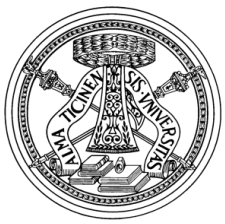

UNIVERSITA' DEGLI STUDI DI PAVIA

DIPARTIMENTO DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE

DOTTORATO DI RICERCA IN MICROELETTRONICA

MONOLITHIC PIXEL SENSORS FOR HIGH RESOLUTION DETECTION SYSTEMS IN ADVANCED CMOS TECHNOLOGIES

Tutor:
Prof. Lodovico Ratti

Coordinatore del Dottorato:
Prof. Franco Maloberti

Tesi di Dottorato
di Stefano Zucca

Anno Accademico 2011/2012

Contents

Introduction

This work will discuss the design and characterization of advanced monolithic active pixel sensors (MAPS) in CMOS technologies. These devices are being proposed as an alternative to other more mature approaches (like hybrid pixels and striplets) in the fabrication of detectors for vertexing applications in the high energy physics (HEP) experiments of the future high luminosity colliders. MAPS have been initially conceived for imaging applications, but, because of their features, their potential for HEP applications are being investigated by several groups in the particle physics community. As compared to current experiments, machines operating at the future colliders will demand for a higher granularity in the vertex detector in order to get more precise information on particle momentum. This leads to a pixel pitch reduction down to 50 μm or even less. Moreover, in order to minimize multiple scattering effects, the substrate of vertex detectors needs to be thinned down. This last requirement makes MAPS very attractive for this kind of applications, since both the collecting elements and the readout circuits are integrated into the same substrate resulting in a significant material budget reduction (as compared for example to hybrid pixels). MAPS can also be thinned down to a few tens of microns, since their operating principle is based on the diffusion of carriers in an undepleted substrate. Another important feature of MAPS lies in the fact that they are based on CMOS deep submicron technologies. Modern CMOS processes enable the design of low power consumption readout circuits, making it possible to use relatively simple cooling systems in order to maintain the temperature to a suitable level during detector operation. In this way, the material budget can be further limited. Moreover, CMOS deep

submicron technologies feature a high tolerance to ionizing radiations, which is a crucial requirement in HEP vertex trackers. One weak point of MAPS is represented by the generally poor tolerance to bulk damage, which reduces the lifetime of minority carriers diffusing in the device substrate. In the last few years, several research groups have started working on the design and fabrication of the SuperB Factory, which is an electron-positron collider targeting a final luminosity of 10^{36} $cm^{-2}s^{-1}$. Due to the accelerator performance and configuration, also a high background rate is expected, demanding for the use of fast readout architectures.
In order to fulfill these requirements, a novel approach to the design of a detector for application to the Layer0 of the SuperB silicon vertex tracker (the innermost of the 6 layers making up the detector) has been proposed. The approach is based on a quadruple well 180 nm CMOS process, called INMAPS, featuring a weakly P-doped epitaxial layer grown upon a heavily P-doped substrate. MAPS designed with the INMAPS technology feature small N-well diffusions acting as collecting electrodes and a fully CMOS front-end readout chain, therefore including P-type transistors with their N-wells. This is made possible by using an additional heavily P-doped implant to shield the N-wells hosting the PMOS transistors from the epitaxial layer. In this way, the charge generated by the passage of a minimum ionizing particle diffuses in the epitaxial layer and reaches the collecting electrode without being collected by parasitic N-wells, leading to a theoretical charge collection efficiency of 100%. Use of both PMOS and NMOS transistors in the front-end electronics makes it possible to implement a sparsified digital readout architecture in order to comply with the high background rate of the Layer0 of the SuperB silicon vertex tracker (SVT). Moreover, the use of a weakly P-doped (i.e. high resistivity) epitaxial layer, is also expected to improve the detector tolerance to bulk damage.The analysis of the radiation performance of the so called deep N-well (DNW) MAPS in 130 nm CMOS process, performed in the second chapter of this work, also serves the purpose of emphasizing the advantages, in terms of radiation tolerance, offered by a high resistivity substrate.
The structure of this work is as follows. The first chapter begins with a short description of the SuperB Factory and of its SVT structure. Then, the specifications for detectors intended for application in the Layer0, and their relevance to the detector and readout electronics design, is discussed. After that, an overview of the technology options for the Layer0 will be given.
The second chapter starts by introducing the reader to the radiation damage mechanisms underlying performance degradation in silicon devices. Results from an irradiation campaign with ionizing radiation performed on deep N-

well MAPS will be shown, paying particular attention to the performance degradation of the analog front-end. These devices have also been irradiated with neutrons from a nuclear reactor in order to evaluate their tolerance to bulk damage. In this case, particular attention has been paid to the charge collection properties degradation of the collecting element. The information collected from these irradiation campaigns has been exploited in the design of the Apsel4well chip, developed with the CMOS INMAPS process, which represents the core of this work.
The third chapter begins with a description of the main features of the INMAPS technology. Then, the design features of the analog front-end will be thoroughly described, along with the performance obtained by post-layout simulations in terms of charge sensitivity, linearity, noise, threshold dispersion and power consumption. The chapter continues with the presentation of the results from physical device TCAD simulations of the pixel, which were used to optimize the layout of the collecting element in terms of charge collection properties. The chapter ends with the description of the Apsel4well chip layout, with particular attention to some design features of the pixel adopted to minimize capacitive coupling phenomena between the in-pixel digital section and the analog section and the collecting electrodes.
The last chapter is aimed at showing the results from the characterization of the analog front-end and of the detector collecting elements. In the first part of the chapter, noise and charge sensitivity measurements will be shown. Also, the results from the study of the charge collection properties of the Apsel4well MAPS, performed by means of a laser and radioactive sources, will be presented, with some emphasis on how the performance is affected by the epitaxial layer features. Then, results from an irradiation campaign performed on Apsel4well chips with neutrons from a nuclear reactor will be shown. Finally, a Monte Carlo model for charge loss in neutron irradiated devices will be discussed, together with the model validation through comparison with experimental data. This model, beside bringing along a better understanding of the mechanisms underlying the charge collection loss after irradiation with heavy particles, also provides the designer with a fast tool able to predict the effects of neutrons on the sensor performance, possibly improving MAPS rad-hard design.

Chapter 1

The silicon vertex tracker for the SuperB Factory

1.1 The SuperB Factory

Everything we can see in the Universe, from a midge to a galaxy, is made up of so called elementary particles. These are collectively referred to as ordinary matter, and represents 4% of the Universe. The obvious question that physicists of the early twentieth century have been asking themselves is what the remaining 96% of the Universe is made off. At present, the most credited hypotheses rely upon dark matter and dark energy, which are extremely difficult to detect, other than through the gravitational forces they exert. Investigating the nature of dark matter and dark energy is one of the biggest challenges since the second half of the twentieth century in the fields of particle physics and cosmology. Dark matter and dark energy apperently are there since the Universe was born. Their existence led physicists to adopt an alternative way to investigate their nature: their generation in a laboratory, by recreating for very short times the Universe initial conditions. This is the reason why, from the beginning of the sixties, physicists begun to build big particle accelerators and colliders able to create, by means of the collision between subatomic particles, these initial conditions featuring extremely high energy densities. When such particles are smashed togheter in a collider at velocities close to light speed, the available energy can be converted into mass according to Einstein relativity theory, generating new particles by exploiting the Terascale energy of the new generation accelerators, scientists are confident they can identify, among the unknown products of collision, dark matter particles [1]. The world's largest and most powerful accelerator, at the moment, is the Large Hadron Collider

(LHC) built at CERN, in Switzerland, which features a circular geometry with a 27 km long circumference. Inside this accelerator, two beams of protons are boosted to very high energies before colliding with one another. Experiments like ATLAS and CMS already started at LHC, exploit the TeV energy of the colliding protons in order to create heavy particles and study their behavior. The approach followed by other HEP experiments, like SuperB, is based on increasing the probability of interaction between particles with lower energy (like electrons and positrons) and maximizing the machine luminosity. In this way, through the study of extremely rare decays of already known particles, it is possible to collect large amount of data on some as yet unkown effect supporting or disproving the Standard Model. By following this second approach, it is possible to reach the goal, with much lower costs, by means of a high luminosity electron-positron collider. However, this involves the design of detectors with higher radiation hardness and able to detect events with a frequency two orders of magnitude higher than in the LHC [2].

Figure 1.1. Floor plan of the SuperB Factory facility, with emphasis on the accelerator ring.

The SuperB Factory is an asymmetric collider (see Fig. 1.1) featuring a linear accelerator (linac) where fifty billions of particles per second are produced: this particles are electrons and positrons. The linac continuously injects (50 times per second) particle bunches in a little ring, called damping ring, in order to compact them and to lower their emittance (the lower is the beam emittance, the higher is its luminosity). Electron and positron bunches are then extracted from the damping ring and sent to two superposed storage rings where they are injected in opposite directions. In both rings, two thousand of bunches are accumulated, each containing fifty billions of electrons and positrons. Bunches collide with each other in the collider hall, where the detector is located, continuously acquiring and processing all the events generated by particle interactions. Thanks to the so called "crabbed waist" collision scheme [3], SuperB is expected to reach a luminosity in the order of 10^{36} cm^{-2} s^{-1}, therefore overcoming some of the issues that have plagued earlier electron-positron collider design. Actually, PEP-II and KEKB asymmetric colliders [4] [5] reached luminosities of 7×10^{35} cm^{-2} s^{-1}, but required wall plug power in the order of 100 MW.

1.2 Design specifications for the SuperB silicon vertex tracker

The SuperB detector concept is based on the *BaBar* detector, with those modifications required to operate at a luminosity of 10^{36} cm^{-2}s^{-1} or more, and with a reduced center-of-mass boost [7]. The *BaBar* experiment was shut down in 2008, after 10 years of operation. The innermost section of the *BaBar* detector consisted in particular of a tracking system with a five layer double-sided silicon strip vertex tracker (SVT). Figure 1.2 shows a schematic cross section of the SuperB detector which, apart from the SVT, also includes several other subsystems, all contributing to an accurate event detection. The Silicon Vertex Tracker (SVT), as in *BaBar*, along with the drift chamber (DCH) and the solenoidal magnet, provides track and vertex reconstruction capability for the SuperB detector. The SuperB SVT design is based on the *BaBar* vertex detector layout with the addition of an innermost layer (Layer0) closer to the interaction point (IP, see Fig. 1.3). The external SVT layers (1-5), with a radius between 3 and 15 cm, will be built with the same technology used for the *BaBar* SVT (double sided silicon strip sensors), which is adequate for the machine background conditions expected in the SuperB accelerator scheme (i.e., with low beam currents). Physics studies and background conditions set stringent requirements on the Layer0 design, radius of about 1.5

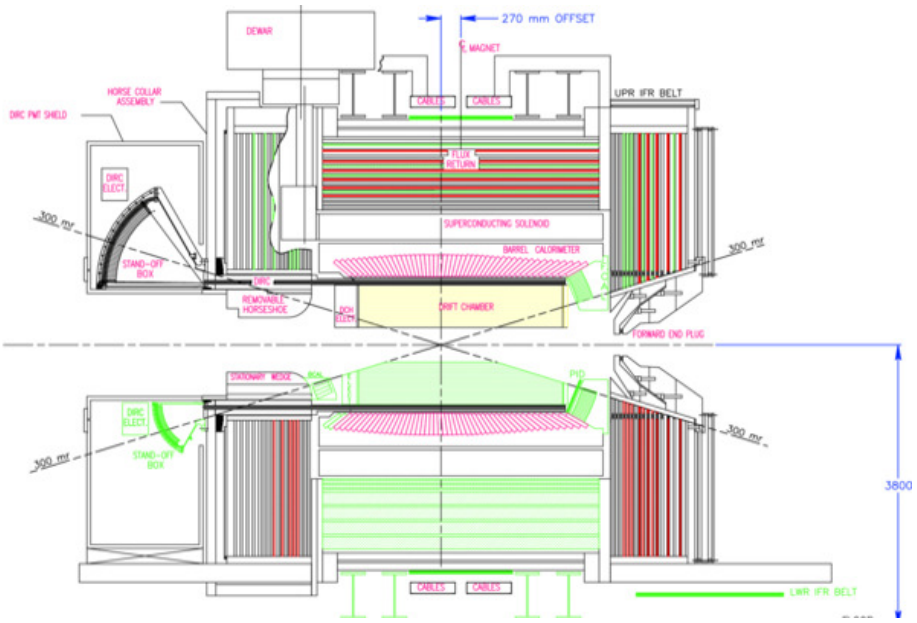

Figure 1.2. Schematic cross section of the SuperB detector.

cm, high granularity (50x50 μm^2 pitch), low material budget (about 1% X_0), and an adequate radiation resistance. This last requirement is of paramount importance for the Layer0 reliability and, for this reason, radiation mechanism on CMOS devices and detectors (which represents one of the candidate approaches for a Layer0 SVT upgrade) will be thoroughly discussed in Chapter 2. The ultra-low emittance beams of the SuperB design enables the use of a small radius beam pipe (1 cm) in the detector acceptance, in order to have the innermost SVT layer very close to the IP. A water cooling channel is foreseen for the beam pipe to avoid overheating. The amount of the beryllium pipe radial material, which includes a few μm of gold foil, and the water cooling channel, is estimated to be less than 0.5% X_0. The six-layer SVT solution for SuperB, with the Layer0 sitting much closer to the IP than the innermost layer in *BaBar*, would significantly improve track parameter determination, therefore matching the more demanding requirements on the vertex resolution. By means of some studies of the detector based on the software Geant4, it was possible to estimate the Layer0 background conditions [8]. According to these studies, the track rate at the Layer0 at a radius of 1.5 cm is at the level of about 5 MHz/cm^2, mainly due to electrons in the MeV energy range. The

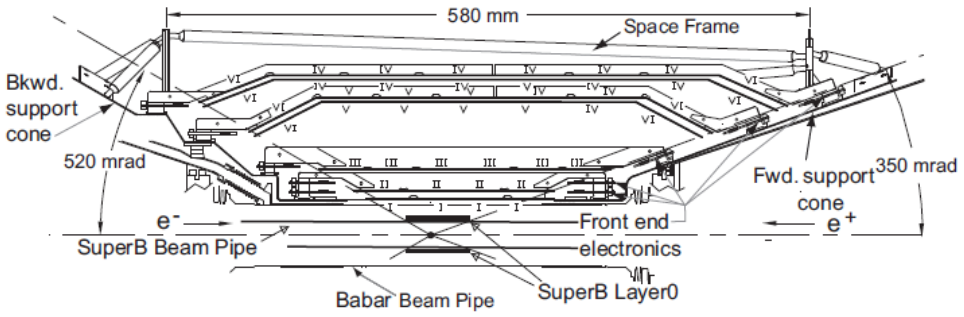

Figure 1.3. Schematic cross section of the SuperB SVT.

equivalent fluence reaches 3.5×10^{12} neutrons·cm^{-2}yr^{-1}, where the expected dose rate is about 3 Mrad/yr. A safety factor of five on top of these numbers has been considered in the design of the Layer0 that should sustain a target hit rate, including an average 4 cluster multiplicity, of 100 MHz/cm^2.

1.3 Technology options for the Layer0

In the last few years, different technological approaches have been investigated in view of an upgrade of the Layer0 to a pixellated detector. The current baseline configuration of the SVT Layer0 is based on the striplets technology (see Section 1.3.1) which, among the available technologies with an acceptable degree of maturity, can provide the best physics performance thanks to the very low material budget associated (striplets have no readout electronics in the active area) [7]. However, for the full luminosity regime, options based on pixel sensors, more robust in high background conditions, are being developed with specific R&D programs in order to meet the Layer0 requirements. The replacement of the Layer0 striplet modules is foreseen for a second phase of the experiment. For this purpose, the SuperB interaction region and the SVT mechanics is being designed to ensure rapid access to the detector for a fast replacement procedure of the Layer0. The schematic drawing of the full Layer0 made of 8 pixel modules mounted around the beam pipe with a pinwheel arrangement is shown in Fig. 1.4. As commented in the previous section, to minimize the undesired effects of multiple scattering on track parameter resolution, the reduction of the material is crucial for all the components of the detector module in the active area. A light carbon fiber support structure with integrated active cooling based on microchannels, is being developed to evacuate

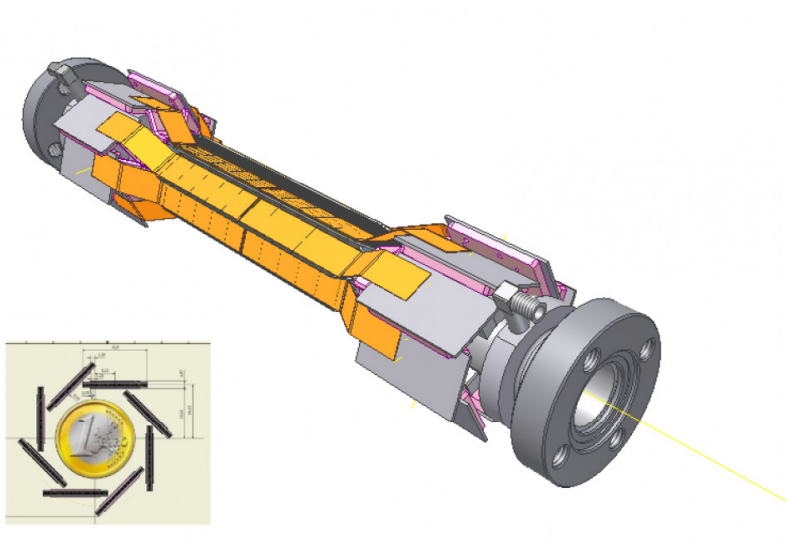

Figure 1.4. Schematic drawing and cross-section of the full Layer0 made of 8 pixel modules mounted around the beam pipe with a pinwheel arrangement.

the power dissipated by the pixel front-end electronics (about 2 W/cm^2) [9]. Measurements on the first support prototypes, with a total material budget of 0.15% X_0, indicate that such approach is a viable solution to the thermal and structural problem of Layer0. The connection of front-end chips to the periphery of the module will be realized by means of a light aluminum/kapton multilayer bus.
The following subsection reviews the different approaches and technology options considered for the Layer0 development, with the exception of deep P-well CMOS monolithic active pixel sensors, thoroughly investigated in Chapter 3.

1.3.1 Striplets

The first approach is based on double-sided silicon strip detectors placed at an angle of $\pm 45°$ to the detector edge on the two sides of the sensor. The strip thickness is 200 μm and the readout pitch is 50 μm, according to Layer0 requirements. The schematic cross section of a double-sided silicon strip detector is shown in Fig. 1.5. It features orthogonally n^+ and p^+ diffusions implanted on both sides of the detector n bulk, in order to get information on both x and y coordinates of the particle impact point. Each n^+ strip is surrounded by a floating p^+ doped implantation to be isolated from any adjacent strip.

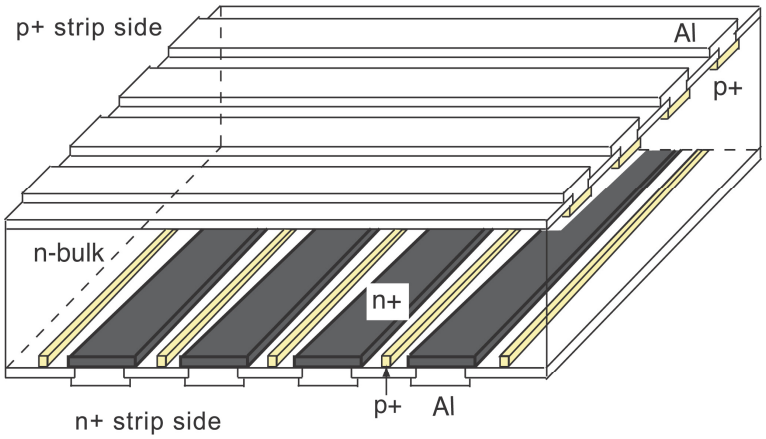

Figure 1.5. Schematic cross-section of a double-sided silicon strip detector.

When a particle crosses the substrate about 80 electron/hole pairs per μm are generated. Electrons and holes drift in opposite directions due to the electric field created by the voltage bias applied between the two sides completely depleting the substrate. Only the strips near to the migrating charges will give a signal [10]. The strips will be connected to the readout electronics through a multilayer flexible circuit glued to the sensor. The readout electronics currenly under design, will features a data-driven, high-bandwidth architecture, similar to that of the FSSR2 readout chip [11]. It will have 128 channels providing a sparsified digital output with address, timestamp and pulse height information for all hits. The chip will be fabricated in a 130 nm CMOS technology for high radiation tolerance.

1.3.2 Hybrid pixels

Hybrid pixels represent a mature technology although they still need some R&D to meet Layer0 requirements, especially in terms of front-end pitch and total material budget. Figure 1.6 shows a schematic cross-sectional view of a hybrid pixel sensor. A hybrid pixel consists of two chips back-to-back assembled by means of bump bonding techniques. Figure 1.6 shows the case of a N-on-N type sensor, featuring a n^+ diffusion in a n-substrate, but also sensors with a p^+ diffusion in a n-substrate (P-on-N) can be used. The first approach is based on the collection of all the electrons generated by the crossing of a

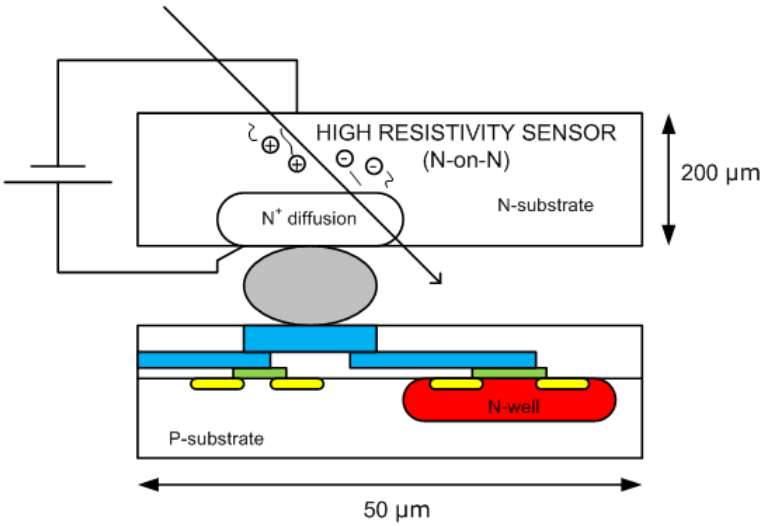

Figure 1.6. Schematic cross-section of a generic hybrid pixel.

minimum ionizing particle (MIP) through the fully depleted high resistivity substrate. In the case of the second kind of sensor, holes are collected by the p^+ diffusion. This second approach features a couple of drawbacks. The first one is related to the lower (with respect to electrons) hole mobility which in turn leads to a slower charge collection. The second issue is related to the lower (with respect to N-on-N sensors) tolerance to bulk damage, especially at very large fluences [12]. Again in Fig. 1.6, the lower tier includes the sensor readout chain, which can be implemented by means of planar or vertically integrated CMOS technologies (described in the next subsection). One of the main advandages of the hybrid pixel approach with respect to monolithic active pixel sensors (MAPS, see next subsection) is related to the high radiation hardness in terms of bulk damage, thanks to the higher resistivity and complete depletion of the substrate. On the other hand, by means of fine pitch bump bonding methods or other, more innovative direct bonding techniques (like the ones offered by Zyptronix or TMicro/ZyCube [13]), it is possible to meet the granularity requirements set for the detector in the Layer0.

A prototype front-end chip called Superpix0 has been developed with a 130 nm planar CMOS process by the SLIM5 collaboration for application to the SuperB Layer0 [14] [15]. The chip is compatible with a 50 μm pitch sensor. The analog cell adopts a shaperless front-end design and is optimized to readout a high resistivity 200 μm thick sensor with a signal to noise ratio of about 100.

The chip contains 4096 pixels grouped in 256 blocks called macropixel, each including 4x4 pixels and adopts a data-push sparsified readout architecture.

1.3.3 CMOS monolithic active pixel sensors

CMOS MAPS were first developed in the domain of imaging applications, where they are also known as CMOS APS (active pixel sensors) or as CIS (CMOS image sensors) [16], in competition with charge coupled devices (CCD). In the last decade, a certain number of research teams working in the high energy physics community began to consider CMOS MAPS in view of using them as charge particle tracking devices with high momentum resolution capabilities, as an alternative to hybrid pixel detectors [17] [18] [19]. As far as high energy physics (HEP) applications are concerned, CMOS MAPS may provide a very convenient solution to the issue of material budget in multilayer detectors, which is the main source of multiple scattering and position measurement inaccuracy. In monolithic sensors, the readout electronics is fabricated in the same substrate as the detector (separate substrates are instead used for the detector and the readout channel in hybrid pixels, as illustrated in the previous section). Moreover, the device substrate can be ground down to a few tens of microns with no significant signal loss, since the operating principle of MAPS is based on the collection of charge diffusing in an undepleted substrate (whereas, in a fully depleted detector, the collected charge grows linearly with the thickness). Although the presence of an epitaxial layer is generally preferred in the design of CMOS MAPS (like in the case of the Apsel4well chip presented later on, which represents the main concern of this work), monolithic pixels have been designed also in epi-less CMOS technologies (see the chip Apsel3T1 described in Chapter 2, but also other cases in the literature [20]).
Figure 1.7 shows a simplified cross-section of CMOS MAPS with and without epitaxial layer. In the first case, the CMOS MAPS features a p^+-substrate, on top of which a p^- epitaxial layer 10-20 μm thick is grown. In this epitaxial layer both the p-well diffusions acting as the substrate of the readout channel NMOS transistors, and an n^+ diffusion operating as collecting electrode, are then implanted. The second structure features a p-substrate in which both p^+ and n^+ diffusions are directly implanted. When a MIP crosses the silicon substrate, it generates 80 electron/hole pairs per μm which diffuse or drift towards the collecting electrode [21]. Let's consider the case of a MAPS with epitaxial layer. At the interface between two differently doped regions (i.e. substrate-epitaxial layer or epitaxial layer-pwell), the generated charge carriers are reflected or transmitted depending on the potential gradient. More in particular, since the

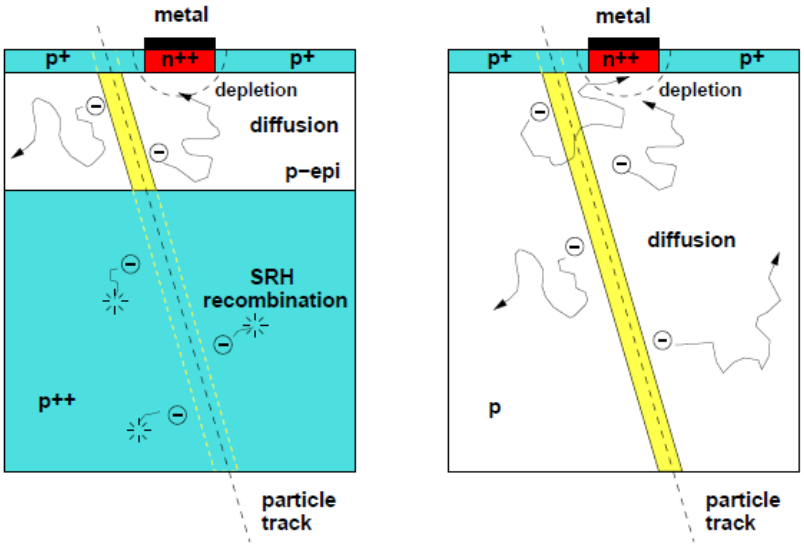

Figure 1.7. Schematic cross-section of CMOS MAPS with (left) and without (right) the epitaxial layer.

Fermi level is constant at the equilibrium (see Fig. 1.8) the potential barrier between the p^- epitaxial layer and the p^+ regions (substrate and p-well) is equal to [22]:

$$qV_{bi} = E_{f,p^-} - E_{f,p^+} = kT\left[ln\frac{N_V}{p_{p^-}} - ln\frac{N_V}{p_{p^+}}\right] = kTln\frac{p_{p^+}}{p_{p^-}} \tag{1.1}$$

where V_{bi} is the built-in potential at the interface, E_{f,p^-} and E_{f,p^+} are the Fermi levels in the epitaxial layer and in the heavily doped region, respectively, k is the Boltzmann constant, T the absolute temperature, q the electron charge, N_V the silicon state density in the valence band and p_{p^+} and p_{p^-} the majority carriers doping concentration at the equilibrium in the considered regions. Since this concentration, for non-negligible doping levels, is almost equal to the acceptor atoms concentration N_A, (1.1) can be re-written as:

$$V_{bi} = \frac{kT}{q}kTln\frac{p_{p^+}}{p_{p^-}}. \tag{1.2}$$

From (1.2) and Fig. 1.8, it is clear that the potential gradient will drive electrons towards the epitaxial layer and, at the same time, will favors their shift

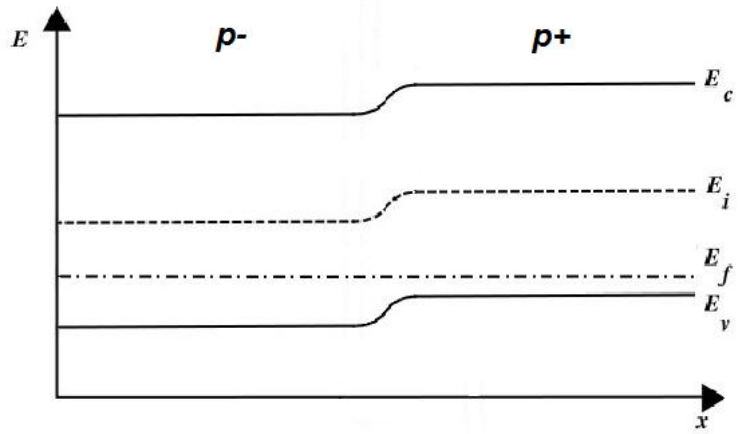

Figure 1.8. Energy band diagram at the interface between two regions with different doping concentration. E_c is the minimum energy of the conduction band, E_i is the intrinsic Fermi level, E_f si the Fermi level and E_v is the maximum energy of the valence band.

from highly doped substrate regions towards the epitaxial layer. In this way, the epitaxial layer acts as a confinement region for the minority carriers (electrons), which are then collected by the N-well sensor at the inversely biased pn junction (see Fig. 1.9). In such a structure, similarly to what happens in a photodiode, the passage of a MIP leads to an inverse current peak through the pn junction. This current signal is then processed by a suitably designed readout chain. In the so far discussed case of a MAPS with epitaxial layer, the active volume of the detector is usually confined within the epitaxial layer, since most of the charge generated in the low resistivity substrate is affected by Shockley-Read-Hall recombination. If an epi-less process featuring a relatively high (10^{15}-10^{16} atoms/cm^3) resistivity substrate is used for MAPS fabrication, the amount of charge not immediately recombining is greater than in the previous case. However, only a small fraction of the charge generated deep in the substrate can reach by diffusion the collecting electrode. The remaining part of the charge can be detected by other pixels of the detector matrix within relatively long times (few hundreds of nanoseconds) or undergoes recombination. The MAPS approaches proposed in Fig. 1.7 are particularly suitable for imaging applications, where usually three transistors are integrated at the pixel-level and data are commonly read out through a sequential scan of the

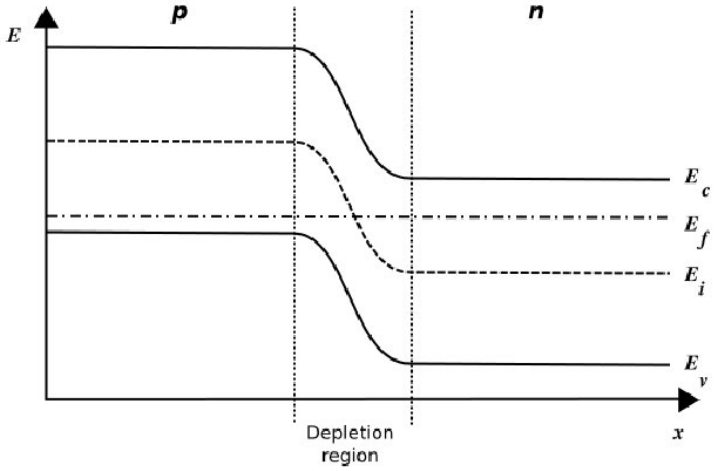

Figure 1.9. Energy band diagram at the interface between an inversely biased pn junction.

sensor matrix. This is possible because of the relatively low operation frequency needed in this kind of applications. However, when application to high luminosity particle accelerators like SuperB is considered, the three transistor scheme does not appear to satisfy the requirements set by the experiment in terms of readout bandwidth. A sparsified readout may fulfill the readout speed specifications, but requires a higher degree of complexity in the pixel-level electronics. The three-transistors approach does not allow the use of PMOS transistors in a more complex readout channel. Actually, PMOS transistors need to be housed in a V_{DD} biased N-well. Such an N-well cannot be implemented inside the pixel because it would behave as a parasitic collecting electrode and subtract charge from the real sensor. A possible solution could be to increase the collecting electrode size, in order to make the ratio between the sensor area and the total N-well area as close as possible to 1. In this way, the charge subtracted by the PMOS parasitic N-wells from the collecting electrode could be made negligible. However, the equivalent noise charge (ENC), which is the charge that, once collected by the sensor, generates at its cathode a voltage variation equal to the square root of the mean square noise at the circuit output, has the following dependence on the detector parasitic capacitance C_D:

$$ENC = \sqrt{K_B T C_D}. \tag{1.3}$$

Then, the signal-to-noise ratio (SNR) is inversely proportional to the square root of C_D:

$$SNR = \frac{Q_{inj}}{\sqrt{K_B T C_D}}. \tag{1.4}$$

The next paragraph discusses the deep N-well MAPS approach, which makes it possible to implement data sparsification techniques in monolithic pixel sensor chips.

Deep N-well CMOS MAPS

The deep N-well MAPS (DNW MAPS) approach is based on the same working principle as standard MAPS described above, where minority carriers generated by charged particles in a P-type substrate (whether an epitaxial layer is included or not) diffuse and are collected by N-type electrodes. The DNW-MAPS sensor, whose schematic cross section is shown in Fig. 1.10, differs from MAPS shown in the previous section in two aspects. First of all, an N-well with a deep junction (deep N-well) is used to collect the charge released in the

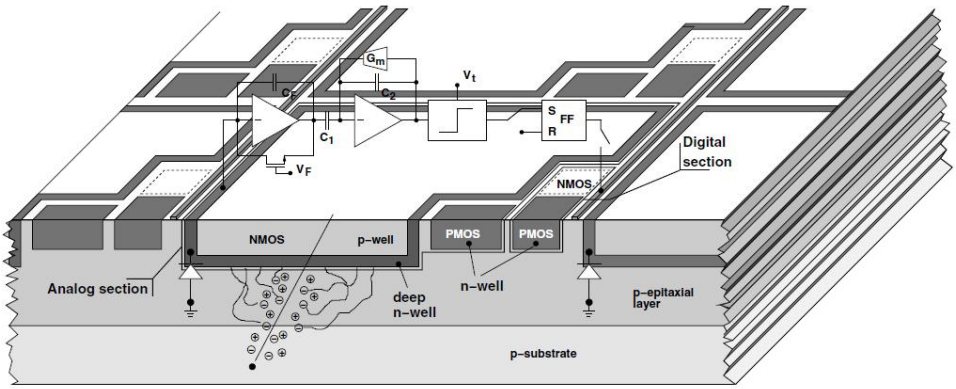

Figure 1.10. Schematic cross-section of DNW MAPS along with an example of readout chain.

substrate by a MIP. Second, a readout chain for capacitive detectors, featuring a charge preamplifier, a shaping stage, a comparator and the in-pixel logic, is used to process the charge signal. The choice of a DNW sensor structure acting as the collecting electrode has a multiple purpose. In modern triple-well CMOS processes, the deep N-well is used to shield NMOS devices from substrate coupled noise in mixed signal circuits. In the case of MAPS, the P-well enclosed in the deep N-well can be used to implement N-type devices, therefore mitigating the constraints set by the readout circuits on the sensor area and geometry. Moreover, if a large area DNW collecting electrode (as compared to standard MAPS) is laid out, PMOS devices may be used in the elementary cell at the expense of a certain amount of charge collection inefficiency, depending on the ratio of the DNW area to the sum of DNW and N-well area. The signal coming from the DNW sensor is processed by means of a charge preamplifier, which makes the charge sensitivity of the readout chain no longer dependent on the collecting electrode capacitance (and, in turn, area). Moreover, the large scale of integration in deep submicron CMOS processes and the area saved by integrating NMOS devices inside the DNW sensor can be exploited to add both analog functions (like signal shaping) and digital functions to the pixel level processing electronics, enabling the implementation of data sparsification techniques. The DNW MAPS approach, while very attractive for high energy physics experiment application, is affected by some partially unsolved issues. First of all, the analog section of the readout chain and the DNW sensor share the same substrate with the noisy digital part of the readout. If the pixel is not carefully designed, the risk of spurious charge injections through parasitic capacitances is very high. Moreover, as described above, this architecture does not allow a complete elimination of the inefficiency related to the presence of parasitic N-wells. The next paragraph discusses the use of vertical integration CMOS processes for the design of DNW MAPS, representing a possible solution to the above mentioned problems.

3D CMOS MAPS

Following Moore's Law by scaling conventional CMOS technologies involves huge investments, especially to mantain or improve delay time performance in digital circuits. In fact, CMOS technologies scaling down leads to an increase of interconnection length between blocks designed in the same chip. 3D monolithic integration is becoming very popular among IC designers as it offers a lot of advantages [23] [24]:

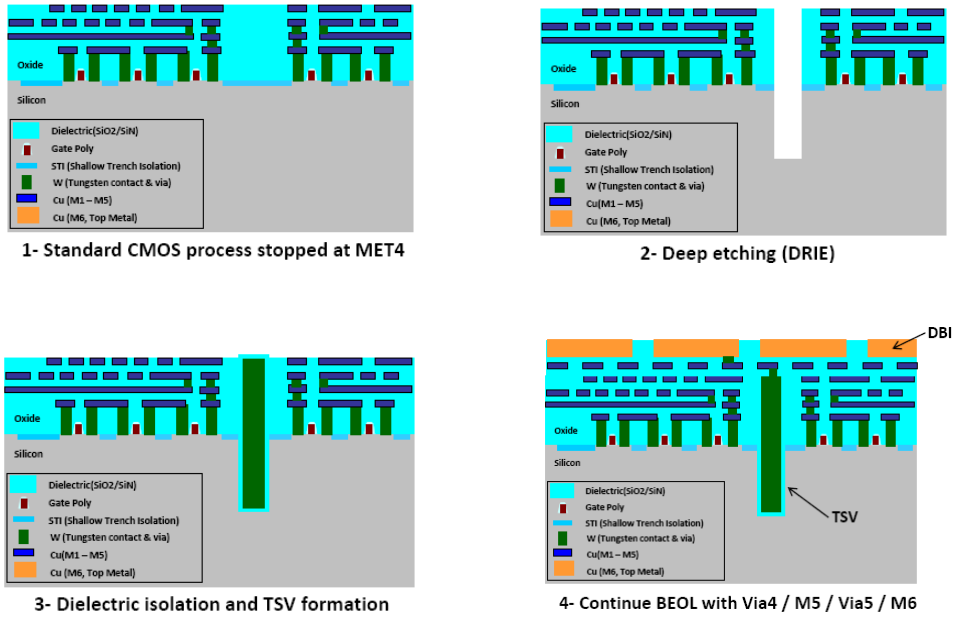

Figure 1.11. Globalfoundries-Tezzaron process: single tier fabrication.

- reduction of line distributed parasitic elements, such as resistance, capacitance and inductance, resulting in short delay times between CMOS logical gates;
- higher fan-out and logical span of control;
- higher functional density: more transistors can be integrated in the same area if they lie on overlapping semiconductor layers;
- lower cross-talk: in mixed-signal circuits, analog and digital sections can be integrated on two different layers;
- lower power dissipation, thanks to the elimination of most of the power hungry signal repeaters.

Vertical integration technologies have been recently proposed for the design of particle sensors for high energy physics (HEP) applications. As mentioned above, it can help obtain high granularity, complex in-pixel logic, low cross-talk between the analog and the digital sections of the sensor channel readout and low power dissipation. A further advantage related to the adoption of 3D

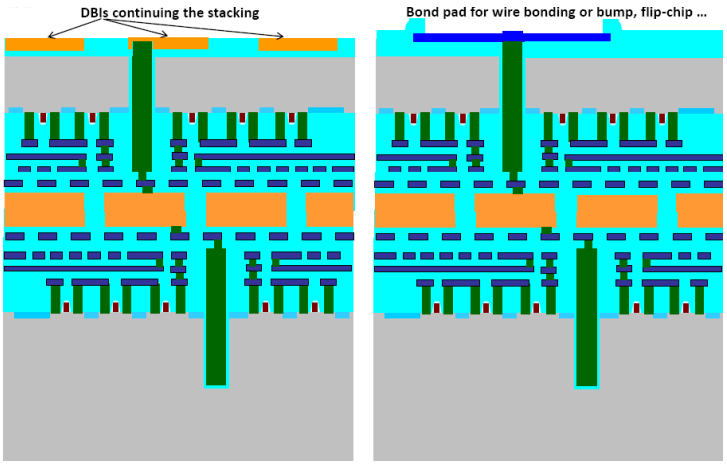

Figure 1.12. Globalfoundries-Tezzaron process: tier-to-tier bonding.

CMOS processes in the design of pixellated particle detectors is the dead area reduction, since the readout electronics can be designed with virtually no-peripheral circuits. In this way it would be possible to obtain a four-side buttable detector. One of the approaches used for the fabrication of a 3D CMOS monolithic active pixel sensor for HEP applications is based on Tezzaron SuperContact technology and Globalfoundries 130 nm CMOS process [25]. This process features two-tiers face-to-face bonded wafers with the top tier exposing through silicon vias (TSV) and backside metal pads for wire bonding. A schematic representation of the process steps is shown in Fig. 1.11 and Fig. 1.12. Figure 1.11 describes the single tier fabrication: the Globalfoundries 130 nm standard CMOS process is stopped at the fourth metal level. Then, a deep reactive ion etching (DRIE) step is performed in order to create the hole housing the TSV. This hole is isolated by means of silicon oxide and then filled with tungsten. The TSV is now formed and the process continue with the deposition of the remaining metal levels and the surface direct bond interface (DBI). The tier-to-tier bonding step is represented in Fig. 1.12, where two layers are face-to-face bonded by means of thermo-compression techniques. The second tier is then thinned down until the TSV emerges from the wafer surface. At this point it is possible to continue stacking the layers by generating other DBIs (on the left of the figure) or create bond pads for wire or bump bonding (on the right). The conceptual step from a DNW MAPS in planar CMOS technology to a vertically integrated version of it is shown in Fig. 1.13, representing a schematic cross-section of a 2D monolithic sensor

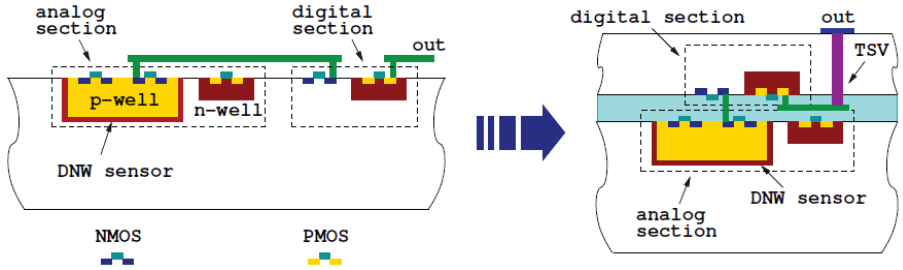

Figure 1.13. Schematic cross-section of a 2D monolithic sensor and of its 3D translation.

and of its 3D translation. The DNW sensor and the analog section are integrated in the bottom tier, while the digital front-end is housed in the top layer. This separation has been conceived to prevent digital signals from interfering with the analog stages by injecting charge into the sensor through parasitic capacitive coupling. Moreover, integrating all the PMOS devices used in the digital blocks in a different substrate from the sensor, significantly reduces the amount of N-well area in the surroundings of the collecting electrode and the relevant parasitic charge collection effects, therefore improving the detector charge collection efficiency. It is also possible to integrate in the top tier the PMOS transistors of the analog front-end in order to minimize the presence of parasitic N-wells in the bottom tier. However, this choice has to be carefully evaluated by the designer, since in this way part of the analog readout channel would share the same substrate with the digital part (a trade-off between cross-talk and collection efficiency performance has to be found).
The first 3D MAPS prototype has been tested with encouraging results [26]. However, the technology has not yet proved mature enough for application to real experiments. Several problems have been experienced in the fabrication of the first prototypes, indicating that the process still needs to be tuned. Also, while a complete characterization still needs to be performed, radiation hardness is expected to be a weak point of vertically integrated DNW MAPS, as in their 2D counterparts. For these reasons, the SuperB collaboration has started investigating a new CMOS technology called INMAPS, featuring a quadruple well structure and a high resistivity epitaxial layer, which is supposed to provide high collection efficiency as well as a high degree of tolerance to ionizing radiation and bulk damage. This technology and its properties will be described in detail in Chapter 3, along with the design features of the Apsel4well chip, developed in the CMOS INMAPS process.

Chapter 2

Radiation damage effects in deep N-well CMOS MAPS

The study of the tolerance both to ionizing radiation damage and to reticular displacement (or bulk) damage, in CMOS monolithic active pixel sensors is a crucial issue. The motivation lies in the harsh environment where these devices will be operated. Actually, the results of these studies may be useful also for applications to the space and medical environments.
The specifications of Layer0 in the SuperB SVT have been exposed in section 1.2, including the requirements in terms of radiation hardness for the detector components. The following chapter begins by describing the radiation damage mechanisms in silicon devices. Then, ionizing radiation damage effects in CMOS transistors and diodes are discussed with emphasis on thedegradation of key parameters like threshold voltage and leakage current. In the second part of the chapter, the ionizing radiation damage effects on a chip prototype of DNW MAPS developed in a planar CMOS process will be discussed. Finally, the third section will deal with the bulk damage effects on the prototypes of the same MAPS family.

2.1 Radiation damage mechanisms in silicon devices

Radiation-matter interaction can give rise to different kinds of events: ionization and reticular dislocation are the main damage mechanisms. The origin of these phenomena can be ascribed directly to the impinging particle or to secondary products. The latter is the most frequent case. Moreover, they depend on the impinging particle nature and on its energy. It is worth saying that

neutral heavy particles are the main cause of atom reticular dislocation from their lattice, while photons and electrons are mostly responsible for ionization phenomena. Heavy charged particles, like protons or ions, are responsible for both ionization and reticular displacement phenomena.

2.1.1 Ionizing radiation damage

Ionization in a generic material, whether a semiconductor or a dielectric, is responsible for electron-hole pairs generation. The number of these pairs is proportional to the energy amount adsorbed by the material, which can expressed in terms of total ionizing dose (TID). The valence band electrons are excited to the conduction band and, if an electric field is applied, they are swept through the material. In general, holes are less mobile than electrons. The production and trapping of holes in oxide films causes serious degradation in MOS devices, where oxides are used as a barrier to stop charge motion between two layers of semiconductor or metal. Actually, due to ionizing radiation:

- may temporarily lower that barrier;
- some of the charge travelling across the oxide may be caught and frozen in place, resulting in a semi-permanent charge sheet;
- the concentration of trapping levels at insulator-semiconductor interfaces may increase.

