is high and it is hard to memorize its offset, the static latch works in two steps, permanently tracking step and latching step, which gives us the possibility to memorize and compensate its offset. Another advantage is that the dynamic consumption of static latch is constant (which is determined by the bias current). For the dynamic latch, its dynamic consumption is hard to predict and a large amount of current is often needed in order to fast develop the output logic level.

Derived from the static latch designed by Song [14], the static latch used in the new column level comparator is given in Fig. 5.34. It works in two operation mode:

- The first one is that the LATCH signal is low. The static latch is in resetting step. The latch operates as a differential amplifier so that the voltages of X and Y track the variation of the input. The transistors M11 and M13 are added in order to adjust the common mode of the equivalent amplifier. The gain of this amplifier is approximately equal to:

$$A_{reset} = \frac{g_{m1,2}}{g_{m9,10} - g_{m7,8}} \qquad (5.16)$$

where $g_{m1,2}$, $g_{m7,8}$ and $g_{m9,10}$ are the transconductances of transistors M1, M2, M7, M8, M9 and M10, respectively.

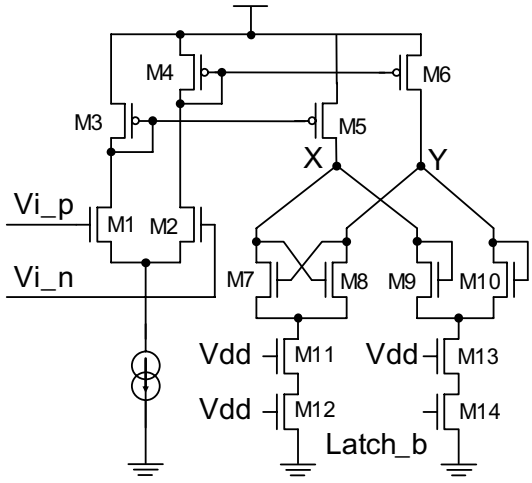


Figure 5.34: Static latch for the new column level comparator.

Under this operation mode, the difference between the inputs Vi_p and Vi_n are pre-amplified and the result is available on the nodes X and Y. This value, when the input

signal is equal to zero, is the offset value which will be memorized for offset cancellation.

- The second is when the LATCH signal is high. Then, the diode connected transistors M9 and M10 turn off and the former amplifier turns into a positive feedback latch thanks to the cross-coupled transistors M7 and M8. Fully developed logic signal is available during the LATCH is high.

The complete schematic of the auto offset cancellation column level comparator is shown in Fig. 5.35. Compared to the column level comparator used in MIMOSA 8 and MIMOSA 16 (Chapter 2), only one stage of the amplifier is used in this new structure. The offset of the static latch is cancelled by OOS techniques. The total input referred residual offset of this comparator is given by:

$$V_{OSR} = \frac{\Delta V_{off,Latch}}{A_0(1+A_{reset})} + \frac{\Delta Q}{A_0 C} \quad (5.17)$$

where A_0 and A_{reset} are the gain of preamplifier and the static latch (during resetting step), $\Delta V_{off,Latch}$ is the offset of the static latch, ΔQ is the total injected charge mismatch from S2 and C is the equivalent series capacitance.

