Università degli Studi di Torino
Scuola di Dottorato

Development of Ultra Fast Silicon Detector for tracking in 4 dimensions

Marco Ferrero

Università degli Studi di Torino
Scuola di Dottorato

Dottorato in Fisica ed Astrofisica

Development of Ultra Fast Silicon Detector for tracking in 4 dimensions

Marco Ferrero

Tutor: Nocolò Cartiglia

Abstract

A significant luminosity increase will characterize the High Luminosity upgrade of the LHC accelerator complex. This new environment will cause a much higher level of pile-up events and radiation in the experiments. To cope with this harsh environment, the experiments need substantial upgrades: specifically, the CMS detector has decided to measure the time of flight of each particle to distinguish the particles associated with interesting events from those originated in uncorrelated, but overlapping, scattering processes. For this reason, the Minimum Ionizing Particle Timing Detector (MTD) is part of the CMS upgrade, and it will cover the CMS barrel and endcap regions.

This thesis is focused on the development of the Ultra-Fast Silicon Detector (UFSD), chosen to instrument the MTD endcap regions. UFSDs are thin silicon sensors based on the Low Gain Avalanche Diode technology; they have moderate internal gain and their design is optimized for precise timing measurements in an environment with a high radiation level. The UFSD active thickness, about 50 μm, combined with an internal gain of about 20 allows achieving a time resolution of ~ 40 ps and to operate up to an irradiation fluence of $3 \cdot 10^{15}$ n_{eq}/cm^2.

The first chapter of this thesis discusses the High Luminosity-LHC environment, in particular, the MTD detector of the CMS experiment, focusing on the endcap timing layer. The second and third chapters describe, respectively, the basic properties of silicon devices and their operation principles. The optimizations of the UFSD design to achieve the time resolution of about 30 ps are discussed in chapter four, while in the fifth chapter the specification of UFSD productions manufactured by Fondazione Bruno Kessler and Hamamatsu Photonics are reported. The sixth chapter provides an overview of the laboratory setups used in this thesis and the methodologies of measurement used for the UFSDs characterization are discussed. The seventh and

last chapter reports the core of the measurements and studies performed on the UFSD including studies on different gain layer designs, the inter-pad region of the devices, the radiation hardness, the time resolution, the yield, and the uniformity of UFSD productions.

Table of Contents

List of Tables

List of Figures

Chapter 1

The Timing of the CMS experiment Detector

This chapter describes the Timing Detector that will be installed at the Compact Muon Solenoid (CMS) experiment for the Large Hadron Collider (LHC) upgrade to High Luminosity Large Hadron Collider (HL-LHC). The timing information will play a fundamental role in events reconstruction and background rejection for several of the physics decay channels to be studied at the HL-LHC. The CMS Timing Detector is designed for the detection of Minimum Ionizing Particle (MIP). The main focus will be given to the detector that will be installed in the endcap regions, which will be instrumented with Ultra Fast Silicon Detectors, a new technology of thin silicon sensor with good time precision and radiation resistance.

1.1 High luminosity LHC

1.1.1 From LHC to High Luminosity LHC

The Large Hadron Collider (LHC) was commissioned in 2010 for proton-proton collisions with a center-of-mass energy of 7 TeV. During the years of activity, LHC delivered center-of-mass energy of $7-8$ TeV from April 2012 to the end of 2013, and an energy of 13 TeV from 2015, delivering a peak luminosity of 1.8×10^{34} $cm^{-2}s^{-1}$ and an integrated luminosity of 40 fb^{-1} per year in 2016-2018. The announcement given by CERN on 4 July 2012 about the discovery of Higgs boson was the first fundamental discovery at LHC. Thanks to LHC, Europe regained the leadership in High Energy Physic (HEP); the full exploitation of LHC is the main priority of the European strategy for HEP. After 2020 the statistic gain in running the accelerator without an increment of luminosity will become marginal; for example, the time re-

quired to halve the statistical error in measurements will be more than ten years. Therefore, to maintain scientific progress and to extend the discovery potential, LHC will be upgraded such to increase the peak luminosity by a factor five and the overall integrated luminosity by a factor ten. The new configuration of LHC is known as High Luminosity LHC (HL-LHC) [1, 2], with targets in term of luminosity of:

- peak luminosity of $5 \times 10^{34}\ cm^{-2}s^{-1}$
- integrated luminosity of 250 fb^{-1} per year and a total of 3000 fb^{-1} in a dozen years.

This increment in luminosity will be achieved with a two-step upgrade, (see figure 1.1).

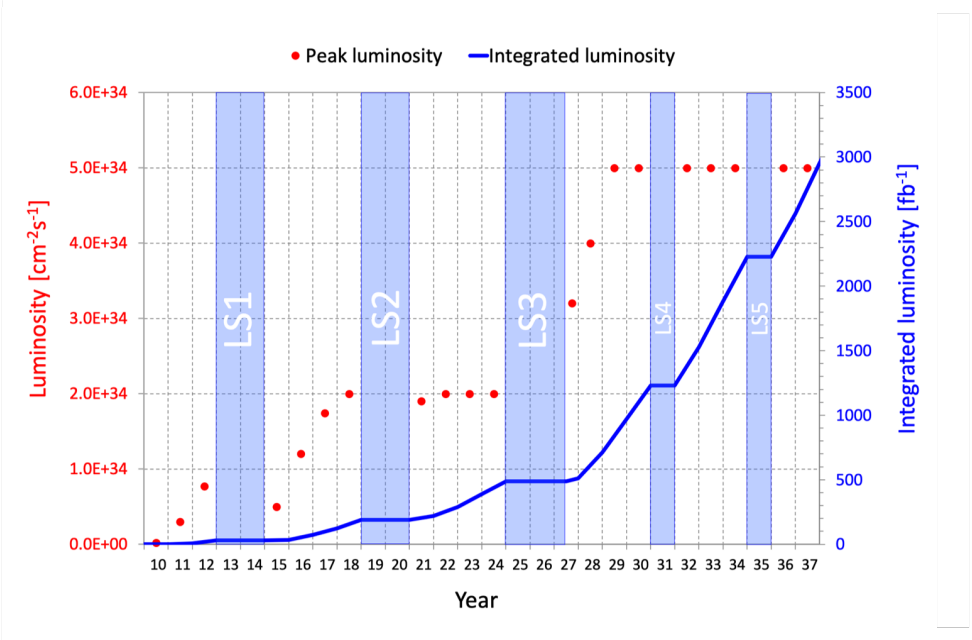

Figure 1.1: LHC luminosity plan for the next two decades, peak (red dots) and integrated (blue line) luminosity. Shutdown periods are indicated.

Starting from 2019 there is a two-year shutdown of LHC, called long shutdown 2 (LS2), dedicated to the upgrade of LHC injector complex to produce brighter (smaller emittance) beams. After a running period of four years, from 2025 there will be

another shutdown of about two and a half years, known as long shutdown 3 (LS3), to upgrade the optics in the interaction regions, to produce more tightly focused and overlapping beams at collision. The LHC will resume operation in the second half of 2027. From the second half of 2027, HL-LHC will be able to provide an instantaneous luminosity of $2 \times 10^{35}\ cm^{-2}s^{-1}$ at the beginning of each fill; in a nominal scenario, HL-LHC will operate at a steady peak luminosity of $5 \times 10^{34}\ cm^{-2}s^{-1}$ with an average number of interactions per bunch crossing of 140. The ultimate operational scenario of HL-LHC forecasts a peak luminosity of $7,5 \times 10^{34}\ cm^{-2}s^{-1}$ and an average number of interactions per bunch crossing of 200. This luminosity level can provide an integrated luminosity of 4000 fb^{-1} in a dozen years ($300 - 350\ fb^{-1}$ per year).

1.1.2 Physics motivations for HL-LHC

The physic program at LHC had a remarkable start during Run1, in 2011 and 2012, when LHC reached a peak of luminosity of $7,7 \times 10^{33}\ cm^{-2}s^{-1}$, more than 75% of its design luminosity. The vast amount of data collected in Run1 led to the discovery of a new 125 GeV mass particle by the ATLAS and CMS collaborations. This new particle was identified with the Standard Model (SM) Higgs boson, figure 1.2.

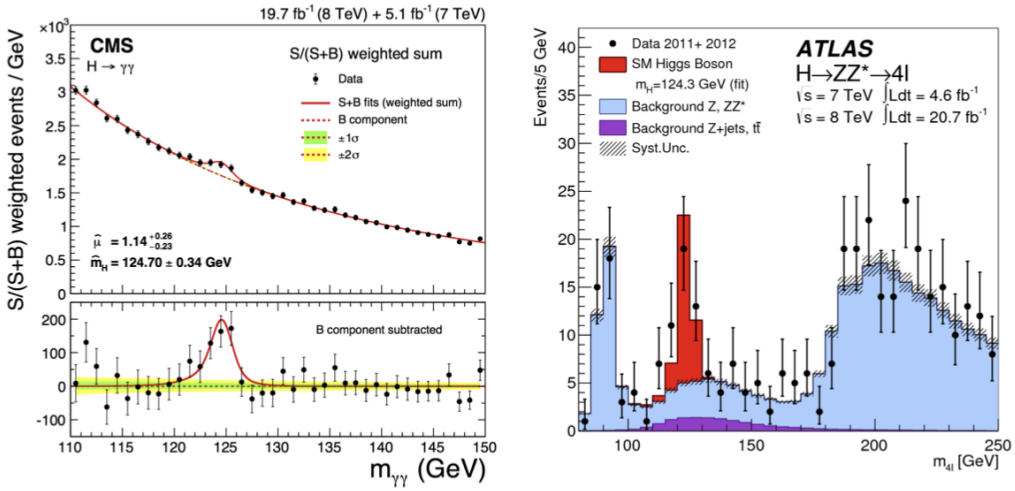

Figure 1.2: Observation of Higgs Boson by the CMS Collaboration in the $\gamma\gamma$ decay mode (left) and by the ATLAS Collaboration in the ZZ^* decay mode (right).

In addition to the discovery of this new particle, ATLAS and CMS have been able to begin an accurate study on its properties to prove that it was indeed the Higgs boson. Decays of the Higgs in SM gauge bosons W, Z, and in photons were established with a significance greater than 5σ; coupling in two-leptons decay $\tau^+\tau^-$ has been studied and is consistent with SM expectation; studies on Higgs decays provided the evidence on spin-parity 0^+ of the new boson.

The HL-LHC program aims at carrying out studies that involve the recent discovery of Higgs boson and the main target is to perform precise measurements of its scalar properties to validate the SM. The ATLAS and CMS collaborations estimate a comparable precision, with an uncertainty of 2-5%, for many of the Higgs boson couplings to elementary fermions and bosons, in a scenario with an integrated luminosity of 3000 fb^{-1}.

Moreover, the huge luminosity at HL-LHC will make it possible to investigate rare processes involving Higgs boson, such as $H \to Z\gamma$, or to study other rare processes that need high statistics, like the Higgs self-coupling through di-Higgs production, dark matter observation, fundamental forces unification and QCD behaviour under extreme condition.

1.1.3 Performance requirements at HL-LHC experiments

The high luminosity LHC upgrade will pose two main challenges, that the experiments will have to overcome to maintain the current excellent performances of the detectors in terms of efficiency, resolution and background rejection. These challenges are: (i) the radiation damage to the detector due to the high integrated luminosity and (ii) the huge pile-up of interactions due to the high instantaneous luminosity [2]. In this section, each of these challenges is discussed.

Radiation damage

The collisions rate in LHC is $5 \times 10^9\ s^{-1}$; the particles generated by the collisions cause damages to the detectors and the readout electronics, degrading their performances. The annual dose absorbed by detectors in HL-LHC will be comparable with the total dose absorbed from the beginning of LHC operation until LS3; therefore it is obvious that the maintenance of detector and electronic performances, in this harsh environment, will be challenging.

In LHC, the proton-proton collisions are the primary radiation source, generating a myriad of particles of various kinds:

- charged particles (mainly pions). They produce ionization into the detectors and undergo nuclear interactions producing showers of particles, which are added to the radiation load.

- photons. They either interact with the material of the beam pipe and with the tracker, producing electron-positron pairs or are adsorbed by the calorimeter, where they produce electromagnetic showers.

- neutrons. Neutron can cross long distances undergoing several scattering into the detectors and producing photons and electrons.

Figure 1.3 shows a simulation of the radiation level in the CMS detectors, for an integrated luminosity of 3000 fb^{-1}.

The damages produced by the radiation vary depending on detectors technology. In silicon, the radiation creates defects in its crystal lattice. One consequence is that the leakage current increases. The charge collection efficiency (CCE) decreases due to the creation of trapping centres for charge carriers, resulting in lower signals. The full depletion voltage of the sensors (voltage required to deplete the full thickness of the sensor) increases, forcing the operation of the sensors in partial depletion regime

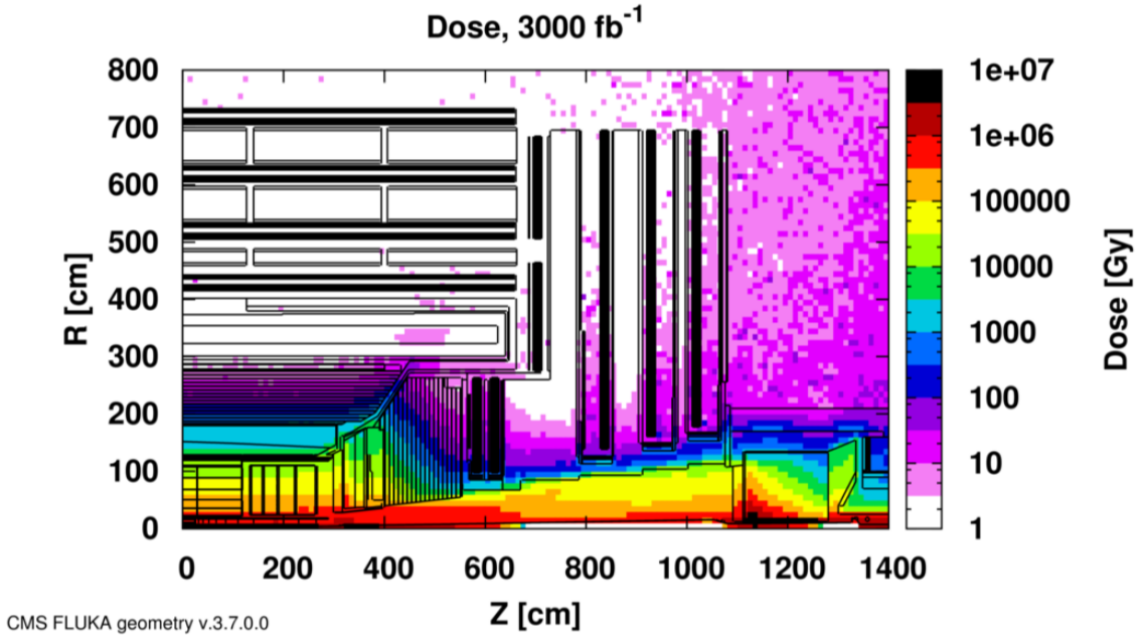

Figure 1.3: Absorbed dose in the CMS detectors after an integrated luminosity of 3000 fb^{-1}. R is the transversal distance from the beam-line, z is the distance along the beam-line from the interaction point ($z = 0$). Figure from [2].

and resulting in lower signals. For the calorimeter, that in CMS is mainly composed of scintillating PbWO4 crystals or plastic scintillating tiles with wavelength-shifting fibers embedded in them, the main problem is the loss of transmission of the media through which the scintillation light or wavelength-shifted light must pass.

Therefore, the strategies to mitigate the effect of the radiation vary from detector to detector.

High pile-up

Under the expected luminosity at HL-LHC, the major challenge for the experiments will be the high number of pile-up interactions, i.e. the superposition of several proton collision events within a typical interaction time frame. At LHC, each bunch of protons has a length with rms of ~ 5 cm, a transversal dimension of 10 μm and a population of 10^{11} protons. The temporal separation between two successive bunches is 25 ns, which in space is equivalent to 750 cm. Each beam contains about 2700

filled bunches and the collision between two bunches ("bunch crossing") occurs with a rate of 40 MHz.

At HL-LHC nominal luminosity, an average of 140 interactions will occur per bunch crossing, figure 1.4; most of these interactions are "soft" collisions, which do not contribute to the search for new physic. A small fraction of interactions will be "hard" collision, which will produce high transversal momentum particles that may contain high mass objects.

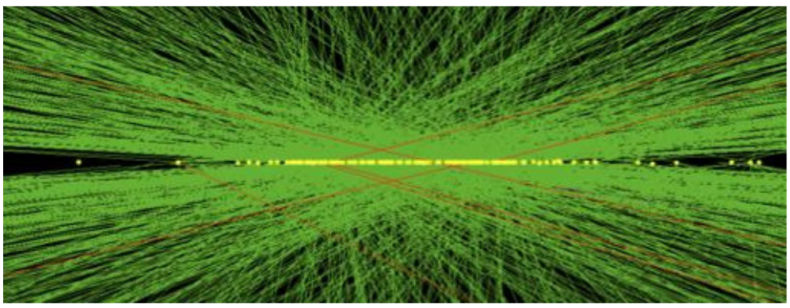

Figure 1.4: z-vertex distribution at pile-up conditions of HL-LHC. The expected amount of overlapping events will lay between 140 and 200. Figure from [2].

Since the number of bunches can not be increased by much it is obvious that an increase of luminosity involves an increase in a number of interactions per bunch crossing, i.e. in pile-up; as a result, there will be more hits in tracking detectors and addition of energy in calorimeter measurements. The high pile-up can confuse the trigger, the offline reconstruction and the interpretation of events; moreover, there will be an increase in event size in each bunch crossing, resulting in an increase of the events reconstruction time in the High Level Trigger and the offline analysis.

1.2 Overview of CMS phase-2 upgrade

The challenging running conditions at HL-LHC require an important upgrade of the CMS detector to maintain its excellent performances. This section is devoted to a brief overview of the main features of the CMS upgrade, which is foreseen for the

HL-LHC period called "Phase-2" upgrade [2].

Tracker: the tracker will suffer radiation damage effects, and it will be completely replaced for Phase-2. To maintain excellent tracking efficiency, at much higher pile-up levels, the pixel system and outer tracker granularity will be increased by a factor 4. In the outer tracker, this will be achieved by shortening the lengths of silicon sensor strips; in the pixel system, it will be implemented by having smaller pixels and thinner sensors. Moreover, in forwards regions, the coverage of the pseudorapidity region will be extended to close to $|\eta| \approx 4$.

Calorimeter endcaps: the calorimeter upgrade, the High Granularity Calorimeter (HGC), will be installed with electromagnetic and hadronic sections in the endcap regions. The electromagnetic section, with an interaction length of $25X_0$, consists of ~ 30 tungsten and copper plates interleaved with silicon sensors as the active material. The hadronic section has a front section of 12 brass and copper plates interleaved with silicon sensors for a depth of 3.5λ. The HGC is characterized by an excellent longitudinal and transverse segmentation.

Muon endcaps: currently, the muon system covers the pseudorapidity region $1,5 \leq \eta \leq 2,4$ and consists of four stations composed of Drift Tube (DT), Cathode Strip Chambers (CSC) and Resistive Plate Chambers (RPC). To maintain the good trigger acceptance in the endcap regions, it has been proposed to extend the coverage with four additional chambers: the first two stations, where the magnetic field is still high, will be Gas Electron Multiplier (GEM) chambers with excellent spatial resolution; the two last ones will be low-resistivity Resistive Plate Chambers (RCP), with lower granularity but with good time resolution to mitigate background effects. By implementing the GEM in the free region behind the endcaps of the calorimeter, the coverage of the muon detector will be close to $|\eta| \approx 3$.

1.3 Implementation of timing in the Phase-2 CMS detector

In the range from 140 pile-up interactions to 200, the peak "line density" dN_v/dz, of the number of collision vertices N_v, along the beam axis z, grows from 1.2 to 1.9 mm^{-1} (figure 1.5), with a rms spread along the beam axis of about 4.5 cm. Therefore, the probability of spatial overlaps grows in all subdetectors and the reconstruction algorithms begin to fail. According to simulation, the optimal selection window to associate a track with a primary vertex is of the order of 1 mm, causing not negligible contamination of tracks from pile-up into the primary vertex for vertex densities approaching 1 mm^{-1}. The resulting degradation in resolution, efficiencies, and misidentification rates impacts several measurements.

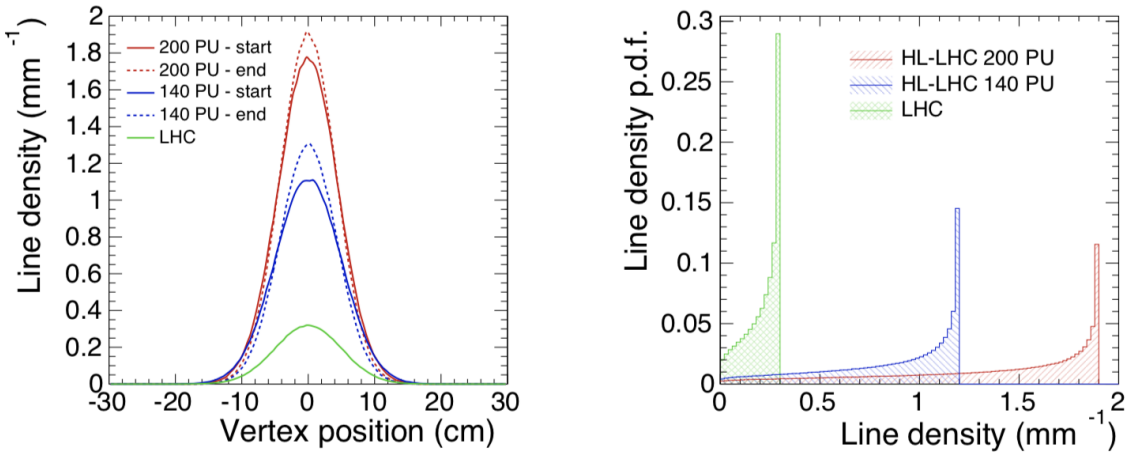

Figure 1.5: Left: density of vertices along the beam axis at the LHC with about 30 pile-up interactions (Run-1 and early Run-2) and at the HL-LHC with 140 and 200 pile-up interactions. The solid (dashed) line refers to the start (end) of the fill, [3]. Right: probability density function of the line density for pile-up values of about 30, 140 and 200. The modes and the means of the three distributions are respectively $0.3, 1.2, 1.9\ mm^{-1}$ and $0.2, 0.9, 1.4\ mm^{-1}$. Figure from [3].

The CMS collaboration has recently approved a technical design review to include the MIP Timing Detector (MTD) in the upgrade plan for the HL-LHC era. The MTD will provide timing information for minimum ionizing particles (MIPs), with a resolution of 30-40 ps at the beginning of the HL-LHC operation in 2027, and it will

improve the particle flow performance ([3], chapter 5) bringing it to a level comparable to the current Phase-1 CMS detector, which is designed to handle a pile-up of about 50 [4]. In the domain of time, the individual interactions of a bunch crossing are distributed with a rms of $\sim$ 150 ps (see figure 1.6), approximately constant during the fill and uncorrelated with the spatial distribution.

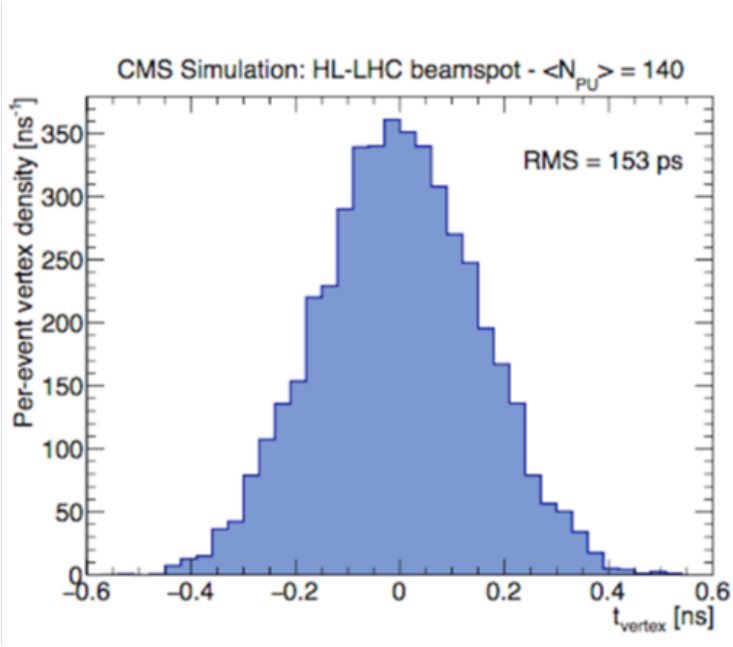

Figure 1.6: Distribution of the interaction time at HL-LHC considering an average pile-up of 140 vertexes. Figure from [5].

If one considers the beam spot sliced in consecutive time exposures of 30-40 ps, then the number of vertices per exposure drops to the current Run-2 LHC pile-up levels of 40-60. The event displayed in figure 1.8 visually demonstrates the power of space-time reconstruction in collisions with 200 pile-up events. According to simulation, the probability of vertex merging is reduced from 15% to 1% in space-time. The addition of track-time information from MTD, with a precision of $30 - 40$ ps reduces wrong associations by a factor two, and the peak line density is reduced from 1.9 mm^{-1} to 0.8 mm^{-1}. Therefore, the addition of MTD timing information has the potential to recover the effective background conditions close to Phase-1 operation

([3], chapter 1).

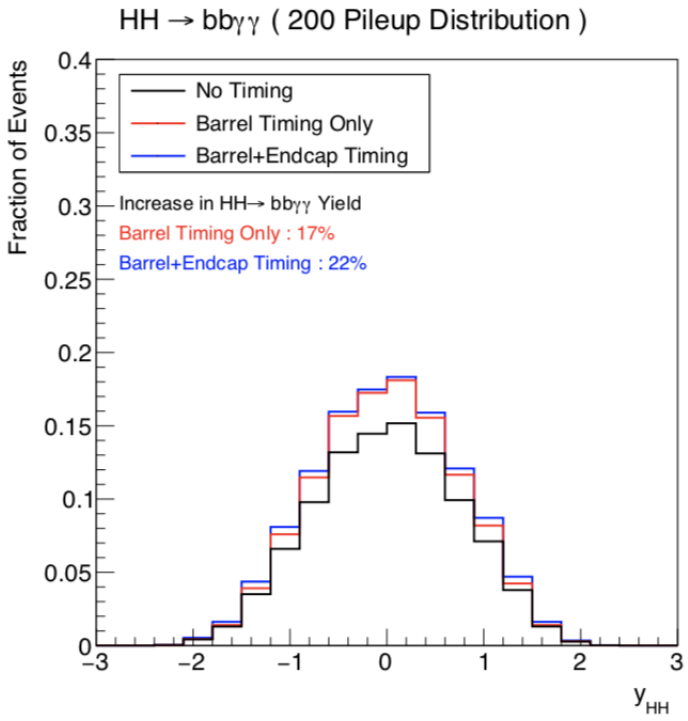

Figure 1.7: Impact of signal efficiency for $HH \to b\bar{b}\gamma\gamma$ for no timing, barrel timing and barrel plus endcaps timing scenarios. The quantity y_{HH} is the rapidity of the di-Higgs system. Figure from [3].

1.3.1 Impact of timing on CMS physics program

The CMS experimental program will benefit greatly from the increase of luminosity provided by the HL-LHC upgrade. The MTD will be crucial to maintain a good resolution and reconstruction efficiency for the physics objects; some of the benefits coming from the time tracking are listed below:

- One of the highest priorities of the HL-LHC physics program is the measurement of Higgs boson self-coupling. Increase in signal yields of 22% for constant background has been predicted for example for $HH \to b\bar{b}\gamma\gamma$ channel (see figure 1.7). Similar increases are also predicted for other significant di-Higgs final states ($bbbb$, $bb\tau\tau$, $bb\gamma\gamma$, $bbWW$, $bbZZ$).

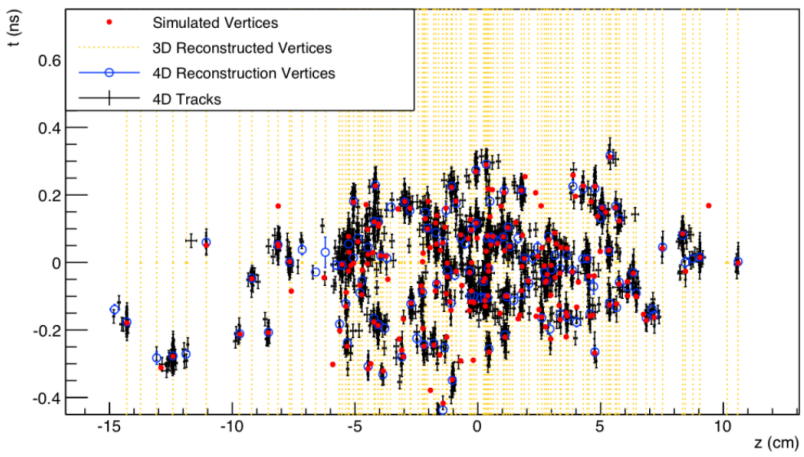

Figure 1.8: Simulated and reconstructed vertices in a bunch crossing with 200 pile-up interactions assuming a MTD with a time resolution of about 30 *ps*, covering the barrel and endcaps regions. The vertical yellow lines are the 3D-reconstructed vertices, the black crosses and the blue open circles represent tracks and vertices reconstructed using the time information (4D-reconstruction). The red dots are the simulated vertices. Figure from [3].

- A significant improvement in the missing transverse momentum (p_T^{miss}) resolution, in the case of final states with $H \to \tau\tau$ in the Vector Boson Fusion (VBF) production mode. The timing information counteracts some of performance degradations observed in the transition from 140 to 200 pile-up events.

- The track-timing reconstruction opens a new future in searches of Long-Lived-Particles (LLPs), postulated in many extensions of the Standard Model like Split-SUSY, GMSB, RPV SUSY, Stealth SUSY and SUSY models with compressed mass spectra. The space-time information associated with the displaced decay vertex, reconstructed from the decay daughters detected, provides the kinematic constraints necessary to get a direct measurement of the LLP mass.

- use of the Time Of Flight (TOF) technique for the particle detector for charged particles identification (PID) of charged hadrons, which is essential in heavy ions physics and in low mass QCD and flavour physics.

1.4 Overview of MTD design

This section is an overview of the design and specifications of MTD ([3], chapter 1). The constraints, linked to the CMS upgrade design and schedule, led the MTD design towards a thin layer located between the tracker and the electromagnetic calorimeter, divided in a barrel and two endcap sections, see figure 1.9.

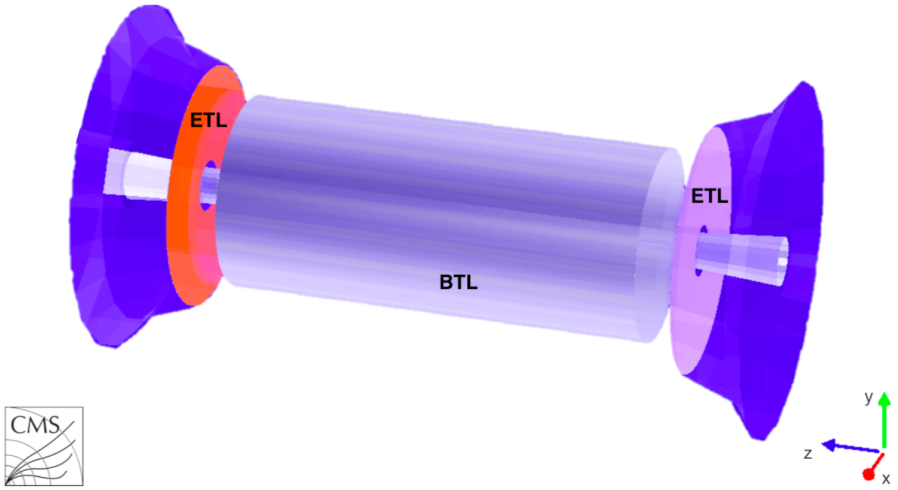

Figure 1.9: A simplified view of the MTD geometry implemented by GEANT for simulation studies comprising a barrel layer (grey cylinder), at the interface between the tracker and the electromagnetic calorimeter, and two silicon endcap layers (orange and violet disks) in front of the endcaps calorimeter. Figure from [3].

The requirements are quite different in the Barrel Timing Layer (BTL) and the Endcap Timing Layer (ETL). The first major difference between ETL and BTL is the radiation level: the expected dose in the inner part of the ETL is 30 times higher than the BTL absorbed dose. Table 1.1 shows the radiation dose for ETL and BTL, at different η and radial distance from the beam pipe. The second difference is the overall surface of the two sections: the BTL surface area is about 2.5 times larger than the two endcap sections. These two differences are at the origin of the selection of different detector technologies for each MTD section. The most suitable technology for the BTL is crystal scintillators read out by silicon Photon Multipliers (SiPMs), which are Avalanche Photo Diode (APD) operating in Geiger breakdown mode. The

Region	$\|\eta\|$	$r[cm]$	$z[cm]$	n_{eq}/cm^2	Dose $[kGy]$
Barrel	0.0	116	0	1.65×10^{14}	18
	1.15	116	170	1.8×10^{14}	25
	1.45	116	240	1.9×10^{14}	32
Endcap	1.6	127	303	1.5×10^{14}	19
	2.0	84	303	3×10^{14}	50
	2.5	50	303	7.5×10^{14}	170
	3.0	31.5	303	1.6×10^{15}	450

Table 1.1: Radiation doses and fluences (normalized to 1 MeV neutron equivalent in silicon) expected in the timing layer (Barrel and Endcaps) after an integrated luminosity of 3000 fb^{-1}, [3].

most suitable technology for the ETL is the Low Gain Avalanche Diode (LGAD), which is a silicon sensor with an internal moderate gain of $\sim$ 10-30. The SiPM technology is not sufficiently radiation hard to operate in the endcap regions, while the cost to instrument the barrel with LGADs is prohibitive. Moreover, the schedule constraints give less time to develop and assemble the barrel, therefore the choice of SiPM technology is dictated also by the fact that it is a well-established technology in the industry. Nevertheless, there are some services that will be shared between BTL and ETL: clock, backend system, cooling, detector low control and safety systems.

1.4.1 Barrel timing layer

The barrel timing layer is a thin, cylindrical detector shown in figure 1.10 (right), that will be housed inside the Tracker Support Tube (TST). This imposes constraints on its design and schedule of installation, which will have to happen in the third quarter of the year 2023. The TST will be inserted inside the CMS detector and will not be removed during the entire lifetime of HL-LHC. Due to the installation schedule, well-established technologies have been selected: LYSO:Ce scintillating crystals readout with SiPMs devices, capable to achieve time resolution below 30 ps. Both of them, crystals and SiPMs, have a proven radiation resistance up to a neutron irradiation fluence of 2×10^{14} n_{eq}/cm^2 (25 kGy as total integrated dose), corresponding to the

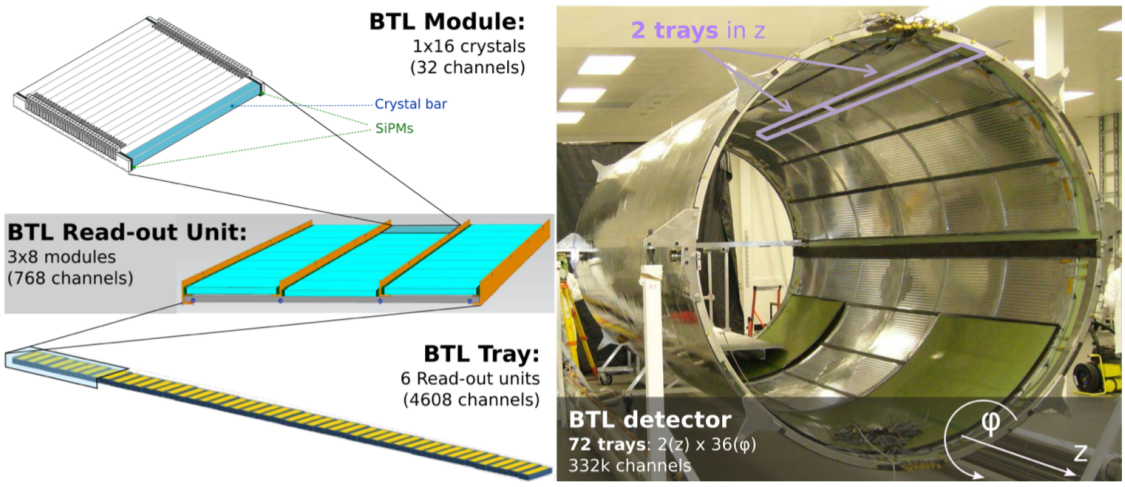

Figure 1.10: Overview of the BTL detector and the hierarchical arrangement of its components: bars, modules, read-out units and trays. Figure from [3].

expected fluence at the end of HL-LHC lifetime, in the BTL region.

The total active surface of the BTL is about 38 m^2, with a coverage of the pseudo-rapidity region up to $|\eta| = 1.48$. The fundamental detecting cell is a LYSO:Ce crystal bar of a length of about 5.7 cm oriented along the ϕ-direction of CMS, a width of 3 mm along the z-direction, and a variable radial thickness from 3.7 to 2.4 mm, (see figure 1.10, left). The total number of SiPMs used will be 331776, which will operate in Geiger mode with a gain of the order of 10^5 at a temperature of $-30\ ^\circ C$, to operate in a low dark current condition. Fluctuations in dark current cause a jitter that degrades the time resolution: the jitter will be the dominant term of the time resolution towards the end of HL-LHC operations.

The BTL is readout by a dedicated ASIC called TOFHIR (Time-of-Flight, High Rate), which readouts 32 SiPMs. TOFHIR is based on discrimination of the leading edge of their pulses followed by measurement with a time-to-digital converter (TDC). In order to achieve a good time resolution, the input of the discriminator should have a fast rise time and a large amplification factor. This technique needs to measure the amplitude of the pulse to correct the time walk (see section 4.2.2).

A more detailed and comprehensive overview of the BTL is described in [3], chap-

ter 2.

1.4.2 Endcap timing layer

The endcap timing layer is composed of a pair of disks on each side of the interaction region, instrumented with MIP-sensitive silicon devices, covering a pseudorapidity range between 1.6 and 3.0. Each pair of disks is located between the end of the tracker and the nose of the endcap calorimeter (see figure 1.11, left), at a distance of 2.98 m from the interaction point. Each disk has silicon devices on both sides, strategically positioned such that the read-out, power and cable infrastructures areas on each side are covered by the sensors on the opposite side. Moreover, to facilitate maintenance, the disks are split down the center vertically, so that they form a "clam shell" around the beam pipe (figure 1.11, right); for this characteristic, the disks can be removed without the need to remove also the beam pipe. The total sensor area of each MTD endcap will be of about 8 m^2.

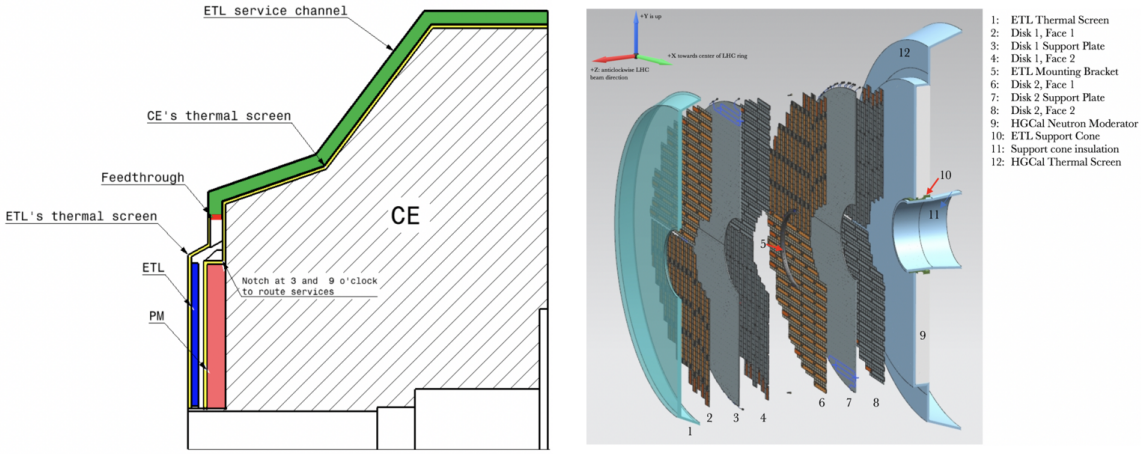

Figure 1.11: Left: placement of ETL (shown in blue) on the calorimeter endcap structure, on the interaction side of the polyethylene neutron moderator (in red). The ETL and endcap calorimeter detectors are in two separate volumes, with each detector having its thermal screen (shown in yellow), [3]. Right: ETL cross-sectional view along the beam axis. Shown are two disks instrumented with modules on each side, and with their support structure. The interaction point is to the left of the image. Figure from [3].

The radiation level expected in ETL is much higher and more non-uniform in η than in BTL. Figure 1.12 shows the expected fluence as a function of ETL radius at three different moments in time during HL-LHC lifetime; the fluence will reach values above 1×10^{15} n_{eq}/cm^2 for radii below 50 cm, in the second half of the HL-LHC lifetime. The expected radiation level in the endcap regions is the main driver for the choice of planar silicon sensors as detector technology.

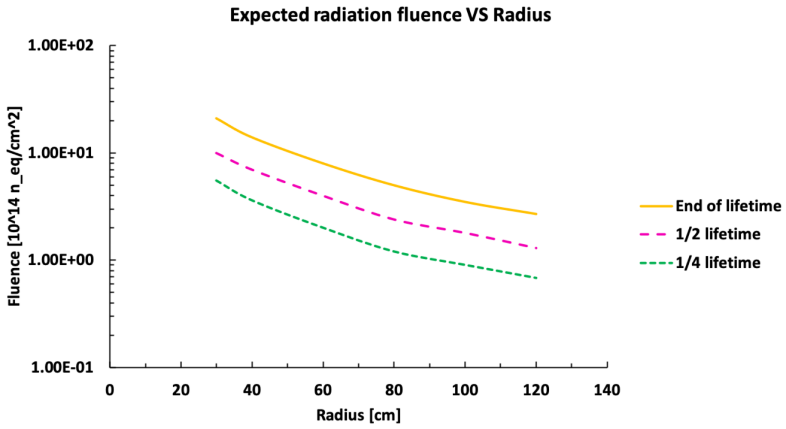

Figure 1.12: Expected radiation fluence, in 1 MeV neutron equivalent per cm^2, in ETL regions, as a function of radius, at three time moments during the operation period of HL-LHC.

Sensors characteristics

The planar silicon sensors chosen to instrument the ETL are Ultra Fast Silicon Detectors (UFSDs), based on LGAD technology, chapter 4. UFSDs are characterized by intrinsic moderate gain, about 10-30, provided by a thin implant with an appropriate doping density; this moderate gain combined with a thin sensor active thickness of $\sim$ 50 μm enables to achieve low-jitter and fast-rising pulses needed for a time resolution of 30 ps. Recent radiation hardness studies demonstrate the feasibility of having a time resolution of 30 and 50 ps at fluences corresponding to η of about 2.5 and 3.0, at the end of the lifetime. Achieving a good time resolution at moderate gain requires pixels (also called pads) with an area of few mm^2, in order to limit the

effect of the sensor capacitance on the time resolution value; therefore, a very large number of pads are needed to cover the total area of 8 m^2 of each endcap timing layer. The sensor design in the MTD technical proposal used very large sensors, of area 5×10 cm^2, with pads of 3×1 mm^2; R&D and optimization design studies led to an update of sensor design that now consists of sensors with area 21.1×42 mm^2, with pads of 1.3×1.3 mm^2. The reduction of sensors and pads area will imply respectively a higher yield (cost saving) and an improvement of time performances due to the lower pad capacitance [6].

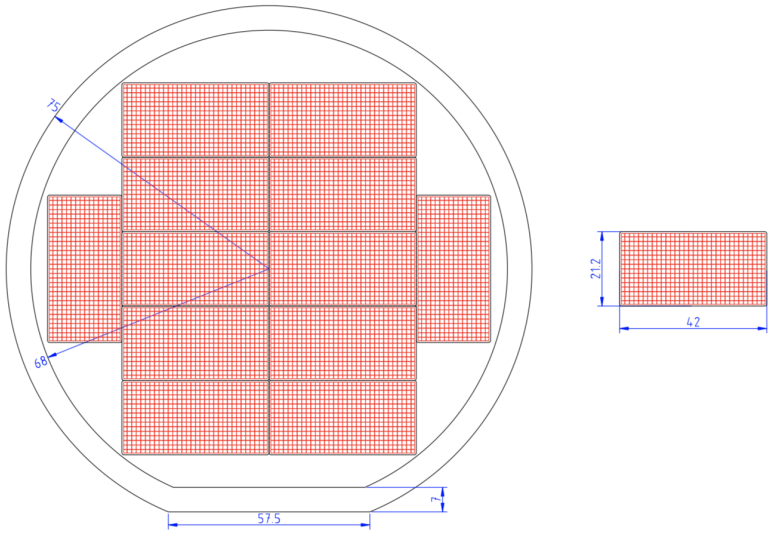

Figure 1.13: Layout of ETL sensors on a 6-inch wafer, with the pads shown. Figure from [3].

The wafer layout proposal composed of twelve sensors is shown in figure 1.13.

Particles crossing the ETL will have on average 1.8 hits per track: the detector design is optimized to achieve the target time resolution for track with two hits. There follow a list of key parameters that need to be studied for the ETL sensors:

- **Fill Factor:** the fill factor is the ratio of the active area to the total sensor area, which can be lower than 1 due to the no-gain region at the pad edges. The optimization of the fill factor increases the probability of two-hits tracks.

- **Sensor uniformity:** the production of the sensors must be as uniform as possible: all sensors must have the same gain to have the same time resolution and high efficiency.

- **Gain and noise:** the target time resolution is achieved only if the sensors provide a signal with high enough charge and low noise.

- **Radiation hardness:** It is necessary to identify the most radiation hard UFSD design. An excellent radiation resistance will allow achieving a good time resolution for the lifetime of the detector.

The ingredients of time resolution and UFSD sensors for high energy physics applications are the topics of this Ph.D. thesis and will be extensively covered in the following chapters.

Read-out electronic and modules

The read-out chip of ETL, called Endcap Timing Read-Out Chip (ETROC), has a design goal for time resolution of 50 ps per hit, to achieve a 35 ps resolution for tracks with one ETL hit in each of the two layers. The ETROC is designed to read out a 16×16 pixel matrix; at the single cell level, each channel consists of a preamplifier, a discriminator, a TDC used to digitize the Time of Arrival (ToA) and the Time over Threshold (ToT) measurements, and a memory for the data storage and readout. In addition, an injection circuit is also implemented for testing and calibration; figure 1.14 (left) shows a schematic block diagram of the ASIC. The ETL modules are composed of sub-assemblies containing a single sensors bump bonded to two ETROCs. The ETROCs are positioned with their short edge oriented along the long edge of the sensors; they extend over the long edge of the sensor forming a "balcony" with wire-bond pads for power connection and input-output signals. In the majority of modules, two sub-assemblies are glued to an Aluminium Nitride

substrate equipped with a cooling path. Figure 1.14 (right) shows the final assembled ETL modules and a view of their parts.

Figure 1.15 shows a front view of one ETL disk, with the layout of the modules and the service hybrids; The service hybrids, mounted between rows of modules, provide power and readout services to the modules via flex circuit connectors. A more detailed overview of ETL modules and their assembly is described in [3], chapter 3.

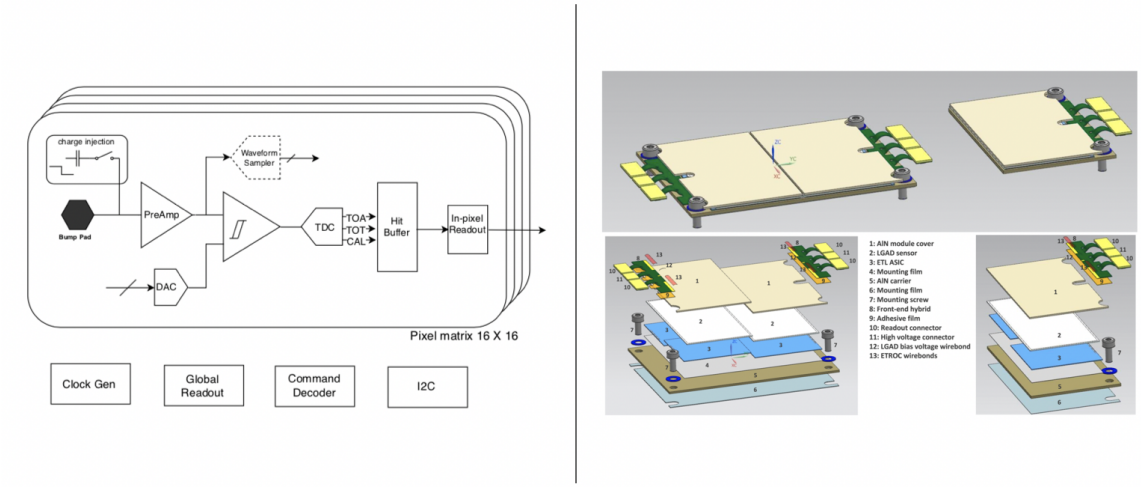

Figure 1.14: Left: schematic of single channel blocks of ETROC. Right: two-sensor and one-sensor assembled modules (top), details of the module parts (bottom). Figure from [3].

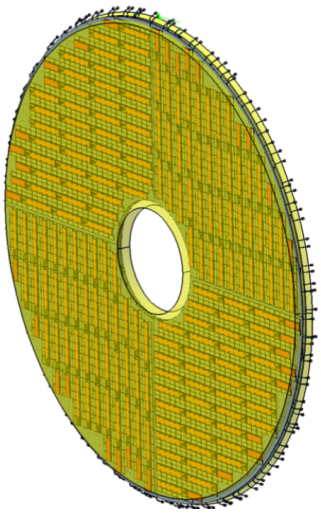

Figure 1.15: Front view of one ETL disk layout, with modules and service hybrids. Figure from [3].

Chapter 2

Basic semiconductor physics and silicon properties

The semiconductor sensors, particularly silicon sensors, play a major role in detectors for HEP applications. Low ionizing energy, good radiation resistance, ease integration with the readout electronic, extensive industrial experience in the development of sensors, and low production costs are the strengths that make these sensors suitable for nuclear physics experiments.

In this chapter will be illustrated the fundamental properties of semiconductor and p-n junctions (chapters 1 and 2 of [7]), the operating principles of silicon sensors for particles detection, their operation in environments with a high level of radioactivity and the effects of radiation on their macroscopic properties [8].

2.1 Semiconductor properties at thermal equilibrium

This section summarizes the basics physics of semiconductor devices and especially presents the main properties of silicon (Si). All the semiconductors, including silicon, are crystalline solid characterized by an energy band structure. This energy band structure is divided into three regions or bands: a forbidden energy region in which allowed states cannot exist and two permitted energy regions above and below the forbidden band. The upper band is called conduction band (CB), the lower one is the valence band (VB) and the forbidden band between the upper limit of the valence band and the lower limit of the conduction band is called band gap. Figure 2.1 shows a simplified representation of the semiconductors energy band structure. In this figure the energy at the bottom level of the conduction band is called E_C, the top level of the valence band is E_V, and the band gap is E_g. The electrons and holes energy is conventionally defined to be positive when measured respectively upward

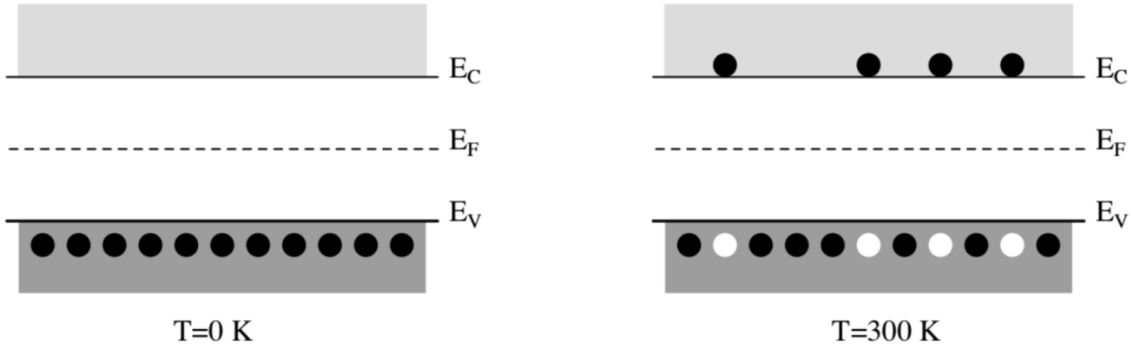

Figure 2.1: Simple sketch of band structure in semiconductors.

and downward whit respect to E_C and E_V. At room temperature ($\sim$ 300 K) and under normal atmosphere, the width of the band gap for high-purity silicon is 1.12 eV. At a temperature of 0 K all the states in VB are occupied while the CB is empty, as a result, there are no free carriers and the material behaves as an insulator; when the temperature increases the electron can acquire enough thermal energy to jump from VB to CB, becoming free carriers.

2.1.1 Intrinsic semiconductor

An intrinsic semiconductor is a material with a negligible amount of impurities. In silicon each atom shares its four valence electrons with the four neighbouring atoms, forming four covalent bonds, as shown in figure 2.2 (left).

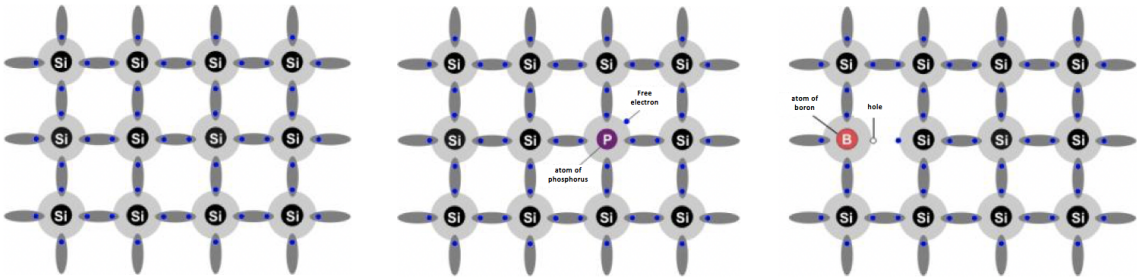

Figure 2.2: Basic bond picture of silicon: (left) intrinsic silicon with negligible impurity; (center) n-type silicon doped with donor impurity (phosphorus); (right) p-type silicon doped with acceptor impurity (boron).

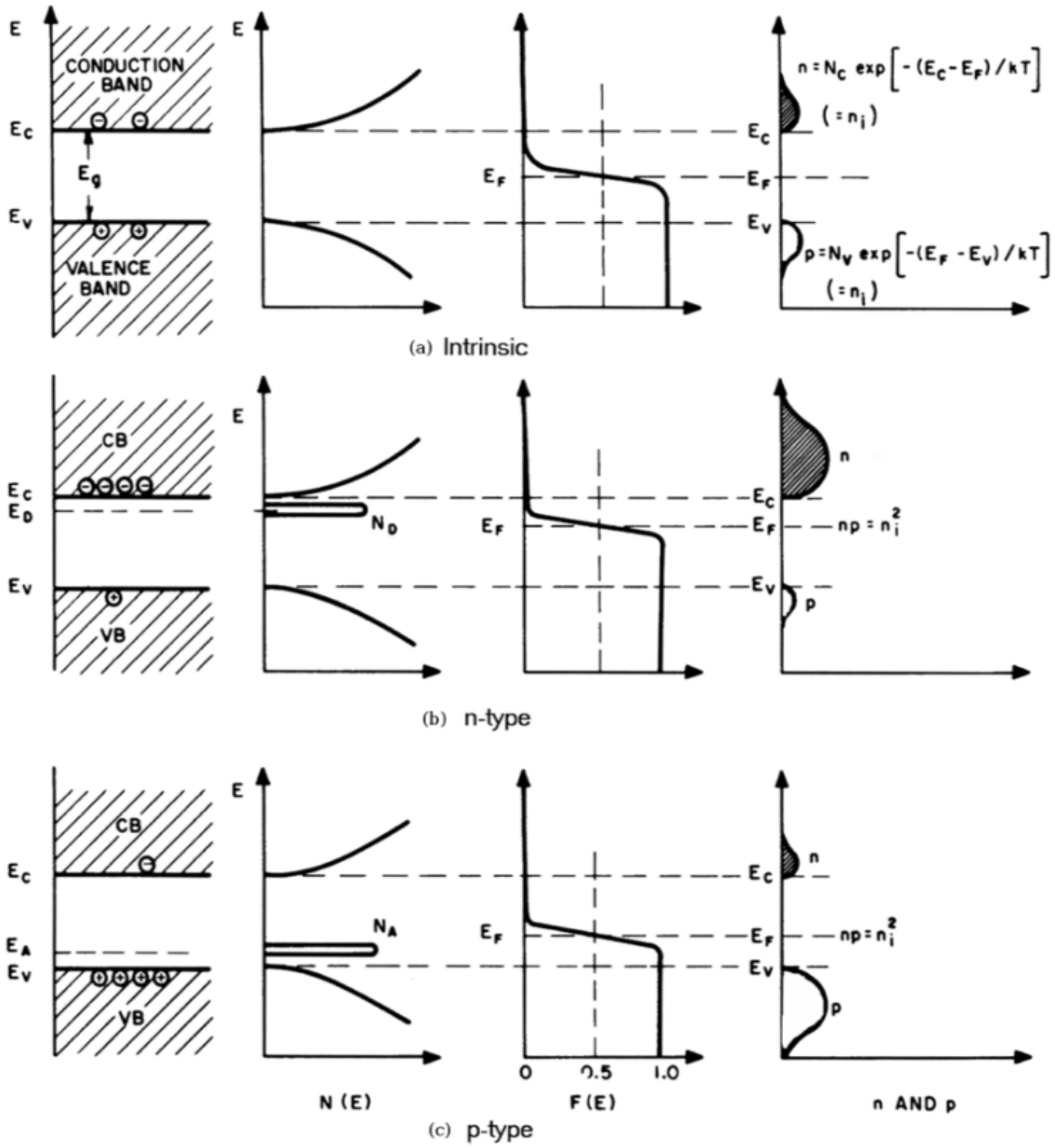

Figure 2.3: Schematic band diagram, density of states, probability distribution, and carrier concentration in an intrinsic (a), n-type (b), and p-type (c) semiconductor at thermal equilibrium. In n-type semiconductors there are electron energy levels close to the lower limit of the CB, so that they can be easily excited into the CB; in p-type semiconductors hole energy levels close to the upper limit of the VB allow excitation of electrons in VB, leaving mobile holes in VB. In n/p-type materials the Fermi level is nearer to the conduction and valence band, respectively. Note that for all the three cases is valid $pn = n_i^2$. Figure from [7].