In device physics, charge trapping defects are usually localized defects at which electrons or holes may be captured. The trapped holes can be sent back in the conduction band of the considered material. In insulators, however, the required energy is generally larger than in semiconductors. In the case of SiO_2, widely used in CMOS processes, the band gap is about 9 eV and trapped holes lie at a level 2.5 eV above the conduction band. Excitation of charges from these traps can be easied by means of thermal annealing. This procedure consists of providing thermal energy to the material, which excites holes back to the conduction band. Figure 2.1 shows the charge trapping mechanism in the oxide depending on the distance from the Si-SiO_2 interface [28]. The most important piece of information contained in this figure is that long-lived hole trapping levels are confined at a distance between 5 and 20 nm from the Si-SiO_2 interface. As it will be shown in section 2.2, this means that CMOS technologies with gate oxide thinner than about 5 nm are less sensitive to charge trapping phenomena at the gate level, since charge can be rapidly annihilated or move away through tunneling effect [29].

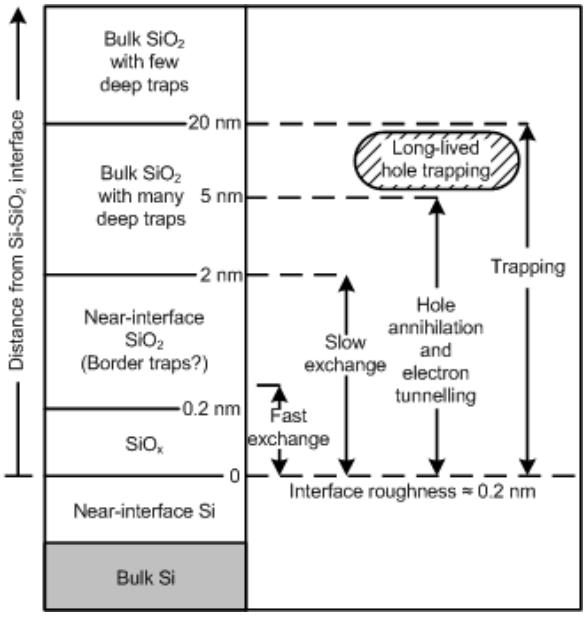

Figure 2.1. Trapping regions at the Si-SiO_2 interface [27].

2.1.2 Reticular dislocation damage

Reticular dislocation occurs when an impinging particle dislodges atoms from the silicon lattice, creating defects and changing the semiconductor properties. The loss of energy taking place as a high-energy particle passes through a solid is a complex series of momentum-sharing events (collisions) with the energy dividing itself into many small packets, some of which are dissipated as heat and others stored as "damage" [27]. Considering the case of a massive particle (proton, neutron, heavy-ion) a sort of "spike" of atomic displacements store energy as a set of terminal subclusters. These subclusters are extended defects generated by aggregation of the basic Frenkel defect pair, featuring the atom displaced by the initial collision with a heavy particle and the vacancy left by it. Many of these defects meet and recombine but a significant number survives for long times in stabilized forms. The microscopic nature depends on the amount of energy transferred to the recoil atom. The damage can be divided into three approximate regions: a region where low energy is absorbed, dominated by Frenkel defects; a region of intermediate energy where Frenkel defects and cluster damage are both important; and a third region where cascade dominates [30]. This situation is well represented in Fig. 2.2. Reticular

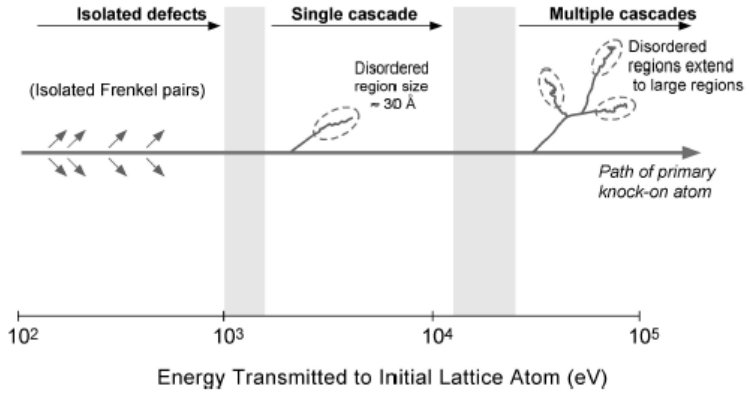

Figure 2.2. Defect structure for various recoil energies in silicon [30].

dislocation manifests itself by means of three important effects:

- mid gap states formation, favoring the transition between valence and conduction band;
- carrier trapping in the states located near the band edges. This phenomenon leads to charge trapping and release with certain time-constants;
- change of the doping profile.

The semiconductor property that is most sensitive to displacement damage is nearly always minority carrier lifetime. Lifetime damage is usully described by the following equation [31]:

$$\frac{1}{\tau} - \frac{1}{\tau_0} = \frac{\Phi}{K}, \tag{2.1}$$

where τ is the lifetime after irradiation, τ_0 is the initial lifetime, K is a damage constant, and ϕ is the particle fluence, expressed in particles per cm^2. An important aspect to be mentioned is that the lifetime damage constant K depends on the resistivity of the silicon, as shown in Fig. 2.3. A change in the minority carrier lifetime may produce different effects. In the case of an inversely biased pn junction, it leads to an increase in the generation component of the leakage current, mostly in the depleted region. In the case of a neutral

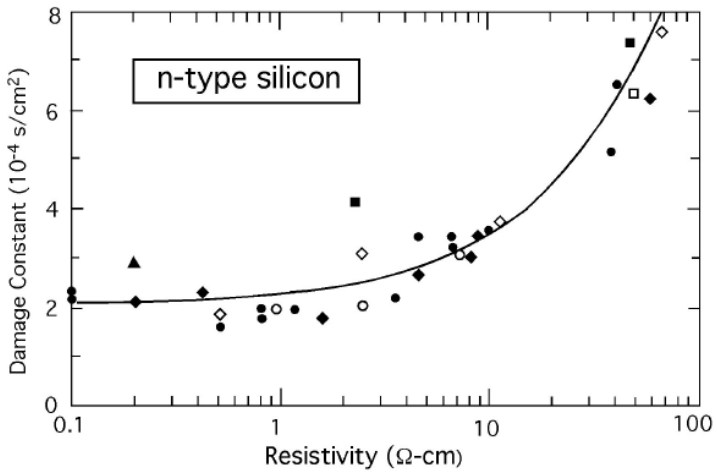

Figure 2.3. Lifetime damage constant for n-type silicon vs. resistivity [32].

semiconductor volume, increase in defect density results in an increased recombination probability for minority carriers.
Another important effect is carrier removal. This occurs when defects produced by radiation damage within the band gap act as traps for majority carriers, therefore reducing the majority carrier density through compensation. In silicon, the majority carrier traps can act as donors, limiting the effectiveness of the compensation. This might cause p-doped regions to become n-type material at very high radiation levels, while n-doped regions would remain n-type, but change their resistivity. For an n-type semiconductor (a dual expression holds for p-type semiconductor), carrier removal can be expressed as:

$$n_0 - n = R_c \Phi, \tag{2.2}$$

where n_0 is the initial doping concentration, n is the doping concentration after irradiation, R_c is the carrier removal rate, and Φ is the particle fluence.
Finally, carrier mobility is also affected by reticular dislocation damage. Changes in mobility can be expressed through the following equation:

$$\frac{\mu}{\mu_0} = \frac{1}{1 + B\Phi}, \tag{2.3}$$

where μ_0 is the initial mobility, μ is the mobility after irradiation, and B is a specific damage constant for mobility.

2.2 Ionizing radiation damage in DNW CMOS MAPS

In Chapter 1, the working principle of the deep N-well MAPS has been discussed, along with their advantages compared to the standard CMOS MAPS approach. The aim of this section is to study the effects of ionizing radiation on DNW MAPS, in particular on the Apsel3T1 chip. The same irradiation campaign involved also NMOS devices and octagonal DNW diodes with the same structure as the MAPS collecting electrode. The outcome of this work has been useful to better understand the effects of total ionizing dose on CMOS DNW MAPS performance and has been taken into account for the design of the Apsel4well chip channel readout described in Chapter 3.

2.2.1 DUTs and irradiation and test procedures

The chip Apsel3T1 was designed and fabricated in a 130 nm CMOS technology and includes a number of structures with different features:

- a 3x3 matrix (called M1) with all analog output accessible, a MIM injection capacitance for the central pixel (C_{inj}=30 fF), collecting electrode featuring a main DNW body and satellite N-well diffusions;
- A 3x3 matrix (called M2) similar to M1, but with a T-shaped DNW collecting electrode;

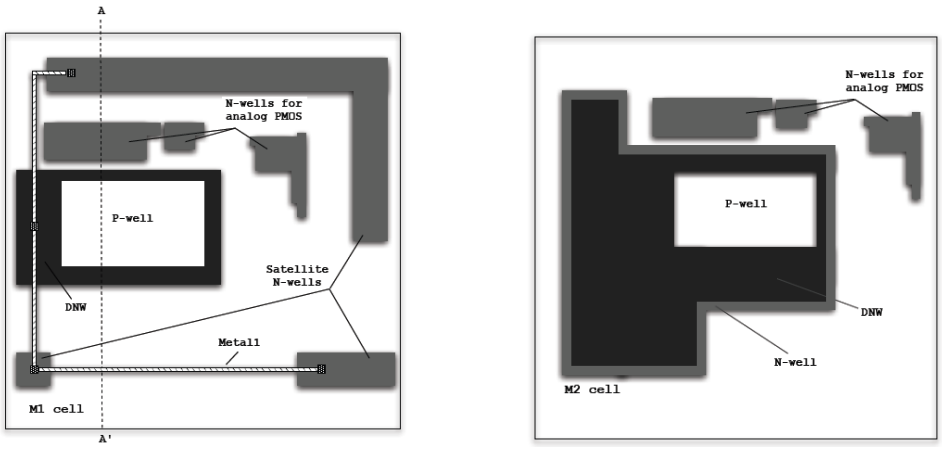

Figure 2.4. Layout of the M1 and M2 cells, with emphasis on the collecting diffusions and the competitive N-well area.

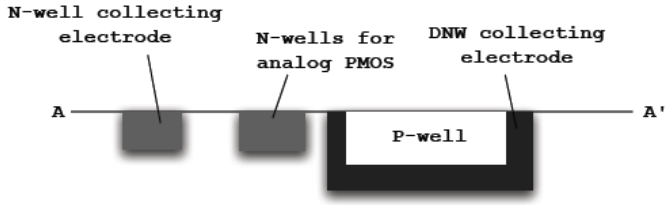

Figure 2.5. A cross section the layout in Fig. 2.4 left along AA'.

- A 8x8 matrix, with row by row, 8 parallel digital readout, one pixel with injection capacitance (C_{inj}=30 fF) and analog output accessible; half of the matrix features M1-like cells, the other half M2-like cells.

Fig. 2.4 shows a simplified layout of the pixels belonging to M1 and M2 matrices, while Fig. 2.5 represents a cross section of the M1 cell in Fig. 2.4 along AA'. In all of the tested devices, the signal coming from the sensor or from an external pulser through the injection capacitance, is processed by means of the readout chain shown in Fig. 2.6, including a charge preamplifier, a unipolar semigaussian shaping stage, a threshold discriminator and a set-reset bistable circuit acting as a memory cell. In Fig. 2.6, C_F is the preamplifier feedback capacitance, continuously reset by means of the M_F NMOS transistor operated in deep subthreshold region. Moreover, A(s) is the transfer function of the shaper gain stage, C_1 is the differentiating capacitance at the shaper input, while G_m and C_2 are the transconductance and the capacitance in the shaper feedback network. By acting on the capacitance in the high impedance node of the shaper gain stage and on G_m, the shaping time can be set to 200 or 400 ns. The signal at the shaper output is compared to a preset threshold voltage V_t. The preamplifier NMOS input device features a channel width of 14 μm and a length of 0.25 μm and is operated at a drain current I_D of 20 μA. All these parameters have been properly chosen in order to minimize the noise with a detector capacitance C_D=300 fF.
The γ-ray cobalt source used for the tests presented here featured a dose rate of about 9.5 rad(SiO_2)/s at the time of irradiation. The irradiation campaign has been performed in three different steps, for a total ionizing dose (TID) of about 10 Mrad(SiO_2). After the last irradiation step, the devices were exposed to high temperature annealing (100°C for 168 h), according to the ESCC radiation test guidelines [33]. The DUTs (devices under test) were biased at the design operating point during both the irradiation and the annealing procedures.

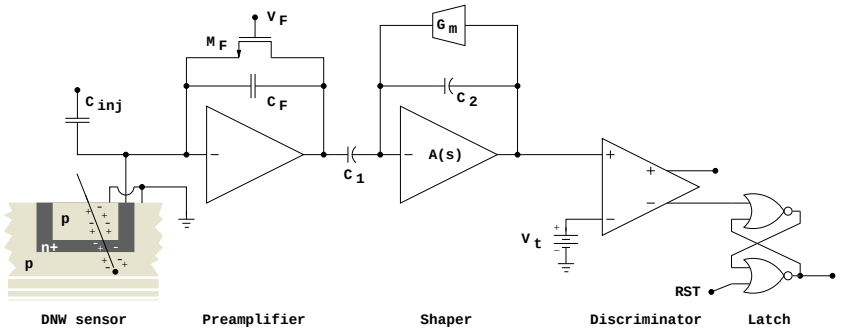

Figure 2.6. Block diagram of the analog readout chain integrated in the DNW MAPS.

2.2.2 Effects of ^{60}Co γ-rays on charge sensitivity

Exposure of the DNW MAPS to ionizing radiation is responsible for a significant decrease in charge sensitivity. This phenomenon can be observed in Fig. 2.7, in which the charge sensitivity as a function of the integrated dose for the accessible pixel of an M2 (left) and M3 (right) matrix is represented. The data, which also include post-annealing measurement results, are plotted for the two available peaking times, 200 and 400 ns. Before irradiation, the charge sensitivity is about 1.5 V/fC at a peaking time t_p=200 ns, about 1.3 V/fC at t_p=400 ns. The decrease after exposure to a 900 krad(SiO_2) γ-ray dose is about 7% at both peaking times. After irradiation with an integrated dose of 9700 krad(SiO_2), the charge sensitivity is found to be around 75-80% of the pre-irradiation value, depending on the peaking time. The annealing cycle brings the charge sensitivity back to a state not far from its pre-irradiation condition. The mechanisms underlying such a behavior are briefly summarized in the following. In particular, charge sensitivity was found to decrese as a result of the threshold voltage shift in the feedback NMOS M_F of the charge preamplifier, only partially compensated by the leakage current increase in the detector, and as a consequence of a change in the feedback transconductance and in the gain-bandwidth product of the shaping stage.

Threshold voltage shift in M_F

Although deep submicron CMOS technologies are acknowledged to possess quite a high degree of radiation tolerance (as mentioned in section 2.1.1), nevertheless, N- and PMOS devices, belonging to the 130 nm CMOS node

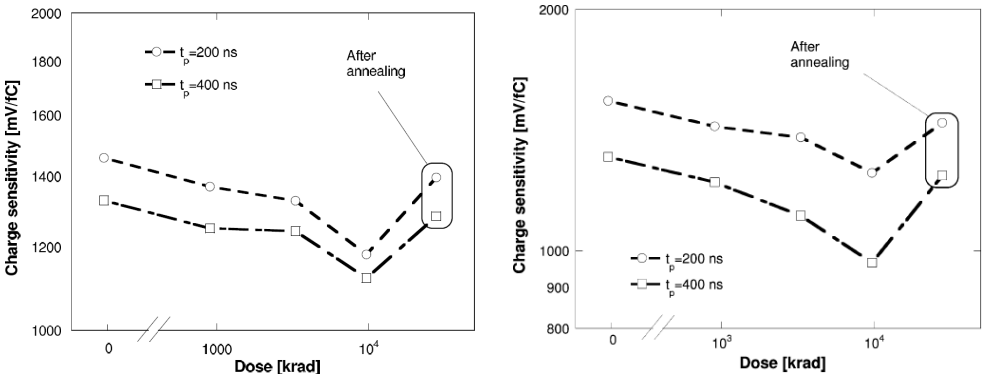

Figure 2.7. Charge sensitivity as a function of the total ionizing dose and after the annealing procedure for the accessible pixel of an M2 (left) and an M3 (right) matrix. Data are plotted for the two available peaking times.

and featuring a channel width W smaller than about 1 μm, have been proven to undergo the so called radiation induced narrow channel (RINC) effect, involving charge trapping in the shallow trench isolation (STI) oxides. This has been proposed as the main cause for a substantial change in the drain leakage current and in the threshold voltage in such devices [34]. This phenomenon is likely to affect M_F in the preamplifier feedback network, featuring W=0.18 μm and channel length L=10 μm. In order to find evidence of the RINC effect also in the technology used for the Apsel3T1 MAPS devices, a couple of transistors, NM1 and NM2, belonging to the same tecnology and with the same aspect ratio as M_F, have been exposed to different integrated doses of γ-rays. Fig. 2.8 shows the variation in the threshold voltages of NM1 and NM2. The change is in the order of few tens of mV, much larger than found in transistors with larger W from the same process [29]. A reduction of the M_F threshold voltage implies an increase of its output conductance $g_{ds,F}$ and, as a consequence, a reduction of the charateristic reset time constant of the preamplifier given by:

$$\tau_F = \frac{C_F}{g_{ds,F}}. \tag{2.4}$$

A partial recovery can be detected for NM2 just after the last irradiation step, as a possible consequence of the interface state negative charge starting to compete with the oxide-trapped positive charge. This phenomenon has already been observed in the technology where the RINC effect was recognized for the first time [34]. On the other hand, no trace of such a rebound-effect can be

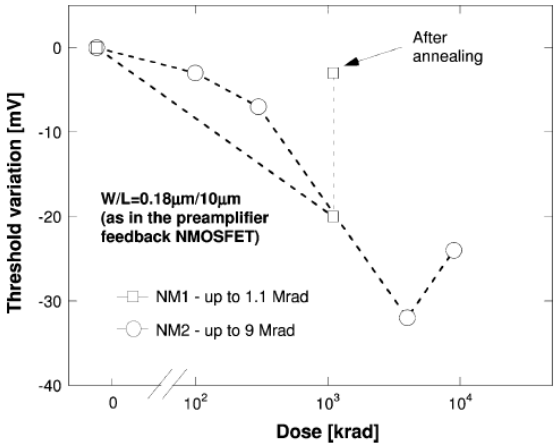

Figure 2.8. Variation of the threshold voltage in a couple of irradiated narrow channel NMOSFETs.

found in Fig. 2.7. This may be due to the operating conditions of NM2 in Fig. 2.8 being different from those of M_F. Also, the starting TID value for the rebound effect may vary by several Mrad from device to device in the same technology [34].

Leakage current increase in the detector

In order to study the ionizing radiation effects on the DNW detector, a set of three octogonal DNW diodes have been tested. Their geometrical features are listed in Table 2.1. Figure 2.9 shows the leakage current in diodes A, B and C for a reverse voltage spanning from 0 to 10 V at varying TID. An important increase can be observed over the whole reverse voltage interval. The buildup of holes in the field oxide and the creation of trapping states at the Si/SiO_2 interface over the PN junction may account for an increase of the surface

Name	**area** [μ**m**^{2}]	**perimeter** [μ**m**]
A	55×10^3	1194
B	16.1×10^3	638
C	4×10^3	318

Table 2.1. Geometrical features of the tested octagonal DNW diodes.

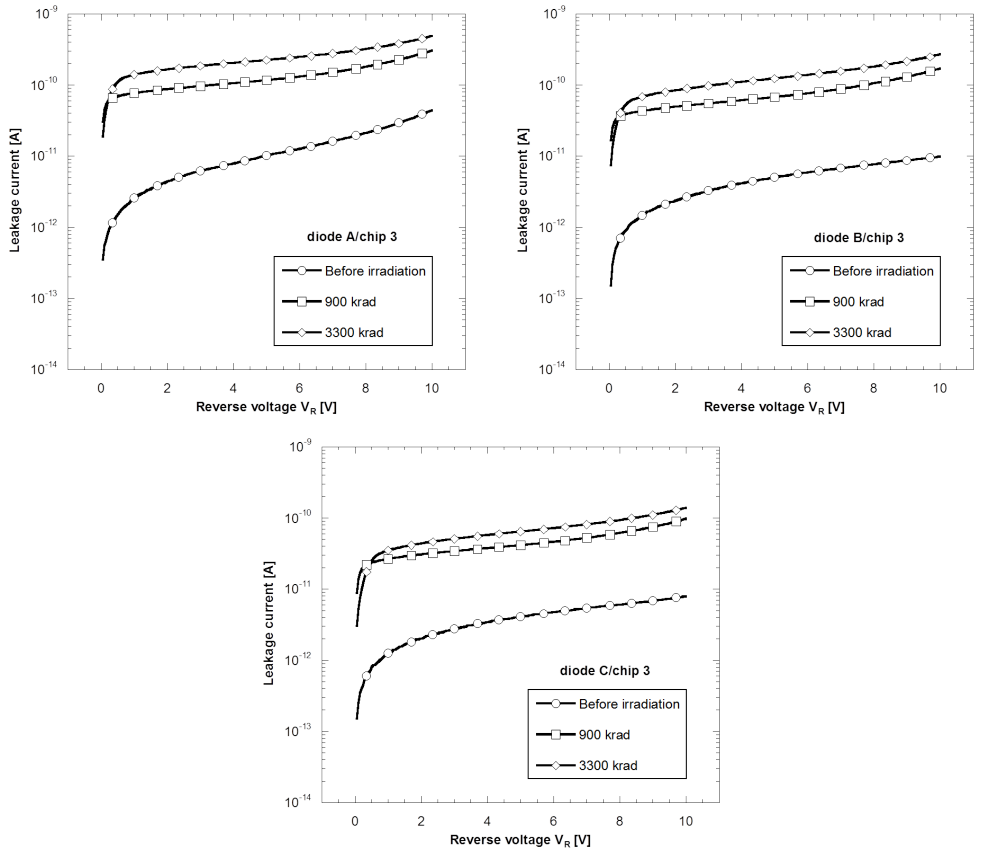

Figure 2.9. Leakage current as a function of the reverse voltage in diodes A, B and C for different irradiation steps.

generation current [35] [36]. Based on the involved damage mechanism, the leakage current is supposed to be proportional to the length of the junction at the silicon surface. This is supported by the results shown in Fig. 2.10, where the data points can be fitted fairly well to a linear function, demonstrating that no significant bulk damage effect, which would be roughly linearly dependent on the junction area, contributes to the leakage current increase. An increase in the leakage current leads to a decrease of $g_{ds,F}$, but this effect only partially compensate the increase of $g_{ds,F}$ given by the M_F threshold voltage decrease.

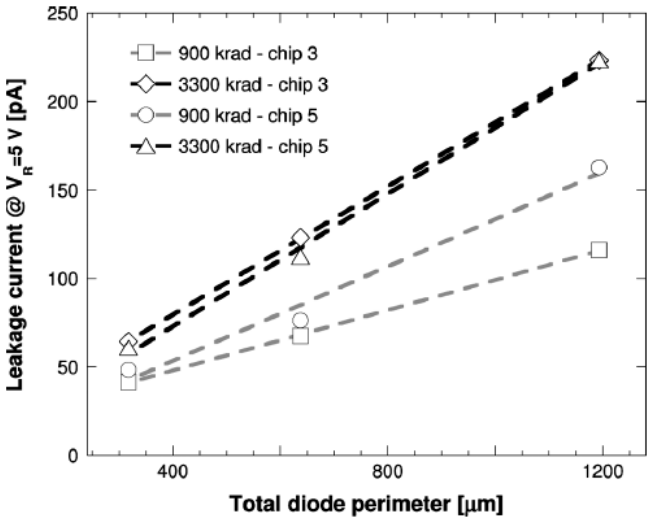

Figure 2.10. Leakage current at a reverse voltage of 5 V is plotted as a function of the junction perimeter length for the three test diodes in Table 2.1.

Change in the feedback transconductance and in the gain-bandwidth product of the shaper

The schematic of the shaping stage, featuring a cascode input stage and an output source follower, is represented in Fig. 2.11. An additional contribution to the reduction of the charge sensitivity is given by the RINC effect affecting some transistors of the shaping filter. Radiation induced change in the threshold voltage of the PMOS current source MP1 (featuring W=0.18 μm) in the shaper input stage is responsible for a decrease in the biasing current and, as a consequence, for a reduction of the gain-bandwidth product:

$$GBP = \frac{k_2 g_{m,MN1,0}}{C_L} = k_2 GBP_0, \tag{2.5}$$

where $g_{m,MN1,0}$ is the transconductance of the input device, C_L is the capacitance loading the output node of the shaper input branch and $k_2<1$ accounts for radiation effects on the shaper input device transconductance. On the other hand, a threshold voltage shift in the transconductor current sources (all featuring W=0.25 μm), may result in an increase of the feedback transconductance:

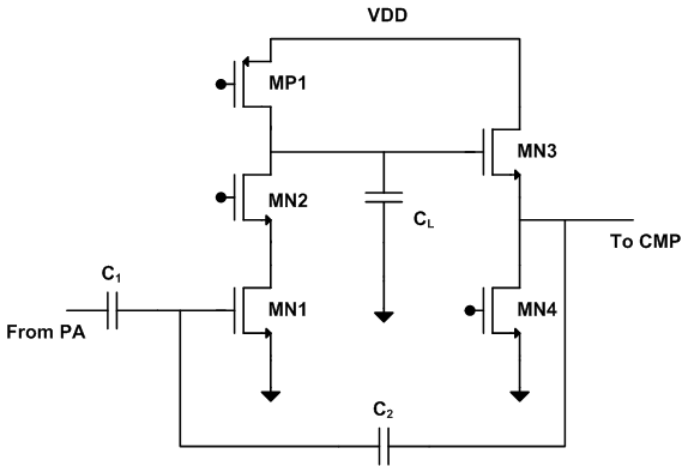

Figure 2.11. Transistor-level representation of the shaping stage.

$$G_m = \frac{qk_1 I_{t0}}{2nk_B T} = k_1 G_{m0}, \tag{2.6}$$

where q is the elementary charge, I_{t0} is the pre-irradiation value of the transconductor bias current, n is a technological parameter, k_B is the Boltzmann constant, T is the absolute temperature and the parameter k_1>1 takes into account the effects of radiation on G_m. Equation 3.1 expresses the relationship between the shaper output voltage and the injected charge Q taking into account the effects of radiation by means of k_1 and k_2:

$$v_{out,s}(t) = -2\frac{C_1 Q}{C_F C_2}\frac{e^{-2k_2 G_{m0}/C_2 t}}{\sqrt{\left(\frac{k_1}{k_2}-1\right)}} sin\left(\frac{t}{\tau_1}\right), \tag{2.7}$$

where:

$$\tau_1 = \frac{C_2}{2G_{m0}}\sqrt{k_1 k_2 - k_2^2}. \tag{2.8}$$

Tts radiation induced relative variation of the charge sensitivity G_Q=max[$v_{out,s(t)}$]/Q,

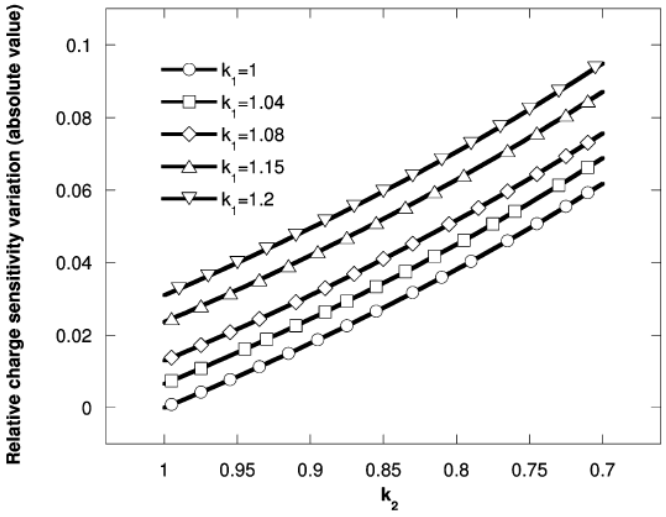

Figure 2.12. Computed relative charge sensitivity variation as a function of the parameter k_2, plotted for different values of the parameter k_1.

$$\frac{\Delta G_{q,s}(k_1,k_2)}{G_{Q,s}} = G_{Q,s}(1,1) - \frac{G_{Q,s}(k_1,k_2)}{G_{Q,s}(1,1)}, \tag{2.9}$$

has been computed and plotted, in the absolute value form, in Fig. 2.12 as a function of the parameter k_2 for different values of k_1. Note that the charge sensitivity decreases both for k_1 increasing and for k_2 decreasing. Finally, the transconductance increase is expected to speed up the return to the baseline of the shaper response. This is actually confirmed by the waveforms in Fig. 2.13, illustrating the radiation effects on a DNW MAPS to an input charge of 750 electrons. Return to baseline grows faster with increasing TID. Note that, after annealing, the pixel response in both the shown cases is restored virtually to its original state.

2.2.3 Effects on noise performance

Equivalent noise charge (ENC) is a figure of merit for analog processors, defined as the charge to be injected at the input of a charge measuring system to have a peak response as high as the output mean square noise [37]. In the MAPS under test, the main noise contribution comes from the fluctuations in

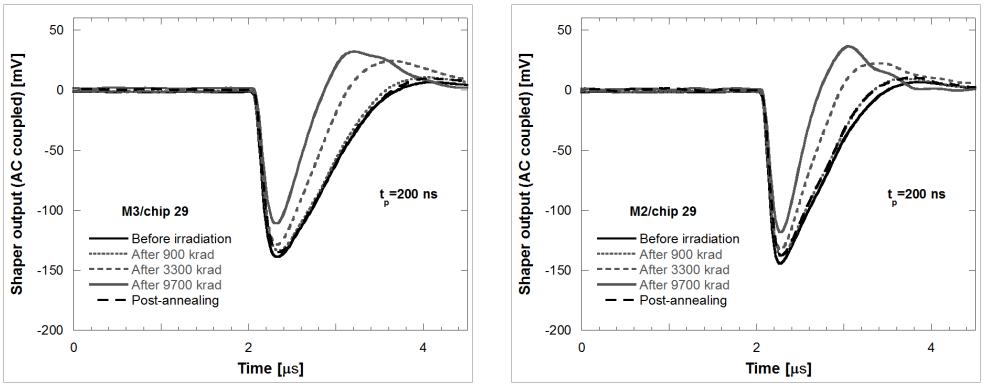

Figure 2.13. Radiation effects on the amplitude and shape of the MAPS response to a 750 electron input charge.

the drain current of the preamplifier input device. The noise in the preamplifier input transistor can be represented by means of a voltage source in series with the device gate. Its power spectral density $S_e{}^2(f)$ is given by:

$$S_e^2 = S_w^2 + \frac{A_f}{f^\alpha} = n\gamma\frac{4k_BT}{g_m} + \frac{K_f}{WLC_{ox}f^\alpha}. \tag{2.10}$$

The previous equation includes a frequency independent component, $S_w{}^2$, mainly accounting for the channel thermal noise in the input device drain current, and a 1/f term, A_f/f^α. In the equation, n is the subthreshold slope coefficient, γ is the channel thermal noise coefficient, depending on the device operating point and polarity, k_B is the Boltzmann's constant, T is the absolute temperature, g_m is the device transconductance , K_f is an intrinsic 1/f noise process parameter, C_{ox} is the gate oxide specific capacitance and α is the 1/f noise slope coefficient. In transistors with gate area of the same order of the preamplifier input device or larger, the low frequency spectrum, both before and after irradiation, was found to be dominated by flicker noise. This is shown as an example in Fig. 2.14 for an NMOS with W/L=20/0.2 irradiated with a 1.1 Mrad(SiO_2) γ-ray dose. ENC has been calculated as the ratio between the noise root mean square measured at the shaper output by means of a digital scope and the charge sensitivity. In an analog processor for capacitive detectors, the ENC takes the following, rather general, expression [38]:

$$ENC^2 = C_T^2\left\{A_1S_w^2\frac{1}{t_p} + A_2(2\pi)^\alpha A_f t_p^{\alpha-1}\right\} + 2qI_{leak}A_3t_p, \tag{2.11}$$

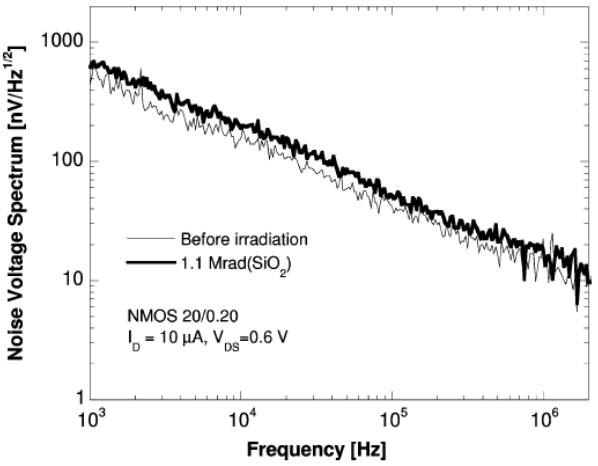

Figure 2.14. Noise voltage spectrum before and after a 1.1 Mrad(SiO_2) γ-ray dose for an NMOS device with W/L=20/0.2 belonging to the 130 nm CMOS process used for DNW MAPS fabrication.

where the two terms in curly brackets are due to channel thermal noise and flicker noise in the preamplifier input device respectively, while the last term accounts for shot noise in the sensor leakage current. In the same equation, C_T is the total capacitance shunting the charge preamplifier input, A_1, A_2 and A_3 are the shaping coefficients, ${S_w}^2$ is the power spectral density of the series white noise in the preamplifier input transistor, dominated by the thermal noise of the device channel, A_f is the flicker noise power coefficient of the same device, with slope coefficient α (about 0.85 for NMOS devices in this 130 nm process [38]), q is the elementary charge and I_{leak} is the leakage current of the detector. Actually, at the considered peaking time, the last term in (3.41) can be safely neglected. The effects of radiation on the ENC are shown in Fig. 2.15, where the ENC is plotted as a function of the absorbed dose and after the annealing cycle for the central pixel of an M1 matrix at the two available peaking times. A significant increase in equivalent noise charge can be detected, from about 40 electrons to between 110 and 150 electrons, depending on the peaking time, after exposure to radiation. The annealing procedure brings the noise performance back to somewhere between the starting DUT state and the post-900 krad step condition. It is worth recalling that the charge sensitivity features a qualitatively similar response, with a remarkable performance recovery after annealing, possibly due to partial annihilation

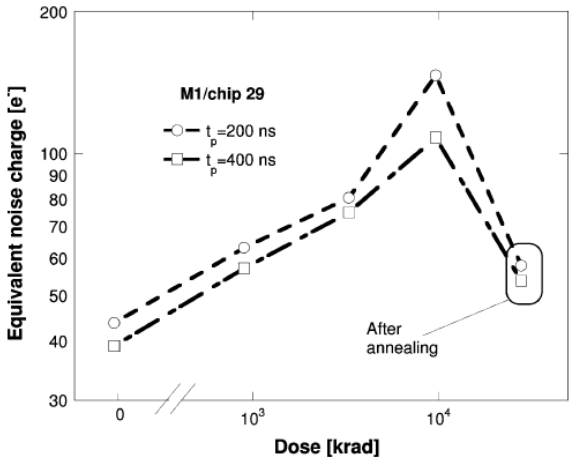

Figure 2.15. ENC as a function of the absorbed dose and after the annealing cycle for the central pixel of an M1 matrix. ENC is plotted for the two available peaking times.

of the positive charge trapped in the STI. ENC degradation after irradiation is likely to originate from low frequency noise increase in the preamplifier input device. Flicker noise gets worse as a consequence of parasitic lateral transistors being turned on by positive charge buildup in the shallow trench isolation oxides and contributing to the overall noise (see Fig. 2.16). The role of STI and parasitic lateral transistors in 1/f noise degradation has been emphasized in deep submicron technologies. It was probably less significant, if not negligible, in older CMOS generations, where the contributions from border trap density increase and charge buildup in the transistor gate oxide were predominant [39], [40]. In irradiated devices, flicker noise increase can be expressed through the ratio between the post-irradiation coefficient $A_{f,post}$ and its pre-irradiation value $A_{f,pre}$:

$$\frac{A_{f,post}}{A_{f,pre}} = \frac{K_{f,m,post}}{K_{f,m,pre}} \frac{1 + \frac{K_{f,lat}}{K_{f,m,post}} \frac{WC_{ox}}{W_{lat}C_{ox,eff}} \frac{g_{m,lat}^2}{g_{m,m,post}^2}}{\left(1 + \frac{g_{m,lat}}{g_{m,m,post}}\right)^2}. \tag{2.12}$$

In (2.12), $K_{f,m,pre}$, is the pre-irradiation intrinsic process parameter for 1/f noise, $K_{f,m,post}$ and $K_{f,lat}$, are post-irradiation intrinsic process parameters in the main and in the parasitic transistor respectively, $C_{ox,lat}$ is the effective ox-

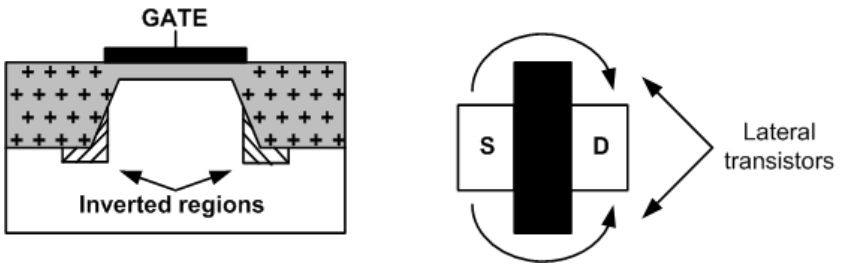

Figure 2.16. Cross-sectional (left) and top (right) view of a MOSFET with positive charge trapping in the STI.

ide capacitance of the parasitic STI sidewall transistor, $g_{m,m,post}$ is the channel transconductance in the main device and $g_{m,lat}=2n_f g_{m,lat,s}$, with n_f the number of finger composing the main transistor and $g_{m,lat,s}$ the transconductance of the single parasitic device. Note that for each finger, two single parasitic devices get switched on by the charge trapped in the STI, each with channel length L and channel width $W_{lat,s}=W_{lat}/2n_f$. Under the simplifying hypotesis, generally valid in the low power operation adopted for DNW MAPS, that the main and the single parasitic devices are made to work in weak and strong inversion respectively [41], then

$$g_{m,lat,s} \simeq \sqrt{2\mu_n C_{ox,eff} \frac{W_{lat,s}}{L} I_{D,lat,s}}, \tag{2.13}$$

$$g_{m,m,post} \simeq \frac{I_{D,m,post}}{n\phi_t}, \tag{2.14}$$

with μ_n the electron mobility in silicon, $I_{D,lat,s}$ the drain current in the single parasitic transistor, $I_{D,m,post}$ the post-irradiation drain current in the main device and ϕ_t the thermal voltage. If $g_{m,lat}/g_{m,m,post} \ll 1$ is also assumed, (2.12) can be rewritten as:

$$\frac{A_{f,post}}{A_{f,pre}} = \frac{K_{f,m,post}}{K_{f,m,pre}} \left(1 + 2n_f \frac{K_{f,lat}}{K_{f,m,post}} \frac{nI_z^*}{L} \frac{W}{I_{D,m,post}} \frac{I_{D,lat,s}}{I_{D,m,post}} \right), \tag{2.15}$$

where $I_z{}^*=2\mu_n C_{ox} n\phi_t{}^2$ is the normalized drain current marking the boundary between weak and strong inversion [42]. Equation (2.15) highlights that,

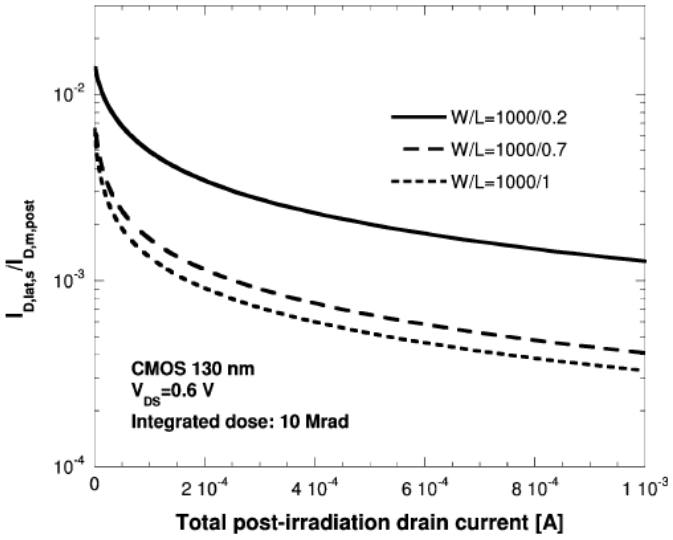

Figure 2.17. Relative variation of the drain current with respect to the post-irradiation value upon absorption of a 10 Mrad γ-ray dose. Data are plotted as a function of the total post-irradiation current.

for a given CMOS process, flicker noise degradation is larger in devices with larger number of fingers n_f and smaller current density $I_{D,m,post}/W$. Radiation induced flicker noise increase was experimentally proven to affect in a more pronounced way NMOSFETs with many fingers and operated at small current densities, where the parasitic devices have a greater impact on the behavior of the main transistor. On the other hand, experimental data show that $I_{D,lat,s}/I_{D,m,post}$ is a decreasing function of the total post-irradiation drain current $I_{D,tot}=I_{D,m,post}+2n_f I_{D,lat,s}$, corresponding to the current forced into the device drain terminal. This phenomenon is shown in Fig. 2.17, where $I_{D,lat,s}/I_{D,m,post}$ is plotted as a function of $I_{D,tot}$ upon absorption of a 10 Mrad γ-ray dose for a set of NMOSFET with W=1000 μm and varying L. Eq. 2.15 provides some guidelines for rad-hard front-end design. In particular, the smaller is the current in the preamplifier input device, the higher is the noise performance degradation with increasing TID. Moreover, since the noise increase with the radiation dose is proportional to the number of fingers n_f, an enclosed layout structure for the preamplifier input transistor is mandatory if low noise performance under irradiation is required.

2.2.4 Gain calibration with ^{55}Fe source

The effects of ionizing radiation on the charge collection properties of the Apsel3T1 chips under examination have been also tested by means of an ^{55}Fe source and a low power, infrared laser source. ^{55}Fe sources can be used to provide charge sensitivity and noise calibration for pixels without injection capacitance and to validate measurements obtained with other methods. ^{55}Fe X-rays release their entire energy in the detector substrate through photoelectric interaction. Photons from the ^{55}Fe 5.9 keV line generate about 1640 electron/hole pairs each. When photons convert into the junction depleted region, the released charge is virtually entirely collected, yielding the peak in the ^{55}Fe spectrum (actually, as discussed later on, also the DNW and N-well collecting electrodes and the P-well inside the DNW seems to be involved). Charge released far from the junction and only partially collected is responsible for the pedestal at lower amplitudes. This mechanism is shown in Fig. 2.18, representing the schematic cross section of a DNW MAPS hit by ^{55}Fe photons, while Fig. 2.19 shows the event count rate distribution for a pixel in an M1 matrix and in an M2 matrix and compares pre-irradiation, post-irradiation and post-annealing spectra. These data are in good qualitative agreement with the charge sensitivity and noise discussed in the previous sections. As the absorbed dose increases, the 5.9 keV peak gets broader as a consequence of the noise increase. At the same time, the peak is shifted towards lower amplitude values, as a result of the charge sensitivity decrease. Immediately after the first irradiation step, a non negligible loss in the count rate can be detected in the peak region. Since the count rate in the peak is proportional to the depleted substrate volume, this result might hint at a reduction in the thickness of the space charge region at the DNW-to-substrate junction. This in turn may be

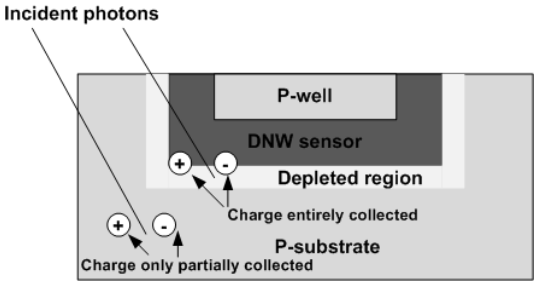

Figure 2.18. Schematic cross section of a DNW MAPS with ^{55}Fe photons releasing energy in different positions of the detector volume.

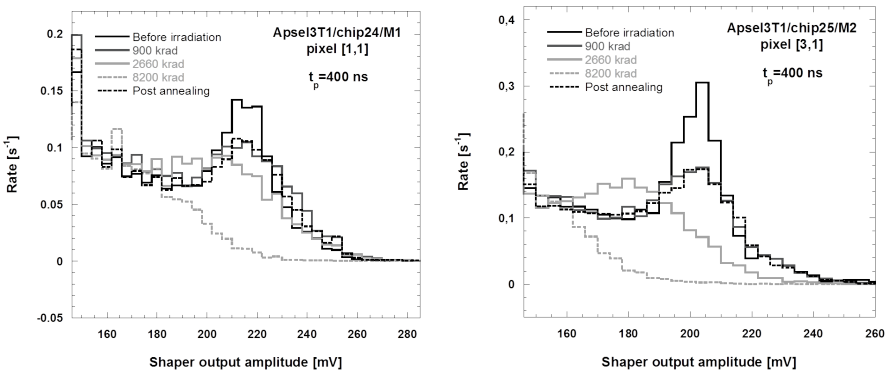

Figure 2.19. Event count rate for a pixel of an M1 (left) matrix and M2 (right) matrix tested with an ^{55}Fe source.

possibly due to a decrease in the DNW potential as a consequence of a change in the gate voltage of the preamplifier input transistor. After the annealing procedure, the ^{55}Fe spectrum recovers almost to the original condition, though in the presence of an event rate loss with respect to the pre-irradiation measurement result. Moreover, the event count rate is higher for a pixel in an M2 matrix than in the case of an M1 matrix. The reason for this difference will be explained in Sec. 2.3.1. It is worth noticing here that the charge sensitivity values obtained from ^{55}Fe measurements, between 800 and 900 mV/fC, while in good agreement with circuit simulations, are quite smaller than those obtained with the charge injection method. Actually, while the amount of charge released by an ^{55}Fe photon, corresponding to the spectrum peak, is precisely known, the knowledge of the charge injected when using the other method is based on the knowledge of the amplitude of the voltage step, also precisely controlled through the front panel of the external pulser, and of the value of the injection capacitance. A discrepancy between the nominal value of C_{inj}, as defined by its as-drawn dimensions and by the foundry models, and its actual value in the test structures may explain the difference in the results provided by the two measurement methods, which was detected in all of the tested samples. No such difference was ever found in DNW MAPS fabricated in previous multi-project runs. This seems to point to the fact that, in this particular run, MIM capacitance parameters might have fluctuated very far from typical models, leading to a larger than the nominal 30 fF injection capacitor value. On the other hand, no stray capacitances of any significant value in parallel with C_{inj} were detected by parasitic extraction tools.