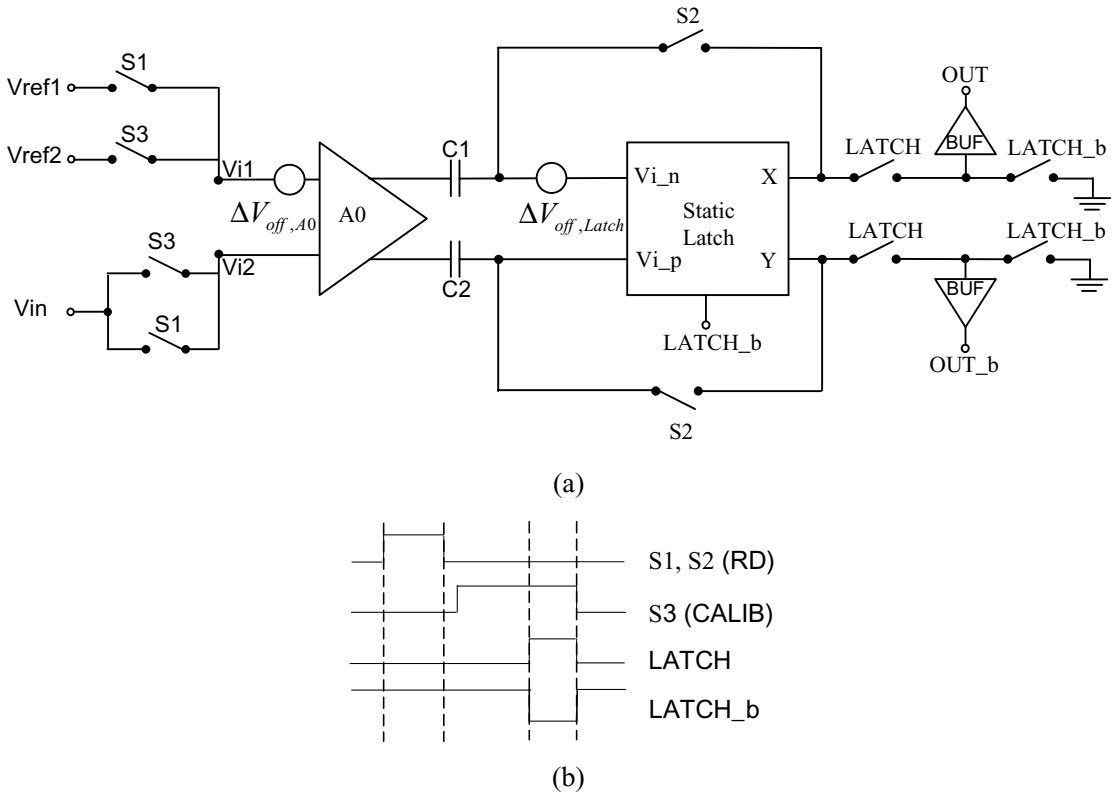


Figure 5.35: New column level comparator: (a) architecture and (b) timing diagram (not scaled).

According to Fig. 5.35 (b), three phases are needed for a complete auto-compensation operation, compatible with the pixel signal readout.

- Firstly, S1 and S2 are switched on, S3 and LATCH are switched off, the residual offset of preamplifier A0 and the static latch is memorized in series capacitors C1 and C2. Because the RD signal, being used for reading the pixel output signal of charged particle (Fig. 2.10), is used to turn on S1 and S2, the pixel output signal is then read out and stored in C1 and C2. Moreover, the pixel output stage offset, which varies from one pixel to another, is also memorized in the series capacitors.

- Secondly, S1 and S2 are switched off, S3 is switched on. The comparator enters its auto-compensation mode. As the CALIB signal is used to control S3, the pixel output stage reference level with its offset value are read out and automatically used to compensate the offset value stored in series capacitors during the first phase. The threshold value (Vref2) of comparator is also sampled during this phase and is compared with the pixel output signal value.

- At last, at the end of the second phase, LATCH activates and makes the static latch to enter its latch step, which rapidly amplifies the difference between the pixel output signal level and the comparator threshold level. A logic signal is output according to the level of difference. If the level is positive, the output of latch is 1; on the contrary, it is 0.

The transient simulation results are shown in Fig. 5.36 below:

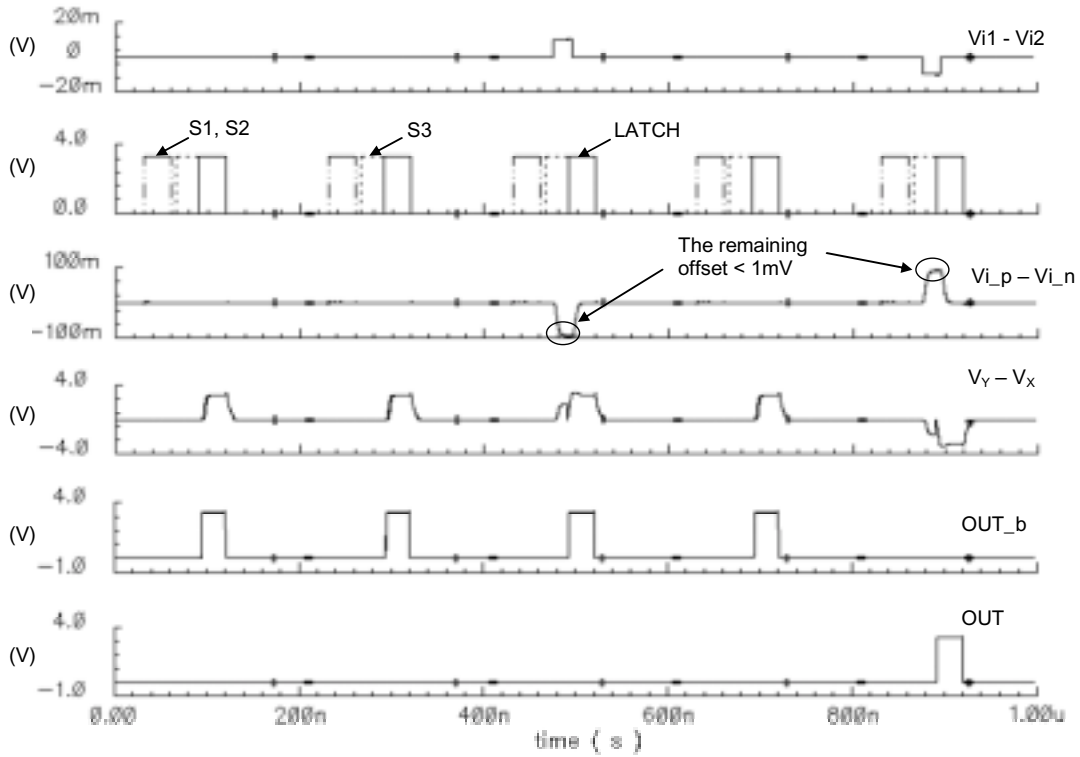


Figure 5.36: Transient simulation of the new column level comparator. The main clock frequency is 100 MHz.

. The main clock frequency is 100 MHz. Vi1 - Vi2 is the differential voltage signal applied to the input of the comparator. An offset of the range from -50 mV to 50 mV is added on the input signal. OUT and OUT_b are the outputs of the comparator. The threshold voltage is set to zero, so that OUT is equal to "1" when the input signal is more important than zero. OUT_b is the complementary output of OUT. During the first two period, no input signal is applied and the state of OUT is "0" (OUT_b is "1" for these periods). Two input values of -10 mV and 10 mV are added consecutively, OUT is "1" when the input signal is equal to 10 mV.

We notice that at the input of the static latch, the remaining offset is less than 1 mV. The bias current of the static latch is 50 µA.

5.2.2 Test of the new column level comparator

The new comparator is fabricated on AMS 0.35 µm OPTO CMOS process. Two prototypes have been fabricated, one is in January 2006 and the other is in November 2006.

Preliminary tests have been performed with the set-up presented in Chapter 3 (Fig. 3.26). Six comparators of different chips have been measured and their transfer curves are shown in Fig. 5.37. Two of the test chips are fabricated in November 2006.

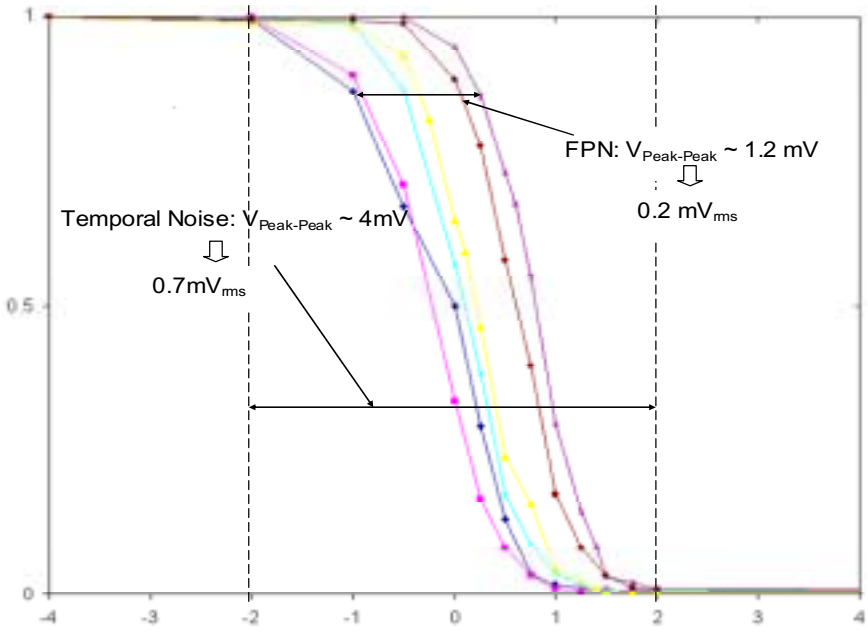


Figure 5.37: Normalised responses of the new column level comparators on six different chips (f_{clock} = 100MHz).