In intrinsic silicon, the number of the occupied levels in the conduction band is given by

$$n = \int_{E_C}^{E_{top}} N(E)F(E)\,dE, \tag{2.1}$$

where E_C is the value of energy at the bottom of the conduction band and E_{top} is the energy at the top. $N(E)$ is the density of states, that can be approximated by the density near the bottom of the conduction band, for low enough carrier densities and temperatures

$$N(E) = M_C \frac{\sqrt{2}}{\pi^2} \frac{(E - E_C)^{1/2}}{\hbar^3} (m_{de})^{3/2}, \tag{2.2}$$

where M_C is the number of equivalent band minima in the conduction band and m_{de} is the density-of-state effective mass for electrons. $F(E)$ is the Fermi-Dirac distribution function, given by

$$F(E) = \frac{1}{1 + exp(\frac{E - E_F}{kT})}, \tag{2.3}$$

where k is the Boltzmann's constant, T the absolute temperature and E_F the Fermi energy, determined by the charge neutrality condition, (see section 2.1.2). The integral of equation 2.1 can be evaluated to be

$$n = N_C \frac{2}{\sqrt{\pi}} F_{1/2}(\frac{E_F - E_C}{kT}), \tag{2.4}$$

where N_C is the effective density of states in the conduction band and it is given by

$$N_C \equiv 2(\frac{2\pi m_{de} kT}{h^2})^{3/2} M_C \tag{2.5}$$

and $F_{1/2}(\eta_f)$ is the Fermi-Dirac integral, given by

$$F_{1/2}(\eta_f) = \int_0^\infty \frac{\eta_{1/2}}{1 + e^{(\eta - \eta_f)}}\,d\eta, \tag{2.6}$$

where η is

$$\eta \equiv \frac{(E_F - E_C)}{kT}. \tag{2.7}$$

For the Boltzmann statistic case (Fermi level kT below E_C), the number of occupied conduction band levels, equation 2.4, becomes

$$n = N_C exp(-\frac{E_C - E_F}{kT}). \tag{2.8}$$

Similarly, we can obtain the density of holes near the top of the valence band

$$p = N_V exp(-\frac{E_F - E_V}{kT}), \tag{2.9}$$

where N_V corresponds to the density of states in the valence band (given by an equation similar to 2.5), and E_V is the value of energy at the top of the valence band. The carrier concentrations (equations 2.8, 2.9) for an intrinsic semiconductor are graphically shown in figure 2.3 (top).

In an intrinsic semiconductor, a continuous thermal agitation exists at finite temperatures, with the results to excite the electrons from the valence to the conduction band. These electrons leave an equal number of holes in the valence band, which is $n = p = n_i$, where n_i corresponds to the intrinsic carrier density. This process is balanced by the recombination of the electrons in the conduction band and the holes in the valence band. The Fermi level for the intrinsic semiconductor in obtained by combining equations 2.8 and 2.9

$$E_F = E_i = \frac{E_C + E_V}{2} + \ln(\frac{N_V}{N_C}). \tag{2.10}$$

The Fermi level E_i in an intrinsic semiconductor is very close to the middle of the band gap, and the intrinsic carrier density is obtained from equations. 2.8, 2.9 and

2.10

$$np = {n_i}^2 = N_C N_V exp(-\frac{E_g}{kT}), \tag{2.11}$$

where E_g is

$$E_g \equiv (E_C - E_V). \tag{2.12}$$

The relation 2.11 is called law of mass action.

2.1.2 Doped semiconductor

Two different types of dopants can bind with the atoms of semiconductor crystalline lattice: donor (phosphorus, arsenic) and acceptor (boron, gallium). A donor is an atom with five valence electrons (Group V of the periodic table of elements), it replaces an atom of the semiconductor crystalline lattice sharing four of its valence electrons, the fifth one is 'donated' to the conduction band. Semiconductor doped with donors is called n-type, figure 2.2 (center). The acceptor is an atom with three valence electrons (Group III) and substituting an atom of the lattice 'accept' an additional electron. A positive charge (hole) is created in the valence band and the semiconductor is called p-type. Figure 2.2 (center and right) shows a sketch of n-type and p-type silicon lattice with the type of dopant respectively phosphorus and boron.

When a semiconductor is doped with donors or acceptors, impurity energy levels are introduced into the forbidden band gap. A donor level is defined neutral if filled by an electron and positive if empty, instead, the acceptor level is defined neutral if empty and negative if filled by an electron. The simplest approach to calculate the impurity energy levels is based on the hydrogen-atom model; the first ionization energy of the hydrogen atom is given by

$$E_H = \frac{m_0 q^4}{32\pi^2 {\epsilon_0}^2 \hbar^2} = 13.6\ eV, \tag{2.13}$$

where m_0 is the mass of the free electron, q is the elementary charge and ϵ_0 is the free-

space permittivity. The ionization energy for the donor E_d can be obtained replacing m_0 with the conductivity effective mass of electron m_{ce} and replacing ϵ_0 with the permittivity of the semiconductor ϵ_S, in equation 2.13:

$$E_d = (\frac{\epsilon_0}{\epsilon_S})^2(\frac{m_{ce}}{m_0})E_H. \tag{2.14}$$

The ionization energy for donors is $0,025\ eV$ for silicon (measured from conductive-band edge). In a similar way it is possible to calculate the ionization energy for acceptors, which is $0,05\ eV$ for silicon (measure from the valence-band edge).

Fermi level

The Fermi level for intrinsic silicon, equation 2.10, is very close to the middle of the band gap, as shown in figure 2.3 (top). In doped semiconductor, the Fermi level must adjust itself to maintain the charge neutrality, figure 2.3 (center, bottom). Consider the case of a concentration N_D of donor impurities added to the semiconductor lattice, to preserve the charge neutrality the number of negative charges must be equal to the positive

$$n = N_D^+ + p, \tag{2.15}$$

where n and p are respectively the densities of electrons and holes in conductive and valence band, and N_D^+ is the density of ionized donors.

$$N_D^+ = N_D[1 - \frac{1}{1 + \frac{1}{g}exp(\frac{E_D - E_F}{kT})}], \tag{2.16}$$

where g is the ground state degeneracy of the donor impurity level, it equals 2 because a donor level can accept one electron with either spin or can have no electron. In the same way, if some of acceptors of density N_A are added to the lattice, we can write a

similar equation of charge neutrality

$$p = N_A^- + n, \tag{2.17}$$

with

$$N_A^- = \frac{N_A}{1 + gexp(\frac{E_A - E_F}{kT})}, \tag{2.18}$$

where the factor g equals 4 because each acceptor impurity level can accept one hole with either spin and the impurity level is doubly degenerate. Rewriting equation 2.15, we obtain

$$N_C exp(-\frac{E_C - E_F}{kT}) = N_D \frac{1}{1 + 2exp(\frac{E_F - E_D}{kT})} + N_V exp(\frac{E_V - E_F}{kT}). \tag{2.19}$$

For a set of values of N_C, N_D, N_V, E_C, E_D, E_V, and a fixed temperature T we can calculate the Fermi level E_F from equation 2.19. In the specific case with $T = 300\ K$ and $N_D = 10^{16}\ atoms/cm^{-3}$ the Fermi level is very close to the edge of the conduction band. At a relatively high temperature, the majority of donors and acceptors are ionized and the condition of charge neutrality can be written as

$$n + N_A = p + N_D. \tag{2.20}$$

Combining equation 2.20 with the law of mass action, equation 2.11, through some mathematical passages we can obtain the concentration of electrons and holes in an n-type semiconductor, at thermal equilibrium:

$$n = N_D, \tag{2.21}$$

$$p = \frac{{n_i}^2}{N_D}, \tag{2.22}$$

and the Fermi level is

$$E_F = E_C - kT \ln(\frac{N_C}{N_D}). \tag{2.23}$$

In the same way, in a p-type semiconductor, the concentration of holes and electrons are:

$$p = N_A, \tag{2.24}$$

$$n = \frac{{n_i}^2}{N_A}, \tag{2.25}$$

and the Fermi level is

$$E_F = E_V + kT \ln(\frac{N_V}{N_A}). \tag{2.26}$$

In an n-type semiconductor, the electrons and holes are called respectively majority and minority charge carriers, their role is reversed in a p-type. A graphic representation of the majority and minority charge carrier concentrations in an n-type and p-type semiconductor are shown respectively in figure 2.3 (center, bottom).

2.2 Carrier transport mechanism

The carrier transport in a semiconductor is permitted by two independent mechanisms, drift, and diffusion. The drift mechanism is originated by the collective movement of charge carriers under the effect of an electric field, the diffusion is due to a different concentration of charges inside the semiconductor.

2.2.1 Drift

Considering a slab of silicon with n electrons per unit of volume, applying an external potential to the slab's borders the electrons drift under the effect of the electric field ε generate by the potential. The current density due to this drift of electrons is given

by

$$j = qnv_d, \tag{2.27}$$

where q is the elementary charge of the electron and v_d is its drift velocity. A moving electron in the semiconductor lattice scatters with the atoms, this scattering reduces to zero the maximum drift velocity reached by the electron before the collision. The time τ between two collisions is called mean free time, and during this time an electron acquires a maximum kinetic energy

$$\frac{1}{2}m_e{v_d}^2, \tag{2.28}$$

where m_e is the effective mass of the electron. This energy is equivalent to the energy acquired by an electron in a constant electric field

$$\frac{1}{2}q\varepsilon v_d\tau. \tag{2.29}$$

Equalizing equations 2.28 and 2.29 we can obtain the maximum drift velocity of an electron in a crystalline lattice

$$v_d = \frac{q\tau\varepsilon}{m_e}. \tag{2.30}$$

The drift velocity is proportional to the electric field ε, the proportionality constant $q\tau/m_e$ is called mobility μ. Electrons and holes have different values of mobility, μ_n, and μ_p. Following equations 2.27 and 2.30, the total current density flowing through the slab is the sum of the electron and hole contributes

$$j = j_n + j_p = q(n\mu_n + p\mu_p)\varepsilon. \tag{2.31}$$

The proportionality constant between the electric field and current density is called conductivity σ

$$j = \sigma\varepsilon. \tag{2.32}$$

The reciprocal value of σ is called resistivity ρ

$$\rho = \frac{1}{\sigma} = \frac{1}{q(\mu_n n + \mu_p p)}. \tag{2.33}$$

In the specific case of n-type silicon ($n \gg p$), equation 2.33 becomes

$$\rho \simeq \frac{1}{q\mu_n n}. \tag{2.34}$$

In the same way in p-type silicon, the resistivity is

$$\rho \simeq \frac{1}{q\mu_p p}. \tag{2.35}$$

The mobility and consequently the resistivity is not constant, several mechanisms affect their values. The two primary ones are phonon and impurity scattering. Phonons are atoms oscillations in the lattice around their equilibrium point, these oscillations are due to the thermal excitation of atoms. The impurities, as the dopant atoms, are ionized and they act as scattering centers: as results, higher is the doping concentration lower is the mobility.

2.2.2 Diffusion

The diffusion mechanism is originated by a different concentration of charges between two regions of the semiconductor. Consider a slab of silicon with a different density of electrons between right and left side: at thermal equilibrium the electrons have the same average velocity and a random direction of movement. Therefore, since the electron concentration is not constant, there will be a greater number of electron

flowing from the higher density region to the lower density one of the slab. The current density due to the electron diffusion is

$$j_n = qD_n \frac{dn}{dx}. \tag{2.36}$$

Similarly, the hole current density is

$$j_p = qD_p \frac{dp}{dx}, \tag{2.37}$$

where $\frac{dn}{dx}$ and $\frac{dp}{dx}$ are the gradients of electron and hole concentration along the x direction and $D_{n/p}$ are the diffusion coefficients of the carriers. At thermal equilibrium, for not degenerate semiconductor ($n \ll N_c$ and $p \ll N_v$), these coefficients are associated with the mobility by the relationships:

$$D_n = (\frac{kT}{q})\mu_n, \tag{2.38}$$

and

$$D_p = (\frac{kT}{q})\mu_p. \tag{2.39}$$

Equations 2.38, 2.39 are known as Einstein relationships. When drift and diffusion act simultaneously, the total current densities for electrons and holes are given by the sum of the two contributes:

$$j_n = qn\mu_n\varepsilon + qD_n \frac{dn}{dx}, \tag{2.40}$$

$$j_p = qp\mu_p\varepsilon + qD_p \frac{dp}{dx}. \tag{2.41}$$

2.3 Generation-ricombination

Whenever the thermal equilibrium condition of a physical system is compromised (i.e., $pn \neq {n_i}^2$), processes exist to restore the system to equilibrium (i.e., $pn = {n_i}^2$). This is the situation in which electrons and holes are generated or recombined, these mechanisms change the numbers of carriers in the conduction and valence band. Generation consists of an electron coming from the conduction band that occupies an empty level in the valence band, the recombination is the inverse mechanism. These two processes are perfectly balanced at equilibrium, while out the equilibrium the algebraic sum of generation and recombination rate, called generation-recombination rate $U = G - R$ is different from zero, $U \neq 0$. The three basic recombination processes are:

- Band-to-band recombination, an electron in conduction band recombines with a hole in the valence band. This transition is made possible by the emission of a photon (radiative process) or by transfer of the energy to another free electron or hole (Auger process).

- Single-level recombination, in which one only trapping energy level is present in band gap. This mechanism can be described by four different processes, electron and hole capture and electron and hole emission.

- Multi-level recombination, in which more than one trapping level is present in band gap.

In addition, in high electric field, the generation of electron-hole pairs happens. As previously mentioned, at low electric field the drift velocity of free carriers is proportional to the electric field and the proportionality constant is the mobility ($\mu_n = 1350\ cm^2V^{-1}s^{-1}$, $\mu_P = 450\ cm^2V^{-1}s^{-1}$). In an enough large electric field condition, effects of nonlinearity in mobility and saturation of drift velocity are ob-

served. At still large electric field, impact ionization or avalanche generation occurs. At thermal equilibrium, the carriers both emit and absorb phonons and the total rate of exchange of energy is zero. In the presence of an electric field, the carriers acquire energy from the field and lose it to phonons by emitting more phonons than are absorbed. At high electric field condition, the most frequent scattering event is the emission of optical phonons. The carriers, on the average, acquire more energy than they have at thermal equilibrium and the drift velocity approaches a saturation value, as shown in figure 2.4

$$v_s = \sqrt{\frac{8E_p}{3\pi m_0}} \sim 10^7 \ \frac{cm}{s}, \tag{2.42}$$

where E_p is the optical-phonon energy ($\sim 0.063\ eV$ for silicon). When the electric field is large enough the carriers gain enough energy to excite electron-hole pairs by impact ionization and charge multiplication occurs. The electron-hole pairs generation rate G from impact ionization is given by

$$G = \alpha_n n v_n + \alpha_p p v_p, \tag{2.43}$$

where α_n is the electron ionization rate defined as the number of electron and hole pairs generated by an electron per unit distance traveled. Similarly, α_p is the hole ionization rate. α_n and α_p are strongly dependent on the electric field. A physical expression for the ionization rate is

$$\alpha(\varepsilon) = \frac{q\varepsilon}{E_I} exp(-\frac{\varepsilon_I}{\varepsilon(1+\frac{\varepsilon}{\varepsilon_p}) + \varepsilon_{kT}}) \tag{2.44}$$

where E_I is the effective ionization threshold energy (in silicon, 3.6 eV for electrons and 5 eV for holes), and ε_{kT}, ε_p, and ε_I are threshold fields for carriers to overcome the decelerating effects of thermal, optical-phonon and ionization scattering, respectively.

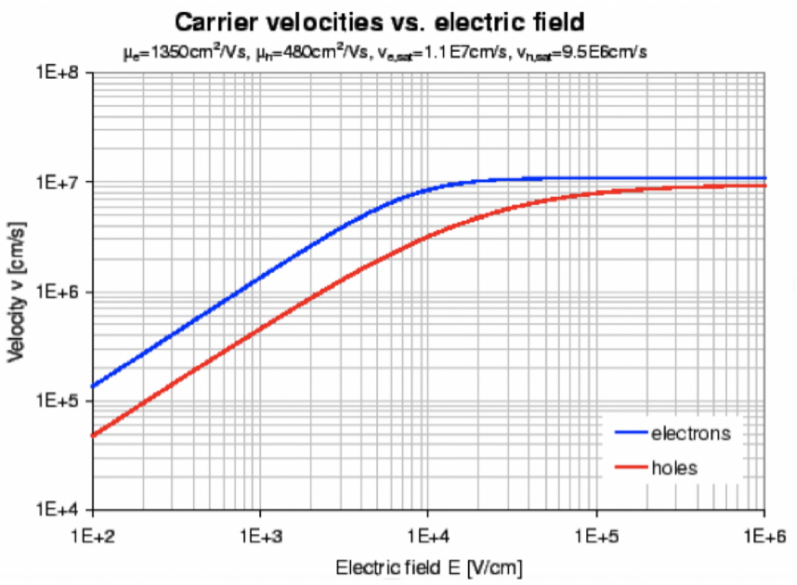

Figure 2.4: Dependence of the drift velocity on the electric field for electrons (blue curve) and holes (red curve).

2.4 Basic equations for semiconductor-device operation

The basic equations for semiconductor device operation describe the static and dynamic behaviour of carriers in semiconductor under the effect of an external field, that causes a deviation from the thermal-equilibrium conditions. The basic equations can be classified into three groups:

- **Maxwell equations for homogeneous and isotropic materials**

$$\nabla \times \varepsilon = -\frac{\partial B}{\partial t} \tag{2.45}$$

$$\nabla \times H = -\frac{\partial D}{\partial t} + J_{cond} = J_{tot} \tag{2.46}$$

$$\nabla \cdot D = \rho(x, y, z) \tag{2.47}$$

$$\nabla \cdot B = 0 \tag{2.48}$$

$$B = \mu_0 H \tag{2.49}$$

$$D(r,t) = \int_{-\infty}^{t} \epsilon_s(t-t')\varepsilon(r,t')dt' \tag{2.50}$$

where ε and $D = \epsilon_s\varepsilon$ are electric field and displacement vector, respectively; H and B are magnetic field and induction vector, respectively; J_{cond} is the conduction current density and J_{tot} the total current density; $\rho(x,y,z)$ is the total electric charge density; ϵ_s and μ_0 are the permittivity and permeability, respectively. The most important, in this list of equations, is the Poisson equation (equation 2.63) that determines the properties of the depletion layer of the p-n junction.

- **Current-density equations**

$$J_n = q\mu_n n\varepsilon + qD_n\nabla n, \tag{2.51}$$

$$J_p = q\mu_p p\varepsilon - qD_p\nabla p, \tag{2.52}$$

$$J_{cond} = J_n + J_p, \tag{2.53}$$

where J_n and J_p are the electron and hole current densities. They consist of the drift components due to the field and the diffusion components due to the current concentration gradient.

- **Continuity equation**

$$\frac{\partial n}{\partial t} = G_n - U_n + \frac{1}{q}\nabla \cdot J_n, \tag{2.54}$$

$$\frac{\partial p}{\partial t} = G_p - U_p - \frac{1}{q}\nabla \cdot J_p, \tag{2.55}$$

where G_n and G_p are the electron and hole generation rates due to external influence as optical excitation with high energy photons or impact ionization under large electric fields, while U_n and U_p are the recombination rates.

2.5 The p-n junction

The basic equations presented in previous sections of this chapter are used to develop the static and dynamic characteristics of an ideal p-n junction. A p-n junction is a two-terminal device, that can perform various terminal functions based on doping profile, device geometry, and biasing condition.

2.5.1 Depletion region and depletion capacitance

We assume an ideal p-n abrupt junction, where the impurity concentration changes abruptly from acceptor impurities N_A to donor impurities N_D, as shown in figure 2.5 (a). We first consider the junction at thermal equilibrium, without applied bias, and without current flow. From equations 2.23 and 2.51

$$J_n = 0 = q\mu_n(n\varepsilon + \frac{kT}{q}\frac{\partial n}{\partial x}) = \mu_n n\frac{\partial E_F}{\partial x}. \tag{2.56}$$

Similarly

$$J_p = 0 = \mu_p p\frac{\partial E_F}{\partial x}. \tag{2.57}$$

The condition of zero electron and hole current density requires that the Fermi level must be constant through the p and n sides of the junction. The diffusion potential, or built-in potential V_{bi}, is equal to

$$\begin{aligned} qV_{bi} = E_g - (qV_n + qV_p) = kT\ln(\frac{N_C N_V}{{n_i}^2}) - [kT\ln(\frac{N_C}{n_{n0}}) + kT\ln(\frac{N_V}{p_{p0}})] \\ = kT\ln(\frac{n_{n0}p_{p0}}{{n_i}^2}) \simeq kT\ln(\frac{N_A N_D}{{n_i}^2}). \end{aligned} \tag{2.58}$$

Since, at equilibrium $n_{n0}p_{n0} = n_{p0}p_{p0} = n_i^2$,

$$V_{bi} = \frac{kT}{q}\ln(\frac{p_{p0}}{p_{n0}}) = \frac{kT}{q}\ln(\frac{n_{n0}}{n_{p0}}). \tag{2.59}$$

Equation 2.59 gives the relationship between the hole and electron densities on either sides of the junction:

$$p_{n0} = p_{p0}exp(-\frac{qV_{bi}}{kT}), \tag{2.60}$$

$$n_{p0} = n_{n0}exp(-\frac{qV_{bi}}{kT}). \tag{2.61}$$

Since at thermal equilibrium the electric field in the neutral regions of the semiconductor must be zero, then the total negative charge per unit area in the p side must be equal to the total positive charge per unit of area in the n side

$$N_A x_p = N_D x_n. \tag{2.62}$$

From Poisson's equation, for the abrupt approximation, we obtain

$$-\frac{\partial^2 V}{\partial^2 x^2} \equiv \frac{\partial \varepsilon}{\partial x} = \frac{\rho(x)}{\epsilon_s} = \frac{q}{\epsilon_s}[p(x) - n(x) + N_D^+(x) - N_A^-(x)] \tag{2.63}$$

or

$$-\frac{\partial^2 V}{\partial^2 x^2} \approx \frac{q}{\epsilon_s} N_D \quad \text{for } 0 < x \leq x_n, \tag{2.64}$$

$$-\frac{\partial^2 V}{\partial^2 x^2} \approx -\frac{q}{\epsilon_s} N_A \quad \text{for } -x_p \leq x < 0. \tag{2.65}$$

By the integration of equations 2.64 and 2.65 we obtain the electric field, as shown in figure 2.5 (b)

$$\varepsilon(x) = -\frac{qN_A(x + x_p)}{\epsilon_s} \quad \text{for } -x_p \leq x < 0 \tag{2.66}$$

and

$$\varepsilon(x) = -\varepsilon_m + \frac{qN_D x}{\epsilon_s} = \frac{qN_D}{\epsilon_s}(x - x_n) \quad \text{for } 0 < x \leq x_n, \tag{2.67}$$

where ε_m is the maximum electric field corresponding to the junction at $x = 0$, it is given by

$$|\varepsilon_m| = \frac{qN_D x_n}{\epsilon_s} = \frac{qN_A x_p}{\epsilon_s}. \tag{2.68}$$

Integrating equation 2.63 once again, we obtain the potential distribution and the built-in potential (figure 2.5 (c)):

$$V(x) = \varepsilon_m(x - \frac{x^2}{2W}), \tag{2.69}$$

$$V_{bi} = \frac{1}{2}\varepsilon_m W = \frac{1}{2}\varepsilon_m(x_n + x_p), \tag{2.70}$$

where W is the total depleted width. Replacing ε_m (equation 2.68) in equation 2.70, we obtain the relationship between the total depleted region and the acceptor and donor densities, for a two-side abrupt junction

$$W = \sqrt{\frac{2\epsilon_s}{q}(\frac{N_A + N_D}{N_A N_D})V_{bi}}. \tag{2.71}$$

Considering the case of a one-side abrupt junction in which $N_D \gg N_A$ or vice versa, equation 2.71 can be reduced to

$$W = \sqrt{\frac{2\epsilon_s V_{bi}}{qN_B}}, \tag{2.72}$$

where $N_B = N_D$ or N_A. A more accurate result for the depletion layer width can be obtained from equation 2.63 by considering the majority-carrier contribution in addition to the impurity concentration, that is $\rho \approx -q[N_A - p(x)]$ on the p side and $\rho \approx -q[N_D - n(x)]$ on the n side. The depletion width is essentially the same as equation 2.72, except that built-in potential that is replaced by $(V_{bi} - \frac{2kT}{q})$. Therefore, the depletion width at thermal equilibrium for one-side abrupt junction becomes

$$W = \sqrt{\frac{2\epsilon_s}{qN_B}(V_{bi} - \frac{2kT}{q})} = L_D\sqrt{2(\beta V_{bi} - 2)}, \tag{2.73}$$

where $\beta = \frac{q}{kT}$ and L_D is the Debye length, which is a characteristic length for semiconductor. The Debye length is given by

$$L_D \equiv \sqrt{\frac{\epsilon_s kT}{q^2 N_B}} = \sqrt{\frac{\epsilon_s}{qN_B\beta}} \tag{2.74}$$

and it is inversely proportional to the square root of the majority doping density on a one-side abrupt junction. For Si, at thermal equilibrium, the depletion width of an abrupt junction is about eight times the Debye length. The Debye length for silicon at room temperature, as a function of the doping density, is shown in figure 2.6. When an external bias V is applied to the junction, the total potential across the junction is given by $(V_{bi} + V)$ and $(V_{bi} - V)$ for reverse and forward bias, respectively. Substituting these total potentials in equation 2.73, we obtain the depletion width as a function of the applied external bias. The depletion capacitance per unit of area is defined as $\frac{dQ_c}{dV}$ where dQ_c is the incremental variation in charge per unit of area

upon the incremental variation of the applied bias dV. In one-side abrupt junction simplification, the capacitance per unit of area is:

$$\begin{aligned} C \equiv \frac{dQ_C}{dV} = \frac{d(qN_BW)}{d[\frac{qN_B}{2\epsilon_s}W^2]} = \frac{\epsilon_s}{W} &= \sqrt{\frac{q\epsilon_s N_B}{2}}(V_{bi} \pm V - \frac{2kT}{q})^{-\frac{1}{2}} \\ &= \frac{\epsilon_s}{\sqrt{2}L_D}(\beta V_{bi} \pm \beta V - 2)^{-\frac{1}{2}}. \end{aligned} \tag{2.75}$$

Equation 2.75 can be rewritten as

$$\frac{1}{C^2} = \frac{2L_D^2}{\epsilon_s^2}(\beta V_{bi} \pm \beta V - 2), \tag{2.76}$$

$$\frac{d(\frac{1}{C^2})}{dV} \cong \frac{2L_D^2\beta}{\epsilon_s^2} = \frac{2}{q\epsilon_s N_B}, \tag{2.77}$$

where the signs $\pm$ are for reverse and forward bias conditions, respectively. It is clear by equation 2.77 the linear relationship between $\frac{1}{C^2}$ and the bias V. The slope of the plotting on $\frac{1}{C^2}$ as a function of V gives the impurity concentration of the substrate N_B, while the intercept at $\frac{1}{C^2} = 0$ gives the built-in potential of the junction. Equation 2.77 holds for a more general distribution than just for an abrupt junction, and for a general distribution we have

$$\frac{d(\frac{1}{C^2})}{dV} = \frac{2}{q\epsilon_s N(W)}, \tag{2.78}$$

with

$$W = \frac{\epsilon_s}{C(V)}, \tag{2.79}$$

where $N(W)$ is the doping density at $x = W$. By equation 2.74 note that the capacitance-voltage data are insensitive to changes in the doping profiles that occur for a distance shorter than the Debye length of the highly doped side. The doping profile determined by the capacitance-voltage method has a spatial resolution of the

order of the Debye length.

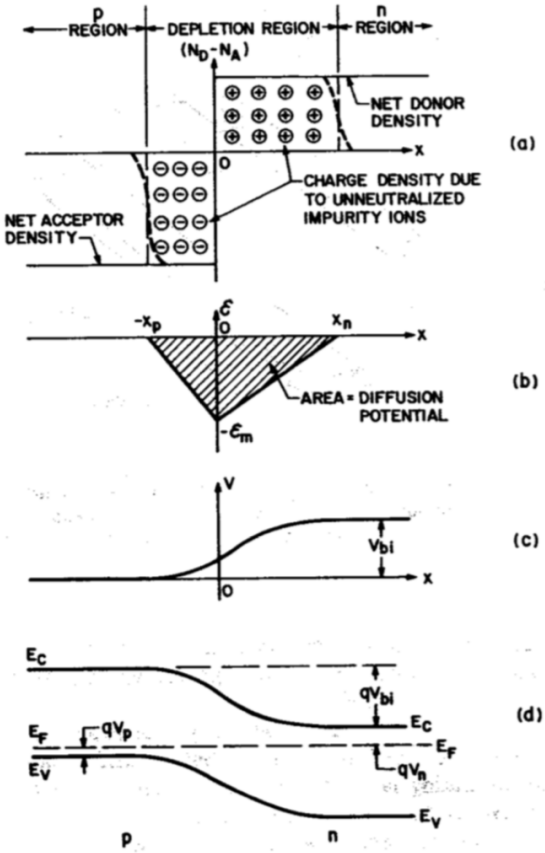

Figure 2.5: *p-n* junction at thermal equilibrium. (a) Space charge distribution; the dashed lines indicate the majority-carrier distribution tails. (b) Electric field profile. (c) Potential variation with the distance, where V_{bi} is the built-in potential. (d) Energy-band diagram. Figure from [7].

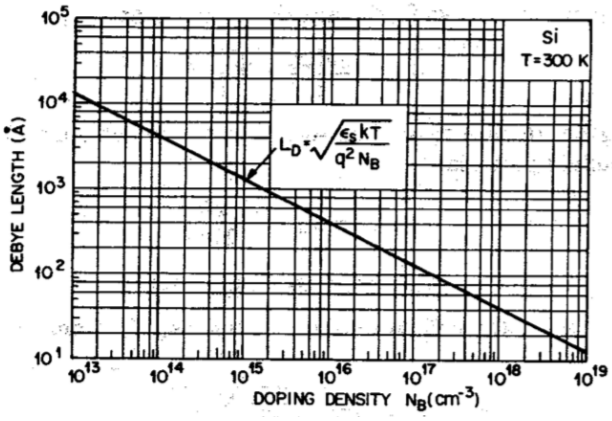

Figure 2.6: Debye length in silicon at room temperature, as a function of the doping concentration. Figure from [7].

2.5.2 Current-Voltage characteristic

The current-voltage characteristic of a p-n junction is based on four assumptions:

- Abrupt depletion layer approximation: built-in potential and applied voltage are supported by a dipole layer with abrupt boundaries, outside the boundaries the semiconductor is assumed to be neutral.
- Boltzmann approximation: through the depletion layer, the Boltzmann relations similar to equations 2.23 and 2.26 are valid.
- Low injection: the injection of minority carrier densities are smaller compared to the majority carrier densities.
- No current generation in the depletion layer: the electron and hole currents are constant through the depletion layer.

We first consider the Boltzmann relation. At thermal equilibrium, this relation is given by

$$n = n_i exp(\frac{E_F - E_i}{kT}) \equiv n_i exp(\frac{q(\psi - \phi)}{kT}), \tag{2.80}$$

and by

$$p = n_i exp(\frac{E_i - E_F}{kT}) \equiv n_i exp(\frac{q(\phi - \psi)}{kT}), \tag{2.81}$$

where ψ and ϕ are the potentials corresponding to the intrinsic level and Fermi level, respectively ($\psi \equiv -E_i/q$ and $\phi \equiv -E_F/q$). At equilibrium the product between p and n is equal to n_i^2. Applying an external bias voltage, the minority carrier densities on both sides of the junction changes, and the relationship $np = n_i^2$ is not anymore valid. we rewrite now the two previous equations as:

$$n \equiv n_i exp(\frac{q(\psi - \phi_n)}{kT}). \tag{2.82}$$

$$p \equiv n_i exp(\frac{q(\phi_p - \psi)}{kT}), \tag{2.83}$$

where ϕ_n and ϕ_p are the quasi-Fermi levels for electrons and holes, respectively. From equations 2.82 and 2.83:

$$\phi_n \equiv \psi - \frac{kT}{q} ln(\frac{n}{n_i}), \tag{2.84}$$

$$\phi_p \equiv \psi + \frac{kT}{q} ln(\frac{p}{n_i}), \tag{2.85}$$

and the pn product becomes

$$pn = n_i^2 exp[\frac{q(\phi_p - \phi_n)}{kT}]. \tag{2.86}$$

For a forward bias $(\phi_p - \phi_n) > 0$ and therefore $np > n_i^2$, whereas for reverse bias $(\phi_p - \phi_n) < 0$ and $np < n_i^2$. From equations 2.51 and 2.82 and from $\varepsilon \equiv - \bigtriangledown \psi$ we can write the electron current density as:

$$\begin{aligned} J_n = q\mu_n(n\varepsilon + \frac{kT}{q} \bigtriangledown n) = q\mu_n n(- \bigtriangledown \psi) + q\mu_n \frac{kT}{q}[\frac{qn}{kT}(\bigtriangledown\psi - \bigtriangledown\phi_n)] = \\ -q\mu_n n \bigtriangledown \phi_n. \end{aligned} \tag{2.87}$$

Similarly, for hole we obtain a current density

$$J_p = -q\mu_p p \bigtriangledown \phi_p. \tag{2.88}$$

Therefore, the electron and hole current densities are proportional to the gradient of the quasi-Fermi-level of electrons and holes, respectively. At thermal equilibrium $\phi_n = \phi_p = \psi = const$, then J_n and J_p are equal to zero. Since the electron density n across the junction varies of many orders of magnitude from the n side to the p side, while the electron current density J_n is constant than ϕ_n must also be constant

over the depletion region. The electrostatic potential difference across the junction is given by

$$V = \phi_p - \phi_n. \tag{2.89}$$

Combining equation 2.86 with 2.89 we obtain the electron density at the boundary of the depletion region on the p side ($x = -x_p$):

$$n_p = \frac{n_i^2}{p_p} exp(\frac{qV}{kT}) = n_{p0} exp(\frac{qV}{kT}), \tag{2.90}$$

where n_{po} is the equilibrium electron density on the p side. Similarly, we obtain the hole density at the boundary of the depletion region on the n side ($x = x_n$)

$$p_n = p_{no} exp(\frac{qV}{kT}). \tag{2.91}$$

The previous equations are the boundary conditions for the ideal current-voltage equation. From continuity equations we obtain, for the steady state, these equations:

$$-U + \mu_n \varepsilon \frac{\partial n_n}{\partial x} + \mu_n n_n \frac{\partial \varepsilon}{\partial x} + D_n \frac{\partial^2 n_n}{\partial x^2} = 0, \tag{2.92}$$

$$-U - \mu_p \varepsilon \frac{\partial p_n}{\partial x} - \mu_p p_n \frac{\partial \varepsilon}{\partial x} + D_p \frac{\partial^2 p_n}{\partial x^2} = 0, \tag{2.93}$$

where U is the net recombination rate, and the charge neutrality condition is valid ($n_n - n_{no} \approx p_n - p_{no}$). Multiplying equation 2.92 by $\mu_p p_n$ and equation 2.93 by $\mu_n n_n$, and combining with the Einstein relation, $D = \frac{kT}{q}\mu$ we obtain

$$-\frac{(p_n - p_{no})}{\tau_a} + D_a \frac{\partial^2 p_n}{\partial x^2} - \frac{(n_n - p_n)}{(\frac{n_n}{\mu_p} + \frac{p_n}{\mu_n})} \frac{\varepsilon \partial p_n}{\partial x}, \tag{2.94}$$

where

$$D_a = \frac{n_n + p_n}{n_n/D_p + p_n/D_n}, \tag{2.95}$$

$$\tau_a = \frac{p_n - p_{no}}{U} = \frac{n_n - n_{no}}{U}, \tag{2.96}$$

From low-injection assumption ($p_n \ll n_n \approx n_{no}$ in the n side region) equation 2.94 becomes

$$-\frac{p_n - p_{no}}{\tau_p} - \mu_p \varepsilon \frac{\partial p_n}{\partial x} + D_p \frac{\partial^2 p_n}{\partial x^2} = 0. \tag{2.97}$$

In the neutral region the electric field is null, then equation 2.97 reduces to

$$\frac{\partial^2 p_n}{\partial x^2} - \frac{p_n - p_{no}}{D_p \tau_p} = 0. \tag{2.98}$$

The solution of equation 2.98 with the boundary condition equation 2.91 and $p_n(x = \infty) = p_{no}$ gives

$$p_n - p_{no} = p_{no}(e^{\frac{qV}{kT}} - 1)e^{-\frac{(x - x_n)}{L_p}}, \tag{2.99}$$

where

$$L_p \equiv \sqrt{D_p \tau_p} \tag{2.100}$$

and at $x = x_n$

$$J_p = -qD_p \frac{\partial p_n}{\partial x}\Big|_{x_n} = \frac{qD_p p_{no}}{L_p}(e^{qV/kT} - 1). \tag{2.101}$$

Similarly, for the p side, we obtain

$$J_n = qD_n \frac{\partial n_p}{\partial x}\Big|_{-x_p} = \frac{qD_n n_{po}}{L_n}(e^{qV/kT} - 1). \tag{2.102}$$

The total current is given by the sum of equations 2.101 and 2.102

$$J = J_p + J_n = J_s((e^{qV/kT} - 1)), \tag{2.103}$$

where

$$J_s \equiv \frac{qD_p p_n o}{L_p} + \frac{qD_n n_p o}{L_n}. \tag{2.104}$$

Equation 2.104 is known as the Shockley equation or ideal diode law. Figure 2.7 (left) shows the current-voltage characteristic curve of an ideal diode. For the silicon, the Shockley equation gives only a qualitative agreement with the current-voltage curve of a real p-n device, the departure from the ideal behaviour is due to the surface effect, the generation and recombination of carriers in the depletion region, the tunneling of carriers between states in the band gap, the high-injection condition that can occur at relatively small forward bias and the series resistance effect. Beside, in high electric field condition, in reverse bias region can occur a breakdown of the junction (for example as a result of avalanche multiplication). The surface effects are due to charges on or outside the semiconductor surface, that induce charges into the semiconductor. These charges create surface channels or surface depleted regions. For silicon, the leakage current generated by surface effects is much smaller than leakage current generates by generation-recombination processes in the depletion region. Consider first the generation current under reverse bias condition, the rate of generation of electron-hole pairs under the conditions $p < n_i$ and $n < n_i$ is

$$U = -[\frac{\sigma_p \sigma_n v_{th} N_t}{\sigma_n exp(\frac{E_t - E_i}{kT}) + \sigma_p exp(\frac{E_i - E_t}{kT})}]n_i \equiv -\frac{n_i}{\tau_e}, \tag{2.105}$$

where σ_p and σ_n are the hole and electron capture cross sections, v_{th} is the carrier thermal velocity, N_t is the trap density, E_t is the trap energy level and τ_e is the effective lifetime. The current due to the generation process in the depletion region is given by

$$J_{gen} = \int_0^W q|U|\, dx \simeq q|U|W = \frac{qn_iW}{\tau_e}, \tag{2.106}$$

where W is the width of the depletion region. At a given temperature J_{gen} is proportional to the depletion layer width, which is proportional to the square root of the reverse bias. The total reverse current can be obtained by the sum of diffusion current in the neutral region of the junction and generation current in the depletion

region

$$J_R = q\sqrt{\frac{D_p}{\tau_p}}\frac{n_i^2}{N_D} + \frac{qN_iW}{\tau_e}. \tag{2.107}$$

For silicon, that has a small n_i, the generation current dominates the diffusion one. A typical result of silicon is shown in figure 2.7 (right). In forward bias condition, where in the depletion region dominates the capture process we have a recombination current in addition to the diffusion current. The depletion layer capacitance considered previously accounts for most of the junction capacitance when the junction is reverse-biased. In forward bias condition, there is a significant additional junction capacitance from rearrangement of minority carrier density, this contribution is called diffusion capacitance.

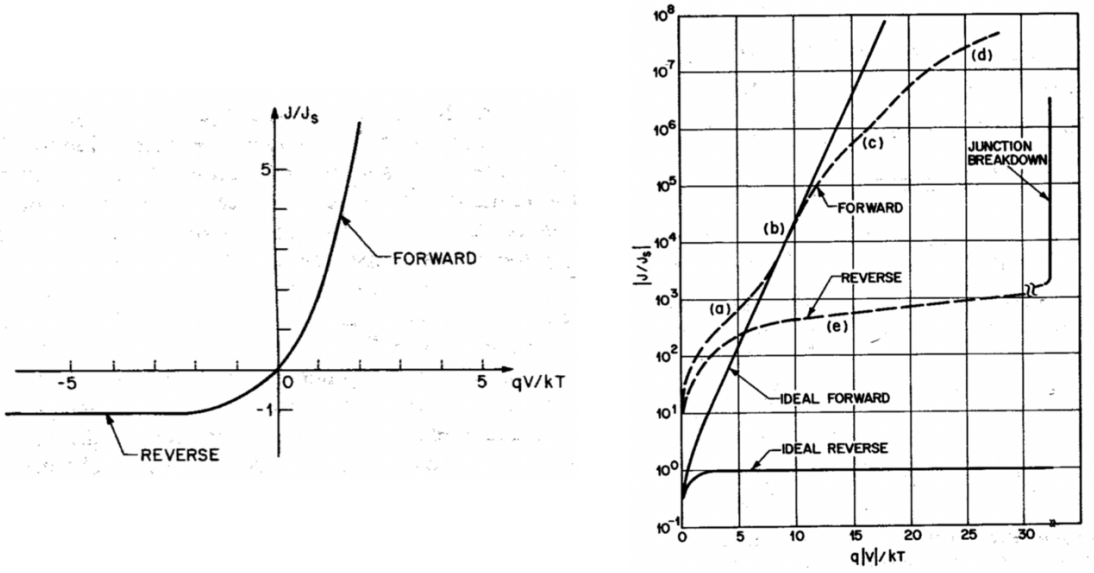

Figure 2.7: Left: current-voltage characteristic of the ideal diode. Right: current-voltage characteristic of a real diode. (a) Generation-recombination current region. (b) Diffusion current region. (c) High injection region. (d) Series resistance effect. (e) Leakage current in the reverse region due to generation-recombination and surface effects. Figure from [7].

2.5.3 Junction breakdown

In a *p*-*n* junction reverse biased and in high field condition, the junction "breakdown" can occur and a large current flux passes through the junction. The breakdown is due to three mechanisms: thermal instability, tunneling effect, and avalanche multiplication. In this section, we will only discuss the avalanche multiplication mechanism.

The avalanche multiplication or impact ionization is the most important mechanism in junction breakdown. We first consider a hole current I_{po} on the left-hand side ($x = 0$) of the depletion region with width W. If the electric field in the depletion region is high enough than pairs electron-hole are generated by impact ionization process and the hole current I_p will increase with distance through the depletion region. the current I_p reaches the maximum value of $G_p I_{po}$ at $x = W$. Similarly, an electron current I_n will increase from $x = W$ to $x = 0$. The total current $I = I_p + I_n$ is constant in a steady condition. The incremental hole current at x is equal to the number of electron-hole pairs generated per second in the distance dx.

$$d(\frac{I_p}{q}) = (\frac{I_p}{q})(\alpha_p dx) + (\frac{I_n}{q})(\alpha_n dx) \tag{2.108}$$

or

$$\frac{dI_p}{dx} - (\alpha_p - \alpha_n)I_p = \alpha_n I, \tag{2.109}$$

where α_p and α_n are the hole and electron ionization rates. The term G_p is the multiplication factor of holes and is defined as

$$G_p \equiv \frac{I_p(W)}{I_p(0)}. \tag{2.110}$$

The avalanche breakdown voltage is defined as the voltage for which the G_p factor approaches infinity. As discussed previously, the electric field and the potential in the depletion region are determined from the solutions of Poisson's equations, knowing

the boundary conditions we can obtain the breakdown voltage of one-side abrupt junction.

$$V_B(breakdown\ voltage) = \frac{\varepsilon_m W}{2} = \frac{\epsilon_s \varepsilon_m^2}{2q}(N_B)^{-1}, \tag{2.111}$$

where N_B is the ionized backgroung impurity concentration, ϵ_s is the semiconductor permittivity and ε_m is the maximum electric field. For the Si junction the maximum field value is

$$\varepsilon_m = \frac{4 \times 10^5}{1 - \frac{1}{3} log_{10}(N_B/10^{16})}, \tag{2.112}$$

where N_B is in cm^{-3}.

Chapter 3

Operating principle of a silicon detector

The operating principle of an n-on-p pad silicon detector is shown in figure 3.1. An external bias voltage polarizes the p-n junction inversely, creating a large depleted volume. When a particle crosses the sensor, it creates along its path electron-hole (e-h) pairs, whose number depends on the particle type, energy, and on the sensor thickness. Under the influence of the electric field in the depletion region, the electrons drift towards the n^{++} contact and the holes towards the p^{++} contact, inducing a current signal on the electrodes. This current signal starts when the impinging particle passes through the sensor and it ends when the last charge carrier is collected by the electrodes.

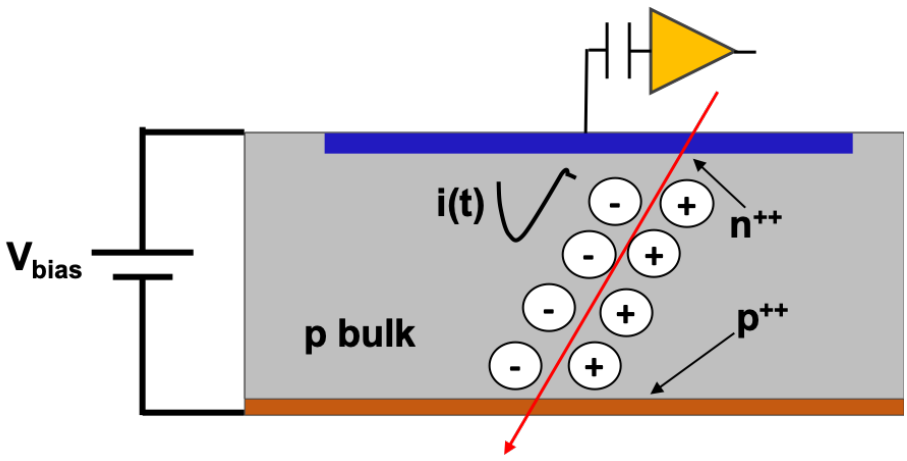

Figure 3.1: Basic operation principle of an n-on-p silicon detector.

3.1 Particle interaction in silicon

Charged particles have electromagnetic interaction with the electrons of the silicon atoms, losing energy gradually and causing two different processes: atomic excitation, displacing an electron to a higher atomic orbital; ionization, producing an electron-ion

pair if the transferred energy is higher than the ionization potential of the silicon. The ionization is the predominant process in silicon sensors, whose operation is mainly based on this effect. The average energy lost per unit length by a particle crossing the matter is called Stopping Power, and it is given with good approximation by the Bethe-Bloch formula [9]:

$$-\frac{dE}{dx} = 4\pi N_A r_e^2 m_e c^2 \rho \frac{Z}{A}\frac{z^2}{\beta^2}[\frac{1}{2}\ln\frac{2m_e c^2\beta^2\gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta}{2}], \tag{3.1}$$

where E is the kinetic energy of the impinging particle of charge z, with moving velocity $\beta = v/c$ and Lorentz factor $\gamma = 1/\sqrt{1-\beta^2}$. I is the mean excitation energy of the target material, characterized by density ρ, and atomic and mass number Z and A. T_{max} is the maximum kinetic energy that could be transferred to a free electron in a single collision. N_A, r_e and m_e are the Avogadro Number, the classic radius of the electron, and the mass of the electron, respectively. At last, δ is the high-energy corrective term for density. According to Bethe-Bloch formula, the ionization loss is proportional to electron density in the medium $\rho Z N_A/A$, to the square of the particle charge, and it strongly depends on the incident particle velocity; at low momenta, the energy loss decreases proportionally to $1/\beta^2$. Figure 3.2 shows the energy loss for a positive muon in copper, as a function of its parameter $\beta\gamma$. Figure 3.3, instead, shows the same curve for different particles in different materials; the stopping power on the y-axis is divided by material density so that it is independent of the medium crossed.

A particle whose energy loss is at the minimum of the Bethe-Bloch function is called Minimum Ionizing Particle (MIP). In silicon, the minimum of ionization occurs at $\beta\gamma = 3$, which corresponds to a stopping power per density unit of about 1.66 $MeVcm^2/g$.

The energy lost by a particle follows the Landau distribution, which is an asym-

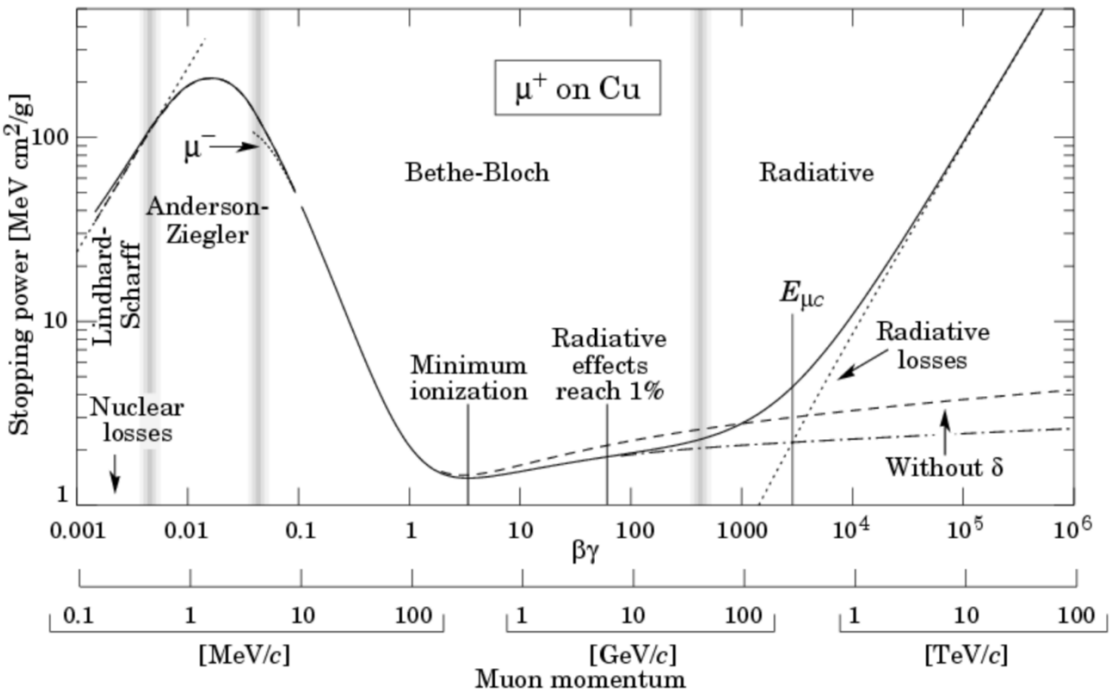

Figure 3.2: The Bethe-Bloch curve for positive muons in copper as a function of $\beta\gamma$.

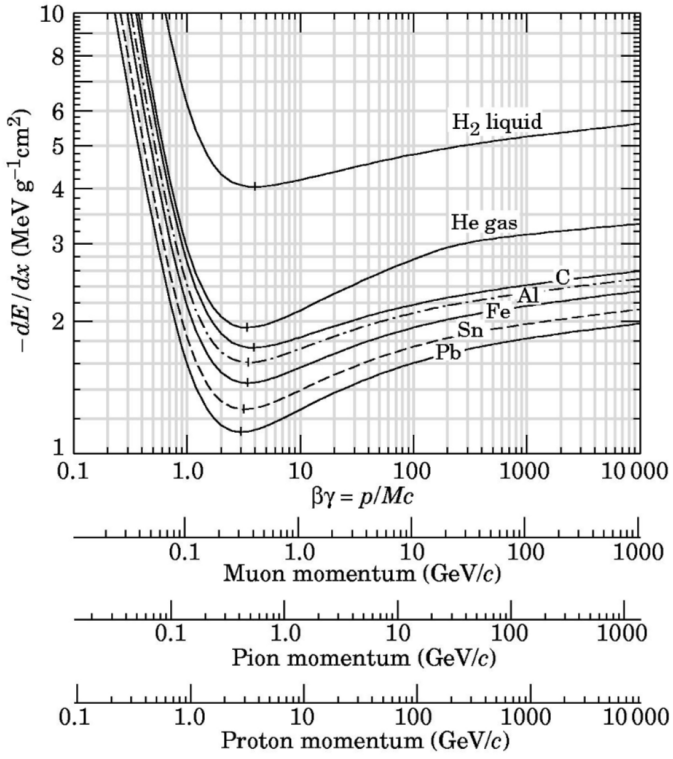

Figure 3.3: The Bethe-Bloch curves for heavy charged particles in different materials.

metric distribution with a not negligible tail at high energies due to delta rays [10]. The delta rays appear when a particle loses a large amount of its energy during an interaction and the electrons produced have enough energy to ionize other atoms. Due to this asymmetry, the mean value of the distribution does not match with the most probable value (MPV), which is 30% lower.

A MIP crossing a silicon detector usually produces, by ionization, 73 electron-hole pairs per micron of thickness; the energy needed to produce a single pair is 3.6 eV, about three times the band gap of silicon since there is a lot of energy lost in lattice oscillations (phonons). Overall a MIP, crossing silicon, loses a small amount of its energy, ensuring the complete crossing of the active thickness of the detector.

Light charged particles such as electrons and positrons behave differently compared to heavy particles; for these particles, another process of energy loss must be considered, the Bremsstrahlung, negligible for heavy particles. The Bremsstrahlung is the emission of photons by electrons and positrons when they are decelerated by the electromagnetic field of nuclei. This contribution must be considered in the total energy loss:

$$(\frac{dE}{dx})_{tot} = (\frac{dE}{dx})_{Brems} + (\frac{dE}{dx})_{Ion}, \tag{3.2}$$

where the contribution of Bremsstrahlung can be written as:

$$-\frac{1}{\rho}(\frac{dE}{dx})_{Brems} = \frac{E}{X_0}, \tag{3.3}$$

where X_0 is the radiation length, that is the distance in the medium needed to reduce the electron energy by a factor e.

Finally, it is interesting to mention the interaction of α particles. The α particles are very heavy, the stopping power is located in the first part of the Bethe-Bloch curve where the energy loss is very strong. Therefore the α particles lose all their energy in a few microns of the silicon detector.

3.2 Signal formation in silicon detector: Shockley-Ramo's theorem

The current signal induced in a parallel plate detector can be described simply considering a charge q moving in the electric field generated in the detector by two electrodes, to which a bias voltage is applied. The charge q induces charges of opposite sign on both electrodes, which are a fraction of q sot that the total sum of the induced charges on both electrodes is equal to $-q$. The induced charge on an electrode varies in time since the charge q is drifting into the volume of the sensor, therefore also the electrostatic flux on a Gaussian surface surrounding the electrode varies in time, generating a current signal. To evaluate the current signal we consider a system of n-electrode, to each of which a certain potential is applied. An electric field is generated in the volume enclosed by the electrodes; a free charge in this volume starts to drift under the effect of this electric field. The induced current on the k-th electrode is given by:

$$i_k(t) = \frac{d}{dt}[\oint_{\Sigma_k} \epsilon \cdot \vec{E}(\vec{r}, t) \cdot d\Sigma]. \tag{3.4}$$

To know the induced current we need to know the electric field, which is not an easy task; Shockley-Ramo's theorem simplifies it. Two possible sets, A and B, of the above electrodes system, are considered; the Green's reciprocity theorem states that: $\sum_j Q_j^A V_j^B = \sum_j Q_j^B V_j^A$, where Q and V are the induced charge and the electrode potential, respectively. If set A includes all the electrodes grounded and a charge q_p present among them, while, set B includes q_p removed and the k-th electrode at a potential V_k^B (all the other electrodes are grounded), the Green's theorem states:

$$q_p \cdot V_p^B + Q_k^A \cdot V_k^B = 0, \tag{3.5}$$

$$Q_k^A = -q_p \frac{V_p^B}{V_k^B} = -q_p \cdot \tilde{V}_w. \tag{3.6}$$

This result states that the induced charge on the k-th electrode has the opposite sign of the charge q_p generated in the sensor bulk, and it is a fraction of q_p given by the adimensional ratio V_w, between the potential of the point where q_p is and the potential of the k-th electrode. The induced current on the k-th electrode is the temporal derivative of Q_k^A, written as $Q_k(t)$.

$$i_k(t) = \frac{dQ_k}{dt} = -q_p \vec{\nabla} \tilde{V}_w \cdot v(\vec{x}(t), \vec{y}(t), \vec{z}(t)) = -q_p \vec{E}_w \vec{v}, \tag{3.7}$$

where $\vec{v}$ is the drift velocity of the charge and $\vec{\nabla} \tilde{V}_w = \vec{E}_w$ is the weighting field. Equation 3.7 is known as Shockley-Ramo's theorem [11, 12]. The weighting field describes the coupling between the charge q_p and the k-th electrode when all the other electrodes of the system are grounded and the k-th electrode is at 1 V of potential; its dimension is the reciprocal of a length $[m^{-1}]$. The weighting field is crucial for signal formation in a silicon detector, indeed if we consider a fixed charge generated by an impinging particle in the sensor volume and a saturated drift velocity of this charge, the weighting field defines the induced current signal.

3.3 Radiation damage in silicon devices

The radiation damage in silicon sensors can be classified in surface and bulk damage, the second one is the limiting factor for the use of silicon detectors in high radiation regions of HEP experiments.

The radiation damage mechanism is based on the interaction between high energy particles (hadrons, leptons, and photons) and silicon atoms of the crystal lattice. This interaction produces displaced atom from lattice site resulting in a silicon interstitial and vacancy, called Frenkel pair. A part of Frenkel pairs annihilate causing no dam-

ages, the remaining interstitials and vacancies migrate through the lattice reacting with other impurity atoms existent in the silicon and producing point defects. However, atomic displacement occurs if the imparted energy by the impinging particle is higher than the displacement threshold energy E_d ($\sim 25\ eV$); therefore, the displaced atom could have a very high energy producing consequently ionization and further atomic displacements. At the end of the recoil range, non-ionizing reactions prevail, which produce a dense agglomeration of defects called cluster, figure 3.4. Point and cluster defects are responsible for the macroscopic deterioration of the detector properties.

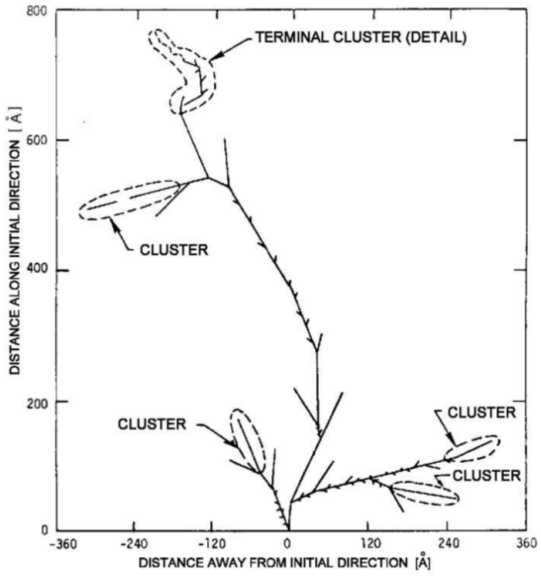

Figure 3.4: Monte Carlo simulation of a recoil atom track with a primary energy E_R of 50 keV. Figure from [8].