2.2.5 Charge collection properties

The available structures have been tested by means of a laser source providing further information about the prototype behavior. The detector performance were assessed by stimulating the central pixels in the M1 and M2 matrices by means of a measurement setup for precision scanning with an infrared laser, with a wavelength λ=1064 nm. The beam spot features a Gaussian intensity profile with a standard deviation of about 3 μm in both dimensions of the plane transversal to the beam propagation. The measurement setup is illustrated in Fig. 2.20 and includes, besides the laser source and the proper optical instrumentation (coupler and fiber focuser), a Newport universal motion controller (3 axis) and a LeCroy Waverunner 64Xi digital storage oscilloscope (600 MHz bandwidth, 10 Gs/s maximum sampling rate), both controlled by means of an NI Labview virtual instrument. The waveform generator, used to control the pulsed operation of the laser source, provides the trigger signal for the oscilloscope through the sync output. The need for an automatic setup arises from the relatively large number of incidence points necessary to perform a full characterization of a MAPS matrix with an incremental step, along both the X and Y directions, not larger than 5 μm. For each laser position point, the virtual instrument acquires the waveforms from three oscilloscope chan-

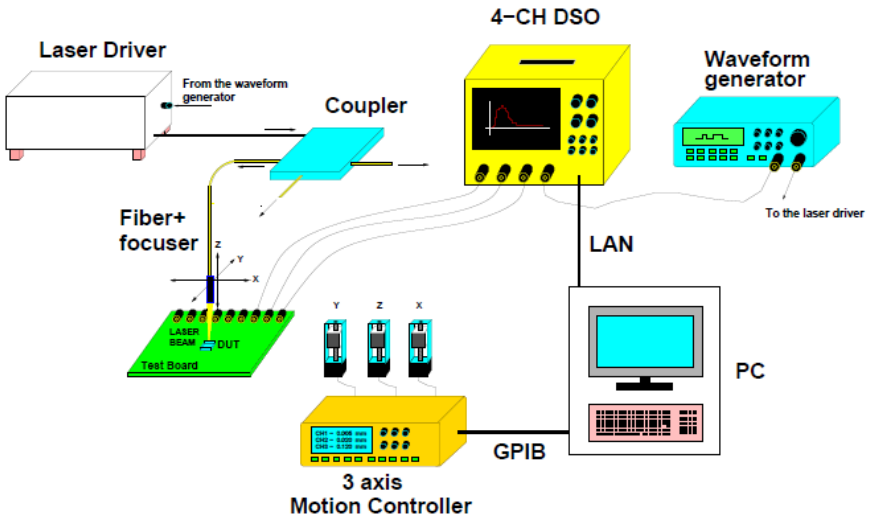

Figure 2.20. Measurement setup used in the laser scan tests of the MAPS sensors

nels and calculates the position of the next point. Moreover, for each channel, it provides the possibility to calculate the average waveform thus improving amplitude measurement accuracy. At the end of the acquisition, measurement results are summarized in a text file with indications about the collision points and the relevant signal amplitude at the shaper output. The DUTs were illuminated from the back side in order to avoid reflections from the intricate net of interconnections laid out on the front side. Note that, in this configuration, the light, after going through the device bulk, is reflected back into the substrate by the metals at the chip top. Since the aim of the laser test is to evaluate the relative variation of the collected charge with the laser spot position inside, reflection is not expected to significantly affect the measurement results, as long as it is uniform over the cell. The latter is a reasonable assumption, as the chip top side is virtually fully covered with metal layers.

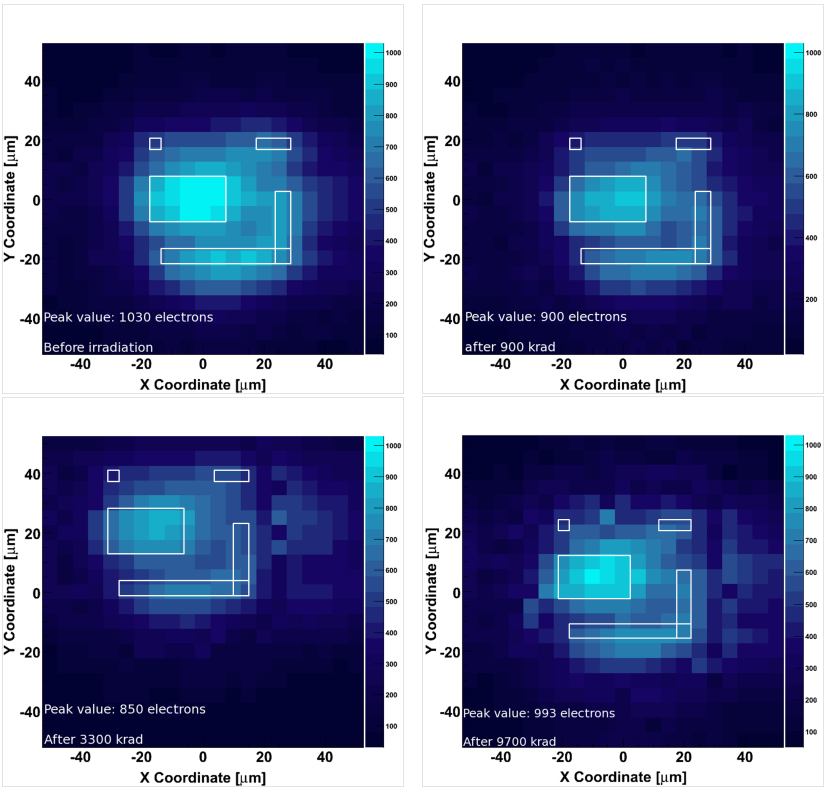

Figure 2.21. Charge collected by the central pixel of an M1 matrix. The figures also show the sensor layout.

Fig. 2.21 shows the charge collected by the central pixel of an M1 matrix as a function of the laser beam position after each irradiation step, up to a 9700 krad(SiO_2) total ionizing dose. The figure also shows the layout of the sensor, with a main body and three N-well satellite diffusions. The collected charge was calculated as the ratio between the amplitude of the shaper output response to the laser pulse and the charge sensitivity measured by means of charge injection through an external pulser. The point corresponding to the maximum collected charge value was taken as the (0,0) coordinate. Before irradiation, the peak value of the collected charge is about 1000 electrons, which is close to the signal from a minimum ionizing particle (MIP). After the last irradiation step, a small decrease (about 5%) in the peak value of the charge collected by the DNW sensor can be noticed. This very limited effect might be ascribed to the small amount of non-ionizing energy released by γ-rays in the sensor substrate, which might be responsible for the creation of traps in the bulk (although with a very small density) through Compton scattering and, as a consequence, for an increase in the carrier recombination probability. Actually, such a small variation might also be explained with errors in the measurement setup, due for instance to slight differences in the distance between the DUT and the laser focuser from measurement to measurement. This latter hypothesis seems to find support in Fig. 2.22, where the peak value

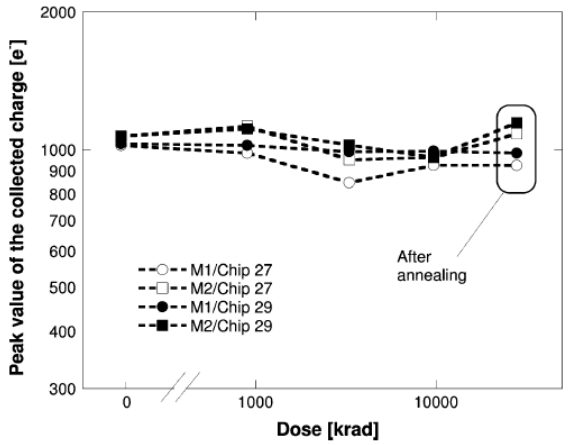

Figure 2.22. Peak value of the collected charge in the central pixel of M1 and M2 matrices illuminated with an infrared laser beam. Data are plotted as a function of the integrated dose and after the annealing cycle.

of the collected charge in the central pixel of M1 and M2 matrices is plotted as a function of the integrated dose and after the 100 °C 168 h cycle. Although the data values at 9700 krad dose are from 5 to 10% lower than the pre-irradiation values for all of the tested devices, nevertheless no clear trend with the dose can be easily determined. Also, the leakage current test results discussed in Section 2.2.2 seem to validate the hypothesis of a very limited, if not completely absent, bulk damage. In fact these tests confirm the linear dependance of the leakage current increase on the diode perimeter, demonstrating no significant bulk damage effect, which would be roughly linearly dependent on the junction area.

2.3 Bulk damage in DNW CMOS MAPS

In section 2.2 the irradiation campaign with ionizing radiation performed on DNW CMOS MAPS has been thoroughly discussed, along with its effects on the circuit performance. Other Apsel3T1 chips have been tested in order to study the bulk damage effects. The devices under test were exposed to neutrons from the reactor core of the Triga MARK II nuclear research facility located at the Jozef Stefan Institute (Ljubljana, Slovenia). The devices were kept unbiased during the irradiation. The final integrated fluence ranges from 6.7×10^{12} to 3.7×10^{13} 1 MeV NIEL-equivalent neutrons/cm^2 for different samples. The maximum fluences have been reached after 4 or 5 irradiation steps, depending on the particular sample. The quarantine period spanned from a few days, after irradiation with a neutron fluence of 2×10^{11} n/cm^2, to about 1 month, after exposure to 3×10^{13} n/cm^2. During this time, the devices were kept refrigerated at about -18 °C to minimize the annealing effects. Similarly to what has been shown in the previous section, measurements of the circuit performance have been performed through a number of different techniques, including electrical characterization of the front-end electronics and of the DNW diodes, laser stimulation of the sensors and tests with ^{55}Fe and ^{90}Sr radioactive sources.

2.3.1 Effects of neutron irradiation on noise and charge sensitivity

Fig. 2.23 shows the equivalent noise charge and the charge sensitivity (both normalized to the pre-irradiation value) as a function of the neutron fluence for the central pixel (the one provided with the injector capacitor) of M1 and M2 operated at t_p=400 ns. As in the previous section, the charge sensitivity has been evaluated by linearly interpolating the peak response of the front-end

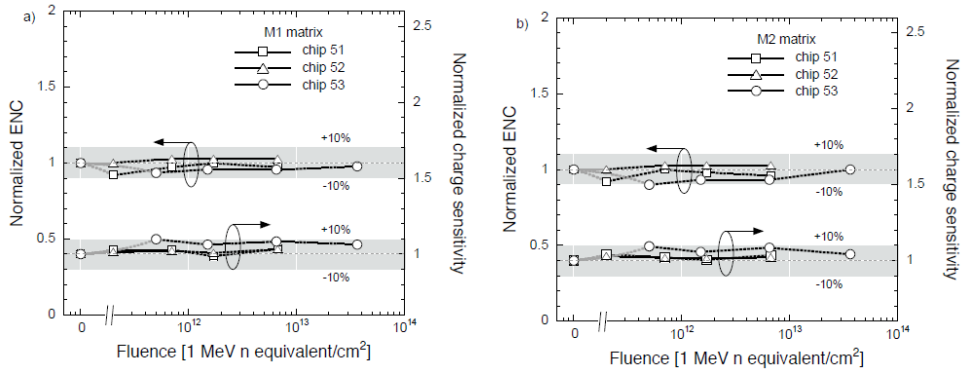

Figure 2.23. Normalized equivalent noise charge and charge sensitivity as a function of the neutron fluence for the central pixel of the (left) M1 and (right) M2 matrices in chips 51, 52 and 53.

channel to a charge pulse with varying amplitude. In the set of available samples, the standard deviation of the measured charge sensitivity was found to be about 7% of the average value both before and after irradiation. Taking into account the expression of the ENC in (3.41), the last term is usually negligible for the considered devices, especially for peaking times lower than 1 μs, but should be monitored in irradiated devices, where the sensor leakage current can increase as a consequence of radiation damage at the silicon/silicon dioxide interface and in the semiconductor bulk, due to ionizing and non-ionizing radiation, respectively. The effects of bulk damage on the leakage current of DNW sensors will be thoroughly discussed in section 2.3.3. Data in Fig. 2.23 show that the equivalent noise charge remains virtually unaffected by neutron irradiation. The limited variations, not exceeding 10% of the pre-irradiation value, around 40 electrons rms on average, are compatible with measurement errors due to relatively small temperature changes. Actually, no neutron induced degradation is expected for the front-end electronics, since MOSFET transistors, whose operation is based on the drift of majority carriers at the device channel surface, are known to be largely insensitive to bulk damage [43] [44]. On the other hand, ionizing radiation may have a negative impact on the noise performance of DNW MAPS. Note that γ-rays are always present in the environment of a nuclear reactor as a byproduct of fission. Therefore, the absence of significant changes in the ENC, besides demonstrating that neutrons do not affect the noise properties of CMOS circuits, also indicates that, during irradiation, background ionizing radiation from the reactor core was below the sensitivity threshold of the considered CMOS process. Fig. 2.23 also shows that

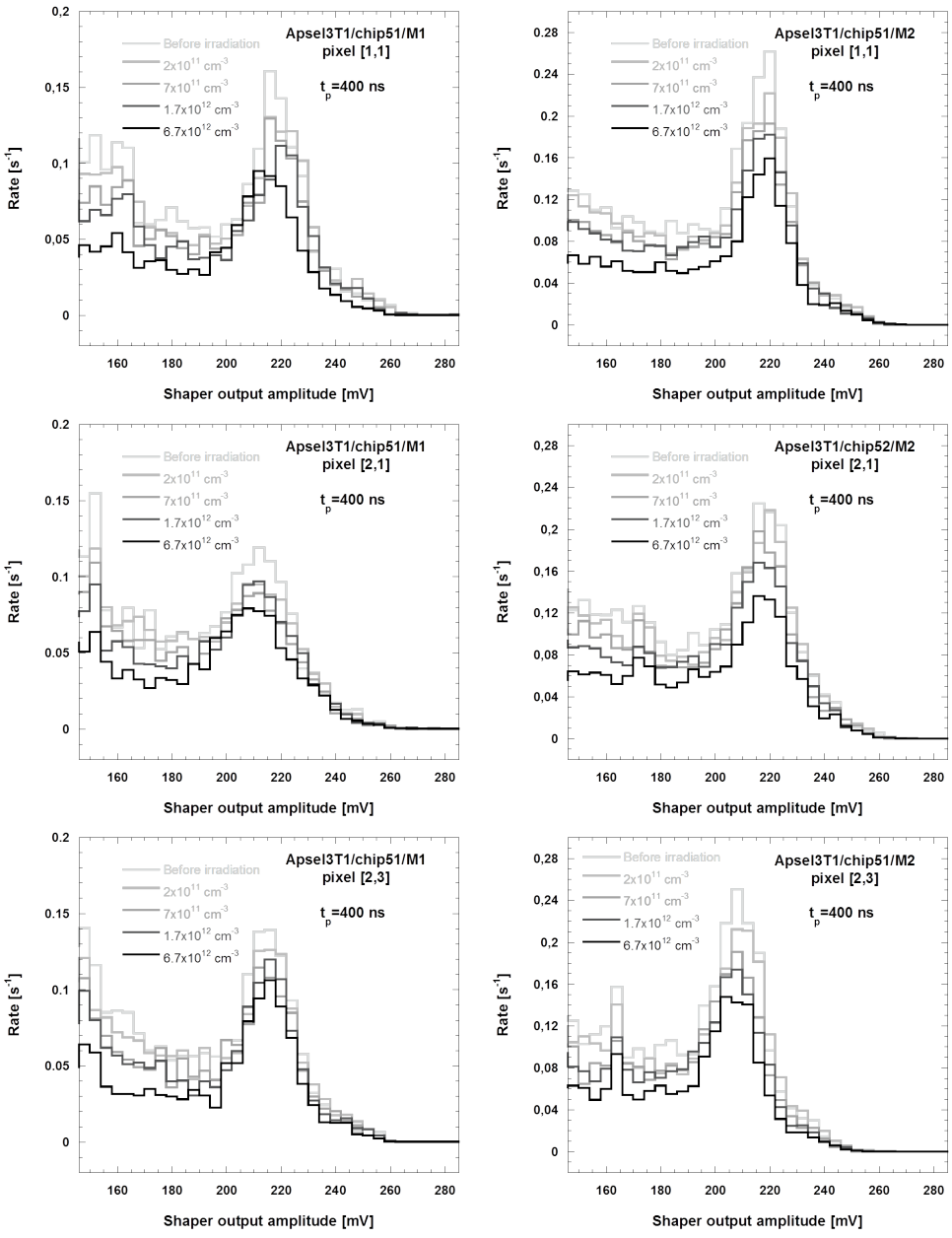

Figure 2.24. Event count rate distribution before irradiation and at different integrated fluences for pixels from M1 (on the left) and M2 (on the right) matrices.

charge sensitivity is not affected by neutron irradiation with variations below 10% not following a specific trend with increasing fluence. Charge sensitivity variations from one step to the other can actually be ascribed again to a change in the temperature at which the measurement was taken, which in particular may affect the threshold voltage in the MOSFET M_F of Fig. 2.6 and impact on the charge preamplifier gain. The electrical properties of the Apsel3T1 analog readout channel were evaluated by front-illuminating the samples under test with a ^{55}Fe source. Fig. 2.24 shows the event count rate distribution for pixels from an M1 and an M2 matrix respectively, for different irradiation steps, up to an integrated fluence of 6.7×10^{12} n/cm^2. Data shown in Fig. 2.24 confirm that the charge sensitivity and noise are substantially independent of the irradiation fluence, as obtained with the injection capacitance method. Actually, as the integrated fluence increases, both the 5.9 keV spectral peak position (which is related to the front-end gain) and the standard deviation σ_n of the distribution around the peak (corresponding to the noise root mean square at the shaper output) show a limited variation, in the order of a few percent.

However, ^{55}Fe spectra shown in Fig. 2.24 deserve further comment. The measured event rate, in particular around the spectral peak, for the monolithic sensor of the M1 kind is smaller than in the case of MAPS from the M2 matrix, a behavior which has been observed in all of the tested samples. This

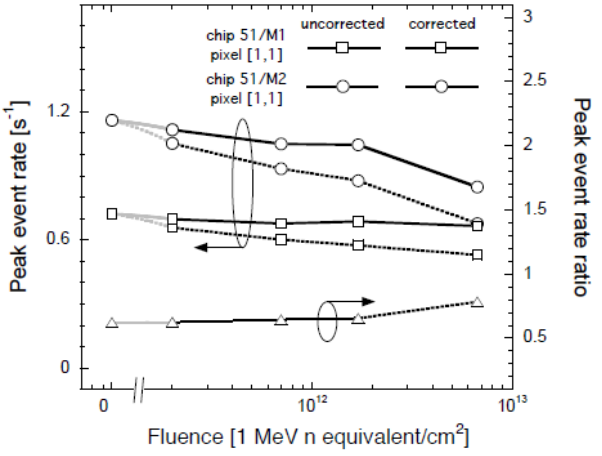

Figure 2.25. Peak event rate as a function of the fluence for the same pixels as in Fig. 2.24 and ratio of the peak event rate in the M1 pixel to the peak event rate in the M2 pixel.

is quite evident in Fig. 2.25 which shows the peak event rate, calculated as the rate of events yielding a shaper output amplitude ranging from $V_{peak}-\sigma_n$ to $V_{peak}+\sigma_n$, as a function of the integrated fluence. The above definition of the peak event rate makes it possible to account for the small noise-induced broadening of the 5.9 keV peak. The figure shows both the peak event rate as obtained from the spectra of Fig. 2.24 (dotted lines curves, denoted as uncorrected) and the data obtained by taking into account the decrease of the source activity with time, according to the 2.73 year ^{55}Fe half-life (continuous line curves , denoted as corrected). Fig. 2.25 also shows the ratio between the peak event rate in the M1 pixel and that in the M2 pixel, which was found to remain almost constant, at least up to a fluence of 1.7×10^{12} n/cm^2, with an average value of 0.64. A slight increase can be detected after the last irradiation step, pointing to a higher sensitivity to bulk damage (from the standpoint of ^{55}Fe count rate) of the M2 structure with respect to M1. The measured peak event ratio was actually found to be very close to the ratio of the device volume M1 to M2 where, for each of the two sensors, the device volume is defined as the region including the DNW diffusion, together with the internal P-well, the N-well diffusion (when used as an extension of the collecting electrode) and the depleted zones of the external junctions.

An estimation of the device volume can be obtained from the data in Table 2.2, showing the extension of the N-well and of the DNW used in the design of the collecting electrodes for the two kinds of MAPS. Based on the doping characteristics of the technology, the junction depth is about 2 μm for the DNW and about 1 μm for the N-well. A depletion width of around 1 μm at the DNW/P-substrate junction of the MAPS in the nominal operating conditions was determined through device simulations. The resulting ratio, also shown in Table 2.2, is about 0.66, very close, as already mentioned, to the peak event rate ratio of Fig. 2.25. This seems to indicate that, when a ^{55}Fe photon

Device	**N-well area** [μm^2]	**DNW area** [μm^2]	**Sensitive volume** [μm^3]
M1	454	222	1570
M2	220	650	2390

Table 2.2. Geometrical features of the collecting electrodes in $M1$ and $M2$ pixels

converts in the N-well, DNW or P-well, the entire amount of released minority carriers manage to diffuse to the depleted regions. Such a result is reasonable, given the small thickness of these three device regions as compared to the carrier diffusion length. Moreover, diffusion of minority carriers to the depleted regions may be aided by the presence of a gradient in the doping concentration. Finally, note that in Fig. 2.25, after the peak event rate data are corrected for the ^{55}Fe source decay, the decrease of the rate after the four irradiation steps is less than 10% in M1 structures, while it is found to be around 25% for M2-like pixels. This could be ascribed to a change in the depletion region width, due to a variation of the doping concentration. However, this seems to be contraddicted by the fact that, given a defect introduction rate of about 0.02 cm^{-1} [45] [46], a doping change in the order of 2×10^{11} n/cm^3 is expected after a neutron fluence of 10^{13} cm^{-2}, much smaller than the minimum doping level in the devices (10^{15} cm^{-3}). Therefore, the effective substrate doping should not be affected to a significant extent by the neutron fluences used for this work. On the other hand, the larger degradation detected in the M2 structures, where, with respect to pixels from M1 matrices, a larger portion of the collecting electrode is made with a DNW layer, seems to point to a higher bulk damage sensitivity of the DNW structure.

2.3.2 Charge collection performance

As in the case of the γ-ray irradiation campaign, the charge collection performance of the Apsel3T1 have been measured after each neutron irradiation step. Since the primary effect of neutron irradiation is the minority carrier lifetime reduction (see Section 2.1.2) a substantial reduction of the charge collected by the sensor is expected. The same two different measurement techniques as for ionizing radiation tests have been adopted: the first one is the laser stimulation already described in Section 2.2.5, the second one is based on a ^{90}Sr radioactive source.

Tests with laser stimulation

While CMOS circuits are known to feature a considerable degree of tolerance to bulk damage, neutron irradiation is expected to have a major effect on the charge collection performance of monolithic sensors. In order to investigate the charge collection properties of the Apsel3T1 DUTs as a function of the neutron fluence, the measurement setup shown in Fig. 2.20 has been used. Fig. 2.26 shows the charge collected by the central pixel of the M1 matrix before irradiation and after each irradiation step (up to a fluence of 6.5×10^{12} neutron/cm^2

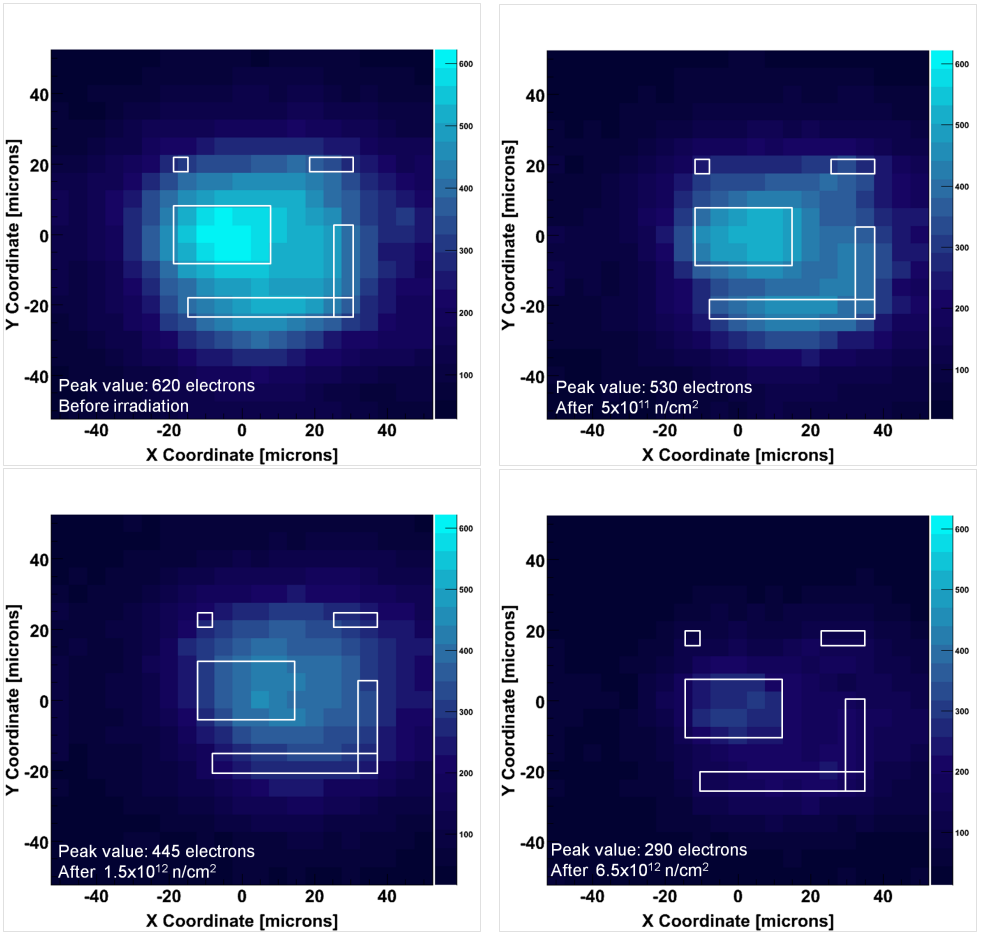

Figure 2.26. Charge collected by the central pixel of an M1 matrix before irradiation and after each irradiation step (up to a fluence of 6.5×10^{12} n/cm^2 neutron) as a function of the position of the laser spot.

fluence as a function of the position of the laser spot. The collected charge was calculated as the ratio between the amplitude of the readout channel response to the laser pulse and the charge sensitivity measured through charge injection from an external pulser (discussed in the previous section). After irradiation, a sizeable change (more than 50%) has been observed, with the peak detected charge decreasing from about 620 to 290 electrons. The primary neutron damage effect in this case is represented by recombination of signal charge carriers at radiation-induced defect sites, resulting in a reduction of the

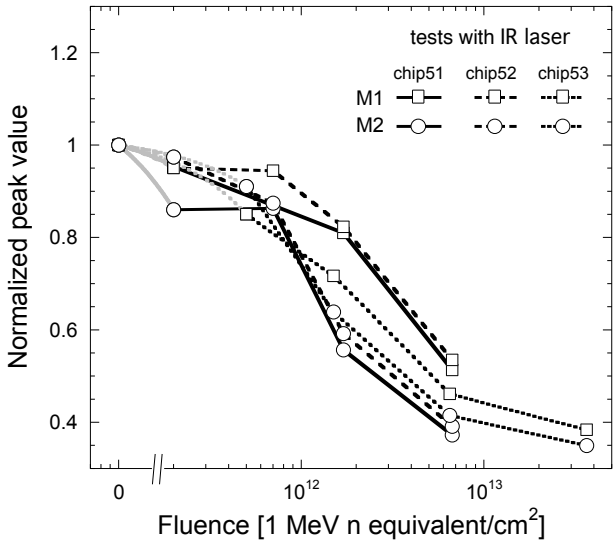

Figure 2.27. Normalized peak value of the collected charge as a function of the fluence in the central pixels of M1 and M2 matrices for the three tested samples (chip 51, 52 and 53) illuminated by an infrared laser (note the log scale for the fluence).

collected charge [47]. Further results on IR laser tests are shown in Fig. 2.27, where the peak value of the collected charge for the central pixels of M1 and M2 matrices of different chips is plotted as a function of the fluence. The reduction of the collected charge with increasing neutron fluence is apparent. The sensitivity of the laser intensity to small temperature changes may justify the not so smooth behavior of the curves.

^{90}Sr radioactive source tests

The measurement setup used for the tests with with ^{90}Sr radioactive source is shown in Fig. 2.28. Electrons released by ^{90}Sr through beta decay have a broad continuous spectrum, with endpoint energy larger than 2 MeV. After detection of an electron, the pulse from a scintillator is used to generate a 500 ns gate signal. If, within this time interval, the signal in the central pixel (the cluster seed) of the 3x3 matrix exceeds $5\sigma_n$ (σ_n being the rms noise at the shaper output), a trigger is issued and the waveforms at the output of the 9 channels are stored. For each event, the amplitudes of the signals from the nine pixels taken at a time t_p (400 ns) after the scintillator pulse are summed, and

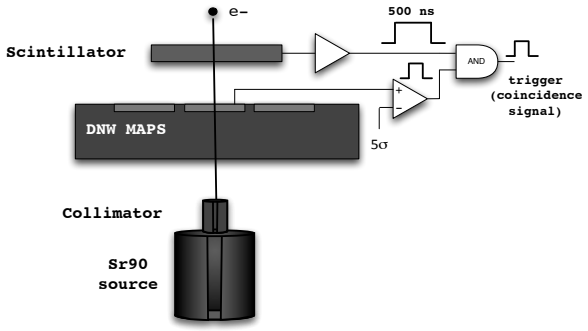

Figure 2.28. Measurement setup for the characterization of the DUTs with a ^{90}Sr source.

the sum is stored as a measurement of the charge collected by the cluster of 9 pixels. Some examples of the ^{90}Sr test results are shown in Fig. 2.29, which represents the spectra for an M1 and an M2 matrix respectively, obtained before irradiation and after exposure to integrated fluences of 1.7×10^{12} n/cm^2 and 6.7×10^{12} n/cm^2. In the two figures, the spectra are normalized to the scintillator detection rate. Based on the spectral characteristic of the ^{90}Sr decay, about 50% of the released electrons are minimum ionizing particles and produce the distinct Landau shaped distribution depicted in these figures. As already mentioned in the previous subsection, the recombination of charge at neutron-induced lattice defects in the neutral substrate region is likely to be the main source of degradation in charge collection efficiency [47]. ^{90}Sr electrons release electron/hole pairs along their track while they pass through the entire substrate. As a consequence, the charge collected by the sensor in a ^{90}Sr experiment always includes some contributions from the undepleted substrate, which depends on the carrier lifetime and, as a consequence, on the neutron fluence in the case of irradiated devices. This results in a general shift to the left of the ^{90}Sr spectrum, obviously including the peak region, consistent with the above mentioned degradation in the charge collection properties of the monolithic sensor.

Figure 2.30 shows the most probable value of the ^{90}Sr spectra normalized to the pre-irradiation value as a function of the fluence for the M1 and M2 matrices of two different chips. A reduction of about 50-60% can be detected after an integrated fluence of 6.7×10^{12} n/cm^2. For comparison sake, the averaged results from IR laser tests are shown in the same figure. While very good agreement can be found in the case of the M1 matrix, M2-like sensors

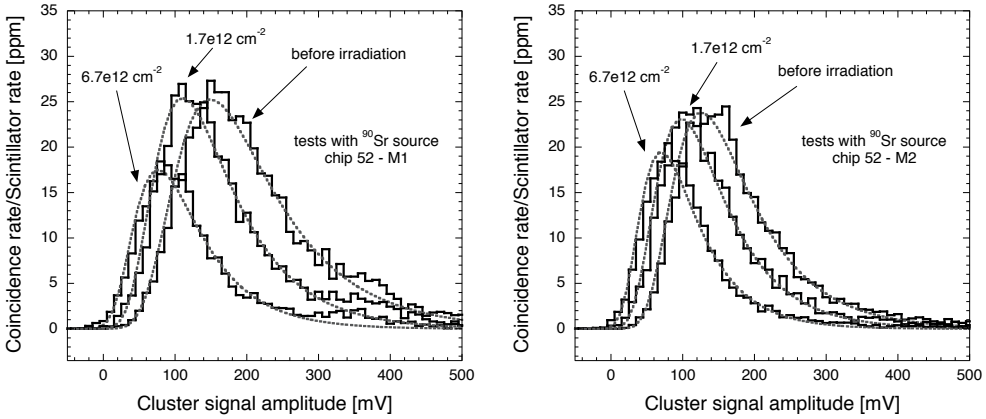

Figure 2.29. Event rate normalized to the scintillator rate in an M1 (left) and M2 (right) matrix tested with a ^{90}Sr source before irradiation and after exposure to a 1.7×10^{12} n/cm^2 and 6.7×10^{12} n/cm^2 neutron fluence.

provided a slightly different (but qualitatively similar) response to the IR laser and ^{90}Sr sources. The signal-to-noise ratio before irradiation is 25 and decreases to less than 15 after the last step, following the charge collection performance emphasized by ^{90}Sr spectrum measurements. As discussed in the last section of this chapter, these results do not comply with the radiation hardness required for chips to be operated in the Layer0 of the SuperB SVT.

2.3.3 Effects on leakage current

As in the case of γ-ray irradiation campaign, the characterization of a set of DNW octagonally shaped diodes has been carried out in order to evaluate neutron irradiation damage effects on the DNW charge collecting electrode. Fig. 2.31 shows the total leakage current of diodes A and B (the sum of the current flowing through the internal and the external junction, see Table 2.1 and Fig. 2.5) as a function of the reverse voltage for different irradiation steps. During all measurements on DNW diodes, the device temperature was in the range 21-23 °C. The three major contributions to leakage current in an inversely biased pn junction are expressed as follows [48]:

- diffusion component:

$$J_{diff,n} = qD_n \frac{n_i^2}{N_A L_n} \left[exp\left(\frac{qV}{kT}\right) - 1 \right], \tag{2.16}$$

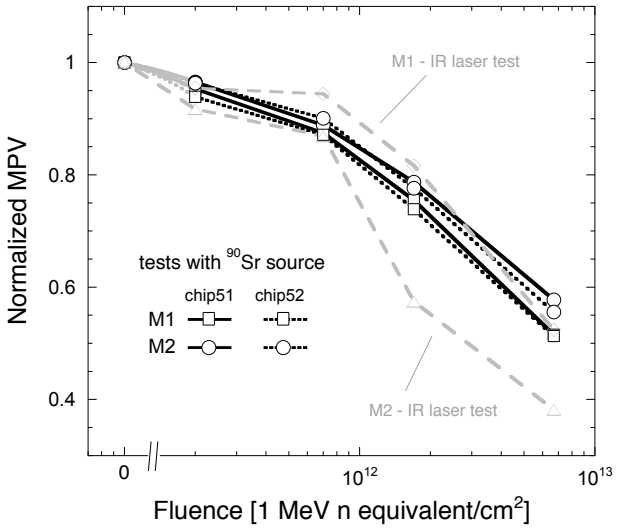

Figure 2.30. Most probable value (MPV) of the ^{90}Sr spectra normalized to the pre-irradiation value as a function of the fluence for M1 and M2 matrices for different chips. Data from IR laser tests are also shown for comparison.

where q is the elementary charge, D_n is the diffusion constant of electrons, n_i is the intrinsic carrier density, N_A is the acceptor density, and L_n is the diffusion length of electrons;

- generation component:

$$J_g = q\frac{n_i}{\tau_g}W\left[exp\left(\frac{qV}{2kT}\right)-1\right], \tag{2.17}$$

where τ_g is the generation lifetime and W is the depletion width;

- surface generation component:

$$J_{sg} = qS_0 n_i \frac{L}{A}W_S\left[exp\left(\frac{qV}{2kT}\right)-1\right], \tag{2.18}$$

where S_0 is the surface generation velocity, W_S is the depletion width at the Si/SiO_2 interface, L is the perimeter length of the junction and A is the junction area.

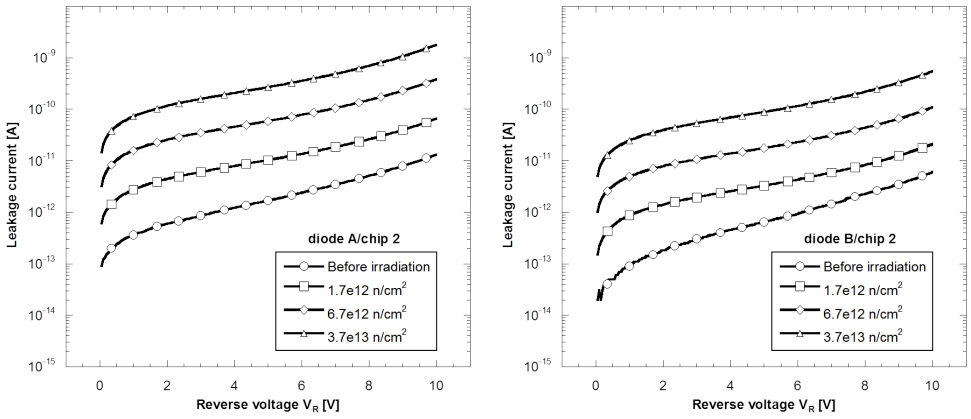

Figure 2.31. Leakage current as a function of the reverse voltage in deep N-well diode A (left) and B (right). The pre-irradiation curve is compared to the ones obtained after exposure to different integrated neutron fluences.

Radiation-induced displacement of atoms from their lattice site is responsible for an increase of the generation current in the depleted region [47] (i.e. τ_g in (2.17) decreases). On the other hand, the leakage current might also be affected by an increase in the diffusion current. This could actually be correlated with the detected decrease in charge collection efficiency (see Section 2.3.2), to be ascribed, as already mentioned, to a reduction in the recombination lifetime τ_n (which is related to the diffusion length by the equation $L_n=\sqrt{D_n\tau_n}$). However, the limited set of available test structures makes it difficult to perform an accurate extraction of the diffusion component [49]. The curves in Fig. 2.31 show an exponential increase (note the logarithmic scale on the current axis) with the reverse voltage V_R for $V_R>2$ V. The increase can be observed to be even more than exponential for $V_R>6$ V. This behavior may indicate the presence of an electric field related current enhancement mechanism, such as Poole-Frenkel emission or phonon assisted tunneling [50] and/or impact ionization [51]. Actually, simply based on the doping concentrations, under the simplifying hypotesis of abrupt junction, for $0<V_R<10$ V a field from 10 kV/cm to a few tens of kV/cm can be estimated for the depleted region of the external junction, about one order of magnitude larger for the internal junction, which actually was found to provide the larger contribution at large V_R values. While both generation and diffusion currents are proportional to the junction area, a third contribution to the leakage current, proportional to

the junction perimeter length, may stem from the surface generation current at the Si/SiO_2 interface. This contribution, as seen in Section 2.2.2 and confirmed by previous tests with γ-ray on DNW MAPS [52], is generally very sensitive to ionizing radiation, due to the involvement of the surface silicon oxides. The relative weight of the area and the edge contributions may be evaluated by expressing the leakage current variation ΔI_{leak} in each of the internal and the exteral junctions as:

$$\Delta I_{leak}(V_R, \Phi) = J_A(V_R, \Phi) \cdot A + J_P(V_R, \Phi) \cdot P \tag{2.19}$$

where J_A and J_P represent the area and the perimeter current density, respectively, A is the junction area and P is the junction perimeter. Both J_A and J_P depend on the reverse voltage and on the fluence Φ. Equation 2.19 can be used to interpolate the measured current value of the three available diodes for given V_R, Φ and junction (internal or external) and extract the current density parameters. Table 2.3 shows the values of J_A and J_P for the maximum fluence value of 3.7×10^{13} n/cm^2 taken, for J_A and J_P in the internal junction, at a reverse voltage of 0.3 V, the same at which the prototype MAPS of the M1 and M2 kinds are operated. The value of J_P in the external junction was taken at V_R=6 V, as the extraction was not sufficiently accurate at smaller V_R values. However, this can be considered a worst case choice as J_P is supposed to be smaller at smaller reverse bias voltages. Based on the coefficients shown in Table 2.3 and on the geometrical features of the M1 and M2 collecting electrodes, the expected dark current increase in the collecting electrode can be estimated to be about 1.5 pA at the maximum fluence, leading to a negligible degradation (about 2 electrons) in the ENC, that corresponds to 5% of the pre-irradiation value. Also, this relatively large current increase is not expected to affect the operating point of the charge preamplifier, which by design can accommodate a current of a few pA through its feedback network.

	J_A [$\mathbf{A}/\mu\mathbf{m}^2$]	J_P [$\mathbf{A}/\mu\mathbf{m}$]
External junction	7.2×10^{-16}	3.2×10^{-15}
Internal junction	1.2×10^{-16}	8.9×10^{-16}

Table 2.3. Area and perimeter current density contributions at a fluence of 3.7×10^{13} n/cm^2.

This is supported by the fact that, in Fig. 2.23, no significant change (which might instead appear in the case of a shift in the operating point) is observed in the charge sensitivity.

2.4 Conclusion on the irradiation campaigns

As already mentioned at the beginning of this chapter, the study of the radiation hardness in CMOS MAPS is of paramount importance, since they are going to be operated in the harshest part of the whole detector, i.e. the silicon vertex tracker. The radiation hardness requirements for the SVT Layer0 already introduced in chapter 1 are summarized in the following [7]:

- equivalent fluence of 3.5×10^{12} n/cm^2 per year;
- total ionizing dose of 3 Mrad(SiO_2) per year.

The detector design has been conceived in order to enable the periodical replacement of Layer0 modules. Even in the case the Layer0 modules are replaced once every two years (it cannot be done more often, since the module replacement implies shutting down the machine for a certain period), Apsel3T1 (or a CMOS MAPS with similar features) might not satisfy the detector requirements in terms of signal-to-noise ratio. In fact, considering the combined effects of ionizing radiation and neutron irradiation, the signal-to-noise ratio is found to drop from 25 before irradiation, to about 6 after irradiation with a TID of 9.7 Mrad(SiO_2) and a fluence of 6.5×10^{12} n/cm^2. The effects of both the irradiation campaigns on the signal-to-noise-ratio over two years of the Apsel3T1 MAPS operation is plotted in Fig. 2.32. The correspondence between time and total ionizing dose (or fluence) is based on the above mentioned SVT Layer0 specifications, assuming that both parameters grow linearly with time. However, it is worth recalling that the irradiation campaign performed with γ-ray features a much higher dose rate than the effective dose rate in the experiment. This means that the radiation hardness test results may underestimate the actual DUT performance, as they do not take possible beneficial annealing effects into consideration.
The results from irradiation campaigns already discussed throughout this chapter, are summarized in the following:

- decrease of about 25% of the charge sensitivity after 9.7 Mrad(SiO_2), likely due to:

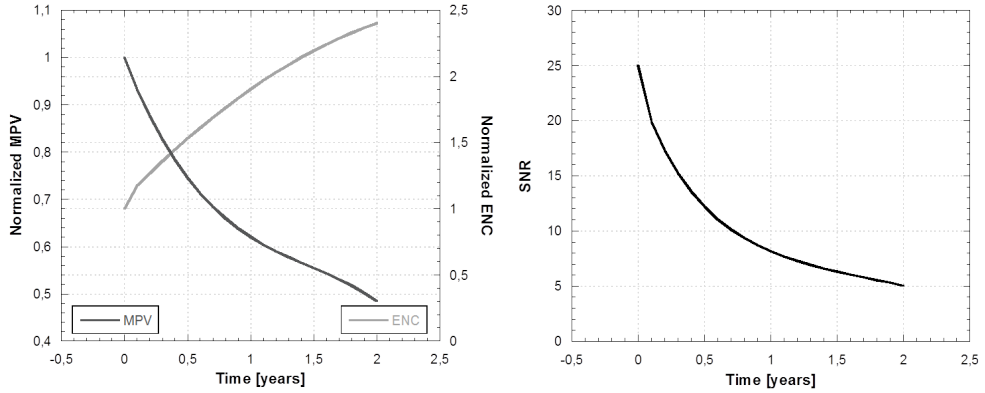

Figure 2.32. Combined effects of γ-ray and neutron irradiation on MPV of the collected charge, ENC and SNR over two years of the Apsel3T1 MAPS operation.

  - threshold voltage shift in the charge preamplifier feedback transistor induced by RINC effect;
  - leakage current increase in the detector, induced by surface generation current increase at the Si/SiO_2 interface;
  - change in the feedback transconductance caused by transconductor tail current increase;
  - change in the shaper GBP induced, again, by the RINC effect on a PMOS current source in the shaper featuring W=0.18 μm;

- noise increase in excess of 100% after 9.7 Mrad(SiO_2) due to noise contribution from STI-related parasitic transistors in the preamplifier input device;
- increase in the detector dark current up to 1.5 pA after 3.7×10^{13} n/cm^2 due to the increase in the generation current and, possibly, in the diffusion current;
- decrease of about 50% of the collected charge after 6.5×10^{12} n/cm^2 caused by the minority carrier lifetime reduction in the undepleted substrate of the DNW sensor.

Since the performance of the Apsel3T1 chip are not satisfying from the radiation hardness point of view, a novel CMOS MAPS approach expected

to guarantee better performance is proposed in Chapter 3. The chip prototype Apsel4well has been developed in a planar 180 nm CMOS process with quadruple well, called INMAPS, featuring, among other attractive characteristics, a high resistivity (1 kΩ·cm) epitaxial layer, which is expected to improve the tolerance to bulk damage [53]. Actually, doping concentration plays a role in determining the equilibrium Fermi level, which in turn influences the effectiveness of neutron-induced defects as recombination centers [44]. If two substrates with different P-type concentration are considered, in the one with smaller doping concentration, the Fermi level is higher and fewer centers (roughly those located in the silicon band-gap between the Fermi level itself and the conduction band) can take active part in the recombination process. In this way, by means of a different technological approach, better charge collection performance, playing an important role in determining the analog channel signal-to-noise ratio, are expected. Moreover, as explained in the next chapter, the presence of the quadruple well enables the design of small N-well collecting electrodes (much smaller than in the case of DNW MAPS). Here the advantage is twofold. First, since a significant portion of the leakage current increase induced by bulk damage is proportional to the collecting electrodes area, smaller sensor shold be less sensitive to neutron irradiation from this standpoind. Second, smaller sensors have smaller capacitance, resulting in a better compromise between power dissipation and noise of the detector. In the chip Apsel4well, also design measures have been adopted in order to reduce the dependence of ENC and charge sensitivity on the total ionizing dose. First of all, an enclosed layout for the charge preamplifier NMOS input transistor has been adopted in order to move the shallow trench isolation oxides away from the channel sides. Also, an attempt has been made to reduce the effects of ionizing radiation on the charge sensitivity by using a larger PMOS transistor, less susceptible to the RINC effect, as current source in the shaper input branch. However, it is worth mentioning that the RINC effect is strongly dependent on the technology through the shape and thickness of the STI oxides.

Chapter 3

The chip Apsel4well

This chapter is aimed at presenting and discussing the design features of the chip Apsel4well fabricated in the CMOS INMAPS technology. First of all, the main features of the 180 nm CMOS INMAPS process will be illustrated, with particular emphasys on the technological measures adopted in order to optimize the sensor charge collection performance (i.e. high resistivity epitaxial layer and quadruple well). Then, the front-end of the sensor will be described, paying particular attention to the analog readout channel. The simulation performance of the analog front-end section will be discussed in detail. The chapter continues by showing results from three-dimensional physical device simulations of the Apsel4well sensor performed by means of the Synopsys TCAD software package. Physical simulations of the Apsel4well sensor have been very useful to predict the effects of the variation of some key parameters (i.e. collecting electrode geometry, epitaxial layer thickness and resistivity) on the charge collection performance, and provided important indications for the optimization of the pixel layout. The layout will be thoroughly described along with all the solutions adopted in order to optimize the charge collection performance and minimize cross-talk phenomena between the analog and the digital sections of the in-pixel electronics and between the digital section and the collecting electrodes. Finally, the whole layout of the Apsel4well test chip, including a number of different structures will be described.