Compared to the corresponding values of the comparator used in MIMOSA 8 and MIMOSA 16, similar performance has been achieved. Considering the results shown in Fig. 5.37, we

can say that the Temporal Noise of the column level comparators is less than 0.7 mV_{rms} and its FPN is about 0.22 mV_{rms}. We can see that the offset due to mismatch and different processes are largely reduced. These excellent results indicate that the integrated offset-cancellation structure works efficiently. However, it is a preliminary result and more statistics is needed for the further studies.

5.3 Conclusion

A 5-bit successive approximation ADC and a new column level comparator have been designed and fabricated on AMS 0.35 µm OPTO CMOS process. Preliminary tests results are obtained.

As the first prototype, the conversion time for each bit is 30 ns. The total ADC conversion period is 270 ns (sampling the pixel signal, auto offset-cancellation and digital data output). The static consumption is about 300 µW. Very small dispersion has been observed for the ADCs tested. However, the ADC suffers from the nonlinearity due to the charge injection of the switches of the capacitor DAC and analog-digital coupling. It is difficult to reach a resolution of 1 mV, what is obtained in simulation. As a result, an amplifier is mandatory for the future design, replacing the current source follower in the S/H circuit. The gain of the amplifier should be at least 7.

A static latch has been implemented inside the new column level comparator. The static latch allows us to sample and memorize its offset during the resetting step, simplifying the structure of the comparator. Thus, only one preamplifier stage is used. Moreover, the dynamic consumption is reduced because the static latch is biased continuously by a 50 µA current source. The preliminary results show that the noise performance is similar to the comparator used in MIMOSA 8 and MIMOSA 16. The Temporal Noise is less than 0.7 mV_{rms} and the FPN is only 0.22 mV_{rms}. However, only six comparators of the different chips have been tested. Further measurements are needed to confirm these results.

Reference

[1] Caso et al., “Particle Data Group : Review of Particle Physics”, Volume 3, Number 1-4, 1998 p. 168.

[2] M. Winter, “R&D on CMOS Sensors for the ILC Vertex Detector,” SOCLE meeting at LAL (Laboratoire de l'Accelérateur Linéaire), Orsay, France, 2006.

[3] **Y. Li**, et al., “CMOS Sensors for the Vertex Detector of the future International Linear Collider,” *Nucl. Instrum. Meth. A 572*, 2007, pp. 300-304.

[4] J.L.McCreary and P.R.Gray, “All-MOS Charge Redistribution Analog-to-Digital Conversion Techniques-Part I,” *IEEE Journal of Solid-State Circuits*, vol. SC-10, Dec. 1975, pp. 371- 379.

[5] Y. Degerli, N. Fourches, M. Rouger, and P. Lutz, "Low-power autozeroed high-speed comparator for the readout chain of a CMOS monolithic active pixel sensor based vertex detector," *IEEE Trans. Nucl. Sci.*, vol. 50, no. 5, October 2003, pp. 1709-1717.

[6] G. Deptuch et al., “Monolithic Active Pixel Sensors With In-pixel Double Sampling Operation and Column-Level Discrimination”, *IEEE Trans. Nucl. Sci.*, vol. 51, No.5, Oct. 2004, pp 2313-2321.

[7] B. Razavi, “Design of Sample-and-Hold Amplifier for High-Speed Low-Voltage A/D,” *IEEE 1997 Custom Integrated Circuits Conference*, 1997.

[8] B. Razavi, “Design of a 100-MHz 10-mW 3-V Sample-and-Hold Amplifier in Digital Bipolar Technology,” *IEEE Journal of Solid-State Circuits*, vol. 30, no. 7, pp. 724- 730, July 1995.

[9] J. Lecoq, “Elément de conception de convertisseurs numériques analogiques et analogiques numériques,” *School of Micro-electronics of IN2P3*, Porquerolles, France, 2005.

[10] A.S. Sedra and K.C. Smith, *Microelectronic Circuits*, 3rd ed. Saunders College Publishing/HRW, NEW Jersey, 1991.

[11] D. A. Johns and K. Martin, *Analog Integrated Circuit Design*, John Wiley & Sons, Inc. 1997.