3.3.1 Non Ionizing Energy Loss scaling hypothesis

Different kinds of particles and different kinds of interactions, depending on the particle energy, produce atomic displacements; the Non Ionizing Energy Loss (NIEL) hypothesis parametrizes how the silicon bulk damage scales with the energy imparted in displacing collisions. The basic assumption of the NIEL is that the displacement-damage induced in the material is linear with the amount of energy imparted in

displacing collisions. In each primary interaction a displaced atom, with a specific recoil energy E_R, is produced; the fraction of this recoil energy that is deposed in form of displacement damage depends on E_R itself and can be analytically calculated by Lindhard partition function $P(E_R)$, [13]. Therefore, thanks to this partition function, the NIEL is expressed by the displacement damage cross section

$$D(E) = \sum_{\nu} \sigma_{\nu}(E) \cdot \int_0^{E_R^{max}} f_{\nu}(E, E_R) P(E_R) dE_R \tag{3.8}$$

where ν indicates all the possible interactions, between impinging particle of energy E and silicon atoms, those produce atomic displacements in the silicon lattice. σ_{ν} is the cross section corresponding to ν interaction, and f_{ν} is the generation probability of a displaced atom with a recoil energy E_R by a particle with energy E, for an interaction type ν. Figure 3.5 shows the displacement damage cross section for neutrons, protons, pions, and electrons in an energy range between few meV and 10 GeV. The damage induced by low energy protons ($< GeV$) is much larger than that for neutrons, since it is dominated by Coulomb interaction; instead at higher energy ($\sim GeV$) the Coulomb contribute becomes very small and the nuclear reactions for protons and neutrons are practically the same, as the result both damage function approach to almost a common value.

Using the displacement damage cross section is possible to define the hardness factor κ, allowing to compare the damage efficiency produced by particles of different kind and energy. It is custom to define this factor by comparing the damage produced by a specific irradiation fluence with the damage which would have been produced by 1 MeV neutrons at the same fluence

$$\kappa = \frac{\int D(E)\phi(E)dE}{D(E_n = 1MeV) \cdot \int \phi(E)dE} \tag{3.9}$$

where $D(E_n = 1MeV)$ is set to 95 $MeVmb$. The equivalent 1 MeV neutron fluence

can be calculated as the product of the hardness factor with the effective irradiation fluence.

$$\phi_{eq} = \kappa\phi = \kappa \int \phi(E)dE \tag{3.10}$$

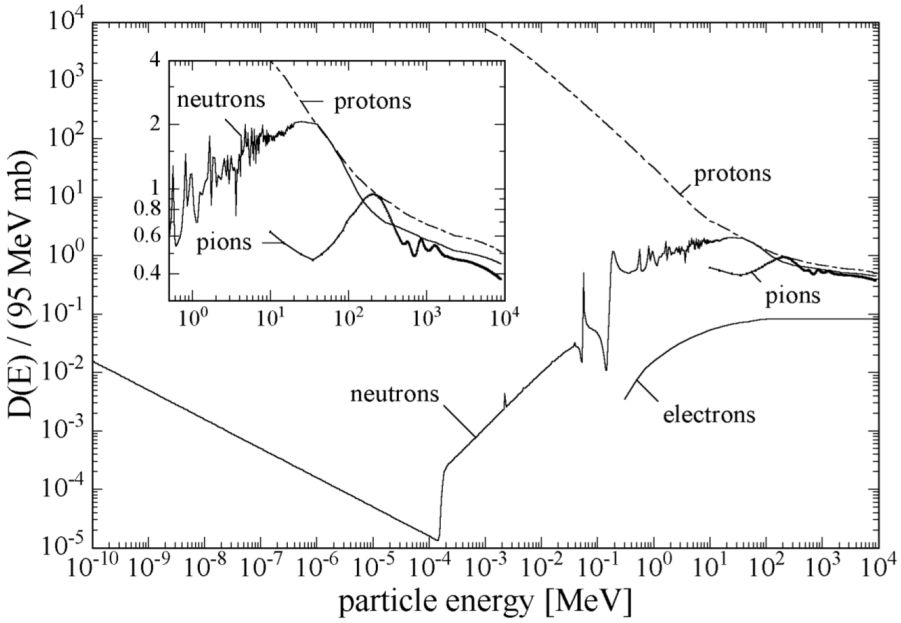

Figure 3.5: Displacement damage cross section $D(E)$ normalized to 95 $MeVmb$ for neutrons, protons, pions, and electrons. Due to the normalization to 95 $MeVmb$ the ordinate axis represents the damage equivalent to 1 MeV neutron. Figure from [8].

3.3.2 Impact of defects on silicon detector properties

As already mentioned in this section, radiation generates point defects and clusters; in this subsection, the impact of these defects on the properties of silicon detectors will be mentioned with special attention to effects that are relevant for UFSDs.

Radiation creates defects as silicon interstitials (Si_i or I), vacancies (V), and clusters, however, with the creation of these defects the radiation damage does not end. These defects are very mobile at room temperature and migrate through the silicon lattice, reacting with impurities and giving rise to further defects. One of the most interesting reactions happens between Si_i and substitutional impurities such

as B_s and C_s. This reaction, called Watkins replacement mechanism, equation 3.11, can eject these impurities from their substitutional sites making them interstitial, electrically active, and mobile inside the silicon lattice.

$$\begin{aligned} Si_i + C_s &\rightarrow Si_s + C_i, \\ Si_i + B_s &\rightarrow Si_s + B_i. \end{aligned} \tag{3.11}$$

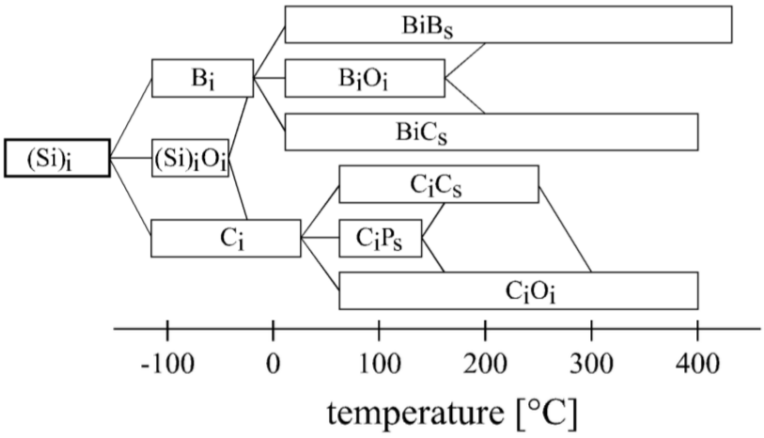

Figure 3.6: Silicon interstitial reactions with impurities and defects stability as a function of temperature. Figure from [8].

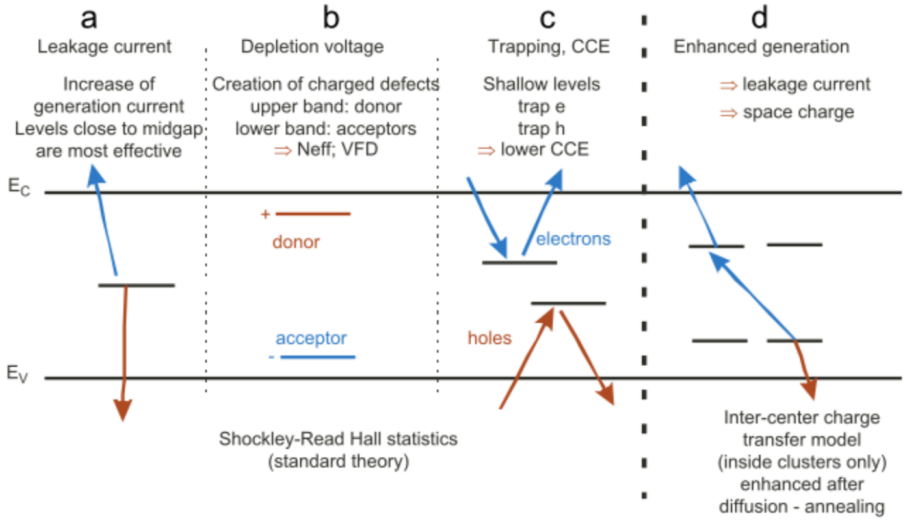

Figure 3.7: Different defect level locations in band gap and their effects on macroscopic properties of detector.

Figure 3.6 shows schematically the formation of B_i and C_i, and their reactions with interstitial oxygen (O_i). The typology of the defect and its level position in the band gap acts on the macroscopic properties of the detector (resistivity, leakage current, depletion voltage, charge collection efficiency, etc.), figure 3.7. Below a brief description of the effects due to radiation on the macroscopic properties more interest for this work, [14]:

- **Leakage current (I_{bulk}):** defects close to the middle of the band gap are generator centers of electron-hole pairs, hence responsible for the increase of the current. The leakage current is often described by the factor called generator lifetime (τ_g)

$$I_{bulk} = V q_0 \frac{n_i}{\tau_g} \tag{3.12}$$

  where V is the depleted volume of the sensor and n_i the intrinsic carrier density. In irradiated sensors the increase in current ($\triangle I$) given by the difference between the leakage currents measured after ($I_{bulk}(\phi)$) and before ($I_{bulk}(0)$) irradiation is proportional to the irradiation fluence, equation 3.13, figure 3.8. $\triangle I$ can be approximated with $I_{bulk}(\phi)$, since $I_{bulk}(0)$ is negligible compared with the current after irradiation, for irradiation fluences above 10^{12} n_{eq}/cm^2.

$$\triangle I = V \alpha \phi \sim I_{bulk}(\phi), \tag{3.13}$$

  where the proportionality factor $\alpha \sim q_0(n_i/\tau_g)\phi$ is called current-related damage rate.

- **Depletion voltage (V_{FD}):** defect levels close to the conduction and valence bands contribute to modify the space charge density (N_{eff}), affecting the depletion voltage of the sensor, equation 2.75.There are two mechanisms that act on the space charge density: the acceptor creation and the donor/acceptor removal. Acceptor creation is due to defects which act as acceptors, and increase N_{eff} in

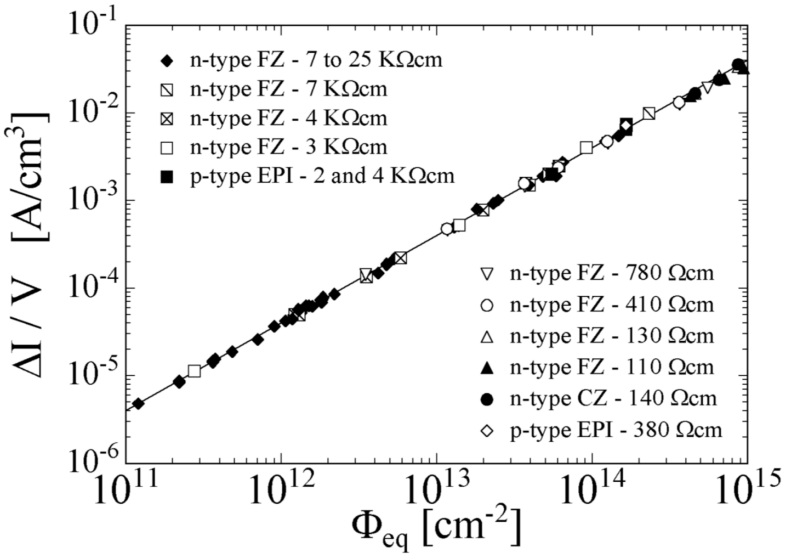

Figure 3.8: Dependence of the leakage current on the irradiation fluence for silicon detectors produced by various technological processes, after an annealing treatment of 80 min at 60 $°C$. Figure from [8]

.

p-type bulk and cause type inversion in n-type bulk, shown in figure 3.9 (left). Donor/acceptor removal mechanism is not fully understood, but experimental pieces of evidence suggest the creation of ion-complexes which react with substitutional donors and acceptors removing them from their lattice sites; thus the effective dopant concentration change, in other words, donors and acceptors have been removed. The space charge density can be parameterized as:

$$N_{eff} = N_{D_0} e^{-c_D \phi_{eq}} - N_{A_0} e^{-c_A \phi_{eq}} - g_{eff} \phi_{eq}, \tag{3.14}$$

where N_{D_0/A_0} are the initial donor/acceptor concentration, $c_{D/A}$ are the donor/acceptor removal rate, and $g_{eff} = 0.02\ cm^{-1}$ is the term of acceptor creation; a different sign is attributed to donors and acceptors. The donor/acceptor removal coefficients are not a constant, they are a function of initial donor/acceptor density:

$$c_{D/A}(N_{D/A_0}) = \alpha (\frac{N_{D/A_0}}{N_0})^{-\beta}, \tag{3.15}$$

where N_0 is a constant term and $-\alpha$, β are empirical coefficients. Figure 3.9 (right) shows the experimental measurement of acceptor removal rate as a function of initial acceptor density; an important observation on c_A is that the acceptor removal rate is less important at high initial doping concentration, the same consideration applies to donor removal rate.

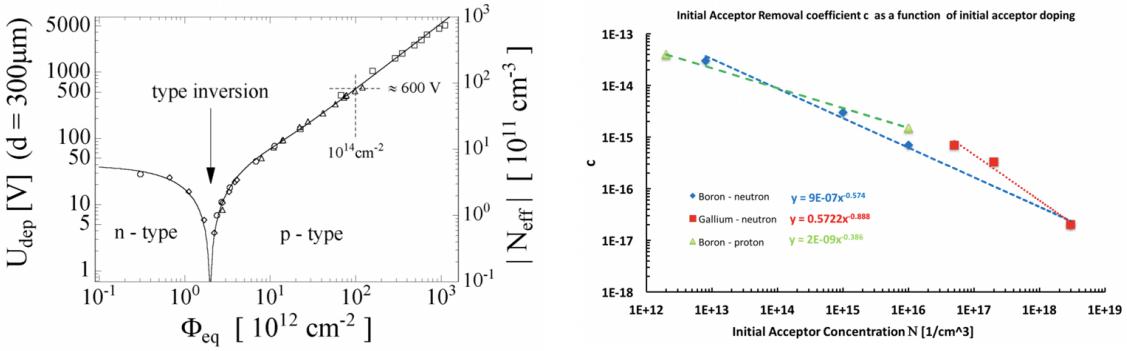

Figure 3.9: Left: depletion voltage and space charge density variation measured after irradiation and annealing treatment for 80 min at 60 $°C$. Figure from [8] .Right: boron and gallium removal coefficient as a function of the initial boron and gallium concentration for neutron and proton irradiation.

- **Charge Collection Efficiency (CCE):** electrons and holes created by ionizing particles into the depleted volume of the sensors drift towards the electrodes, this mechanism is called Charge Collection (CC). Some of these charge carriers are trapped by deep defects along their drift path, this effect is called trapping. As a result, the current signal in an irradiated sensor follows an exponential decrease as a function of time

$$I = I_0 e^{-\frac{t}{\tau_{eff}}}. \tag{3.16}$$

Therefore longer is the drift time, higher is the probability that a charge carrier is trapped. The time constant τ_{eff} is given by

$$\frac{1}{\tau_{eff}} = \phi_{eq}\beta_{e/h}(t), \tag{3.17}$$

therefore, it is a function of the irradiation fluence and the $\beta_{e/h}$ parameter. Moreover, if the de-trapping time is longer than the shaping time of the readout electronic, than the trapping will result in a decrease of the CCE (i.e. the ratio of the number of collected charges over the number of generated ones). Figure 3.10 shows the effect of trapping on the collected charge for a 300 μm thick n-on-p sensors as a function of the irradiation fluence [15].

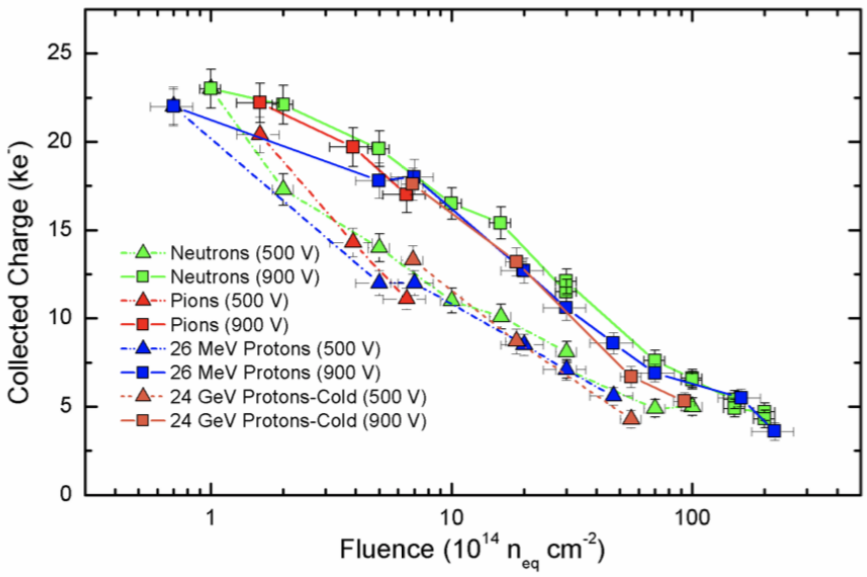

Figure 3.10: Collected charge for 300 μm thick n-on-p sensors as a function of irradiation fluence. Figure from [15].

An important technique that helps to partially recover the irradiated sensor performances is the annealing. Annealing consists of heating a sample at a temperature usually below 100 $°C$, for periods that vary from tenths minute to tenths hours. The provided thermal energy ($E \sim kT$) at lattice reduces considerably part of the radiation damages, improving the sensor performances. The current-related damage rate, the space charge density, and the trapping time constant depend on annealing temperature and time. The current-related damage rate decrease with the increase of annealing temperature and time, figure 3.11 (left); figure 3.11 (right) shows the variation of space charge density with annealing time, for treatment at 60 $°C$; at last,

annealing also has benefits on CCE, resulting in a lower trapping probability [16].

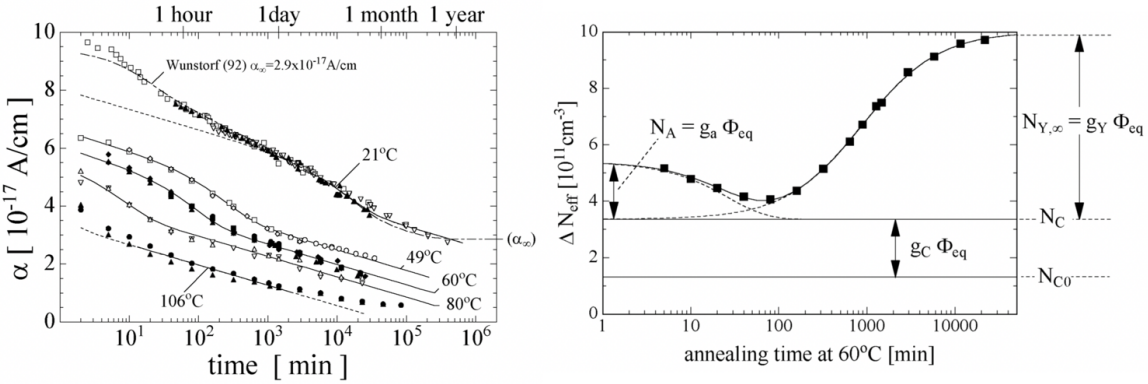

Figure 3.11: Left: current related damage rate as a function of annealing time at different temperatures. Right: variation in effective doping concentration as a function of the annealing time in irradiated sensor annealed at 60 $°C$. Figure from [8].

Chapter 4

Ultra Fast Silicon Detector

Ultra Fast Silicon Detector is a new silicon sensor design able to perform time and space measurements with excellent accuracy; hence, this design is suitable for 4 dimensions tracking. The objective of the UFSD project is to achieve a time resolution of $\sim 30\ ps$ and a spatial resolution of $\sim 10\ \mu m$ due to the combination of a moderate intrinsic gain of ~ 20 implemented in a thin sensor of $\sim 50\ \mu m$ thickness and the possibility of a fine segmentation of the sensor.

In this chapter we will discuss, the main ingredient for a good time resolution measurement, the low Gain Avalanche Diode technology, at the basis of UFSDs and how as this technology has been optimized for time measurements. Afterward, we will describe the UFSD segmentation technology required for the design of large area sensors for tracking in 4 dimensions. Finally, in the last part of the chapter, we will discuss the radiation damage in UFSD, possible technological solutions to improve the radiation hardness, and the ability of performance recovery. The radiation hardness is a key point to operate UFSD in harsh environments such as the HL-LHC scenario.

All the topics covered in this chapter are detailed in [5, 17, 18, 19, 20, 21].

4.1 LGAD technology

Charge multiplication is a well understood mechanism in solid state sensors and it is based on the avalanche process initiated by a charge moving in a high electric field. The gain is defined as the ratio between the number of electron/hole pairs produced on the number of electron/hole pairs produced in the absence of multiplication ($G = N_{e,h}/N_{0;e,h}$). The avalanche process is widely used in semiconductor sensors as Avalanche Photon Diodes (APD) and Silicon Photon Multipliers (SiPM) [22], with a

gain of about 100 and 10000, respectively. The Low Gain Avalanche Diode (LGAD) is a new development of silicon sensor, which merges the best characteristics of standard silicon sensors with those of APDs and SiPMs.

In silicon sensors, the charge multiplication happens when the charge carriers cross a region with an electric field (ε) greater than of 300 kV/cm. Under this condition, the electrons (and to less extent the holes) acquire enough kinetic energy to produce, by impact ionization, additional e/h pairs. In standard silicon sensor (PiN) high value of the electric field is obtained exclusively applying a high external bias voltage, which would cause a breakdown in the device with certainty; in LGAD sensors, the value of $\sim$ 300 kV/cm can be obtained by implanting an additional doping layer of p^+ material (boron or gallium) close to the n-p junction. A schematic drawing of a standard silicon sensor n-in-p and LGAD are shown in figure 4.1. The additional p^+ layer, a few microns wide, is characterized by an appropriate acceptor density ($N_A \sim 10^{16}/cm^3$), that generates locally an electric field high enough to activate the avalanche process. This additional layer is implanted in a high resistivity p-bulk (float zone or epitaxial) with an acceptor concentration of about 10^{11}-$10^{12}/cm^3$.

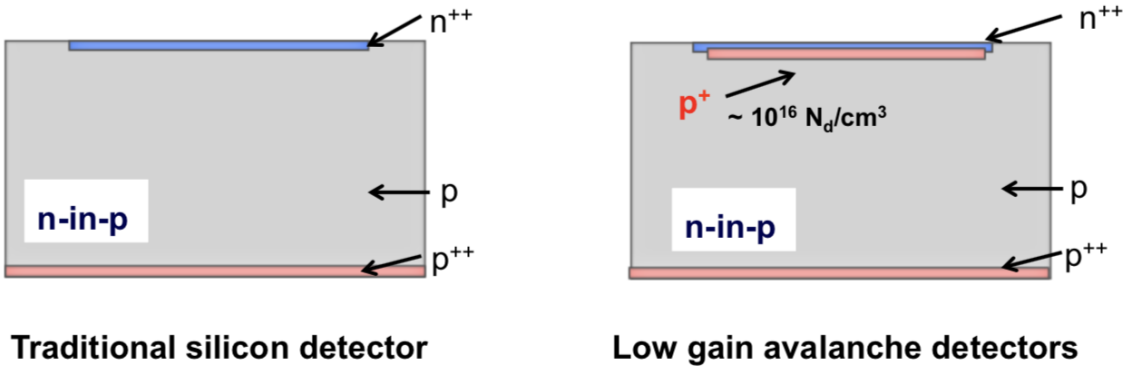

Figure 4.1: Left: traditional silicon sensors n-in-p (no gain). Right: Low Gain Avalanche Diode with the additional p^+ layer close to the n-p junction. Figure from [19].

Figure 4.2 shows the electric field in a 300 μm thick LGAD at three different bias voltage (50 V, 200 V, and 600 V) and in a PiN diode at 600 V, in linear (left) and

logarithmic (right) scale. The electric field in LGAD is divided into three different regions: the drift volume is the region with a rather low electric field ($\varepsilon \sim 30\ kV/cm$); the thin multiplication region, of depth of few microns, with a very high electric field ($\varepsilon \sim 300\ kV/cm$); the zero-field region where the electric field drops to zero, this region spatially overlaps with the n^{++} electrode, which is partially not depleted due to its high doping concentration ($N_D \sim 10^{19}/cm^3$).

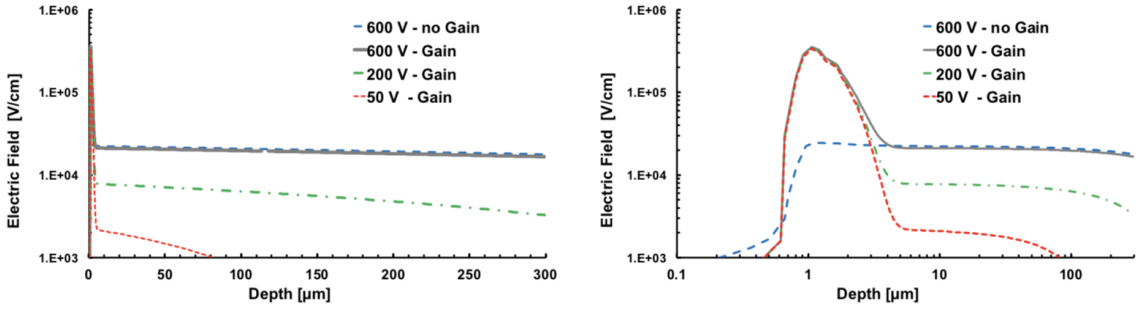

Figure 4.2: Electric Field of a 300 μm thick LGAD at different bias voltage compared to a traditional silicon sensor (no gain), with depth-axis in linear (left) and logarithmic (right) scale. Figure from [5]

4.1.1 Charge multiplication

As already mentioned in section 4.1, the avalanche multiplication mechanism starts when a charge carrier passes through a high electric field region. According to the theory of impact ionization model (equation 4.1), the number $N_{e,h}(d)$ of e-h pairs generated by the avalanche has an exponential dependency on the impact ionization coefficient $\alpha_{e,h}$ and on the length d travelled inside high electric field region

$$N_{e,h}(d) = N_{0;e,h} e^{\alpha_{e,h} d}. \tag{4.1}$$

The first parameter to consider in order to understand the mechanism of charge multiplication is the path length d. Assuming to have two hypothetical LGADs with p^+ layer implanted at two different depths (shallow and deep) from the n^{++} electrode:

assuming equal impact ionization coefficients, the gain will be greater for the deep gain layer. Vice versa, if we want to have the same gain for the two depths of gain layer the higher electric field for the shallow one must be higher than of the deep one. This is due to the longer path length in the multiplication region travelled by the charge carriers in the deep gain layer configuration, figure 4.3.

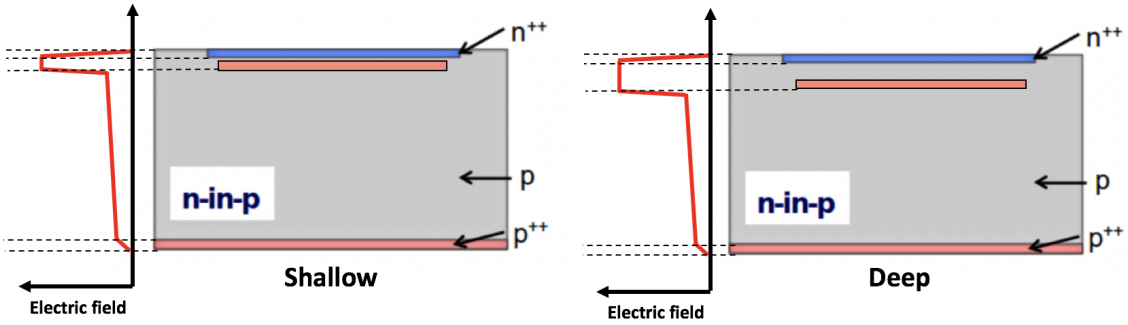

Figure 4.3: Schematic drawing of LGADs with shallow (left) and deep (right) gain layer implant with their respective electric field profiles.

The second parameter in the avalanche multiplication law is the ionization coefficient α, which is the number of e/h pairs generated per length unit by a single charge carrier. This coefficient is different for electrons (α_n) and holes (α_p). The inverse of the ionization coefficient is the mean free path (λ) needed to achieve the multiplication. Impact ionization occurs when the charge carrier travels a distance between two scattering centres (phonons and defects) long enough ($> \lambda$) to acquire kinetic energy greater or equal to the first ionization energy E_i. E_i, using the conservation law of momentum and energy, can be estimated to be about $1.5E_g$, where E_g is the energy gap of the semiconductor (1.12 eV for silicon at 300 K). This mechanism of multiplication suggests a dependency of the ionization coefficient and mean free path on the electric field. Simplified avalanche multiplication model (similar to the ionization in gases) is the Chynoweth model [23]:

$$\alpha_{n,p}(\varepsilon) = \frac{1}{\lambda_{n,p}(\varepsilon)} = A_{n,p} e^{-\frac{B_{n,p}}{\varepsilon}}, \tag{4.2}$$

where the coefficients $A_{n,p}$ are the maximum number of e/h pairs that can be generated in presence of a very high electric field, ε is the electric field and $B_{n,p}$ are coefficients derived from experimental fits and depends linearly on the temperature:

$$B_{n,p}(T) = C_{n,p} + D_{n,p}T. \tag{4.3}$$

Numerical values of $A_{n,p}$, and $B_{n,p}$ can be found in [24, 25, 26]. Equation 4.3 introduces also a relationship between ionization coefficients and temperature: the mean free path necessary to achieve multiplication is lower at low temperature (minor phonon population) causing an increase of gain, for equal electric field condition. More complex and realistic multiplication models are based on numerical simulation [27]; the four most important models are Van Overstraeten-De Man [28] and Massey [29] (both Chynoweth-like expression), Okuto-Cromwell [30], and Perugia model. Figure 4.4 shows the relationship between the mean free path and the electric field at temperatures of 300 K and 250 K for the three models Massey, Van Overstraeten-De Man, and Okuto-Cromwell.

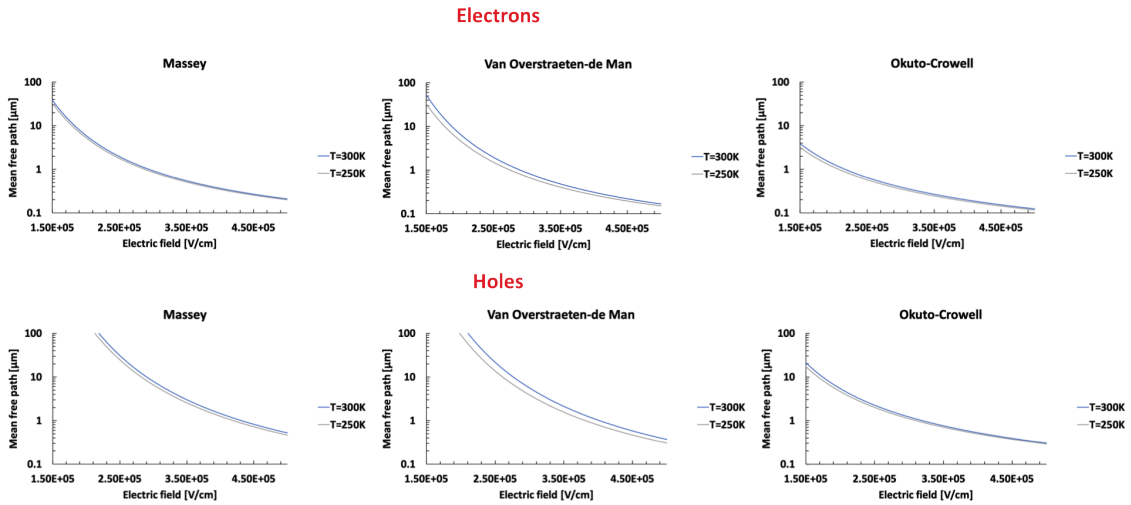

Figure 4.4: Mean free path of electrons (top) and holes (bottom) as a function of the electric field, for the three avalanche multiplication models: Massey, Van Overstreaten-de Man, and Okuto-Crowell. The mean free path at a temperature of 300 K is plotted in blue, while at 250 K in grey.

Additional consideration on the mean free path will be shown in section 4.8.3, where possible solutions to recover the full operation of irradiated sensors are illustrated.

4.2 Time-tagging detector

A time measurement relies on the read-out electronics capability to determine the time of passage of a particle, using as input the signal generated by the sensor. The main feature of the signal, for accurate time measurements, is to have a constant shape that scales with the amount of energy deposed. Figure 4.5 shown a very simple model of a sensor and the main electronic components able to measure the time of arrival of a particle (time-tagging detector), for an up-to-date review of current trends in electronics see [31]. The sensor, modelled as a capacitance (C_d) with a current generator (I_{in}) in parallel, is read-out by a preamplifier that shapes the signal. The preamplifier output is compared with a voltage threshold (V_{th}) to determine the time of arrival of a particle. The output of the comparator is digitized in a Time-to-Digital Converter (TDC). In the following, we will consider this simple model and we will do not use other approaches as waveform sampling, which are more complex and space-consuming. In this model, the time of arrival of the particle is defined as the time when the signal of the preamplifier crosses the threshold set in the comparator: any effects that change the shape of the signal near the threshold value can anticipate or delay the time measured and therefore affect the time resolution (σ_t). The time resolution, equation 4.4, is affected by several contributes:

$$\sigma_t^2 = \sigma_{Jitter}^2 + \sigma_{Ionization}^2 + \sigma_{Distortion}^2 + \sigma_{TDC}^2 \tag{4.4}$$

- Electronics, mainly due to the noise and the slew rate (how fast the signal rises), (σ_{Jitter});

- No uniform energy deposition by the particles determines irregularity and variation in amplitude of the signal, ($\sigma_{Ionization}$);
- Signal distortions, due to non-saturated drift velocity of charge carriers and non-uniform weighting field, ($\sigma_{Distortion}$);
- TDC binning uncertainties, (σ_{TDC});

Each of these contributes will be discussed in detail in the sections, also using simulation results obtained whit the simulator Weightfield2 (WF2) [32].

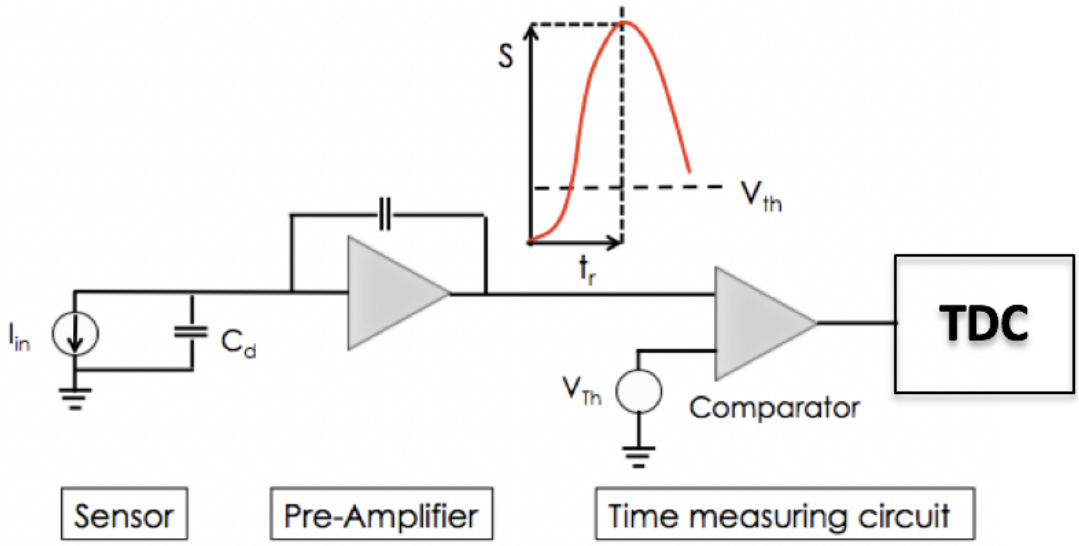

Figure 4.5: Simple block representation of a timing-tagging detector. The arrival time of a particle is measured when the signal crosses the comparator threshold V_{th}. Figure from [18].

4.2.1 Jitter

The presence of noise on the signal itself or in the electronics induces amplitude variations around the threshold value of the comparator, causing the early or late firing of the comparator itself; this generates the term in the time uncertainty expression called jitter. The jitter, equation 4.5, is directly proportional to the noise N of the

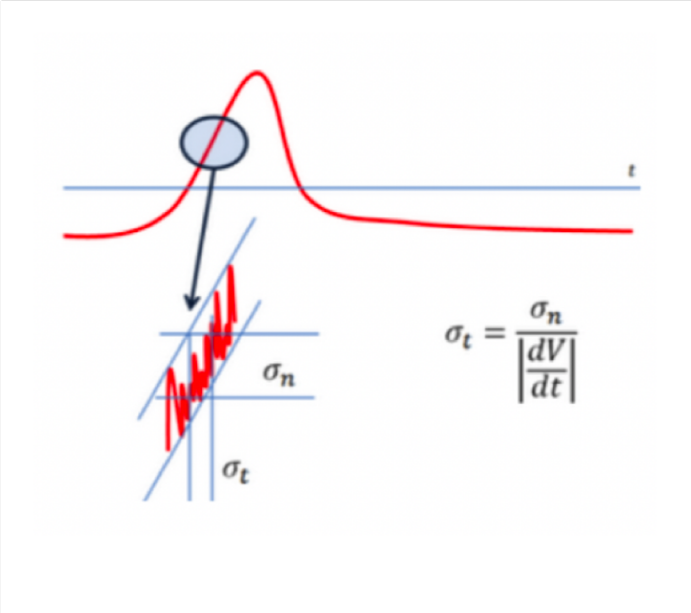

Figure 4.6: Jitter term, that causes the early or late firing of the comparator.

system and inversely proportional to the slope of the signal around the threshold value of the comparator, figure 4.6. Assuming a constant slope of the signal, we can write $dV/dt = S/t_{rise}$, where S is the amplitude and t_{rise} is the rise time of the signal, therefore:

$$\sigma_{Jitter} = \frac{N}{dV/dt} \approx \frac{t_{rise}}{S/N}. \tag{4.5}$$

The jitter can be minimized with sensor and electronic design optimization, which is a balance between two competitive effects: large slew rate requires wide bandwidth, which however increases the noise and therefore the jitter.

4.2.2 Ionization

The ultimate limit to signal uniformity is given by the physics governing the energy deposition when the particle passes through the sensor. The charge distribution generated into the sensor by ionizing particle varies on an event-by-event basis. These variations produce two effects, irregular current signals (Landau Noise) and changes in signal amplitude (Time Walk). These two effects are related to each other (discussion

in section 7.4).

1. **Landau Noise**: Figure 4.7 (top) shows three examples of energy deposition, simulated using the Weightfield2 program (WF2) [32], released by a minimum ionizing particle along the active thickness of a 200 μm thick traditional (no intrinsic gain) silicon sensor; while figure 4.7 (bottom) shows the associated generated current signals and their electron and hole components.

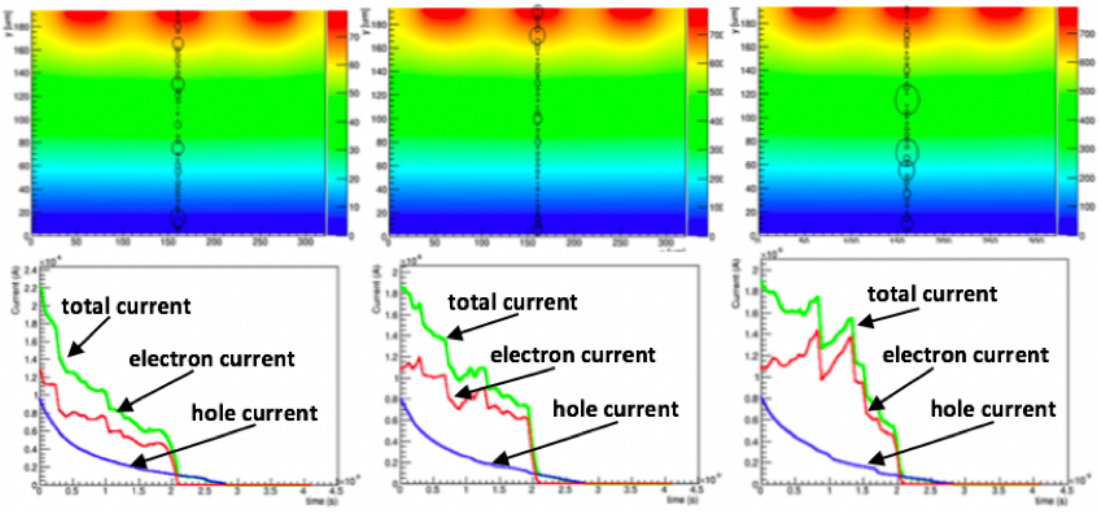

Figure 4.7: Three examples of simulated energy deposition in a traditional silicon sensor (top), and their corresponding total, electron, and hole current signals (bottom), WF2 simulation.

The variations of energy deposition are rather large and they can degrade a lot the achievable time resolution. There are two methods to mitigate landau noise:

- Integrating the output current over times longer than the typical spike length;
- Using thin sensors, its steeper signal is less sensitive to signal fluctuations. The intrinsic limit of Landau Noise is about 25 ps in a 50 μm thick sensor,

while it is about 60 ps in a 300 μm one.

In a well designed system, the landau term is the dominant contribution to the total time resolution.

2. **Time Walk**: the Time Walk term is due to the fact that larger signals cross a fixed threshold earlier than smaller ones, figure 4.8 (left). Let's assume for simplicity a linear signal with amplitude S and rise time t_{rise}; this signal crosses the threshold V_{th} with a delay t_d, figure 4.8 (right).

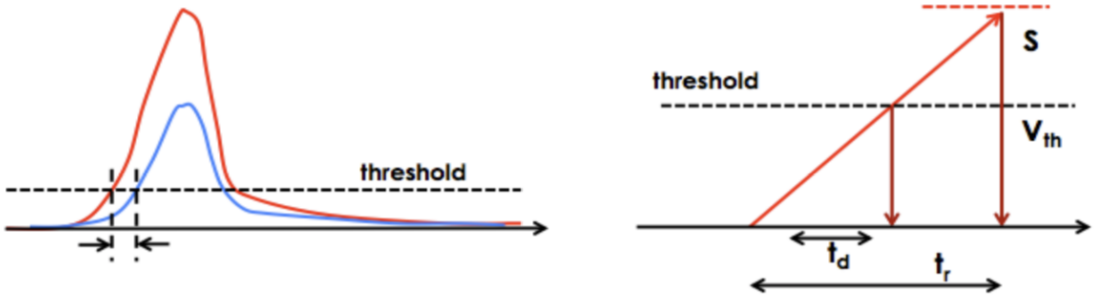

Figure 4.8: Left: two signals with different amplitudes, that cross at different times a fixed threshold, generating Time Walk, [33]. Right: linear signal of amplitude S and rise time t_{rise}, that crosses a threshold V_{th} with a delay t_d. Figure from [33].

Using the geometrical relationship $t_d/t_{rise} = V_{th}/S$, the instant of time when the signal crosses the threshold can be written as $t_d = t_{rise}V_{th}/S$; the Time Walk is then defined as the rms of t_d:

$$\sigma_{TimeWalk} = [t_d]_{RMS} = [\frac{V_{th}}{S/t_{rise}}]_{RMS} \propto [\frac{N}{dV/dt}]_{RMS}. \tag{4.6}$$

In this equation, where we use the relationship of equality $S/t_{rise} = dV/dt$ and we use the custom to express V_{th} as a multiple of the noise N of the system, is evident that the time walk contribution is minimized in systems with low noise and high slew rate.

The Time Walk effect is inevitable, but it can be corrected using appropriate electronic circuits, section 4.5.

4.2.3 Distortion

In each particle detector, the shape of the current signal can be calculated using Ramo-Shockley's theorem ($i \propto qv_dE_w$), equation 3.7. This equation indicates several key points in the design of the sensors to achieve accurate time measurements.

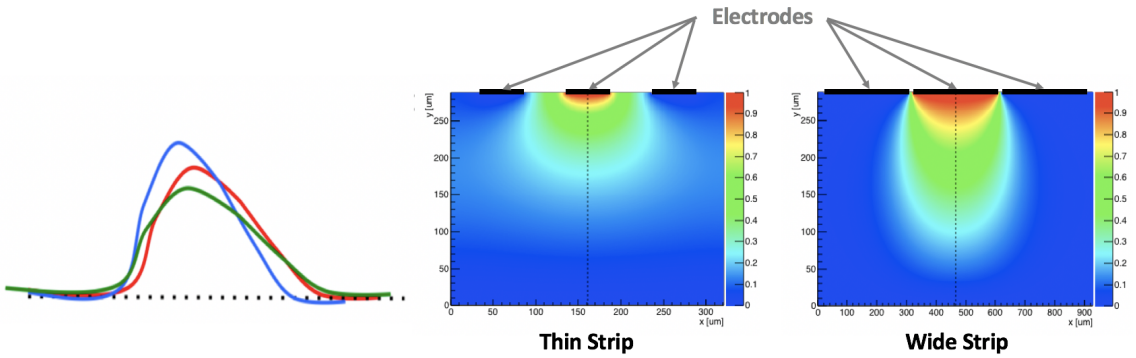

Figure 4.9: Left: the signal shape depends on the drift velocity of electrons and holes generated by the impinging particle; in non-saturated drift velocity condition, the signal shape depends on the particle hit position. Right: weighting field for two different strip geometry: wide strip (right) with a strip width of 290 μm and a pitch of 300 μm, thin strip (left) with a strip width of 50 μm and a pitch of 100 μm. In this second strip geometry, the weighting field is not uniform along the x-axes and a particle hitting near the center of the strip electrode generates a much steeper and faster signal.

1. The drift velocity of the charge carriers must be uniform everywhere in the active volume of the sensor, since non-uniform drift velocity induces variations in the signal shape that depend upon the particle hit position, spoiling the overall time resolution, figure 4.9 (left). The simplest way to obtain a uniform drift velocity in the active volume of the sensor is to have an electric field high enough to saturate the carriers drift velocity. The relationship between drift velocity of charge carriers and the electric field is linear up to a value of the electric field of 30 kV/cm, after which it saturates, figure 2.4.

2. The weighting field represents the capacitive coupling of a charge with the readout electrode. If this coupling changes significantly along the pitch implant, then the signal shape depends on the hit position of the particle. Figure 4.9 (right) shows two different strip geometries: first one, on the right, is a wide strip with an electrode width close to its pitch; the second one, on the left, is a thin strip with strip electrode much narrower of the pitch; in this second case the weighting field is located exclusively below the strip implant.

4.2.4 TDC

The time information is computed by a TDC, where the time of the leading edge of the discriminator signal is digitized and placed in a time bin of width $\triangle T$, given by the TDC least significant bit. This process adds a contribution to time uncertainty equal to $\triangle T/\sqrt{12}$. The TDC term in time resolution, equation 4.4, can be neglected thanks to the fine binning of TDCs commonly used in high energy physics experiments: for instance, the HPTDC [34] has a bin width of 25 ps that will contribute with an overall uncertainty of 7 ps.

4.3 UFSD signal formation

Weightfield2, a full simulation software, has been developed in Torino to assess the timing capability of UFSD sensors and to simulate the current signal of LGADs. The software has been validated by comparing the simulated signal for MIP and alpha particle with TCAD simulation and with experimental measurement performed at beam test, finding excellent agreement in both cases. Hence, WF2 is a very useful tool to investigate the main UFSD signal properties.

To understand the characteristics of the UFSD current signal it is necessary to study first the signal formation in a traditional (no-gain) silicon sensor. Using Ramo-Shockley's theorem equation 3.7, it is possible to calculate the maximum current in

a pad diode (without internal gain) of active thickness d, and in the condition of carriers drift velocity saturation:

$$I_{max} \propto Nq\frac{1}{d}v_{sat} = (n_{e-h}d)q\frac{1}{d}v_{sat} = n_{n-e}qv_{sat}, \tag{4.7}$$

where we assumed the weighting field E_w proportional to $1/d$ and where N is the number of e/h pairs created along the sensor thickness. N is given by the product between the sensor thickness d and the number of pairs $n_{e/h}$ created for unit length, assuming a uniform charge creation. Equation 4.7 shows a very interesting result about the no-gain silicon sensor: the maximum of current does not depend on the sensor thickness. In thicker sensors more e/h pairs are created (N is bigger), however, the charges contribute less to the initial current generation, due to the lower weighting field ($E_w \propto 1/d$). In thicker sensors, compared to thinner ones, the capacitive coupling between the generated charges and the readout electrodes is smaller. The relationship between current and sensor thickness in traditional silicon sensors yields an always equal peak current, $I_{max} \sim 1.5\ \mu A$. Equation 4.7 imposes a limit to the reachable time resolution by silicon sensors without gain, due to the current peak limited by the saturated value of carriers drift velocity and a fixed number of e/h pairs generated per length unit. In conclusion, traditional silicon sensors are not suitable to perform an accurate time measurement and it is, therefore, necessary to boost the current using gain mechanism.

Let's focus on the signal formation in a UFSD sensor, figure 4.10. An impinging particle generates primary electrons and holes in the depleted region, which drift towards n^{++} and p^{++} electrodes, respectively. The primary electrons cross the gain layer, activating the avalanche multiplication mechanism and producing secondary e/h pairs, called gain electrons and holes. Since gain electrons have been created close to the cathode, they are collected faster than gain holes, which drift in the

bulk before being collected by the anode. Hence, the signal in UFSD continues to increase for the collection time of gain electrons. The drift of gain holes generates a large induced current constituting the major part of the total current. The signal formation in UFSD takes a longer time than in no-gain sensor: a traditional diode has a faster rise time because the signal rising edge starts as soon as the e/h pairs have been created, while in UFSD the electrons must first reach the multiplication layer.

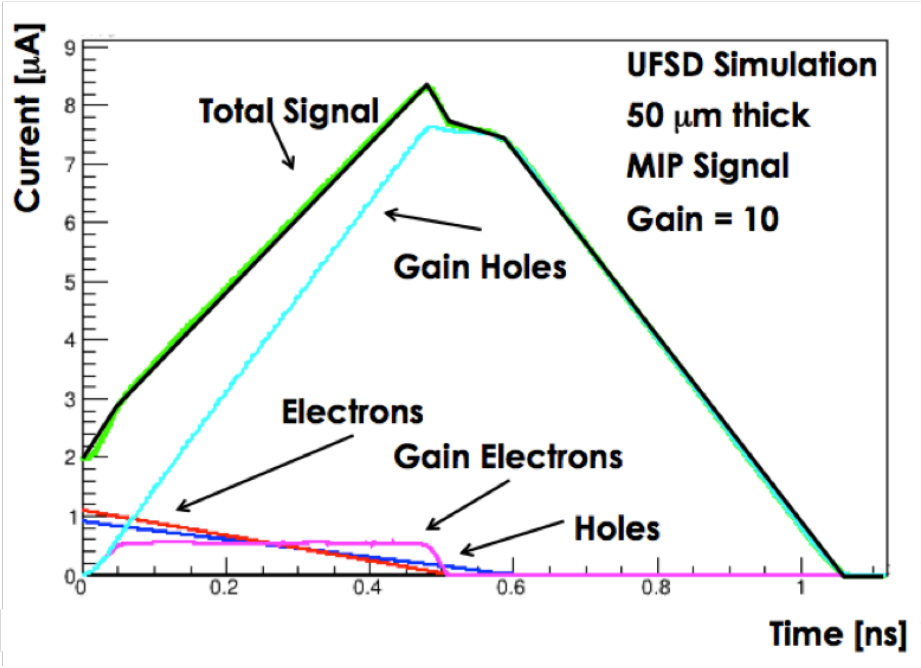

Figure 4.10: Simulated MIP current signal for a UFSD 50 μm thick. In red and blue electron and hole current, in violet and light blue gain electron and gain hole current, respectively. Figure from [32].

The current generation by the multiplication can be estimated from the number of electrons entering the gain layer in a time interval dt, with a drift velocity v_{sat}. The amount of these primary electrons is $n_{e-h}v_{sat}dt$ and they generate a number of e-h pairs $dN_{Gain} \propto n_{e-h}(v_{sat}dt)G$. Using the Ramo's theorem and assuming a parallel plate geometry ($E_w = 1/d$), it is possible to calculate the current signal induced by

these secondary charges:

$$dI_{gain} = dN_{gain} q v \frac{1}{d} \propto \frac{G}{d} dt, \tag{4.8}$$

which lead to:

$$\frac{dI_{gain}}{dt} \sim \frac{dV}{dt} \propto \frac{G}{d}. \tag{4.9}$$

Equation 4.9 shows a key feature of UFSD: the increase of the current signal (slew rate) is proportional to the ratio between gain and sensor thickness (G/d). This implies that thin sensors with high gain are suitable to perform high precision time measurements. WF2 simulation, in figure 4.11, shows the slew rate for UFSDs with different thicknesses and gains: a 300 μm thick sensor, with gain 15, has a slew rate two times higher than a tradition silicons sensor, while this factor becomes six times higher in a 50 μm sensor.

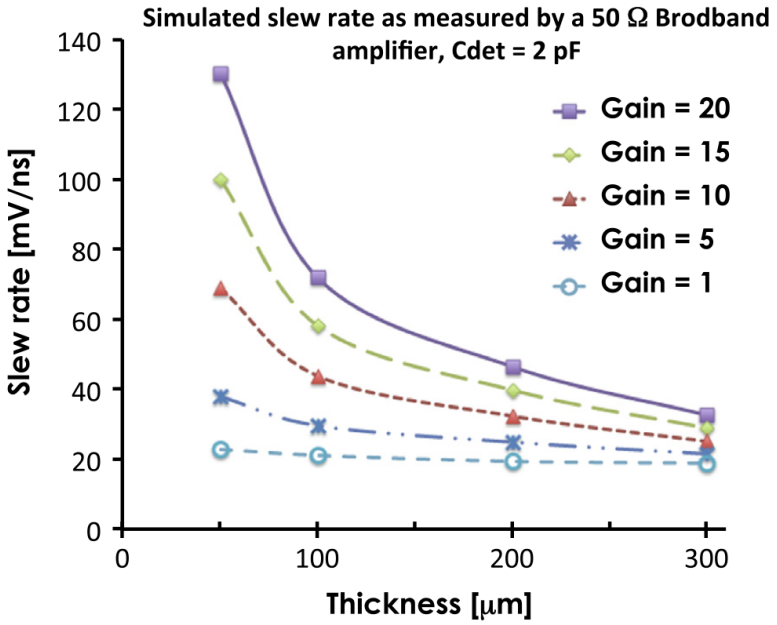

Figure 4.11: WF2 simulation of signal slew rate as a function of sensor thickness, for 5 different values of gain. Figure from [17].

As it was done for a no-gain sensor at the beginning of this section, it is possible to

calculate the maximum current signal in UFSD. Since each primary electron generates G e/h pairs, we can write:

$$I_{max} \propto N_{max} q \frac{1}{d} v_{sat} = (N_{e-h} dG) q \frac{1}{d} v_{sat} = n_{e-h} G q v_{sat}. \tag{4.10}$$

Equation 4.10 shows that, in drift velocity saturation mode, the current signal peak only depends on the gain and not on the sensor thickness; on the other hand the sensor thickness determines the rise time of the signal. Figure 4.12 schematically shows the signal shapes for sensors with equal gain and different active thicknesses.

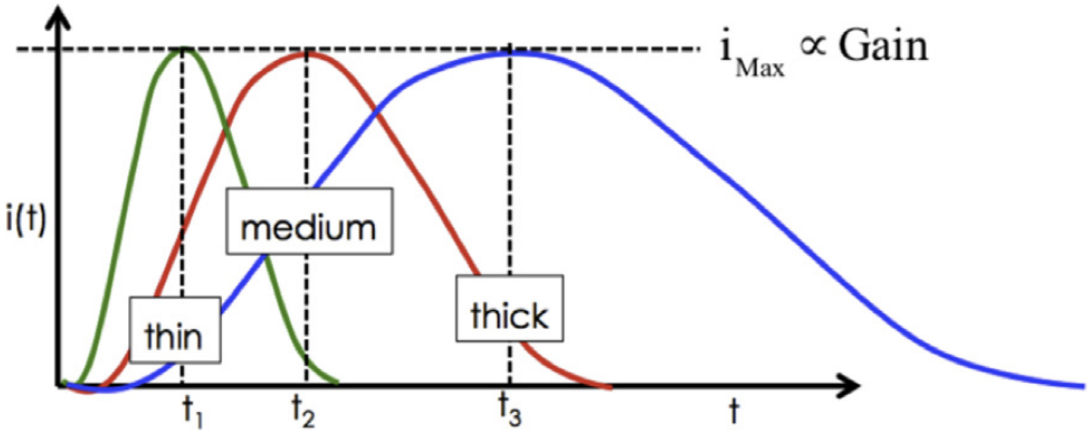

Figure 4.12: Current signal shapes from UFSDs with the same gain and different active thicknesses. Figure from [17].

In conclusion, we can summarize that sensors with high internal gain and thin active thickness are suitable to achieve excellent time resolution.

4.4 Noise

The Noise is defined as spontaneous fluctuations of current or voltage across the semiconductor or bulk device. If the device is used to measure small physical quantities or to amplify a small signal, then the fluctuations in current and voltage set a lower limit to the quantities to be measured or the signal to be amplified. The noise can be

classified into thermal, flicker, and shoot noise. The thermal noise occurs in any semiconductor and it is due to the random motion of charge carriers. The mean-square voltage of thermal noise is given by

$$\left\langle V_n^2 \right\rangle = 4kTBR \tag{4.11}$$

where k is the Boltzmann constant, T is the absolute temperature in Kelvin, B is the bandwidth in Hz and R is the real part of the impedance between terminals. The Flicker noise is due to surface effect and it becomes important at a lower frequency, it is proportional to $1/f^{\alpha}$ where α is close to unity. The shot noise is the main noise in the most of semiconductor devices, it is frequency independent (white spectrum) at medium and low frequencies. The mean-square noise current of shot noise for a p-n junction is given by

$$\left\langle i_n^2 \right\rangle = 2qB|I| \tag{4.12}$$

where I is the current, which is positive in the forward and negative in the reverse direction. For low injection, the total mean-square noise current is given by

$$\left\langle i_n^2 \right\rangle = \frac{4kTB}{R} - 2qB|I| \tag{4.13}$$

Experimental measurements confirm that the mean-square noise current is proportional to the saturation current I_s, which varies with the irradiation level.