3.1 The INMAPS process

In section 1.3.3, the operating principles of MAPS have been discussed. The first and still most typical approach to MAPS design, which found application and is still largely in use in the imaging field [16], features a small N-well

diffusion acting as the collecting electrode and read out by a three NMOS transistor chain. The signal is then sent to the periphery for further processing. One of the main issues with this approach is that it can hardly be made compatible with the implementation of a sparsified readout architecture, which is the most straightforward solution to comply with bandwidth requirements set by the experiments at the future high luminosity colliders. Actually, the implementation of these architectures requires the design of complex readout electronics, also including PMOS transistors. However, N-wells designed to host PMOS transistors, by acting as parasitics charge collectors, may negatively affect the pixel charge collection efficiency. In order to address this important issue, the chip Apsel4well, implementing a new approach to MAPS design, has been fabricated in the INMAPS process. This new approach is based on a 180 nm CMOS technology with quadruple well in which, by means of an additional processing step, a high energy deep P-well implant is deposited beneath the PMOS competitive N-wells, as shown in Fig. 3.1 [54]. Since this additional step involves the deep implantation of dopant species, it cannot be made too small in size. This could be a limitation for very small pixels (few square microns, such as the ones in digital cameras), but it does not pose any significant problem in the case of large pixels (i.e. 50x50 μm^2) for HEP applications. The deep P-well implant creates a barrier for the charge diffusing in the epitaxial layer, preventing it from being collected by the positively biased N-wells hosting PMOS transistors of the in-pixel readout circuits. In this way, a theoretical charge collection efficiency of 100% can be obtained. The NMOS transistors are designed in heavily doped P-wells located over a lightly P-doped epitaxial layer about 10 μm thick, which has been grown upon a low

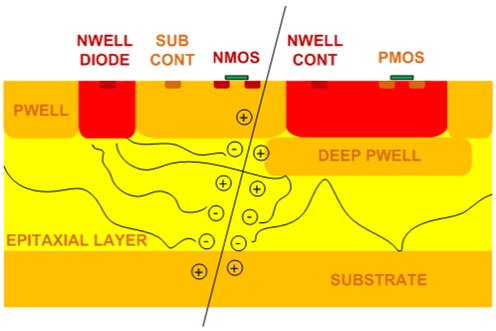

Figure 3.1. Schematic cross section of a MAPS developed with INMAPS process.

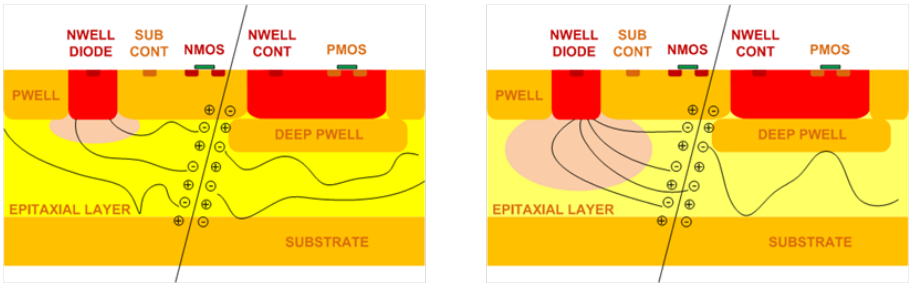

Figure 3.2. Illustration of the depletion region (pink) versus resistivity: standard resistivity silicon (left) and high resistivity silicon (right).

resistivity P-doped substrate (with doping concentration in the order of 10^{18} cm^{-3}). The epitaxial layer, featuring a resistivity higher than both the deep P-well and the substrate, plays an important role in improving the charge collection properties. Actually, the presence of two small potential barriers (deep P-well/epitaxial layer or P-well/epitaxial layer and epitaxial layer/substrate) keeps the carriers within the epitaxial layer, preventing them from diffusing through the substrate. The foundry provides different options for the epitaxial layer resistivity and thickness. For the fabrication of the Apsel4well chip, three different epitaxial layer typologies have been adopted:

- 5 μm thick standard resistivity (about 10 Ω·cm) epitaxial layer;
- 12 μm thick standard resistivity epitaxial layer;
- 12 μm thick high resistivity (about 1 kΩ·cm) epitaxial layer.

The advantage of using a high resistivity epitaxial layer is twofold. First, it is expected to provide better performance (with respect to the case of a lower resistivity layer) in terms of charge collection properties, thanks to the higher depleted volume under the inversely biased collecting electrode (see Fig. 3.2), to the higher carrier lifetime and to the higher potential barriers at the boundaries with P-well, deep P-well and substrate. Second, the adoption of a high resistivity epitaxial layer is expected to improve the device tolerance to bulk damage [44], as required by the Layer0 specifications discussed in chapter 1. The INMAPS process also features 6 metal layers, passive components for analog design (i.e. polysilicon resistors, accumulation capacitors, metal-insulator-metal capacitors, metal fringe capacitors) and multiple gate-oxide thickness (including core devices, with V_{DD}=1.8 V, and thick oxide transistors, available in two flavors, with V_{DD}=3.3 V and V_{DD}=5 V).

3.2 Analog front-end

In the Apsel4well design, the in-pixel analog readout chain, similar to that of Apsel3T1 (see Fig. 2.6) except for the shaping stage feedback network, includes a charge preamplifier, a shaper and a threshold discriminator. The block diagram of the circuit is represented in Fig. 3.3. In some of the test structures an injection capacitance C_{inj} has been implemented outside the pixel in order to enable charge sensitivity measurements by means of charge injection techniques. The charge collected by the inversely biased N-well collecting electrode (featuring a parasitic capacitance C_D) is converted into a voltage signal by the charge preamplifier. The feedback capacitance of the charge preamplifier, C_{FB}, is continuously discharged by a PMOS transistor biased in deep-subtreshold region by the V_{FB} voltage. The aim of the shaping stage, featuring a current mirror block in its feedback network to discharge the capacitance C_2, is to maximize the signal-to-noise ratio. The main drawback of the current mirror feedback configuration lies in the constant slope discharge of the shaper feedback capacitance C_2. This implies a non-constant return-to-baseline time of the signal v_{bl} for different amplitudes of the signal at the charge preamplifier input. Actually, the first order unipolar semi-Gaussian (RC-CR) shaping stage adopted for Apsel3T1, including a transconductor in the shaper feedback network, features a constant shape discharge of the C_2 capacitance. Actually, the RC-CR shaping solution was found to yield a higher noise and baseline volt-

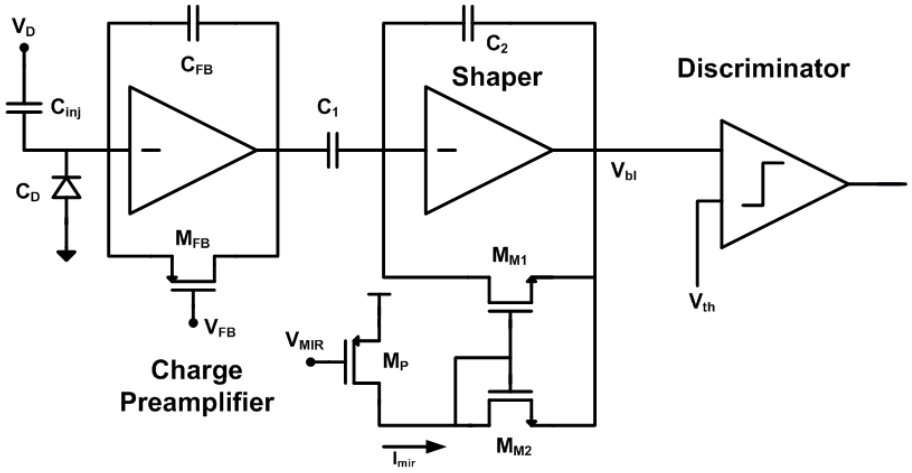

Figure 3.3. Schematic representation of the Apsel4well analog front-end.

Component	W/L [μm]	Capacitance	Value [fF]
M_{FB}	2.5/0.22	C_{inj}	30
M_P	0.32/10	C_{FB}	7.5
M_{N1}	0.4/1.5	C_1	25
M_{N2}	8/0.25	C_2	165

Table 3.1. Features of the components shown in the front-end channel of Fig. 3.3.

age (V_{bl}) dispersion with respect to the current mirror solution. The signal at the shaper output is then compared with a threshold voltage V_{th} by means of a discriminator. Binary information at the discriminator output is then fed to the in-pixel digital section. Table 3.1 summarizes the features (device dimensions and capacitance values) of the components shown in Fig. 3.3.

3.2.1 Charge preamplifier

The schematic diagram of the charge preamplifier forward stage is shown in Fig. 3.4. The input device M1 features an aspect ratio of 10/0.25: the channel width was chosen in such a way to optimize the noise performance of the preamplifier at the design input branch current (5 μA) by applying a capacitive matching criterion. Transistors M2 and M5 implement a folded cascode configuration. Moreover, in order to increase the impedance seen at the node C, two local feedback networks have been designed: the first is implemented by M4 and M5, the second by M6 and M7. The output stage is a source follower composed of the M10 and M11 NMOS transistors. The MOS capacitor M12 has been included with the aim of limiting the bandwitdth of the amplifier and reducing high frequency noise contributions. Voltages V_A and V_B are given to the circuit by means of simple in-pixel NMOS voltage references, while the drain current of M2 is set by means of a current mirror, whose main branch, including transistors M13 and M14, is located outside the pixel. By means of this design choice, the current in M13 can be made large enough to make its contribution to the pixel noise negligible without significantly increasing the per pixel current consumption. Moreover, V_{rif_in} can be externally regulated for a more accurate test of the channel performance. The dimensions of the transistors included in the charge preamplifier are summarized in Table 3.2.

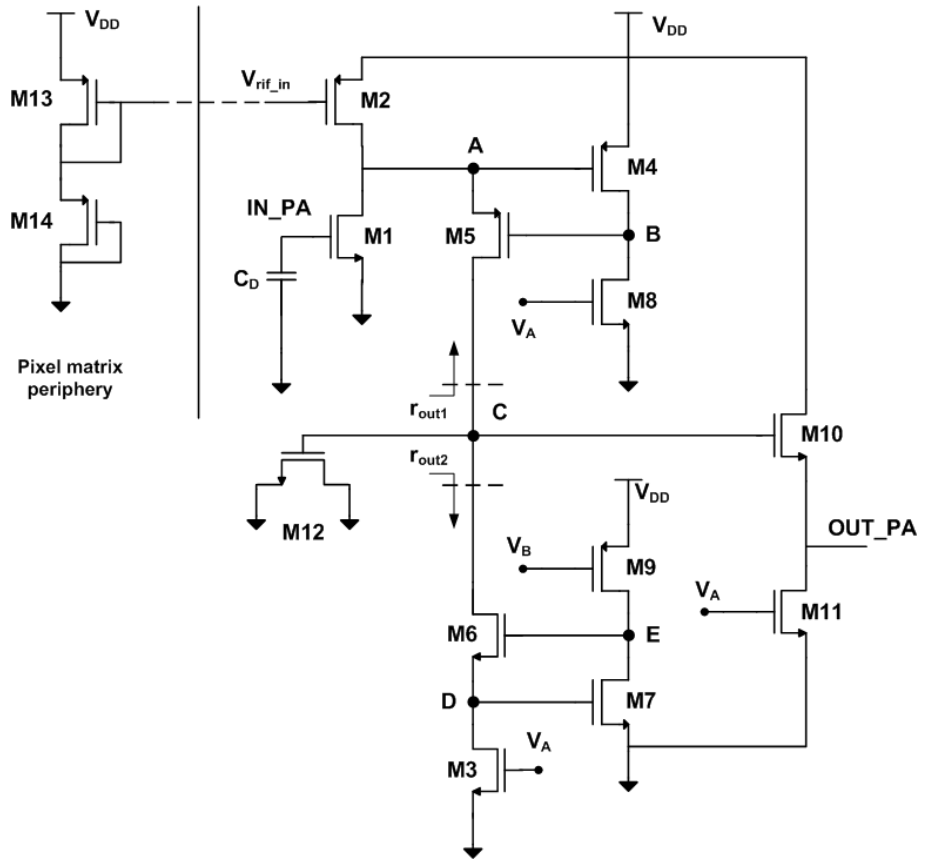

Figure 3.4. Schematic circuit diagram of the Apsel4well charge preamplifier forward stage.

Open loop gain

Small signal analysis is carried out with reference to the equivalent circuit shown in Fig. 3.5. In the following analysis, r_{di} is the drain to source resistance of the M_i transistor, g_{mi} is the channel transcounductance of the same device, C_A and C_C are the small signal capacitance at the node A and node C respectively. Since M10 is in source follower configuration, the gain of this stage will be considered close to 1. His contribution to the transfer function will be considered later. Under these assumptions, the transfer function of the folded cascode stage can be obtained by considering the simplified small signal circuit shown in Fig. 3.6. The output voltage v_c can be written as:

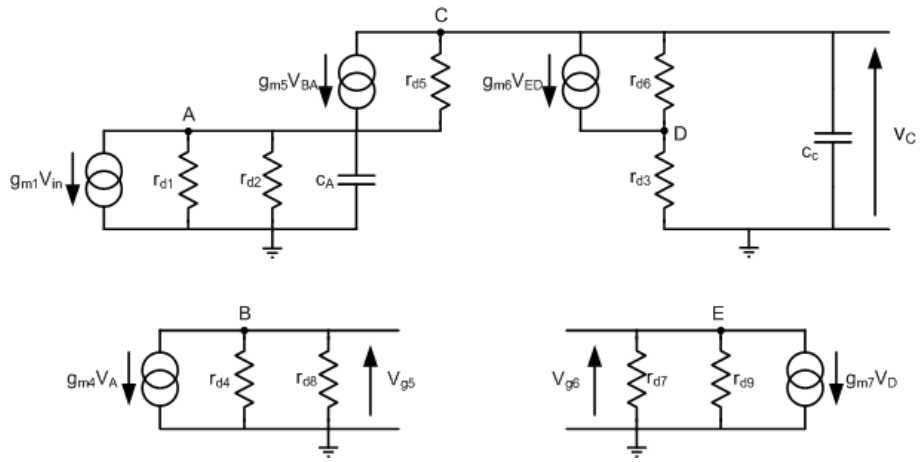

Figure 3.5. Small signal schematic circuital diagram of the Apsel4well charge preamplifier.

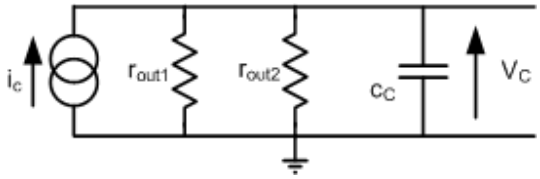

Figure 3.6. Small signal schematic circuital diagram of the Apsel4well charge preamplifier.

$$v_c = Z_c \cdot i_c, \tag{3.1}$$

where:

$$Z_c = r_{out1}||r_{out2}||\frac{1}{sC_C}, \tag{3.2}$$

and

$$i_c = -g_{m1}v_{in}. \tag{3.3}$$

Transistor	**W/L [μm]**
M1	10/0.25
M2	2/1
M3	1.3/0.3
M4	0.4/0.2
M5	0.5/0.2
M6	0.5/2
M7	0.5/4
M8	0.6/0.2
M9	0.22/0.8
M10	5/0.3
M11	2/1.5
M12	5/2.2
M13	4/1
M14	2.6/1

Table 3.2. Aspect ratio of the MOSFETs used in the charge preamplifier design

Equation (3.2) includes also the terms r_{out1} and r_{out2} which can be calculated as

$$r_{out1} = r_d + r_{d5} + g_{m5} r_{d5} r_d (1 + A_1), \tag{3.4}$$

$$r_{out2} = r_{d3} + r_{d6} + g_{m6} r_{d6} r_d (1 + A_2), \tag{3.5}$$

from Fig. 3.5. As a result of the negative local feedback, the voltage at the gate of M5 is equal to its source voltage multiplied by A_1. Similar considerations hold for the M6 gate voltage, where the amplification factor is A_2. Factors A_1 and A_2, along with the resistance r_d, can be calculated as follows:

$$r_d = r_{d1} || r_{d2}, \tag{3.6}$$

$$A_1 = g_{m4}(r_{d4} || r_{d8}), \tag{3.7}$$

$$A_2 = g_{m7}(r_{d7}||r_{d9}). \tag{3.8}$$

Table 3.3 summarizes the values of the small signal resistances, capacitances and transconductances introduced in this section. The transfer function can be approximated by means of a single pole equation. Actually, since $r_{out1} \ll r_{out2}$ (r_{out1}=360 MΩ, r_{out2}=285 GΩ, thanks to the high gain A_2 of the second local feedback):

$$\frac{v_c}{v_{in}}(s) = -g_{m1}\frac{r_{out1}}{1 + sr_{out1}(C_{d5} + C_{d6} + C_{M10} + C_{M12})}. \tag{3.9}$$

where C_{d5} and C_{d6} represent the parasitic drain capacitances at the M5 and M6 drain, respectively, and C_{M10} and C_{M12} are the gate capacitance of M10 and M12. The sum of all these capacitances equals C_C.
In the following, the expression of the dc gain G_0 and the cut-off frequency of the pole f_{P1} of the folded cascode stage are summarized:

$$G_0 = -g_{m1}r_{out1}, \tag{3.10}$$

$$f_{P1} = \frac{1}{2\pi r_{out1}(C_{d5} + C_{d6} + C_{M10} + C_{M12})}. \tag{3.11}$$

$\mathbf{r}_{di}$	**Value [Ω]**
r_{d1}	0.46
r_{d2}	2.44
r_{d3}	6.2
r_{d4}	6.3
r_{d5}	2.9
r_{d6}	20
r_{d7}	323
r_{d8}	5.2
r_{d9}	152

$\mathbf{C}_{xi}$	**Value [fF]**
C_{d5}	0.13
C_{d6}	0.12
C_{d11}	8
C_{s10}	3.1
C_{M10}	7.7
C_{M12}	100.4

$\mathbf{g}_{mi}$	**Value [μS]**
g_{m1}	112
g_{m4}	8.9
g_{m5}	12.5
g_{m6}	9.2
g_{m7}	2.3
g_{m10}	23.7
g_{m11}	16.9

Table 3.3. Small signal resistances, capacitances and transconductances in the charge preamplifier.

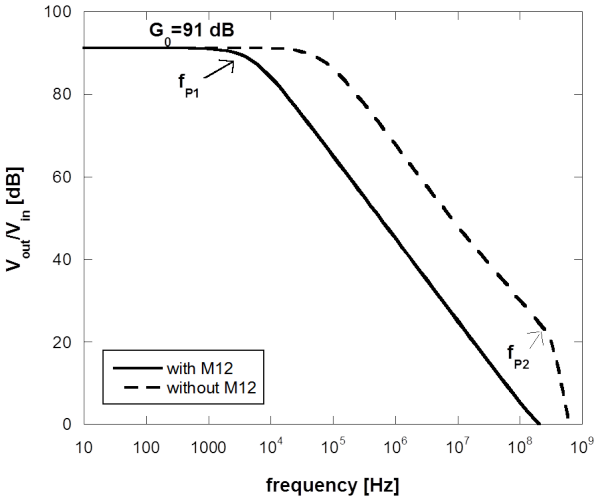

Figure 3.7. open loop ac response of the charge preamplifier.

The source follower output stage, is responsible for a high frequency pole f_{P2} (under the assumption that $r_{d11} \gg 1/g_{m10}$) at frequency

$$f_{P2} = \frac{1}{2\pi \frac{C_{s10}+C_{d11}}{g_{m10}}}, \tag{3.12}$$

where C_{s10} and C_{d11} are the overall source and drain capacitances of transistors M10 and M11, respectively. Figure 3.7 shows the simulation results of the open loop ac response of the charge preamplifier. The dotted curve has been obtained by simulating the circuit without M12, while the continuous line shows the circuit behavior with M12. In the latter case, both from calulations and simulations, the dc gain of the circuit is equal to 91 dB, while f_{P1}=4 kHz and f_{P2}=350 MHz.

Time domain analysis

The analysis of the preamplifier time response to a current pulse will be performed with reference to Fig. 3.8. In this figure, the transistor M_{FB} operating in deep subthreshold region represented in Fig. 3.3 has been replaced by means of the equivalent resistance R_F in order to simplify the analysis. As already shown in section 3.2.1, the charge preamplifier transfer function can be approximated with a single pole transfer function,

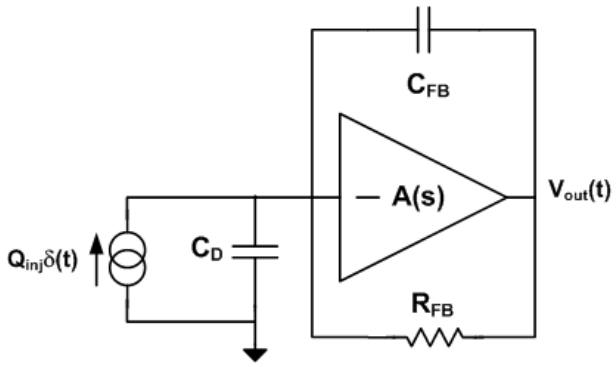

Figure 3.8. Schematic representation of the charge preamplifier for time domain response analysis.

$$A(s) = \frac{A_0}{1 + s\tau}. \tag{3.13}$$

With this assumption, the Laplace transform of the circuit response to a current pulse with area Q_{in} is given by:

$$V_{out}(s) = \frac{Q_{in} A_0 R_F}{s^2 \tau C_D R_F + s[\tau + R_F(C_D + A_0 C_F)] + A_0 + 1}. \tag{3.14}$$

Since $A_0 \gg 1$, $A_0 C_F \gg C_D$ and $R_F A_0 C_F \gg \tau$, (3.14) can be re-written as:

$$V_{out}(s) \simeq \frac{Q_{in} GBP}{C_D} \cdot \frac{1}{s^2 + s \cdot GBP \cdot \frac{C_F}{C_D} + GBP \frac{1}{C_D R_F}}, \tag{3.15}$$

where GBP=A_0/τ is the gain-bandwidth product of the preamplifier. By introducing the following time constants:

$$\tau_r = \frac{C_D C_F R_F}{GBP \cdot C_F^2 R_F - C_D}, \tag{3.16}$$

$$\tau_f = R_F C_F, \tag{3.17}$$

equation (3.15) can be further re-written as:

$$V_{out}(s) = \frac{Q_{in}GBP}{C_D} \cdot \frac{1}{\left(s + \frac{1}{\tau_r}\right) \cdot \left(s + \frac{1}{\tau_f}\right)}. \tag{3.18}$$

If $v_{out}(t)$ is the Laplace antitransform of $V_{out}(s)$, then

$$v_{out}(t) = \frac{Q_{in}}{C_F} \cdot H(t) \cdot \left[exp\left(-\frac{t}{\tau_f}\right) - exp\left(-\frac{t}{\tau_r}\right) \right], \tag{3.19}$$

where *H(t)* is the Heaviside function. Equation (3.19) corresponds to the superposition of two exponential signals with time constants τ_r and τ_f. This equation, for very high value of *GBP* and R_{FB}, can be approximated as:

$$V_{out}(s) = \frac{Q_{in}}{C_F} \cdot H(t), \tag{3.20}$$

which corresponds to a step signal with amplitude Q_{in}/C_F. In conclusion, under the assumed condition about the value of R_{FB} and *GBP*, the charge sensitivity of the preamplifier does not depend on the collecting electrode capacitance C_D. As shown by the open loop analysis, the gain-bandwidth product of

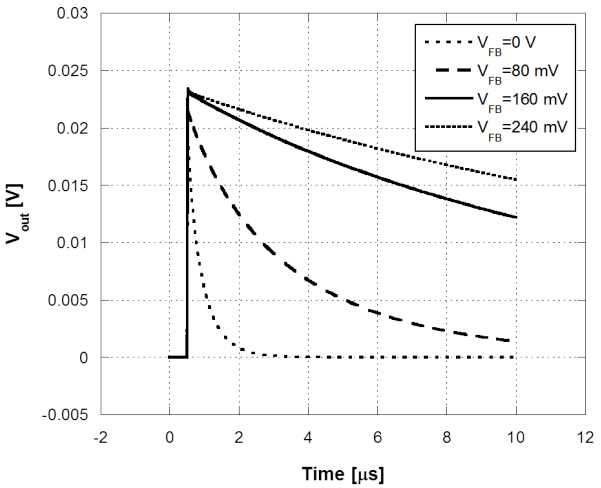

Figure 3.9. Charge preamplifier output waveforms at different V_{FB} values.

the circuit is quite high (about 140 MHz). The small signal output resistance of the transistor M_{FB} shown in Fig. 3.3 (modeled by means of R_{FB}) has been chosen as a compromise between the minimization of the charge preamplifier parallel noise contribution (requiring a large resistance value) and a feedback capacitance discharge fast enough to comply with high event rates while avoiding output saturation. Figure 3.9 shows the charge preamplifier output waveforms at different values of V_{FB} (i.e. at different small signal resistance values) for an input charge signal of 750 electrons. The V_{FB} value chosen for the design of the preamplifier is V_{FB}=160 mV, corresponding to a small signal resistance R_{FB} of about 9 GΩ. As stated in the first chapter, the expected event rate is equal to 100 MHz/cm^2, including a safety factor 5. This means that the per pixel event rate is equal to 2.5 kHz, which is compatible with the discharge time constant of the charge preamplifier, τ=$R_{FB}C_{FB}$=67 μs.

3.2.2 Shaper

The schematic circuit diagram of the shaper forward stage is shown in Fig. 3.10. It features a cascode input stage (MN1 and MN2) with active load (MP1) and a source follower output stage (MN3 and MN4). The input branch current is set by means of MP1, whose gate is biased by means of V_{rif_in} (the same biasing M2 in Fig. 3.4), while V_{casc} and V_{sf} are given by simple in-pixel NMOS voltage references. From the standpoint of time domain analysis, the MN5 gate capacitance value sets the output waveform peaking time. In the input branch, MN1 is a thick oxide transistor. The use of this kind of device, featuring a high threshold voltage value, increases the dc voltage level at the shaper output and, as a consequence, improves output dynamic range performance of the stage. In fact, due to the adopted feedback network, the dc volotage at the shaper output is set by the gate voltage of MN1. Table 3.4 shows the dimensions and the type of the transistors in the shaper, along with their threshold voltage. By assuming that the small signal gain of the source follower stage is 1, the open loop gain of the shaping stage can be written as:

$$G_{0s} = -g_{mN1} \cdot r_{out,s}, \tag{3.21}$$

where g_{mN1} is the transconductance of MN1 and $r_{out,s}$ is the output resistance of the cascode stage, given by:

$$r_{out,s} = r_{dP1} || \left(r_{dN1} + r_{dN2} + g_{mN2} r_{dN1} r_{dN2} \right), \tag{3.22}$$

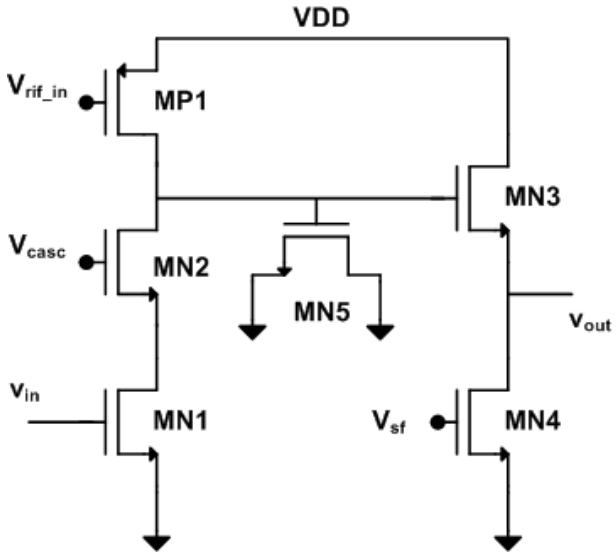

Figure 3.10. Schematic circuti diagram of the shaper forward stage.

with r_{dNi} standing for the small signal resistance of the transistor Ni. The open loop shaper transfer function can be approximated with a single pole expression. The pole frequency can be calculated as

$$f_P = \frac{1}{2\pi r_{out,s} C_{MN5}}, \tag{3.23}$$

where C_{MN5} is the gate to bulk capacitance of MN5 (about 50 fF). Other capacitance contributions (like the one from the gate of MN3) can be safely

Transistor	**Kind of device**	**W/L [μm]**
MN1	thick oxide	2.5/5
MN2	core	1/0.5
MN3	core	2/0.2
MN4	core	1/0.5
MN5	core	2.8/2
MP1	core	0.9/5

Table 3.4. Dimensions and kind of the transistors used in the shaper design.

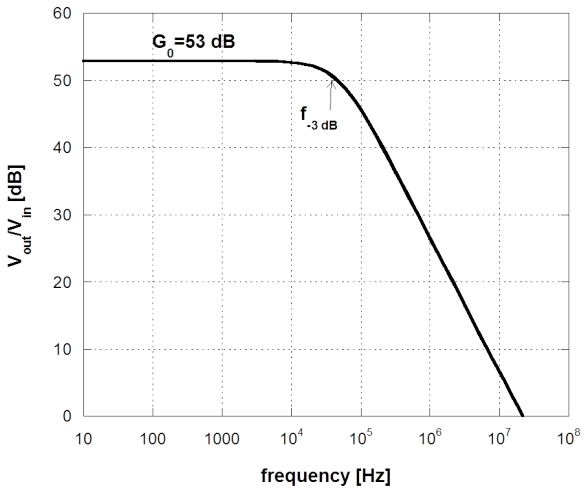

Figure 3.11. Open loop ac response of the shaper.

neglected, since they are much smaller than C_{MN5}.
Figure 3.11 shows the simulated open loop transfer function of the shaper. Both from calculations and simulations, the dc gain is equal to 53 dB, while the the cut-off frequency is about 50 kHz.

Time domain analysis

The time domain analysis of the shaper is performed with reference to Fig. 3.12. In this figure, the resistor R_P and the capacitor C_P are used to model the impedance in the shaper high gain node (i.e., the gate of MN3 in Fig. 3.10). The unity gain block represents the source follower stage. The input signal is modeled by means of a voltage source generating a step signal with an amplitude signal V_{PA} equal to the peak value of the charge preamplifier output signal. This approximation can be done if, as in the considered case, the discharge time of the charge amplifier output signal is much higher than the signal processing time of the analog readout (i.e. the peaking time). The NMOS current mirror can be considered off until a signal is sent to the shaper input and the output voltage v_{out} start decreasing. The reset current i_F can be expressed as:

$$i_F = \alpha \left(v_{out} - v_{in} \right) I_F, \tag{3.24}$$

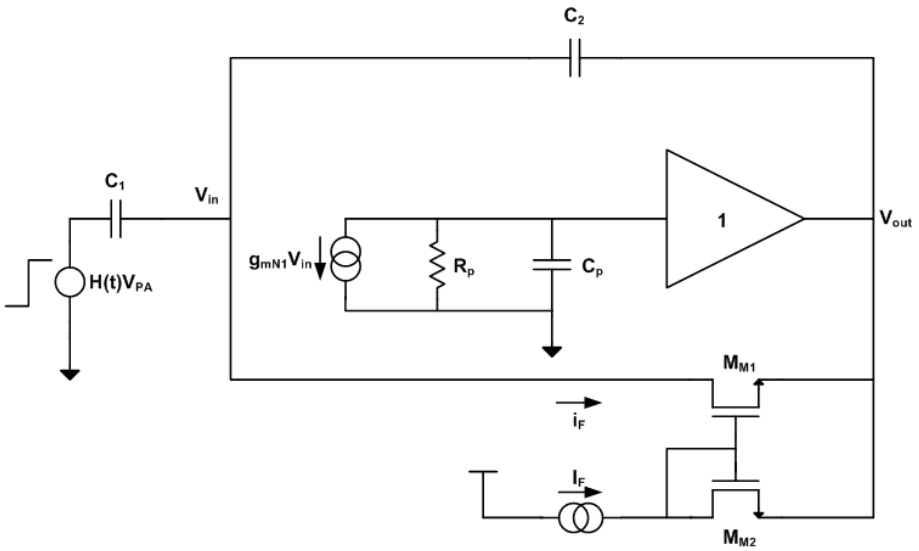

Figure 3.12. Shaper model for the time domain analysis.

where I_F is the reference current source in the current mirror and $\alpha(v_{out}\text{-}v_{in})$ is a monotonically decreasing function of $(v_{out}\text{-}v_{in})$. In particular, α is assumed equal to 0 for $v_{out}\text{-}v_{in} \geq 0$, when the mirror is off. On the other hand, α reaches a constant value α_0 when $v_{out}\text{-}v_{in} \leq -V_{DS,sat}$, where $V_{DS,sat}$ is the V_{DS} voltage at the edge between the triode and the saturation region in M_{M1}. Therefore, the response of the circuit to a step signal is ruled by the following equation:

$$\frac{d^2 v_{out}}{dt^2} + \frac{g_{mN1}\left(1 + A_0 \frac{C_2}{C_1+C_2}\right)}{A_0 C_P}\frac{dv_{out}}{dt} - \frac{g_{mN1}}{C_P\left(C_1+C_2\right)}\alpha\left(v_{out}-v_{in}\right)I_F + \frac{g_{mN1}C_1}{C_P\left(C_1+C_2\right)}V_{PA}\delta\left(t\right) = 0 \quad (3.25)$$

where g_{mN1} is the transconductance of the shaper input device and $A_0 = g_{mN1}R_P$ is the dc gain of the circuit. In the following, M_{M1} is assumed to be in saturation, and $\alpha = \alpha_0$, for $v_{out}\text{-}v_{in} < 0$. The larger is V_{PA}, the more accurate is the previous approximation. Moreover, by assuming:

$$V_{PA} = G_q Q, \quad (3.26)$$

G_q being the sensitivity of the charge preamplifier and Q the charge injected at the preamplifier input, the solution of (3.25) can be written as:

$$v_{out}(t) = H(t)\left[\frac{G_q Q C_1}{C_2}\left(e^{-t/\tau} - 1\right) + \frac{\alpha_0 I_F}{C_2} t\right], \tag{3.27}$$

where

$$\tau = \frac{R_P C_P}{1 + A_0}. \tag{3.28}$$

In order to take into account in (3.27) the nonlinear behavior of α, $v_{out}(t)$ has to be 0 for $t>t_0$, t_0 being such that:

$$\frac{G_q Q C_1}{C_2}\left(e^{-t_0/\tau} - 1\right) + \frac{\alpha_0 I_F}{C_2} t_0 = 0. \tag{3.29}$$

This is needed to take into account the fact that, when $v_{out}(t)$ returns to zero, the current mirror is switched off and the discharge phase ends. From (3.27), the peaking time of $v_{out}(t)$ changes with the injected charge according to the following equation, showing that t_p is a monotonically increasing function of Q:

$$t_p(Q) = \tau \cdot ln\left(\frac{G_q Q C_1}{\tau \alpha_0 I_F}\right). \tag{3.30}$$

In figure 3.13, representing the simulated waveforms at the shaper output for injected charge values up to 2100 electrons, the peaking time increase with the injected charge can be observed. It is worth noticing that (3.27) predicts a constant slope of v_{out} after the peaking time, virtually independent of the injected charge. This behavior can be detected for all the waveforms of Fig. 3.13, except for the one referring to an injected charge value of 160 electrons, where the slope is significantly lower. This discrepancy is a consequence of the approximation made in order to linearize (3.25) (i.e., $\alpha=\alpha_0$ *for* v_{out}-$v_{in}<0$). Actually, for small values of Q, M_2 works in the triode region instead of in saturation (as assumed with that approximation), where $\alpha<\alpha_0$. As a consequence, in this case the discharge of C_2 is slower than for higher Q values.

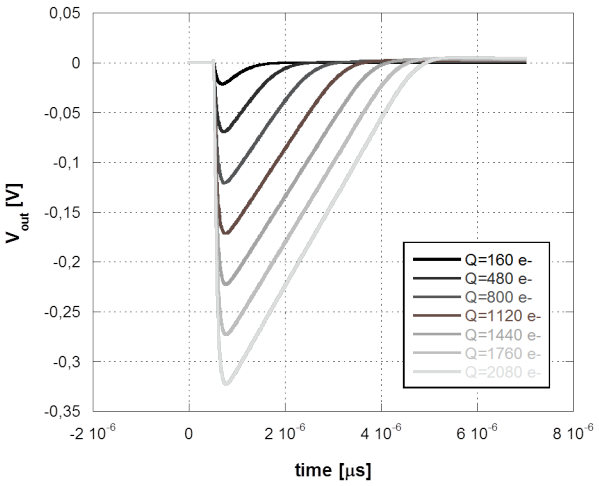

Figure 3.13. Shaper output waveforms simulated at different values of the injected charge.

3.2.3 Discriminator

The schematic circuit diagram of the threshold discriminator is shown in Fig. 3.14. It is based on a differential pair with active load followed by a common source PMOS gain stage. Transistors MC1 and MC2 make up the differential pair, while MC3 and MC4 are used in the active load. The area of these devices is made relatively large with respect to the other analog front-end transistors in order to minimize the threshold dispersion, as discussed in section 3.3.2. Transistors MC6 and MC7 make up the second gain stage. The branch including transistor MC8 generates the reference current, which is properly mirrored in the differential pair and in the common source gain stage. For small differential signals at the input terminals, the overall dc gain is given by

$$\frac{V_{out}}{V_{in} - V_{th}} = g_{m,MC1} \cdot g_{m,MC6} \left(r_{d,MC2} || r_{d,MC4}\right) \cdot \left(r_{d,MC5} || r_{d,MC6}\right). \quad (3.31)$$

The differential pair bias current is equal to 450 nA, while the overall comparator current consumption reaches 650 nA.

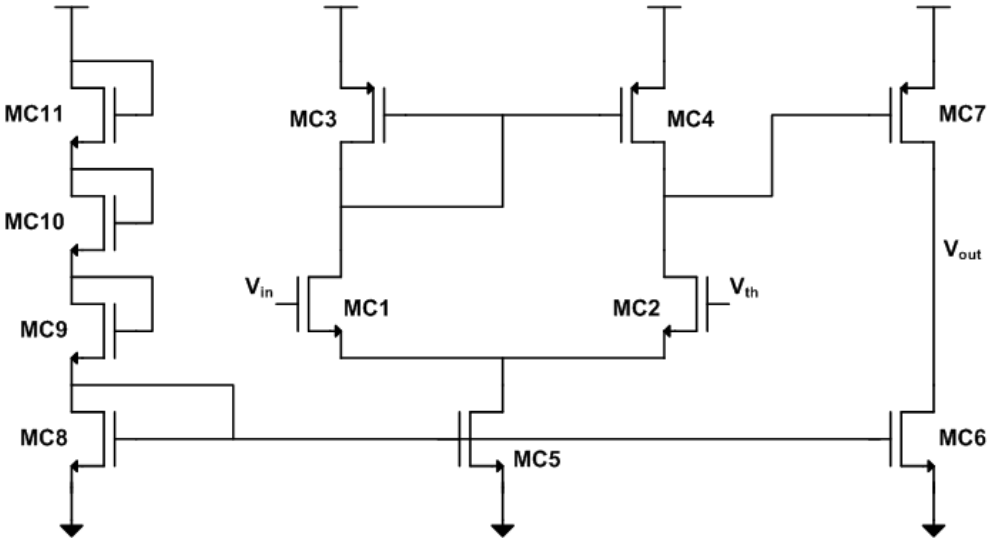

Figure 3.14. Schematic circuit diagramof the threshold discriminator

3.3 Analog front-end simulation

This section is aimed at summarizing the analog front-end simulation results in terms of charge sensitivity, linearity, threshold dispersion, noise and power dissipation. Moreover, simulation results of the charge sensitivity and threshold dispersion at varying temperature will be shown.

3.3.1 Charge sensitivity and INL

Figure 3.15 shows the peak value of the shaper output waveform at different injected charge values Q, up to about 5000 electrons. Since the channel will tipically operate for Q lower than 2000 electrons (in this interval lie about 90% of the events), its charge sensitivity G_q has been defined as the slope of the straight line interpolating all the points of the characteristic for Q between 0 and 2000 electrons, forcing its passage through the axis origin. Anyway, as it can be seen from Fig. 3.15, about the same value can be found considering data points for injected charge up to 5000 electrons. The extracted charge sensitivity is 930 mV/fC. The linearity of the analog readout channel has been again evaluated by means of Fig. 3.15. The integral non-linearity (INL) parameter is defined as

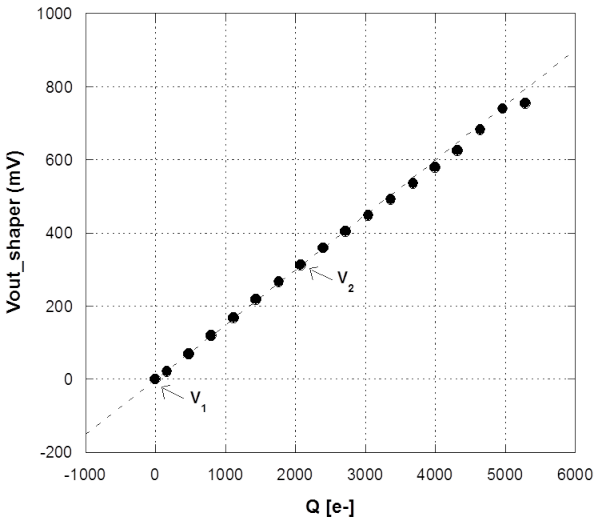

Figure 3.15. Input-output characteristic of the analog channel.

$$INL = \frac{\Delta V_{max}}{V_2 - V_1} \cdot 100\%, \tag{3.32}$$

where ΔV_{max} is the maximum difference between the peak value obtained by simulation and the straight line interpolating all the points of the graph, V_1 is assumed to be 0 and V_2 is the peak value at Q=2100 e^-. The INL calculated by means of (3.32) is 1%. Since the analog channel is not required to perform amplitude measurements, a good linearity is not mandatory for correct system operation. Actually, the analog chain operates in a binary channel readout, which simply detects a hit when the output level of the shaper exceeds a given threshold. However, a good linearity in a charge interval significantly larger than the most probable value of the collected charge, can be of some help for the characterization of the pixel gain and noise by means of the ^{55}Fe techniques already shown in Chapter 2.

3.3.2 Threshold dispersion

High granularity detection systems need to provide high performance parallel processing. In the case of a binary front-end like the one described in this chapter, channel-to-channel differences in the baseline level at the output of

the shaper (i.e., threshold dispersion) may lead to a degradation of the system collection efficiency. In the considered front-end, the main contributions to threshold dispersion are given by mismatch in the shaper and in the discriminator. The schematic diagram of Fig. 3.16 shows the main dispersion sources. Differences in the threshold voltage V_t and in the current gain factor β are the predominant mismatch sources in MOS transistors [42]. However, under usual operating conditions, the contribution from V_{th} variation is more significant than the one from β mismatch [56] and it becomes largely dominant in devices operated in weak inversion. As a consequence, only the effect of V_t will be considered in this study. The threshold voltage variation ΔV_t is assumed to have a normal distribution with zero mean and a variance $\sigma^2(\Delta V_t)$ inversely proportional to the device gate area [58]:

$$\sigma^2(\Delta V_t) = \frac{A_{vt}^2}{W \cdot L}, \tag{3.33}$$

where A_{vt} is a proportionality constant obtained from the characterization of a statistically significant number of device pairs and generally provided by the foundry along with the device models. In Fig. 3.16, mismatch in the transistor

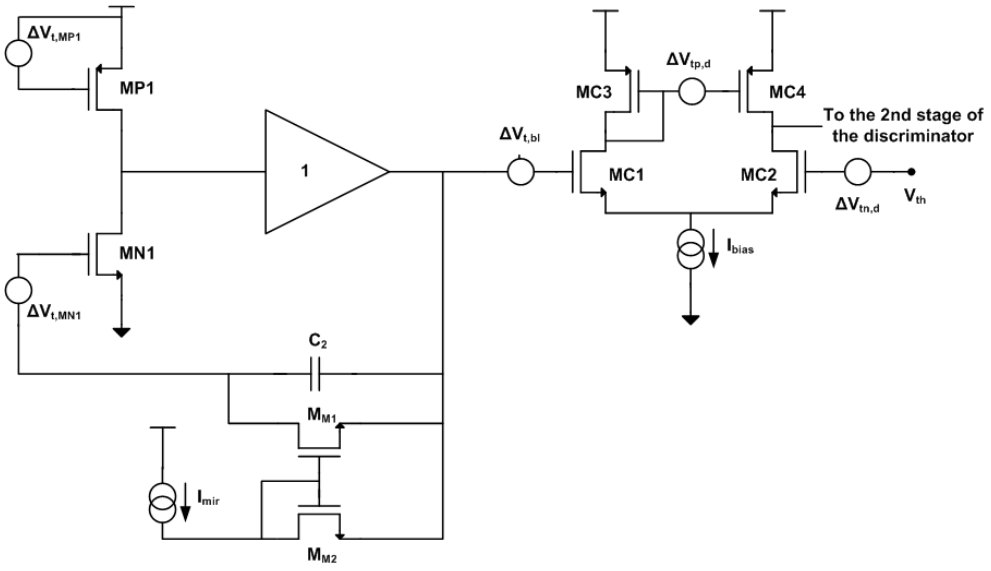

Figure 3.16. Simplified circuit schematic of the shaper and of the discriminator with voltage sources for threshold dispersion modeling.

pairs and V_t fluctuations from channel to channel are modeled by means of the voltage sources $\Delta V_{t,MP1}$, $\Delta V_{t,MN1}$, $\Delta V_{tp,d}$ and $\Delta V_{tn,d}$. Channel-to-channel V_t random fluctuations in transistors MN1 and MP1 in the shaper can change the shaper output level between different pixels. Assuming small mismatch values, small signal analysis of the circuit can be performed to calculate the resulting contribution to the baseline dispersion:

$$\sigma(\Delta V_{t,bl}) = \sqrt{\sigma^2(\Delta V_{t,MN1}) + \frac{g^2_{m,MP1}}{g^2_{m,MN1}} \cdot \sigma^2(\Delta V_{t,MP1})}, \tag{3.34}$$

where $g_{m,MP1}$ and $g_{m,MN1}$ are the transconductances of MP1 and MN1, respectively. The effect of mismatch in MC1-MC2 and MC3-MC4 discriminator transistor pairs on the overall threshold dispersion can be modeled as a variation $\Delta V_{t,d}$ in the externally set threshold voltage V_{th} or, equivalently, as a change -$\Delta V_{t,d}$ of the shaper output baseline. Assuming linear operation for the discriminator, it is possible to calculate the resulting threshold dispersion. In fact, it is true that the discriminator almost always works in the nonlinear region of its characteristic, but transistor mismatch effects become appreciables during 1-to-0 or 0-to-1 transitions, where the circuit operates in linear condition. Therefore, by performing small signal analysis of the circuit, the equivalent shaper output dispersion due to mismatch in the comparator MOSFET pairs can be expressed as:

$$\sigma(\Delta V_{t,d}) = \sqrt{\sigma^2(\Delta V_{tn,d}) + \frac{g^2_{m,MC4}}{g^2_{m,MC2}} \cdot \sigma^2(\Delta V_{tp,d})}, \tag{3.35}$$

where $g_{m,MC4}$ and $g_{m,MC2}$ are the transconductances of MC4 and MC2, respectively. The total equivalent threshold dispersion can be written as:

$$\sigma(\Delta V_{t,eq}) = \sqrt{\sigma^2(\Delta V_{t,bl}) + \sigma^2(\Delta V_{t,d})}. \tag{3.36}$$

Threshold dispersion can be referred to the analog channel input and directly compared to the input charge signal, by taking into account the charge sensitivity,

$$\sigma(\Delta Q_t) = \frac{\sigma(\Delta V_{t,eq})}{G_q}, \tag{3.37}$$

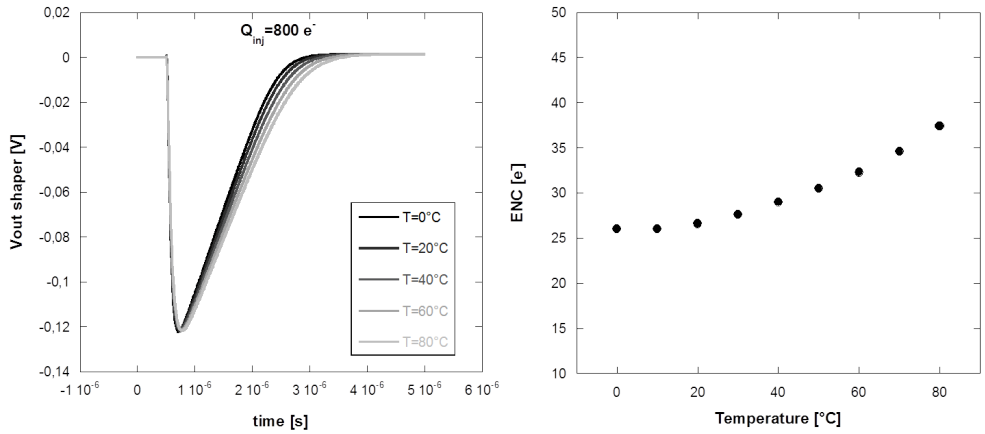

Figure 3.17. Shaper output waveforms (left) and ENC (right) for different temperatures.

where $\sigma(\Delta Q_t)$ is the variation in the equivalent input threshold charge $Q_t = V_{t,eq}/G_q$ due to device mismatch in the shaper and in the comparator. Circuit simulations of the considered front-end give a threshold dispersion equal to 23 electrons.