[12] T. Kobayashi et al., “A current-controlled latch sense amplifier and a static power-saving input buffer for low-power architecture,” *IEEE Journal of Solid-State Circuits*, vol. 28, No. 4, April 1993, pp. 523-527.

[13] B. Razavi and B. A. Wooley, “Design techniques for high-speed, high-resolution comparators”, *IEEE J. Solid-State Circuits*, vol. 27, Dec. 1992, pp. 1916–1926.

[14] B. Song et al., “A 10-b 15-MHz CMOS Recycling Two-Step A/D Converter”, *IEEE J. Solid-State Circuits*, vol. 25, No. 6, Dec. 1990, pp. 1328–1338.

Conclusion and perspectives

The strict requirements (high readout speed, high spatial resolution, tolerance to radiation hardness and low-power dissipation) of the vertex detector for the future ILC and the very small signal level of MIP (only a few mV, the same order of the noise level for standard CMOS process) make the design of high performance sensors a real challenge. Two prototypes developed for this purpose are studied in this work.

The two prototypes have the same architecture: a 32 × 128 pixel matrix which is divided into sub-arrays with the photo diode pixel of different diode sizes. 8 columns are directly output, allowing characterization of pixel performance. The other 24 columns are connected with column level auto-zero discriminators. As very high readout speed is required for this application, the prototypes are optimized to work at a main clock frequency of 100 MHz, corresponding to a readout speed of 20 µs/frame. Now, it is the fastest readout speed ever achieved for sensors designed for the vertex detector of the future ILC. 1-bit binary outputs are also realized for the first time thanks to the column level discriminators. Moreover, because of the very small input signal (only a few mV) for the High Energy Physics application, CDS structures are successfully integrated, for the first time, both inside the pixel with signal amplifying stage and inside the column level discriminator, which largely reduce the noise and offset levels. The successful implementation of CDS structures opens the door for the development of the auto on-chip sparsification.

The first prototype, MIMOSA 8, was fabricated in TSMC 0.25 µm digital process with an epitaxial layer of 8 µm in the year early 2004. The second prototype, MIMOSA 16, was fabricated in AMS 0.35 µm OPTO process in the year 2006. Two versions are realized with the different thickness of epitaxial layers: 14 µm and 20 µm.

Thorough tests have been carried on for these two prototypes since 2004. Excellent results are obtained and similar behaviors are observed for the two prototypes.

With a ^{55}Fe X ray source (energy peaks: 5.9 keV and 6.49 keV), the important parameters, such as the noise level (Temporal Noise and FPN), the Charge-to-Voltage conversion Factor (CVF) and the Charge Collection Efficiency (CCE), are obtained by the direct analog outputs. The results are excellent:

- Extremely low input referred noise levels are observed for these two prototypes: the Temporal Noise is only tens of electrons (equivalent to 1 mV_{rms}) and the FPN, which is crucial to on chip auto sparcification, is less than 6 electrons (equivalent to 0.45 mV_{rms}).

- The CCE for MIMOSA 8 is very good with a value above 80% (for 5×5 cluster). Relative low CCE are observed for MIMOSA 16: about 35% for the 14 µm version and about 28% for the 20 µm version. However, more charge (compared to MIMOSA 8 with only 8 µm of epitaxial layer) will be created when MIPs cross through because the charge generated is proportional to the thickness of the epitaxial

layer (about 80 e-h/µm). For this reason, the performance of MIMOSA 16 needed to be evaluated with high energy beams.

- Moreover, all the above results are tested as the function of the main clock frequency, which determines the readout speed. We noticed that remarkably stable performances are obtained for these two prototypes on a rather large frequency range (from 24 MHz to 140 MHz for MIMOSA 8 and from 10 MHz to 170 MHz for MIMOSA16). The corresponding readout speed is about 14.6 µs for MIMOSA 8 at 140 MHz and about 12 µs for MIMOSA 16 at 170 MHz. With a slightly degraded CCE values (25% for sub-array S2), the fastest readout speed, 12 µs/frame, has been obtained with MIMOSA 16 of 20 µm epitaxial layer working at 170 MHz. In the application of CMOS sensor for MIPs detection, it is the first chip that achieves such a high readout speed (approaching the requirement of the final vertex detector).