An undesired effect related to the gain mechanism, called excess noise factor F, can limit the accuracy of time measurement of UFSD sensors. The excess noise factor is an additional noise induced by the multiplication mechanism: each primary electron entering the gain layer generates a number of secondary charges that on average is equal to G. However, each electron can generate more or less secondary charges, introducing a fluctuation in the multiplication mechanism, which induces a noise. In

sensors with internal gain the signal is multiplied by G, while the noise current by $\sqrt{F}$, where F is a function of G and it is given by, [35]:

$$F \sim G^x = Gk + (2 - \frac{1}{G})(1 - k) \tag{4.14}$$

where x is called excess noise index and k is the ration (α_p/α_n) between impact ionization coefficients of hole and electron. Hence, the gain improves the ratio signal to noise SNR only if the sensor noise current is not the dominant source of the noise. In particular, in order to keep the noise low, also factor F must be kept low; this is possible with a low internal gain and reducing as much as possible the hole ionization coefficient. This last requirement is satisfied by n-on-p sensor design, in which it is possible to tune the electric field in such a way that only the electrons are multiplied, whereas this is not possible with p-on-n design, [17].

4.5 Read-out electronic

The time performances of UFSDs depend not only on their design but also on read-out electronic, as mentioned in section 4.2. Read-out preamplifiers can be classified into two main architectures: current amplifier called Broad-Band Amplifier (BBA) and Charge Sensitive Amplifier (CSA). The first one amplifies the signal without any additional shaping, the second one integrates and strongly shapes the signal. Figure 4.13 shows the responses of a broad-band and charge sensitive amplifier to an input current signal generated by a MIP in an LGAD 300 μm thick: in red the BBA output signal with a shape very similar to the input one; in blue the CSA output.

A BB amplifier converts with some gain the current signal $i(t)$ in a voltage signal $V(t)$ so that $V(t) \sim i(t)$. This kind of amplifier needs a rather wide bandwidth to follow the time evolution of the current signal, hence, it is advantageous for thin sensors with a large signal slew rate. The noise of BBA tend to be large, due to the

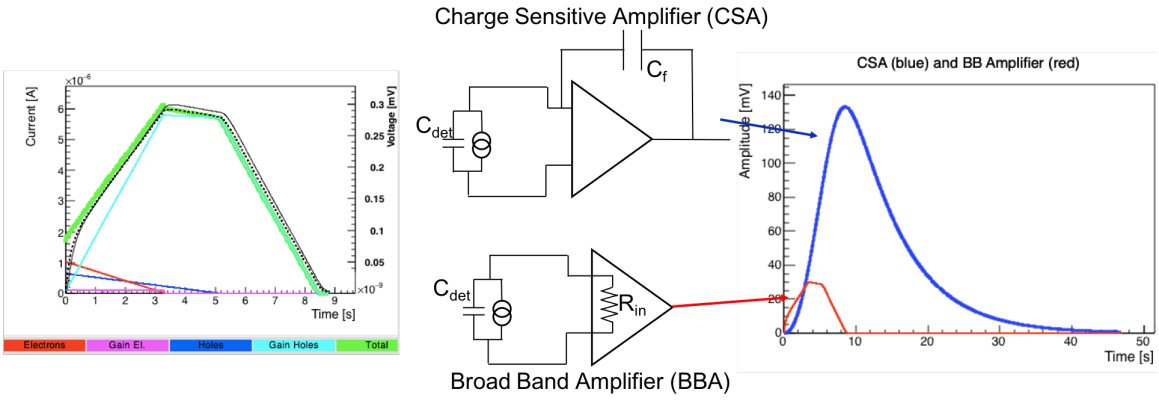

Figure 4.13: Weightfield2 simulation of the BBA (red) and CSA (blue) response to an input current signal generated by a MIP, in an LGAD 300 μm thick.

wide bandwidth, however, the jitter term is minimized by large dV/dt term. The sensor-BBA system can be modelled as a RC circuit with a current generator in parallel and, as such, it has a time constant $\tau = C_{det}R_{in}$, where C_{det} is the sensor capacitance and R_{in} is the read-out input impedance. This implies that, in order to follow the large slew rate provide by UFSD, the time constant must be shorter or at most comparable with the rise time of the signal ($\tau \sim t_{rise}$). It is evident that the design of this kind of amplifier is strongly connected to the development of the sensor.

CSA integrates the current signal generating an output voltage signal of amplitude proportional to the charge collected on the feedback capacitance C_f of the preamplifier stage, and the integration time is given by the time constant R_fC_f of the feedback circuit. This preamplifier architecture minimizes the jitter thanks to the low noise, while the slew rate of the signal is slower compared to the BBA. A CSA output signal is governed by two time constants: (i) the rise time $t_{rise} \sim (C_{det} + C_{load})/g_m$, which depends on the sensor capacitance C_{det}, on the capacitance C_{load} of the loading circuit and on the input transconductance g_m; (ii) the fall time $t_{fall} \sim R_fC_f$, which depends on the time constant of the feedback circuit. The optimal operation of this amplification mode requires a signal fall time longer than the rise time, otherwise, the charge on the feedback capacitance will discharge before reaching the maximum

amplitude of the output signal; this effect is called ballistic deficit [36]. The sensor capacitance has a strong influence on the CSA performances, because, when it is too large, it increases the noise, decreases the signal rise time and output amplitude.

The choice of preamplifier architecture is strongly determined by the thickness of the sensor. As mentioned in section 4.2.2, the energy deposed by an impinging particle is not uniform along the sensor thickness, resulting in Landau fluctuations; if the unavoidable integration due to the sensor capacitance is enough to mitigate these fluctuations, then a BBA is the best choice, otherwise, the CSA provides the best performance. Moreover, Landau fluctuations generate time walk effect, 4.2.2, which can be corrected with an appropriate electronic circuit.

The "time walk correction" aims at mitigating the fact that larger signals cross a fixed threshold earlier than smaller signals. There are two common approaches to this problem, figure 4.14: (i) Constant Fraction Discriminator (CDF) or (ii) Time over Threshold (ToT). The CFD fixes the time of arrival of the particle when the signal crosses a given fraction of the maximum amplitude of the signal in this way the time of arrival depends only on the rising edge of the signal and its maximum. In this method, the effect of different signal amplitudes is removed by using a fixed signal fraction and not a fixed voltage as a definition of time of arrival. The ToT uses two time points to determine the time of arrival of a particle. This method measures the time duration of the signal above a set threshold value, and it uses it to correct the time point when the signal exceeds the threshold value with a formula optimized for the given electronics. In this method, the effect of different signal amplitudes is removed by applying this correction. Since BBA and CSA shape the current signal differently, the effectiveness of CFD and ToT is different for the two amplifier architectures. Both methods CDF and ToT can be used with CSA architecture since its output signal has amplitude and duration proportional to the input charge; while with the BBA the best method is CFD since the output signal width is practically constant.

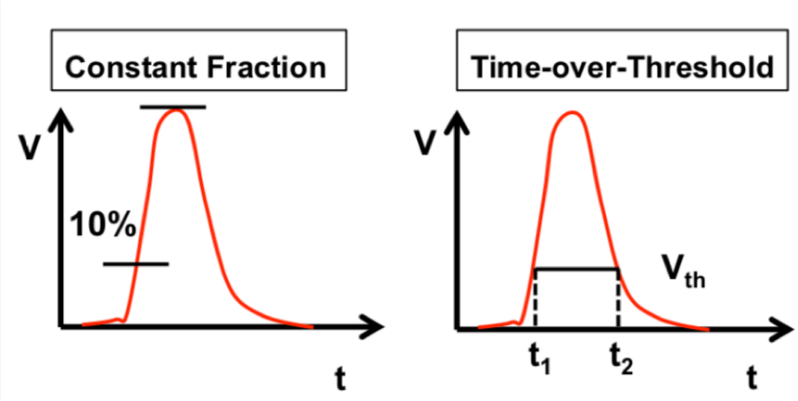

Figure 4.14: Basic methods to correct signal amplitude fluctuations. Left: Constant Fraction Discrimination. Right: Time over Threshold. Figure from [17].

4.6 Time performances

It is now possible to conclude that UFSDs interfaced with an appropriate read-out electronic have the capacity to achieve very good time resolution. The moderate gain of the LGAD technology (maximization of SNR), the reduced thickness (mitigation of Landau fluctuations and fast slew rate of the signal), the appropriate sizing of the sensor capacitance and its coupling with a suitable read-out electronic, provide the key ingredients for a good time resolution.

Figure 4.15 shows one of the most important results on the simulation of the time resolution as a function of the active thickness of LGAD. The WF2 simulation shows that UFSD 50 μm thick, 6 pF of capacitance, achieves a time resolution of 30-40 ps, for a gain of about 20 (BB amplifier of gain 40 dB, bandwidth 2 GHz, and the CFD at 15% have been set in the simulation). In addition, an interesting result is also shown: for active thickness below 150 μm the jitter and Landau noise contribute equally to the total time resolution, while the jitter term dominates for those above.

Another interesting simulation is the interplay of the gain layer doping with the external bias in the determination of the time resolution. As discussed in section 4.1 a certain value of gain can be achieved with different combinations of gain layer

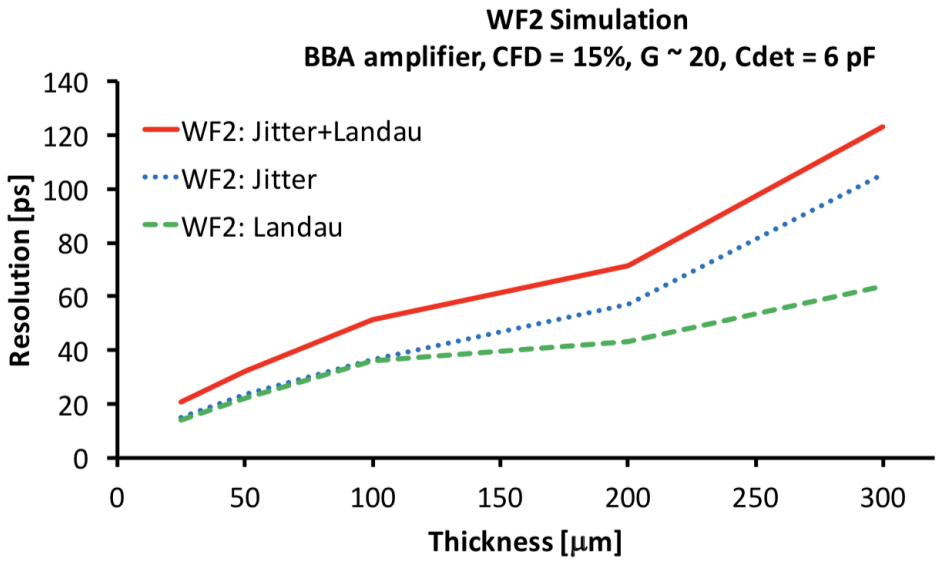

Figure 4.15: Weightfield2 simulation of jitter and landau noise contribution to the time resolution as a function of LGAD thickness. Figure from [5]

doping (depth of implant) and external bias voltage. Figure 4.16 shows the simulated time resolution for a UFSD 50 μm thick with gain 25, with different combinations of external bias and gain layer doping, where the doping values are relative to the one for which gain 25 is obtained at bias 160 V.

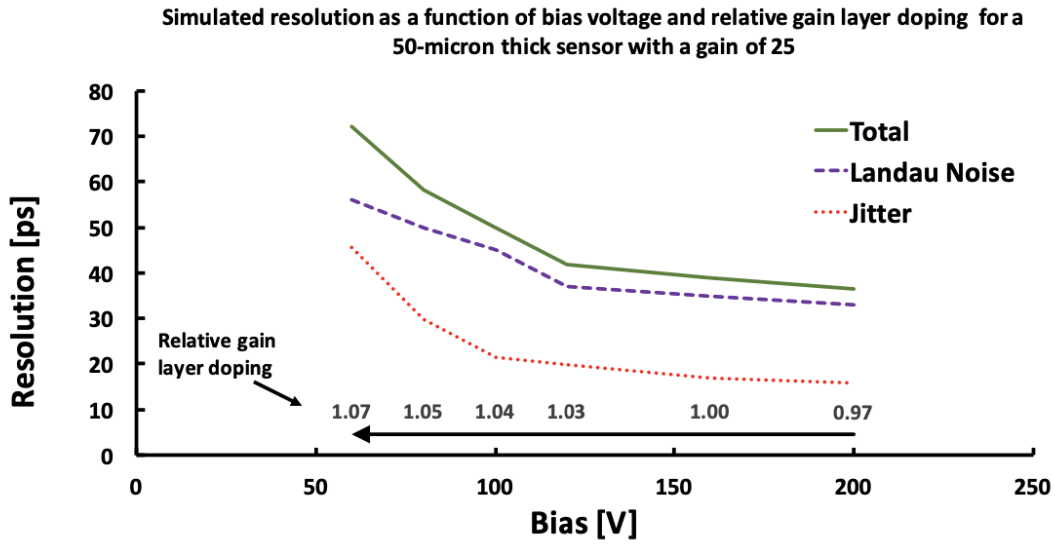

Figure 4.16: Time resolution simulation for different combinations of gain layer doping and external bias voltage, of a UFSD 50 μm thick with a constant gain of 25. Figure from [5]

At high relative doping values of the gain layer (1.04-1.07), the gain of 25 is reached at rather low bias voltage and the time resolution is worse: this result is due to a not saturated drift velocity of the charge carriers (low bias condition) and it demonstrates the need to work in a saturated regime to obtain a good time resolution.

4.7 Large area sensors, UFSD building block

The UFSDs were chosen to instrument the ETL of the MTD of the CMS experiment. Each ETL sensors will have 16×32 pads, each pad 1.3×1.3 mm^2. This design requires the optimization of the area between pads in order to prevent premature breakdown and to electrically isolate adjacent pads. Figure 4.17 shows the cross section of the UFSD building block (not in scale), with the termination structures required for multi-pad sensor operation: Junction Termination Extention (JTE), p-stop and Guard-Ring (GR).

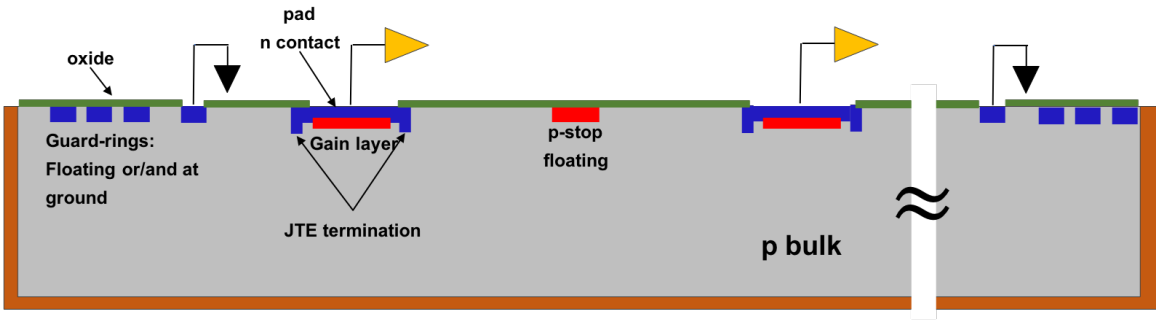

Figure 4.17: Cross Section (not in scale) of the UFSD multi-pad sensor building block, with edge termination structures: Junction Termination Extention (JTE), p-stop, and Guard-Ring (GR).

JTE is a deep n^{++} implant (usually with the same concentration of n electrode), equipped with metal field plate, implanted around each pad in single and multi-pad sensors. The reason for this implant is to ensure that e/h pairs, generated by particles hitting in the inter-pads region, do not reach the gain layer and that they are collected by this protection ring on the peripheral region of the pad. This is necessary since

the drift time from the inter-pad region to the gain layer is rather long and it would generate an "out of time" signal.

The p-stop is a p^{++} implant, with a doping concentration between 10^{16}-$10^{18}/cm^3$, it is located between pads; this implant is necessary to keep pads electrically isolated.

The guardring (GR) is a termination structure to prevent premature breakdown at the periphery of the device. The GR consists of one or more n^{++} rings, equipped with metal field plates, which surround the active area of the sensor. These rings are floating, except for the inner one, which is biased in order to collect charge carriers generated outside the core region of the device. Each GR is separated from adjacent n^{++} implants by p-stop implants.

The JTE and p-stop terminations introduce a no-gain region, tenths of μm wide, between adjacent pads: thus decreasing the fill factor of the device. Figure 4.18 shows 2D TCAD simulation of the pad periphery, where it is possible to distinguish the drift lines, followed by the drifting charges [37].

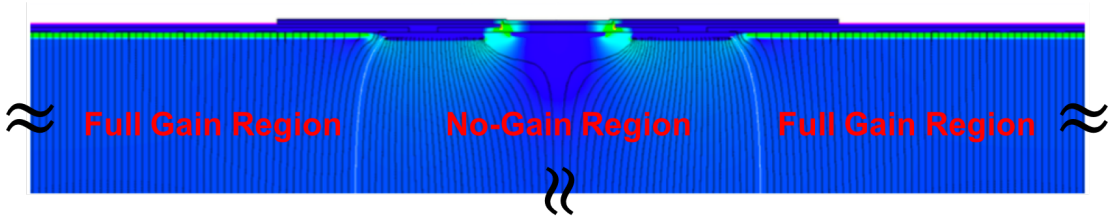

Figure 4.18: TCAD 2D-simulation of the electric field (intensity color map) and drift lines in the inter-pad region of a UFSD sensor.

Several different types of UFSDs, based on different segmentation technologies to increase the fill factor, are under development. Two interesting technologies, trenches isolation [38] and resistive AC-couple silicon detectors [37, 39], have the potential to obtain the 100% fill factor.

4.8 Radiation effects on UFSD

In chapter 1 it was illustrated that UFSD sensors in the ETL detector of CMS will be subject, in their lifetime, to an irradiation fluence of $1.5 \cdot 10^{15}$ n_{eq}/cm^2. Considering a safety factor of 2, the radiation hard requirement for the UFSD sensors installed in CMS is that they survive to a fluence of $3 \cdot 10^{15}$ n_{eq}/cm^2. This target will be discussed in this chapter.

The main effects of radiation damage on traditional silicon sensors have been discussed in detail in section 3.3; hence, only the specific effects on UFSDs performances will be discussed below.

4.8.1 Effects of current increase: power consumption and shot noise

In sensors with internal gain, the leakage current generated in the bulk is multiplied by the gain factor G before being collected at the electrodes. This unavoidable effect increases power consumption:

$$i_{gain} = G \cdot i_{no-gain} \tag{4.15}$$

leading to

$$P_{gain} = G \cdot P_{no-gain}. \tag{4.16}$$

Power consumption can be reduced using thin sensors since both leakage current and operating voltage are lower in thin sensors and by cooling the sensor, since the leakage current depends on the temperature:

$$i(T) = i_0 T^2 \exp \frac{1.2eV}{2kT}. \tag{4.17}$$

A temperature variation of 7 degrees leads to a current variation of a factor of two.

Another effect due to the high leakage current in irradiated sensors is the increase

in shot noise. Shot noise occurs when the charge carriers cross a potential barrier, as it happens in silicon sensors. In UFSD sensors this noise is enhanced by the internal gain and it can become the dominant source of the noise. As mentioned in section 4.4, the shot noise depends on the leakage current of the sensor, which can be divided into two components, figure 4.19 (left): surface current, that does not pass through the gain layer; and bulk leakage current, that is multiplied by the avalanche multiplication mechanism. When the charge carriers undergo multiplication, there is an additional mechanism that increases the shot noise, this is the excess noise factor F, discussed in section 4.4. F causes a peculiar effect in sensors with gain: with increasing gain, the ratio signal to noise becomes smaller since shot noise grows faster than the signal, figure 4.19 (right). Hence, in UFSD sensors, the gain is beneficial only if it remains below the electronic noise floor.

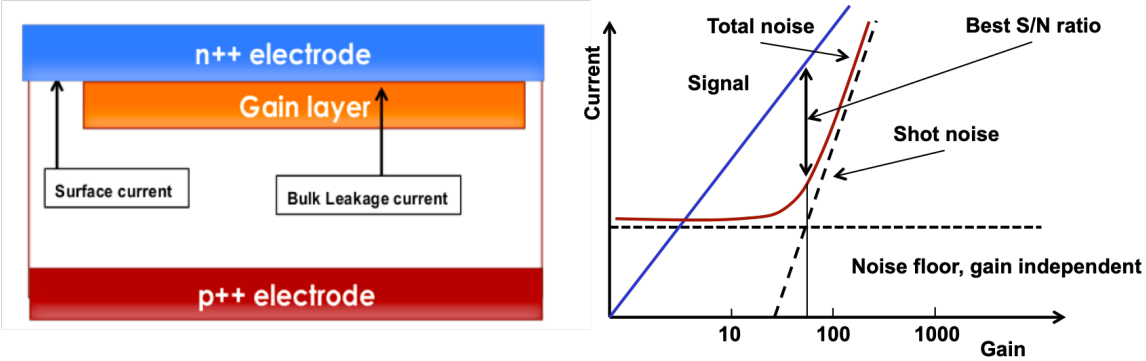

Figure 4.19: Left: schematic representation of the shot noise mechanism in sensors with internal gain: bulk leakage current is multiplied by the gain, while the surface current is not. Right: signal and shot noise growth as a function of the sensor internal gain. Figure from [18].

The shot noise can be written as:

$$i^2_{shot} = 2q(I_{surface} + I_{bulk}G^2G^x). \tag{4.18}$$

Shot noise is usually lower than electronic noise in un-irradiated sensors, but it becomes the dominant source of noise in irradiated ones. Figure 4.20 shows the shot

noise as a function of the irradiation fluence for a 4 mm^2 50 μm thick sensor, assuming an integration time of 2 ns; the plot on the left shows the strong impact of the gain on the shot noise, while on the right the effect of the temperature. In conclusion, figure 4.20 suggests operating irradiated sensors at low gain and temperature. Additionally, shot noise can be kept low by operating sensors with small active volume, since the leakage current is proportional to the depleted volume of the sensors, equation 3.13.

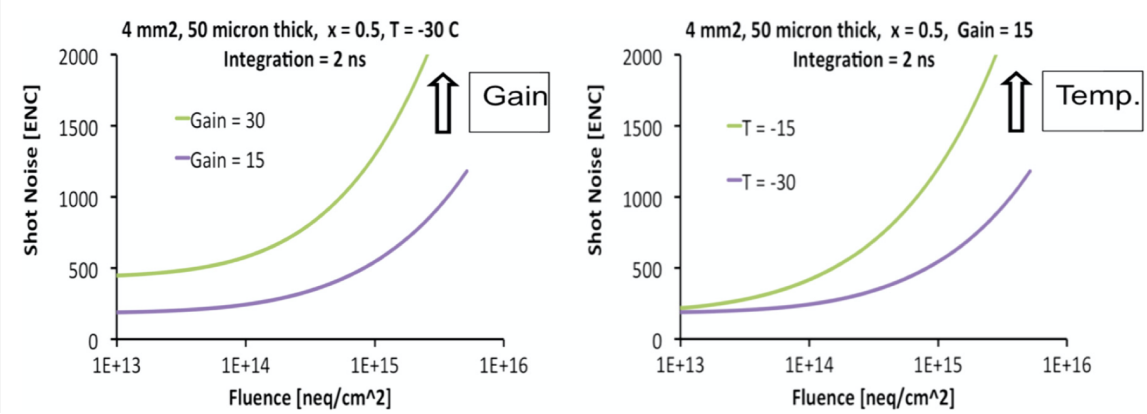

Figure 4.20: Left: shot noise as a function of the irradiation fluence for two different values of internal gain. Right: shot noise as a function of the fluence for two different operating temperatures. Figure from [19].

4.8.2 Variation in doping concentration

Another effect of the radiation is the variation of doping concentration as a function of the fluence, equation 3.14. This mechanism has been discussed in section 3.3, and it consists of two opposite and concurrent contributions: initial acceptor removal and acceptor creation. Initial acceptor removal causes an exponential decrease in the concentration of acceptors (the rate of initial acceptor removal is faster for lower initial acceptor density, figure 3.9), while, acceptor creation creates acceptor-like defects that increase the p-type doping concentration. Both mechanisms can be applied to both gain layer and bulk of an LGAD. Figure 4.21 shows the evolution of the boron concentrations as a function of fluence, in gain layer and bulk, with initial boron densities of $2 \cdot 10^{16}/cm^3$ and $2 \cdot 10^{12}/cm^3$, respectively. At sufficiently high fluences

($> 1 \cdot 10^{16}$ n_{eq}/cm^2) the curve of the gain layer doping converges on the straight line of the bulk, indicating complete removal of the initial gain layer doping.

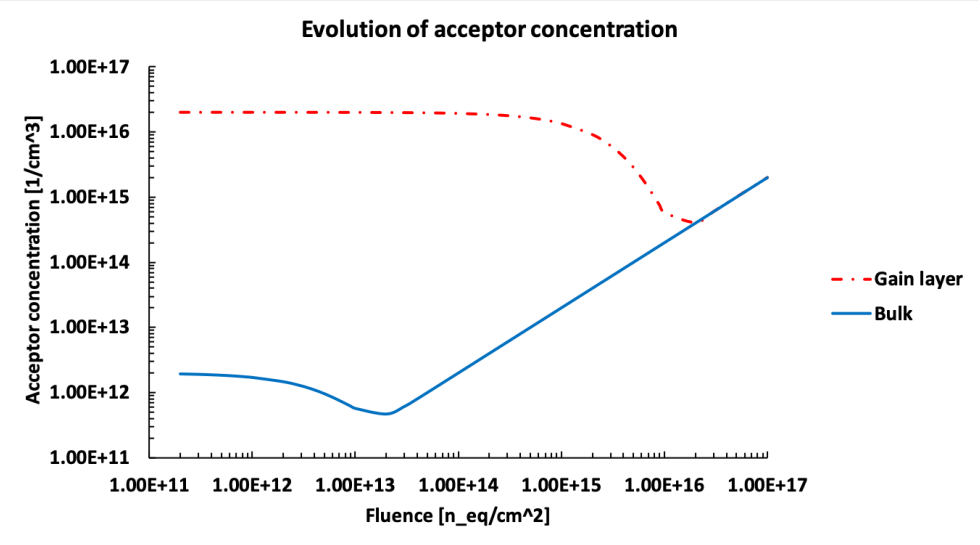

Figure 4.21: Evolution of acceptor concentrations as a function of the irradiation fluence, for a gain layer and bulk typical of an LGAD.

Considering the gain layer and bulk doping concentration plot in figure 4.21, it is possible to calculate the full depletion voltage of UFSD at different irradiation fluence values. This value is the sum of two terms: depletion voltage of the gain layer and depletion voltage of the bulk, and it is shown in figure 4.22 for a 45 μm thick sensor.

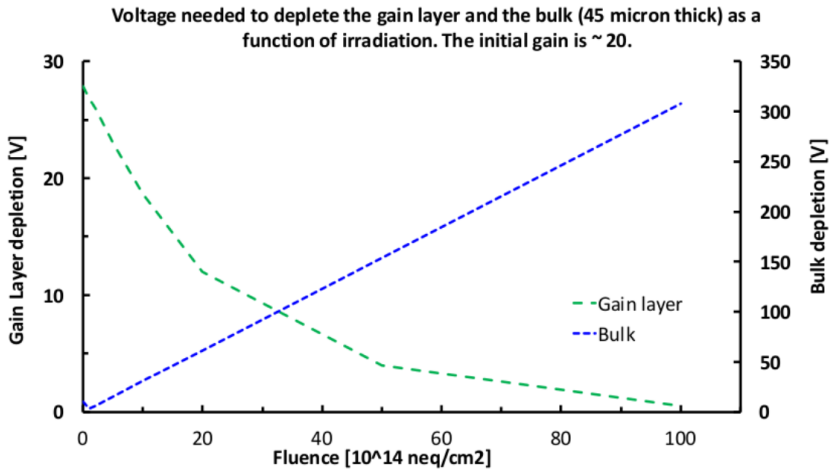

Figure 4.22: Evolution of the depletion voltage of the gain layer and bulk, with the fluence, for a typical 45 μm thick UFSD. Figure from [5].

The microscopic origin of initial acceptor removal [40] has not been fully understood and it is still under investigation. An important consideration is that the removal of boron atoms has been excluded from kinetics considerations and Secondary Ion Mass Spectrometer (SIMS) measurements: SIMS has been made on two twins UFSDs, the first (M83) un-irradiated, the second (M80) heavily irradiated with a fluence of $1 \cdot 10^{16}$ n_{eq}/cm^2, where the gain layer has completely disappeared. The purpose of the SIMS is to measure the boron density (substitutional and interstitial) as a function of the depth; the SIMS results show identical boron profiles for the two samples, indicating that the disappearance of the gain does not correspond to the removal of the boron atoms, only to their inactivation, figure 4.23.

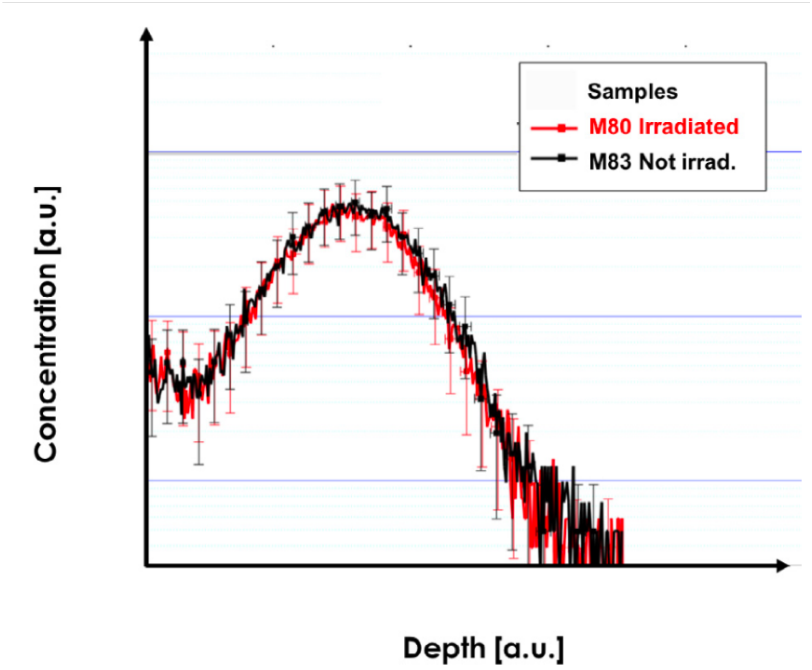

Figure 4.23: SIMS measurements of the density of boron atoms forming the gain layer as a function of the depth, in a not-irradiated (M83), and heavily irradiated (M80, irradiated to $1 \cdot 10^{16}$ n_{eq}/cm^2) UFSDs. Although the gain layer of the sample M80 is almost completely deactivated, its doping profile is identical to that of the not-irradiated sensor. Figure from [41]

A possible explanation of the initial acceptor removal mechanism is based on the ion-acceptor complexes formation with irradiation, the topic discussed in section 3.3:

irradiation creates interstitial silicon, which subsequently deactivates the boron via kick-out reactions, figure 4.24 (top). This interpretation of initial acceptor removal mechanism motived two research lines within RD50 collaboration [42], with the goal of mitigating the deactivation of the gain layer: the first one aims at reducing the concentration of interstitial defects available for capturing boron atoms by using implanted carbon atoms, which replace boron in ion-defect complexes formation, figure 4.24 (middle); the second one consists in replacing the boron atoms with gallium, which is predicted to have a slower acceptor removal rate, figure 4.24 (bottom).

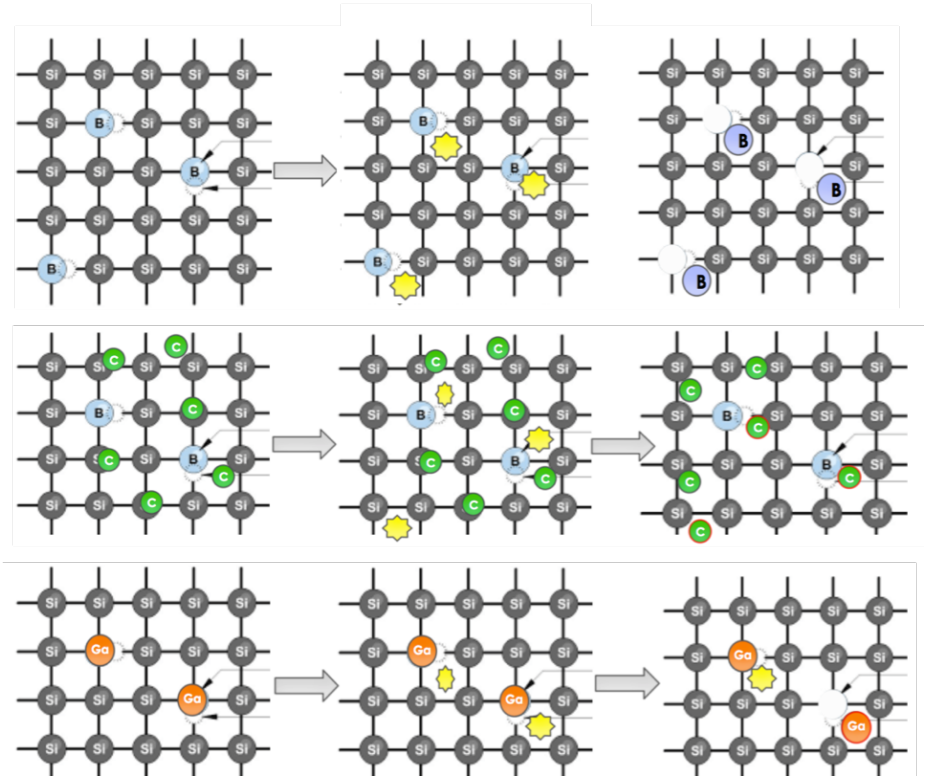

Figure 4.24: Sketch of initial acceptor removal mechanism for gain layers with active dopants: boron (top), boron enriched with carbon (middle), and gallium (bottom).

4.8.3 Gain recovery in irradiated UFSD

Radiation damage has a dramatic effect on the gain of UFSDs, severely reducing it for fluence of the order of $1 \cdot 10^{15}$ n_{eq}/cm^2. However, gain loss can be recovered by increasing the electric field inside the device (equation 4.1), by increasing the external

bias voltage as shown in the WF2 simulation in figure 4.25. Before the irradiation, the gain layer provides a sufficient electric field to start the multiplication mechanism. With irradiation, while the doping concentration of the gain layer decreases the bulk doping increases (acceptor creation, section 3.3) making it possible to start the multiplication mechanism also in the bulk. This effect is called bulk gain. However, it is important to stress that the doping of the bulk, at a fluence of the order $\sim 10^{15}\ n_{eq}/cm^2$, does not reach the high doping concentration values of the gain layer; therefore, since the electric field is given by the sum of the field resulting from the external bias and the doping concentration, the bias voltage should be raised to higher values than in not-irradiated condition. As an example, in sensors irradiated with a fluence of $\sim$ 2-3 $\cdot\ 10^{15}\ n_{eq}/cm^2$, bulk gain happens at an electric field value of $\sim 200\ kV/cm$, which is reachable with a bias of 600 V on 50 μm of active thickness.

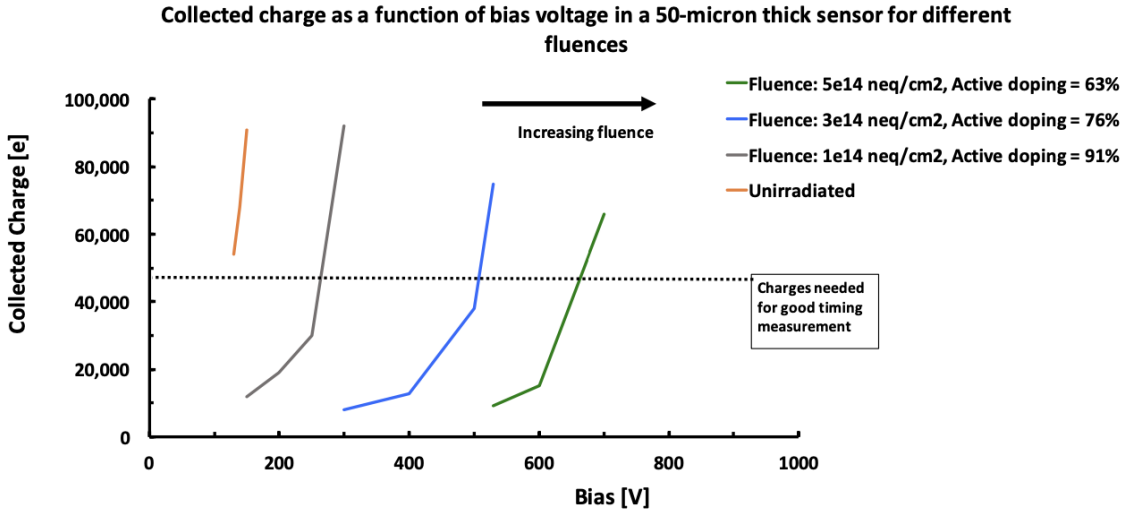

Figure 4.25: WF2 simulation of Gain recovery by increasing the bias voltage for different irradiation fluences, in a 50 μm thick UFSD. Figure from [5].

Another parameter that affects the gain recovery, in addition to the bulk gain, is the depth of the multiplication layer. As discussed in section 4.1, the same gain in LGADs with shallow and deep gain layers requires different working points of the electric field and therefore different mean free paths λ to achieve multiplication,

equation4.2. Figure 4.26 (left) shows the field working points for shallow and deep gain layer: $\sim$ 400 kV/cm for shallow GL and $\sim$ 300 kV/cm for deeper GL. In the multiplication models discussed in section 4.1.1, λ saturates at high electric field values, thus limiting the gain recovery capability provided by the electric field.

The deep gain layer (low field working point) has a higher gain recovery capability than a shallow one (high field working point), due to the faster variation of λ as a function of the electric field, figure 4.26 (right).

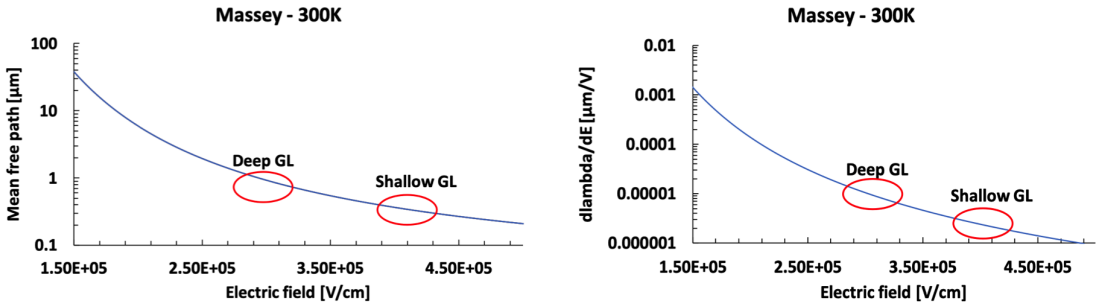

Figure 4.26: Left: electric field working point for a deep and shallow gain layer. Right: saturation effect of the mean free path at the increase of electric field shown through its derivative.

In conclusion, a good balance between the radiation resistance of the gain layer (acceptor removal rate) and its depth of implant (gain recovery capability) are the two main parameters that determine the radiation hardness of the UFSD design.

4.8.4 Effect of trapping on output signal shape

In irradiated sensors, the charge carriers undergo trapping, which causes a decrease of CCE and affects the output signal shape. The trapping increases with the irradiation fluence and with the drift length of the charge carriers. For irradiation fluence of the order of 10^{15} n_{eq}/cm^2 the trapping length is $\sim$ 50 μm. For this reason, trapping is less problematic in thin sensors than in thicker ones. Figure 4.27 shows a WF2 simulation of how the signal shape changes as a function of fluence, for a 50 μm-thick UFSD [43]; the effect of trapping is rather small up to a fluence of 10^{15} n_{eq}/cm^2, while it

becomes more important above this level.

It is interesting to note that in the simulated signal, figure 4.27, the rising edge used for timing is not affected much by trapping. The rise time becomes shorter due to the bulk gain which immediately generates charge carriers in the bulk without the electrons having to reach the gain layer.

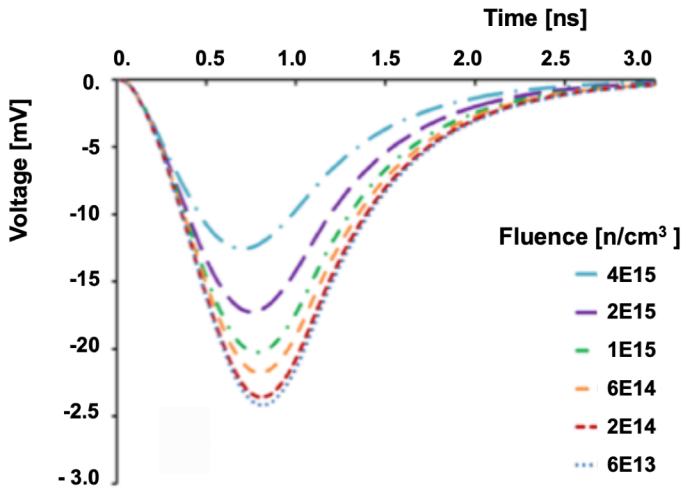

Figure 4.27: Effect of trapping on signal shape as a function of the irradiation level, for a 50 μm UFSD (in this plot the only effect of trapping is considered). Signals are the result of a WF2 simulation with the BB amplification stage. Figure from [6].

In conclusion, the possibility of accurate time measurements persists even after high levels of irradiation.

Chapter 5

UFSD productions

The UFSD sensors are produced by three main vendors: Centro Nacional de Microelectronica (CNM) in Barcelona [44], Hamamatsu Photonics (HPK) [45] and Fondazione Bruno Kessler (FBK) in Trento [46]. The first production and measurements of LGADs were presented by CNM in 2014 [47]; CNM itself produced the first production of UFSDs (active thickness of 50 μm) in 2016. The second vendor was HPK, whose first results on UFSDs have been shown at the TREDI 2017 conference [48]. A the end of 2018, HPK produced prototypes of UFSD sensors for CMS and ATLAS MTD R&D activity. Finally, FBK manufactured its first production of LGADs in 2016 [49], followed by two different UFSD productions, UFSD2 in 2017 and UFSD3 in 2018. The main targets of these two UFSD productions were the radiation hardness improvement (UFSD2) and the large area sensors design for CMS and ATLAS Endcap Timing Layer (UFSD3). Experimental measurements and results discussed in this thesis involve sensors produced by FBK and HPK, whose main features are presented below.

5.1 FBK productions

5.1.1 UFSD1

FBK, in 2016, completed its first production of LGAD sensors, named UFSD1. These sensors are manufactured on Si-on-Si high resistivity ($\rho > 5000\ \Omega cm$) p-type FZ 6-inch wafers, with an active thickness of $\sim 275\ \mu m$. The goal of this production was to investigate the device performances at different doses of gain implant; for this reason 12 wafers with 5 different gain doses of boron spaced by 2% between them, have been

Gain split	Gain dose normalised
1	1.00
2	1.02
3	1.04
4	1.06
5	1.08

Table 5.1: UFSD1, boron gain dose splits.

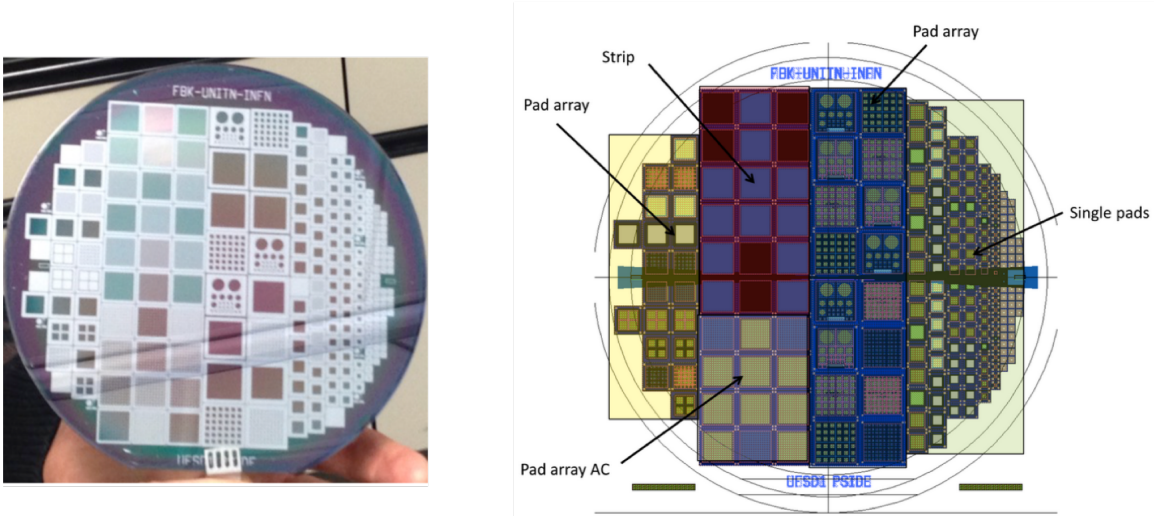

Figure 5.1: UFSD1 wafer layout.

produced, table 5.1.

UFSD1 include a large variety of structures, figure 5.1: strip sensors with different pitches, pixel sensors and single pad with areas ranging from 0.25 mm^2 to 25 mm^2. Furthermore, all the segmented devices have been manufactured in different versions of the LGAD technology [49]: front junction segmentation and back side ohmic contact.

5.1.2 UFSD2

The first production of thin LGADs by FBK has been completed in 2017 and it is called UFSD2. This production consists of 18 high-resistivity ($\rho > 3000\ \Omega cm$) Si-on-Si 6-inch p-type FZ wafers, with a nominal active thickness of $\sim 60\ \mu m$, thermally

Wafer #	Dopant	Gain dose	Carbon dose	Diffusion
1	Boron	0.98		Low
2	Boron	1.00		Low
3	Boron	1.00		High
4	Boron	1.00	Low	High
5	Boron	1.00	High	High
6	Boron	1.02	Low	High
7	Boron	1.02	High	High
8	Boron	1.02		High
9	Boron	1.02		High
10	Boron	1.04		High
11	Gallium	1.00		Low
12	Gallium	1.00		Low
13	Gallium	1.04		Low
14	Gallium	1.04		Low
15	Gallium	1.04	Low	Low
16	Gallium	1.04	High	Low
18	Gallium	1.08		Low
19	Gallium	1.08		Low

Table 5.2: UFSD2 production by FBK.

bonded on support wafer (500 μm thick); the thermal bonding reduced the wafer active thickness from the nominal value to ~ 55 μm. The main goals of this production were to establish a reliable design of thin UFSDs and to test solutions to improve the radiation resistance of the gain layer; for these reasons, different flavours of gain layers have been implemented. The acceptor dopants used for gain layer implants are boron and gallium and for the first time, the gain layer of some wafers has been enriched with carbon. The carbon enrichment has been done strictly in the volume of the gain layer to avoid a sharp increase in the leakage current. In this production, there are five different configurations of gain layers: Boron High Diffusion (B HD), Boron Low Diffusion (B LD), Gallium (Ga), Carbonated Boron High Diffusion (B HD + C) and Carbonated Gallium (Ga + C), high-diffusion and low-diffusion refer to the thermal load of gain layer activation. Ten wafers of the 18 processed have a gain layer boron-doped, the 8 remaining are gallium doped. Boron-doped has 4 splits of gain dose, while gallium-doped only 3 ones. Boron splits have dose separation steps

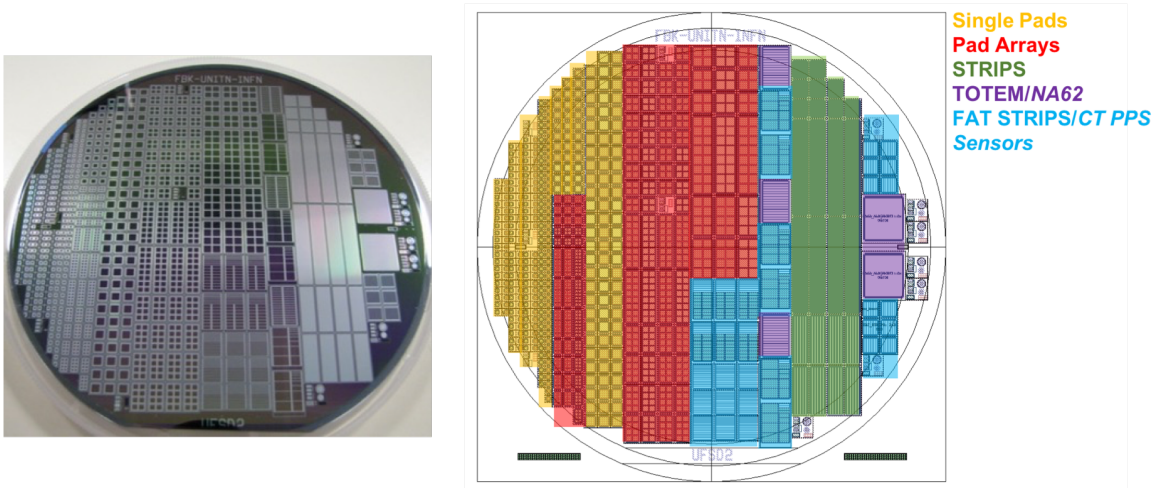

Figure 5.2: Left: picture of a UFSD2 wafer. Right: UFSD2 wafer layout and its variety of devices.

of 2%, whereas gallium splits have 4% steps; the gain dose value 1.00 is considered as reference. Two splits of gain layer B-doped and one Ga-doped are enriched with carbon. There are two carbon doses implanted into the gain layer: Low and High, which differ in an order of magnitude in carbon density. All the wafers with gain layer Ga-doped have been processed at a lower thermal load than B HD wafers due to the greater diffusivity of gallium compared to boron. Nevertheless, the width of the Ga implant is larger than the B HD implant. All information on dopants, gain layer doses, carbon doses, and diffusion are summarised in table 5.2.

Since UFSD technology has been proposed in several experiments, which have different targets, this production is characterized by devices with a large variety of geometries and dimensions, as shown in figure 5.2. The wafer layout of UFSD2 production includes:

- Single pad devices, with pad active area of 1 mm^2, 4 mm^2, and 9 mm^2, with and without a gain implant. These sensors are designed as UFSD technology demonstrators;

- Pad arrays 2x2, 2x4 and 2x8, as precursors of segmented and large area sensors

for HEP applications;

- Strips, suitable for counting in medical physics applications [50];
- Fat strips, developed for time measurements in TOTEM and CT-PPS experiments [51, 52];
- Pixels matrix of 45x40 pads, each $300 \times 300\ \mu m^2$, as first demonstrator of UFSD pixel sensors usable with the bump bonding layout of the NA62 readout chip.

5.1.3 UFSD3

The following production by FBK, called UFSD3, has been designed to investigate specific features required in sensors for ETL of CMS at HL-LHC, such as radiation hardness, a narrow inactive area between sensor pads and uniformity of high segmented large area sensors. This production consists of twenty 6-inch wafers (active thickness $\sim$ 55 μm), 16 of which with high-resistivity FZ substrate (UFSD2-like), while the remaining 4 with Epitaxial (Epi) substrate. There are four gain layer configurations implemented in UFSD3: Boron High and Low Diffusion with and without co-implantation of carbon (B HD/LD, B HD/LD + C). In this production, the gallium dopant has been excluded, since experimental measurements on UFSD2 have not shown any improvement in radiation hardness compared to boron (see section 7.3). The carbon enrichment has been maintained since the UFSD2 production demonstrated the mitigation of radiation damage in carbonated sensors. The wafers of this new production have five splits of gain dose in steps of 2% (1.00 is the dose reference, the same of UFSD2) and they have four splits of carbon dose labeled with A, B, C and D. Carbon dose A corresponds to the lowest dose in UFSD2, while the next three doses are two, three and five times dose A, respectively. The wafers of UFSD3 are listed in table 5.3.

Wafer #	Substrate	Gain dose	Carbon dose	Diffusion
1	FZ	0.98		Low
2	Epi	0.96		Low
3	FZ	0.96	A	Low
4	Epi	0.96	A	Low
5	FZ	0.98	A	Low
6	FZ	0.96	B	Low
7	FZ	0.98	B	Low
8	FZ	0.98	B	Low
9	FZ	0.98	C	Low
10	FZ	1.00	C	Low
11	FZ	1.00	D	Low
12	FZ	1.02		High
13	Epi	1.00		High
14	FZ	1.02	A	High
15	Epi	1.00	A	High
16	FZ	1.02	B	High
17	FZ	1.02	B	High
18	FZ	1.04	B	High
19	FZ	1.02	C	High
20	FZ	1.04	C	High

Table 5.3: UFSD3 production by FBK; UFSD3 W1, W12 and W14 correspond to UFDS2 W1, W8 and W6 respectively (reference wafers).

Besides splits of process, the main difference between UFSD2 and UFSD3 productions is the lithographic technique used. In UFSD2, the lithographic process has been carried out with the mask aligner, while, in this last production, the stepper technology has been used. The stepper technology uses an area $25 \times 19\ mm^2$ called reticle, which is repeated several times on the surface of the wafer. The stepper technology is substantially different from mask aligner used in UFSD1 and UFSD2, where a single pattern covers the entire surface of the wafer. The two stepper reticles used in UFSD3 are shown in figure 5.3 (right). The strength of the stepper is the greater spatial precision, which allows to create shaper implant borders and to reduce the distance between implants. Moreover, the stepper allows the photocomposition of devices to be performed, that is, to compose a device combining images from different exposures. This technique is necessary to produce sensors with an area larger

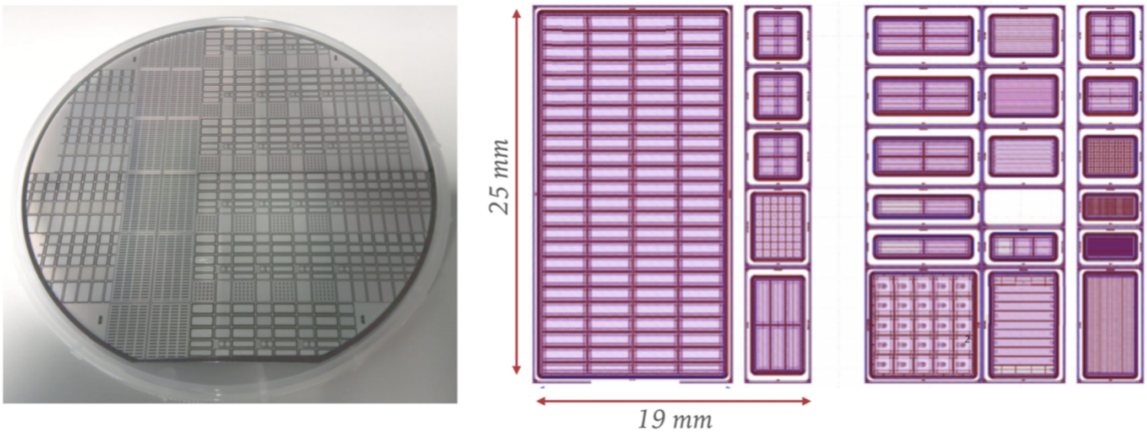

Figure 5.3: Left: picture of a UFSD3 wafer. Right: layout of the two reticles, with all the UFSD3 structures, repeated on the wafer surface.

than that of the reticle, for example the future final full-size CMS sensors for ETL. The photocomposition has been introduced in UFSD3 producing photocomposed test structures.

Figure 5.4: The three strategies of gain layer implant termination in UFSD3, from top to bottom: Aggressive, Intermediate, and Safe design.

Finally, in UFSD3 different new strategies for the gain layer termination implants (JTE and *p*-stop) have been pursued, to minimize the no-gain region between pads in multi-pads sensors. Four different termination layouts with different distances

between adjacent gain layers have been implemented: Aggressive, Intermediate, Safe and Super Safe, with a nominal distance between gain layers of $\sim 10\ \mu m$, $\sim 20\ \mu m$, $\sim 30\ \mu m$, and $\sim 40\ \mu m$, respectively; three of the four termination layouts are drawn in figure 5.4. An inter-pads region with a width lower than 50 μm is what makes these devices innovative compared to UFSD2 silicon sensors.

5.2 HPK productions

5.2.1 ECX20840

HPK, in 2017, completed its first production of UFSD sensors, called ECX20840. This production consists of 16 high-resistivity p-type wafers, of physical thickness 150 μm and two active thickness, 50 and 80 μm. Being the first HPK production based on LGAD technology, four different doping conditions have been implemented; the four gain doses are indicated in table 5.4 with numbers from 1 to 4, where 1 corresponds to the lowest and 4 to the highest one.

Very simple device geometries have been implemented in this production: single pad and pad-array 2×2, with pad size of $1 \times 1\ mm^2$.

Wafer #	Active thickness [μm]	Dopant	Doping condition
1/2	50	Boron	1
4/5	50	Boron	2
7/8	50	Boron	3
10/11	50	Boron	4
13/14	80	Boron	1
16/17	80	Boron	1
19/20	80	Boron	3
22/23	80	Boron	4

Table 5.4: ECX20840 production by HPK (2017).

5.2.2 EXX28995

EXX298995 was the second UFSDs production by HPK, completed in 2018; which was focused on R&D of sensors for ETL of CMS and High Granularity Timing Detector

(HGTD) of ATLAS. This production aimed at demonstrating the feasibility of highly segmented large area UFSD sensors. It consists of twenty 6-inch high-resistivity wafers, with an active thickness of $\sim 45\ \mu m$.

In this production gain layers with single gain dose are implanted at two different depths: shallow GL, called Type 3.1 (wafer 1-10), and deep GL, called Type 3.2 (wafer 11-20).

This production is divided equally between ATLAS and CMS, figure 5.5 shows the wafer layout and its division in two main blocks. On the left block the ATLAS devices: single pads (with and without gain), 2×2, 3×3, 5×5 and 15×15 arrays (pad dimension of $1.3 \times 1.3\ mm^2$); on the right one the CMS devices: single pads (with and without gain), 2×2, 3×3, 4×4 and 4×24 arrays (pad dimension of $1 \times 3\ mm^2$).

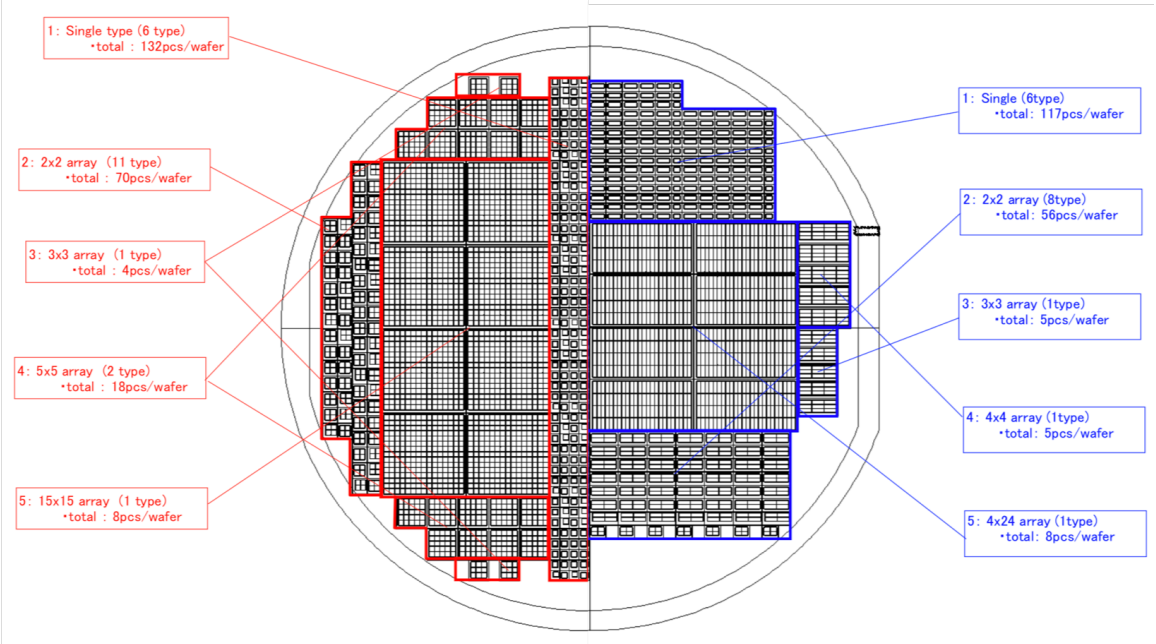

Figure 5.5: HPK EXX28995 wafer layout. On the left side of the wafer, there are devices for HGTD of ATLAS (red), on the right side the devices for ETL of CMS (blue).