3.3.3 Temperature variation

Another aspect to be taken into account in the development of high granularity detection systems is the effect of temperature variation within the detector module on the channel charge sensitivity (especially for systems performing amplitude measurements), noise and threshold dispersion. For this reason, the channel behavior has been investigated taking into account temperature variations in transient and Monte Carlo simulations in order to evaluate their effect on the channel performance. Figure 3.17 on the left shows the shaper output waveforms at different temperatures in the range between 0 °C and 80 °C, for an injected charge of 800 electrons. From this simulation, the charge sensitivity temperature coefficient $\Delta G_{q,T}$ can be calculated as the ratio between the G_q variation, ΔG_q, and the temperature range ΔT:

$$\Delta G_{q,T} = \frac{\Delta G_q}{\Delta T} = -420\ \mu V/(^{\circ}C \cdot fC). \tag{3.38}$$

Taking into account a temperature variation of 10 °C for each 10 cm module

and considering a 1 cm chip width, the variation per chip is equal to 1 °C, corresponding to a charge sensitivity variation of -0.05% with respect to the nominal value of 930 mV/fC, which is negligible with respect to variation induced by process fluctuations. The plot on the right of Fig. 3.17 shows the ENC at varying temperatures. Considering the same temperature variation per module as in the case of charge sensitivity and assuming the operating temperature of the module being maintained between 30 and 40 $°C$, the ENC variation can be estimated in 0.15 electrons/$°C$. This value corresponds to an ENC per chip variation of 0.5% with respect to the nominal value of 27 electrons. Similar considerations hold for the threshold dispersion variation with the temperature. Actually, by performing Montecarlo simulations at different temperature steps in the considered range, the baseline temperature coefficient $\Delta V_{bl,T}$ is found to:

$$\Delta V_{bl,T} = \frac{\Delta V_{bl}}{\Delta T} = 560\ \mu V/°C, \tag{3.39}$$

where ΔV_{bl} is the variation of the shaper output dc voltage in the considered temperature range. When referred to the input, the shaper output dc voltage variation becomes

$$\Delta Q_{t,T} = \frac{\Delta V_{bl,T}}{G_q} = 4\ e^-/°C, \tag{3.40}$$

which, again in the case of a 1 $°C$ variation, is negligible with respect to the threshold dispersion performance from the previous section.

3.3.4 ENC and power dissipation

The noise from the analog channel, represented as the root mean square value of the output noise (v_{rms}) as provided by simulations, is 4.06 mV. This result has been obtained with an estimated collecting electrode capacitance C_D of 30 fF. Table 3.5 summarizes the main noise contributions in the circuit. The three most noisy transistors are M_{M1} in the shaper feedback network, (see Fig. 3.3), the input device of the charge preamplifier (M1 in Fig. 3.4) and the input transistor of the shaper (MN1 in Fig. 3.10). It is interesting to notice that the thermal noise contribution of the charge preamplifier feedback transistor M_{FB} is only about 3% of the total amount. In order to compare the noise performance of different channels, it is useful to express the channel noise in terms of ENC. This figure of merit (already defined in section 2.2.3) refers

Transistor	Noise contribution [mV]	Noise contribution [%]
M1	2.08	26
M_{M1}	2.06	26
MN1	1.06	6.75
M3	1.05	6.68
M8	0.8	3.7
M_{FB}	0.7	3.6

Table 3.5. Main noise contributions in the analog front-end of the Apsel4well MAPS (transistor names as in Fig. 3.3, Fig. 3.4 and Fig. 3.10.

the overall channel noise directly to its input. In this way it is possible to obtain an indication of the channel noise that does not depend on its charge sensitivity and can be immediately compared to the input charge signal. The ENC expressed in electrons, can be calculated as:

$$ENC = \frac{v_{rms}}{G_q} \cdot 6250, \tag{3.41}$$

corresponding to about 27 electrons in the case of the analog front-end of the Apsel4well MAPS. Since the most probable value of the collected charge is about 800 electrons (as shown in chapter 4), the channel SNR can be estimated to be about 30. It is worth mentioning that the simulated ENC value has been obtained with a total current dissipation of 10 μA. About one half of this current is dissipated in the charge preamplifier input device in order to minimize its thermal noise contribution. Since the channel voltage supply is 1.8 V, the power dissipation is 18 μW. As already mentioned in section 1.3, the maximum power dissipation per unit area allowed for a pixel module in the SuperB SVT is about 2 W/cm^2. Given a pixel pitch of 50 μm, the power dissipation per unit area of the designed pixel is about 0.7 W/cm^2, less than one half of the limit set by the specifications, therefore leaving enough room in terms of power for the digital in-pixel and peripheral circuits. Table 3.6 summarizes the main performance of the analog front-end.

Series noise contribution from the preamplifier input device

As mentioned before, one of the main noise contributors of the analog front-end is the charge preamplifier input transistor M1. Its mean square noise contribution at the shaper output can be expressed as

Performance	Value
G_q	930 mV/fC
INL	1%
ENC	27 e^-
$\sigma(\Delta Q_t)$	23 e^-
Power dissipation	18 μW

Table 3.6. Analog front-end performance.

$$\overline{v_{N,e_s}^2} = \frac{1}{2\pi}\int_0^\infty S_{M1}|T(j\omega)|^2\,d\omega, \tag{3.42}$$

where S_{M1} is the power spectral density of the noise M1 and T(s) is the transfer function of the circuit shown in Fig. 3.18, given by the product between T_1(s) and T_2(s). Some approximations are made here in order to make the overall transfer function calculation easier without any significant loss of accuracy. First of all, the open loop gain functions of the charge preamplifier and of the shaper A_1(s) and A_2(s) respectively, are expressed as single-pole transfer functions:

$$A_1(s) = \frac{A_{1,0}}{1+s\tau_1}, \quad A_2(s) = \frac{A_{2,0}}{1+s\tau_2}. \tag{3.43}$$

In the previous equations, $A_{1,0}$ and $A_{2,0}$ are the dc gain of the charge preamplifier and of the shaper, respectively, while τ_1 and τ_2 are the relevant time constants. Second, both the M_{N1} transistor in the shaper feedback network and M_{FB} in the preamplifier feedback network have been replaced by means of the relevant small signal equivalent resistance R_{MIR} and R_{FB}, respectively. The computation of T_1(s) has been made under the simplifying hypothesis that $R_{FB}\to\infty$:

$$T_1(s) = \frac{V_{out1}}{e_s} = \frac{\frac{C_T+C_{FB}}{C_{FB}}}{1+s\frac{\tau_1}{A_{1,0}}\frac{C_T+C_{FB}}{C_{FB}}}, \tag{3.44}$$

where C_T includes both the detector parasitic capacitance C_D and the gate capacitance of M1. The transfer function T_2(s) is calculated under the assumption of $A_{2,0}\gg 1$ and can be expressed as

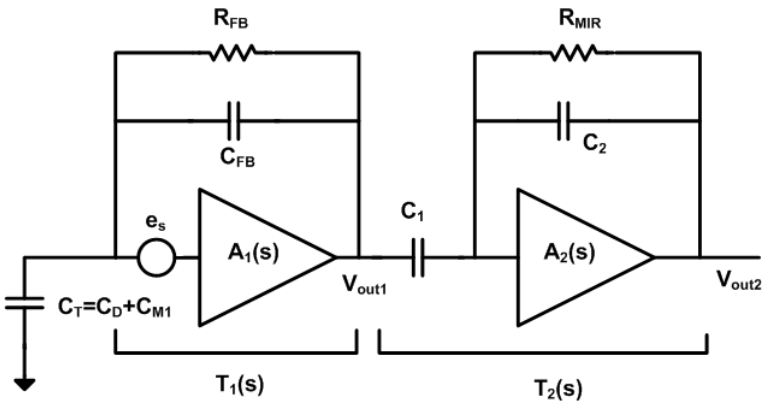

Figure 3.18. Analog front-end network used for the computation of T(s).

$$T_2(s) = \frac{V_{out2}}{V_{out1}} = \frac{sA_2R_{MIR}C_1}{s^2(C_1 - C_2)R_{MIR}\tau - sA_2R_{MIR}C_2 - A_2}. \tag{3.45}$$

The overall transfer function T(s) is computed as the product between T_1(s) and T_2(s):

$$T(s) = T_1(s)T_2(s) = \frac{K}{s^3P + s^2Q - sR - A_2}, \tag{3.46}$$

where K, P, Q and R have the following expression:

$$\begin{aligned}
K &= \frac{C_T + C_{FB}}{C_T} R_{MIR}C_1, \\
P &= \frac{\tau_1\tau_2}{A_1}(C_1 - C_2)R_{MIR}, \\
Q &= (C_1 - C_2)R_{MIR}\tau_2 - \frac{A_2}{A_1}\tau_1\frac{C_T + C_{FB}}{C_{FB}}C_2R_{MIR}, \\
R &= A_2\left(C_2R_{MIR} + \frac{\tau_1\frac{C_T+C_{FB}}{C_T}}{A_1}\right).
\end{aligned}$$

The gate referred noise power spectral density S_{M1} of the transistor M1 computed by the simulator is $3.65{\cdot}10^{-16}$ V^2/Hz and can be included in (3.42) in order to calculate the output noise mean square value:

Parameter	Value
A_1	$3.55 \cdot 10^4$
A_2	447
τ_1	$3.9 \cdot 10^{-5}$ [s]
τ_2	$3.2 \cdot 10^{-6}$ [s]
R_{MIR}	$12.7 \cdot 10^6$ [Ω]
R_{FB}	$9 \cdot 10^9$ [Ω]
C_D	30 [fF]
C_{FB}	7.5 [fF]
C_{M1}	15 [fF]

Table 3.7. Parameter values of the circuit in Fig. 3.18.

$$\begin{aligned}\overline{v_{N,e_s}^2} &= \frac{3.65 \cdot 10^{-16}}{2\pi} \int_0^\infty |T(j\omega)|^2 \, d\omega \\ &= \frac{3.65 \cdot 10^{-16}}{2\pi} \cdot 7.86 \cdot 10^{10} \quad V^2 = 4.56 \cdot 10^{-6} \quad V^2 \end{aligned} \tag{3.47}$$

The root mean square value of the computed output noise is equal to 2.1 mV, which is quite similar to the value obtained by simulations shown in Tab. 3.5 (2.08 mV). This points to the fact that all the assumptions made in order to calculate the transfer function T(s) are reasonable and do not invalidate the calculation results. Table 3.7 summarizes all the parameter values for the circuit in Fig. 3.18.

Noise contribution from the shaper feedback network

The other main noise contribution in the analog front-end is the transistor M_{N1} in the shaper feedback network. In this case, the noise mean square at the shaper output can be expressed as:

$$\overline{v_{N,M_{N1}}^2} = \frac{1}{2\pi} \int_0^\infty S_{M_{N1}} |H(j\omega)|^2 \, d\omega. \tag{3.48}$$

where H(s) is the transfer function between the current generator i modeling the transistor M_{N1} noise source and the shaper output, as represented in Fig. 3.19. $S_{M_{N1}}$ is the power spectral density of the noise in the M_{N1} transistor, which can be again considered independent of the frequency. Once again, the

transistor M_{N1} has been replaced by its small signal resistence R_{MIR}, while the transfer function H(s) has been calculated by considering $A_{2,0}(s) \rightarrow \infty$. Therefore, H(s) can be expressed as

$$H(s) = \frac{V_{out2}}{i} = \frac{R_{MIR}}{1 + sR_{MIR}C_2}. \tag{3.49}$$

The white noise power spectral density S_{MN1} of the transistor M_{N1} computed by the simulator is $3.48 \cdot 10^{-26}$ A^2/Hz. Substituting this value of the noise power spectral density in (3.48), it is possible to get the mean square output noise:

$$\begin{aligned} \overline{v_{N,M_{N1}}^2} &= \frac{3.48 \cdot 10^{-26}}{2\pi} \int_0^\infty |H(j\omega)|^2 \, d\omega. \\ &= \frac{3.48 \cdot 10^{-26}}{2\pi} \cdot 7.98 \cdot 10^{20} \quad V^2 = 4.4 \cdot 10^{-6} \quad V^2 \end{aligned} \tag{3.50}$$

The root mean square value of the computed output noise is about 2.1 mV, quite close to the value obtained by simulations shown in Tab. 3.5 (2.06 mV). This points again to the fact that, also in this case, all the assumptions made in order to compute the transfer function H(s) are plausible.

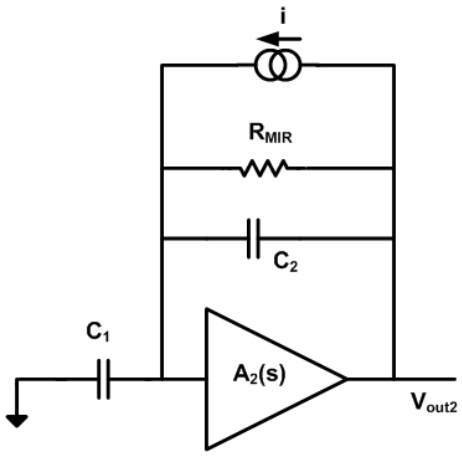

Figure 3.19. Schematic representation of the circuit used for the computation of H(jω).

3.4 Device simulations

Technology CAD (TCAD) software packages represent a powerful tool for the physical simulation of the fabrication process and the operation of semiconductor devices. The software used for this work is the *Synopsys* TCAD package, employing finite element method (FEM) for the solution of the Poisson and transport equation for charge carriers. The Synopsys TCAD tool allows the designers to define and simulate the steps required for the device fabrication and to simulate the electrical response under given operating conditions. This software package is equipped with a couple of tools (*Mdraw* and *Devise*) that can be used for either the visualization of the devices obtained through process simulations and the direct generation of a new geometry with the relevant doping profile. The electrical or thermal characterization of a device can be carried out with the simulator *Dessis*, which, as already mentioned, numerically solves the equations governing the motion and the spatial distribution of charge carriers. Since it employs FEM, it requires that the area (2D simulations) or the volume (3D simulations) of the device be partitioned into a finite number of elements, producing, in this way, an appropriate grid file. For this purpose, two different mesh engine, called *Mesh* and *NOFFSET3D*, are available. The accuracy degree of the simulations can be improved by increasing the number of the elements in which the device is subdivided. This, in turns, leads to an increase in computation time. Typically, a high granularity mesh should be produced where the change in doping is abrupt, in order to avoid convergence problems. The results of the simulation can be displayed through the tools *Tecplot* and *Inspect*. The first allows the representation of quantities

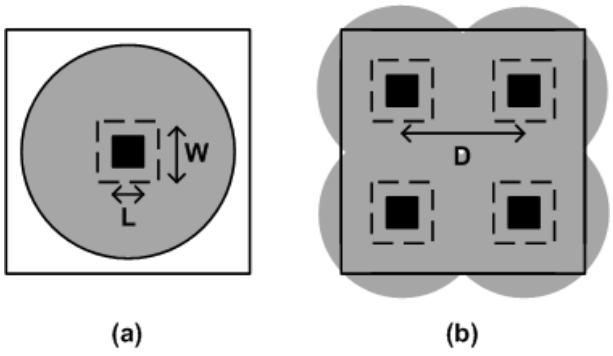

Figure 3.20. Qualitative representation of the pixel coverage with one collecting electrode (a) and four interconnected collecting electrodes (b).

like charge distribution, current density and electrostatic potential in every point of the device. *Inspect* is suitable for the representation of curves like I-V device characteristics and transient simulation results.

For the purpose of this work, the software has been used to build a sensor model enabling the prediction of its charge collection properties. This has been done by defining both 2D and 3D geometries of the device. Simulations of 2D geometries, while providing a less accurate representation of the real device, features shorter simulation times. For the design of the Apsel4well pixel, four squared interconnected N-well diffusions are used as collecting electrodes. This solution has been chosen as the best compromise between noise performance (depending on the collecting electrode capacitance), pixel area coverage and charge sharing between adjacent pixels. Figure 3.20 shows a comparison between the pixel coverage obtained by adopting just one small N-well collecting electrode in the pixel center (a) and the case of the solution adopted in the Apsel4well pixel with four interconnected collecting electrodes (b). The two solutions feature the same value of electrodes parasitic capacitance. A better coverage is obtained in (b) which, on the other hand, may result in charge sharing with adjacent pixels. This latter phenomenon is the more pronounced the higher is the distance D between the collecting electrodes. The deep P-well layer is deposited all over the pixel except for the area around the collecting electrodes, which in the figure is represented by means of a dashed square having side W. In the following, results from 2D simulations aimed at optimizing L, W and D in Figure 3.20 in order to maximize the charge collected by the pixel and minimize charge sharing with adjacent pixels will be shown. Finally, the results from 3D simulation of the charge collection in a 3x3 matrix with different epitaxial layer thickness and resistivity will be presented. Figure 3.21

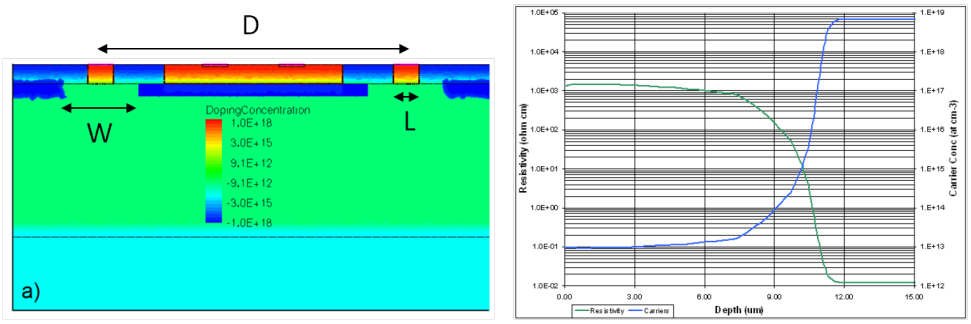

Figure 3.21. Cross section of the simulated structure (on the left) and substrate doping profile and resistivity (on the right).

represents on the left the cross section of the simulated structure (that is the same for both 2D and 3D simulations) and on the right the substrate doping profile and resistivity. The cross section features two lateral N-well collecting electrodes and a central N-well shielded from the epitaxial layer by means of a deep P-well implant about 2 μm deep. In all the performed simulations, a MIP crosses the pixel in its center and releases an amount of energy generating about 80 electron/hole pairs per micron. This event has been simulated by means of the *Heavy Ion* model supported by *Dessis*.

3.4.1 Effects of electrode dimensions on the charge collection properties

The dependence of the charge collection performance on the collecting electrodes area have been evaluated in device simulations by varying the N-well side L from 1.5 to 5 μm. Simulation results with reference to a structure featuring a high resistivity epitaxial layer are summarized in Fig. 3.22. The collected charge and collection time normalized values are shown on the left, while the parasitic capacitance value of four squared N-well collecting electrodes having side L is plotted on the right. This figure shows that, in the considered L range, the collected charge features a quite limited increase (about 2%), while the collection time decreases with L of about 20% over the considered L range. These results seem to lead to the conclusion that the collected charge value mainly depends on the diffusion length of the charge carriers which, in turns, is determined by the epitaxial layer doping concentration and is obviously independent of the electrode side. However, this geometrical parameter affects

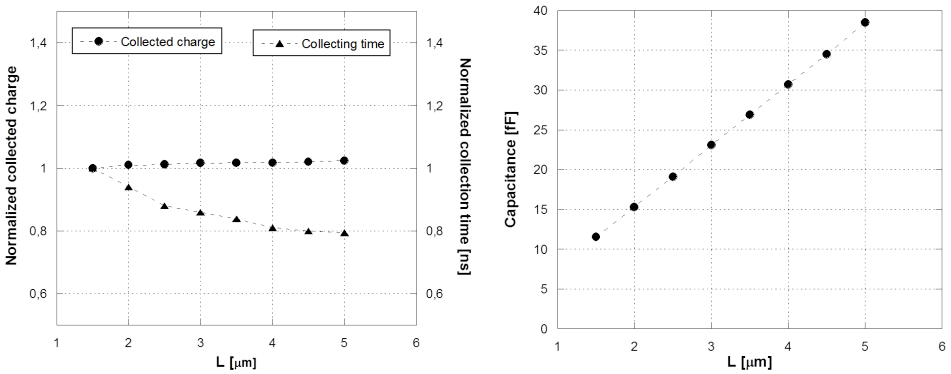

Figure 3.22. Charge collection performance (left) and parasitic capacitance of a four-electrode sensor (right) for L varying between 1.5 and 5 μm.

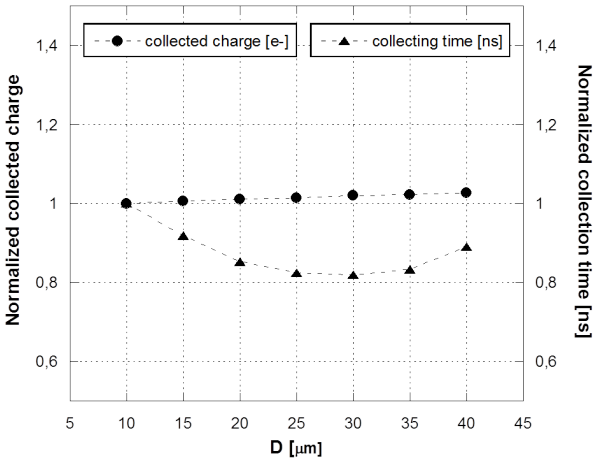

Figure 3.23. Collected charge and collection time normalized values for D between 10 and 40 μm.

the depleted volume under the electrodes. As a consequence, by increasing L (and the depleted volume in proportion) the collection time decreases. Finally, the plot on the right of Fig. 3.22 shows a linear increase of the parasitic capacitance value with L. The increase is not proportional to the collecting electrode area, as one would expect, since in pixels with high resistivity epitaxial layer, the area capacitance contribution is found to be negligible with respect to the perimeter contribution. This is compatible with the fact that the depletion region volume is much larger under the collecting electrode towards the weakly P-doped epitaxial layer than around its perimeter junction (about 1.5 μm deep), where the surface P-doping concentration is higher. The area contribution remains negligible as longer L is not much larger than the N-well junction depth

The effects of the collecting electrodes mutual distance on the charge collection performance have been evaluated by varying the parameter D of Fig. 3.21 between 10 and 40 μm. Simulations have been performed with W=8 μm and L=2 μm. Figure 3.23 shows the normalized values of collected charge and collection time. The plot points to a negligible increase (less than 5%) of the collected charge with D. This increase is probably due to the reduction of the amount of charge collected by adjacent pixels. Results concerning the collection time feature a minimum for D values between 25 and 30 μm.

The last parameter taken into account in these simulations is the side W of the

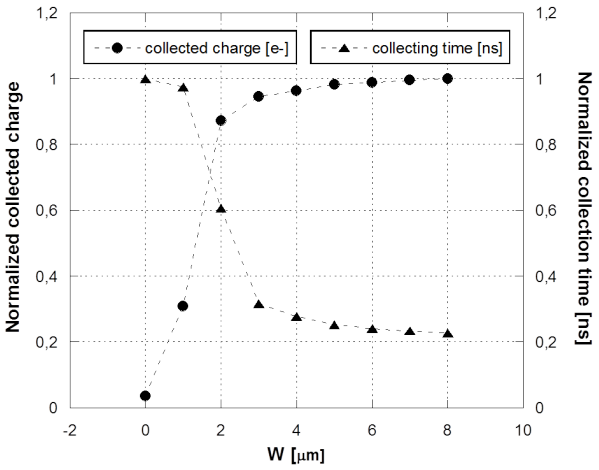

Figure 3.24. Collected charge and collection time values normalized to their maxima for W between 0 and 8 μm.

square area without deep P-well under the collecting electrodes. The charge collection performance of the pixel have been simulated for W values between 0 (i.e., deep P-well implanted all over the pixel) and 8 μm. Simulations have been performed setting D=25 μm and L=2 μm. Simulation results are summarized in Fig. 3.24. The collected charge reaches values close to the maximum for W=L=2 μm, while, in these conditions, the collection time is still far from the minimum. For W>6 μm, the charge collection performance (amount of collected charge and collection time) is very close to the optimum condition, obtained for W=8 μm.

3.4.2 3x3 matrix charge collection properties

Synopsys TCAD simulations of a 3x3 matrix featuring the same geometry of the Apsel4well pixel, shown in section 3.6.1, have been performed. The doping profile of the pixel featuring a 12 μm high resistivity epitaxial layer is the same as the one shown in Fig. 3.21. Structures with 5 and 12 μm thick standard resistivity epitaxial layer have also been simulated. The typical p-doping concentration of the standard resistivity epitaxial layer is 10^{15} cm^{-3}. Figure 3.25 shows the simulated structure along with the doping profile in the case of a high resistivity epitaxial layer. Black circles highlight the collecting electrodes of the central pixel, while the other orange areas model the N-wells

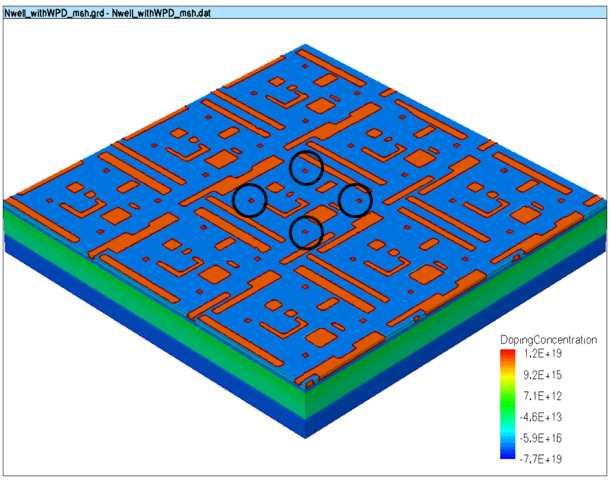

Figure 3.25. 3x3 matrix layout and doping profile. Black circles highlight the collecting electrodes of the central pixel.

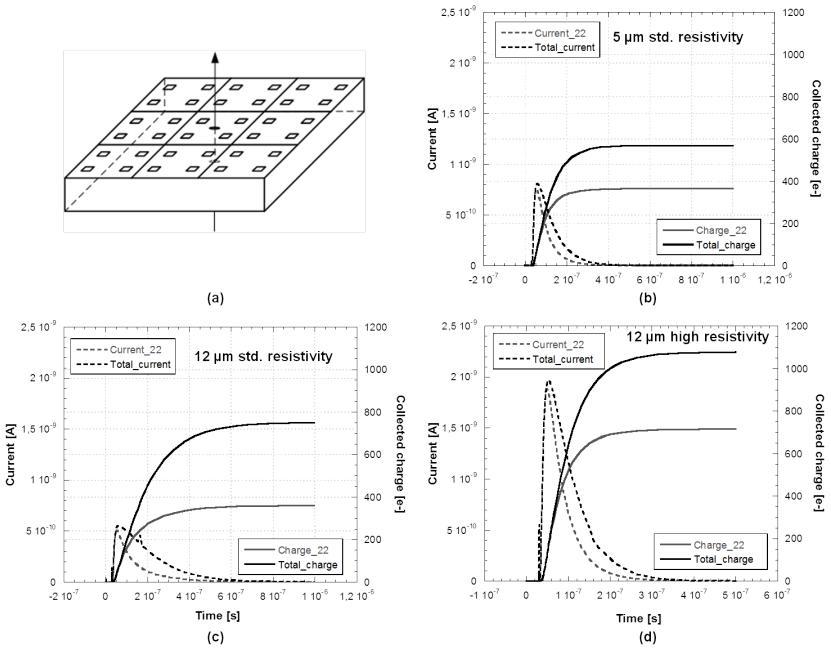

Figure 3.26. Current and charge collected by the central pixel and the whole matrix after the crossing of an impinging MIP in the matrix center (a).

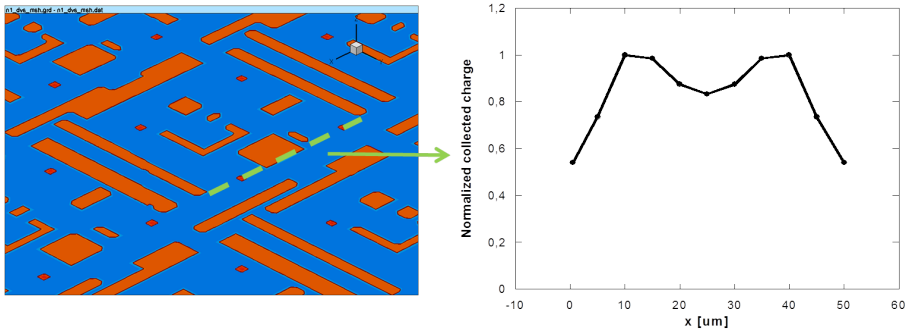

Figure 3.27. Charge collection normalized to the maximum value along the pixel cross-section passing through its collecting electrodes.

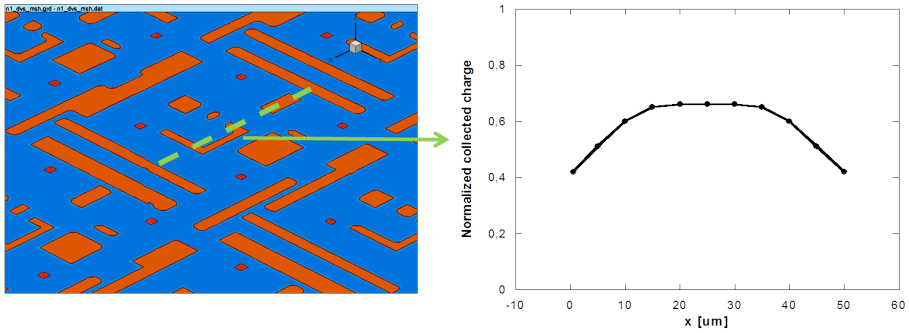

Figure 3.28. Charge collection normalized to the maximum value of Fig. 3.27 along the pixel section passing through its center.

hosting PMOS transistors. The deep P-well layer (barely visible in the figure) is deposited all over the 3x3 matrix except for the area under the collecting electrodes.
Figure 3.26 shows the current flowing at the output of and the charge collected by the central pixel (22) after the passage of a MIP along a track orthogonal to the matrix center (as shown in (a) in the figure). Moreover, also the total matrix current and charge are represented. The three plots refer to the cases of (b) a 5 μm standard resistivity epitaxial layer, (c) 12 μm standard resistivity epitaxial layer and (d) 12 μm high resistivity epitaxial layer. As expected, results show a better behavior in the case (d), in terms of both collected charge and collection time.

In order to evaluate the charge collection uniformity of this last solution, the MIP impact point has been moved along two cross sections of the central pixel (both parallel to the matrix side). The first section passes through the collecting electrodes (see Fig. 3.27), the second through the pixel center (shown in Fig. 3.28). The plot in Fig. 3.27 highlights the beneficial effects of the deep P-well implant, preventing the charge from being collected by the parasitic N-wells located under the green line. The minimum collected charge between the two peaks along this section is about 80% of the peak value, corresponding to the collecting electrodes position. The same comment about the deep P-well effectiveness can be done with reference to Fig. 3.28. The pixel features a good charge collection uniformity also along its center, where the charge collected reaches about 65% of the maximum value (obtained in correspondence of the collecting electrodes) along a large portion of the considered device section.

3.5 Digital readout

As described in the next section, the chip Apsel4well includes a 32x32 matrix with digital sparsified readout. The readout logic implemented in this matrix belongs to the SQUARE (Sequenced QUery Architecture for REad out) family [59]. This kind of peripheral readout technique requires a specific in-pixel

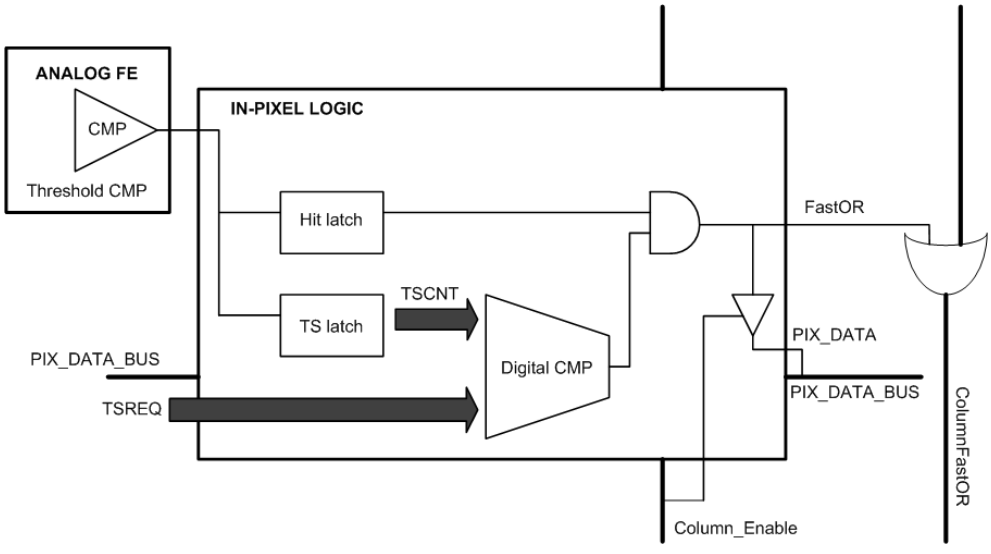

Figure 3.29. Conceptual representation of the in-pixel logic.

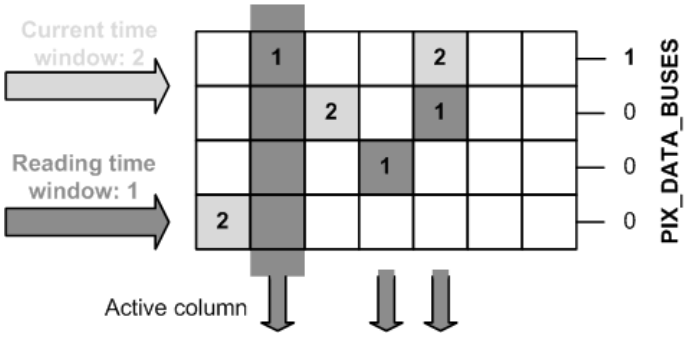

Figure 3.30. Conceptual representation of the hit extraction procedure.

digital logic and signal routing over the matrix. Figure 3.29 shows a conceptual representation of the in-pixel logic. A digital latch stores the hit information generated by a threshold crossing within each pixel. Along with the hit information, the pixel also contains an 8-bit time stamp. An 8-bit grey-coded time counter, driven by the peripheral readout, is distributed over the matrix with a dedicated bus called time stamp counter (*TSCNT*). The in-pixel logic is provided also with an 8-bit digital comparator. The peripheral readout can request a specific time stamp to all the matrix pixels with a dedicated bus called time stamp request (*TSREQ*), which is compared inside each pixel to the latched *TSCNT* value. When *TSREQ* matches the in-pixel latched *TSCNT*, the pixel *FastOR* signal goes high. Each pixel column implements a cascaded *Column-FastOR* connected to the peripheral readout. When a temporal time window (ticked by a dedicated clock called *BC*) is closed, the *TSCNT* grey-counter is incremented by 1. Each pixel is provided with a three-state buffer activated by the *Column_Enable* signal to drive the corresponding data bus line. When a column is active for readout, only the pixels with active hit latch and *FastOR* drive the corresponding pixel data bus line. The pixels automatically reset the hit latch after a reading operation. There are two other ways to reset a hit latch: one is by using a global reset, the other, meant to be used in triggered operating mode, is a time-dependent reset. When this signal is set high, only the pixels with an active *FastOR* are reset. The component responsible for the matrix interrogation and the hit extraction is the sweeper (shown later on in Fig. 3.31). A scheme of the hit extraction logic is shown in Fig. 3.30. In summary, the peripheral readout performs the hit extraction following these steps:

1. query the matrix with a closed *TSCNT* value;

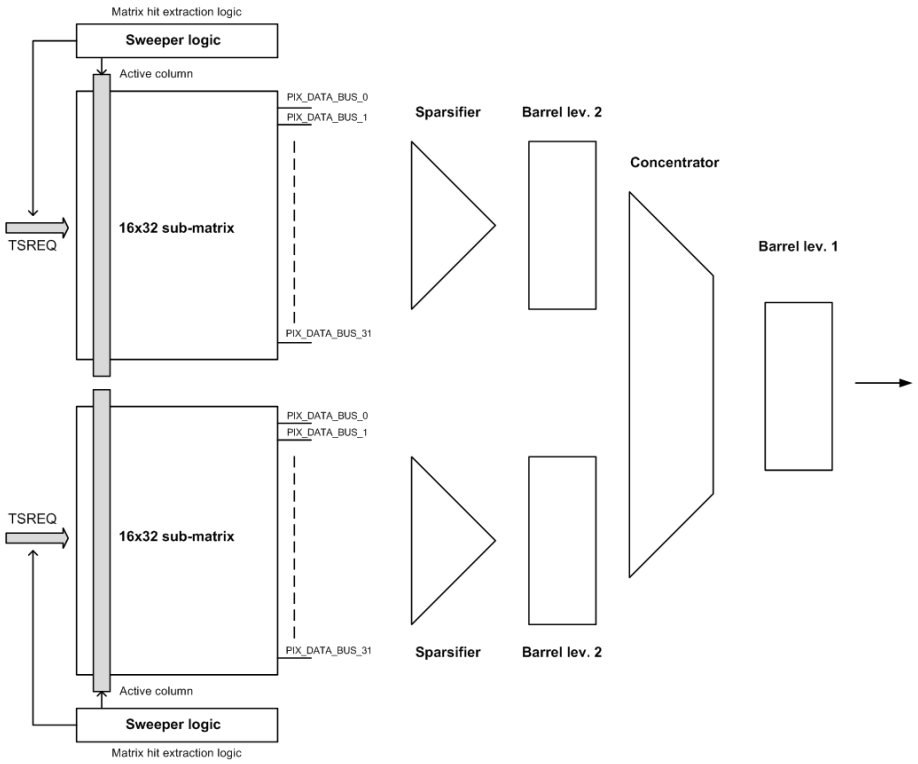

Figure 3.31. Conceptual representation of the readout logic.

2. look for the lowest index column with an activated *ColumnFastOR*;

3. enable the output data drivers of the pixels belonging to that column; the pixels in a row share the same *PIX_DATA* bus line;

4. go back to point 2 and loop until all *ColumnFastOr* are low.

The readout can work in two operating modes: data-push and triggered. In a typical data-push run, all the closed time windows are requested to the matrix and hits are extracted in a time sorted way. In triggered mode, only the hits belonging to a triggered time stamp are extracted from the matrix. The latched hits belonging to a time stamp which has not been triggered within a certain latency are erased from the matrix and discarded. The whole 32x32 sensor matrix is divided into two 16x32 logically independent units, called sub-matrices managed in parallel by the peripheral readout (see Fig. 3.31). The

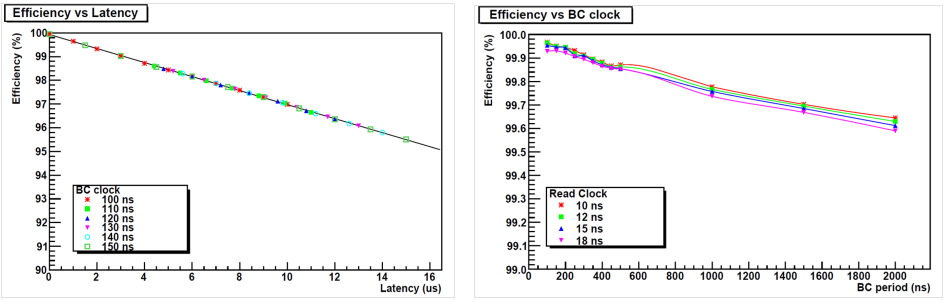

Figure 3.32. Simulation of the readout efficiency as a function of the trigger latency in the triggered (left) and data-push (right) operating modes. Both simulations have been performed at different readout clocks.

PIX_DATA bus of each sub-matrix is connected to the sparsifier component, that encodes the spatial and temporal information into a digital word. The storage element next to each sparsifier, called level 2 Barrel, can store the arbitrary number of encoded words produced by the coupled sparsifier, in the succesive clock cycle. It basically acts as a FIFO memory. Thanks to this feature it is possible to store many hit words in the same clock cycle, allowing a very fast read/reset sweep over the hit pixels. Only the column with an active *ColumnFastOR* are read out one after the other, and reading/resetting a column takes only one clock cycle independently of the column occupancy. The data stored in the two level 2 Barrel is then conveyed on a level 1 Barrel by a concentrator that merges the flux of data preserving the time order.

As already mentioned before, the described digital matrix interface was designed both for a data-push and triggered operating mode. In this second case the pixel latches act as data buffers during the trigger latency, saving silicon area usually dedicated to the storage of sparsified hits as well as dramatically reducing the chip output bandwidth. Since the pixel occupancy becomes a function of the trigger latency, the triggered working mode affects the detection efficiency. However, an efficiency of 98% has been simulated by introducing the 6 μs latency and the 100 MHz/cm^2 hit rate expected for Layer0 (including a safety factor of 5), which is not far from acceptable values [60]. Simulation results obtained with a Monte Carlo hit generator and a VHDL architecture model operated in trigger mode are represented on the left of Fig. 3.32. On the right, results obtained by operating the architecture in data-push mode are shown. Inefficiencies in this case are much lower since the hits are read out of the matrix as soon as a time window is closed. The occupancy is improved and the efficiency is constantly above 99,6% in a wide range of BC clock period.

3.6 Layout of the chip Apsel4well

This section is aimed at introducing and discussing the layout of the Apsel4well chip. First of all, the single pixel layout will be thoroughly described, paying attention on some design features aimed at minimizing the capacitive coupling between the input and the output of the analog channel and to improve the immunity to cross-talk phenomena between the analog and the digital in-pixel sections. Then, the 3x3 test matrix layout will be shown, with focus on the position of the digital lines used to test the immunity achieved with the above mentioned design. Finally the layout of the whole chip will be introduced with the description of the included test structures.

3.6.1 Pixel layout

Figure 3.33 shows a simplified layout of the Apsel4well pixel highlighting the deep P-well and N-well layers (top picture), and the complete pixel layout (bottom picture). As already mentioned in the previous sections, the deep P-well layer has been deposited all over the pixel, except for the area around the collecting electrodes. In this way, all the competitive N-wells are shielded from the epitaxial layer, therefore preventing the charge from being subtracted from the actual sensor. The collecting electrode dimensions have been chosen as the best compromise between charge collection performance of the pixel, charge sharing between adjacent pixels and collecting electrode capacitance by means of the device simulations shown in the section 3.4. The mutual distance between the square N-well diffusions is D=27 μm, their side is L=1.5 μm, while the side of the square area without the deep P-well layer is equal to W=8 μm. From the complete pixel layout shown at the bottom of Fig. 3.33, it is possible to observe the separation between the analog front-end section, located in the region among the collecting electrodes, and the in-pixel logic, located in the surrounding area. This separation, along with other design features shown later on, is mandatory in order to avoid capacitive coupling between the analog and the digital sections, that would be detrimental for the correct operation of the analog section. The pixel supply distribution is addressed by means of vertical metal lines. Since the analog power supply is distributed to a large number of pixels (at least in a large pixel matrix) and each pixel drains a certain amount of dc current, the design of analog supply lines has been addressed by minimizing their resistance in order to avoid undesired voltage drop along the lines. The positive analog supply (AVDD) is distributed to the pixel through a top metal line (metal6) 15.5 μm wide, while the ground (AGND) is distributed through a couple of top metal lines 7.75 μm wide at the two sides of the AVDD

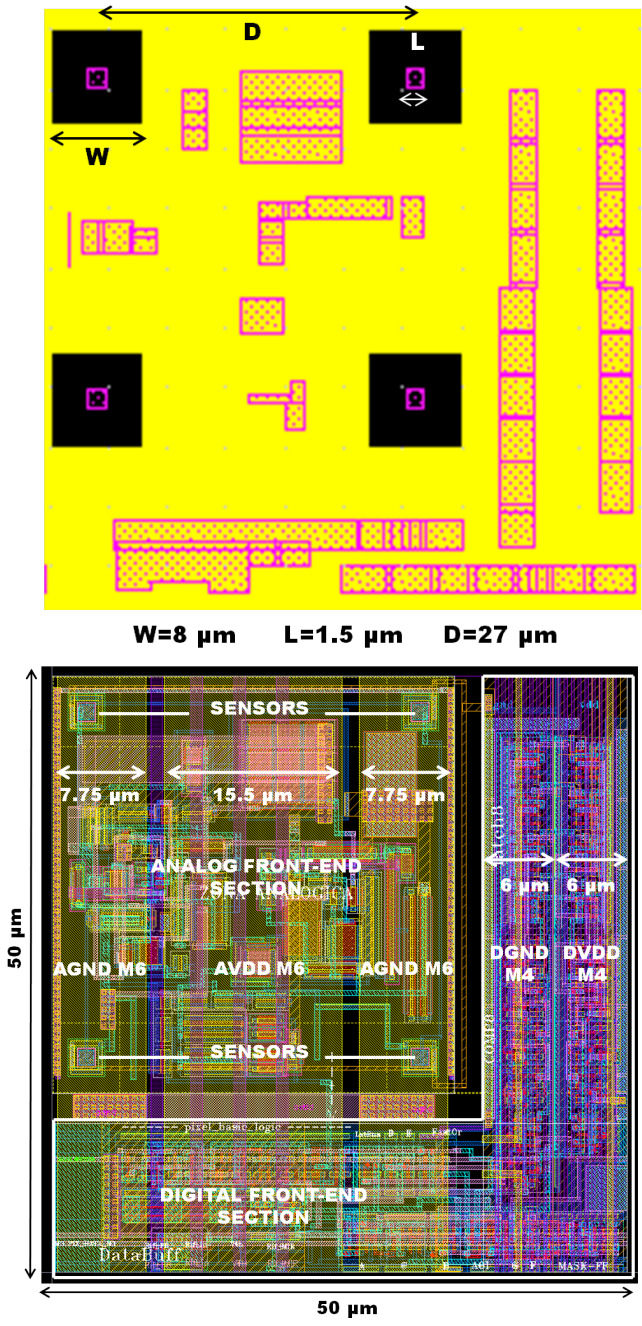

Figure 3.33. Simplified pixel layout with deep P-well and N-well layers (top) and complete pixel layout (bottom).

line. The choice of the top metal level is motivated by its lower sheet resistance (40 mΩ/square) with respect to the other metal layers (80 Ω/square). The digital supply distribution is less critical under this point of view and has been addressed by means of two wide metal4 vertical lines (DVDD and DGND). The use of two different metal layers for the pixel analog and digital power supplies results in a further reduction of their capacitive coupling.