The performances of the analog outputs for these two prototypes are re-summarized in Tab. 6.1 in the following:

Chip	Sub array	CVF (µV/e-)	Input Referred Noise (ENC)	Input Referred FPN (ENC)	S/N	Charge Collection Efficiency	
					(^{55}Fe)	3x3	5x5
MIMOSA 8	S2 (1.2 x 1.2 µm²)	66	11 e-	4 e-	130	66%	80%
	S3 (1.7 x 1.7 µm²)	60	12 e-	5 e-	110	73%	85%
	S4 (2.4 x 2.4 µm²)	52	14 e-	6 e-	90	82%	92%
MIMOSA 16 14 µm version	S2 (2.4 x 2.4 µm²)	59	11 e-	2 e-	152	36%	-
	S3 (2.4 x 2.4 µm²) Rad. Tol.	54	12 e-	2 e-	127	37%	-
MIMOSA 16 20 µm version	S2 (2.4 x 2.4 µm²)	60	11 e-	2 e-	151	25%	-
	S3 (2.4 x 2.4 µm²) Rad. Tol.	53	13 e-	2 e-	132	29%	-

Table 6.1: Characteristics of MIMOSA 8 and MIMOSA 16. Results obtained at 100 MHz. The measurement error for the noise is less than 2% and the error for the CCE values is negligible.

The digital outputs are also been characterized in laboratory for a main clock frequency of 40 MHz. Excellent performance is obtained:

- Very small Temporal Noise and FPN have been observed. The value of the Temporal Noise is about 1 mV_{rms} for MIMOSA 8 and 0.6 mV_{rms} for MIMOSA 16. The FPN values are similar for both of the prototypes: about 0.3 mV_{rms}. These values are compatible with the results obtained by the analog outputs.

- With the ^{55}Fe X ray source, the sensitivity of comparators is studied under different threshold values (from 4 mV to 6 mV). Uniform responses (accumulated results for 100 K events) of the comparators are observed and the fake hits due to noise are efficiently cut off when a higher threshold value is applied.

- Moreover, the photon detection efficiency is compared between the analog outputs and the digital outputs. Compatible results are obtained (Fig. 3.36).

The excellent results of laboratory tests indicate that the CDS structures integrated in both pixel and column level work extremely efficiently.

The noise performance remains stable with the neutron flux of 5×10^{12} neutrons/cm².

The prototype MIMOSA 8 has been tested with high energy particle beams at DESY with 5 GeV electrons in 2005 and at CERN with 180 GeV pions in 2006, respectively. The main clock frequency used during the tests is about 40 MHz.

For the analog outputs, similar temporal noise levels are observed (only tens of electrons) for the two beam tests. A high detection efficiency of above 98% has been achieved for all the three sub-arrays. The highest value is obtained by sub-array S3 for the tests made at DESY: about 98.6±0.3 %. The results are re-summarized in Tab 6.2 in the following:

Beam Test	Sub Array	Temporal Noise (e-)	Charge in 3×3 cluster (e-)	S/N of seed pixel	Detection efficiency to MIPs
DESY 2005	S2 (1.2 x 1.2 µm²)	11±2	353±4	8.6±0.2	98.2±0.4 %
	S3 (1.7 x 1.7 µm²)	12±2	392±4	9.8±0.1	98.6±0.3 %
	S4 (2.4 x 2.4 µm²)	14±2	433±5	9.4±0.2	98.0±0.4 %
CERN 2006	S2 (1.2 x 1.2 µm²)	10±2	306±4	6.2±0.5	97.6±0.6 %
	S3 (1.7 x 1.7 µm²)	11±2	328±4	6.3±0.2	98.1±0.5 %
	S4 (2.4 x 2.4 µm²)	12±2	381±5	6.6±0.6	97.5±0.5 %

Table 6.2: The beam test performance of the analog outputs for MIMOSA 8.
Measured at DESY in 2005 and CERN at 2006.

For the digital outputs, the average hit multiplicity, the fake hit rate and the detection efficiency are studied under different discriminator threshold cut values. Considering the results of the detection efficiency, the average fake hit rate and the hit multiplicity, the most appropriate threshold cut should be set between 3.5 and 5. For the threshold cuts near this range, a high detection efficiency (more than 95%) is achievable with very small fake hit rate (around 10^{-4}) and an average hit multiplicity of above 2. Tab. 6.3 below summarizes the average hit multiplicity, the fake hit rate and the MIPs detection efficiency when a threshold SNR cut of 3.5 is applied.