Moreover, to maximize the fill factor of devices, four different inter-pad gaps have been implemented: 95, 70, 50, and 30 are the factors that indicate the four inter-pad

layouts.

5.2.3 EXX30327-EXX30328-EDX30329

In 2019, HPK produced three small batches of UFSDs (15 wafers), with an active thickness of $\sim 35\ \mu m$ and with the same wafer layout of the production EXX28995. In these batches, three different combinations of gain layer-bulk resistivity, called Type 1.1, 1.2, and 2 (table 5.5), have been implemented. Type 1.1 has a deep gain layer implant and a low-resistivity bulk ($\rho < 100\ \Omega cm$); Type 1.2 has the same deep gain layer implant of the previous one, but a high-resistivity bulk ($\rho > 3000\ \Omega cm$); while Type 2 has a shallow gain layer and a low-resistivity bulk ($\rho < 100\ \Omega cm$).

Production	Type	Gain layer implant	Bulk resistivity
EXX30327	1.1	Deep	Low
EXX30328	1.2	Deep	High
EDX30329	2	Shallow	Low

Table 5.5: UFSDs productions by HPK, in 2019.

Chapter 6

Experimental techniques

Most of the characterization of UFSD sensors have been performed in the Laboratory of Innovative Silicon Sensors in the Physics Department of the University and INFN of Turin. In this section, we will describe the two main laboratory setups used for the characterization of the devices: (i) a probe station, interfaced with a power device analyzer for static DC electrical test, and (ii) the Transient Current Technique (TCT) setup, to study the dynamic properties of the UFSD signals. Methodology and examples of typical measurements obtained with these setups are discussed in the following.

6.1 Experimental setup for static characterization of UFSD sensors

A probe station, connected with a curve tracer analyzer, is the setup used for DC electrical characterization of UFSD sensors. Current-voltage (IV), capacitance-frequency (Cf) and capacitance-voltage (CV) are the three basic DC measurements performable with this setup. The laboratory in Turin is equipped with two probe stations, each with a Keysight power device analyzer B1505 [53]. In figure 6.1 (left) is shown one of the two setups with the probe station on the left and the analyzer on the right. The probe station setup is equipped with support called chuck, where the device under test (DUT) is placed; the DUT is kept still on the chuck by a vacuum system. An optical microscope equipped with different magnifications and a video camera positioned in a dedicated opening is used to visualize the DUT. The DUT biasing occurs via the chuck and coaxial needles placed on specific contact points on the front side of the device, positioned using manipulators equipped with micrometric screws, figure 6.1 (right). Chuck and needles are connected with triaxial cables to the modules of

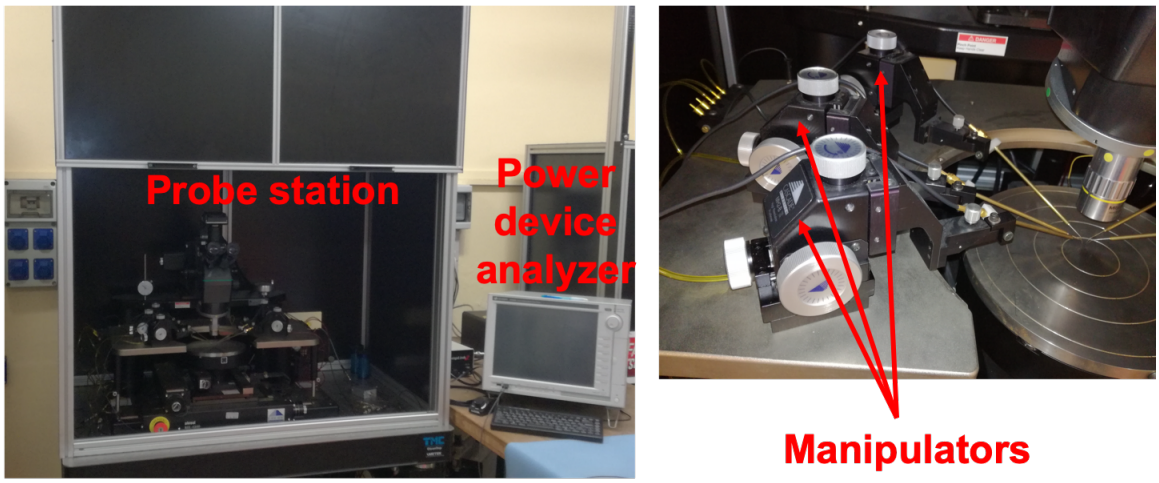

Figure 6.1: Left: probe station and Keysight B1505A power device analyzer used to perform DC electrical characterization measurements. Right: manipulators equipped with micrometric screws and contact needles.

the curve tracker; generally, the chuck is supplied with a negative bias, while the needles are grounded. The Power device analyzers in Turin are Keysight B1505A mainframes, equipped with three different types of modules:

- High Voltage Source Monitor Unit B1513C (HVSMU), with a range up to 3000 V and 8 mA;
- Medium Power Source/Measure Unit B1511B (MPSMU), with a range up to 100 V/0.1 A and a minimum measurement resolution of 10 fA/0.5 μV;
- Multi-Frequency Capacitance Measurement Unit Module B1520A (MFCMU), with a frequency range from 1 kHz to 5 MHz and an AC signal level of amplitude up to 250 mV;

6.1.1 Current-Voltage

We consider, for simplicity, as DUT a single pad device to explain the setup for IV measurements: the chuck and a needle positioned on the guard-ring are connected to the HV-SMU, while a second needle, positioned on the pad, is connected to the

MP-SMU. The MP-SMU has higher resolution and it is therefore used to measure the current flowing in the pad. A sketch of the IV setup is shown in figure 6.2. In multi-pad devices, all the pads are contacted with needles: the pads under test are wired to MPSMUs, while the remaining pads are short-circuited to the guard-ring. This configuration allows measuring with high precision several pads during a single IV scan.

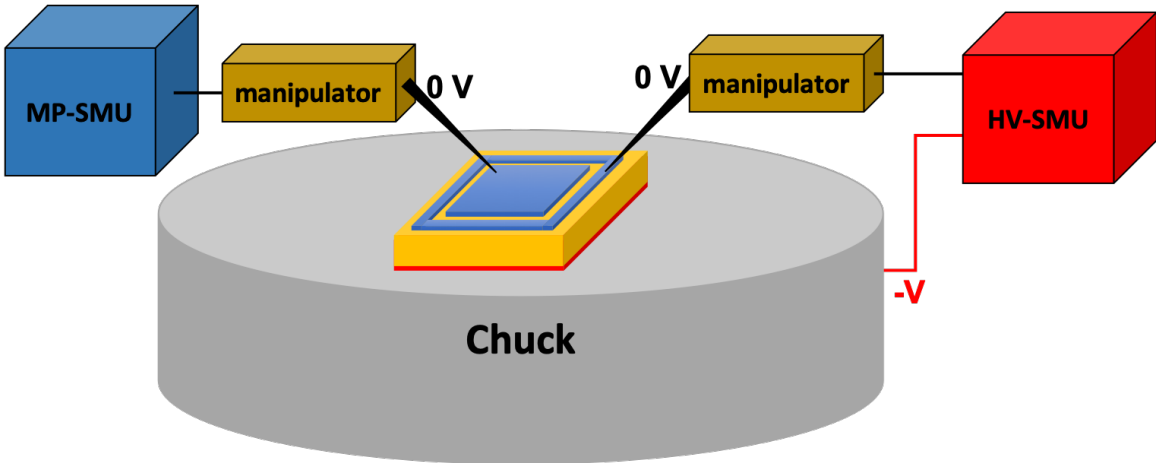

Figure 6.2: Sketch of electrical connections for current-voltage measurement on a single pad device.

The characteristics IV curves of UFSD and PiN can be compared in figure 6.3, LGAD in black, PiN in red. The PiN curve has a behaviour predicted by the Shockley equation 2.103, for a *pn* junction in reverse polarization: the leakage current reaches a saturation value and it remains roughly constant up to the breakdown voltage. The LGAD IV curve, instead, has an exponential trend above a certain bias voltage value, due to the presence of the gain layer: the gain increases exponentially as a function of the electric field (equation 4.1). At low voltage, the IV curve of an LGAD differs from one of the PiN due to the presence of a knee after tens of Volts which indicates the depletion voltage of the gain layer.

The IV curves are a good tool to distinguish UFSDs with different gains since sensors with higher gain have steeper exponential growth of the leakage current.

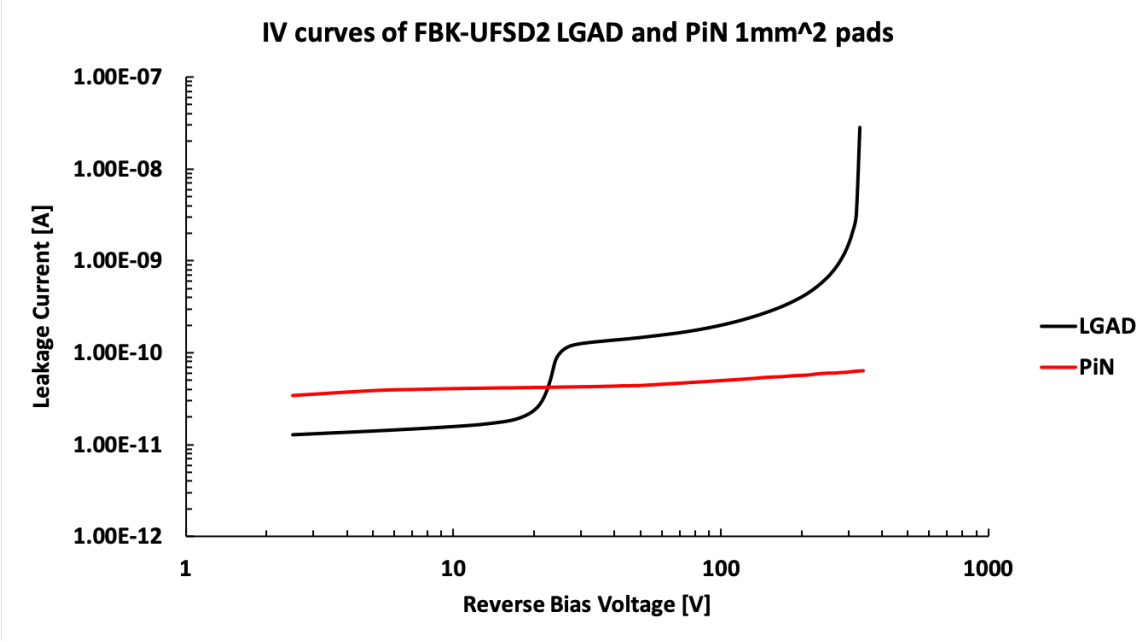

Figure 6.3: Current-Voltage characteristic curve of LGAD (black) and PiN (red).

They also identify not working devices (with premature breakdown) and can be used to characterize the yield of production.

6.1.2 Capacitance-Frequency/Voltage

The Capacitance-Voltage/Frequency measurements (CV/Cf) are performed using the HV-SMU and MF-CMU modules: the two modules are interfaced with each other via a bias-T, which receives as input the DC bias voltage from the high voltage module and the AC signal from the capacitance bridge. The bias-T output is the sum of the two inputs, split in high ($\pm V$) and low ($0V$) voltage level. For a single pad device testing, the chuck and the needle contacting the pad are wired to the high and low voltage level, respectively, while the GR is grounded, as shown in figure 6.4. For a multi-pad device, all the pads not under test must be grounded.

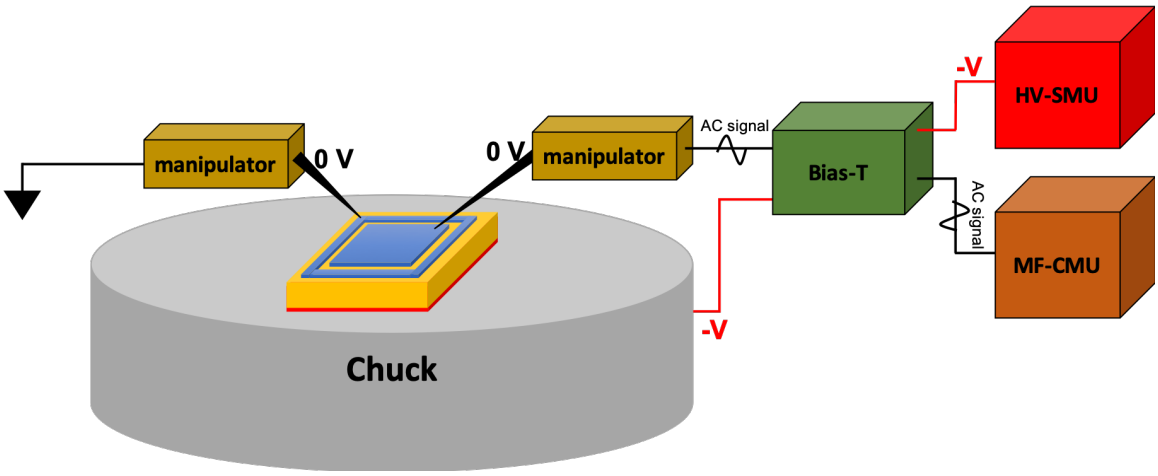

Figure 6.4: Sketch of electrical connections for capacitance-frequency/voltage measurements on a single pad device.

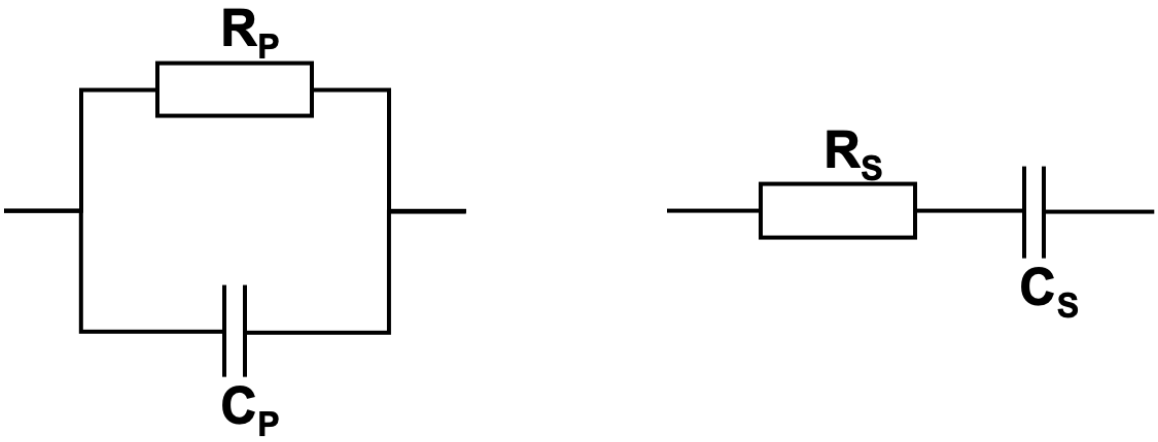

Figure 6.5: Capacitance-resistance parallel (left) and series (right) equivalent circuit.

A silicon detector with parallel plate geometry can be modelled with a parallel or series capacitance-resistance equivalent circuit, C_p-R_p, and C_s-R_s, respectively, figure 6.5. In the case of a parallel equivalent circuit, the capacitance can be obtained from the imaginary part of the admittance Y

$$Y = \frac{1}{R_p} + j\omega C_p, \tag{6.1}$$

dividing the imaginary part of equation 6.1 by $\omega = 2\pi f$, where f is the frequency of

the AC-signal. On the other hand, for a series equivalent circuit, the capacitance is obtained from the imaginary part of the impedance Z ($Z = 1/Y$)

$$Z = R_s - j\frac{1}{\omega C_s}. \tag{6.2}$$

The C_p-R_p is a good model for un-irradiated sensors: the leakage current of these devices is very low, implying that the conductivity is low. In this case, the measured admittance Y is given almost exclusively by the capacitance (see equation 6.1). In irradiated sensors, the high leakage current implies a low parallel resistance, while the high resistivity of the not depleted bulk demands a series resistance; for this reason, the capacitance measurement is difficult to understand and neither the parallel and series model represents good modelling of the detector.

Figure 6.6 shows the CV curves of a PiN and LGAD. The LGAD curve in black shows a capacitance drop (knee) at few dozen volts, this knee corresponds to the depletion voltage V_{GL} of the gain layer, which is proportional to the active doping concentration N_B and to the square of the width W of the gain layer implant. From equation 2.72 it can be written:

$$V_{GL} \propto \frac{qN_B}{2\epsilon}W^2. \tag{6.3}$$

The full depletion voltage V_{FD} of the device occurs when the capacitance becomes constant. The difference between V_{FD} and V_{GL} is the depletion voltage of the bulk, called V_{bulk}, which corresponds to the full depletion voltage of the PiN (red curve in figure 6.6).

The CV measurements are performed using a fixed frequency of the probing AC-signal provided by the MF-CMU: the selection of the optimal frequency of measurement is done by performing a capacitance-frequency (Cf) scan, using the CV setup. The sensor can be approximated to an RC network and it has a frequency-dependent

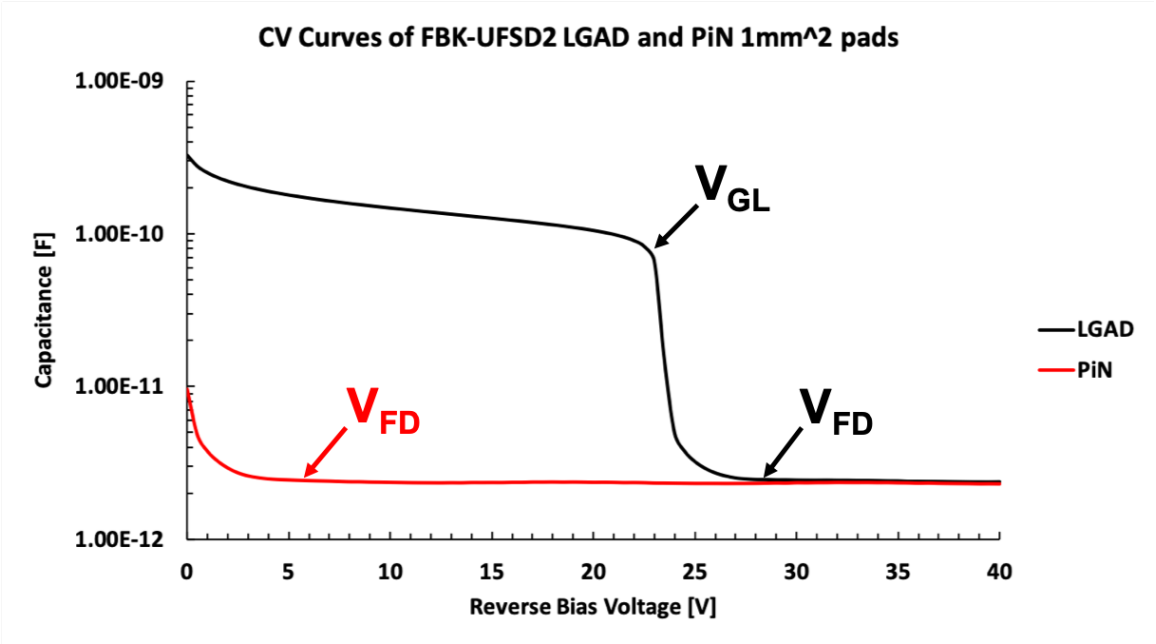

Figure 6.6: Capacitance-Voltage characteristic curve of LGAD (black) and PiN (red).

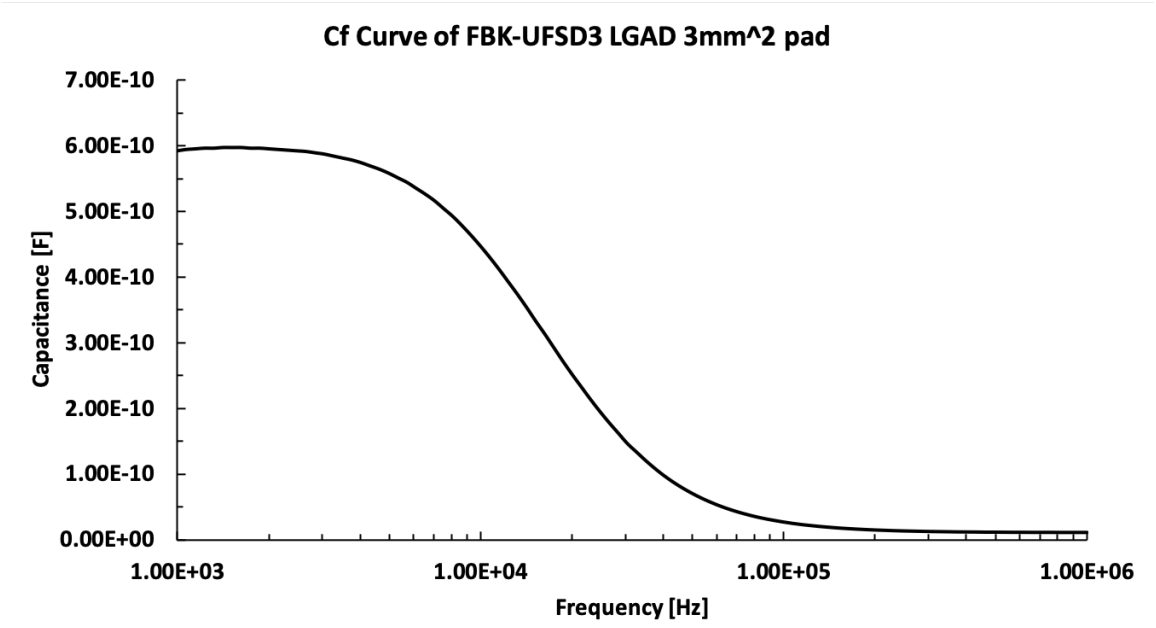

Figure 6.7: Capacitance-frequency characteristic of an LGAD at a fixed reverse bias voltage of $-2\ V$.

behaviour similar to that of a low pass filter. The Cf measurement in figure 6.7 shows a capacitance more or less constant up to the frequency of $\sim$ $10kHz$, while it decreases above this value. The drop indicates the use of a too high frequency leading to a reduction of the area probed by the measurement. The optimal measurement frequency belongs to the low frequency range, where the capacitance values are constant.

The optimal measurement frequency to perform CV measurement depends on several parameters: the sensor active area, the resistivity of the gain layer and bulk, the irradiation level, and the temperature of measurement; for these reasons, it is essential to perform the Cf characterization before any CV measurement. In the specific case of irradiated sensors it is advisable to perform CV measurements at frequency and temperature values lower than those suitable for a not irradiated one [54]. We now consider UFSD sensors irradiated up to a fluence of $3 \cdot 10^{15}$ n_{eq}/cm^2. Figure 6.8 shows Cf curves of irradiated and not irradiated UFSDs, measured at room temperature, and at a bias voltage of -10 V corresponding to a condition of partial depletion of the gain layer. For irradiated UFSDs, compared to not irradiated one, the frequency range suitable to perform CV measurements of the gain layer, at room temperature, is reduced below 2 kHz.

The frequencies suitable to explore, at room temperature, the gain layer of irradiated sensors are not suitable to explore their bulk. Figure 6.9 shows the $C^{-2}(V)$ curves of irradiated PiN diodes, performed at room temperature: increasing the irradiation fluence, the curves deviate much more from the linear trend expected from the depletion of the sensor (see equation 2.76). This deviation from a linear trend is not due to a non-uniformity of defects concentration, created by irradiation along the active thickness of the devices, but it is an artefact of the measurement due to an inappropriate measurement frequency and temperature.

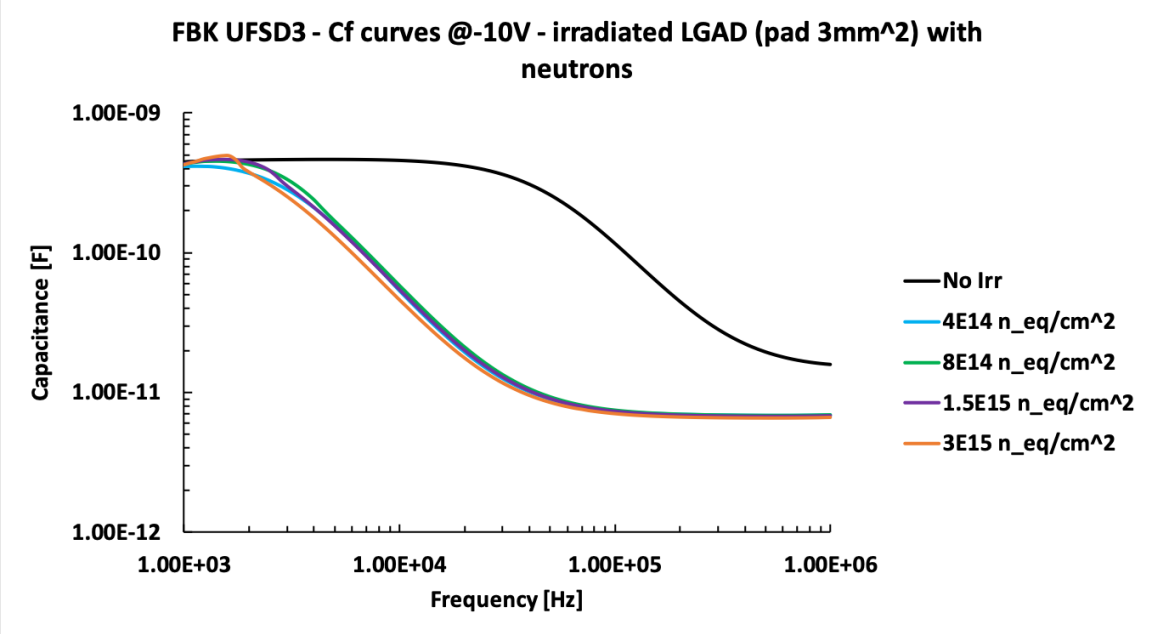

Figure 6.8: Comparison between capacitance-frequency characteristics of LGADs irradiated with neutrons up to a fluence of $1.5 \cdot 10^{15}$ n_{eq}/cm^2; measurement bias voltage fixed at -10 V, measurement at room temperature.

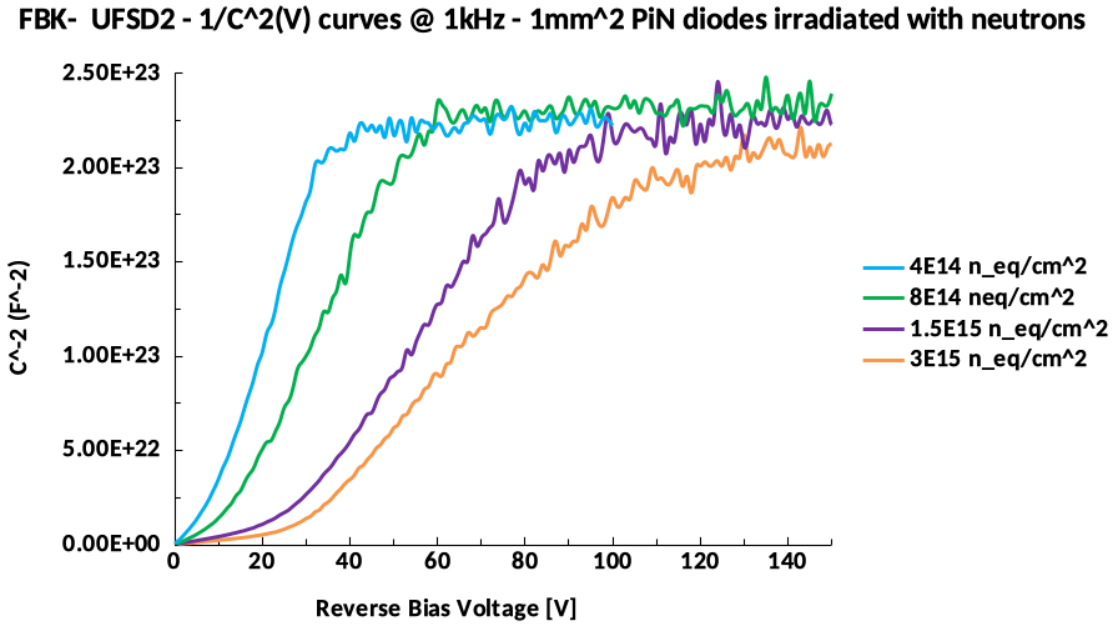

Figure 6.9: $C^{-2}(V)$ characteristic curves of PiN diodes irradiated with neutron to fluences of 0.4, 0.8, 1.5 and $3 \cdot 10^{15}$ n_{eq}/cm^2 performed at room temperature. Frequency measurement of 1 kHz.

In conclusion, a good CV measurement on an irradiated UFSD is performable using 1 kHz AC-signal at room temperature to explore the gain layer, and a lower frequency and temperature to explore the bulk.

The CV measurement is a very powerful tool to characterize LGADs. From this measurement is possible to: (i) extract the profile of the gain layer implant (equation 2.78), from which it is possible to obtain pieces of information on amplitude, width and depth of the implanted profile; (ii) evaluate the uniformity of the gain implant (equation6.3) within a production, and (iii) extract the sensor active thickness and bulk resistivity (equation 2.78). The characterization of LGADs, based on CV measurements, will be shown and discussed in detail in chapter 7, where all the CV curves on irradiated sensors have been performed at frequency values between 1 and 2 kHz, and at room temperature since the laboratory in Turin is not equipped with a cold probe station.

6.1.3 Multi-pad sensor test

The IV and CV measurement setups describe in the previous sections are used to test devices with a single pad or with a small number of pads/strips. The test of multi-pad devices, with tens or hundreds of electrodes, need the use of a multi-needle probe card, connected via a switching matrix to the power device analyzer. The probe card is a Printed Circuit Board (PCB) instrumented with needles to contact electrodes of the devices under test; the probe card replaces the manipulators in figure 6.1. The switching matrix allows selecting which pads on the sensors are measured and which pads are connected to the ground.

The probe cards in the laboratory in Turin are provided by Technoprobe [55], while the switching matrix is a Keithley 7002 Switch System equipped with four 7058 low current scanner cards, with ten channels each. A DAQ software to perform automatic measurements has been developed in LabView. In figure 6.10, it is shown

on the left a photo of a probe card, and on the right a schematic representation of the electrical connections of the measurement setup.

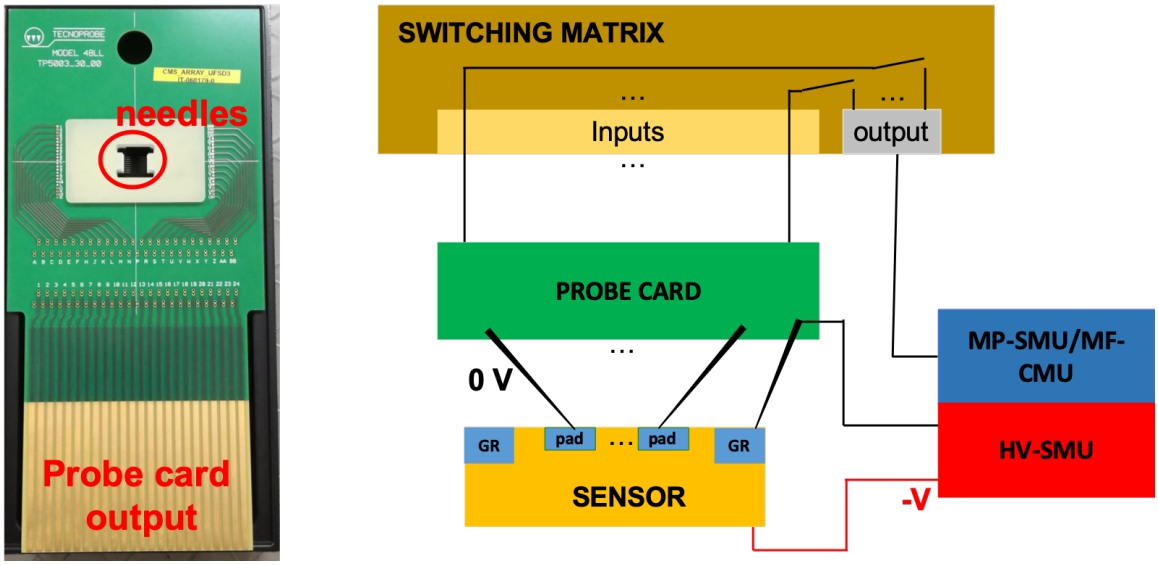

Figure 6.10: Left: probe card. Right: a sketch of the setup for multi-pad sensors testing.

6.2 Transient Current Technique (TCT) setup

The TCT setup has been developed by Particulars [56] in Ljubjana, and it is extensively used for the characterization of silicon devices, including LGADs and UFSDs. In this section, the operation principle of the TCT will be described in detail, together with some measurements on LGADs possible with this setup.

A sketch of the TCT setup is shown in figure 6.11. The transient current technique is based on Ramo's theorem (section 3.2): a focused laser penetrates the sensor generating electron-hole pairs, which induce a current signal as a function of the time on the read-out electrode; the signal is amplified by an amplifier and stored in an oscilloscope/digitizer, for offline analysis.

The TCT setup in Turin was prepared to characterize UFSD sensor: the laser system, the amplification and read-out chain have been optimized to generate and manage fast signals, with a duration in time of the order of 1-2ns. The duration of the

laser shot is 50 ps. The amplifiers usually used are Cividec [57] BB amplifiers, with a gain of 20 or 40 dB; they have input/output impedance of 50 Ω and a bandwidth of 2 GHz; the oscilloscope for the signals storage is a LeCroy with a bandwidth of 4 GHz and a maximum sampling capability of 40 Gs/s.

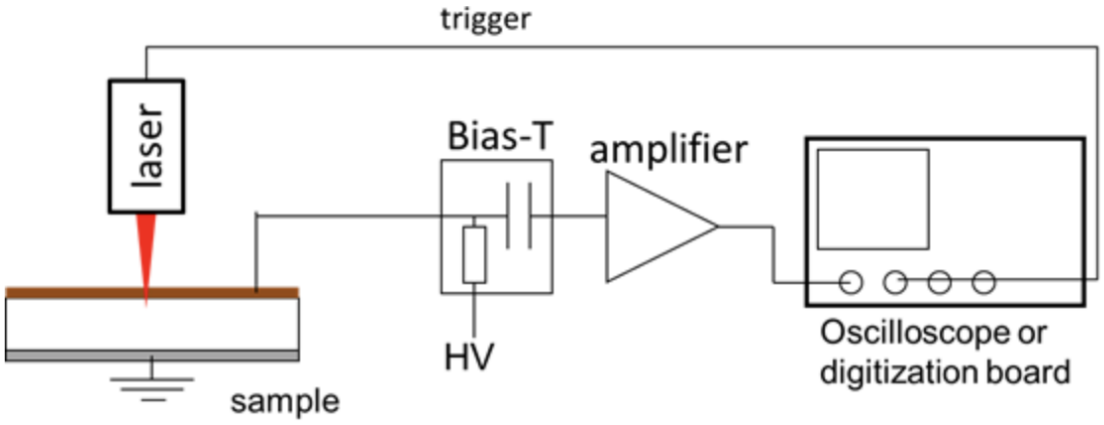

Figure 6.11: Schematic view of a typical front-TCT setup. Figure from [56].

To perform the measurements, the sensors are usually glued with conductive glue or tape on a PCB (figure 6.12) designed at the Electronics Laboratory of the National Institute of Nuclear Physics in Turin; the DUT biasing usually occurs by providing a negative bias voltage to the back of the device, while the front is wire-bonded to the output connectors on the board. All the TCT measurements in this Ph.D. thesis have been performed with the read-out chain just discussed.

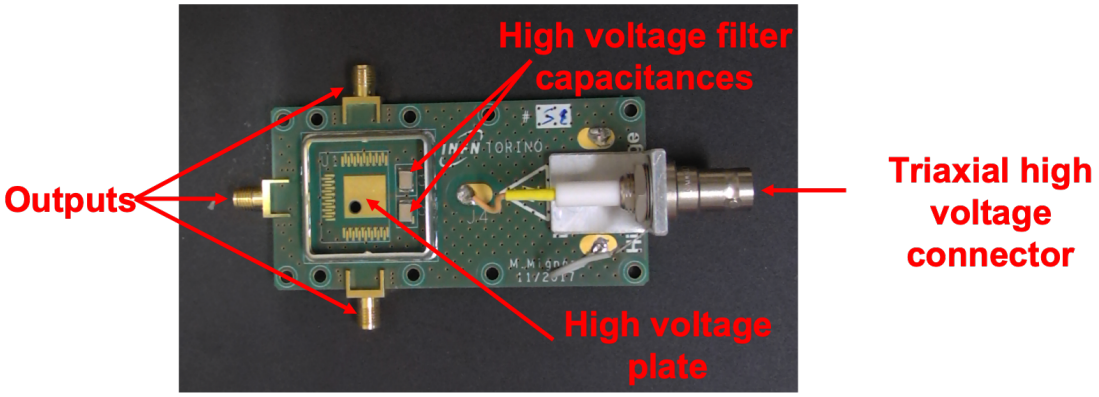

Figure 6.12: PCB designed in Turin, for the UFSD testing.

An important feature of the TCT setup is the possibility of mounting the DUT on a Newport M-ILS100PP x-y stage, which can move with a micrometric precision over the range of ten centimeters. The stage is equipped with a cooling system composed of a Peltier element and a cooled block, which is connected to a chiller Lauda-ECO-RE-1050. This cooling system allows performing measurement over a range of temperatures between -20 to 80 $°C$. In addition, the TCT setup is also equipped with a dry-air inlet in order to lower the dew point while performing low temperature measurements. The possibility to perform measurements at low and constant temperature is necessary to test irradiated sensors.

The focusing optical system is placed above the x-y stage on a z-translator with micrometric precision in movement. This system achieves a minimum laser beam spot size of $\sim 10\ \mu m$. The setup in Turin, shown in 6.13, is equipped with a pulsed laser (pulse duration of dozens ps) of two wavelengths (1060 and 400 nm) with selectable intensity. A more detailed discussion on laser pulse characteristics and focusing technique is given in chapter 3 of [58]. The infrared laser (1060 nm) has an absorption length in silicon of $\sim 1\ mm$, therefore it crosses with little attenuation the thin ($\simeq 50\ \mu m$) active thickness of UFSDs, generating via photoelectric effect e-h pairs along its path simulating the passage of a MIP. Instead, the blue laser (400 nm) has an absorption length of few microns, simulating energy deposition of an α-particle. It should be noted that the laser creation of e/h pairs has some differences compared to particles: (i) the photons interactions do not follow the landau distribution; (ii) the laser beam has a finite spot size, therefore the charges generated by a laser pulse are spread in a larger volume compared to the same ones generated by a particle. For this reason, the charges created by the passage of a particle are subject to a stronger screening effect.

The laser intensity can fluctuate during a measurement, the fluctuations are monitored by splitting the laser 10%-90%: the 10% branch is direct to an InGaAs reference

diode whose signal amplitude is monitored during the measurement, while the 90% branch goes directly to the device.

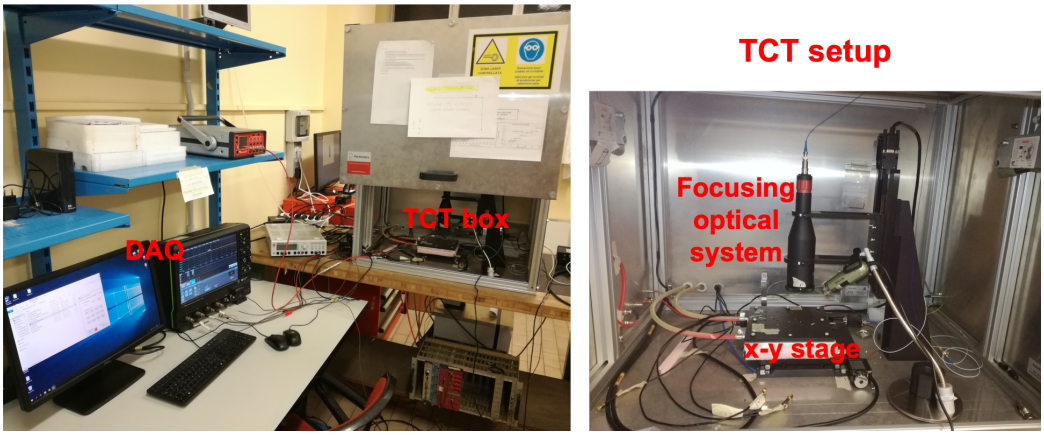

Figure 6.13: TCT setup in the laboratory of the Innovative Silicon Sensor, in Turin.

Another feature of the TCT system is the data acquisition software: the DAQ developed in LabView code by Particulars is interfaced to the oscilloscope and the x-y-z translators. It performs measurements both in manual and automatic mode (1D and 2D maps of the device) and it storages a large number of signal waveforms; the stored files can be analyzed offline with ROOT macros developed by Particulars. Figure 6.14 on the left, shows the user interface of the DAQ software, and on the right a signal waveform induced by the TCT system in a UFSD3 sensor, using $1060nm$ focused laser.

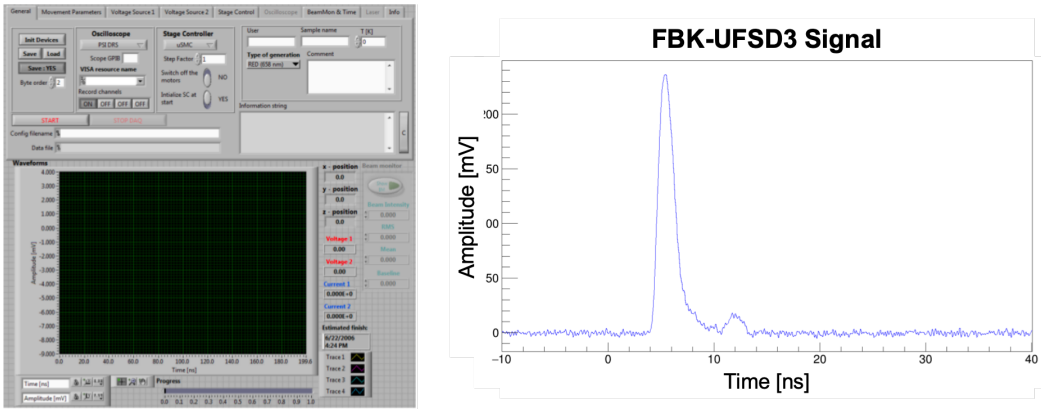

Figure 6.14: Left: the user interface of the TCT data acquisition software. Right: typical UFSD signal acquired with the TCT system.

The most relevant TCT measurements for UFSDs characterization are charge collection, internal gain, the extent of the inactive area between adjacent pads, and time resolution. The methodologies used to perform these measurements are discussed below.

- **Collected charge**: The charge Q induced by a laser shot and collected by the electrodes of the device can be obtained using this equation:

$$Q = \frac{A_{signal}}{G_A \cdot R_{in}}, \tag{6.4}$$

where A_{signal} is the area of the induced signal, G_A is the gain of the amplifier and R_{in} is the input impedance of the instrument used for signal acquisition (oscilloscope, digitizer, etc.). In the presence of fluctuations of the laser intensity, it is necessary to apply a correction factor to the collected charge Q: the correction factor can be obtained from the fit of the calibration curve in figure 6.15.

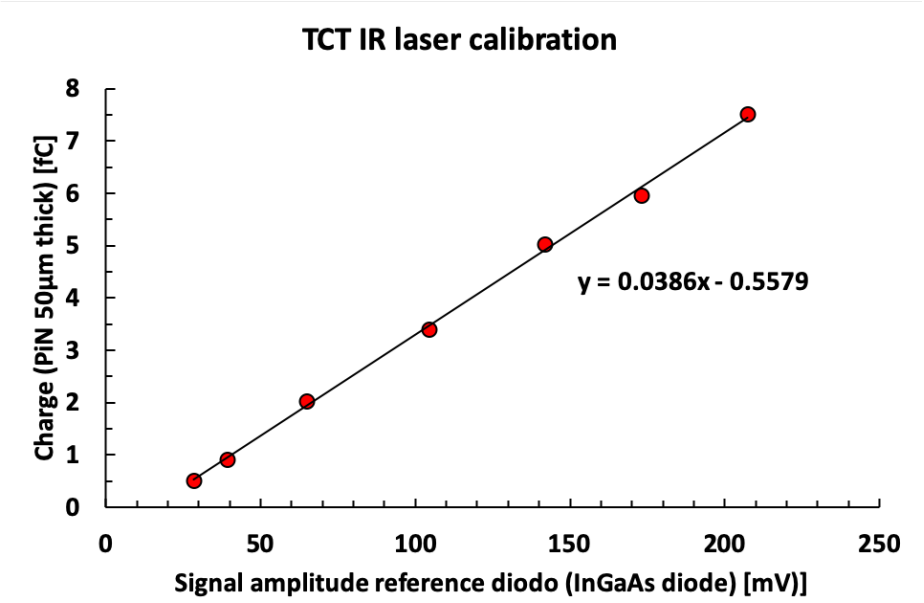

Figure 6.15: TCT IR laser calibration, induced charge in a 50 μm thick PiN diode as a function of the amplitude of the signal generated in a reference diode InGaAs, related to the Particulars laser intensity with 1 kHz frequency.

This calibration curve shows how the collected charge in a 50 μm thick PiN diode changes as a function of the signal amplitude of the reference diode InGaAs. This correction is necessary while performing measurements that require a precise determination of the collected charge.

The conversion equation 6.4 is only valid for a current amplifier; for a charge sensitive amplifier the physical quantity proportional to the collected charge is not the area of the signal, but its amplitude.

- **Gain**: the gain G is defined as the ratio between the charge collected in an LGAD (Q_{LGAD}) and that collected in a PiN diode (Q_{PiN}) with the same active thickness and in same generation condition (laser intensity and pulse duration).

$$G = \frac{Q_{LGAD}}{Q_{PiN}} \tag{6.5}$$

The charge is generated using the IR laser set to a fixed intensity; the laser intensity is usually such as to generate a number of charges equal to the one generated by few MIPs. A correction factor may also be applied to the gain measurement to correct for the laser fluctuations.

- **Inactive area width**: The measurement of the inactive area width between two adjacent pads (strips) is made by performing a TCT scan between them and acquiring the collected charge profile as a function of the laser position. Using the DAQ of the TCT, this scan is performed a hundred times for each single acquisition, aiming at measuring a mean and an error value of the physical quantities under study. The laser used for this measurement is the IR laser, focused, with a spot of about 10 μm and with an intensity that does not saturate the dynamic range of the amplifier. The expected charge profile from the acquisition consists of two S-curves, one for each readout pad; the width in the inactive area is defined as the distance between the 50% of the maximum

of the fits of these two S-curves. The S-curve shape is due to the convolution of a step function representing the response of the sensor when shooting the laser in an area where the gain layer is present, with a gaussian function, which represents the profile of the laser pulse intensity. Figure 6.16 shows a sketch of the inter-pad measurement.

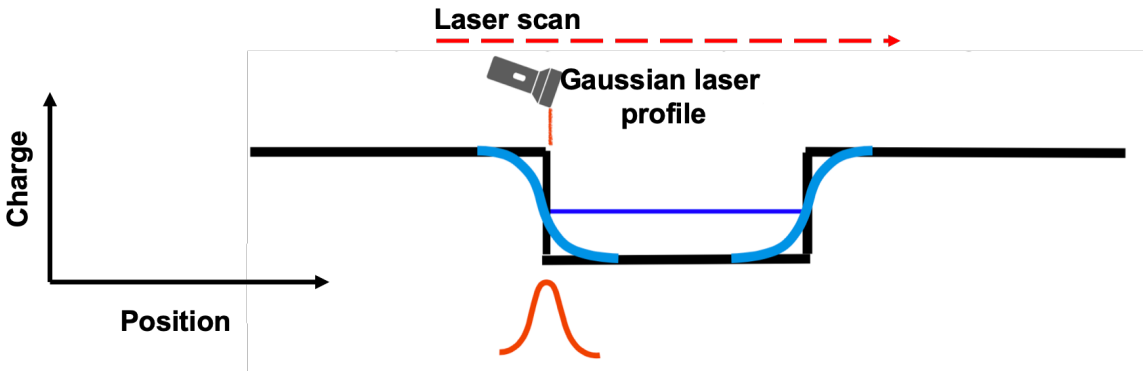

Figure 6.16: Sketch of the inter-pad width measurement based on a laser scan between two adjacent pads or strips. In blue the two charge profile collected from the nearby electrodes; the typical S shape is due to the convolution of the step function that represents the gain layer, with a gaussian that represents the profile of the laser spot intensity.

- **Time resolution**: with the TCT setup, the time resolution of a UFSD is usually measured as the standard deviation σ_t of the time distance between the induced signal in the sensor and the trigger signal generated by the laser controller, as sketched in figure 6.17. The term σ_t can be written as

$$\sigma_t = \sqrt{\sigma^2_{t-UFSD} + \sigma^2_{t-trigger}}, \tag{6.6}$$

where $\sigma_{t-trigger}$ and σ_{t-UFSD} are the time uncertainties of the trigger signal and the UFSD, respectively. σ_t can be considered a good approximation of the UFSD time resolution under the condition $\sigma_{t-trigger} \ll \sigma_{t-UFSD}$. To perform this measurement, it is important to calibrate the laser intensity (calibration curve in figure 6.15) to generate the same amount of charge in the sensor as

would be generated by the passage of a MIP. The time resolution measured with this methodology gives a good approximation of the jitter σ_{jitter} of the total time resolution of the device, equation 4.4 since the signal generated by the laser is not affected by Landau fluctuations.

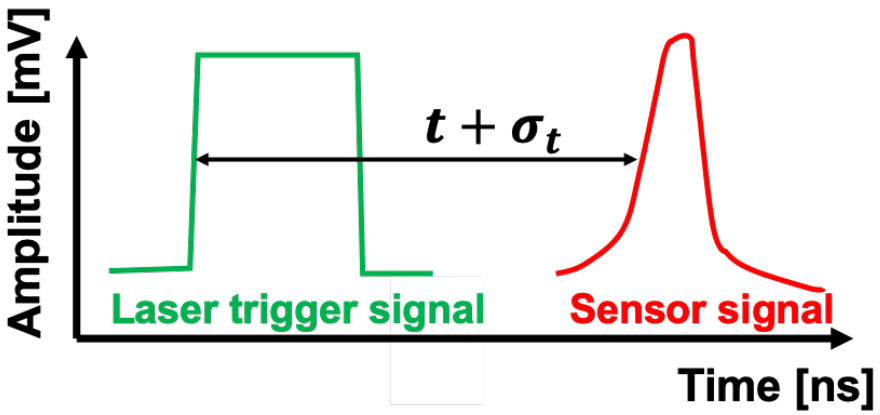

Figure 6.17: Sketch of the time resolution measurement as the standard deviation of the time distance between a laser trigger signal in green and the sensor signal in red.

Chapter 7

Laboratory measurements

After having described the main properties of UFSD or LGAD sensors (from now on these two terms will be used as synonyms), this chapter illustrates the laboratory measurements performed on devices manufactured by FBK and HPK, with the goal of fulfilling the requirements of the CMS and ATLAS experiments at HL-LHC. Hence, this chapter is focused on the measurements performed to investigate the different gain layer strategies of FBK and HPK, the radiation hardness and inter-pad inactive region of their devices, the timing performances, and finally the uniformity of large area sensors.

A selection of key laboratory measurements, based on their importance, will be shown and discussed in this chapter.

7.1 Gain layer characterization

This section focuses on the gain layer characterization using IV and CV curves and gain measurements. The main differences between the gain layer designs used by FBK and HPK will be discussed.

The depth, width, peak height of the profile of the gain layer implant, the thickness of the bulk, and the working temperature are five parameters that determine the working voltage range of a UFSD sensor. Measurements on a few sensors used as case studies will be shown to discuss the effect of these parameters on the optimal operation of a UFSD.

During the R&D activity of UFSDs, FBK produced sensors with gain layer enriched with carbon, aiming at improving the radiation hardness. In the last part of this section, the effect of adding carbon in the gain layer will be discussed.

7.1.1 Gain layer strategy by FBK and HPK

FBK and HPK R&D on UFSDs led to different designs of the gain layer region, which comprises the type of acceptor dopant, the implant depth, and the activation thermal load. CV measurements have been performed on FBK UFSDs with gain layer type B-HD/LD, Ga, and on HPK UFSDs Type 3.1/3.2. The CV curves in figure 7.1 show different depletion trends of the gain layer and V_{GL} that varies between $\sim$ 20 V and ~ 60 V, suggesting different gain layer implants.

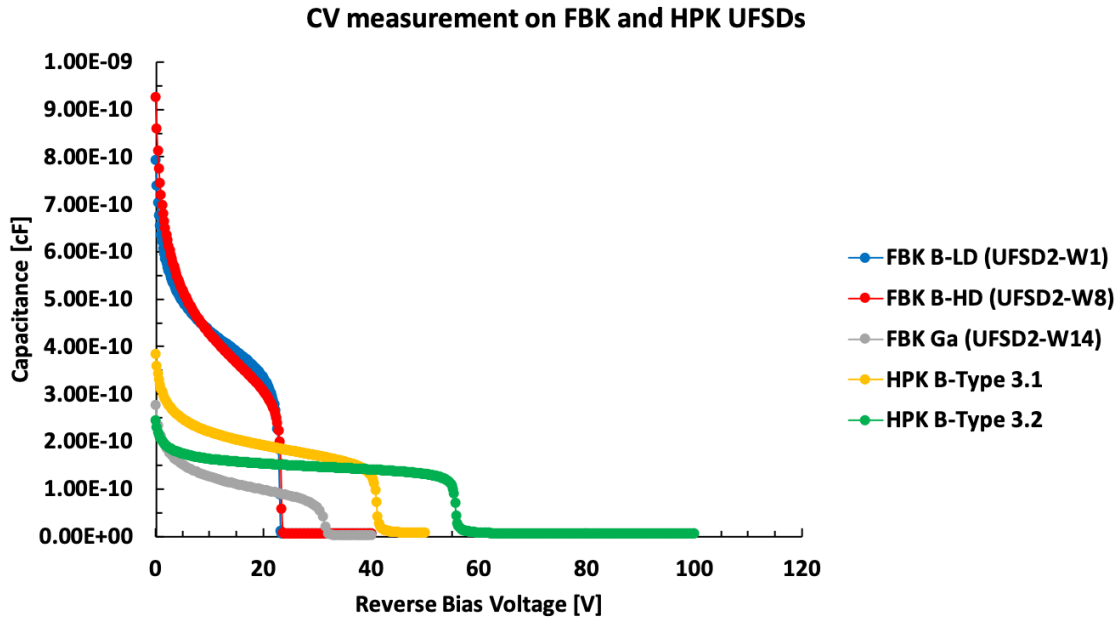

Figure 7.1: CV measurements on UFSDs with five different strategies of gain layer: FBK B-HD/LD/Ga and HPK Type 3.1/3.2.

Using equations 2.78 and 2.79, the implant profiles of each gain layer has been extracted from the CV measurements. Figure 7.2 shows the extracted profiles: on the y-axis (linear scale) there is the active doping concentration, while on the x-axis (linear scale) the depth from the edge of the depleted region in the n^{++} implant; the peaks of implants are in a depth range between 0.5 μm and 2.5 μm. The numerical values of the implants are not shown in figure 7.2 since they are sensitive data of the foundries and can not be disclosed. The FBK gain layers B-HD (UFSD2-W8) and

B-LD (UFSD2-W1) have profiles aligned in depth, but the LD is narrower and higher than the HD: this difference is due to the different activation thermal load used. The FBK gain layer with the Ga dopant (UFSD2-W14) has a deeper peak of the implant than B one. The depth of the gallium implant is different from that of the boron implant since the gallium has been implanted by an external company and not by FBK. Moreover, the Ga-profile is wider than the B one due to its higher diffusivity in the silicon lattice. The gain layer implants by HPK are both much deeper than the FBK ones.

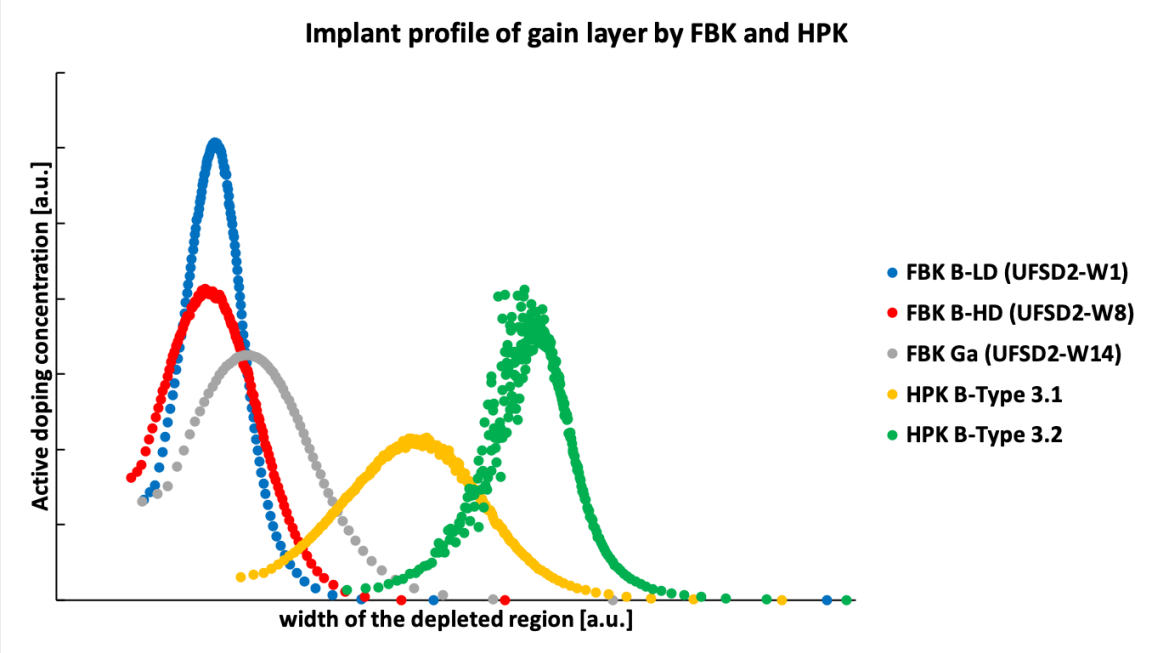

Figure 7.2: Gain layer profiles extracted from the CV measurements shown in figure 7.1.