3.6.2 3x3 test matrix layout

The layout of a 3x3 test matrix of the chip Apsel4well is shown in Fig. 3.34. The transistors making up the digital front-end section of the pixels included in this matrix are disconnected from the digital power supply. This has been done for two reasons. First, this test matrix has been designed for the characterization of the analog front-end section and of the sensor charge collection properties. Second, the disconnected digital metal lines can be used to test the immunity of the analog section to digital line switching.
In Fig. 3.34 these groups of externally accessible digital lines can be seen to be located along the horizontal and vertical digital lanes. The measurement results relevant to the immunity tests will be shown in the next chapter. However, the effect of digital line switching on the analog front-end section has been evaluated by means of post-layout simulations during the design phase.

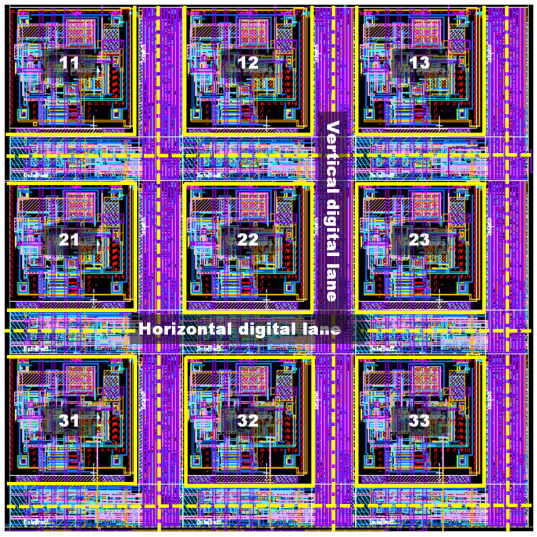

Figure 3.34. Layout of a 3x3 test matrix of the Apsel4well chip.

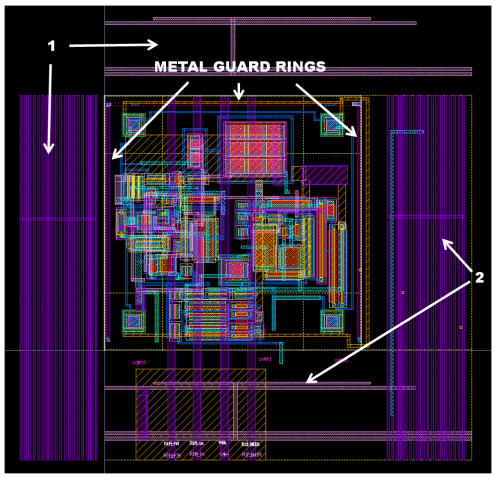

Figure 3.35. Simplified pixel layout, including the digital lines of one of the adjacent pixels (1), the digital lines of the pixel itself (2) and the metal guard ring.

Figure 3.35 shows the simplified version of the layout of a pixel, including the digital lines of one of the adjacent pixels in the matrix (1) and the digital lines of the pixel itself (2). Moreover, the layout also includes a metal guard ring structure made up by stacking five metal layer strips all interconnected and connected to ground. The aim of this structure is to further decouple the analog section from the digital section. Its effectiveness has been studied in post-layout simulations by stimulating all the digital lines with a square voltage waveform (ranging from 0 to $+V_{DD}$) with a period of 2 μs and a 50% duty cycle while injecting an 800 electron pulse at the preamplifier input. The post-layout simulation results are shown in Fig. 3.36. On the left, both the charge preamplifier input and the shaper output waveforms are shown in the case of the pixel without metal guard ring. Waveforms on the right refer to the case where this structure is included in the pixel layout. The beneficial effects of the metal guard ring are quite evident. Actually, in the case where the guard ring is absent, the value of the parasitic coupling capacitance between the charge preamplifier input and the digital lines has been estimated by the parasitic extractor QRC as 400 aF. When the guard ring is present, this capacitance is estimated to be less than 1 aF and no disturbances superimposed on the circuit response to the 800 electron input pulse can be detected.

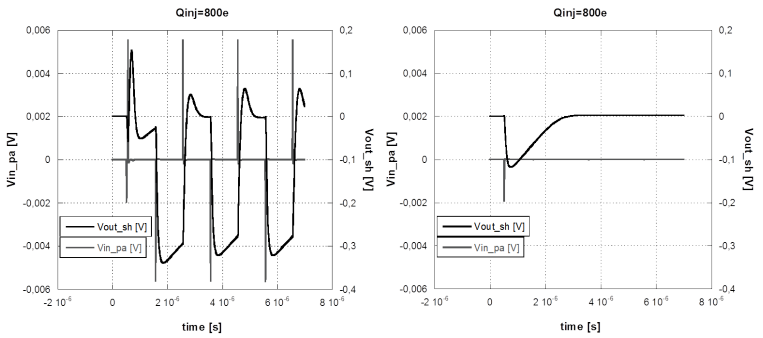

Figure 3.36. Post-layout simulations of the waveforms at the preamplifier input and shaper output without (left) and with metal guard ring (right).

3.6.3 Chip layout

The layout of the chip Apsel4well is shown in Fig. 3.37. The chip size is 5x5 mm^2. It features a 32x32 matrix (identified with F in the figure) with a sparsified digital readout architecture and four 3x3 matrices designed with the aim of studying the analog front-end behavior and the sensor performance in terms of charge collection. The four matrices differ in the number of collecting electrodes, presence or absence of the deep P-well implant and the layout of the charge preamplifier input device (with or without enclosed layout).This was done to evaluate the influence of each of these parameters on the main pixel performance (signal-to-noise ratio, charge collection efficiency and radiation hardness). Moreover, the chip also includes some test structures, namely octagonal pn diodes and analog channels featuring different input capacitances emulating the sensor capacitance. The octagonal diodes have been included with the aim of obtaining more information on the collecting electrode behavior, especially in terms of leakage current, after exposure to radiation, while single analog channels have been included in order to evaluate the analog front-end performance with varying detector parasitic capacitance. Table 3.8 summarizes the features of each structure included in the Apsel4well chip, with reference to Fig. 3.37.

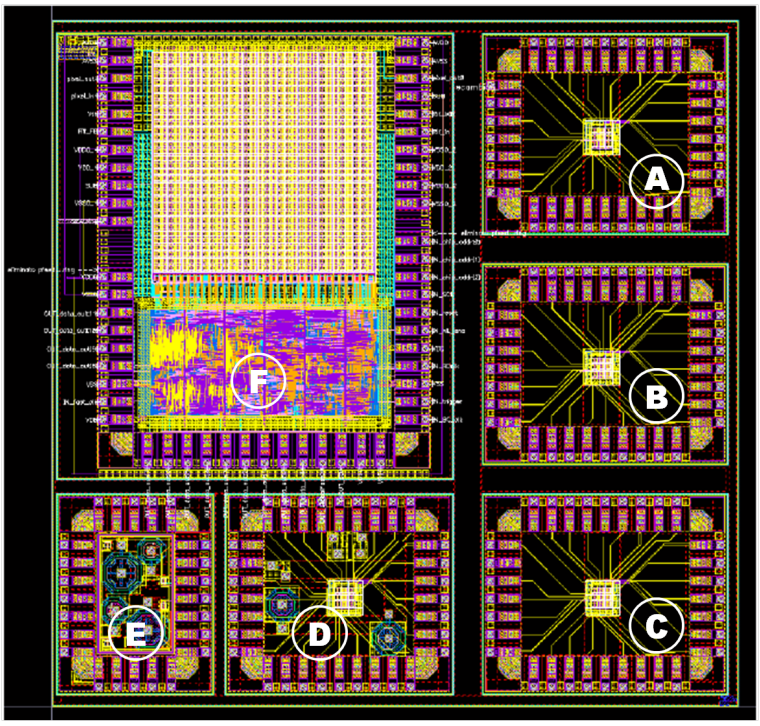

Figure 3.37. Layout of the Apsel4well chip.

Structure	Description
A	3x3 matrix. Pixel with 4 collecting electrodes, deep P-well and charge PA with enclosed layout input transistor.
B	3x3 matrix. Pixel with 4 collecting electrodes, deep P-well and charge PA with open layout input transistor.
C	3x3 matrix. Pixel with 2 collecting electrodes, deep P-well and charge PA with enclosed layout input transistor.
D	3x3 matrix. Pixel with 4 collecting electrodes, no deep P-well and charge PA with enclosed layout input transistor.
E	Octagonal pn diodes with different junction area and single analog channels.
F	32x32 matrix with peripheral sparsified digital readout architecture.

Table 3.8. Description of the structures included in the Apsel4well chip.

Chapter 4

Apsel4well characterization

This chapter is aimed at showing and discussing the experimental results obtained from the characterization of the Apsel4well chip. The first section is dedicated to the results from the characterization of the analog front-end readout, in terms of charge sensitivity and equivalent noise charge. The experimental data will be compared with simulation results also to assess the reliability of the device models provided by the foundry. Then, results from the characterization of the cross-talk between the digital section and the analog section and between the digital section and the collecting electrode will be shown. A significant part of this chapter is aimed at showing the results from the characterization of the Apsel4well sensor in terms of charge collection performance. Results obtained from the test of chips with different epitaxial layer thickness and resistivity will be discussed and compared. Finally, data from a non-ionizing irradiation campaign performed on different samples having 12 μm thick standard and high resistivity epitaxial layer will be presented, along with the modelization of the radiation induced loss in the collected charge, implemented in a Monte Carlo simulation tool.

4.1 Analog front-end characterization

In this section results from the characterization of the Apsel4well analog front-end will be shown. The measured parameters include the charge sensitivity and the equivalent noise charge also as a funtction of the channel bias condition. All of these measurements have been performed on the 3x3 test matrix of the A and D kind described in Tab. 3.8. Moreover, results concerning the threshold dispersion performance have been obtained from the characterization of the 32x32 matrix.

4.1.1 Charge sensitivity measurements with charge injection

The setup used for charge sensitivity measurements by means of charge injection techniques features a waveform generator Agilent 33250A, an Agilent E3631A power supply and a LeCroy Waverunner 64Xi digital scope having a bandwidth of 600 MHz and sampling frequency of 10 gigasamples per second. The test board used for 3x3 matrix characterization and powered at ±7 V is shown in Fig. 4.1. The chip under test is located in the center of the board. Here a circular hole can be noticed which has been used for charge collection measurements performed by means of a laser source, shown in section 4.3. All the chip reference voltages mentioned in the previous chapter (V_{RIF_IN}, V_{RIF_FB} and V_{RIF_MIR}) and the chip supply voltages can be independently set by means of dedicated trimmers. The characterization of the analog front-end has also been carried out for different values of some of the reference voltages mentioned above in order to test the dependence of the front-end performance on some key parameters. The nominal bias voltages of the chip are summarized in Tab. 4.1.

The charge sensitivity has been measured by injecting a 1 kHz square waveform

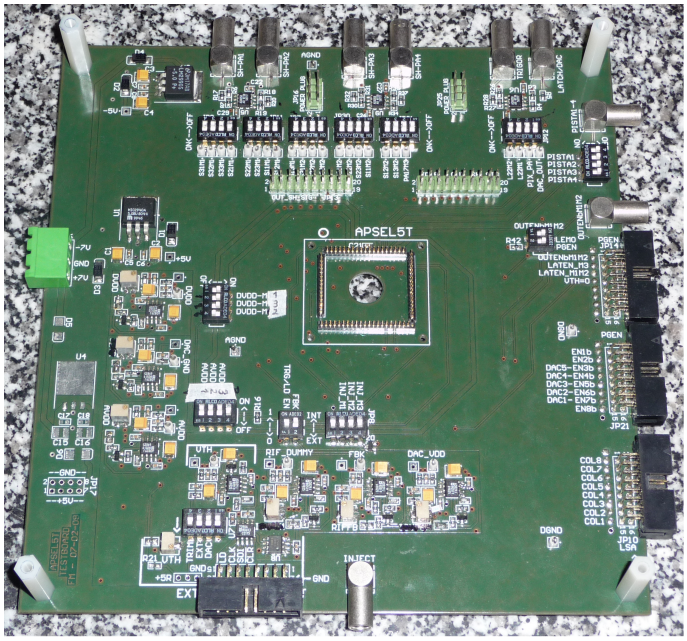

Figure 4.1. Test board used for Apsel4well 3x3 matrix characterization.

Voltage reference	**Nominal value** [V]
V_{RIF_MIR}	1.311
V_{RIF_IN}	1.045
V_{RIF_FB}	0.16
AVDD	1.8
AGND	0

Table 4.1. Nominal bias voltages of the Apsel4well chip.

with amplitude between 0 and 8 mV at the input of the 3x3 matrix central pixel, featuring an injection capacitance C_{inj}=30 fF. The relationship between the waveform amplitude V_{in} and the injected charge Q_{inj} is expressed as follows:

$$Q_{inj} = C_{inj} \cdot V_{in}. \tag{4.1}$$

Therefore, the injected charge ranges from 0 to 1500 electrons. This interval has been chosen after estimating (as shown in section 4.3) the most probable value of the collected charge relevant to a MIP to be about 750 electrons. In order to plot the input-output characteristic, the shaper output (through a buffer stage included in the test board) has been acquired at each V_{in} step by means of the digital scope. Eight different chip prototypes have been characterized: three of these (chip nuber 2, 4 and 5) feature a 5 μm thick standard resistivity epitaxial layer and a deep P-well layer. One other prototype, referred to as chip 6, features the same epitaxial layer, but without deep P-well implant. Finally, two samples (chip 31 and 32) feature a 12 μm thick standard resistivity epitaxial layer and two other chips (41 and 42) feature a 12 μm thick high resistivity epitaxial layer (all with a deep P-well implant). Table 4.2 summarizes the tested sample features, while Tab. 4.3 shows the measured charge sensitivity both in the nominal conditions shown in Tab. 4.1 and by setting the V_{RIF_MIR} voltage to such a value as to obtain the same return to baseline time t_{rb} of the output waveform as in circuit simulations. Actually, during the characterization phase, in nominal bias conditions t_{rb} was found to be smaller than in simulation. This can be clearly seen in Fig. 4.2, where the measured shaper output waveform from different chips, obtained by injecting 750 electrons at the channel input, is compared to the simulated waveform. All the waveforms have been obtained with nominal bias applied. This discrepancy between experimental data and simulations is probably due to a higher

DUT	**Epi. thick. [μm]**	**Epi. res. [Ω·cm]**	**DPW**
2	5	10	Yes
4	5	10	Yes
5	5	10	Yes
6	5	10	No
31	12	10	Yes
32	12	10	Yes
41	12	10^3	Yes
42	12	10^3	Yes

Table 4.2. Features of the tested samples

DUT	**Gq [mV/fC] V_{RIF_MIR}=1.311 V**	**V_{RIF_MIR} [V] t_{rb}=2 μs**	**Gq [mV/fC] t_{rb}=2 μs**
2	793	1.405	970
4	798	1.405	959
5	758	1.405	906
6	719	1.405	883
31	746	1.425	965
32	744	1.425	947
41	1014	1.38	1123
42	809	1.405	1018

Table 4.3. Measured charge sensitivity at different V_{RIF_MIR} values.

discharge current I_{MIR} in the actual circuits (see Fig. **??**) than the 20 nA set in the design phase. A faster C_2 discharge also leads to a smaller waveform peak value, as discussed in section 3.2.2. Again in Fig. 4.2, a significant difference can be detected between the response in chip 41 and the waveforms from the other chips. This behavior is probably related to mismatch phenomena between transistors MN1 and MN2 making up the current mirror in the shaper feedback network.

Table 4.3 shows the measured charge sensitivity values in the case of samples featuring standard resistivity epitaxial layer, taken at a V_{RIF_MIR} value such that t_{rb}=2 μs. The experimental data are very similar to the simulation results shown in chapter 3 (930 mV/fC). In samples 41 and 42, featuring a high resistivity epitaxial layer, the measured charge sensitivity is up to 20% higher

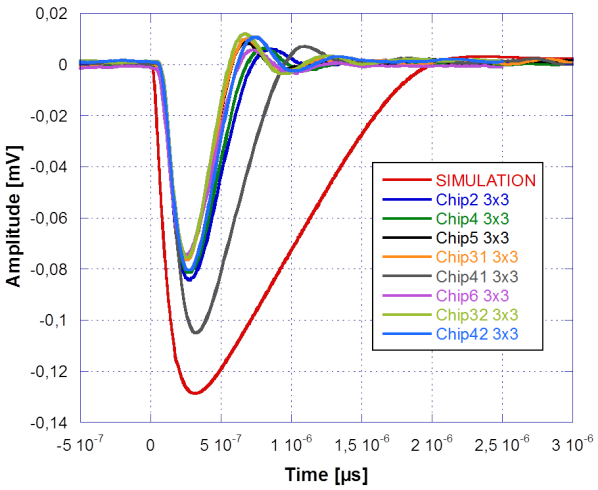

Figure 4.2. Shaper output waveform measured in nominal conditions in different Apsel4well samples compared to simulation results.

than in simulation. This effect seems to point to a dependence of the analog front-end performance on the epitaxial layer resistivity. The reason for this phenomenon could lie in a slightly different channel doping concentration in the devices integrated in the two different epitaxial layers, which may lead to different I-V characteristics. Further charge sensitivity measurements have been performed by varying the reference voltages listed in Tab. 4.1. In the next section, results obtained as a consequence of V_{RIF_MIR} variations (which led to the most interesting results) will be shown.

Charge sensitivity for varying V_{RIF_MIR}

As already explained in chapter 3, the reference voltage V_{RIF_MIR} controls the shaper feedback current I_{MIR}. Figures 4.3 and 4.4 show the impulse responses (left) and the normalized charge sensitivity (right) of chips 2, 4, 5 and 6 and 31, 32, 41 and 42 respectively. In the considered interval of V_{RIF_MIR} values, corresponding to an I_{MIR} value ranging from about 3 nA to 70 nA, a remarkable charge sensitivity increase can be detected (in the order of 60%). Moreover, it is possible to observe a variation in the t_{rb} from chip to chip for varying V_{RIF_MIR}. This variation is quite evident if one considers chip 32 and chip 41 wavefoms with V_{RIF_MIR}=1.405 V. Actually, in the first case, t_{rb}=1.5 μs, while in the second case t_{rb}=3 μs. This different behavior has already been

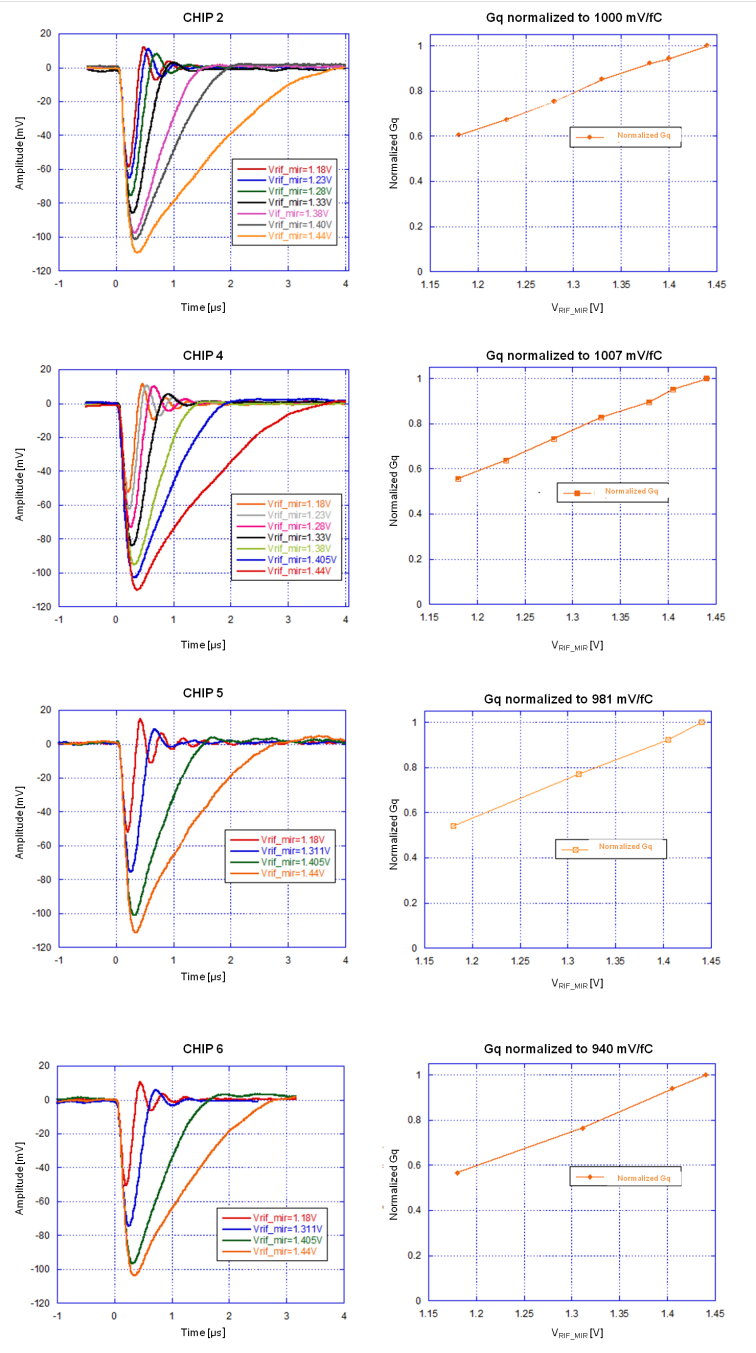

Figure 4.3. Channel response and normalized charge sensitivity of chips 2, 4, 5 and 6 for varying V_{RIF_MIR}.

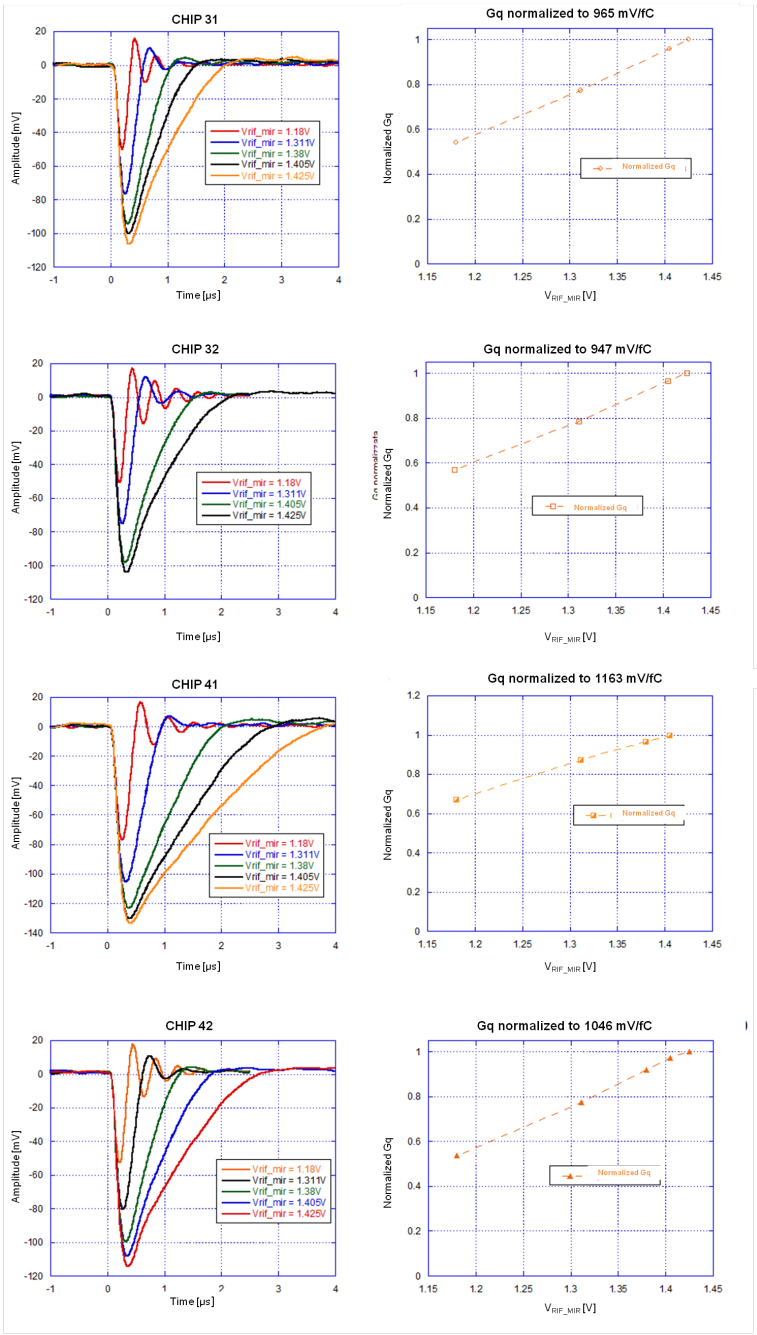

Figure 4.4. Channel response and normalized charge sensitivity of chips 31, 32, 41 and 42 for varying V_{RIF_MIR}.

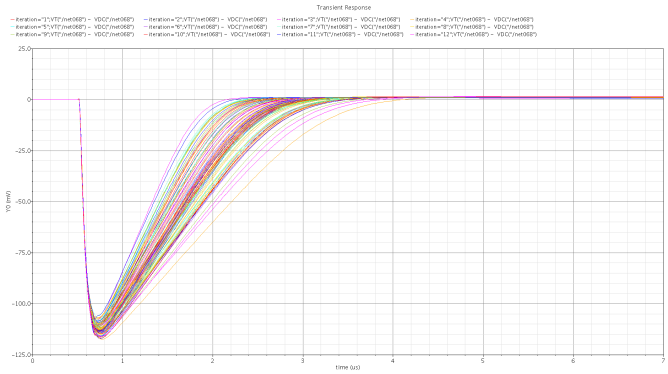

Figure 4.5. Monte Carlo simulation of the analog front-end response to an input charge Q_{inj}=750 electrons.

emphasized in Fig. 4.2 concerning the chips biased in nominal conditions. However, this behavior, in particular the spread in t_{rb}, could be accounted for through Monte Carlo simulations. The simulation results are shown in Fig. 4.5 in the case of a 750 electron input charge. This issue will have to be taken into account in the design of the next version of the Apsel4well analog front-end, for example by increasing the active area of the current mirror transistors in order to reduce their parameter mismatch.

4.1.2 ^{55}Fe charge sensitivity measurements

Charge sensitivity measurements shown in the previous section have been performed by means of charge injection techniques exploiting the presence of an injection capacitance at the input of the central pixel of the 3x3 matrix. This technique features a couple of drawbacks. First of all, the measured charge sensitivity value depends on the injection capacitance value, which may fluctuate because of process parameter variations. Second, this technique applied to the 3x3 matrix enables charge sensitivity measurements on the 3x3 matrix central pixel, the only one equipped with the injection capacitance, but is not suitable for gain measurements on the other matrix pixels. These are some of the reasons why measurements through ^{55}Fe radioactive source have been performed on the 3x3 matrix of Apsel4well. Actually, ^{55}Fe measurements also provide information on the sensor collection properties. The principles of this gain calibration technique have already been described in section 2.2.4. Figure 4.6 shows the event count rate distribution for all the pixels of the 3x3 ma-

trix in chip 42. From the plots, it can be clearly seen that the shaper output corresponding to the spectrum peak of the central pixel is about 30% higher than in the other eight pixels. In order to better understand this behaviour, new post-layout simulations of the 3x3 matrix have been performed, in order to account for the possible parasitic capacitances introduced by the matrix structure with its additional metal lines, not included in the layout of the single pixel. Figure 4.7 shows the charge preamplifier (left) and the shaper output waveforms (right) of the pixels 12 and 22 (the central one) of the simulated matrix. An overshoot at the output of the preamplifier, leading to a 25% higher peak value of the waveform at the shaper output, can be clearly detected in the case of pixel 22. All the pixels in the matrix (whose waveforms are not shown) feature a very similar behavior to pixel 12.

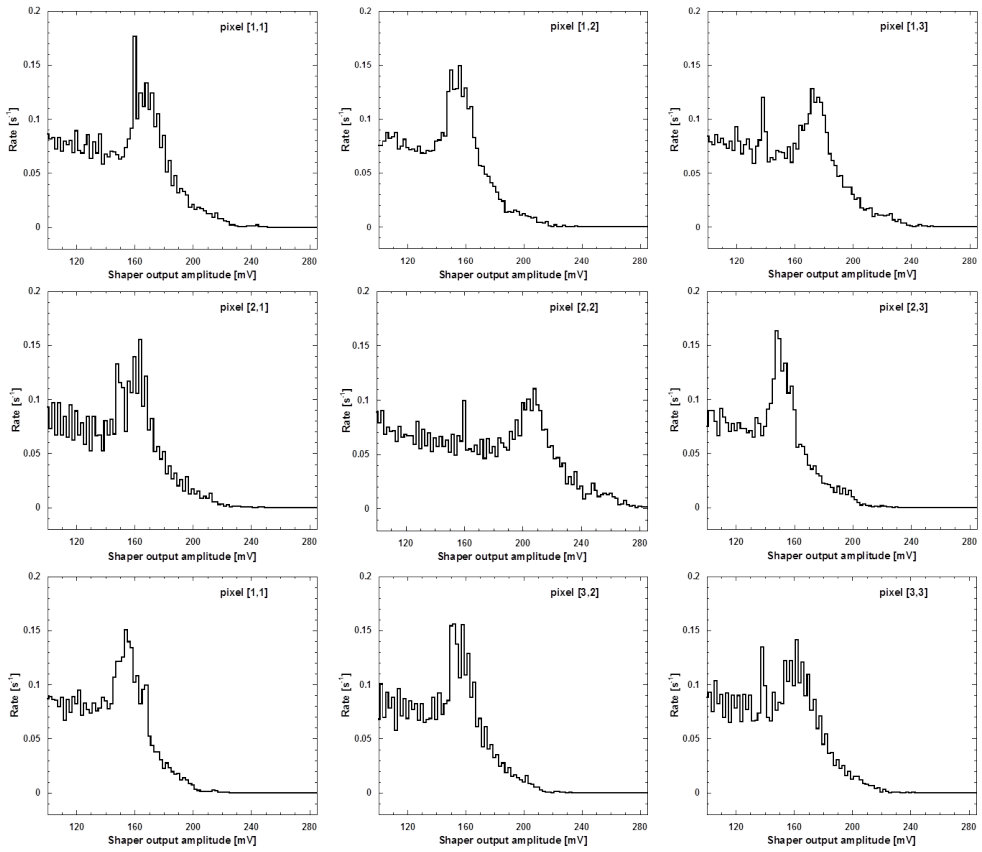

Figure 4.6. Event count rate distribution of the 3x3 matrix pixels of chip 42.

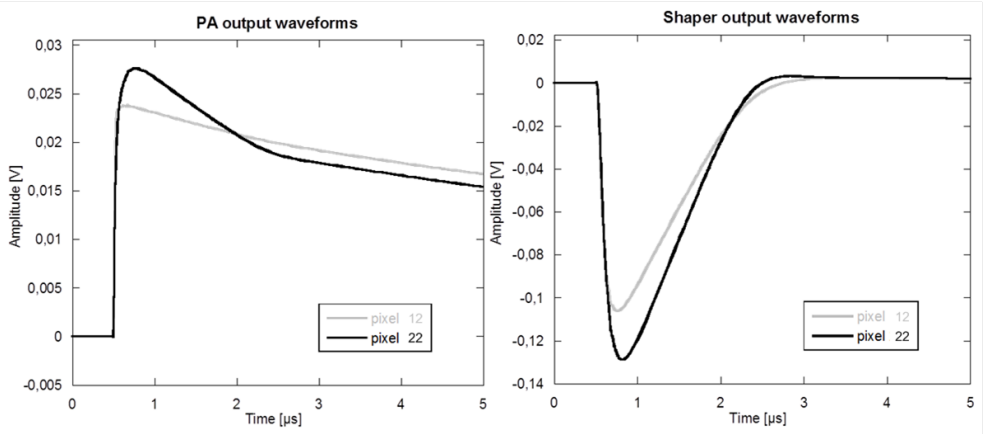

Figure 4.7. Post-layout simulation results for pixels 12 and 22 of a 3x3 matrix structure: preamplifier (left) and shaper (right) response to a 750 electron input charge.

Post-layout extraction shows that the coupling capacitance between the shaper output and the preamplifier input is identical for all the pixels of the first two rows (40 aF), except for the central pixel, where it is larger (200 aF). The pixels of the third row feature a coupling capacitance lower than 1 aF. The higher coupling capacitance value of the central pixel leads to the charge preamplifier overshoot and, as a consequence, to a higher charge sensitivity value. Figure 4.8 shows the 3x3 matrix layout along with the position of the injection capacitance and of the analog buffers connected to the pixel shaper outputs and used to drive the pad capacitance. From an inspection of the layout, it can be observed that pixel shaper outputs of the first two rows cross the metal 1 connection between two electrodes (inside the red circles). This metal crossing accounts for the 40 aF coupling capacitance contribution. In the central pixel, the coupling between the metal 6 input line and the metal 5 shaper output line is added (160 aF). This last contribution is responsible for the overshoot in the charge preamplifier output. Pixels from the third row do not feature any of these metal crossings.

4.1.3 Input-output characteristics

Figure 4.9 shows the input-output characteristic of the analog front-end of chips 31, 32, 41 and 42. Virtually the same trend have been found for the other tested chips (not shown). In this figure, the plots on the left represent

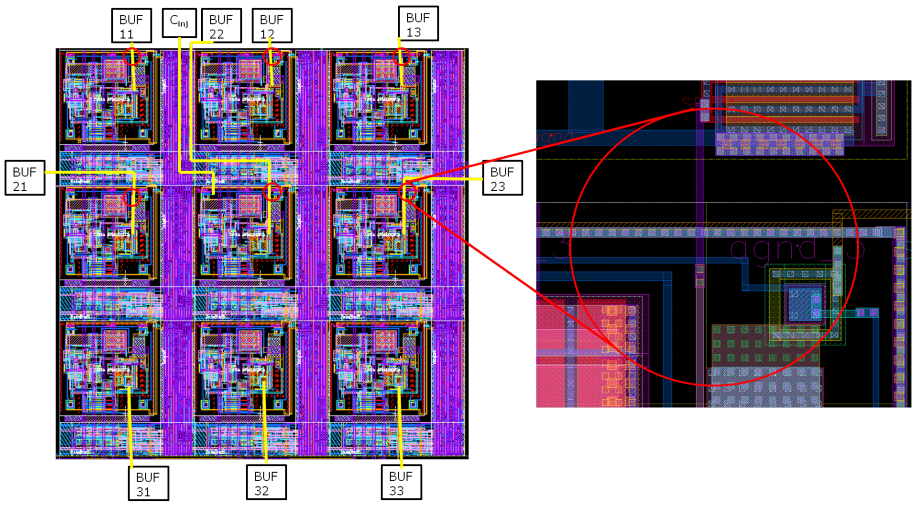

Figure 4.8. 3x3 matrix layout with details of the shaper output lines and their connection with analog buffers (left) and of the injection line of the central pixel (right).

the characteristic measured in nominal bias conditions, while on the right the characteristics have been obtained by setting V_{RIF_MIR} in order to get t_{rb}=2 μs with an injected charge of 750 electrons. In the nominal conditions, the INL ranges from about 4.5% to 6.5%, quite far from the 1% obtained in simulation. For the plots on the right, the INL range is narrower and lies between 1.5% and 2%, close to simulation results.

4.1.4 SLOWFAST simulation models

In Fig. 4.10, the shaper output response measured in nominal conditions on the one hand and with the V_{RIF_MIR} such that t_{rb}=2 μs on the other, obtained with an injected charge of 750 electrons, are shown along with the simulated waveform again in nominal conditions. The higher peak value of the simulated waveform can be clearly detected. However, by simulating the circuit in one of the device model corners, in particular in the SLOWFAST corner (slow NMOS - fast PMOS), a fairly good agreement between simulation and experimental results can be found. This is clearly shown in Fig. 4.11, displaying on the left the comparison between the shaper output as a response to a 750 electrons pulse simulated in SLOWFAST model corner and the one obtained from device characterization at V_{RIF_MIR}=1.33 V. On the right of the figure,

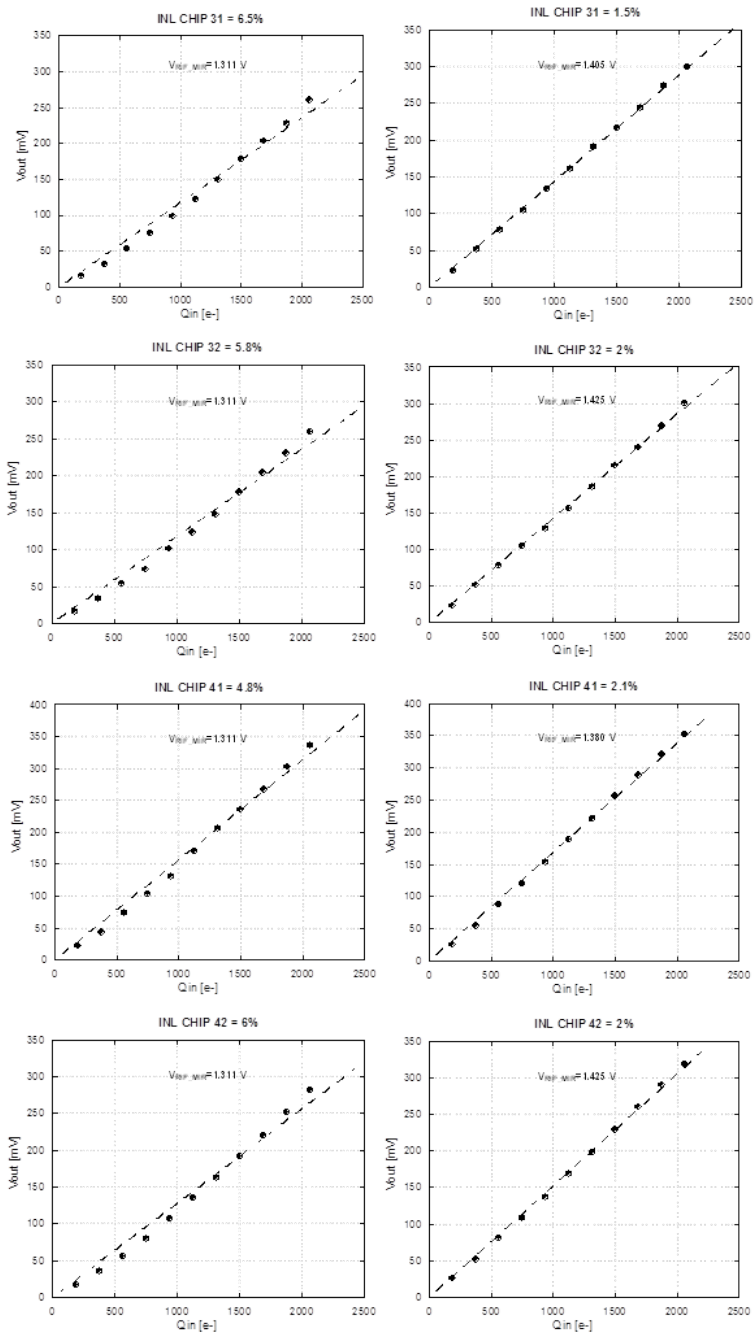

Figure 4.9. Measured input-output characteristics of the analog front-end for the central pixel in 3x3 matrices of chips 31, 32, 41 and 42: in nominal bias conditions (left) and with V_{RIF_MIR} such that t_{rb}=2 μs (right).

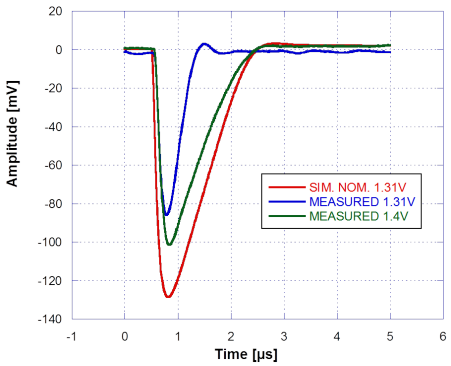

Figure 4.10. Comparison between simulated and measured shaper response with Q_{inj}=750 electrons.

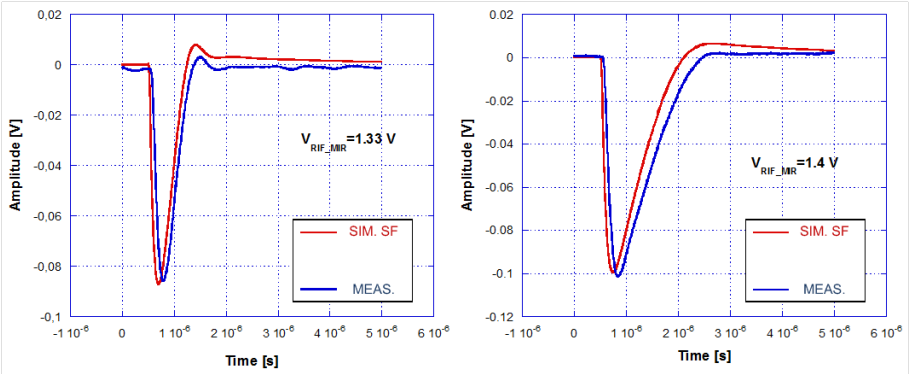

Figure 4.11. Comparison between simulations performed in the SLOWFAST (SF) model corner and measurements at different V_{RIF_MIR} values.

the same comparison is performed at V_{RIF_MIR}=1.4 V. Actually, the characterization of devices from two different runs provided very similar results, with the SLOWFAST corner models quite accurately reproducing the circuit behavior. This seems to point to a poor device modelization in the design kit provided by the foundry, since it seems very unlikely that devices from two different production batches be found to lie in the same model corner, far from the expected average behavior.

4.1.5 ENC measurements

The equivalent noise charge, already introduced in chapter 2, can be operationally defined by means of the following equation:

$$ENC = \frac{V_{rms}}{G_q}, \tag{4.2}$$

where V_{rms} is the root mean square value of the noise voltage at the output of the shaper. V_{rms} has been evaluated by observing the analog output of each pixel of the 3x3 matrix and taking advantage of the functions for statistical analysis available in the digital scope. A time scale of 2 ms/division, much larger than the typical processing time of the analog chain, makes it possible to include in the measurement also low frequency noise contributions. As in the case of charge sensitivity measurements, ENC measurements have been performed by varying the reference voltage V_{RIF_MIR} in order to observe the dependence of the ENC on the shaper feedback current. Figure 4.12 shows the ENC as a function of V_{RIF_MIR}. Chips 2 and 4 seems to feature a lower noise (ranging between 25 and 30 electrons) than the other chips. Simulations performed both with typical models and in the SLOWFAST corner provided an ENC value of 37 electrons, about the same value as in chip 41. A careful analysis of Fig. 4.12 leads to the conclusion that almost all the ENC plots show an optimum point in the considered interval. The ENC variation is probably

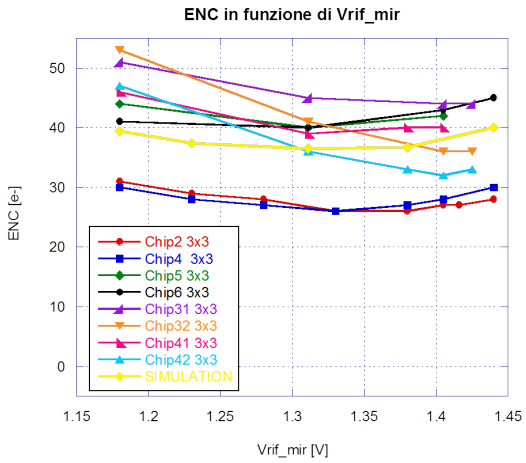

Figure 4.12. ENC of the tested chips for varying V_{RIF_MIR}.

related to the variation of the signal shaping time with V_{RIF_MIR}. For short shaping times (i.e., low V_{RIF_MIR}), the predominant ENC contributions come from series noise in the charge preamplifier input transistor. For higher shaping times, parallel noise contributions, coming from the preamplifier feedback network play a more significant role. Inside the considered V_{RIF_MIR} range, both terms contribute to ENC in a comparable fashion and give rise to the minimum in the ENC characteristic [61]. This consideration is confirmed by results of ENC simulations performed in the considered V_{RIF_MIR} range (also shown in Fig. 4.12).

4.2 Tests of immunity to cross-talk

As already mentioned in chapter 3, the layout of the Apsel4well 3x3 test matrix has been arranged to enable the test of the immunity to cross-talk due to coupling between the digital in-pixel circuits and the analog section and/or between the digital lines and the collecting electrodes. Figure 4.13 shows a simplified layout of the 3x3 matrix with emphasis on two groups of externally

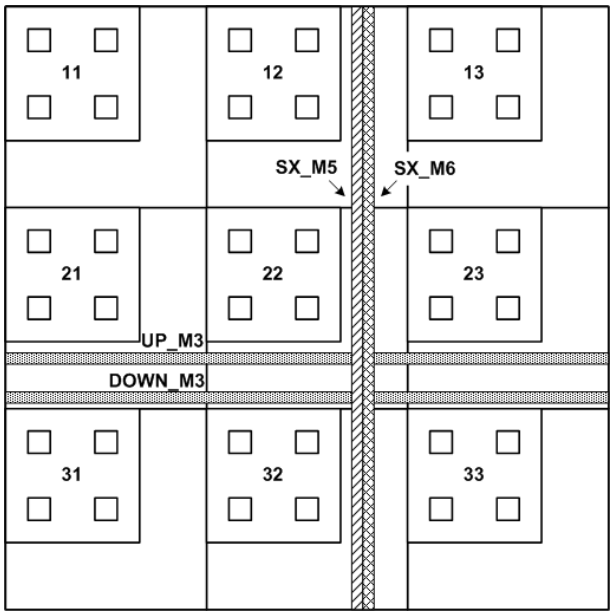

Figure 4.13. Simplified layout of the 3x3 matrix with emphasis on the externally accessible digital lines for immunity tests.

Figure 4.14. Response of the central pixel at the negative (left) and positive edge (right) of a square waveform simultaneously injected into all the digital test lines.

accessible digital lines. In order to perform the test, each of the lines has been fed by a square voltage waveform between 0 and 1.8 V. The response of the analog section next to the square waveform positive and negative edges has been evaluated. Moreover, in order to test the pixel immunity in the worst case conditions, the square voltage waveform has been simultaneously injected into all the digital lines and the analog shaper output of the central pixel (the only one surrounded on two of its four sides by digital lines) has been acquired with the scope. The results of this latter test, as the most significant example of the tests performed, is shown in Fig. 4.14. On the left of the figure, the response of the pixel analog section to the square waveform negative edge is shown, on the right the response to the positive edge. In the first case, an induced signal at the shaper output with an amplitude of about 5 mV can be detected. In the second case, the induced signal is almost negligible. Anyway, these results are satisfying and in good agreement with simulation results shown in section 3.6.2, confirming the effectiveness of the metal ring structures in limiting the capacitive coupling between the digital and the analog section and between the digital lines and the collecting electrodes. Actually, if one considers the 5 mV signal at the shaper output induced by a 1.8 V square voltage amplitude injected in all the digital lines and assumes a charge sensitivity of about 900 mV/fC, the overall parasitics coupling capacitance between the digital lines and the charge preamplifier input can be estimated to be in the order of 3 aF, not far from the 1 aF obtained by using the QRC extractor (and much lower than the 400 aF obtained by simulating the pixel without the metal guard ring structures).