The spatial resolution is studied for the tests made at CERN. The center of gravity method is used to obtain the spatial resolution under different discriminator threshold cut values. Quite stable results are obtained for the threshold cut values from 3.7 to 6. For digital outputs, the average spatial resolution for all the sub-arrays is around 9.14 ± 0.10 µm for U coordinate and 7.27 ± 0.10 µm for V coordinate. Under a threshold cut of 4.8, the best spatial resolution is found with sub-array S3 (about 6.6 µm for the U coordinate, Fig. 4.24), which has the lower fake hit rate with the best detection efficiency.

Beam Test	Sub Array	Average Hit multiplicity	Fake hit rate	Detection efficiency to MIPs
DESY 2005	S2 (1.2 x 1.2 µm²)	3.7	1.6×10^{-3}	99.6±0.1%
	S3 (1.7 x 1.7 µm²)	3.6	0.8×10^{-3}	99.5±0.1%
	S4 (2.4 x 2.4 µm²)	3.1	1.6×10^{-3}	99.5±0.1%
CERN 2006	S2 (1.2 x 1.2 µm²)	3.7	1.1×10^{-3}	99.7±0.1%
	S3 (1.7 x 1.7 µm²)	3.4	2.5×10^{-3}	99.6±0.1%
	S4 (2.4 x 2.4 µm²)	3.2	1.8×10^{-3}	99.5±0.2%

Table 6.3: The beam test performance of the digital outputs for MIMOSA 8. Threshold cut is set to 3.5 Measured at DESY in 2005 and CERN at 2006.

The very encouraging laboratory tests and beam tests show that the two prototypes perform remarkably stable and efficient performance. The noise level, which is crucial for the application of vertex detector for the future ILC, is extremely well controlled thanks to the CDS structures integrated. Therefore, the following development should be based on the actual pixel architecture.

The digital outputs are only tested with a main clock frequency of 40 MHz, corresponding to a readout speed of 51.2 µs per frame. Thorough tests with a higher readout speed are needed for evaluating the digital outputs performance.

Relative low CCE are observed for MIMOSA 16 (Tab. 6.1). However, beam tests are needed for correctly evaluating the performance of MIMOSA 16 because more charge will be created due to its thicker epitaxial layer. A beam test for MIMOSA 16 is programmed on September 2007. Moreover, enlarge the diode size is another solution. A submission for another prototype with larger diode size is made recently. The tests for this new chip are needed.

Despite the low noise level, the SNR for the seed pixel of clusters associated to charged particle hits is relatively low with a value less than 10. The main reason is due to the limited thickness of the epitaxial layer: only 8 µm indicated by the manufactory but we found that the real thickness is probably less than this value: only 6.6 ± 0.4 µm. The average total charge generated inside the epitaxial layer is then quite small: about only 520 e-. Considered a CCE of about 90%, the maximum charge that can be collected is around 475 e-, which is a very limit value. Therefore, the search for new CMOS process with appropriate thickness of epitaxial layer and doping profile will never stop.

The spatial resolutions we found approach to the prediction of uniform distribution: around 7 µm in both U and V directions. The reason is that only two values ("1" and "0") are available for the digital outputs. With an average hit multiplicity of 3.5, it is very difficult to identify the center of the hit cluster. Thus, a column level ADC is needed to replace the actual 1-bit discriminator. The challenge of this ADC is that the CDS structures must be integrated despite the very limit width of the column (about 25 µm).

A 5-bit ADC of successive approximation architecture was designed and fabricated by AMS 0.35 µm OPTO CMOS process in 2006. Combined with CDS structures, very small dispersion is observed among the different channel. The size is very compact: 25 µm × 1 mm. The total ADC conversion period is 270 ns (sampling the pixel signal, auto offset-cancellation and digital data output). The static consumption is about 300 µW. Preliminary tests indicate that this ADC works correctly, but with a LSB more important than previewed. For this reason, an amplifier is mandatory to replace the source follower used actually in the sample-and-hold stage.

A new column level comparator is also designed and fabricated by AMS 0.35 µm OPTO CMOS process. Using a static latch, which allows sampling and memorizing its offset during the resetting step, the structure of the comparator is large simplified. Only one preamplifier stage is used. The preliminary results show that its performance approaches the comparator used in MIMOSA 8 and MIMOSA 16. The Temporal Noise is less than 0.7 mV_{rms} and the FPN is only 0.22 mV_{rms}. Test with pixel matrix is needed for father evaluating its performance.