From a comparison between CV curves and extracted gain layer profiles, it is possible to note that the relationship between V_{GL} and the concentration of acceptors at the peak expressed by equation 6.3 is incomplete; for example, the B-profiles of FBK have a peak concentration greater or equal than the HPK ones, however, V_{GL} is tens of volts lower. This difference of V_{GL} between the two vendors can be explained considering the effect of the implant depth on equation 6.3. The electric field ε_{GL} in

the gain layer rises due to the presence of free charges. At the end of the gain layer, the field is

$$\varepsilon_{GL} = \frac{dV_{GL}}{dW} = 2\frac{qN_B W}{2\epsilon}. \tag{7.1}$$

The electric field in the gap d between the gain layer and the n^{++} implant can be considered constant (assuming no doping)

$$\varepsilon_{Gap} = \frac{dV_{Gap}}{dW} = 2\frac{qN_B W}{2\epsilon}, \tag{7.2}$$

while the voltage increases linearly as a function of d

$$V_{Gap} = \varepsilon_{Gap} d = 2\frac{V_{GL}}{W} d, \tag{7.3}$$

as shown schematically in figure 7.3. The voltage of the knee in the CV curves (V_{knee}) is given by the sum

$$V_{knee} = V_{GL} + V_{Gap} = V_{GL}(1 + 2\frac{d}{W}). \tag{7.4}$$

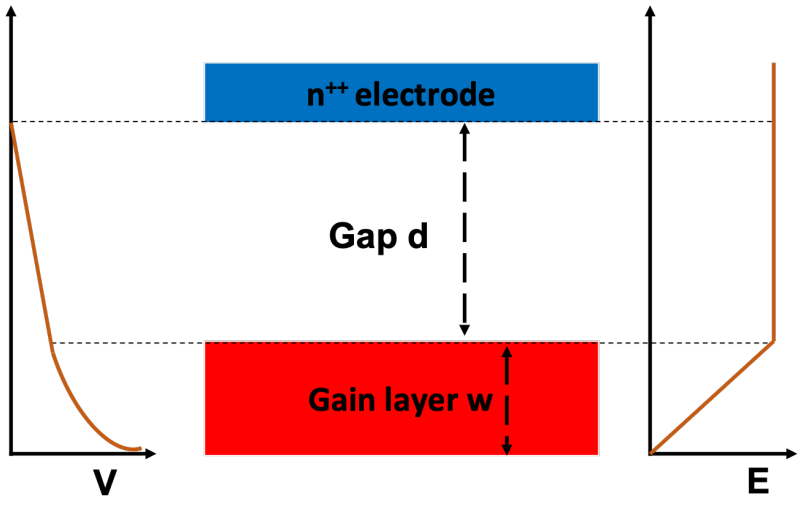

Figure 7.3: Sketch of the electric field and voltage in the gap d between n^{++} electrode and the gain layer of thickness W.

Type of gain layer implant	$V_{GL}[V]$	$V_{Gap}[V]$	$V_{knee}[V]$	V_{knee} measured $[V]$
FBK-UFSD2 B LD-W1	~ 4	~ 19	~ 23	22.4
FBK-UFSD2 B HD-W8	~ 6	~ 18	~ 24	22.6
FBK-UFSD2 Ga-W14	~ 9	~ 25	~ 34	30.5
HPK type 3.1	~ 10	~ 33	~ 43	40.4
HPK type 3.2	~ 5	~ 48	~ 53	55.2

Table 7.1: V_{GL}, V_{Gap}, and V_{knee} calculated using equations 6.3, 7.3 and 7.4 and comparison with V_{knee} measured.

Replacing in equation 7.4 the width at half height and the depth of the FBK and HPK gain layer profiles extracted from the CV measurements, it is possible to estimate the voltages of the knee. V_{knee} calculated using equation 7.4 are summarized in table 7.1; they show a good agreement with the CV curves in figure 7.1. In addition, the comparison between V_{Gap} and V_{GL} shows that the voltage of the knee is mostly due to the field in the gap.

The depth and width of the gain layer are two of the main parameters that determine the multiplication factor, the breakdown voltage due to the gain, and the working voltage of an LGAD. Comparing the IV curves in figure 7.4, which refer to the five sensors characterized, two conclusions can be drawn: (i) in the HPK sensors, with deep gain layer, the breakdown due to gain occurs at lower bias voltage than in FBK ones with shallow gain layer, as expected from the multiplication mechanism discussed in section 4.1.1 (deep gain layer to operate at high bias must be not much doped); (ii) LGADs with narrow gain layer have a lower BD voltage due to higher gain, than sensors with wide one. As case of study we consider the two sensors FBK-UFSD2-W1-(B LD) and FBK-UFSD2-W8-(B HD): from the gain measurements, figure 7.5, the first one has higher gain than the second one, despite a 4% lower gain dose; this result is in agreement with the BD voltage from the IV curves of the two sensors under test.

The third determining factor for the optimal design of a UFSD is the gain dose of the multiplication layer, which also determines the gain and the breakdown voltage of

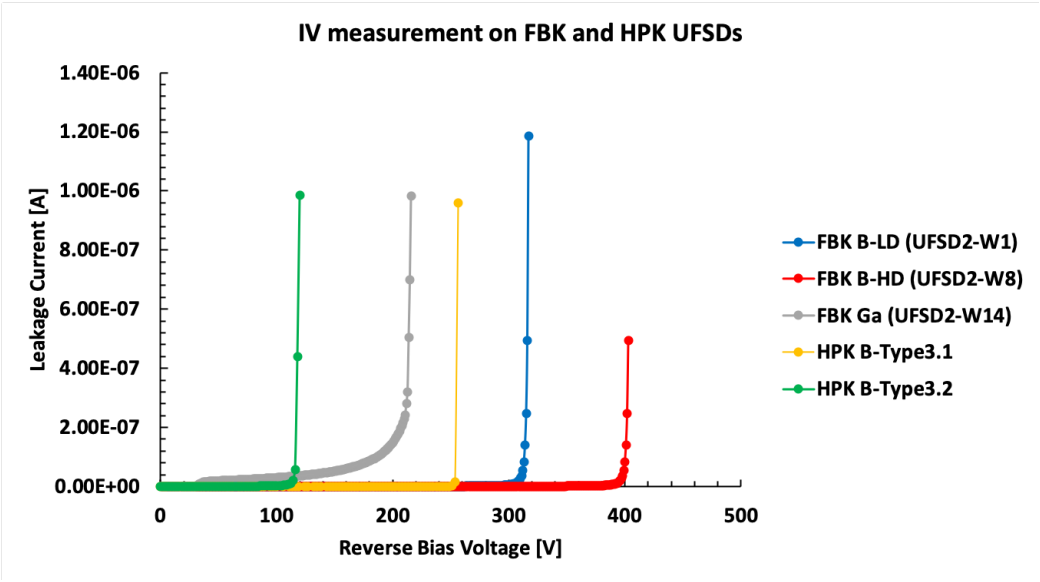

Figure 7.4: IV measurements on sensors with the five different types of implants of the gain layer shown in figure 7.2.

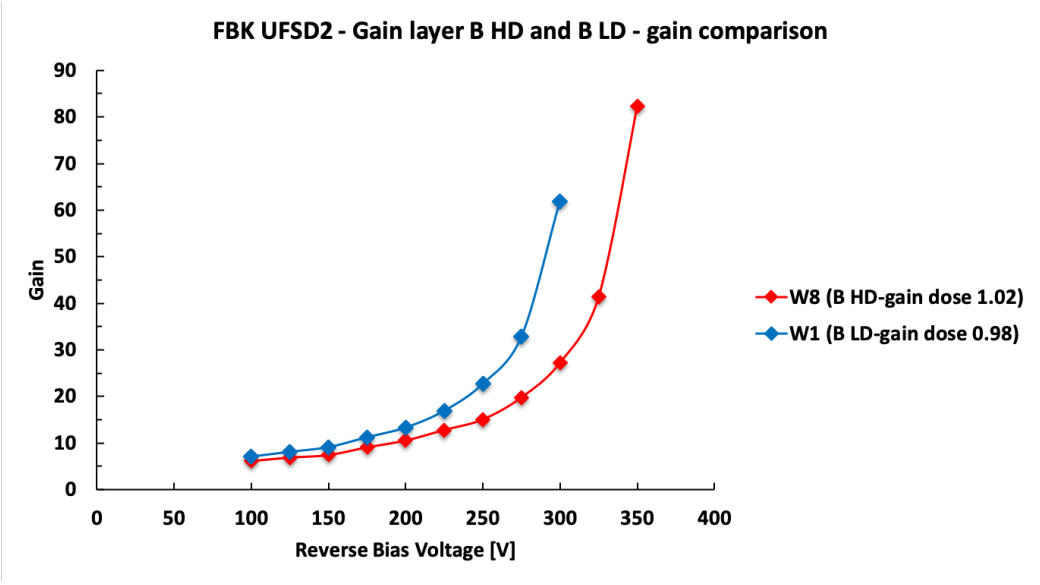

Figure 7.5: Gain measurements on FBK-UFSD2 sensors with gain layer type B HD (W8) and B LD (W1). W1 has a higher gain than W8 despite a 4% lower gain dose. These gain measurements have been performed with the front TCT setup.

the device. Figure 7.6 shows a comparison between W3 and W8 of the FBK-UFSD2 production: both wafers have a gain layer type B HD, differing by 2% in the gain dose. The top panel of the figure shows the CV curves, where it is possible to note a slight difference ($< 1\ V$) between the voltages of the knee; on the bottom panel, the gain measurements as a function of the bias. It is interesting to note that for a variation of 2% in gain dose, the gain curves shift by about $\sim 25\ V$ (only valid statement for an FBK-B HD gain layer). Hence, a small variation in gain dose can have a rather large effect on the gain and the working voltage of a UFSD.

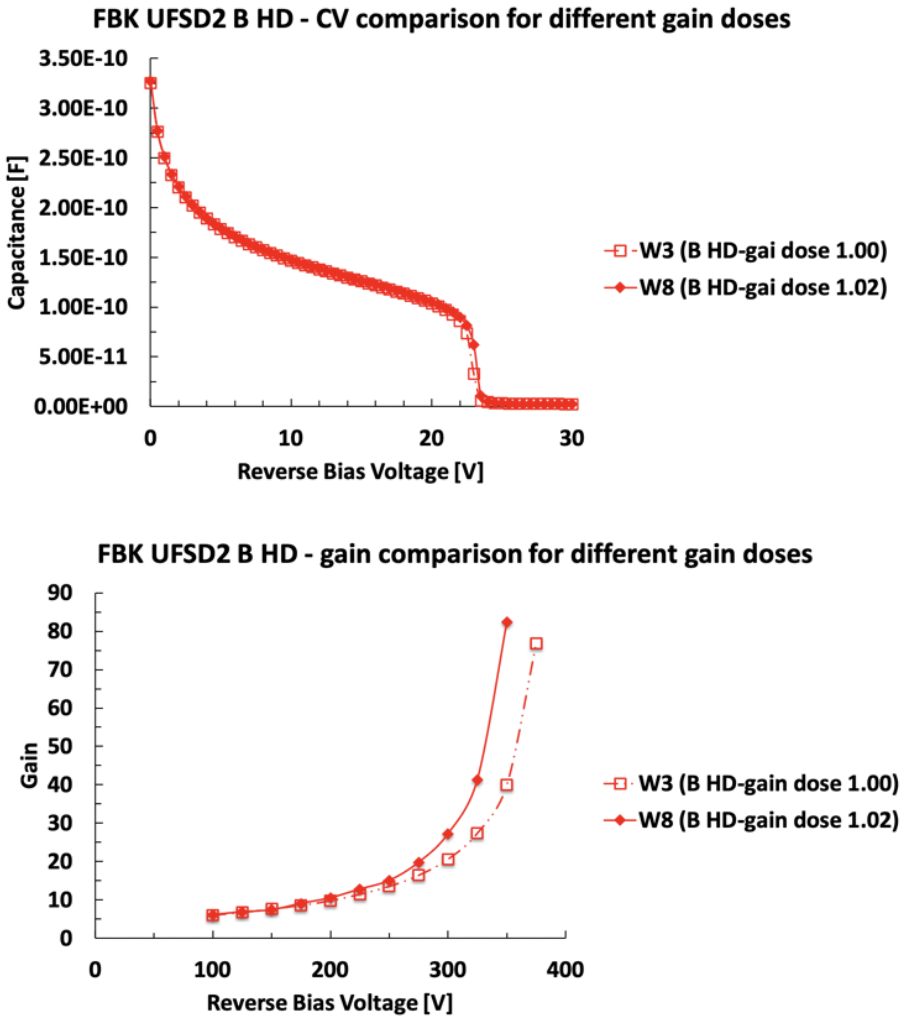

Figure 7.6: CV (top) and gain (bottom) measurements on W3 and W8 of FBK-UFSD2 production. W3 and W8 have the same gain layer type (B HD) with a difference in gain dose of 2%.

The fourth factor that determines the working bias voltage of a UFSD is the active thickness of the substrate: in sensors with the same gain layer implant and external bias voltage applied, the UFSD with thinner active thickness will have a higher gain (higher internal electric field) and a lower BD voltage. The two IV curves in figure 7.7 show the BD difference between HPK UFSDs (table 5.4), with the same doping condition but an active thickness of the bulk that differs by 30 μm. The breakdown voltage due to gain in the 50 μm thick UFSD is $\sim$ 250 V lower than in the 80 μm thick one.

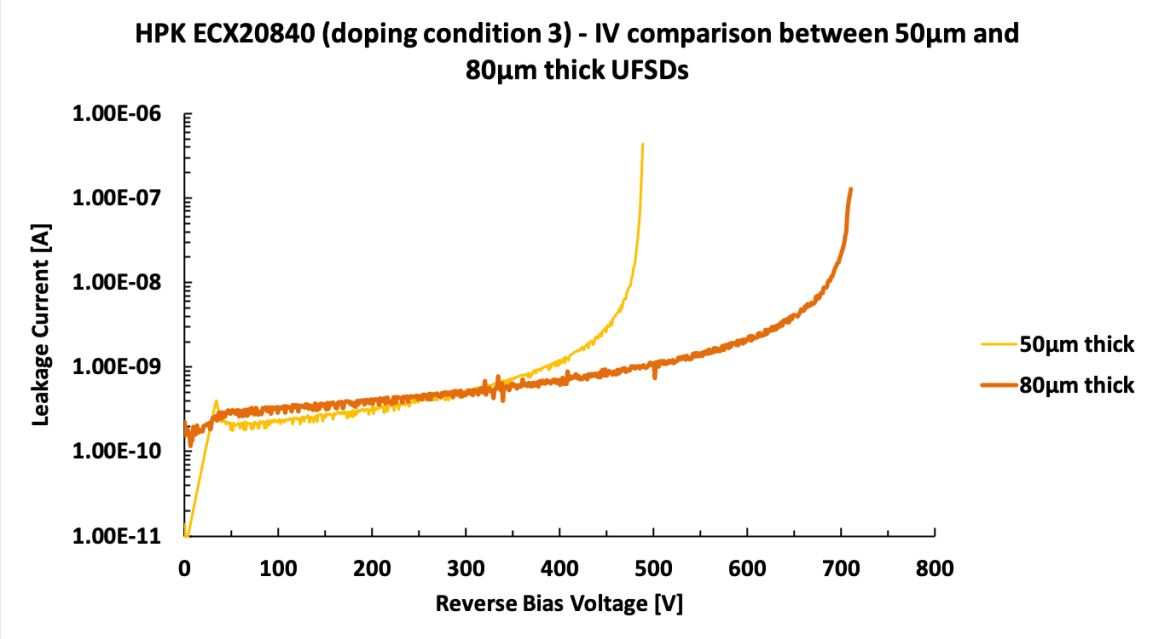

Figure 7.7: IV measurements on two HPK UFSDs, table 5.4, with the same gain layer doping concentration (doping condition 3) and bulk thickness of 50 and 80 μm.

The fifth and last factor determining the working bias of a UFSD is the temperature at which the device operates. As discussed in section 4.1.1, the avalanche multiplication mechanism is temperature-dependent (see equations 4.2 and 4.3). As shown in figure 4.4 the mean free path λ needed to achieve multiplication moves to lower electric field values as the temperature decreases; therefore, with the same external bias voltage condition, the gain of an LGAD is higher by lowering the working

temperature. Figure 7.8 shows the effect of the temperature on the gain of a HPK-Type 3.1 and a FBK-UFSD3-W5 device: by lowering the working temperature from 300 K to 250 K, the same gain value is achieved decreasing the bias voltage of 50 V and 90 V for the HPK and FBK sensor, respectively. The bias difference required to compensate for a change in temperature is greater for the FBK sensor due to the shallower gain layer, which operates in a higher electric field region than the HPK one (see section 4.8.3): λ saturates at high electric field values (see figure 4.26), reducing the gain variation capability as a function of external bias; this explains the difference in bias compensation observed in HPK and FKB sensors.

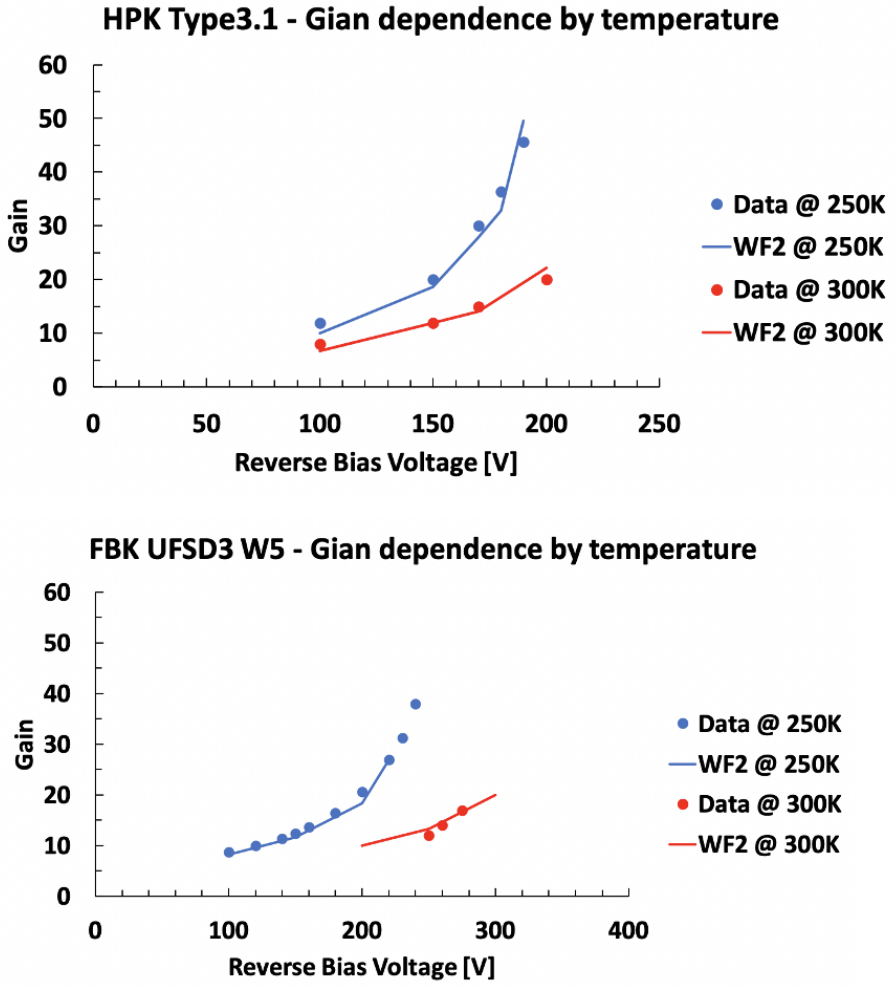

Figure 7.8: Temperature effect on the gain of a HKP-Type 3.1 and FBK-UFSD3-W5 LGAD. Markers represent experimental data, while solid lines represent WF2 simulations.

Concluding the discussion on the gain: for the same internal gain condition, a sensor working at higher bias is better to achieve good time resolution than a sensor working at a lower one. The drift velocity of the holes, contrary to electrons, never saturates, (see figure 2.4); therefore, the higher the bias, the better the drift velocity of the holes and the better the slew rate of the signal, as result, a good time resolution is achieved at lower gain [59].

In conclusion, the optimal design of UFSDs to generate an electric field inside the sensors high enough to saturate the drift velocity of charge carriers away from the BD voltage goes through a tuning work of the amplitude, width and depth of the gain layer profile, combined with the active thickness of the bulk and the temperature at which the device operates.

7.1.2 Carbon effect on the gain layer

In the UFSD2 and UFSD3 productions of FBK (tables 5.2 and 5.3), the gain layer of some wafers has been enriched with carbon atoms, in order to reduce the acceptor removal rate and to improve the gain layer radiation resistance. As discussed in section 4.8.2, the interstitial defects created by the radiation form complexes with boron atoms making them inactive, reducing the gain layer active fraction. The carbon atoms also form complexes with the interstitial defects introducing a competitive process that reduces the probability of the boron atoms to be deactivated. Therefore, with carbon addition, the deactivation of the gain layer is slowed down. This mechanism of protection will be discussed in detail in section 7.3.3.

The carbon enrichment produces also additional effects: (i) the leakage current increases, as shown in figure 7.9, without however causing degradation in sensor performances, and (ii) the capture of a fraction of the implanted boron atoms, leading to a less doped gain layer. The carbon-boron capture can be appreciated by comparing the depletion voltages of FKB sensors with the same initial active doping

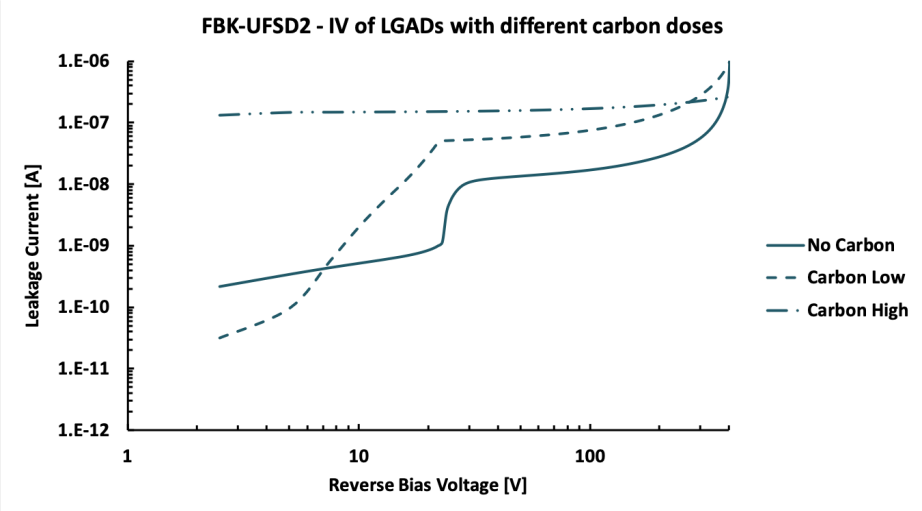

Figure 7.9: Carbon effect on the leakage current in UFSD2 LGADs: carbon enrichment increases the sensor leakage current.

concentration in the gain layer, but with different carbon densities.

A first comparison has been made on UFSD2 sensors, between wafer 6 (carbon) and wafer 8 (no carbon): the depletion voltage of the carbonated wafer is lower by $\sim 0.3\ V$ than not carbonated one. A similar result, with a difference of $\sim 0.7\ V$, has been obtained in the comparison of wafer 3 (no carbon) with wafer 4 (carbon). Given the precision of the CV measurements, $\sim 0.1\ V$, it is possible to conclude that the differences in depletion voltage are due to the presence of carbon in the gain layer. The fraction of inactive boron due to carbon has an important effect: the gain is lower in carbonated LGADs at the same external bias voltage. Figure 7.10 shows a comparison between the gain curves measured on W3, W4, W6, and W8: above 300 V, a clear difference of gain is evident in carbonated wafers with same initial boron concentration; moreover, the gain of W6 is very similar to that of W3, as result, the low carbon dose in UFSD2 has the same effect as decreasing the initial active boron concentration by $\sim 2\%$ (difference in gain dose between W8 and W3).

The carbon-boron capture has been studied using CV measurements also in the UFSD3 production, figure 7.11. UFSD3 wafers have been enriched with four different doses of carbon: the lowest one (dose A) corresponds to dose Low in UFSD2, the

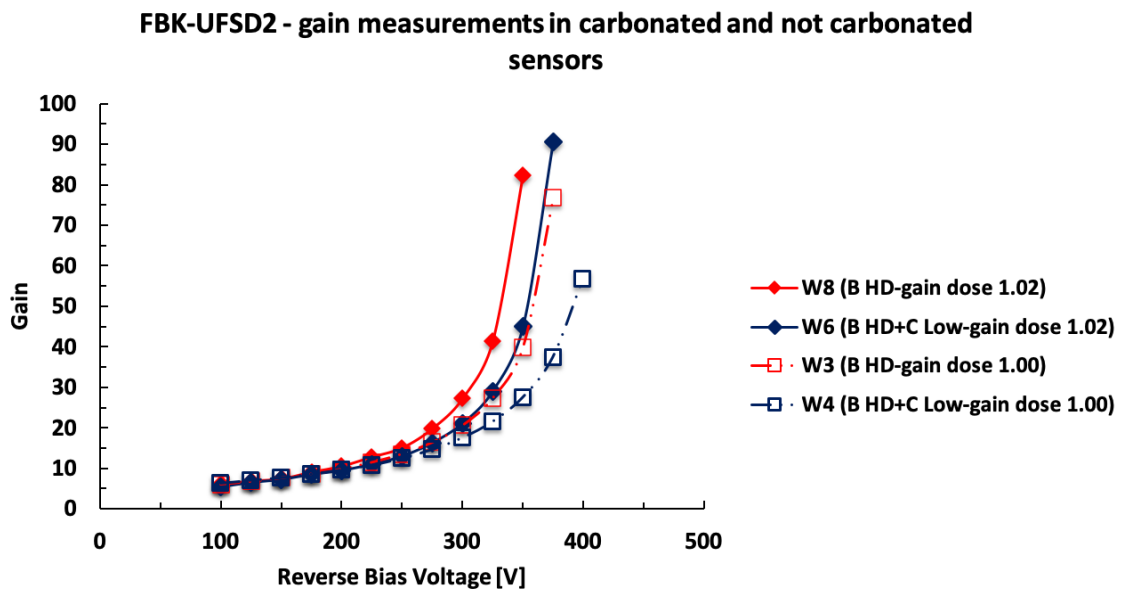

Figure 7.10: Gain measurements on UFSD2 W3, W4, W6 and W8. Carbon enrichment decreases the internal gain of UFSDs.

highest one (dose D) is five time the dose A, and the two intermediate doses (B and C) are two and three times the dose A, respectively.

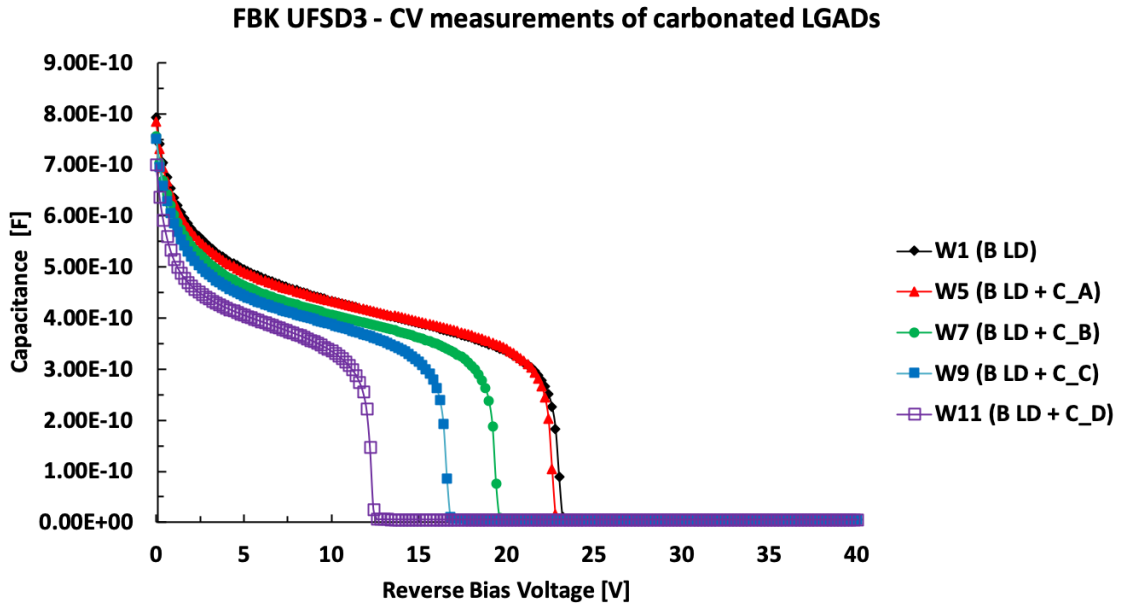

Figure 7.11: CV measurements on UFSD3 sensors enriched with different carbon doses: the depletion voltage of the gain layer shifts at lower values due to the carbon enrichment.

A difference of $\sim 0.5\ V$ has been measured from the comparison of the wafer 5 enriched with carbon dose A and the wafer 1 not-carbonated, validating the results obtained on UFSD2. Measurements on UFSD3 sensors enriched with dose B, C and D, and on UFSD2 sensors enriched with carbon dose High show a stronger boron deactivation: at these carbon doses, the depletion voltage of the gain layer with respect to the not carbonated sensors decreases by $\sim 3\ V$, $\sim 5\ V$, $\sim 10\ V$, and $\sim 20\ V$, respectively.

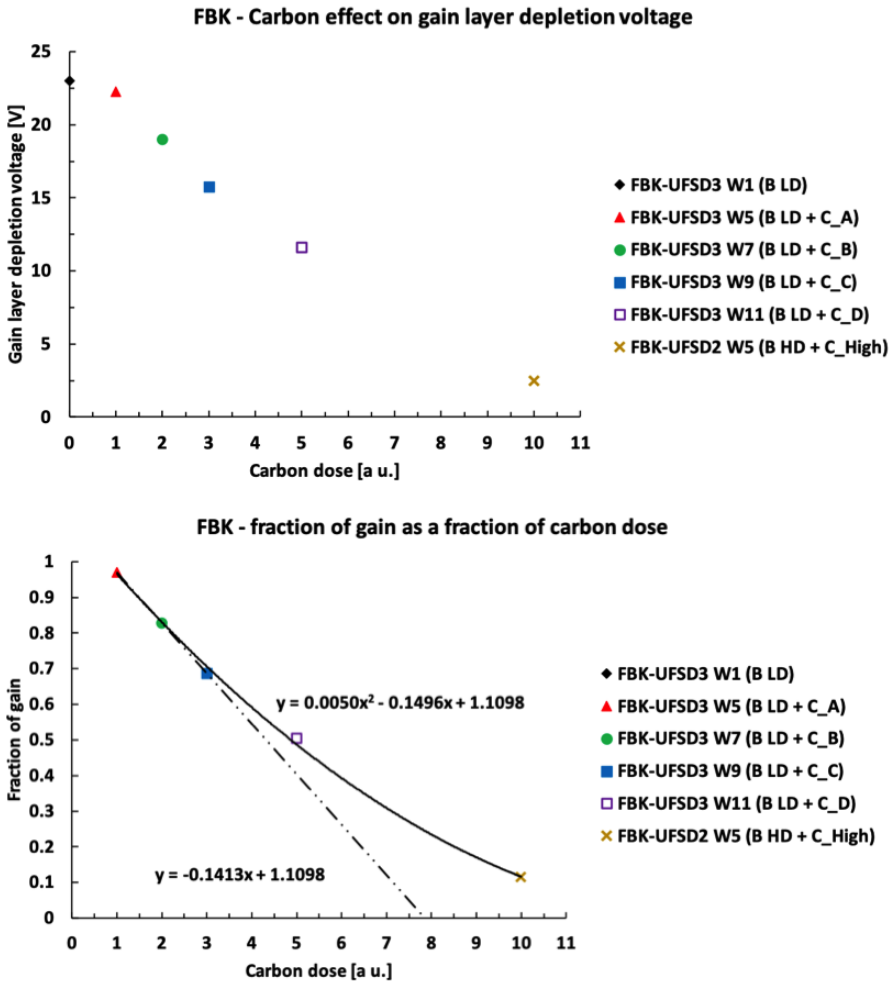

Figure 7.12: Depletion voltage (top) and the active fraction of the gain layer (bottom) as a function of the carbon dose for FBK-UFSD2 and UFSD3 sensors.

Figure 7.12 (top) shows the gain layer depletion voltage as a function of the carbon dose normalized to the dose A, while the plot on the bottom shows the active fraction

of the gain layer ($V_{GL}(carbon)/V_{GL}(no\ carbon)$, from equation 6.3) as a function of the carbon dose. From the plot on the right it is possible to observe that the deactivation of the gain layer increases linearly as a function of the carbon dose in a dose range between A and C. For doses below A, the boron deactivation is weaker, suggesting that the carbon-boron capture could be a threshold mechanism, linear above a minimum carbon dose of $\sim$ 0.8A; while for carbon doses above C, the carbon-boron capture tends to saturate.

The carbon enrichment has a secondary effect on the gain layer: it limits the lateral diffusion of the boron implant during the activation. The gain layer profile, extracted from CV measurements, on UFSD3-W5 (B LD + C-A), is higher and narrower concerning that of a sensor with the same type of gain layer without carbon, UFSD3-W1 (B LD) without carbon. Figure 7.13 shows a comparison between the gain layer profiles for these two wafers. This effect has a significant impact on the performances of the irradiated UFSDs, see section 7.3.3.

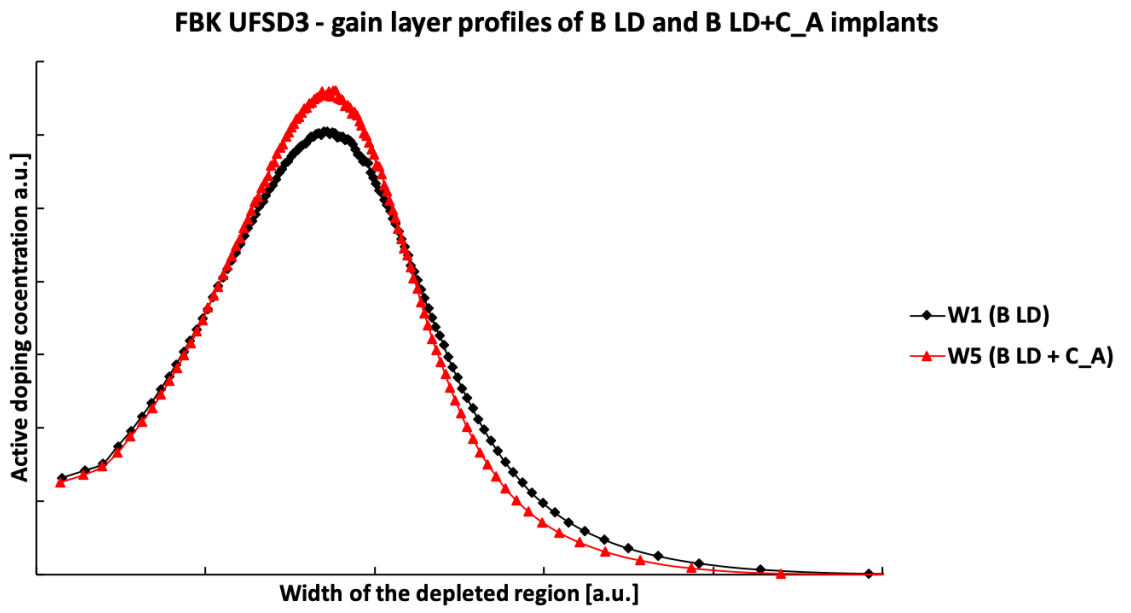

Figure 7.13: Gain layer profiles of a UFSD3-W5 (B LD + C-A) and UFSD3-W1 (B LD), the implant profile in the carbon enriched wafer is narrower and higher than in the not enriched wafer.

7.2 Inter-pad region studies

Several types of inter-pad designs in FBK-UFSDF3 and HPK-EXX2899 productions have been characterized. In this paragraph, the study of the design of the inter-pad region will be presented. The measurements of the extent of the inactive region (no-gain) between neighbouring pads have been performed with TCT setup on devices from both foundries. On the UFSD3 devices, this study has been completed with the use of a low noise CCD camera, in order to identify the location of premature breakdown. The paragraph ends with a discussion of the pop-corn noise, that is, intermittent microdischarges very similar to signals generated by the passage of particles.

7.2.1 Breakdown and inter-pad in FBK-UFSD3

Two types of UFSD3 devices from W1 have been selected to characterize the four inter-pad layouts used in this production: devices CMS 2×2 array with Aggressive, Intermediate and Safe layout, figure 7.14 (left); a device MoVeIT strips (fat strips) with Super Safe layout, figure 7.14 (right). These devices have a special optical window (i.e. a region without metalization) of dimension $\sim 20\ \mu m \times 200\ \mu m$ between two adjacent pads or strips, which allows shooting the laser in the inter-pad region.

IV measurements have first been performed on these devices, connecting all pads or strips and guardring to the ground level of the measuring instrument. Sensors with the Aggressive, Intermediate and Safe layouts have an earlier breakdown voltage compared to the Super Safe design, and it is visible in the IV measurements shown in figure 7.15. The BD of Super Safe sensor occurs at $\sim 300\ V$ and it is due to the internal gain, this can be deduced from the exponential trend of the current as a function of the bias. For other layouts, the BD voltage is strongly related to the width of the inactive area, since it decreases for narrower inter-pad regions. These

measurements have been repeated on several devices, showing good uniformity.

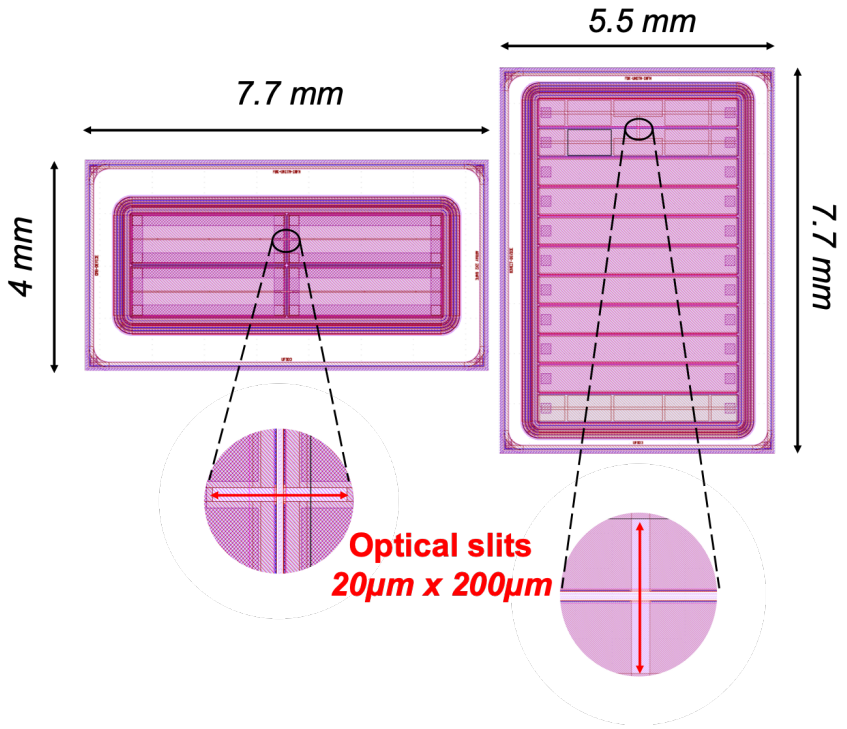

Figure 7.14: The layout of the FBK-UFSD3 sensors selected to characterize different inter-pad regions. On the left, the CMS 2×2 array sensor (Aggressive, Intermediate, and Safe layout), on the right the MoVeIT strips sensor (Super Safe layout); the circular windows show a zoom of the optical slits present on these devices.

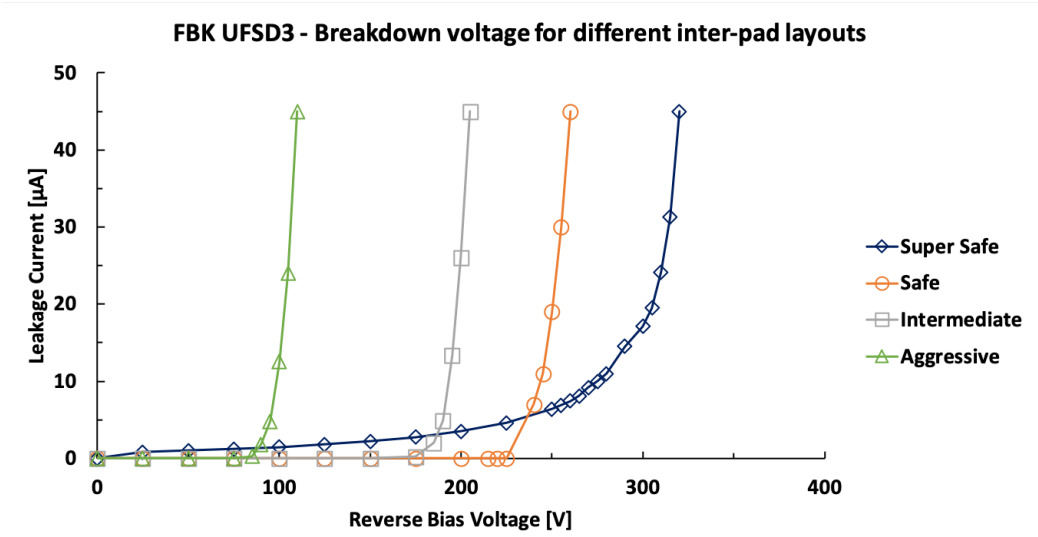

Figure 7.15: Breakdown measurements on FBK-UFSD3 termination layouts: Super Safe, Safe, Intermediate and Aggressive.

Aiming at investigating the phenomenon of premature breakdown, each device underwent a full x-y TCT scan in the region between pads, at bias voltages close to BD. Figure 7.16 shows an example of TCT scans performed on a CMS 2×2 array-Safe at three different bias: 200 V, 250 V and 260 V (BD voltage); the three 2D maps show the amount of charges (colour scale) collected from the top right pad (pad 1) of the sensor. From these maps of charge, it is noticeable that as the bias increases, becoming close to the BD voltage, the amount of collected charge increases in the corner of the pads in the center of the array.

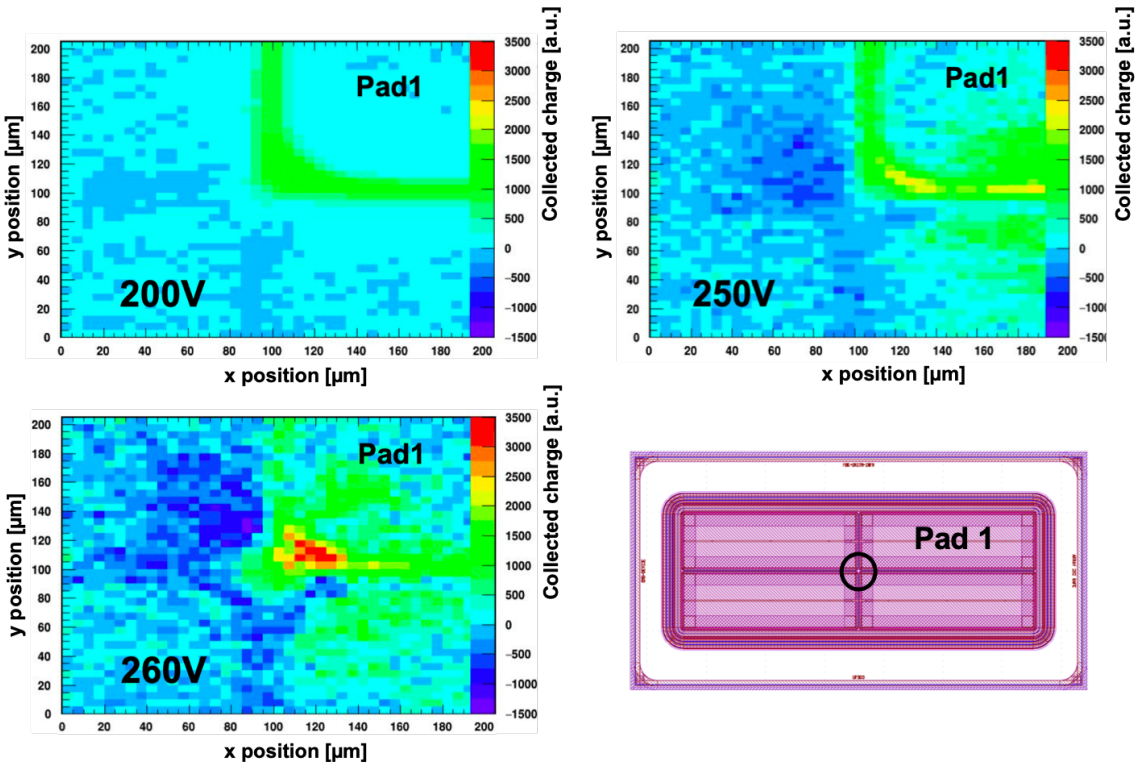

Figure 7.16: Maps of collected charge in a pad of W1-CMS 2×2 array-Safe device, at bias voltages of 200 V, 250 V, and 260 V (breakdown voltage). In the bottom-right corner the device layout, with the TCT scan area highlighted.

This phenomenon is anomalous since in this region there are only termination structures. The "hot spot" (region with high current density) in the corner of the pad is due to a high surface electric field, which generates a local gain that leads to early breakdown. In these maps, the effect of charge multiplication due to the gain

layer is not visible because it is covered with a metal layer. The same measurements have been performed on MoVeIT strip-Super Safe device, close to the corner of the strips, at bias 300 V, 320 V and 330 V (BD voltage). The maps of the collected charge in figure 7.17 do not show any hot spots on the strip under test, as expected from the high BD voltage seen in the IV curves.

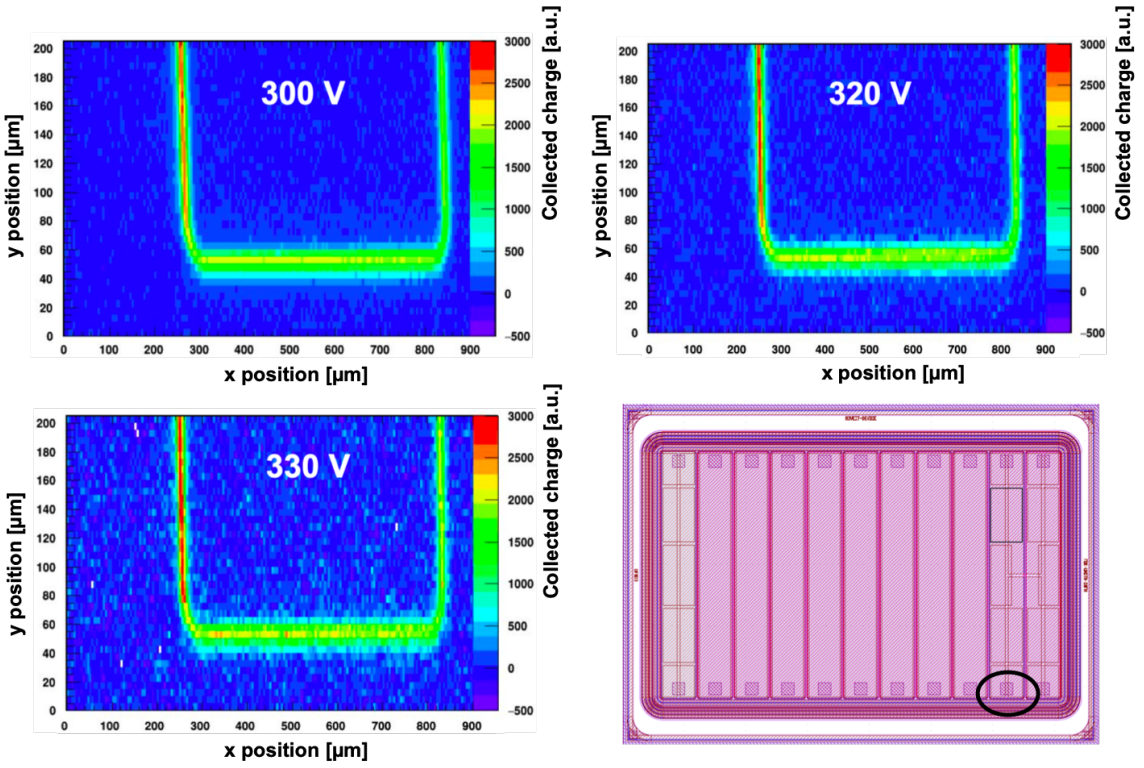

Figure 7.17: Maps of collected charge in a strip of W1-MoVeIT strip-Super Safe device, at bias voltages of 300 V, 320 V and 330 V (breakdown voltage). In the bottom-right corner, the device layout with the TCT scan area highlighted.

From the charge maps in figures 7.16 and 7.17, the collected charge along the edge of the readout pad and strip have been extracted. On the left side of figure 7.18, at the corner of the readout pad, the profiles obtained from the Safe sensor shown an increase with increasing bias. On the right side, the same curves from the Super Safe sensor do not show this trend. Similar measurements have also been performed on Intermediate and Aggressive devices, showing the same behaviour of the Safe one.

To further explore the hot spots phenomenon, the TCT measurements have been

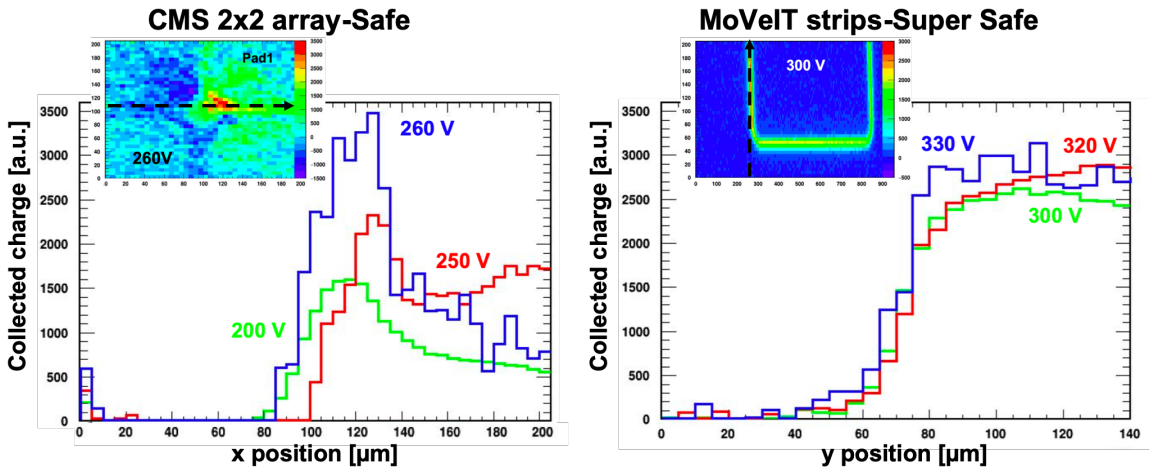

Figure 7.18: Projections of the collected charge along the edge of a pad in an UFSD3-CMS 2×2 array-Safe device and of a strip in a UFSD3-MoVeIT strips-Super Safe device, at three different value of bias voltage.

complemented by similar measurements performed with a CCD camera mounted on the optic of the probe station. The CCD camera used is the Hamamatsu C11090-22B EM-CCD camera, able to detect the photons emitted by the hot spots in the visible spectrum. For each sensor under test, several photos have been taken in complete darkness, at the same bias voltages of the TCT setup. These photos have been superimposed to a photo taken with light, in order to locate the hot spots on the surface of the device. Figure 7.19 shows the photos taken on the sensors CMS 2×2 array-Safe, the same one used in figure 7.16. The hot spots appear in the corners of the pads, confirming the results obtained with the TCT. Similar images have been obtained for the Intermediate and Aggressive devices, finding identical results. For the Super Safe sensor, no hot spots appear close to the BD voltage, in agreement with the TCT measurements. In conclusion both TCT and CCD camera measurements, on Safe, Intermediate and Aggressive devices, show a breakdown due to high electric fields at *p*-stop and JTE termination structures.

Premature BD has also been studied in irradiated sensors, showing as the irradiation cures it in Aggressive, Intermediate and Safe layouts. Figure 7.20 shows

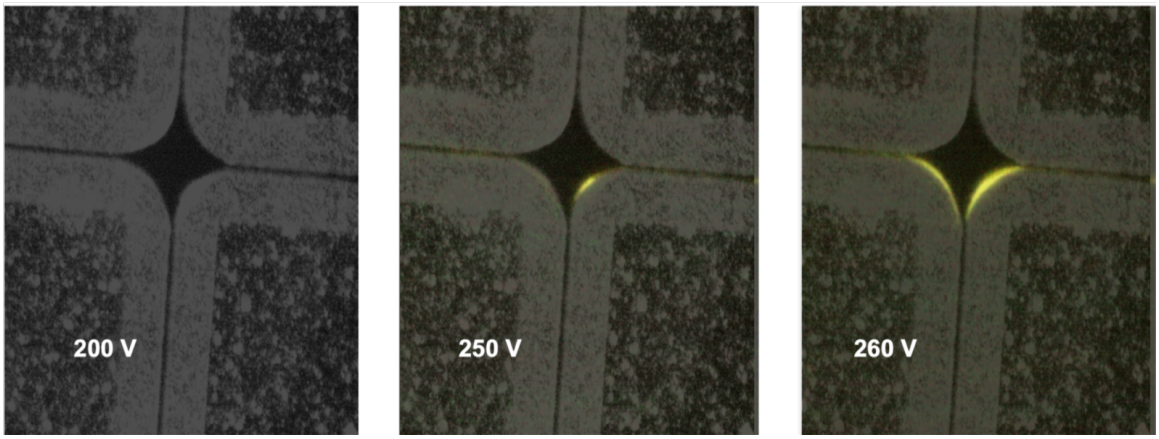

Figure 7.19: Photos taken with a CCD camera on a UFSD3-CMS 2×2 array-Safe sensor, at three different values of bias voltage: 200 V, 250 V and 260 V. Hot spots (yellow spots) appear close to the breakdown voltage.

the IV curves of Aggressive sensors irradiated with neutrons to fluences 0.4, 0.8 and $1.5 \cdot 10^{15}$ n_{eq}/cm^2, which do not show any premature BD. The irradiation generates scattering centres that slow down the charge carriers and they inhibit the avalanche multiplication mechanism, that in non-irradiated sensors is caused by high surface electric fields (see section 4.1.1).

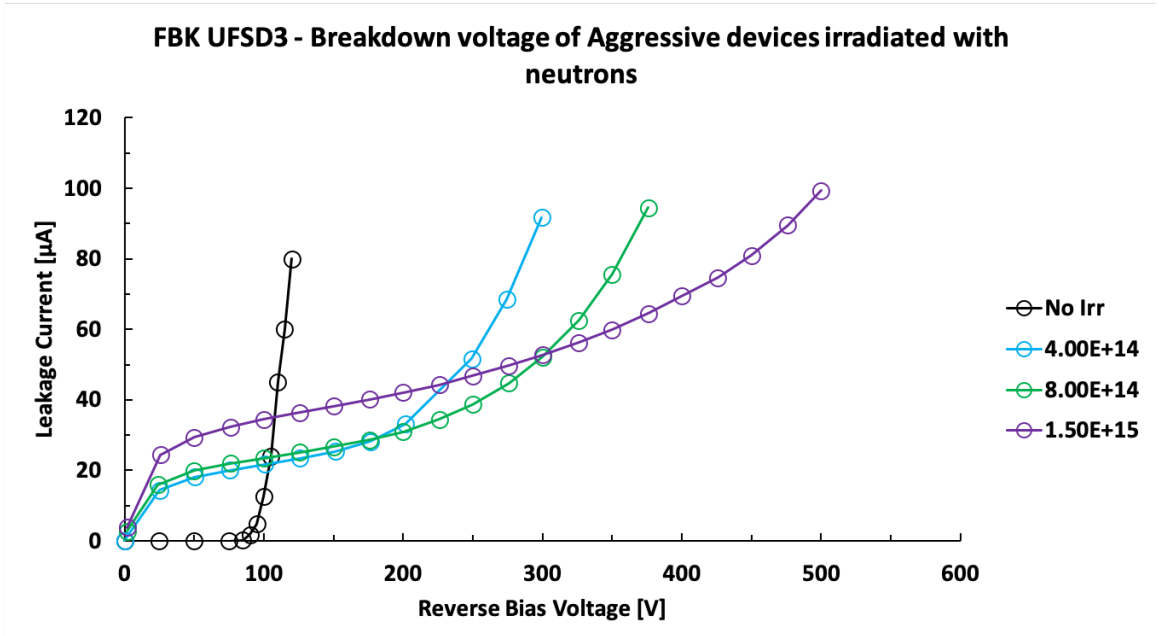

Figure 7.20: Breakdown measurements on FBK-UFSD3 Aggressive sensors irradiated up to a fluence of $1.5 \cdot 10^{15}$ n_{eq}/cm^2; measurement at room temperature.

After having investigated and studied the phenomenon of the earlier breakdown in UFSD3 devices, the same typology of sensors has been used to perform the measurements of the extent of the inactive region between adjacent pads. The measurement methodology is described in section 6.2. Figure 7.21 shows these measurements on the four flavours of termination layouts: the solid lines with the typical S shape are the fits of the charge profiles collected by the two pads or strips, on each side of the optical window. Each fit is obtained with the convolution of a Step and a Gaussian function (dashed lines).