4.3 Charge collection measurements

Charge collection measurements have been performed on the 3x3 test matrices integrated in the Apsel4well chip. The measurement procedure involves a laser source emulating the effects produced by a MIP crossing the sensor. The main purpose of these tests is to evaluate the collection efficiency of the sensor (at least in relative terms) and to detect possible inefficient regions inside the pixel. The energy released by the laser beam in the device substrate generates electron/hole pairs. A portion of this charge is collected by the sensors, while the remaining part recombines in the substrate. The charge collected by each single pixel of the 3x3 test matrix has been measured by illuminating the structure with the laser beam and acquiring the shaper output waveform. A simplified scheme of the measurement setup is shown in Fig. 4.15, already shown in chapter 2, but shown again here for sake of clarity. The beam source is a double eterojunction InGaAs/GaAlAs/GaAl laser diode adopting a Fabry-Perot configuration, operating at a 1064 nm wavelength. The beam at the focuser output features a Gaussian intensity profile with a standard deviation of about 3 μm in both dimensions of the plane transversal to the beam direction. In the laser tests, the samples are back-illuminated in order to avoid reflections from the large metal layers used to distribute power and ground to the front-end circuits, almost completely covering the silicon substrate, as shown in section 3.6.1. The laser diode is driven by means of the Agilent 33250A waveform generator, which also provides the generation of the

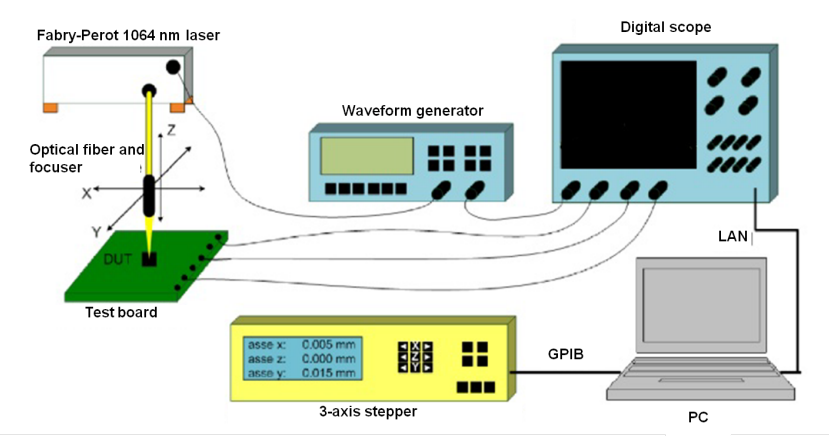

Figure 4.15. Charge collection measurement setup with laser source.

trigger signal for the digital scope. The waveform driving the laser diode is a periodic pulse with amplitude, frequency and duty cycle properly tuned in order to provide the necessary energy amount for generating about 80 electron/hole pairs per micron in the substrate (i.e. the energy released by an impinging MIP). The period of the pulse train is much larger than the typical processing times of the analog readout chain (i.e., the front-end peaking time). The measurement setup features also a Newport ESP300 precision scanning 3-axis stepper enabling micrometric control of the beam position over the device under test (DUT) along the x, y and z-axis. The shaper output waveform is acquired by means of the digital scope LeCroy Waverunner 64Xi having a bandwidth equal to 600 MHz and sampling frequency of 10 gigasamples per second (already introduced in section 4.1.1). The micrometric stepper and the digital scope are both controlled through a software interface developed in the LabVIEW environment. This interface makes it possible to program and automatically operate the setup, therefore enabling long (possibly overnight), high spatial resolution measurements with no need for human presence. In the measurements presented in this section, the incremental step in both x and y directions is 2.5 μm. Since the single pixel features an area of 50x50 μm^2, the number of scanning points per pixel is equal to 441. For each point, the software acquires the shaper output waveform coming from the digital scope and stores its peak value, along with the point coordinates, in a text file. Then, it moves the laser beam to the next point. Given the output waveform peak value V_{peak} and the pixel charge sensitivity G_q, it is possible to calculate the charge collected for each scanning point. In fact, if Q_{coll} is the collected charge at a given coordinate, then:

$$Q_{coll}[e^-] = \frac{V_{peak}[mV]}{G_q[mV/e^-]}. \tag{4.3}$$

4.3.1 Charge collection in a 3x3 matrix

The peak value of the collected charge for each pixel of a 3x3 test matrix of chip 5 is shown in Fig. 4.16. This sample features a 5 μm thick standard resistivity epitaxial layer. The considered matrix is the A structure of Table 3.8 featuring four collecting electrodes for each pixels and the presence of the deep P-well layer. It can be observed that the charge collected by the central pixel is 25-30% higher than the other pixels (except for pixel11). Actually, in order to evaluate the charge collected by all the 9 pixels, the assumption was made that all the pixels in the matrix have the same charge sensitivity as the

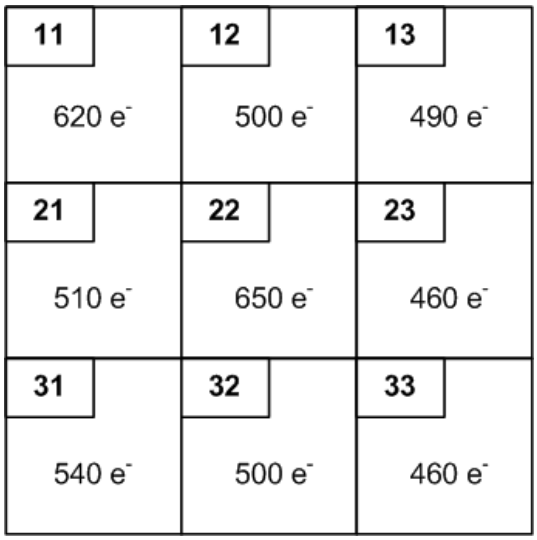

Figure 4.16. Peak value of the collected charge for each pixel of a 3x3 test matrix in chip 5.

central one, which instead was found to feature a larger gain due to the input-output capacitive coupling in the analog channel already discussed in section 4.1.2. This accounts for the apparently smaller amount of charge detected by the peripheral pixels in the matrix. A further confirmation of this behavior can be found in Fig. 4.17, showing the channel response to the laser stimulus for the 9 pixels of a 3x3 matrix in chip 4. In those measurements, the laser beam was in such a position to maximize charge collection. Again, the higher waveform amplitude for pixel22 with respect to those of the other pixels, all produced with the same amount of charge generated in the substrate by the laser beam, can be easily detected.
A more detailed representation of the charge collection distribution inside each pixel as a function of the laser beam impact point is shown in Figs. 4.18, 4.19 and 4.20, each representing the pixel response in a different column of the 3x3 matrix of chip 5. The in-pixel N-well layout has been superimposed to each color plot in order to emphasize the fact that charge collection reaches the maximum in the surrounding of the four square collecting electrodes. Moreover, it can be easily seen that, as expected, the amount of collected charge does not drop significantly around the competitive N-wells. As shown in more details in the next section, this is related to the presence of the deep P-well layer.

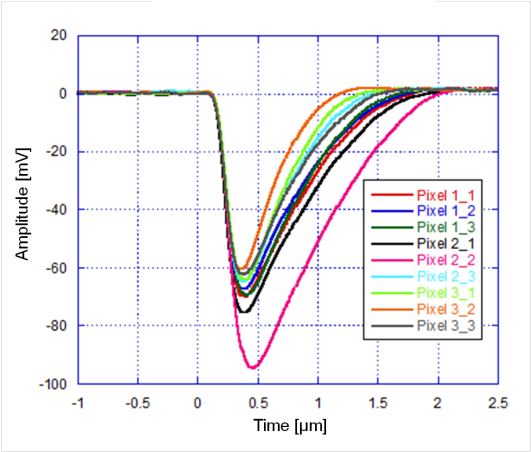

Figure 4.17. Channel response to the laser stimulus for the 9 pixels of a 3x3 matrix in chip 4 in conditions of maximum charge collection.

4.3.2 Advantages of using the deep P-well

Figure 4.21 represents the collected charge as a function of the laser beam position for pixels in the central row of 3x3 matrices in chip 6 (left) and chip 5 (right). Chip number 6 differs from chip 5 for the absence of the deep P-well implant underneath the competitive N-wells. As a consequence, in chip 6 the charge collection is confined around the four collecting electrodes. On the other hand, in the case of chip 5, a significant amount of charge is collected also when the laser beam is not close to the electrodes. This result proves the effectiveness of the deep P-well in shielding the competitive N-wells from the epitaxial layer. The absence of the deep P-well determines a dramatical drop of the charge collected by the sensor. Note that also the peak value of the collected charge in chip 6 drops to about 50% of the charge collected by chip 5.

Some interesting information about the pixel charge collection uniformity can be found by analyzing the charge collection behavior as a function of the laser beam position along two different sections of a 3x3 matrix, as shown in Fig. 4.22. The selected sections are located, one along the collecting electrodes (1) and the other along the pixel center in parallel to the first one (2). Figure 4.23 shows, on its top left side, the charge collected along section (2) and, on its top right side, the charge collection along section (1) in a 3x3 matrix belonging to chip number 6, the one without deep P-well implant. The amount of the

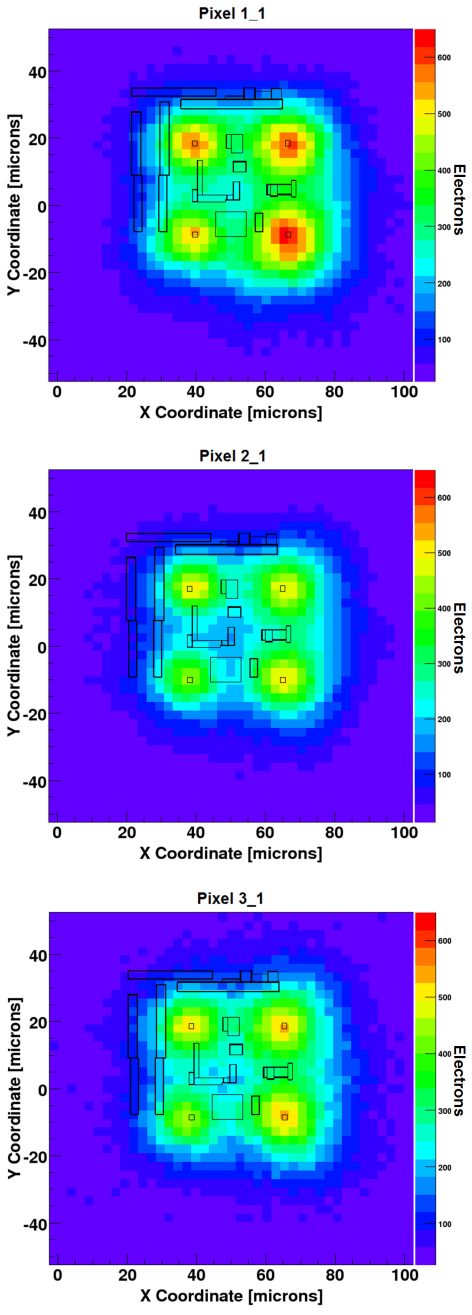

Figure 4.18. Collected charge as a function of the laser beam position in the first column of a 3x3 matrix of chip 5.

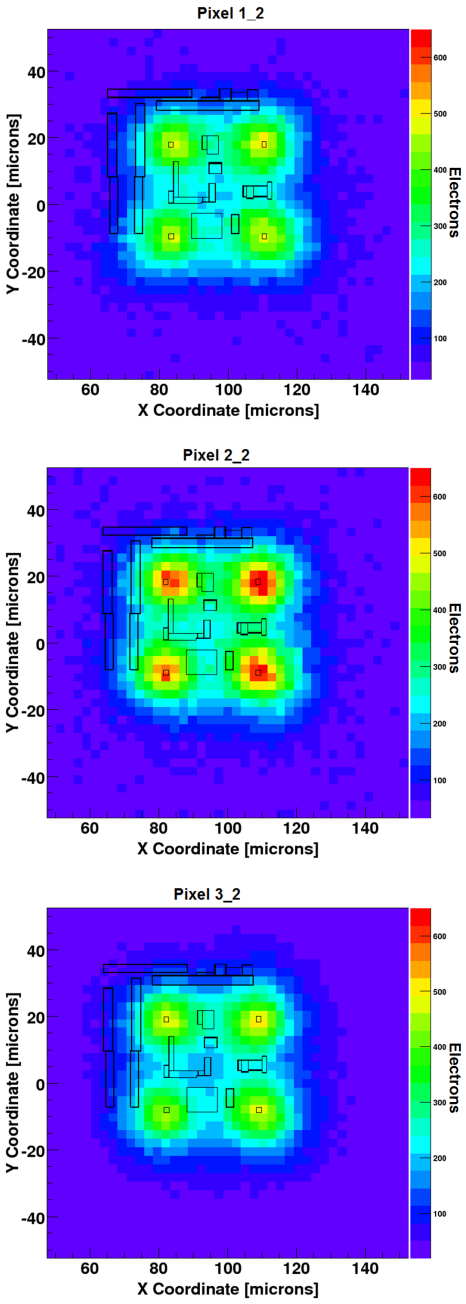

Figure 4.19. Collected charge as a function of the laser beam position in the second column of a 3x3 matrix (same as in Fig. 4.18) of chip 5.

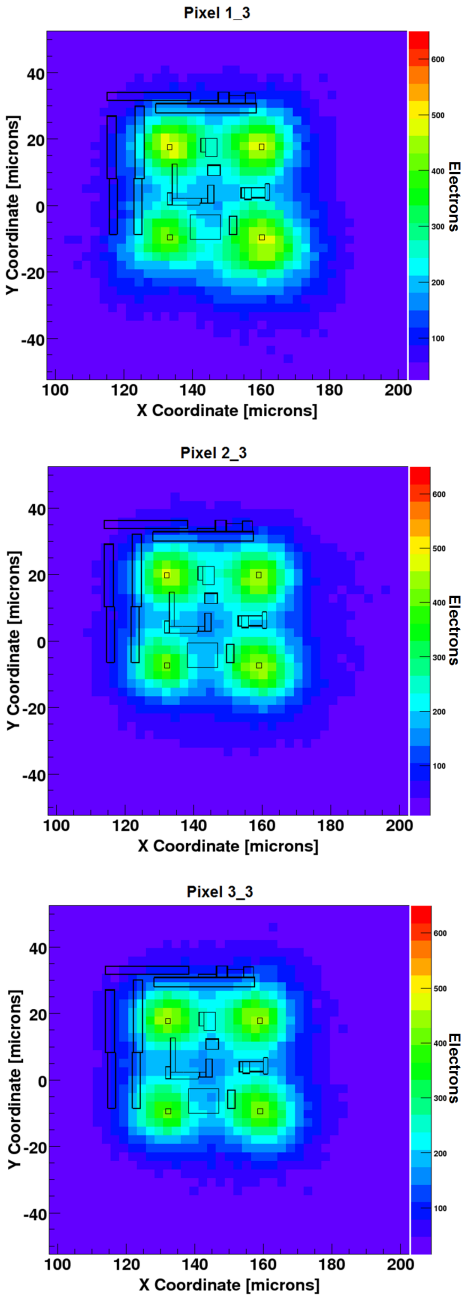

Figure 4.20. Collected charge as a function of the laser beam position in the third column of a 3x3 matrix (same as in Fig. 4.18) of chip 5.

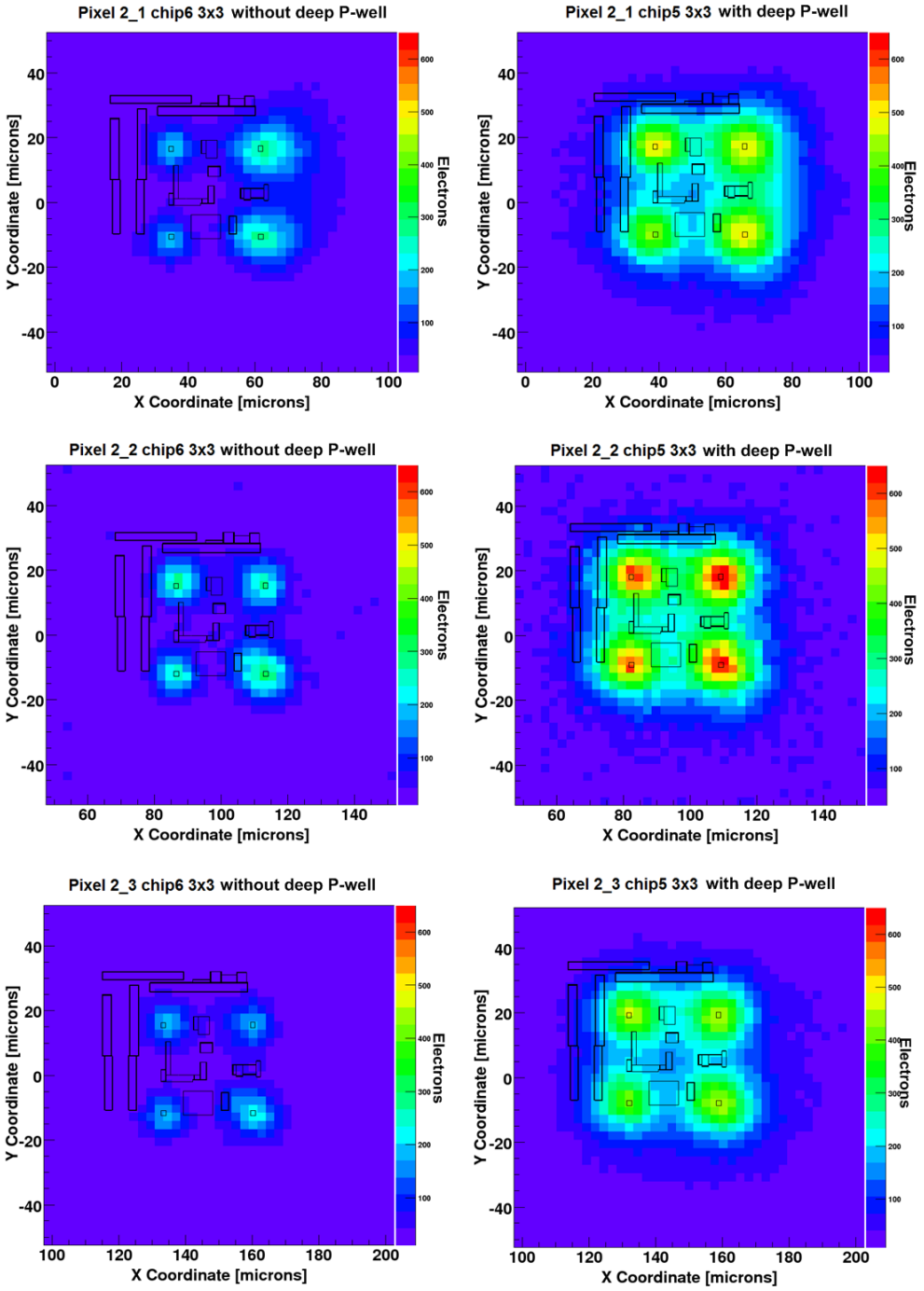

Figure 4.21. Collected charge as a function of the laser beam position in the second row of a matrix in chip 6 (left) and chip 5 (right).

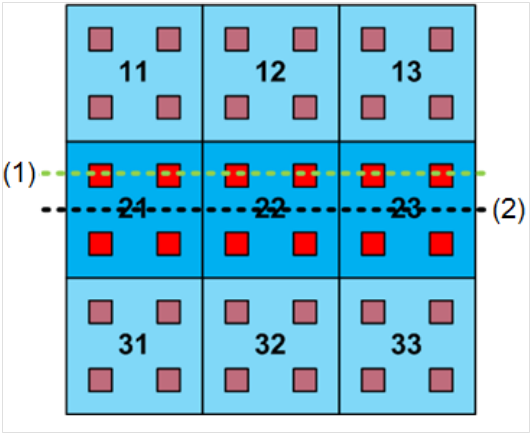

Figure 4.22. 3x3 matrix sections for charge collection distribution analysis.

collected charge along section (2) is quite small, not higher than 85 electrons. In the plot on the right, the charge collection peak, about 310 electrons, can be found next to the collecting electrode. The charge collection decreases very fast as soon as the laser moves away from the electrodes. The plots at the bottom of Fig. 4.23 show the same measurement results for chip 5, featuring the presence of the deep P-well implant. In this latter case, the charge collection is much more uniform inside the pixels (it never drops below 20% of the pixel maximum). The charge collected along the section (2) of the matrix varies between 30% and 40% of the peak value (which is about 650 electrons), while the charge collection along section (1) never drops below 50% of the peak value in each pixel.

4.3.3 Dependence of the charge collection on the epitaxial layer thickness

A study of the dependence of the charge collection on the epitaxial layer thickness has also been performed. For this purpose, the charge collection in a 3x3 matrix of chip 2 featuring a 5 μm thick, standard resistivity epitaxial layer has been compared to that of a matrix of chip 32, featuring a 12 μm thick, standard resistivity epitaxial layer. In both the samples, the deep P-well implant is present. Figure 4.24 shows the collected charge as a function of the laser beam position for pixels of the 3x3 matrix central row in both the chips. In chip 2, the charge is more confined around the collecting electrodes with respect to chip 32. Results in Fig. 4.24 lead to the conclusion that the use

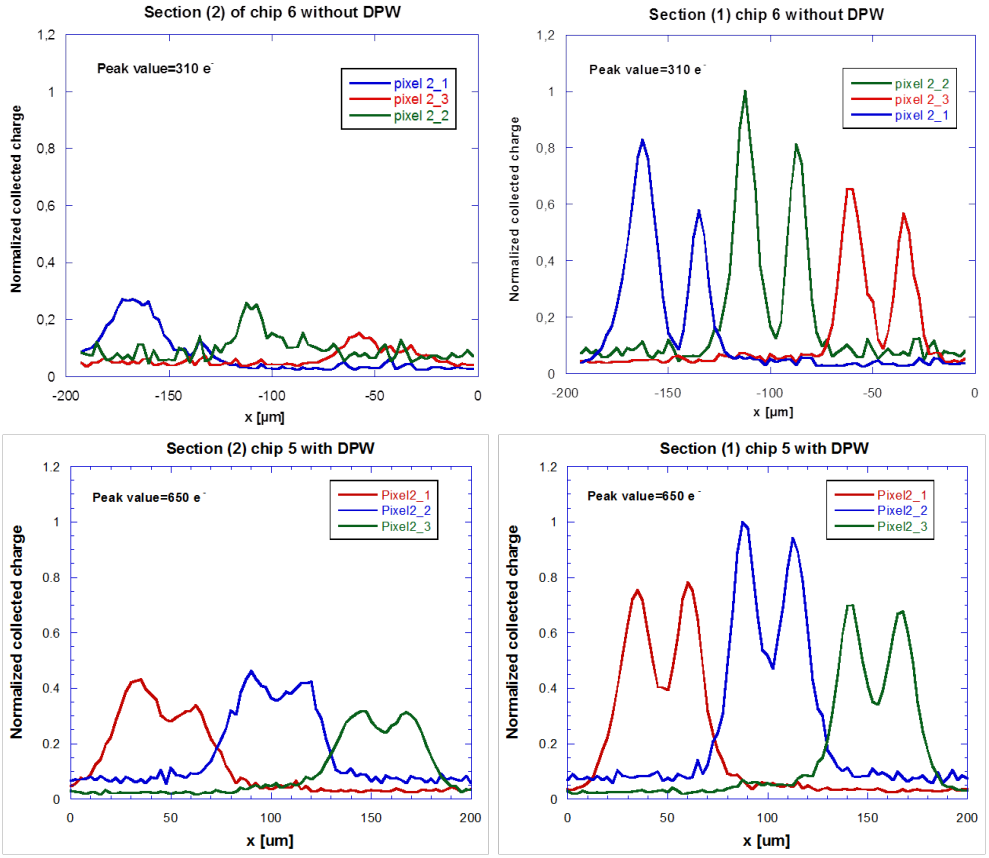

Figure 4.23. Collected charge in the 3x3 matrices of chip 6 (top) and chip 5 (bottom) along the sections shown in Fig.4.22. Plots are normalized with respect to the peak value detected in the relevant chip.

of a thicker epitaxial layer helps improve the charge collection uniformity inside the pixel. Moreover, the collected charge peak value of chip 32 is about 70% higher than that of chip 2. Further confirmation to these results comes from charge collection measurements performed by means of a $^{90}Sr/^{90}Y$ radioactive source. The measurement setup is the one described in section 2.3.2. Figure 4.25 represents the spectra obtained from measurements performed on chip 2 and chip 32. Chip 2 features a most probable value (MPV) of 45.4 mV, while the MPV measured in chip 32 is about 90% higher. This result was expected, since the amount of collected charge is supposed to increase with the epitaxial layer thickness.

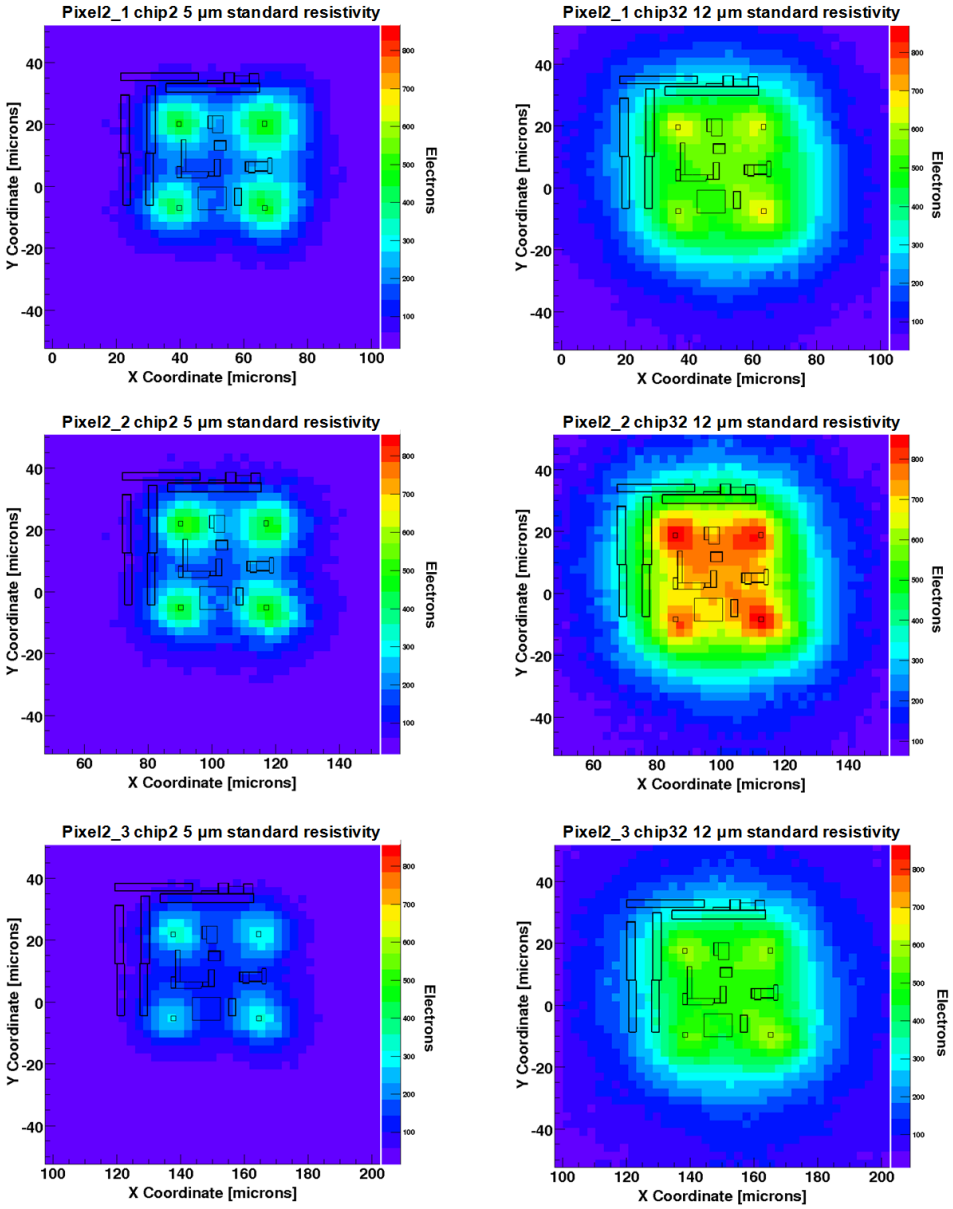

Figure 4.24. Collected charge as a function of the laser beam position in the second row of a 3x3 matrix in chip 2 (left) and chip 32 (right).

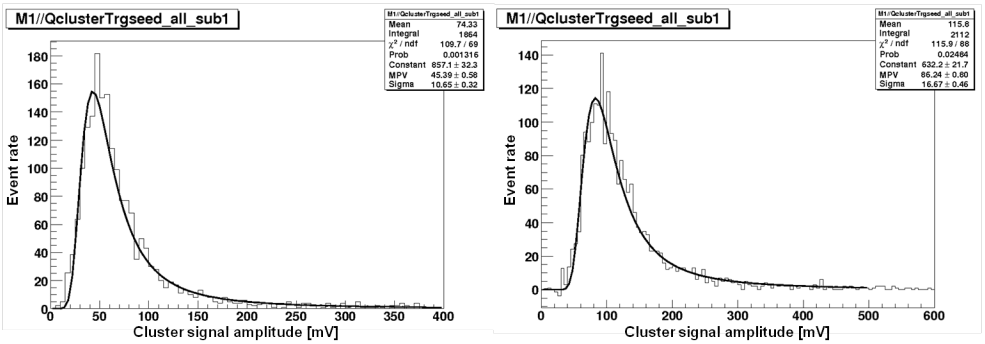

Figure 4.25. Event rate distribution in the case of chip 2 (left) and chip 32 (right) tested with a $^{90}Sr/^{90}Y$ source.

As done in the previous section, also for chip 2 and chip 32 the collected charge along the two sections of Fig. 4.22 has been extracted from the relevant laser scans (see again Fig. 4.24). Figure 4.26 shows, on the top row the charge collected by pixels in chip 2, which again appears to be more confined around the collecting electrodes (a significant decrease in the collected charge can be detected between the collecting electrodes). Moreover, the charge collected along the pixel mid-section is between 40 to 50% of the peak value of the same pixel. The central pixel of chip 2 features a collected charge maximum equal to 515 electrons. In the case of chip 32, featuring a 12 μm thick epitaxial layer, a higher charge collection uniformity can be detected, as confirmed by the plots at the bottom of Fig. 4.26. Note that the charge collection behavior along the pixel mid-section is close to that detected along the section passing through the collecting electrodes.

4.3.4 Dependence of the charge collection on the epitaxial layer resistivity

The dependence of the charge collection on the epitaxial layer resistivity was also evaluated. The study was performed by comparing the behavior of monolithic sensors featuring the same epitaxial layer thickness (12 μm), but different resistivity. For this purpose, Fig. 4.27 shows on the left the collected charge for the 3x3 matrix central pixel of chips 31 and 32 featuring a standard resistivity (10 Ω·cm) epitaxial layer. On the right, the collected charge is shown for chips 41 and 42, fabricated with a high resistivity (1 kΩ·cm) epitaxial layer. The peak value in chips 31 and 32 is found to be about 15% higher than that of

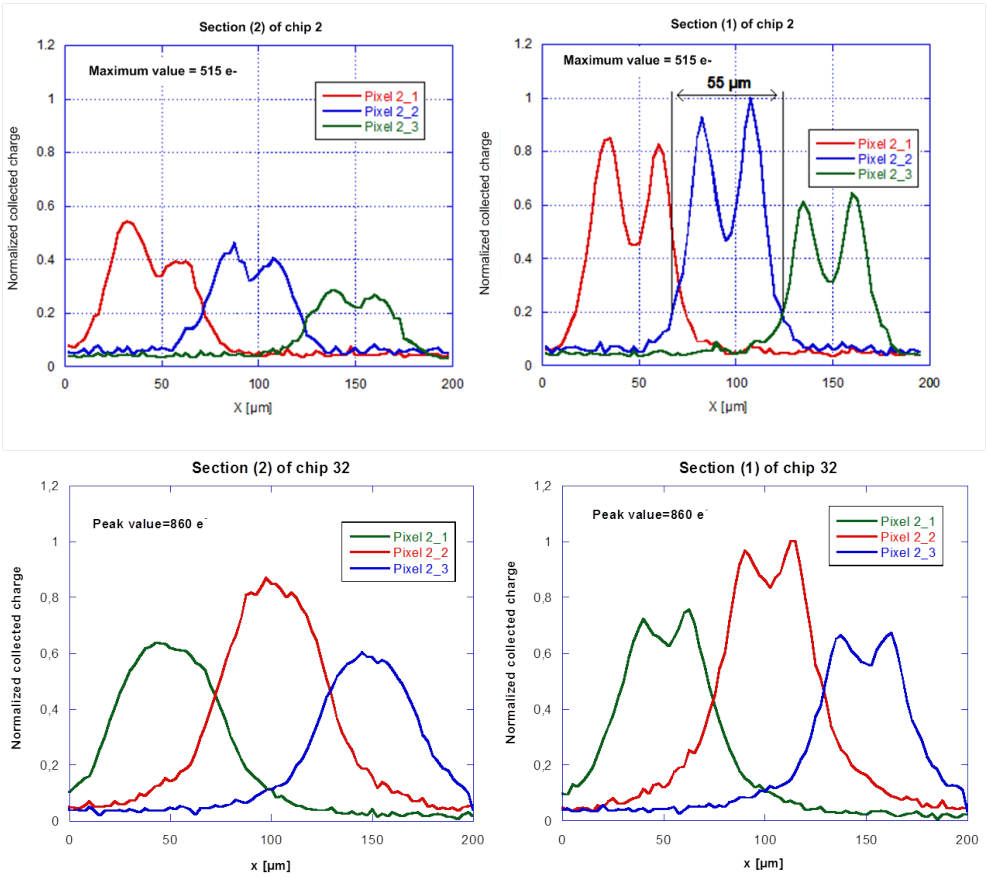

Figure 4.26. Collected charge in the 3x3 matrices of chip 2 (top) and chip 32 (bottom) along the sections shown in Fig.4.22. Plots are normalized with respect to the peak value detected in the relevant chip.

chips 41 and 42. This is an unexpected result, since the high resistivity epitaxial layer is supposed to improve charge collection performance with respect to the standard resistivity option. This latter consideration is actually confirmed by charge collection measurements performed by means of a $^{90}Sr/^{90}Y$ radioactive source. Figure 4.28 shows the spectra obtained from measurements performed on chip 32 (already shown in Fig. 4.25) and chip 42. The MPV value of the Landau distribution obtained for chip 42 is 115 mV, about 35% higher than that detected in chip 32 (85 mV).
$^{90}Sr/^{90}Y$ measurements find further support in physical device simulations performed on the 3x3 matrix structure already shown in Fig. 3.25 and repro-

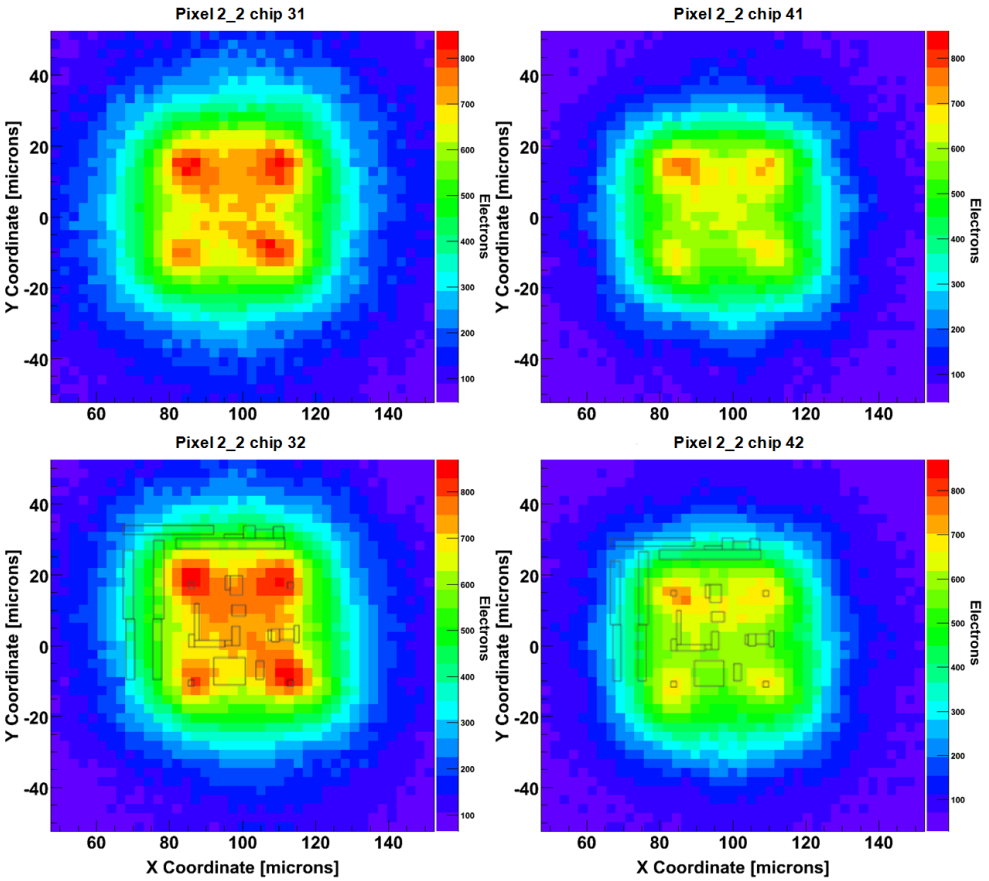

Figure 4.27. Collected charge as a function of the laser beam position in the central pixel of a 3x3 matrix in chips 31 and 32 (left) and chips 41 and 42 (right).

duced on the left side of Fig. 4.29 for sake of clarity. The same figure also shows the simulation results concerning the sum of the current signals at the matrix collecting electrodes and the collected charge as a function of time after the passage of a MIP, in the case of standard and high resistivity epitaxial layer. The maximum value of the collected charge in the case of a high resistivity epitaxial layer is about 40% higher than in the case of standard resistivity epitaxial layer. This result is in good agreement with measurements shown in Fig. 4.28. Note also, in the same simulations, the faster charge collection

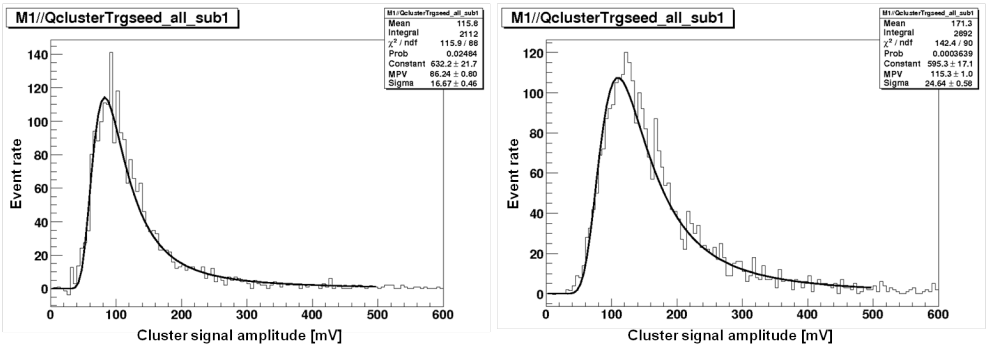

Figure 4.28. Event rate distribution in the case of chip 32 (left) and chip 42 (right) tested with a ^{90}Sr/^{90}Y source.

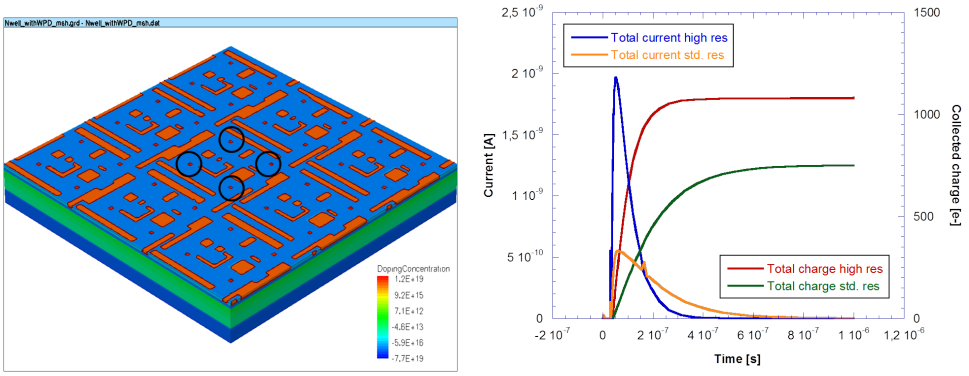

Figure 4.29. 3x3 matrix geometry and doping profile (left) and physical device simulation results concerning the charge collection performance (right) in the case of standard resistivity and high resistivity epitaxial layer. Black circles highlight the collecting electrode position in the central pixel.

in the case of the high resistivity epitaxial layer. Actually, in this latter case the depleted volume under the collecting electrodes is increased with respect to the standard resistivity case and, as a consequence, the charge collection speed is improved.

Although laser measurements do not seem to be in agreement with ^{90}Sr/^{90}Y tests, they provide some information about charge sharing. Actually, in the case of the high resistivity epitaxial layer, Fig. 4.27 seems to show that the collected charge is more confined inside the pixel. It is worth recalling here

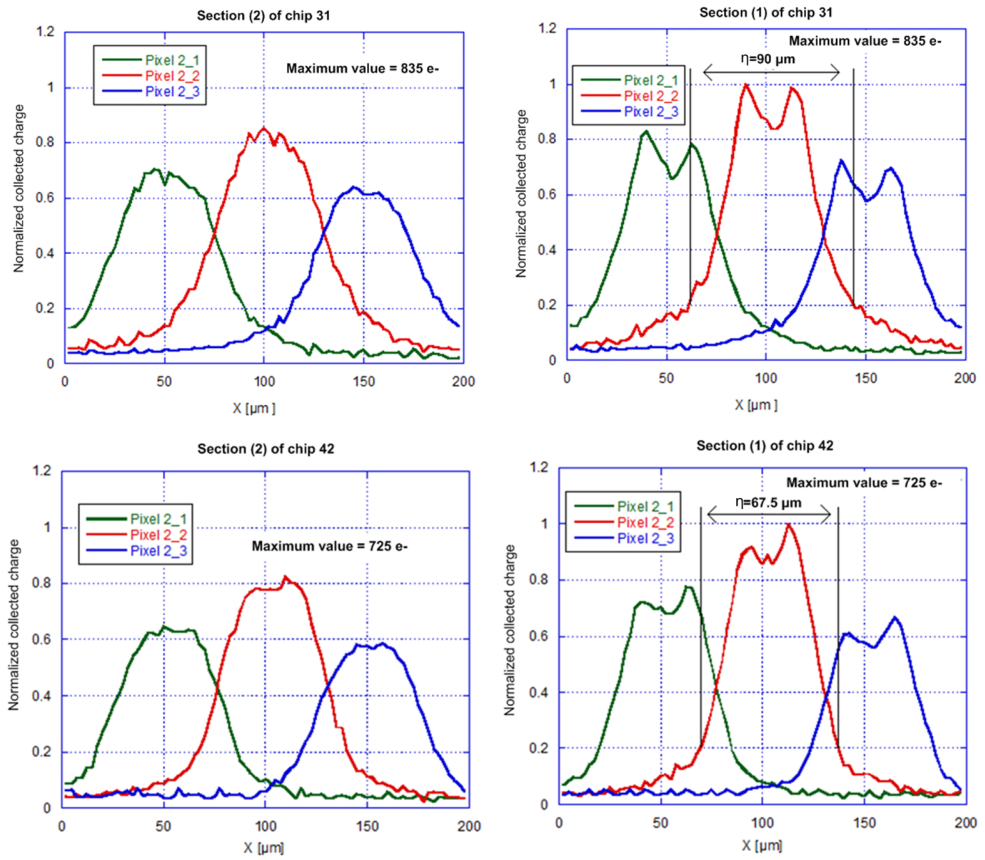

Figure 4.30. Collected charge in the 3x3 matrices of chip 31 (top) and chip 42 (bottom) along the sections shown in Fig.4.22. Plots are normalized with respect to the peak value detected in the relevant chip.

that in the case of purely binary readout, charge sharing may actually degrade the spatial resolution of the detector. The charge collection along the matrix section of Fig. 4.22 are shown for chip 31 and chip 42 in Fig. 4.30. A better charge collection uniformity can be detected in the case of chip 42. Also, in the same figures a parameter η is introduced to evaluate charge confinement inside the pixel. This parameter measures the width of the region in which the collected charge is higher than 20% of the maximum value. In the case of chip 31, η is equal to 90 μm, while in the case of chip 42, featuring a high resistivity epitaxial layer, it is reduced to 67.5 μm. Comparison between these results and physical device simulations of the 3x3 matrix central pixel is shown

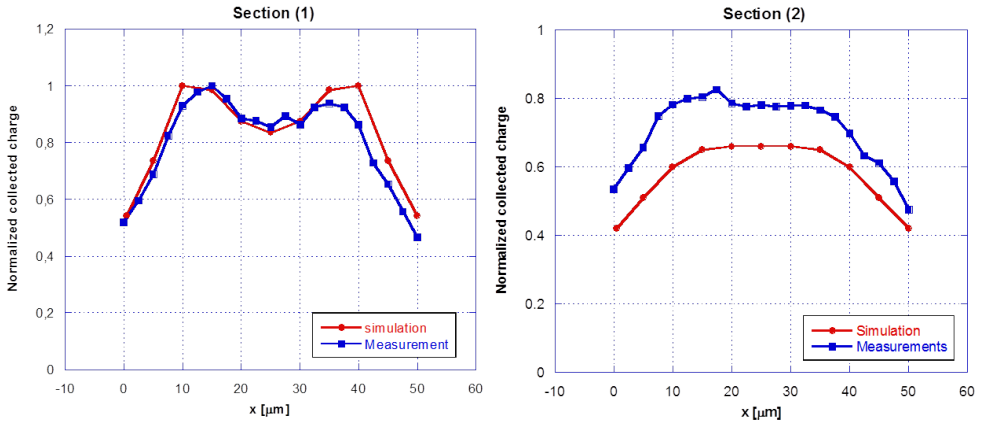

Figure 4.31. Comparison between simulation and measurement results of the charge collected in the central pixel for the two pixel sections of Fig. 4.22.

in Fig. 4.31. The plot on the left shows the charge collected along the section passing through the collecting electrodes, while the one on the right refers to the section passing in the middle of the pixel. A fairly good agreement can be detected between simulations and experimental data in the first case. In the second comparison, the simulation results seem to underestimate the charge collection, though the charge collection trend through the x coordinate is quite similar to the experimental data.

4.3.5 Conclusions on charge collection performance

The study of charge collection performance led to some conclusion summarized in the following.

- Deep P-well has beneficial effects both in terms of amount of collected charge and of charge collection uniformity.
- Adopting a 12 μm thick epitaxial layer leads to a significant improvement in charge collection performance, even in the case of standard resistivity epitaxial layer.
- The collected charge is more confined inside the pixel in the case of a high resistivity epitaxial layer, leading to a lower charge sharing with adjacent pixels. Moreover, the high resistivity epitaxial layer option further improves the charge collection uniformity. ^{90}Sr/^{90}Y measurements

DUT	Epi. thick.	Epi. res.	DPW	Pix21 max	Pix22 max	Pix23 max	Pix22 center	η [μm]
2	5 μm	std.	Yes	480 e^-	515 e^-	350 e^-	235	-
5	5 μm	std.	Yes	510 e^-	650 e^-	460 e^-	300	-
6	5 μm	std.	No	290 e^-	310 e^-	205 e^-	80	-
31	12 μm	std.	Yes	695 e^-	825 e^-	625 e^-	710	90
32	12 μm	std.	Yes	655 e^-	860 e^-	575 e^-	755	82
41	12 μm	high	Yes	505 e^-	725 e^-	470 e^-	640	70
42	12 μm	high	Yes	585 e^-	730 e^-	500 e^-	620	67

Table 4.4. Summary of the charge collection measurements performed with laser beam.

seem to point to a significant improvement also in terms of amount of collected charge.