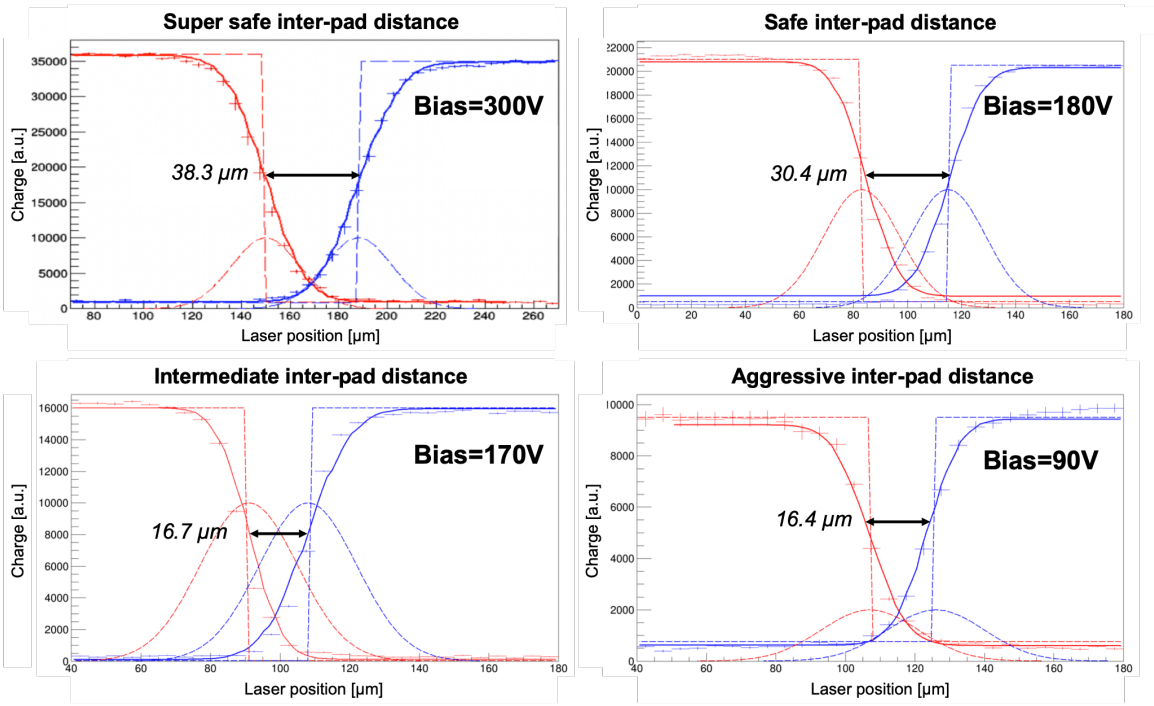

Figure 7.21: Collected charge profile along the optical window between two nearby pads for the four different termination layout designs in UFSD3: Super Safe (top-left), Safe (top-right), Intermediate (bottom-left) and Aggressive (bottom-right).

The first interesting result is that even for the device with the largest inter-pad layout, the width of the inactive region is less than 40 μm wide; this result is very important for the development of large area sensors with a fill factor of at least 95%, as required by the HL-LHC experiments. The results of the inter-pad width region of the UFSD3 layouts are summarized in table 7.2. It can be noticed that these results are in

Inter-pad layout	Measured distance [μm]	Nominal distance [μm]	Laser spot [μm]
Super Safe	38.3	41	13.4
Safe	30.4	31	10
Intermediate	16.7	20.5	10.1
Aggressive	16.4	11	7.7

Table 7.2: Results of the inter-pad (no-gain) region extension performed with Front-TCT setup on FBK-UFSD3 sensors.

good agreement with the nominal distances expected from the sensor layout, except for the aggressive configuration. The inter-pad measurement of the Aggressive sensor has been the only one performed at a bias voltage not high enough to saturate the drift velocity of the charge carriers since this type of device has a breakdown voltage of $\sim 100\ V$. This fact has a not negligible effect on the charge collection lines at the edge of the gain layer, increasing the inactive region. The measurements on the Safe and Intermediate devices have been repeated at four different bias voltages, making sure that the devices were fully depleted and the drift velocity of the charge carriers saturated. Both types of devices show a decrease of the inter-pad distance as the bias voltage increases: the inter-pad width varies of about 1 μm for a bias variation of 30 V and 45 V for the Safe and Intermediate sensors, respectively. This result further supports the hypothesis that the measurement performed on the Aggressive sensor is affected by a bias voltage that is too low.

Pop-Corn noise

An undesirable effect has been observed during the characterization of the UFSD3 devices. Microdischarges appear at certain values of bias and cause an increase in device noise, we call this effect "Pop-corn noise". Pop-corn noise had already been observed in UFSD2 devices, but it did not represent a problem since it appeared just a few volts before BD; it has always been considered as an indication of the beginning of BD. In UFSD3 devices, the Pop-corn noise appears at much lower voltages than the BD, preventing the operation of the devices at an appropriate voltage. An example

of Pop-Corn noise is shown in the oscilloscope screen reported in figure 7.22 (left): the yellow noise baseline of a sensor with Pop-Corn noise shows several spikes, while the pink one is the noise of a sensor without this effect. According to the literature [60], Pop-corn noise is generated by a too sharp p-n junction. In a UFSD device, this junction is located between the p-stop and the bulk: positive charges in the oxide induce a layer of electrons underneath it, which creates an inversion layer that acts as n-doped silicon. This inversion layer creates a p-n junction with the p-stop, as shown in figure 7.22 (right). In the UFSD3 production, all the implants have been created with the stepper technique, instead of the mask aligner used in UFSD2. The stepper creates much sharper implants, with better defined edges. This positive aspect, however, makes the p-n junction sharper and more prone to generate Pop-corn noise. The doping of the p-stop has also an effect: the higher the doping, the sharper the p-n junction.

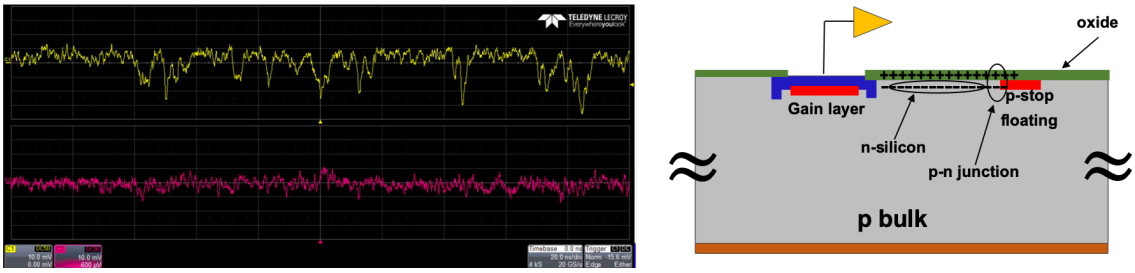

Figure 7.22: Left: comparison between baselines noise with and without Pop-Corn; Pop-Corn (yellow), no Pop-Corn (pink). Right: a sketch of a possible interpretation of the Pop-Corn noise in UFSD sensors.

7.2.2 Inter-pad in HPK UFSDs

Measurements of inter-pad region have been also performed on HPK UFSDs. In the HPK-EXX28995 production (section 5.2.2) four different configurations of inter-pad region are implemented, labelled with the inter-pad factor 95, 70, 50 and 30. The first label represents the safest configuration, while the last one the more aggressive. The devices selected to study the inactive area between pads are 2×2 array (FBK-like);

a front view of the layout of one of these devices is shown in figure 7.23, on the left. Since, the 70, 50, and 30 devices do not have inter-pad optical slits on the front side, all the inter-pad measurements have been performed shooting the IR laser from the back side of the devices, where there is a grid of metal-optical windows, shown in figure 7.23, on the right.

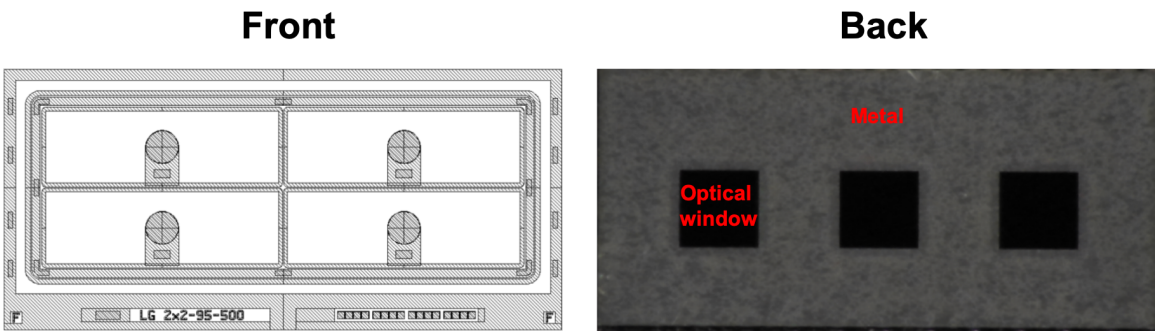

Figure 7.23: Picture of a HPK-UFSD 2×2-array selected to characterize different inter-pad regions on EXX28995 production. On the left, the sensor layout of the front side, while on the right a photo of the back side of the sensor with an optical window grid.

The first measurements performed are IV characterizations to study the BD voltage for the different termination layouts. The BD occurs at very similar voltage values in each of the 4 types of devices, not showing a dependency on the inter-pad layout: 250 V is the BD voltage of the devices 95, 70 and 50, while 220 V it is of the 30. It was decided to perform the measurements on the width of the no-gain region at voltages 30 V lower than the BD one.

Figure 7.24 shows the four measurements of inter-pad width, while the results are summarized in table 7.3. It is possible to notice that the extension of the inactive area is between a maximum value of 130 μm and a minimum of 72 μm. The inter-pad measurement has also been performed on the front side of the array 2×2-95 device which has optical slits, obtaining a result consistent with that one obtained from the back side. This result validates the back side method performed on the other types of devices.

In conclusion, the results show that all the FBK layouts have lower inter-pad regions than the HPK ones, however at the expense of a BD voltage that depends on the termination layout type. The design of FBK-UFSD3 needs to be improved to remove the presence of the Pop-corn noise and increase the BD voltage.

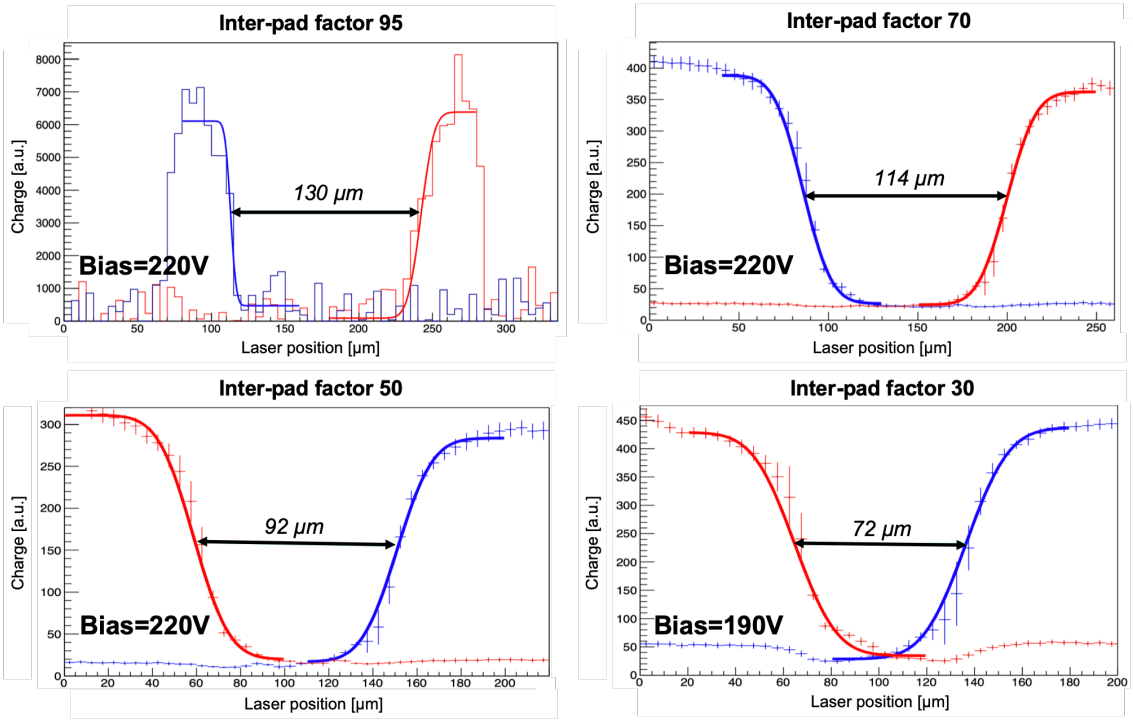

Figure 7.24: Inter-pad measurements on HPK-Type 3.1 sensors for 4 different inactive region layouts: inter-pad factor 95 (top-left), inter-pad factor 70 (top-right), inter-pad factor 50 (bottom-left) and inter-pad factor 30 (bottom-right).

Inter-pad factor	Measured distance [μm]	Nominal distance [μm]	Laser spot [μm]
95	130	unknown	15
70	114	unknown	17
50	92	unknown	16
30	72	unknown	20

Table 7.3: Inter-pad region extension measurements on HPK-Type 3.1-EXX28995

7.2.3 Trench isolated LGADs

An alternative segmentation technology to the standard one (JTE and p-stop), which has the potentiality to reduce the extent of the inter-pad region, is the Trench Isola-

tion (TI) technology. This technology is based on deep and narrow trenches, which physically separate and electrically isolate the multiplication layers of adjacent pads; Trench isolation is extensively used in CMOS image sensors [61] and in silicon photomultipliers [22]. The trenches replace the JTE and p-stop implants, they are dug with deep reactive ion etching technique and filled with silicon oxide; they have a width less than 1 μm. Figure 7.25 (left) shows a sketch of a cross section of a TI-LGAD.

The first LGAD devices with a trench isolation (TI-LGADs) have been produced by FBK in an own internal production [38]. These devices have been fabricated on epitaxial substrate $\sim$ 55 μm thick; to facilitate the electrical characterization, sensors with two pads (250 $\mu m \times$ 375 μm) have been fabricated. Figure 7.25 (right) shows a photo of a device, where the two pads and guardrings that surround them can be identified; moreover, an optical slit (15 $\mu m \times$ 200 μm) between the two pads allows shooting the laser in the inter-pad region to perform inactive area measurement.

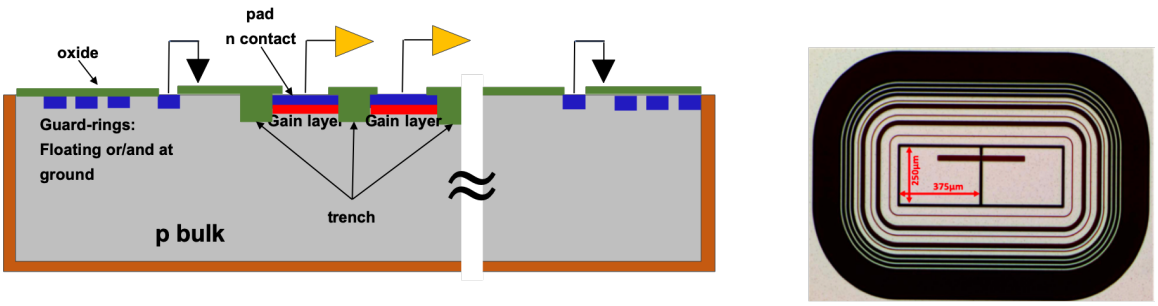

Figure 7.25: Left: cross section, not to scale, of a trench isolated LGAD. Right: picture of a TI-LGAD with two pads surrounded by guardrings.

Two different trench designs have been implemented: the first one with a trench grid between pads (version T1); the second one where each pad is surrounded by a trench ring (version T2). The nominal distance between the gain layer in T1 and T2 designs is $\sim$ 4 μm and $\sim$ 6 μm, respectively. IV characterization on both designs shows a breakdown due to gain and a very good pads insolation. TCT measurements of the width of inter-pad region have been performed: figure 7.26 shows the compari-

son between the inter-pad of a TI-LGAD T2 (red) and the inter-pad of a FBK-UFSD3 Intermediate layout (green). The inter-pad profile of the TI-LGAD proves excellent electrical isolation of the two pads; the collected charge from the read-out pad is zero when the laser shots on the nearby one. The inter-pad distance measured is $\sim 7\ \mu m$. In conclusion, the TI-LGAD is a very promising technology, it reduces the inter-pad width by a factor of five compared to the Safe design in UFSD3 and by a factor of two compared to the Intermediate one. The successful production of TI-LGAD prototypes paves the way to the production of UFSD pixel sensors with pitch down to 50 μm.

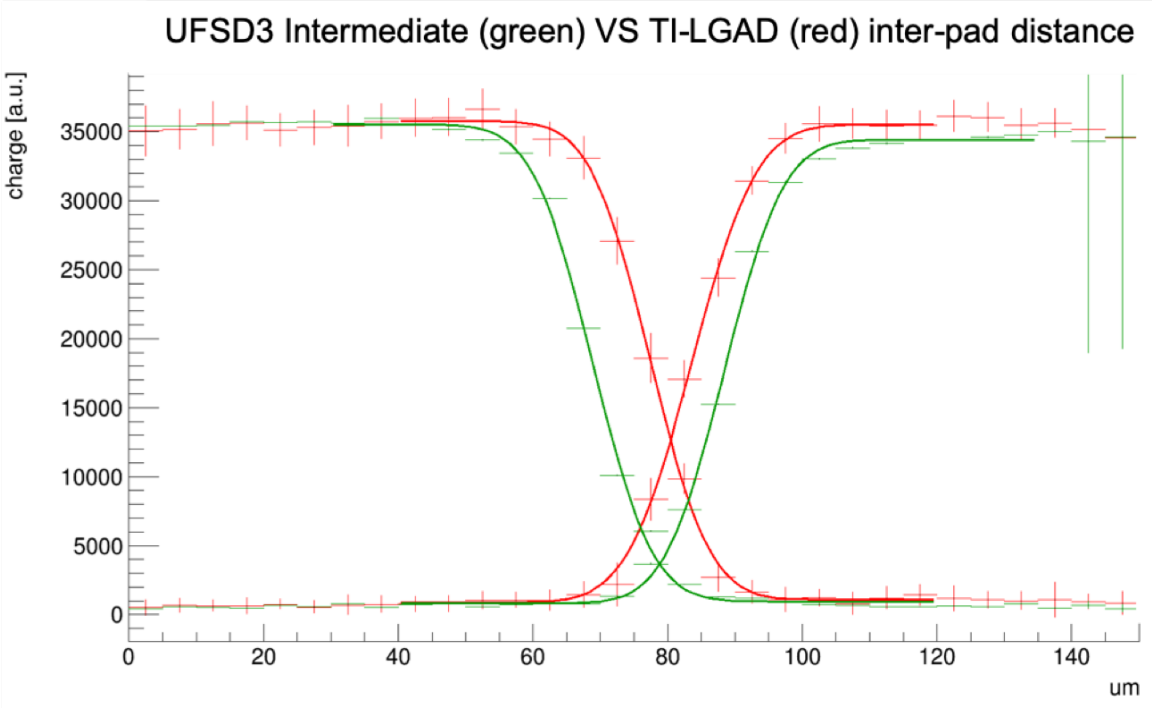

Figure 7.26: Inter-pad measurement of a TI-LGAD T2 (red) compared with the inter-pad of a UFDS3 Intermediate device (green).

7.3 Radiation hardness

This section focuses on the study of radiation damage produced by neutrons and protons in LGAD sensors with an active thickness of about 50 μm, and with different

types of gain layers.

Measurements of leakage current increase, of acceptor removal in the gain layer, of acceptor creation in the bulk and of the decrease in charge collection efficiency will be shown and discussed below.

7.3.1 Irradiation campaign and handling of irradiated sensors

The UFSD and PiN devices of UFSD2 and UFSD3 productions by FBK and of EXX28995 production by HPK have been irradiated, without bias, with neutrons and protons. The neutron irradiation has been performed at JSI research reactor of TRIGA type in Ljubljana, of which neutron spectrum and flux are very well known [62]. The irradiation fluence is expressed in 1 MeV neutrons equivalent per cm^2 (n_{eq}/cm^2) and for this irradiation campaign, the range of fluences chosen is between $1 \cdot 10^{14}$ and $1 \cdot 10^{16}$ n_{eq}/cm^2. Table 7.4 reports the wafers, their corresponding gain layer types and the steps of irradiation chosen for the neutron irradiation campaign.

The protons irradiations have been performed in three facilities: the IRRAD facility located at the T8 beamline at the CERN PS East Hall, with 24 GeV/c protons; the Cyclotron and Radioisotope Center (CYRIC) at Tohoku University, with 70 MeV/c protons; KIT irradiation facility in Karlsruhe, with 23 MeV/c protons. The purpose of these irradiation campaigns is to study the radiation damage (acceptor removal) as a function of proton energy. The NIEL factor decreases as proton energy increases, as shown in figure 7.27, therefore the expected damage is strongest for low energy protons. The NIEL factors of the three proton irradiation facilities are reported in table 7.5 and also marked with red circles in figure 7.27. Table 7.6 reports all the information about the proton irradiation campaigns; the irradiation fluence is expressed in number of protons per cm^2 (p/cm^2), the corresponding value in n_{eq}/cm^2 can be obtained by multiplying the proton fluence by the NIEL factor.

The sensors selected for neutrons and protons irradiation campaigns are single

pad, couple PiN-LGAD and LGAD-LGAD, 2×2, and 5×5 array devices; the active area of each pad of the devices varies from 1 mm^2 to 3 mm^2.

Foundries	Production	Gain layer type	Wafer	Fluence $[n_{eq}/cm^2]$
FBK	UFSD2	B LD	1	0.2, 0.4, 0.8, 1.5, 3.0, 6.0, $10 \cdot 10^{15}$
		B HD	3	
			8	
		B HD+C-A	6	
		Ga	14	
			18	
		Ga+C-A	15	
	UFSD3	B LD	1	0.4, 0.8, 1.5, $3.0 \cdot 10^{15}$
			2	
		B LD+C-A	4	
			5	
		B LD+C-B	7	
		B LD+C-C	9	
		B LD+C-D	11	
		B HD	12	
			13	
		B HD+C-A	14	
			15	
		B HD+C-B	18	
		B HD+C-C	20	
HPK	EXX28995	Type 3.1	8	0.1, 0.4, 0.8, 1.5, $3.0 \cdot 10^{15}$
		Type 3.2	18	

Table 7.4: Summary of wafers and fluences used in the neutron irradiation campaign performed at JSI research reactor of TRIGA type in Ljubljana.

Irradiation facility	Proton energy $[MeV/c]$	NIEL factor
KIT	23	2
CYRIC	70	1.47
IRRAD	$24 \cdot 10^3$	0.67

Table 7.5: Summary of proton irradiation facilities with the respective proton energy and NIEL factor.

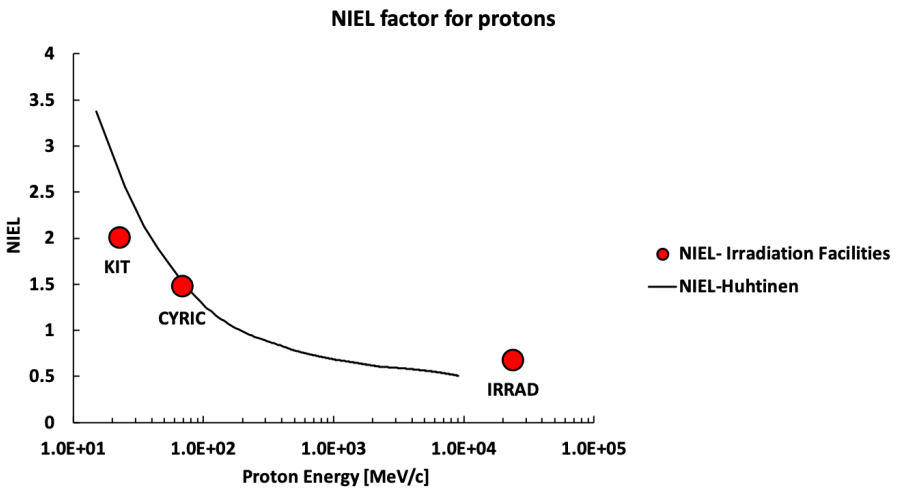

Figure 7.27: NIEL factor tabulated in literature as a function of protons energy (black line) and NIEL factors of the three protons irradiation facilities (red circles).

Foundries	Production	Gain layer type	Wafer	Proton energy $[MeV/c]$	Fluence $[p/cm^2]$
FBK	UFSD2	B LD	1	23	$0.4, 0.8 \cdot 10^{15}$
		B HD	3	$24 \cdot 10^3$	$0.167, 1.0, 1.67, 5.0 \cdot 10^{15}$
			8	23	$0.4, 0.8 \cdot 10^{15}$
		B HD+C-A	6		
				$24 \cdot 10^3$	$0.167, 1.0, 1.67, 5.0 \cdot 10^{15}$
		Ga Ga+C-A	14 15		
	UFSD3	B LD	1	70	$0.34, 0.68 \cdot 10^{15}$
		B LD+C-A	4	23	$0.34, 0.8, 1.6 \cdot 10^{15}$
			5		
HPK	EXX28995	Type 3.1	8	23	$0.34, 0.8, 1.6 \cdot 10^{15}$

Table 7.6: Summary of wafers, fluences and proton energy used in proton irradiation campaign.

Once the irradiation has been completed and the sensors are considered no-more radioactive, before starting the characterization tests, it is necessary to perform the annealing procedure (section 3.3). The annealing recipe for irradiated UFSDs consists to heat them 80 min at 60 $°C$, this temperature is above the recrystallization one of the silicon, while it is below the melting one. Afterward, the devices are constantly

kept in a cold box at $-20\ ^{\circ}C$. All irradiated sensors characterized in this Ph.D. thesis have undergone this annealing treatment.

7.3.2 Leakage current in irradiated 50 μm thick PiN diodes

Once the annealing procedure is complete, it is possible to start the sensor characterization. The first measurement performed is the IV on PiN diode (no-gain), to measure the bulk leakage current increment due to the irradiation. Figure 7.28 shows the IV curves measured on FBK-UFSD2 W1 devices of active area 1 mm^2, irradiated with neutrons in a fluences range from $2 \cdot 10^{14}$ to $1 \cdot 10^{16}$ n_{eq}/cm^2. The radiation increases the bulk leakage current by several orders of magnitude compared to that of an unirradiated sensor (black IV).

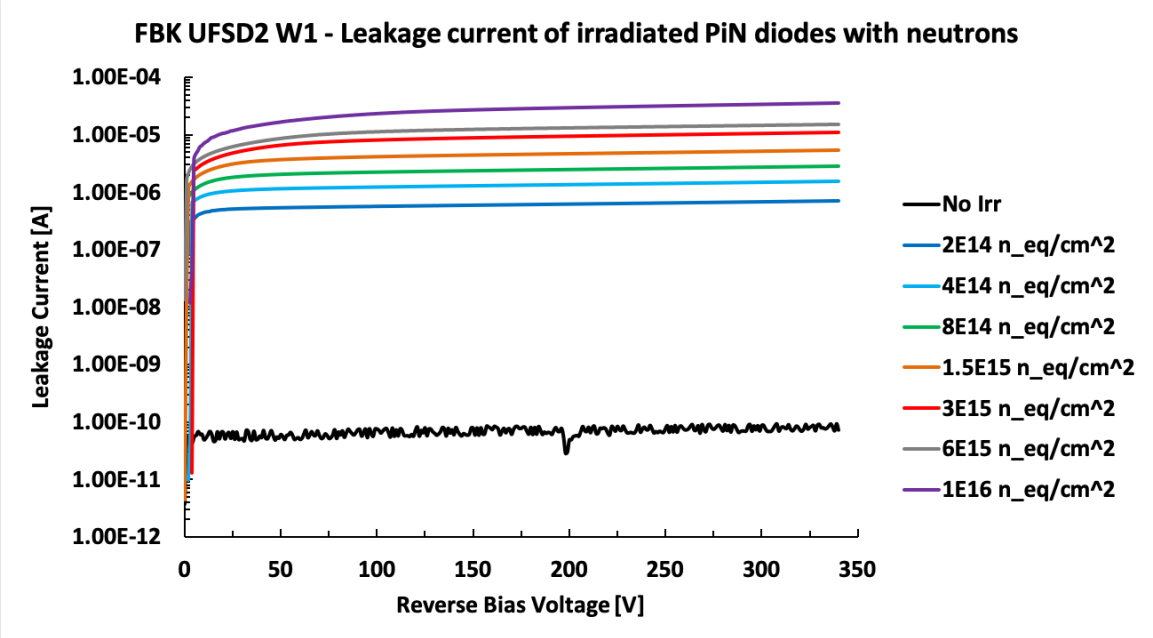

Figure 7.28: Bulk leakage currents measured at room temperature on FBK-UFSD2-W1 PiN diodes, of active area 1 mm^2 and active thickness of $\sim$ 50 μm, irradiated with neutrons.

Figure 7.29 shows the increase of the leakage current measured at the external bias voltage of 340 V, as a function of the irradiation fluences. Since the bulk current increases linearly with the fluence, as discussed in section 3.3, from the linear fit of

the data it is possible to estimate the current-related damage rate (α) as the ratio between the angular coefficient of the linear fit and the depleted volume of the devices (see equation 3.13).

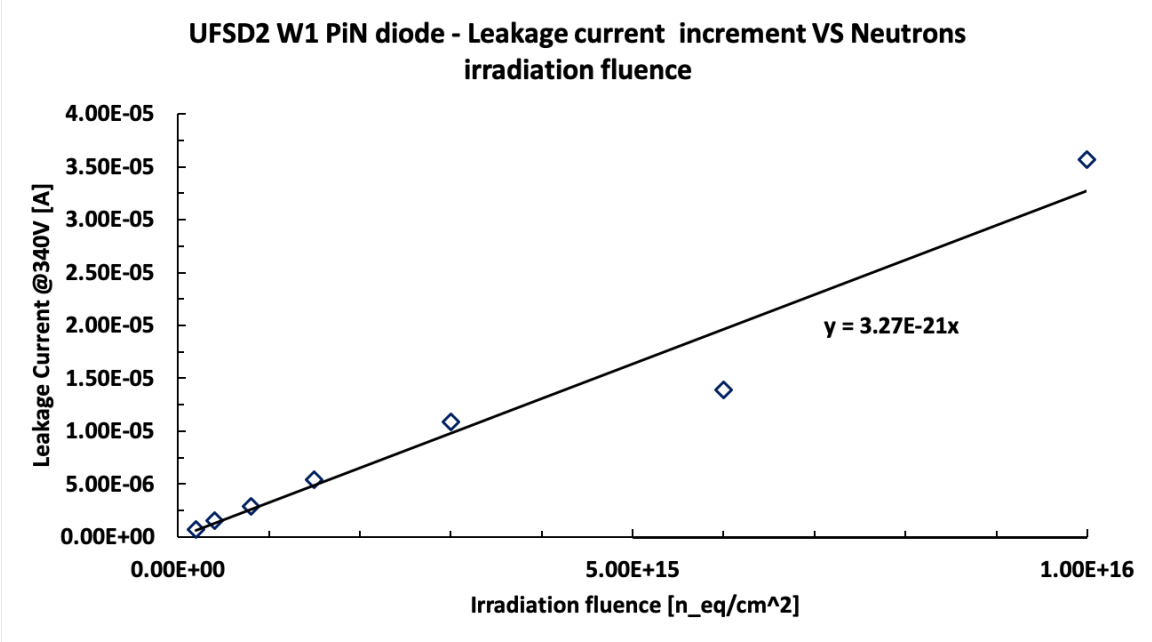

Figure 7.29: Linear increase of the bulk leakage current as a function of the irradiation fluence for FBK-UFSD2-W1 PiN diodes, of active area 1 mm^2 and active thickness of ~ 50 μm, irradiated with neutrons.

The α factor has been extracted from five FBK-UFSD2 wafers, the values obtained are reported in table 7.7. The depleted volume of the sensors, used to extract these five values, is $5 \cdot 10^{-5}$ cm^3; this value is given by the product of the active area of the sensors with the bulk active thickness of ~ 50 μm. The average value $\alpha_{mean} = (6.3 \pm 0.4) \cdot 10^{-17}$ A/cm has been obtained from values in table 7.7. The value of α_{mean}

UFSD2 wafer	α [10^{-17} A/cm]
1	6.6
6	6.9
8	6.0
14	6.1
15	5.9

Table 7.7: Current-related damage measured on irradiated FBK-UFSD2 PiN diodes with a float zone bulk of thickness ~ 50 μm.

is $\sim 30\%$ higher than the value of $(3.99 \pm 0.03) \cdot 10^{-17}$ A/cm reported in the literature (see [8], chapter 5). The overestimation of α_{mean} could be due to edge effects not considered in the calculation of the depleted volume of the sensors.

7.3.3 Acceptor removal rate in different gain layer designs

The acceptor removal mechanism (see sections 3.3 and 4.8.2) has been studied on different gain layer designs, aiming at identifying the most radiation resistant. The experimental technique used to quantify the acceptor removal rate is the CV measurement. As mentioned in section 6.1.2, from the CV curve it is possible to extract the depletion voltage of the gain layer, which is proportional to the amount of the active doping into the gain layer, equation 6.3. The radiation reduces the active acceptors concentration in the gain layer, therefore reducing its depletion voltage. Figure 7.30 shows the evolution of the CV curve in FBK-UFSD2-W8 (B HD) single pad sensors, of active area 1 mm^2, irradiated with neutrons.

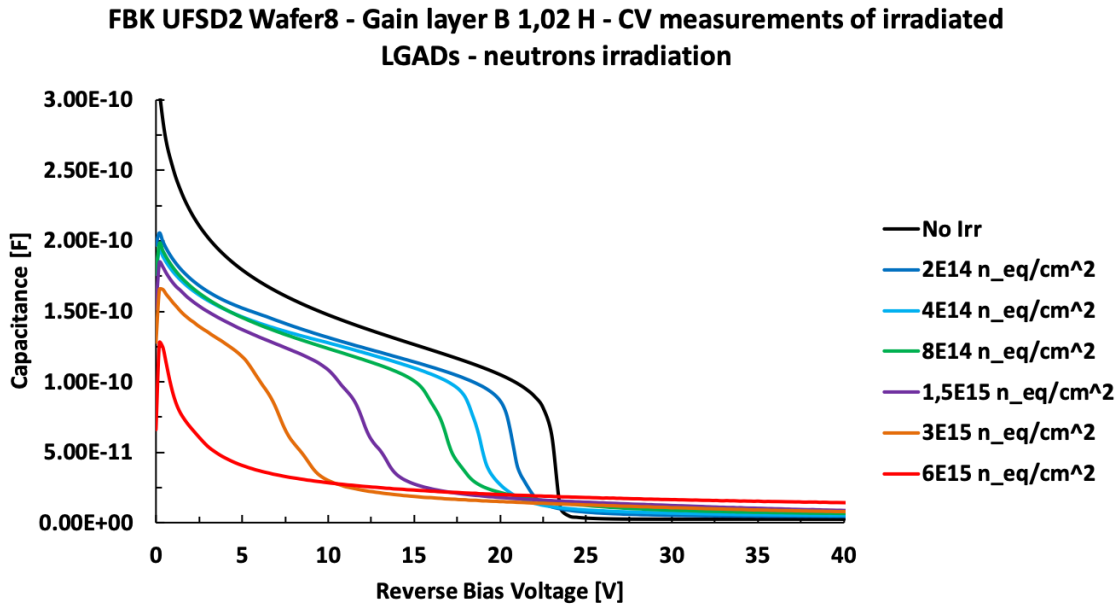

Figure 7.30: Evolution of CV curves for UFSD FBK-UFSD2-W8 sensors, of active area 1 mm^2, irradiated with neutrons. Irradiation fluence starts at $2 \cdot 10^{14}$ n_{eq}/cm^2 and doubles at each step up to $6 \cdot 10^{15}$ n_{eq}/cm^2. Measurements performed at room temperature and 1 kHz frequency.

This plot clearly shows that V_{GL} decreases as the radiation fluence increases, in a fluence range between 0 (black curve) to $6 \cdot 10^{15}$ n_{eq}/cm^2 (red curve); For the sensor irradiated to $6 \cdot 10^{15}$ n_{eq}/cm^2 the typical knee in the CV curve is completely disappeared, proving the complete deactivation of the gain layer.

The acceptor removal rate $c(N_A(0))$ (see equation 3.14) can be obtained from the ratio between the values of the depletion voltage of the gain layer of irradiated ($V_{GL}(\phi)$) and not irradiated ($V_{GL}(0)$) sensors; this ratio is equivalent to the ratio of their active acceptors densities

$$\frac{V_{GL}(\phi)}{V_{GL}(0)} = \frac{N_A(\phi)}{N_A(0)} = e^{-c(N_A(0))\phi}. \tag{7.5}$$

Figure 7.31 shows this ratio (fraction of still active gain layer) as a function of the irradiation fluence, for the wafer 8 of FBK-UFSD2 production. Fitting these data with an exponential function it is possible to obtain the c coefficient for this specific type of gain layer. The smaller the value of c, the higher the gain layer radiation hardness (acceptor removal curve less steep).

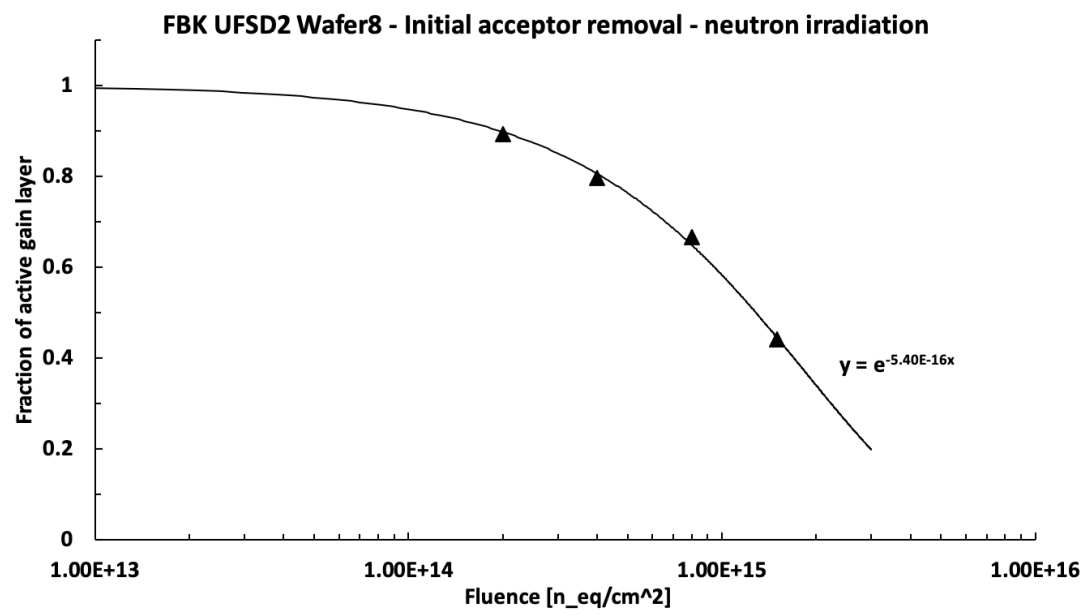

Figure 7.31: Fraction of still active gain layer as a function of the neutron irradiation fluence, for FBK-UFSD2 wafer 8 (B HD).

Modelling the sensor with the equivalent circuit $R_P - C_P$, there are two methods to extract the value of the depletion voltage of the gain layer: the first one is based on the $1/C^2(V)$ curve, the second one using the $R_P(V)$ curve. The $1/C^2(V)$-method exploits the sharp change in slope of this curve corresponding to the transition from the gain layer to the bulk region; the intercept of the linear fits of the $1/C^2(V)$ curve before and after its change in slope gives the depletion voltage of the gain layer called V_{GL}^C. The $R_P(V)$-method exploits the presence of a cusp in correspondence of the depletion voltage of the gain layer, allowing easy identification of V_{GL}^R: the sharp decrease of the R_P value after V_{GL}^R is due to the transition of the depletion volume from the gain layer to the bulk; the depletion volume initially increases both laterally and deeply generating an undershoot in the $R_P(V)$ curve. Figure 7.32 (top) shows the $1/C^2(V)$ and $R_P(V)$ curves for a FBK-UFSD2 W1 sensor irradiated to $2 \cdot 10^{14}$ n_{eq}/cm^2, with neutrons; it is evident the good correspondence between V_{GL}^C and V_{GL}^R obtained with the two extraction methods described above. Figure 7.32 (bottom), instead, validates the use of R_P-method to identify V_{GL}: from a comparison between the R_P trend with the gain layer profile extracted from CV measurement, there is a very good agreement between the R_P cusp and the depth where the concentration of gain layer implant approximates the doping value of the bulk.

The extraction of V_{GL}^C becomes more and more difficult as the irradiation fluence increases: acceptor removal in the gain layer and acceptor creation effect in the bulk make their p-doping concentrations closer to each other, making the change of slope in the $1/C^2(V)$ curve less and less evident; on the other hand, the cusp in $R_P(V)$ curve is also evident to any fluence, as shown in figure 7.33. The comparison between V_{GL}^C and V_{GL}^R values, in irradiated sensors with neutrons up to fluence of $3{\cdot}10^{15}$ n_{eq}/cm^2, shows differences between them of about 1-2 V; these differences affect the acceptor removal coefficient measurements with differences of $\sim 10\%$ considering only irradiation fluences of the order of 10^{14} n_{eq}/cm^2 and with differences of 20%

considering also fluences of the order of 10^{15} n_{eq}/cm^2. In order to standardize as much as possible the acceptor removal measurements, all acceptor removal rates that will be shown below, have been measured using the R_P-method.

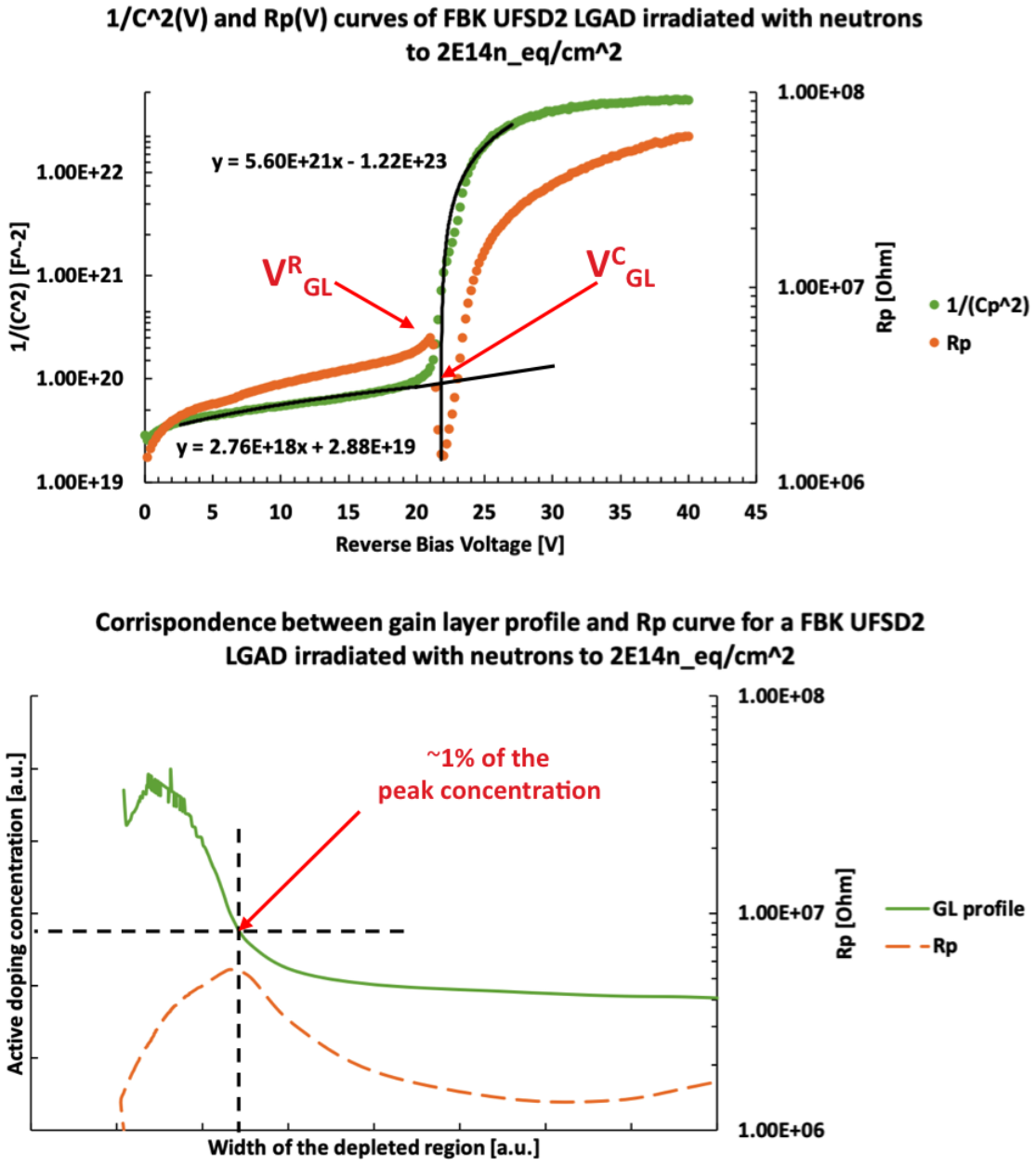

Figure 7.32: Left: correspondence between the depletion voltage of the gain layer V_{GL}^R and V_{GL}^C extracted from R_P and C_P measurements respectively, for a FBK-UFSD2-W1 sensor irradiated with neutrons to $2 \cdot 10^{14}$ n_{eq}/cm^2. Right: correspondence between the gain layer profile and its depletion point from R_P curve.

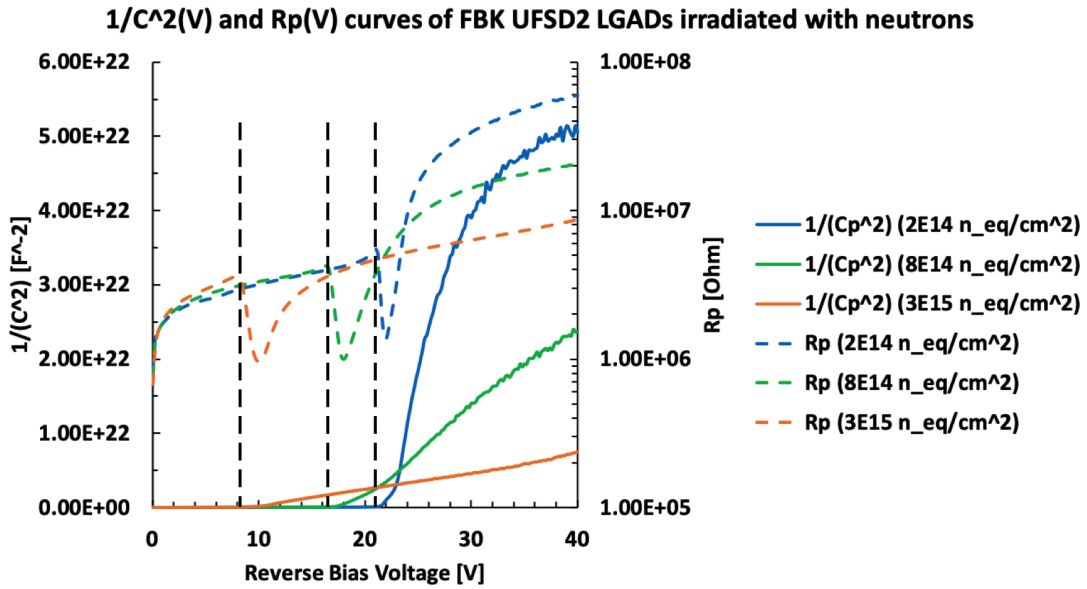

Figure 7.33: $R_P(V)$ and $1/C_P^2(V)$ curves of FBK-UFSD2-W1 sensors irradiated with neutrons to fluences 0.2, 0.8 and $3 \cdot 10^{15}$ n_{eq}/cm^2. Measurements at room temperature.

Figure 7.34 shows acceptor removal curves on twenty different types of gain layers, irradiated with neutron up to a fluence of $1.5 \cdot 10^{15}$ n_{eq}/cm^2, acceptor removal coefficients are reported in table 7.4. From exponential fits to the data it can be seen that the gain layers enriched with carbon (red and green data) are more radiation hard than not enriched ones (blue data). It is interesting to note how the radiation resistance does not improve by increasing the carbon dose: sensors enriched with carbon doses B,C and D (green data) have acceptor removal rates faster than those enriched with dose A (red data).

From data, the FBK-B LD+C-A gain layer design is the most radiation hard; to a fluence of $1.5 \cdot 10^{15}$ n_{eq}/cm^2 still has the 80% of active gain layer. The carbon enrichment (dose A) doubles the radiation resistance of a gain layer. Figure 7.35 shows a comparison between two gain layers with same initial acceptor density, but that differ in the presence of carbon: the FBK-UFSD2 W1 without carbon keeps the 80% of active gain layer to a fluence of $\sim 7 \cdot 10^{14}$ n_{eq}/cm^2, this fluence doubles

($\sim 1.5 \cdot 10^{15}$) for the FBK-UFSD2 W5 enriched with carbon dose A.

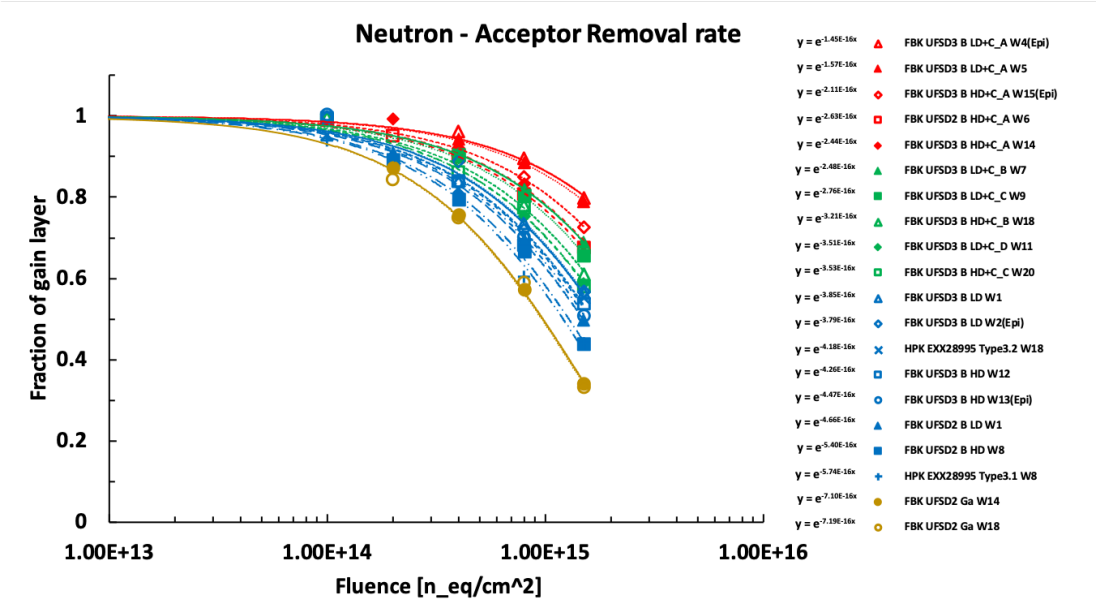

Figure 7.34: Fraction of still active gain layer as a function of neutron fluence for twenty different types of gain layers. Gain layers enriched with carbon (red and green), and without carbon with boron (blue) and gallium (gold) dopants. The lines are the acceptor removal fits; the equations of these fits are reported next to the legend of gain layer types.

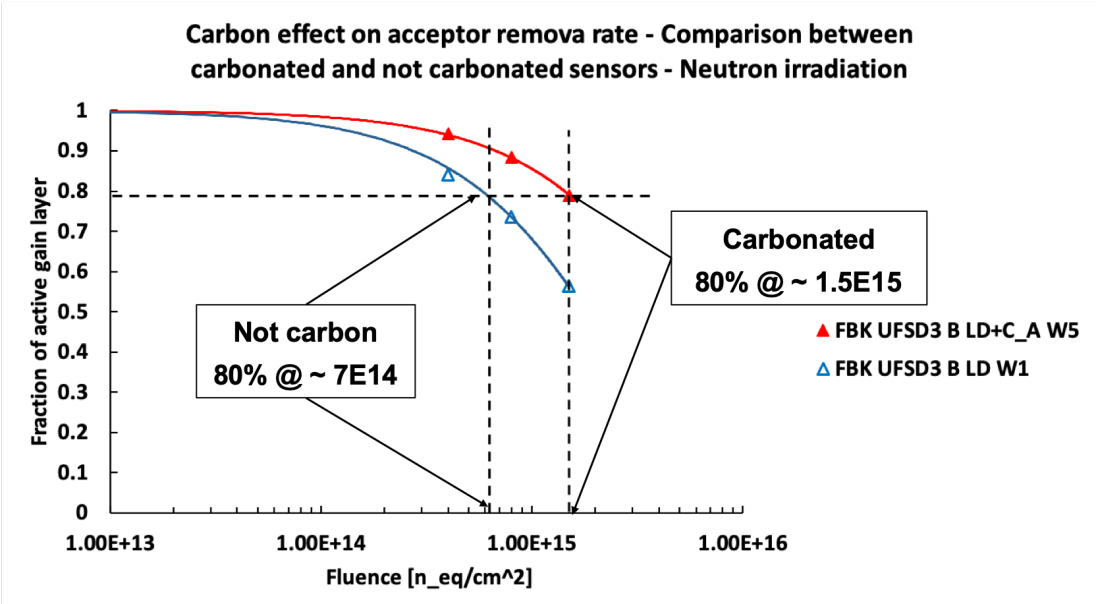

Figure 7.35: Comparison between acceptor removal curves of gain layers B LD+C-A (FBK-UFSD3-W5) and B LD (FBK-UFSD3-W1).

Finally, a last interesting conclusion from figure 7.34, the gain layers with gallium dopant (gold data) are no more radiation resistant than those with boron dopant (blue data).

To complete the acceptor removal studies, sensors irradiated with protons at three different energies, 23, 70 and $24 \cdot 10^3$ MeV/c have been characterized (see table 7.6). Figure 7.36 reports the acceptor removal measurements on two gain layer categories, on top those without carbon, while on bottom those carbonated. These two plots show that less energetic 23 MeV/c protons (grey data) cause faster acceptor removal than the more energetic ones (orange and violet data). This result suggests a relationship between the acceptor removal mechanism and the energy of the irradiation particles, in this case, protons. Figure 7.37 shows more explicitly this relationship: the acceptor removal coefficients for two type of gain layers, B HD+C-A (green) and B LD+C-A (grey), have been plotted as a function of the proton energy; instead, the two dashed lines indicate the respective values of c coefficients for neutrons. The acceptor removal rate decreases for high energy protons; for energies above $\sim GeV/c$ is approximately the same one caused by 1 MeV neutron equivalent. The dependence of the acceptor removal rate from protons energy is probably due to the stronger electromagnetic interaction that the less energetic protons have with the atoms of the silicon lattice, compared to more energetic ones.

The acceptor removal coefficients of all types gain layers are summarized in table 7.8, where c_n indicates those for neutron, while c_p those for proton irradiations. Each coefficient has been obtained by averaging the active gain layer fraction measurements over a population of two or more sensors; considering the error on the average and the uncertainty of our measuring setup, it is possible to attribute approximately 10% of uncertainty to each coefficient.

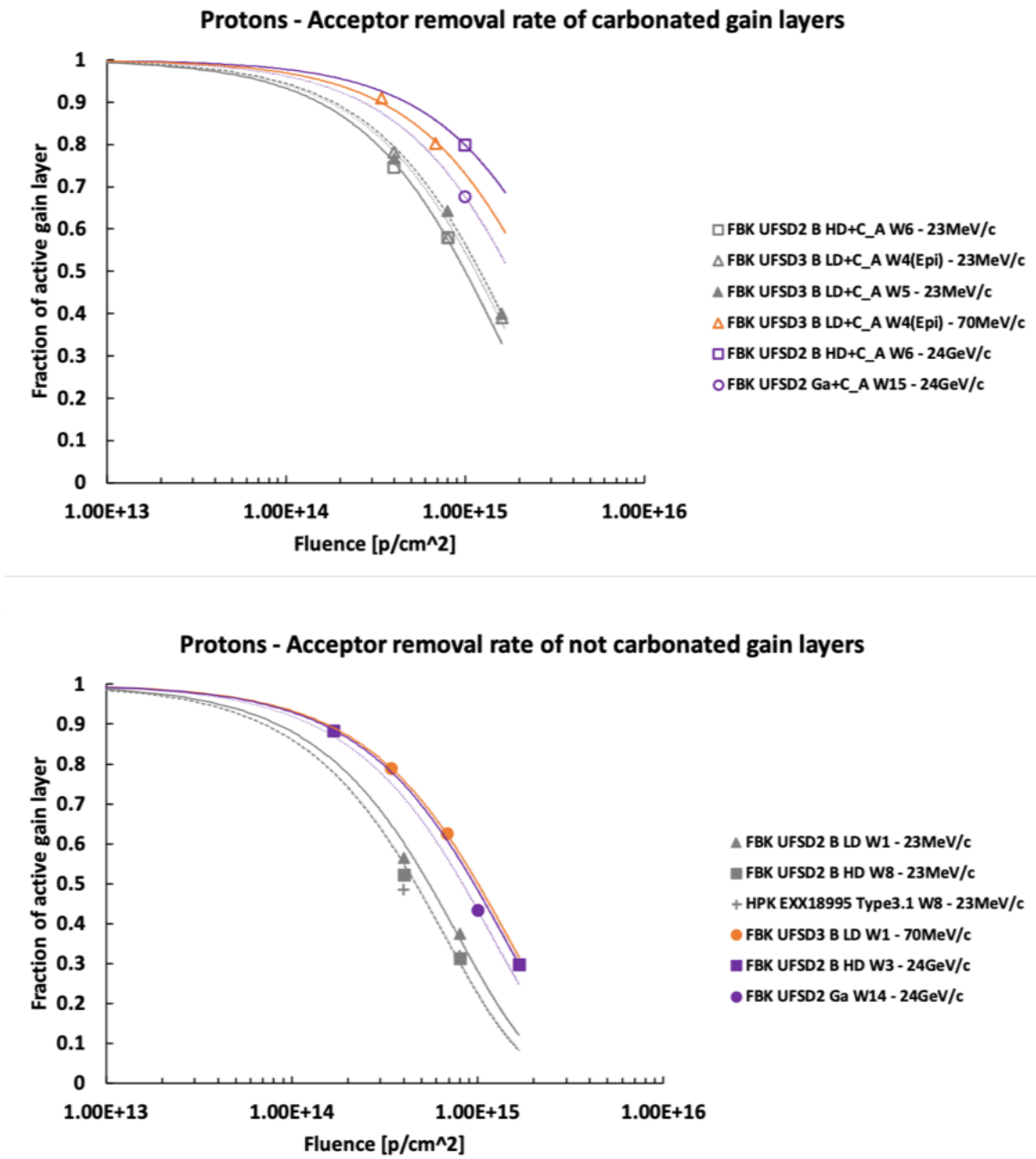

Figure 7.36: Fraction of still active gain layer as a function of the proton fluence for the carbonated gain layer (top), and not carbonated (bottom). Irradiation with 24 GeV/c protons in violet, 70 MeV/c protons in orange and 23 MeV/c protons in grey.