- A good agreement was found between measurements and TCAD physical device simulations in the case of a 3x3 matrix featuring a 12 μm thick epitaxial layer, both in terms of amount of collected charge and in terms of charge collection behavior along different pixel sections.

Table 4.4 summarizes the results from laser tests. The next section will be dedicated to discuss the experimental results obtained from a non-ionizing irradiation campaign performed on the Apsel4well MAPS.

4.4 Non-ionizing irradiation campaign

As remarked throughout this work, a critical design constraint for sensors intended for application to the Layer0 of the SuperB SVT and, more in general, in HEP experiments, is the capability to withstand high levels of ionizing and non-ionizing radiation during their lifetime. This is of paramount importance in order to maintain acceptable performance, especially in terms of SNR during the experiment life. In this section, results from a non-ionizing irradiation campaign performed on a number of Apsel4well samples featuring a 12 μm thick epitaxial layer, both in the version with standard and high resistivity option, will be shown. This chips have been irradiated with neutron from a Triga Mark II nuclear reactor at the Josef Stefan Institute in Ljublijana. The maximum fluence was 10^{14} 1 MeV neutron equivalent/cm^2.

4.4.1 Analog front-end performance

This section is aimed at showing the effects of the irradiation campaign on the analog front-end performance. Figure 4.32 shows the charge sensitivity and the equivalent noise charge of Apsel4well chips featuring a 12 μm thick epitaxial layer as a function of the fluence. The figure on the left shows the results in the case of a standard resistivity epitaxial layer. Charge sensitivity has been measured through charge injection techniques, with the reference voltage V_{RIF_MIR} always set to the design value of 1.311 V for all the irradiation steps. As expected, both the charge sensitivity and the equivalent noise charge do not show any clear trend with the fluence. Actually, as already discussed in chapter 2, both the gain and the ENC depend on the characteristics of MOSFETs making up the analog readout chain. Since MOSFET operation principle is based on the drift of majority carriers in the surface region of the silicon substrate, its performance are not expected to change after irradiation with neutrons, whose damage effects are mainly confined in the device bulk. In the case of devices with high resistivity epitaxial layer, shown in the figure on the right, a charge sensitivity drop with increasing fluence can be detected , with the point corresponding to the last irradiation step about 20% smaller than in the non irradiated device. This phenomenon can be explained with an increase in the leakage current of the pixel detector, which affects the charge preamplifier gain by reducing its feedback impedance (i.e., by increasing the

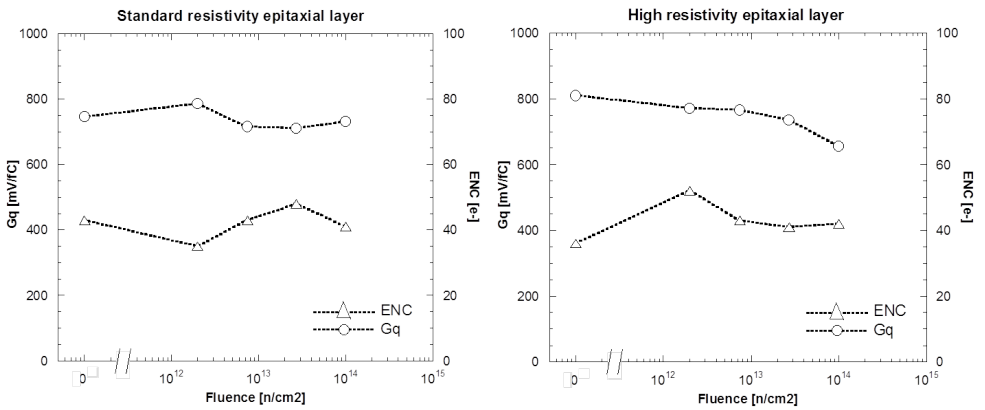

Figure 4.32. Charge sensitivity and equivalent noise charge as a function of the fluence for samples featuring standard (left) and high resistivity (right) epitaxial layer.

transconductance of its feedback transistor). The fact that this phenomenon is detectable only in devices with high resistivity epitaxial layer can be ascribed to the higher depleted volume under the collecting electrodes, which implies a larger contribution of the radiation induced generation component to the overall leakage current in the collecting electrodes. As far as the ENC is concerned, like in the case of standard resistivity epitaxial layer, no particular trend with the fluence can be clearly detected. Fluctuations in the ENC are likely to be related to process variations, as the points in Fig. 4.32 were obtained from the characterization of different samples irradiated at different fluences.

4.4.2 Charge collection performance

The effects of neutron irradiation on the charge collection properties of the Apsel4well MAPS have been evaluated by means of measurements with infrared laser and $^{90}Sr/^{90}Y$ radioactive source. Figure 4.33 and 4.34 show the collected charge as a function of the laser beam position in the central pixel of 3x3 matrices in chips fabricated with standard and high resistivity epitaxial layer. As already discussed in the previous section, the pre-irradiation maximum value of the chip featuring a standard resistivity epitaxial layer seems to be higher than its high resistivity counterpart. However, after irradiation, the decrease in the charge collection is significantly larger in the case of the samples with standard resistivity epitaxial layer. Figure 4.35 shows the collected charge normalized to the pre-irradiation value as a function of the fluence in samples illuminated with a laser beam. In the figure, devices with different epitaxial layer resistivity are compared. In the case of the standard resistivity epitaxial layer, the peak collected charge is 25% of the pre-irradiation value, while it is 50% in the case of the high resistivity counterpart. The high resistivity epitaxial layer seems to guarantee a higher degree of radiation hardness. This is actually in agreement with the scientific literature on the subject [32]. Results concerning laser characterization of the samples with high resistivity epitaxial layer are in good agreement with those coming from $^{90}Sr/^{90}Y$ measurements shown in Fig. 4.36. On the left, the figure shows the event rate distributions of Apsel4well samples with high resistivity epitaxial layer, one non-irradiated, the other two irradiated with neutrons at two different fluences. In the same figure, on the right, the normalized MPV as a function of the fluence is shown. The detected decrease in collected charge is in fairly good agreement with laser measurement data.

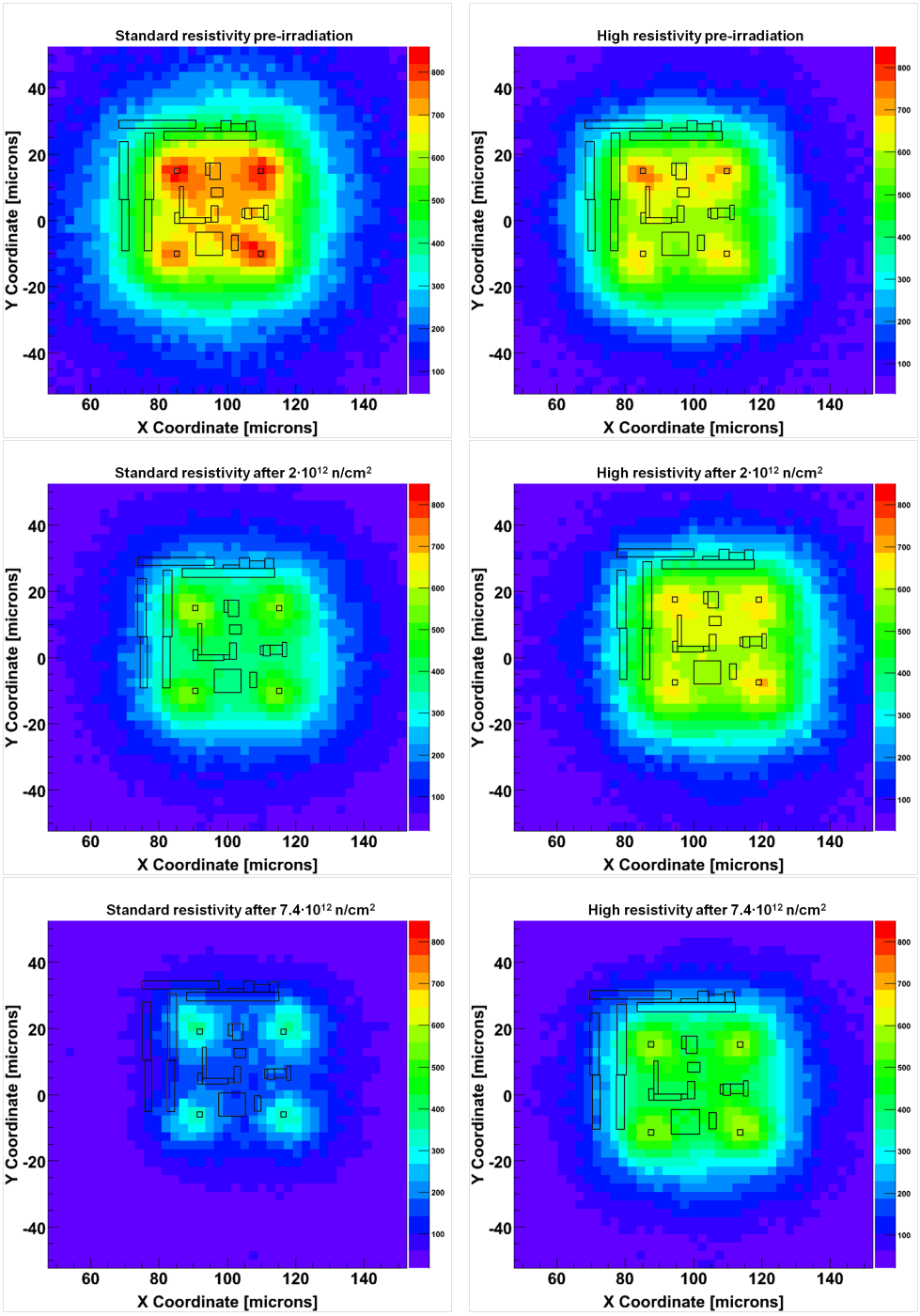

Figure 4.33. Collected charge as a function of the laser beam position in the central pixel of 3x3 matrices at different fluence values for samples fabricated with standard (left) and high resistivity (right) epitaxial layer (pre-irradiation, first and second irradiation steps).

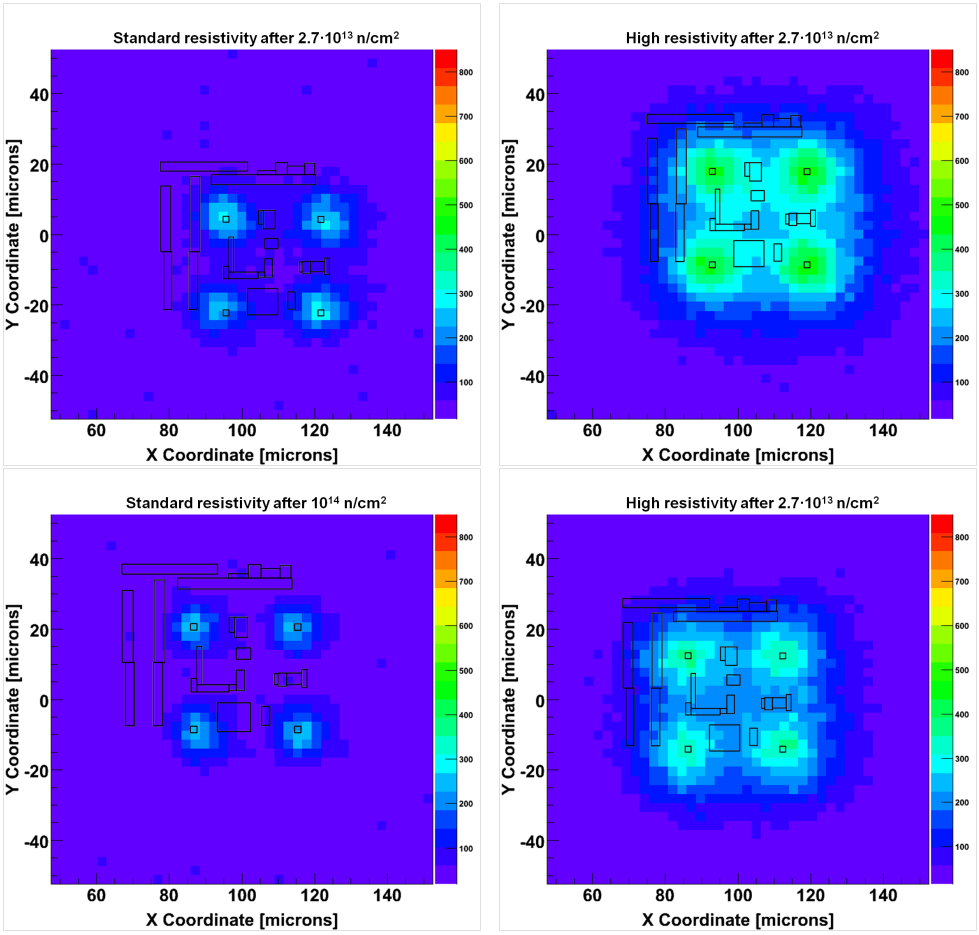

Figure 4.34. Collected charge as a function of the laser beam position in the central pixel of 3x3 matrices at different fluence values for samples fabricated with standard (left) and high resistivity (right) epitaxial layer (third and fourth first steps).

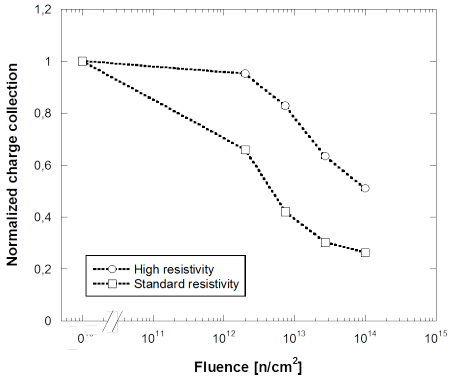

Figure 4.35. Peak collected charge normalized to the pre-irradiation value as a function of the fluence.

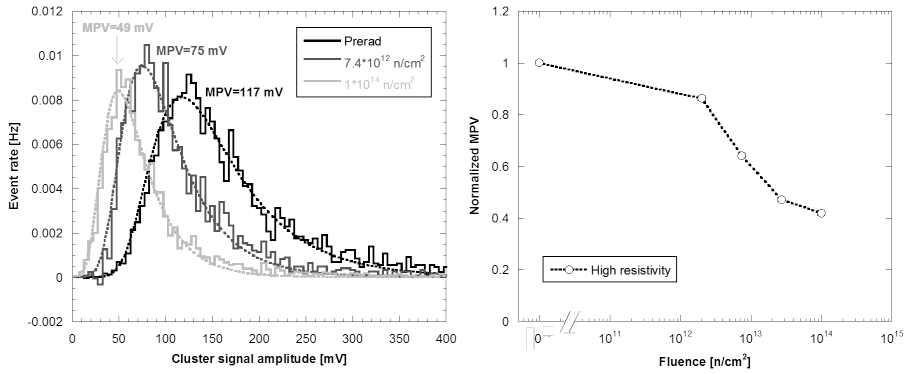

Figure 4.36. Event rate distribution (left) and normalized MPV (right) at varying neutron fluence for Apsel4well samples with high resistivity epitaxial layer.

4.4.3 A comparison between Apsel3T1 and Apsel4well performance

In order to complete this section, a comparison between charge collection properties of Apsel4well MAPS with high resistivity epitaxial layer and the Apsel3T1 monolithic sensor introduced in chapter 2 has been performed. Figure 4.37 shows the normalized MPV of the ^{90}Sr/^{90}Y spectra as a function of the time both for Apsel4well and Apsel3T1 chips. In this last case, no dis-

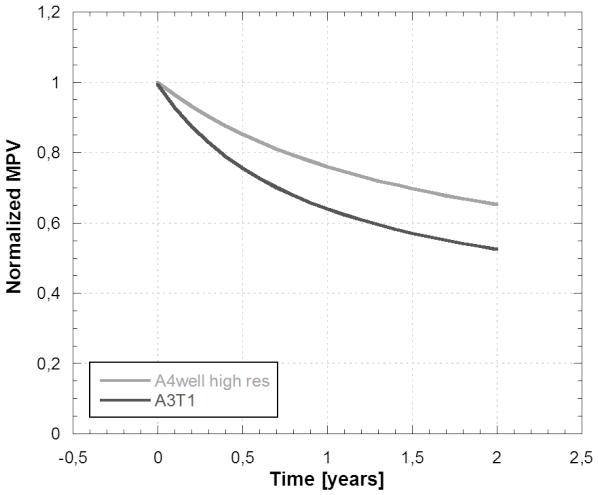

Figure 4.37. Normalized MPV as a function of the time of operation in the Layer0. A comparison between Apsel4well and Apsel3T1 (A3T1) sensorsis shown.

tinciton have been made between the M1 and M2 structures, since, as shown in Fig. 2.30, no significant differences have been detected in the two cases. In each curve, the MPV has been normalized with respect to the relevant pre-irradiation value. As already discussed in section 2.4 for the irradiation campaign performed on the Apsel3T1 sensor, the correspondence between time and fluence is based on the SVT Layer0 specifications, and on the assumption that fluence grows linearly with time. From this plot, the higher degree of hardness to bulk damage of Apsel4well with respect to Apsel3T1 is apparent, especially with increasing time interval (i.e. integrated fluence). After two years of operation in the Layer0, the charge collection drop in the case of the Apsel3T1 chip is about 50%. This drop reduces to about 35% in the case of Apsel4well with high resistivity epitaxial layer.

4.5 Charge loss modeling in irradiated MAPS

Both to improve the understanding of the mechanisms underlying charge loss and to provide a tool for predicting performance degradation in neutron irradiated DNW MAPS and quadruple well MAPS, an already existing Monte Carlo (MC) model, developed for physical simulation of monolithic sensors

[62], has been expanded to account for bulk damage effects. The model benefits from the fact that, in most MAPS devices, doping concentrations and voltages are unsuitable for full depletion of the sensor substrate and diffusion remains as the main transport mechanism for the minority carrier released by an impinging particle. Therefore, a Monte Carlo algorithm, simulating carrier motion through random walk, is expected to be accurate enough to obtain information about the charge collection properties of a monolithic pixel sensor. With respect to TCAD sofware packages like the one provided by Synopsys, typically based on finite element methods, MC tools can capitalize on their much shorter computational times.
In the considered MC model, the motion of each minority carrier released in the substrate is represented by means of random spatial steps: for every time interval Δt, each carrier moves by a step with length Δl and direction uniformly distributed over the full solid angle of 4π steradians. The generic random step $\vec{r_q}$ for the single carrier can be expressed as

$$\vec{r_q} = \vec{i}x_q + \vec{j}y_q + \vec{k}z_q, \tag{4.4}$$

where $\vec{i}$, $\vec{j}$ and $\vec{k}$ are the unit vectors of a Cartesian coordinate system and the point $\vec{r_q}$ is uniformly distributed over the surface of a sphere of radius $\Delta l = \sqrt{x_q^2 + y_q^2 + z_q^2}$. After N steps, the distance d_{rw} from the starting point is

$$d_{rw} = \left| \sum_{q=1}^{N} \vec{r_q} \right| \tag{4.5}$$

and its mean square value is

$$\begin{aligned} < d_{rw}^2 > &= N \cdot (< x_q^2 > + < y_q^2 > + < z_q^2 >) = \\ &= N \cdot \Delta l^2, \end{aligned} \tag{4.6}$$

with $N = \frac{t}{\Delta t}$, t being the time elapsed from carrier release. Moreover, minority carriers have to obey the diffusion equation (no interaction between particles is assumed here), which, in three-dimensional space, can be expressed as:

$$\frac{\partial n(x,y,z,t)}{\partial t} = D\nabla^2 n(x,y,z,t), \tag{4.7}$$

with *n(x,y,z,t)* the carrier concentration as a function of the three spatial coordinates and of the time t, and D the diffusion constant. Therefore,

$$\int_{R^3} n(x,y,z,t)dxdydz = N_{tot}, \tag{4.8}$$

where N_{tot} is the total number of released carriers. If carriers are assumed to diffuse from $(x,y,z)=(0,0,0)$, then a displacement d_d can be defined as

$$d_d = \sqrt{x^2+y^2+z^2}. \tag{4.9}$$

From (4.7) and (4.8), it can be demonstrated that

$$< d_d > = \frac{1}{N_{tot}} \int_{R^3} d_d \cdot n(x,y,z,t)dxdydz = 0, \tag{4.10}$$

$$< d_d^2 > = \frac{1}{N_{tot}} \int_{R^3} d_d^2 \cdot n(x,y,z,t)dxdydz = 6Dt. \tag{4.11}$$

Now, by equating (4.6) and (4.11), a relationship is found between the elementary step Δl and the associated time lapse Δt,

$$\Delta t = \frac{\Delta l^2}{6D}, \tag{4.12}$$

for the MC model to comply with the constraints set by the diffusion equation. Δl is chosen to be much smaller than the dimensions of the simulated structure. The model also has to account for the fact that, as already discussed in chapter 2, in a doped semiconductor, minority carriers tend to recombine with a characteristic time (i.e., the carrier lifetime) τ_l. As a consequence, the total number of excess charge carriers N_{exc} resulting from a particle crossing the sensor at t=0, decays in time according to the following equation,

$$N_{exc}(t) = N_{exc}(0) \cdot \exp\left(-\frac{t}{\tau_l}\right), \tag{4.13}$$

where $N_{exc}(0) = N_{tot}$ and the lifetime expression,

$$\tau_l(N_{dop}) = f(N_{dop}) \cdot \tau_0, \tag{4.14}$$

also takes into account the influence of the substrate doping concentration N_{dop} on carrier recombination by means of the empirical relationship [63]

$$f(N_{dop}) = \frac{1}{1+\frac{N_{dop}}{N_{ref}}}. \tag{4.15}$$

In 4.14, τ_0 defaults to 10 μs in the case of electrons in P-doped silicon. In (4.15), $N_{ref} \simeq 10^{16}$ cm^{-3}. As already discussed in chapter 2, under neutron irradiation,

the reciprocal of minority carrier lifetime in silicon increases proportionally to the fluence up to neutron levels in the order of at least 10^{15} cm^2. To account for bulk damage effects, the lifetime has to be expressed as [32]

$$\tau_l(N_{dop}, \Phi) = \left[\frac{1}{f(N_{dop}) \cdot \tau_0} + \frac{\Phi}{K(N_{dop})} \right]^{-1}, \tag{4.16}$$

where Φ is the fluence and K is a silicon damage constant, depending on the doping concentration in the substrate. Actually, doping concentration (as well as injection levels) plays a role in the effectiveness of neutron-induced defects as recombination centers. Equation 4.13, which holds in the absence of a current flow and in low injection level conditions [64], is accounted for in the MC model by assigning to each single released carrier a specific time of life T_i,

$$T_i = \tau_l \ln \left[\frac{N_{tot}}{i} \right], \quad i = 1, \ldots, N_{tot}. \tag{4.17}$$

In the case of Apsel4well chips with high resistivity epitaxial layer, diffusion cannot be considered the only charge transport mechanism. Actually the depleted volume around the collecting electrode and its effect on the charge collection properties of the device can no longer be neglected. To account for these effects in the MC model, the assumption is made that, when a random walking carrier enters the depleted volume under an N-well, it is istantaneously collected by the same N-well. This is a reasonable approximation, since the transit time of a carrier in the space charge region of a PN junction is much shorter than the time involved in the diffusion process, and the probability of its recombining in the same region is negligible. The width of the depletion region (which will mainly extend into the more lightly doped epitaxial layer), to be included in the device has been obtained from TCAD simulations of the structure shown in Fig. 3.21 modeling the collecting electrode configuration in Apsel4well MAPS. The results of this simulations show a maximum depth of the space charge region of about 7 μm in the case of an inverse voltage of 0.5 V applied to the collecting electrode in a high resistivity epitaxial layer. Figure 4.38 shows the space charge density around the collecting electrode in the case of standard (left) and high resistivity (right) epitaxial layer.

4.5.1 Monte Carlo model validation

This subsection will discuss the comparison between MC simulations and the experimental data obtained by means of ^{90}Sr and laser stimulation both on

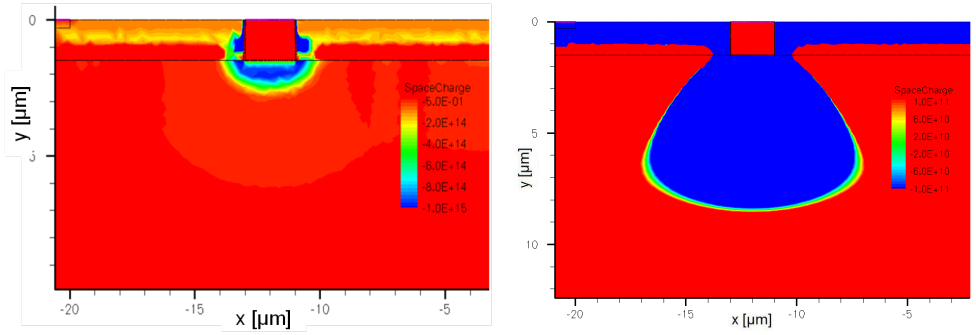

Figure 4.38. Space charge density around the Apsel4well collecting electrode for the standard (left) and high resistivity (right) epitaxial layer.

Apsel3T1 chips (discussed in chapter 2) and on Apsel4well samples with a 12 μm thick epitaxial layer. ^{90}Sr measurements have been performed by means of the measurement setup already shown in section 2.3.2. This setup (with the exception of the scintillator and the block for gate signal generation) was reproduced in the Monte Carlo device simulations, including the model for the MAPS DUTs based on the N-well and deep N-well geometries and taking into account the threshold used to trigger the waveform capture in the acquisition system. The left plot of Fig. 4.39 shows the MPV of the ^{90}Sr spectra normalized to the pre-irradiation value as a function of the fluence for the M1 and M2 matrices of the Apsel3T1 chip. The experimental data, averaged over the set of tested devices, are compared to the MC simulation results. Horizontal error bars account for the uncertainty in the neutron fluence, about $\pm$10% of the fluence level. The figure inset shows the spectra obtained for one of the samples (matrix M1 type) before irradiation and after exposure to a fluence of 6.7×10^{12} cm^{-2} (already shown in chapter 2). In the MC simulations, a value of 3×10^{5} cm^{-2}·s was used for the damage constant K, as provided by the literature data for electrons in a P-type substrate with a 10 Ω·cm resistivity operated in low injection conditions [32]. Simulation and measurement results are in fairly good agreement, with the model replicating, in the case of the last irradiation step (corresponding to a fluence of 6.7×10^{12} cm^{-2}), the slightly different radiation response of the two kinds of sensors. Figure 4.39 on the right shows the MPV of the ^{90}Sr spectrum as a function of the fluence normalized to the pre-irradiation value for an Apsel4well 3x3 matrix (the same as in Fig. 4.36). The experimental data are again compared to the results from MC simulations, where a damage constant K equal to 2×10^{6} cm^{-2}·s was used, as found in the literature in the case of a P-doped substrate with a 1 kΩ·cm

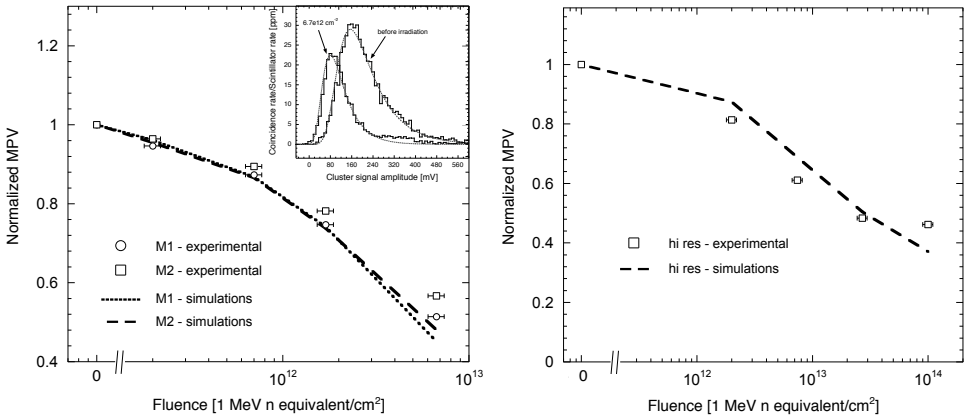

Figure 4.39. MPV of the ^{90}Sr spectra normalized to the pre-irradiation value as a function of the fluence for the Apsel3T1 MAPS (left) and Apsel4well samples with high resistivity epitaxial layer (right). Experimental data are compared to MC simulation results.

resistivity. The agreement between simulation and experimental data is again quite good. As a general consideration, the consistency between the model and the experimental data, both in this latter case and in the previous one, seems to confirm that radiation induced increase in the recombination rate is the predominant source of charge collection degradation in irradiated MAPS and can by itself explain the decrease in the amount of collected charge with increasing neutron fluence.

The left plot of Fig. 4.40 shows the peak value of the collected charge, normalized to the pre-irradiation value, as a function of the fluence in the central pixels of the M1 and M2 matrices of the Apsel3T1 chip tested with a laser source. The vertical cross bars account for the inaccuracies in charge sensitivity and peak value measurements used to determine the points in the plot. Measurement results are compared to the outcome of the MC simulations, where the DUT geometry and substrate characteristics and mechanism of charge release of the laser beam in silicon have been implemented. In the case of the elementary cell of the M1 kind, the agreement between the experiment and the simulation results is quite good at all the considered fluences. In the case of the M2 matrix, the effect of bulk damage seems to be underestimated by the simulations at the intermediate fluence steps, between 10^{12} and 10^{13} cm^{-2}. The right plot of Fig. 4.40 shows the normalized peak value of the collected charge as a function of the fluence for the central pixel of an Apsel4well matrix (also illuminated with a laser beam). Experimental data (already shown in Fig.

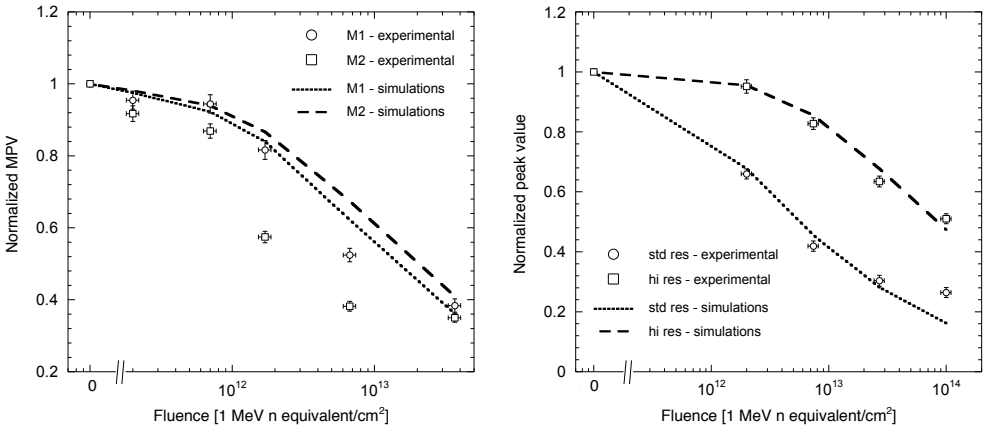

Figure 4.40. Normalized peak value of the collected charge as a function of the fluence for the Apsel3T1 MAPS (left) and Apsel4well samples (right) tested with a laser source. Experimental data are compared to MC simulation results.

4.35) are now compared to the MC simulation results, performed again using K equal to 3×10^5 cm^{-2} ·s for the devices with standard resistivity epitaxial layer and K equal to 2×10^6cm^{-2}·s for the high resistivity case. As a general consideration, the MC simulation is capable of reproducing, with a fairly good approximation, the dependence of the device charge collection properties on the neutron fluence and on the epitaxial layer doping concentration. The simulation results are found to be in good agreement with the experiment over the considered fluence interval, for both the standard and the high resistivity cases, with some overestimation of the bulk damage for data point at the largest fluence in the device with standard resistivity epitaxial layer.

Conclusion

In this thesis work, the design and the experimental characterization of a CMOS monolithic active pixel sensor (MAPS) developed for application to the Layer0 of the SuperB silicon vertex tracker (SVT) have been discussed. In the first chapter the main specifications for the SuperB SVT have been introduced. In the second chapter, the results from the irradiation campaigns performed on Apsel3T1 deep N-well MAPS intended for the same application have been illustrated and thoroughly analyzed. Such tests were peformed with the aim of studying the tolerance of the Apsel3T1 MAPS both to ionizing radiations and neutrons. Measurement techniques used for the chip characterization include tests with benchtop instrumentation, as well as the use of an infrared laser and ^{55}Fe and ^{90}Sr/^{90}Y radioactive sources. Measurement results showed an increase of the equivalent noise charge in excess of 100% of the pre-irradiation value for a total ionizing dose of about 10 Mrad(SiO_2). The irradiation of Apsel3T1 samples with neutrons from a nuclear reactor with fluences up to 3.7×10^{13} 1 MeV n_{eq}/cm^2 led to a decrease in the collected charge of about 60% of the pre-irradiation value. The combination of the mentioned effects leads to a signal-to-noise ratio degradation not suitable for the considered applications.
In the third chapter the new MAPS prototype designed in the 180 nm CMOS INMAPS process, called Apsel4well, has been introduced. The proposed approach, featuring the use of a deep P-well implant, enables the integration of an in-pixel readout channel with data sparsification and time stamping capabilities. Moreover, in this technology, an epitaxial layer is made available with different thickness (5 and 12 μm) and resistivity (10 Ω·cm or 1 kΩ·cm)

options. By taking advantage of these features, Apsel4well chips with different epitaxial layer characteristics have been designed and fabricated. Among the various solutions, the one featuring a 12 μm thick high resistivity epitaxial layer was found to provide the best performance in terms of charge collection and bulk damage tolerance, especially as compared to the Apsel3T1 prototype. The design of the Apsel4well front-end channel was carried out paying particular attention to its noise and charge collection properties, also with the help of TCAD simulation tools, which made it possible to optimize the collecting electrode layout. Rad-hard design techniques, like enclosed layout of transistors, have been adopted to improve the tolerance of the analog front-end to ionizing radiation. Results obtained by means of benchtop instrumentation were found to be in fair agreement with circuit simulations in one of the four model corner with an ENC of 37 electrons and a charge sensitivity of between 900 and 1000 mV/fC. Charge collection properties of the Apsel4well MAPS were investigated by means of an infrared laser source. The results proved the beneficial effects in terms of charge collection properties provided by the deep P-well layer. The measured charge collection inside the pixel was found to be in good agreement with TCAD simulation results. Charge collection was also characterized by means of a ^{90}Sr/^{90}Y source, providing a most probable value of the collected charge signal around 800 electrons.
Finally, an irradiation campaign with neutrons on Apsel4well samples with 12 μm thick standard and high resistivity epitaxial layer with fluences up to 10^{14} 1 MeV n_{eq}/cm^2 was performed. Devices with high resistivity epitaxial layer were proven to be significantly more radiation tolerant than their standard resistivity counterparts and the Apsel3T1 MAPS. Results from the irradiation campaigns performed on the Apsel4well and the Apsel3T1 chips were found to be quite in good agreement with results from a Monte Carlo physical simulation tool modeling the charge loss after irradiation with neutrons in CMOS MAPS structures. This tool will be used in the design of the next generation of monolithic pixel sensors in the CMOS INMAPS process in order to further increase the tolerance to bulk damage.

Bibliography

[1] G. Polesello, "E allora fabbrichiamo..."
http://www.asimmetrie.it/materiali/8.pdf

[2] M. Giorgi, "Una fabbrica di fascino e bellezza"
http://www.asimmetrie.it/modules/smartsection/item.php?
idnumero=83&itemid=150&categoryid=84&keywords=superb

[3] P. Raimondi, M. Zobov, D. Shatilov, "Beam-beam simulations for particle factories with crabbed waist", 2007 IEEE *Particle Accelerator Conference Records (PAC)*, 25-29 June 2007, pp. 1469 - 1471.

[4] "PEP-II Status report," PAC 2005 Knoxville, TN.

[5] "KEKB Status report," PAC 2005 Knoxville, TN.

[6] "ILC Reference Design Report"
http://www.linearcollider.org/about/Publications
/Reference-Design-Report

[7] "SuperB Progress Reports," June 30, 2010.
http://arxiv.org/abs/1007.4241v1

[8] M. Bona et al., "SuperB: A High-Luminosity Heavy Flavour Factory. Conceptual Design Report"
http://arxiv.org/abs/0709.0451v2

[9] F. Bosi, "Micro-channel cooling for pixel detectors" Poster presented at Workshop on Intelligent Trackers 2012, April 3^{rd}-5^{th}, 2012, Pisa (Italy).

[10] L. Hubbeling et al., "Measurements of spatial resolution of a double-sided AC-coupled microstrip detector", Nucl. Instrum. Meth., sec. A, vol. 310, pp.197-202, Dec. 1991.

[11] V. Re et al., "FSSR2, a Self-Triggered Low Noise Readout Chip for Silicon Strip Detectors", IEEE Trans. Nucl. Sci., vol. 53, pp. 2470-2476, 2006.

[12] L. Andricek et al., "Single-sided p^+n and double-sided silicon strip detectors exposed to fluences up to $2\times10^{14}/cm^2$ 24 GeV protons", Nucl. Instrum. Meth., sec. A, vol. 409, pp. 184-193, May 1998.

[13] "Handbook of 3D Integration: Technology and Application of 3D Integrated Circuits", edited by P. Garrou, C. Bower, P. Ramm, WILEY-VCH, Weinheim, 2008

[14] G. Traversi, "Charge Signal Processor in a 130 nm CMOS Technology for the Sparse Readout of Small Pitch Monolithic and Hybrid Pixel Sensors, IEEE Trans. Nucl. Sci., vol. 58, no. 5, pp. 3949-3955, Oct. 2011.

[15] G. Casarosa, S. Zucca et al. "Thin pixel development for the Layer0 of the SuperB Silicon Vertex Tracker" 2010 IEEE Nuclear Science Symposyum Conference Records (NSS/MIC), pp.1901-1905.

[16] E.R. Fossum,"CMOS Image Sensors: Electronic Camera On-A-Chip, IEEE Trans. El. Dev., vol. 44, no. 10, pp. 1689-1698, Oct. 1997.

[17] Y. Degerli, M. Besanon, A. Besson, G. Claus, G. Deptuch, W. Dulinski, et al., "Performance of a Fast Binary Readout CMOS Active Pixel Sensor Chip Designed for Charged Particle Detection, IEEE Trans. Nucl. Sci., vol. 53, no. 6, pp. 3949-3955, Dec. 2006.

[18] M. Barbero, G. Varner, A. Bozek, T. Browder, F. Fang, M. Hazumi et al., "Development of a B-Factory Monolithic Active Pixel Detector -The Continuous-Acquisition Pixel Prototypes, IEEE Trans. Nucl. Sci., vol. 52, no. 4, pp. 1187-1191, Aug. 2005.

[19] D. Contarato, J.-M. Bussat, P. Denes, L. Greiner, T. Kim, T. Stezelberger et al., "CMOS monolithic pixel sensors research and development at LBNL, PRANAMA Journal of Physics, vol. 69, no. 6, pp. 963-967, Dec. 2007.

[20] 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, no.4, pp.1613-1617, Aug. 2004

[21] R. Turchetta et al., "CMOS Monolithic Active Pixel Sensor (MAPS): Developments and Future Outlook", Nuclear Instruments & Methods in Physics Research A, vol. 582, no. 3, pp. 866-870, December 2007.

[22] R.S. Muller, T.I. Kamins, *Device electronics for integrated circuits*, John Wiley & Sons, New York, 1986.

[23] S. J. Koester et al., "Wafer-Level 3D Integration Technology", IBM J. Res. & Dev., vol. 52, no. 6, Nov. 2008.

[24] J. U. Knickerbocker et al., "Three-dimensional silicon integration", IBM J. Res. & Dev., vol. 52, no. 6, Nov. 2008.

[25] R.S. Patti, "Three-Dimensional Integrated Circuits and the Future of System-on-Chip Design, Proceedings of the IEEE, vol. 94, No. 6, June 2006.

[26] G. Rizzo, "Recent developments on CMOS MAPS for the SuperB Silicon Vertex Tracker, Oral presentation at the 12^{th} Pisa Meeting on Advanced Detectors, May 20-26, 2012, La Biodola (Italy).

[27] A. Holmes-Siedle, L. Adams, "Handbook of radiation effects, second edition, edited by Oxford University Press (2002)

[28] D. L. Griscom, D. B. Brown, and N. S. Saks, "The Physics and Chemistry of SiO2 and Si-SiO_2 Interface, edited by C. R. Helms and B. E. Deal (Plenum, New York, 1988).

[29] V. Re et al., "Total ionizing dose effects on the noise performance of a 0.13 μm CMOS technology", IEEE Trans. Nucl. Sci., vol. 53, no. 6, pp. 1599-1606, Jun. 2006.

[30] S. Wood et al., "Simulation of Radiation Damage in Solids, IEEE Trans. Nucl. Sci., 28(6), pp. 4107-4122 (1981).

[31] O.L. Curtis et al., "Effect of Irradiation on the Hole Lifetime of Germanium, J. Appl. Phys. 28, pp. 1161-1165, Oct. 1957.

[32] G.C. Messenger, "A Two-Level Model for Lifetime Reduction Processes in Neutron-Irradiated Silicon", IEEE Trans. Nucl. Sci., 14(6) pp. 88-102 (1967).

[33] ESCC radiation test methods and guidelines: Total dose steady-state irradiation test method. [Online]. Available: https://escies.org/ReadArticle?docId=229.

[34] F. Faccio and G. Cervelli, "Radiation-induced edge effects in deep submicron CMOS transistors", IEEE Trans. Nucl. Sci., vol. 52, no. 6, pp. 2413-2420, Dec. 2005.

[35] J. Bogaerts et al., "Total dose and displacement damage effects in a radiation-hardened CMOS APS", IEEE Trans. Electron Devices, vol. 50, no. 1, pp. 84-90, Jan. 2003.

[36] E. Eid et al., "Design and characterization of ionizing radiation-tolerant CMOS APS image sensors up to 30 Mrd (Si) total dose", IEEE Trans. Nucl. Sci., vol. 48, no. 6, pp. 1796-1806, Dec. 2001.

[37] E. Gatti and P.F. Manfredi, "Processing the signals from solid-state detectors in elementary particle physics", La Rivista del Nuovo Cimento, vol. 9, pp. 1147, 1986.

[38] L. Ratti et al., "Minimum Noise Design of Charge Amplifiers with CMOS Processes in the 100 nm Gate Length Regime", IEEE Trans. Nucl. Sci., vol. XX, no. X, pp.235-242, Feb. 2009.

[39] D. M. Fleetwood et al., "Effects of oxide traps, interface traps, and border traps on metal-oxide-semiconductor devices", J. Appl. Phys., vol. 73, no. 10, pp. 50585074, May 1993.

[40] V. Re et al., "Comprehensive study of total ionizing dose damage mechanisms and their effects on noise sources in a 90 nm CMOS technology", IEEE Trans. Nucl. Sci., vol. 55, pp. 32723279, Dec. 2008.

[41] V. Re et al., "Impact of lateral isolation oxides on radiation-induced noise degradation in CMOS technologies in the 100-nm regime", IEEE Trans. Nucl. Sci., vol. 54, pp.2218-2226, Dec. 2007.

[42] Y. Tsividis, "Operation and modeling of the MOS transistor", Edited by Oxford University Press Inc., Sep. 2003.

[43] T.P. Ma, P.V. Dressendrofer, "Ionizing radiation effects in MOS devices and Circuits", John Wiley and Sons, New York, 1989.

[44] G.C. Messenger, "A summary review of displacement damage from high energy radiation in silicon semiconductors and semiconductor devices", IEEE Trans. Nucl. Sci., vol. 39, no. 3, pp. 468-473, 1992

[45] G. Lindsrtöm et al., "Radiation hard silicon detectors - developments by the RD48 (Rose) collaboration", Nucl. Instrum. Methods, vol. A466, pp. 308-326, 2001.

[46] V. Cindro et al., "Radiation damage in p-type silicon irradiated with neutrons and protons", Nucl. Instrum. Methods, vol. A599, pp. 60-65, 2009.

[47] J.R. Srour et al., "Review of displacement damage effects in silicon devices", IEEE Trans. Nucl. Sci., vol. 50, no. 4, pp. 653-670, Jun. 2003

[48] Y. Murakami and T. Shingyouji, "Separation and analysis of diffusion and generation components of pn junction leakage current in various silicon wafers", J. Appl. Phys. 75, 3548 (1994); doi: 10.1063/1.356091

[49] E. Simoen et al., "Accurate extraction of the diffusion current in silicon p-n junction diodes", Appl. Phys. Lerr., vol. 72, pp. 1054-1056, 1998

[50] M.J.J. Theunissen et al., "Analysis of the soft reverse characteristic of n+p drain diodes", Solid-st. Electron., vol. 28, pp. 417-425, 1985

[51] W. Maes et al., "Impact ionization in silicon: A review and update", Solid-st. Electron., vol. 33, pp. 705-718, 1990

[52] L. Ratti et al., "TID Effects in Deep N-well CMOS Monolithic Active Pixel Sensors ", IEEE Trans. Nucl. Sci., vol. 56, no. 4, pp. 2124-2131, Aug. 2009

[53] M. Deveaux et al., "Radiation tolerance of a column parallel CMOS sensor with high resistivity epitaxial layer", JINST 6 C02004, 2011

[54] J.A. Ballin et al., "Monolithic Active Pixel Sensors (MAPS) in a Quadruple Well Technology for Nearly 100 Fill Factor and Full CMOS Pixels", Sensors 2008, 8, 5336-5351; DOI: 10.3390/s8085336

[55] A. Leona et al., "AToM: the front-end chip for the BaBar Silicon Vertex Tracker Detector", Proceedings of the 2^{nd} IEEE-CAS Region 8 Workshop on Analog and Mixed IC Design, 12-13 September 1997, pp. 59-64.

[56] P. Kinget, M. Steyaert, "Impact of transistor mismatch on the speed-accuracy-power tradeoff of analog CMOS circuits", Proceedings of the IEEE Custom Integrated Circuits Conference (CICC), 5-8 May 1996, pp. 333-336.

[57] K. Laker, W. Sansen, "Design of Analog Integrated Circuits and Systems", McGraw-Hill, New York, 1994.

[58] K.R. Lakshmikumar, R.A. Hadaway, M.A. Copeland "Characterization and modeling of mismatch in MOS transistors for precision analog design", IEEE J. Solid-State Circuits, vol. 21, no. 6, pp. 1057-1066, Dec. 1986.

[59] A. Gabrielli et al., "High efficiency readout for large matrices of pixels", Nucl. Instrum. Meth., sec. A, vol. 658, pp. 141-144, Dec. 2011.

[60] F. Giorgi et al., "2D and 3D Thin Pixel Technologies for the Layer0 of the SuperB Silicon Vertex Tracker" 2011 IEEE Nuclear Science Symposyum Conference Records (NSS/MIC), pp. 1324-1328.

[61] L. Ratti et al., "Design Optimization of Charge Preamplifiers With CMOS Processes in the 100 nm Gate Length Regime" IEEE Trans. Nucl. Sci., vol. 56, no. 1, pp. 235-242, Feb. 2009

[62] E. Pozzati et al., "MAPS in 130 nm triple well CMOS technology for HEP applications" Proceedings of the Topical Workshop on Electronics for Particle Physics, Prague (CZ), 3-7 Sep. 2007, pp.492-496.

[63] J.G. Fossum, "Computer-aided numerical analysis of silicon solar cells", Solid-State Electronics, vol. 19, pp. 269-277, 1976.

[64] R.S. Muller, T.I. Kamins "Device electronics for integrated circuits", John Wiley & Sons, New York, 1986.