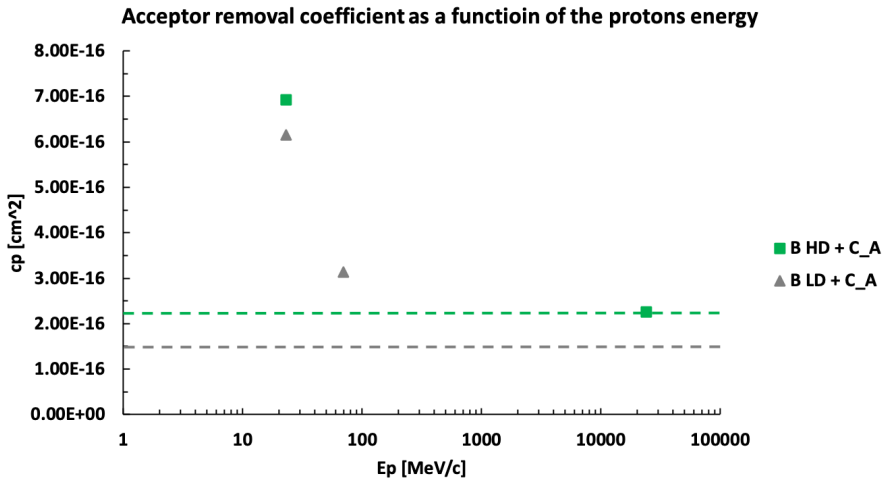

Figure 7.37: Acceptor removal coefficients as a function of protons energy for B HD+C-A gain layer (FBK-UFSD2-W6) in green, and B LD+C-A gain layer (FBK-UFSD3-W5) in grey.

Gain layer Type	Wafer	c_n $[10^{-16}\ cm^2]$	c_p $[10^{-16}\ cm^2]$		
			$p_{23\ MeV/c}$	$p_{70\ MeV/c}$	$p_{24\ GeV/c}$
B LD	$W1_{FBK-UFSD2}$	4.70	12.69		
	$W1_{FBK-UFSD3}$	3.85		6.91	
	$W2_{FBK-UFSD3}$	3.79			
B LD+C-A	$W4_{FBK-UFSD3}$	1.45	6.15	3.14	
	$W5_{FBK-UFSD3}$	1.57	6.24		
B LD+C-B	$W7_{FBK-UFSD3}$	2.48			
B LD+C-C	$W9_{FBK-UFSD3}$	2.76			
B LD+C-D	$W11_{FBK-UFSD3}$	3.51			
B HD	$W8_{FBK-UFSD2}$	5.40	14.87		
	$W12_{FBK-UFSD3}$	4.26			
	$W13_{FBK-UFSD3}$	4.47			
B HD+C-A	$W6_{FBK-UFSD2}$	2.35	6.93		2.25
	$W14_{FBK-UFSD3}$	3.21			
	$W15_{FBK-UFSD3}$	3.53			
B HD+C-B	$W18_{FBK-UFSD3}$	2.48			
B HD+C-C	$W20_{FBK-UFSD3}$	2.76			
Ga	$W14_{FBK-UFSD2}$	7.10			8.35
	$W18_{FBK-UFSD2}$	7.19			
Ga+C-A	$W15_{FBK-UFSD2}$	3.76			3.91
Type 3.1	$W8_{HPK-EXX28995}$	5.74	15.12		
Type 3.2	$W18_{HPK-EXX28995}$	4.18			

Table 7.8: Compilation of the acceptor removal coefficients for neutron c_n and proton c_p irradiation. The error on these coefficients has been estimated to be 10%.

The NIEL factor of protons can be extracted from the proton acceptor removal coefficients. Applying the NIEL to the acceptor removal curve means multiplying the axis of fluence by the NIEL value. For example, figure 7.38 shows two acceptor removal curves of the same type of gain layer (B HD+C-A) with and without the NIEL factor applied. Considering the NIEL factor, the acceptor removal equation for protons can be written as

$$N_A(\phi_p) = N_A(0)e^{-c_p\phi_p} = N_A(0)e^{-c_{n_{eq}}\cdot NIEL\cdot\phi_p} = N_A(0)e^{-c_{neq}\phi_{neq}}, \qquad (7.6)$$

where the fluence in neutron equivalent is $\phi_{n_{eq}} = NIEL \cdot \phi_p$. From equation 7.6 it is possible to obtain the NIEL value as the ratio between the c coefficient for protons and neutrons ($NIEL = c_p/c_{n_{eq}}$).

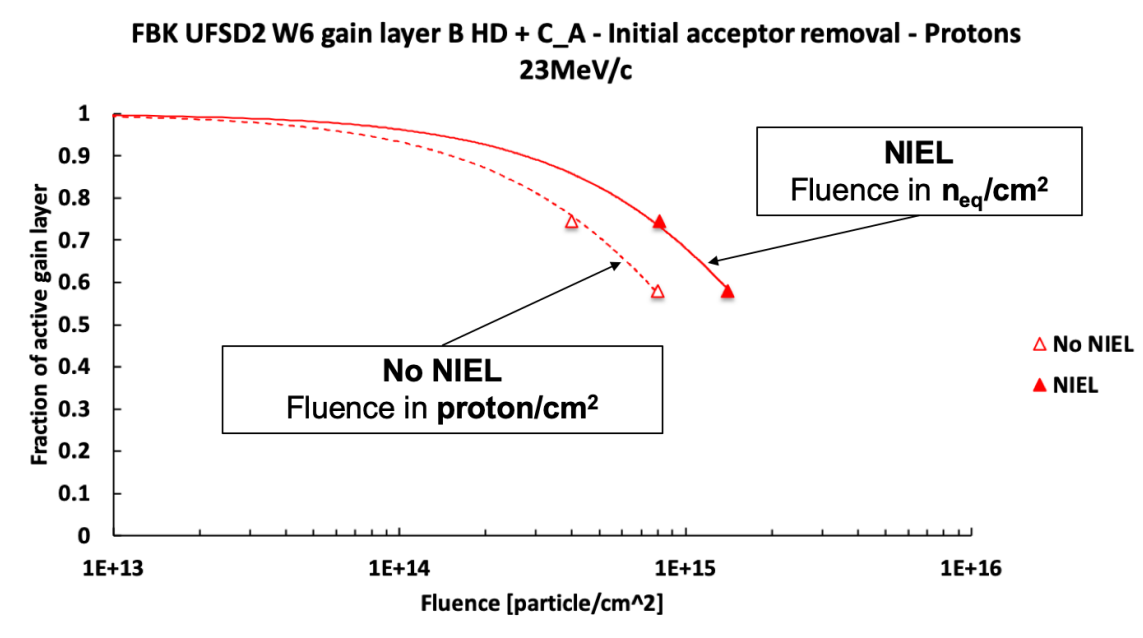

Figure 7.38: Effect of NIEL factor applied to acceptor removal curve of a B HD+C-A gain layer (FBK-UFSD2-W6).

Table 7.9 reports the NIEL values obtained from acceptor removal rates for the protons energy 23, 70 and $24 \cdot 10^3$ MeV/c. In figure 7.39, these NIEL values have been compared with those tabulated in literature [42] and with those of the irradiation

facilities (table 7.5). NIEL measured from acceptor removal is higher than NIEL tabulated in literature.

Gain layer Type	Wafer	$NIEL = c_p/c_n$		
		$p_{23\ MeV/c}$	$p_{70\ MeV/c}$	$p_{24\ GeV/c}$
B LD	$W1_{FBK-UFSD2}$	2.7		
	$W1_{FBK-UFSD3}$		1.79	
B LD+C-A	$W4_{FBK-UFSD3}$	4.24	2.17	
	$W5_{FBK-UFSD3}$		3.98	
B HD	$W8_{FBK-UFSD2}$	2.75		
B HD+C-A	$W6_{FBK-UFSD2}$	2.95		0.96
Ga	$W14_{FBK-UFSD2}$			1.18
Ga+C-A	$W15_{FBK-UFSD2}$			1.04
Type 3.1	$W8_{HPK-EXX28995}$	2.63		

Table 7.9: Compilation of the ratio c_p/c_n for three different protons energy 23, 70 and $24 \cdot 10^3$ MeV/c.

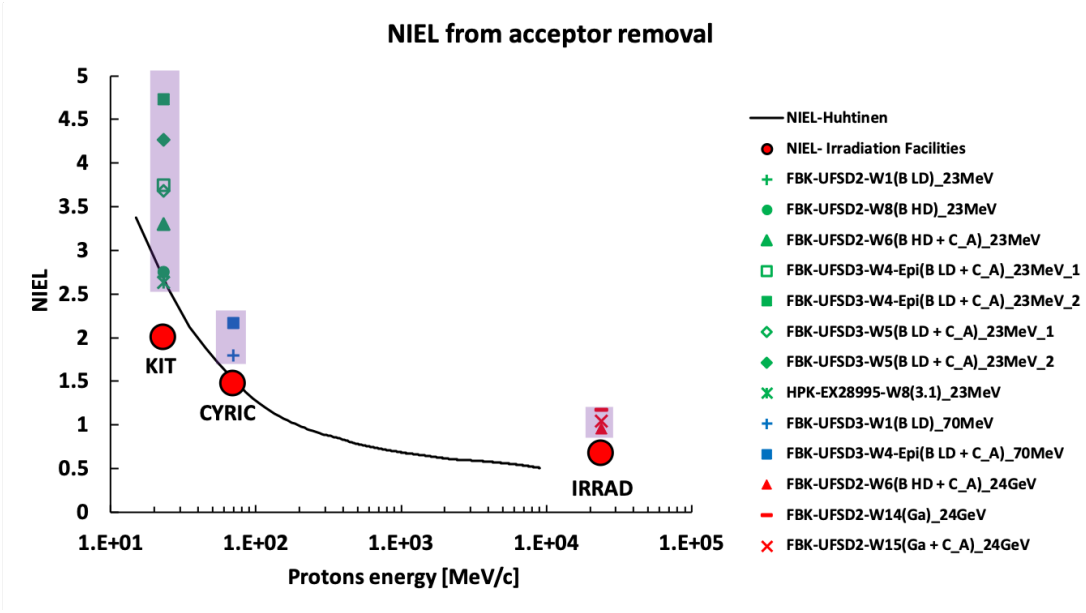

Figure 7.39: NIEL obtained from acceptor removal rate measurements (violet bands), compared with NIEL values tabulated in the literature (black lines), and with NIEL values of protons irradiation facilities (red circles), table 7.5.

Parametrization of acceptor removal mechanism

Experimental measurements of acceptor removal rate show a dependence from initial acceptor density, as mentioned in section 3.3; here a parametrization of this depen-

dency is proposed [41]. The acceptor removal mechanism is described by this equation

$$N_A(\phi) = N_A(0)e^{-c\phi}, \tag{7.7}$$

where the acceptor removal coefficient c can be rewritten as $\phi_0 = 1/c$; ϕ_0 is the fluence needed to reduce the initial doping density $N_A(0)$ to $1/e$ of its initial value. In a simple model of acceptor removal, the density of initial acceptor atoms deactivated by radiation is given by the product of the fluence ϕ_0 times the silicon atomic density ρ_{Si} times the cross section of impinging particle to deactivate an acceptor σ_{Acc}

$$(1 - \frac{1}{e})N_A(0) = \phi_0 \rho_{Si} \sigma_{Acc}, \tag{7.8}$$

where $\rho_{Si} = 5 \cdot 10^{22}\ cm^{-3}$. The term σ_{Acc} can be written as the product of the cross section between radiation and silicon atoms (σ_{Si}) times the number of the interstitial generating in the scattering (N_{Int}) times the probability of capturing an acceptor (k_{cap})

$$\sigma_{Acc} = k_{cap} N_{Int} \sigma_{Si}. \tag{7.9}$$

The probability of acceptor capture is influenced by the presence of impurity as carbon, oxygen, hydrogen, etc.; those might intercept the interstitial atoms before they reach the acceptors.

Equation 7.8 assumes that each interstitial atom is created near an acceptor, however, this might not be the case at low acceptor density. Therefore a proximity function D must be included in equation 7.8. This function describes the probability that an interstitial atom is located in the vicinity of an acceptor; the analytic form of this function is not unique, for example

$$Dn = \frac{1}{1 + (\frac{N_{A_0}}{N_A(0)})^{n/3}}, \tag{7.10}$$

where N_{A_0} is a fit parameter that indicates the acceptor density at which an interstitial state has a probability of 0.5 of being in the vicinity of an acceptor, and n is an exponent that needs to be determined experimentally. Combining equation 7.8 with equation 7.10 we obtain:

$$\frac{1}{c} = \phi_0 = 0.63 \frac{N_A(0)}{\rho_{Si}\sigma_{Acc}} (1 + (\frac{N_{A_0}}{N_A(0)})^{n/3}), \tag{7.11}$$

where σ_{Acc} and N_{A_0} are fit parameters. The analytic functions $D1$, $D2$ and $D3$ have been tried, finding the best agreement between model and experimental data with $n = 2$, $\sigma_{Acc} = 7.5 \cdot 10^{-22}\ cm^2$ and $N_{A_0} = 2.5 \cdot 10^{16}/cm^3$. Using the three proximity function ($D1$, $D2$ and $D3$) and the values of fit parameters, the ϕ_0 in equation 7.11 has been plotted and superposed to the experimental data from the literature in figure 7.40. The experimental points of B-neutrons (boron gain layer irradiated with reactor neutrons) are taken from [40, 63, 64], the B-protons (boron gain layer irradiated with 800 MeV/c protons) from [40, 63, 64], while Ga-neutrons (gallium gain layer irradiated with reactor neutrons) from [65, 66].

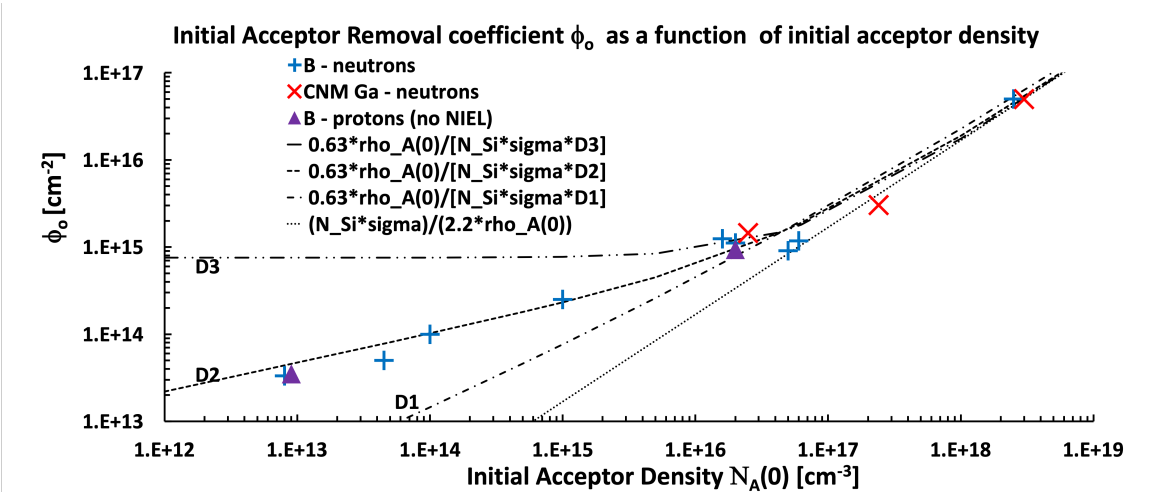

Figure 7.40: Parametrization of ϕ_0 from equation 7.11 together with experimental data from the literature as a function of initial acceptor density. The plot shows the parametrization of equation 7.11 with and without the effect of the proximity function. The best agreement with experimental data is obtained with the $D2$ proximity function. Figure from [41].

The effect of the proximity function is more important at low initial acceptor density, where the overlap probability of interstitial atoms and acceptors is smaller. It is important to note that the acceptor removal rate depends upon the type of irradiation (neutrons, protons, etc.), the irradiation energy and the type of acceptor dopant; however, for lack of experimental, data the parametrization shown in figure 7.40 is a single common fit.

By applying the parametrization in equation 7.11 to the specific case of LGADs, it is possible having an overall overview of the acceptor removal mechanism on different types of gain layers. In figure 7.41 the acceptor removal coefficients c_n, tabulated in table 7.8, have been plotted as a function of the initial acceptor density $N_A(0)$, in a density range from $1 \cdot 10^{16}/cm^3$ to $1 \cdot 10^{17}/cm^3$. The $N_A(0)$ values plotted are the amplitudes of the peak of the gain layer implant profiles obtained from CV measurements; for each type of gain layer, the initial acceptor density is averaged over measurement on several sensors.

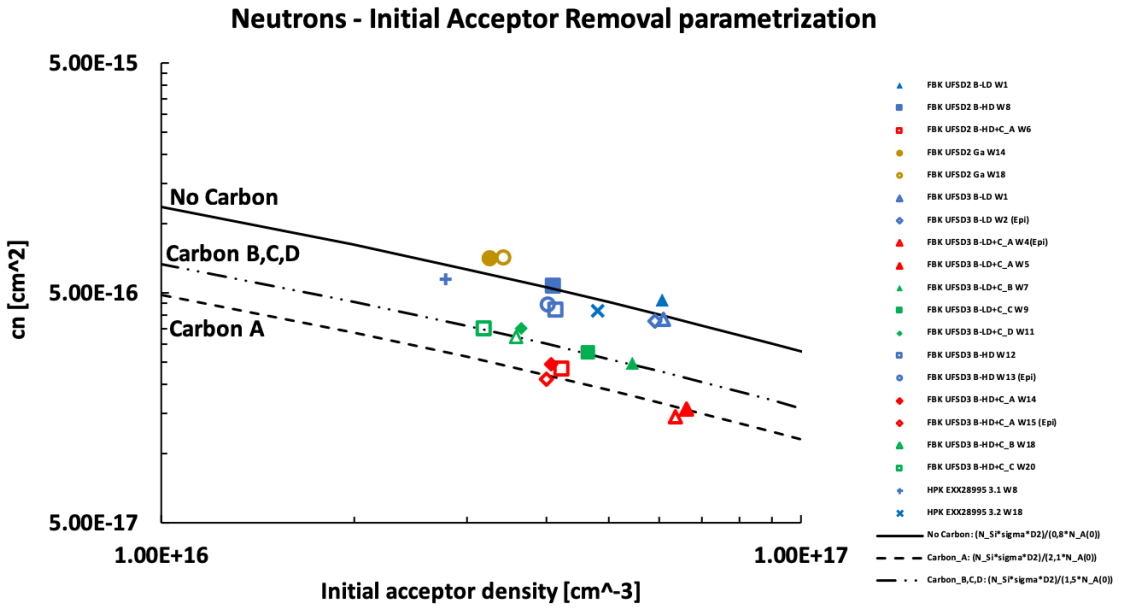

Figure 7.41: Acceptor removal coefficients tabulated in table 7.8 as a function of acceptor removal density. Gain layer with boron dopant in blues, gallium in gold, enriched with carbon dose A in red, and enriched with carbon doses B, C and D in green.

From the data shown in figure 7.41, it is clear that there are three main families of gain layers (blue, red and green markers), which are related to as many initial acceptor removal parametrizations (black dashed lines). Gain layers with c coefficients on the same parametrization have the same radiation resistance, and their acceptor removal rate varies only for a different initial acceptor density; vice versa, belonging to different parametrizations implies a different radiation resistance, probably due to different mechanism of acceptor removal.

Four important consideration on experimental data in figure 7.41: (i) gain layer with boron dopant (FBK-B HD/LD, HPK-Type 3.1/3.2), indicated with blue markers, have the same radiation resistance (same acceptor removal parametrization); (ii) gain layers with gallium dopant (gold markers) have the same radiation resistance than those with boron; (iii) gain layers enriched with carbon dose A (red markers) follows a different acceptor removal parametrization than the not enriched ones; finally, an unexpected result, (iv) gain layers enriched with carbon doses B, C and D (green markers) do not follow the parametrization of those enriched with carbon dose A. Table 7.10 reports the numerical values of the parameters used in the three different parametrizations No Carbon, Carbon-A and CarbonB,C,D plotted in figure 7.41.

Parametrization	$\rho_{Si}[cm^{-3}]$	$\sigma_{Acc}[cm^2]$	$N_{A_0}[cm^{-3}]$	Free parameter
No Carbon	$5 \cdot 10^{22}$	$7.5 \cdot 10^{-22}$	$2.5 \cdot 10^{16}$	0.85
Carbon-A				2.05
Carbon-B,C,D				1.5

Table 7.10: Numerical values of the parameters used in the No Carbon, Carbon-A and CarbonB,C,D parametrizations in figure 7.41.

In conclusion, gain layers with co-implantation of carbon have a higher intrinsic radiation resistance than gain layer without carbon. Increasing the carbon dose the radiation resistance does not improve, but rather worsens compared to the gain layers enriched with carbon dose A; this unexpected effect is still not understood and needs

more studies. Finally, the gallium dopant shows the same intrinsic radiation hardness of boron.

7.3.4 Acceptor creation and charge collection efficiency in PiN diode, 50 μm thick, irradiated with neutrons

Two effects that radiation has on the bulk of the sensors, in addition to the increase of the leakage current, are the acceptor creation and decrease of the CCE.

The acceptor creation (sections 3.3 and 4.8.2) has been studied on FBK-UFSD2 devices without gain, with an active area of 1 mm^2 and an active thickness of $\sim$ 50μm. The acceptor density created with the irradiation can be estimated from the full depletion voltage V_{FD} of the device

$$N_A = \frac{2\epsilon_{Si} V_{FD}}{qd^2}, \tag{7.12}$$

where d is the active thickness of the device. The experimental technique chosen to extract the full depletion voltage of the device is the front-TCT equipped with IR laser. The number of charges created shooting the laser and collected in the depleted region of the sensor increases as the external bias voltage increases, and becomes constant when the depletion voltage is reached. Figure 7.42 (left) shows an example of CC measurement as a function of external bias voltage, in a PiN device irradiated to a fluence of $1.5 \cdot 10^{15}$ n_{eq}/cm^2. From the change of slope in the CC curve, the full depletion voltage of the device is clearly distinguished. The collected charge is proportional to the depleted thickness in the device, which in turn it is proportional to the square root of the external bias voltage (equation 7.12); then the full depletion voltage of the device can be obtained from the intercept on two linear fits on CC($\sqrt{V}$) curve, made before and after its change of slope, as shown in figure 7.42 (right).

Figure 7.43 shows the CC curves, as a function of the bias voltage, measured on

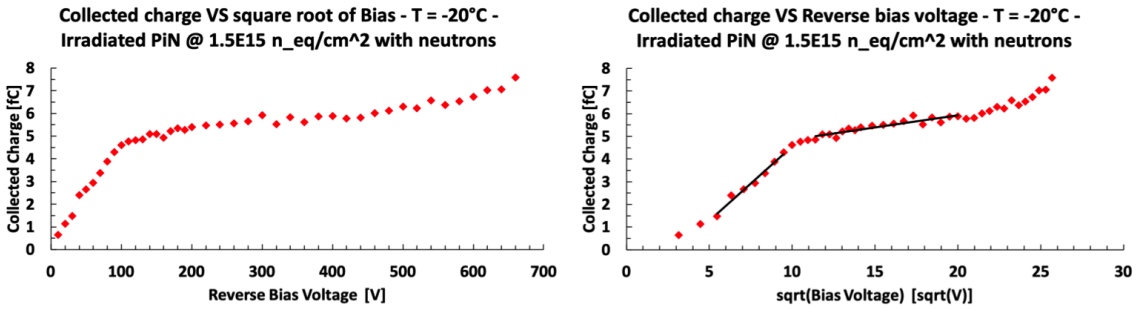

Figure 7.42: Collected charge measurement ($T \sim -20\ ^{\circ}C$) on a FBK-UFSD2 PiN diode irradiated with neutrons to a fluence of $1.5 \cdot 10^{15}\ n_{eq}/cm^2$. On the left, the CC as a function of external bias voltage. On the right, the CC as a function of the square root of the external bias voltage; in this plot, the two black lines represent the linear fits used to extract the full depletion voltage of the device.

not irradiated (black) and irradiated PiN diodes with fluences from $8 \cdot 10^{14}\ n_{eq}/cm^2$ to $1 \cdot 10^{16}\ n_{eq}/cm^2$. These measurements have been performed at a temperature of $-20\ ^{\circ}C$ to keep as low as possible the bulk leakage current. All CC curves have been normalized to the voltage value of 400 V. These measurements have been performed with two different laser intensities, in order to exclude any dependence on the number of generated charges.

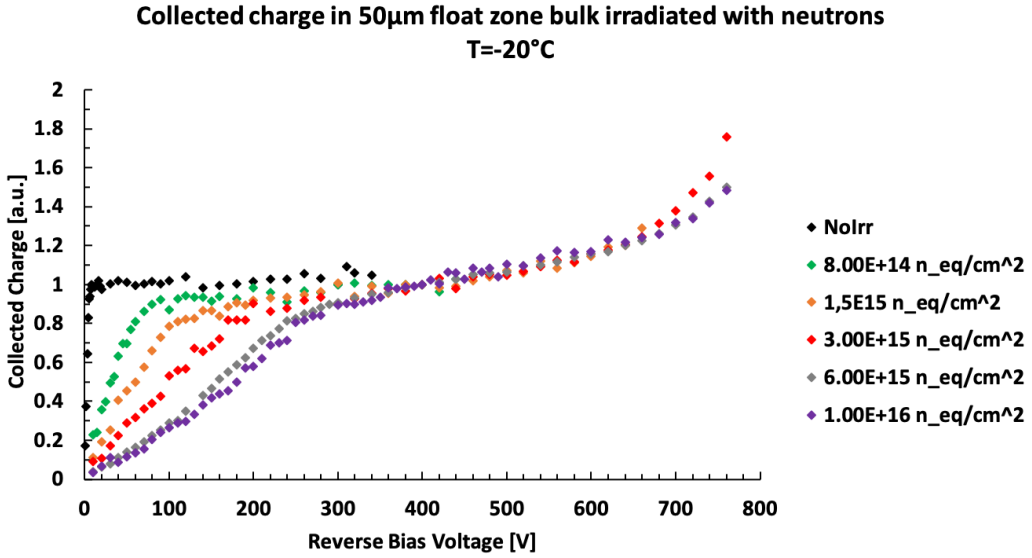

Figure 7.43: Collected charge as a function of external bias voltage measured, at a temperature of $T \sim -20\ ^{\circ}C$, on FBK-UFSD2 PiN diodes irradiated with neutron up to a fluence of $1 \cdot 10^{16}\ n_{eq}/cm^2$.

Table 7.11 reports the full depletion voltage values measured on the irradiated PiN diode.

Irradiation fluence $[n_{eq}/cm^2]$	Full depletion voltage [V]	
	Laser intensity 1	Laser intensity 2
0	5.3	6
$8 \cdot 10^{14}$	71.6	69.5
$1.5 \cdot 10^{15}$	111.8	112.3
$3 \cdot 10^{15}$	205.1	194.9
$6 \cdot 10^{15}$	258.6	
$1 \cdot 10^{16}$	294.4	337.4

Table 7.11: Full depletion voltage values of irradiated PiN diodes.

The Acceptor concentration obtained from CC curves has been plotted as a function of the irradiation fluence, in figure 7.44. The data, for both laser intensities, are in good agreement with the acceptor creation model up to fluence 3-6 $\cdot 10^{15}$ n_{eq}/cm^2, while they show a clear saturation effect above $6 \cdot 10^{15}$ n_{eq}/cm^2. The two dashed lines in figure 7.44 represent the acceptor creation model with a 2% and 3% growth of the concentration of acceptors, with the irradiation fluence.

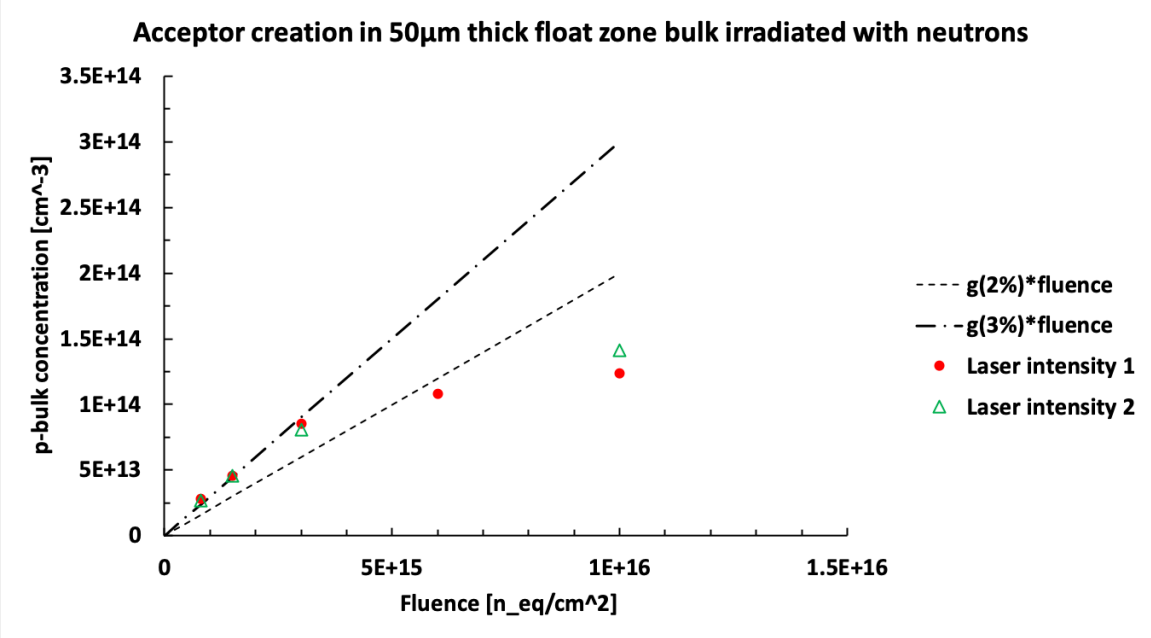

Figure 7.44: Bulk acceptor concentration as a function of neutron irradiation fluence, measured using two different laser intensity. The dashed lines represent the acceptor creation model with a 2% and 3% growth of the acceptor density with the fluence.

The CCE in irradiated sensors is an effect closely correlated to the acceptor creation in the bulk. At irradiation fluences high enough the deep traps created act as acceptors, and they can trap the charges created by the passage of a particle before they are collected by the sensor electrodes. As a result, the number of charges collected by the electrodes in an unirradiated sensor is larger than those collected in an irradiated one. The CCE measurements have been performed on the same devices used for the acceptor creation study. Using front-TCT setup with IR laser, measurements at three different laser intensities equal to $\sim$ 4, $\sim$ 11 and $\sim$ 24 MIPs have been performed. Figure 7.45 shows the CCE measurement as a function of irradiation fluence: the black markers represent the weighted average of the measurements performed at the three different laser intensities, while the grey band indicates the maximum uncertainty band of these measurements. All the CCE measurements have been performed on fully depleted sensors, at a bias voltage of 350 V. This plot shows that the CCE decreases up to a fluence of $\sim 3 \cdot 10^{15}$ n_{eq}/cm^2, while for higher fluences it seems to be saturated to a value of $\sim$ 70%.

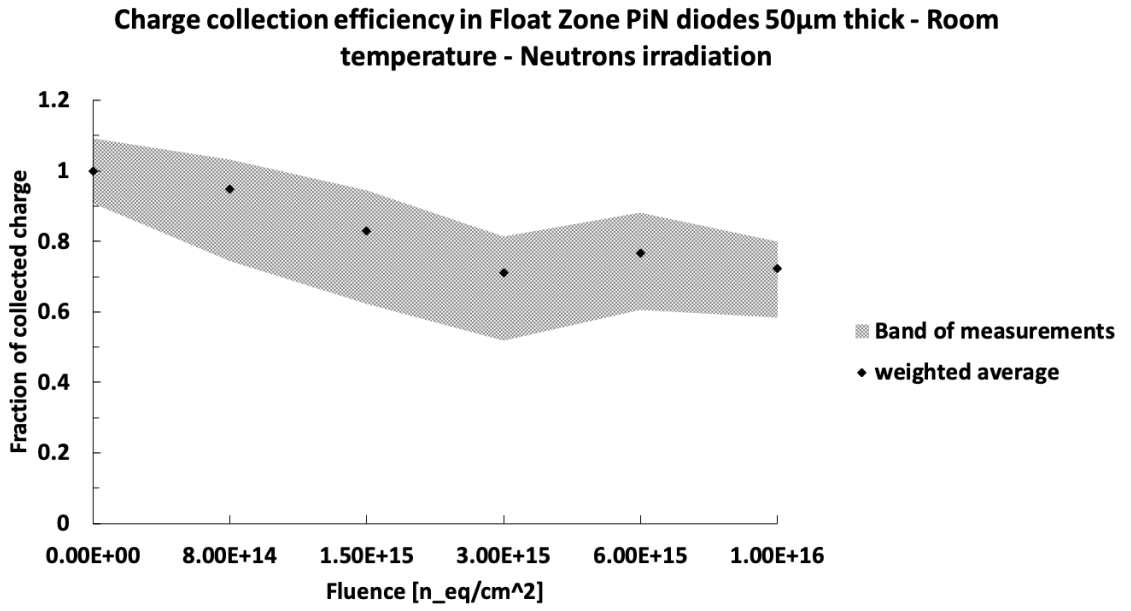

Figure 7.45: Charge collection efficiency as a function of neutron irradiation fluence for 50 μm FBK-UFSD2 PiN diodes.

In conclusion, the saturation effects observed in acceptor creation and CCE, in 50 μm thick PiN diode irradiated with fluence above $3 \cdot 10^{15}$ n_{eq}/cm^2, are unforeseen effects that need to be investigated in future studies. These saturation effects could extend the maximum fluence up to which devices of 50 μm thick can operate.

7.3.5 Measurement of the gain due to gain layer after a neutron fluence of 0.8, 1.5 and $3 \cdot 10^{15}$ n_{eq}/cm^2

Using the TCT setup with IR laser, measurements of gain as a function of the bias voltage have been performed on irradiated sensors. The gain has been measured on FBK-UFSD2 devices with three different types of gain layer, B LD (W1), B HD (W8) and B HD+C-A (W6), irradiated with neutrons at fluences of $8 \cdot 10^{14}$, $1.5 \cdot 10^{15}$ and $3 \cdot 10^{15}$ n_{eq}/cm^2. The devices chosen for these measurements are pair PiN-LGAD of active area 1 mm^2 shown in figure 7.46.

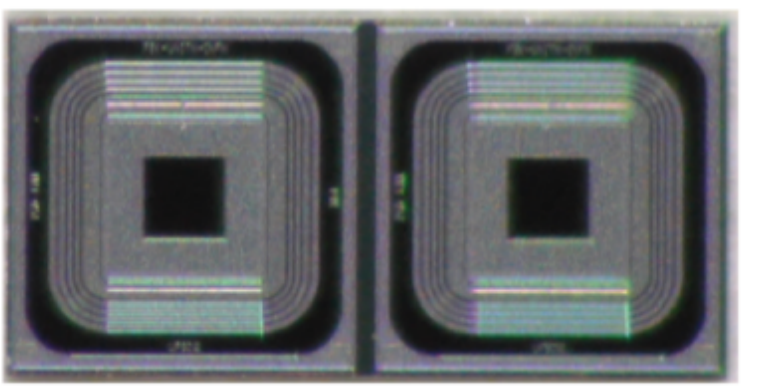

Figure 7.46: Example of FBK-UFSD2 PiN-LGAD device used for the gain measurements. Each sensor is 1 mm^2 wide and 50 μm thick.

The gain value is obtained as the ratio between the charge collected in an LGAD and in a PiN diode irradiated at the same fluence (see section 6.2); since both devices are irradiated at the same fluence, the effect of CCE discussed in the previous section is not included in this measurement. The results of gain measurements are shown in figure 7.47: the plot on the top left shows the curves of gain on the three types of gain

layers before irradiation, while the three other plots show the gain normalized to the respective unirradiated gain at the bias of 150 V. As expected from acceptor removal measurements the gain layer enriched with carbon is the more radiation resistant. After a fluence of $8 \cdot 10^{14}$ n_{eq}/cm^2 the measured gain at a bias voltage of 500 V is the same one measured in not irradiated device at 150 V. Moreover, this device is the only one among the three that still has gain at a fluence of $3 \cdot 10^{15}$ n_{eq}/cm^2.

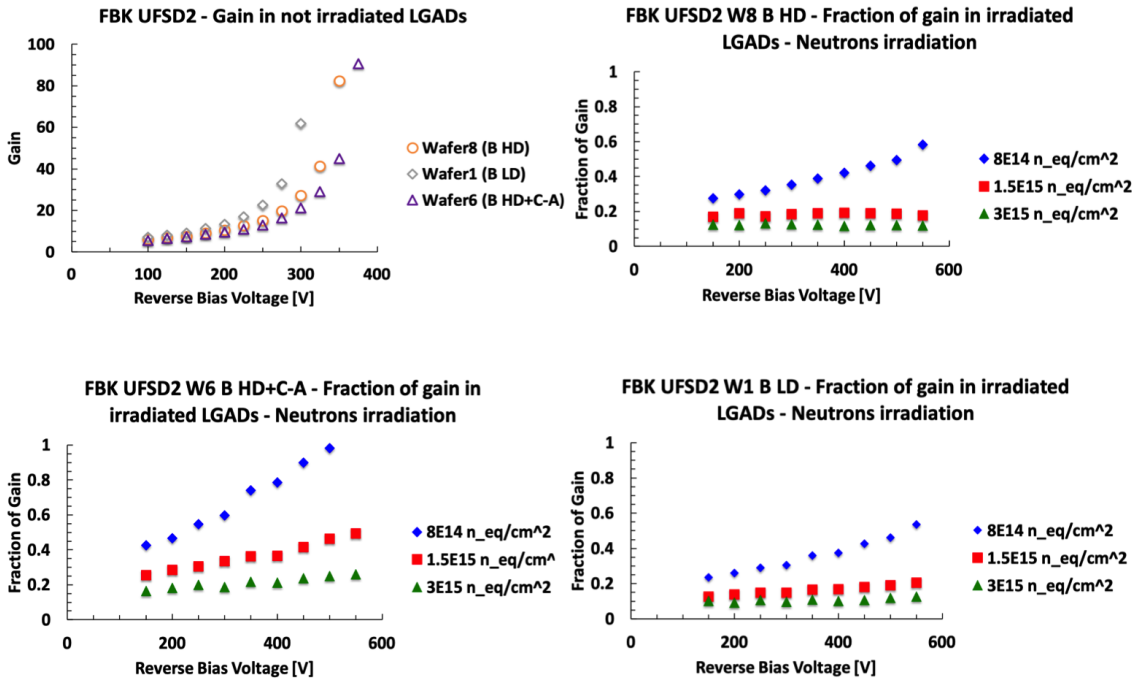

Figure 7.47: Top left plot: gain curves before irradiation. Following 3 plots for gain layer types B LD (W1), B HD (W8) and B HD+C-A (W6). The plots show the fraction of gain at 3 fluences, normalized to each respective gain at a bias voltage of 150 V.

7.4 Time resolution

Several beam tests have been performed, on UFSD sensors manufactured by FBK and HPK, at CERN and Fermilab facilities with the goal to measure their time performances. In this section, the most relevant results obtained during beam test campaigns are shown.

The beam test setup usually consists of one or more UFSD planes coupled to a

fast trigger (SiPM or UFSD); more detail on CERN and Fermilab beam test setup can be found in [67] and [68], respectively. The sensors under test are glued and wire-bonded on three different types of read-out board: the one developed in Turin shown in figure 6.12, a single channel board developed at the University of California Santa Cruz (UCSC) [67], and a multi-channels board developed at the Fermilab laboratory. The output signal from these boards is usually amplified by a current amplifier with gain 20 dB or 40 dB, and acquired using oscilloscopes with a horizontal sampling of 20 $GSample/s$ or 40 $GSample/s$. Figure 7.48 (left) shown an example of beam test frame with four UFSD planes and the fifth plane as trigger consisting of a SiPM sensor; while on the right, typical UFSD signals acquired during a beam test at CERN H8 site with π-mesons with a momentum of 180 GeV/c are shown. The UFSD signals shown are very fast (width of ~ 2 ns and rise time < 1 ns) with low noise, ideal for timing studies. Using this type of setup the time resolution is usually measured from the time difference between the trigger signal and the UFSD signal, as described in section 6.2.

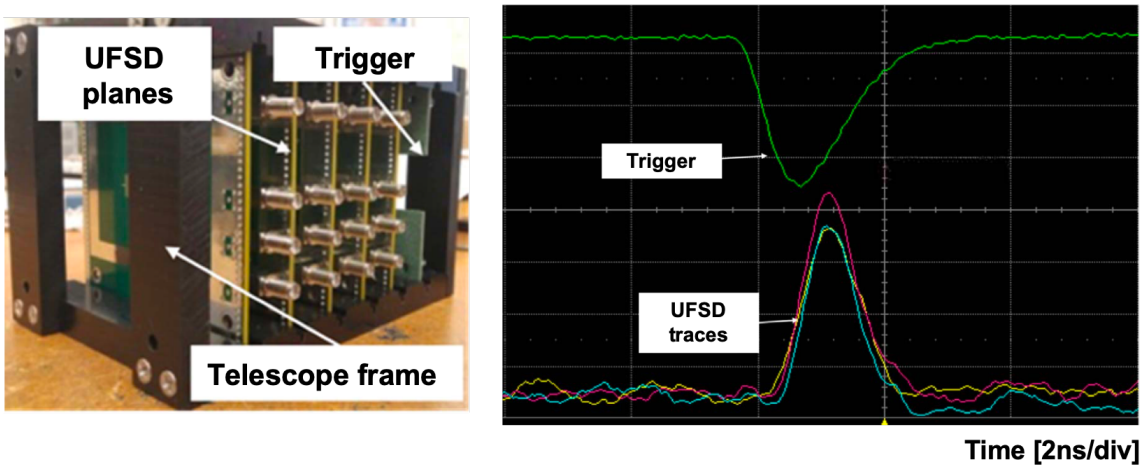

Figure 7.48: Left: picture of a frame with four UFSD planes and the fifth plane as trigger. Right: typical UFSD signals acquired during a beam test at CERN H8 site, with π-mesons with a momentum of 180 GeV/c. Figure from [67].

Figure 7.49 shows time resolutions measured on HPK-ECX20840-W10 (top) and

FBK-UFSFD2-W6 (bottom) UFSD devices, new and irradiated with neutrons to fluences of 0.6, 1 and $3 \cdot 10^{15}$ n_{eq}/cm^2. The time resolution decreases as the bias increases and falls below the 40 ps value for new and irradiated sensors to the two lower fluences, while it falls below the 50 ps value for those irradiated to $3 \cdot 10^{15}$ n_{eq}/cm^2. These results meet the CMS time resolution request for ETL sensors, which is a time resolution in the interval 30-40 ps throughout the HL-LHC lifetime, to degrade to 40-50 ps at a fluence of $3 \cdot 10^{15}$ n_{eq}/cm^2 ([3], chapter 3).

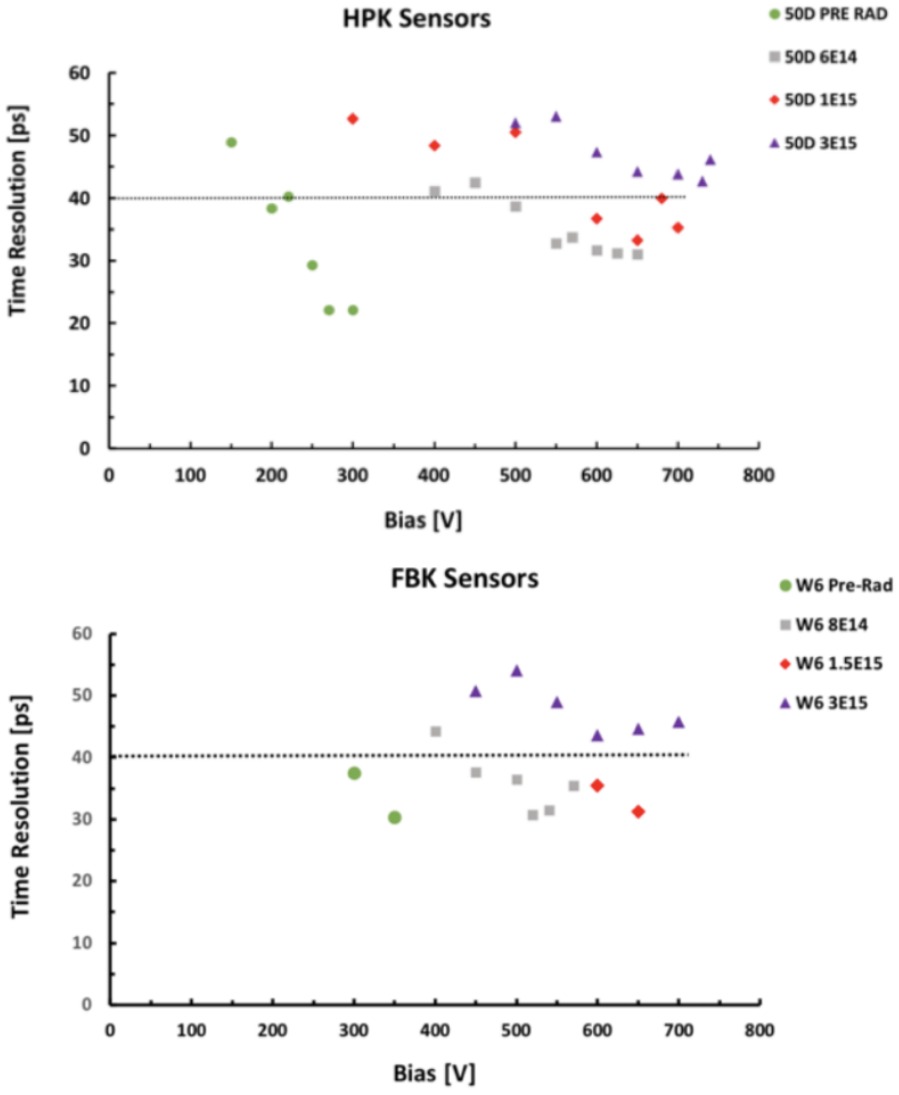

Figure 7.49: Time resolutions of HPK-ECX20840-W10 (top) and FBK-UFSD2-W6 (bottom) UFSDs, new and irradiated with neutrons to fluences of 0.6, 1 and $3 \cdot 10^{15}$ n_{eq}/cm^2. Figure from [3].

7.4.1 Unusual effects on time resolution measurements

Three unusual effects that lead to a worsening of the time resolution have been observed. The first two, called "metal effect" and "hole effect", occur in devices without metal cover on the n^{++} electrode; the third one is due to particle ionization (see section 3.1).

Metal effect occurs in sensors with an optical window. The Time of Arrival (ToA) of the signal generated in the middle of the optical window is delayed compared to those generated close to or underneath the metal cover. Figure 7.50 shows a 2D map, performed with the TCT setup, of the ToA of the signal related to the generation position on the optical window; the trigger signal generated by the laser controller was taken as a time reference to measure the ToA of the signal. The 2D map has been performed on a FBK-UFSD2-W1 4 mm^2 sensor, using x-y steps of 40 μm.

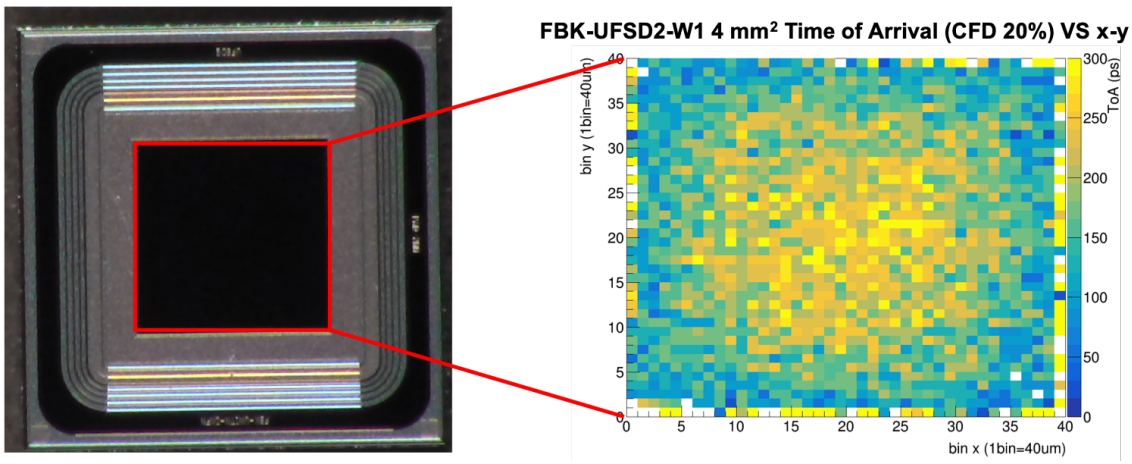

Figure 7.50: 2D map of the time of arrival of the signal, shooting an IR laser on the optical window of a FBK-UFSD2-W1 4 mm^2 sensor. The map was obtained with x-y steps of 40 μm.

The ToA of the signal decreases moving from the center to the periphery of the optical window. Figure 7.51 shows the ToA projection along x-direction of the 2D map crossing the center of the sensor; each measurement has been obtained as an average of hundred signals. From this ToA profile a maximum, time difference of

$\sim 100\ ps$ has been measured from the center to the periphery of the metal opening. The metal effect has also been observed during beam tests at Fermilab, where a tracker with a spatial resolution between 20 μm and 50 μm is available. This effect does not occur in fully metallized devices.

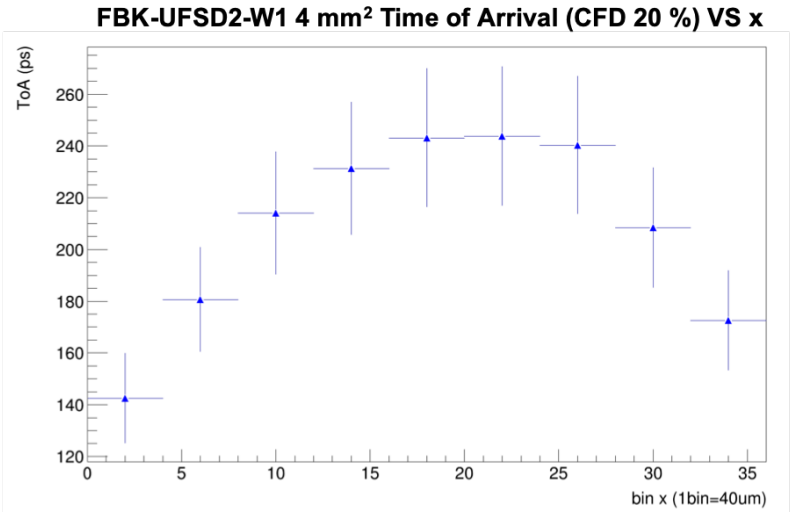

Figure 7.51: ToA projection along the x-axis of the 2D map in figure 7.50, crossing the center of the sensor.

Hole effect refers to the fact that UFSD sensors with surface metallization show a difference in gain between metallized and not metallized region, spoiling the uniformity and the time resolution of the sensors. This effect has been observed for the first time during beam tests at Fermilab National Laboratory: a gain variation of 10% (before irradiation) in [69] and of 50% (after irradiation) in [68] has been measured, worsening the time resolution of dozens of ps. Several TCT measurements campaign has been performed in laboratory reproducing this effect and validating the results obtained during the beam tests; a detailed discussion about these TCT measurements can be found in [58]. The results obtained during beam tests and with TCT measurements led the foundries to produce fully metallized sensors.

Metal and Hole effects have no a simple explanation: a hypothesis is that the created charges are initially collected by the n^{++} electrode and in a second time have

to move towards the metallic contact; if the no metal region has a slight different resistivity compared to the metal one, then the charges created in these different regions can acquire different group velocities leading to a different time of arrival and maybe to different gain.

The third unusual effect on the time resolution is due to the particle ionization process. The sensor on which this effect has been studied is a HPK-Type 3.1 4×4 array, with pads of active area $1 \times 3\ mm^2$. This sensor has been tested during a beam test at Fermilab. Figure 7.52 (left) shows the signal amplitude distribution acquired at 210 V of bias; as expected this distribution has the typical Landau distribution due to the ionization mechanism discussed in section 3.1. The amplitude distribution has been divided into three ranges in order to study the effect of ionization on time resolution. All time resolution measurements shown below have been obtained with CFD20% technique. The relationship between time resolution and ionization is shown in figure 7.52 (right).

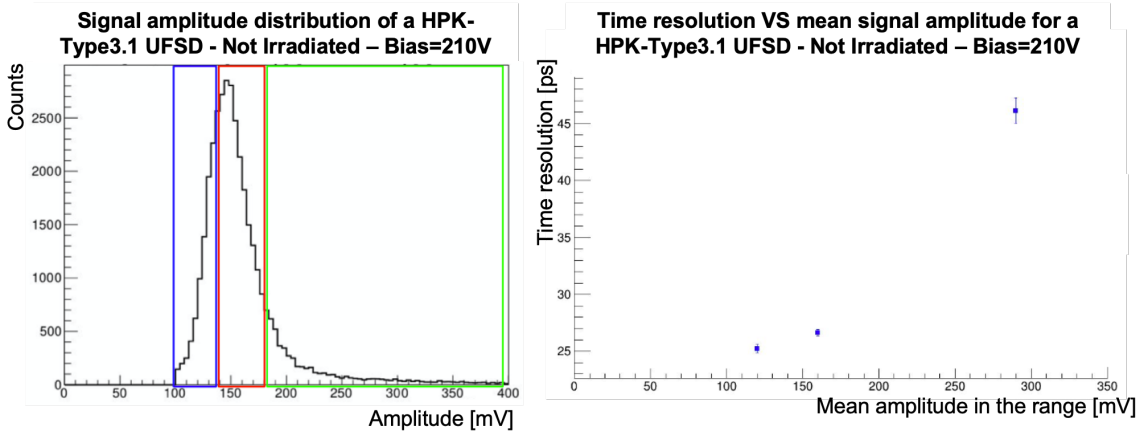

Figure 7.52: Left: Signal amplitude distribution for a HPK-Type 3.1 4×4 array at bias 210 V. The three colored frames represent the three regions of the distribution where the time resolution has been studied. Right: relationship between the time resolution and signal amplitude.

The time resolution for the first two range of landau distribution has values between 25 ps and 30 ps. In the last range of the distribution, the signals have large

amplitudes and rise times due to secondary ionizations, which cause a worsening of time resolution. On this sensor the time resolution has been measured at four bias voltages (175, 185, 205 and 210 V) showing the same dependency from the ionization; the results are plotted in figure 7.53.

From these data, it is clear that the time resolution does not only depend on the absolute value of the signal amplitude, but rather from the combination of gain and ionization. Groups of signal with the same average amplitude but different gain (different external bias) have a different time resolution. This analysis has been performed using different CFD threshold obtaining similar results. In conclusion, the relationship between time resolution and the combination of ionization and gain demonstrates a correlation between the time walk and landau terms discussed in section 4.2.

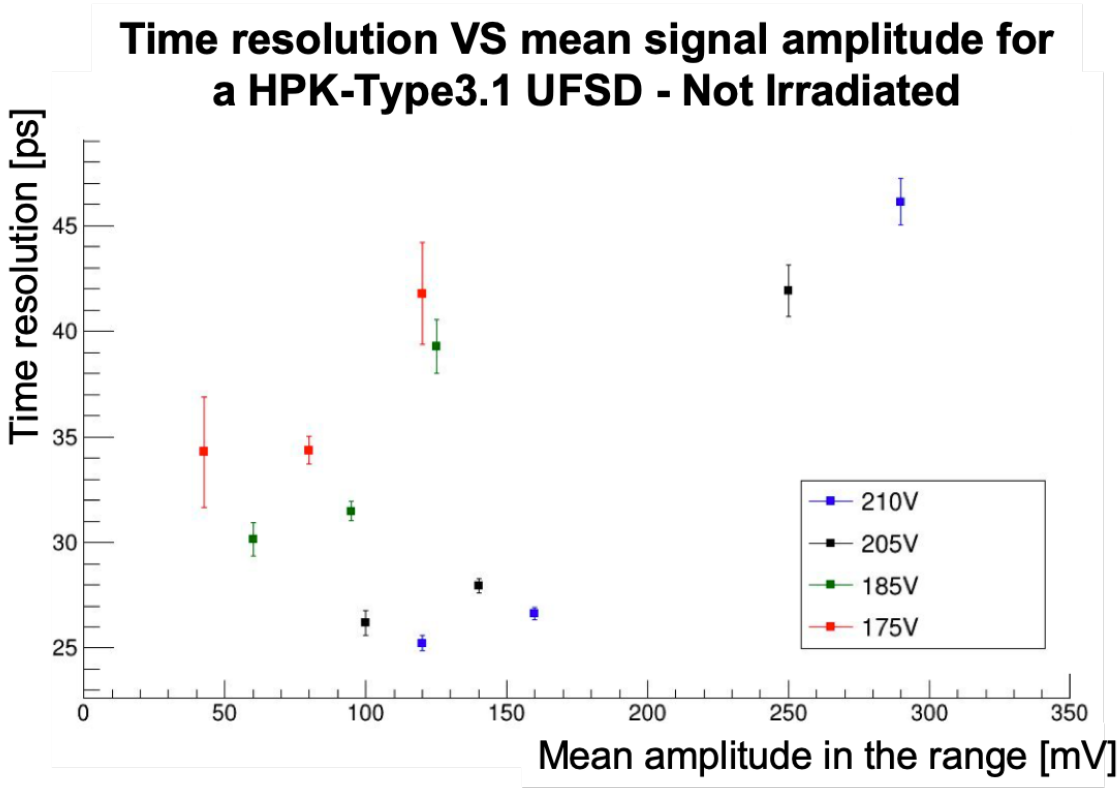

Figure 7.53: Relationship between time resolution and signal amplitude at different bias voltages.

7.5 Yield and uniformity of UFSD productions

The UFSD productions for the timing layer of ATLAS and CMS experiments at HL-LHC must satisfy the requirements of high yield and uniformity. The dimension of the final sensors for the timing layer of these two experiments will be larger than all the devices characterized in this chapter: they will have 30×15 (ATLAS) or 32×16 (CMS) pads of active area $1.3 \times 1.3\ mm^2$. Therefore, it is fundamental that the yield and uniformity of production are as high as possible, i.e. the working properties of each sensor and each pad must be as similar as possible. Characterization measurements on several wafers from FBK-UFSDF3 and HPK-EXX28995 productions have been performed, aiming at studying the yield of the sensors and the uniformity of the implanted doping concentration in the gain layer. The types of characterizations used for these studies are the IV and CV curves.

7.5.1 Yield and leakage current uniformity

The devices selected for the yield and leakage current uniformity studies have two different layouts, shown in figure 7.54. The first type of sensor has 96 pads with active area $1 \times 3\ mm^2$ each, arranged in an array of 4×24 pad, for a total size of $1.3 \times 2.5\ cm^2$. The second type is a 5×5 pad array, with full size $0.75 \times 0.75\ cm^2$, the single pad has an active area of $1.3 \times 1.3\ mm^2$. These devices represent an important R&D step towards the future production of full size ETL sensors; they allow us to evaluate the yield and the quality of two UFSD productions (FBK-UFSD3 and HPK-EXX28995), with a large number of identical multi-pad sensors.

The characterization of these devices consists of performing an IV measurement on each pad that composes the device. The BD voltage and leakage current uniformity are used as quality indexes of the wafers production. The measurement setup used consists of a multi-needle probe card connected via a switching matrix to the IV

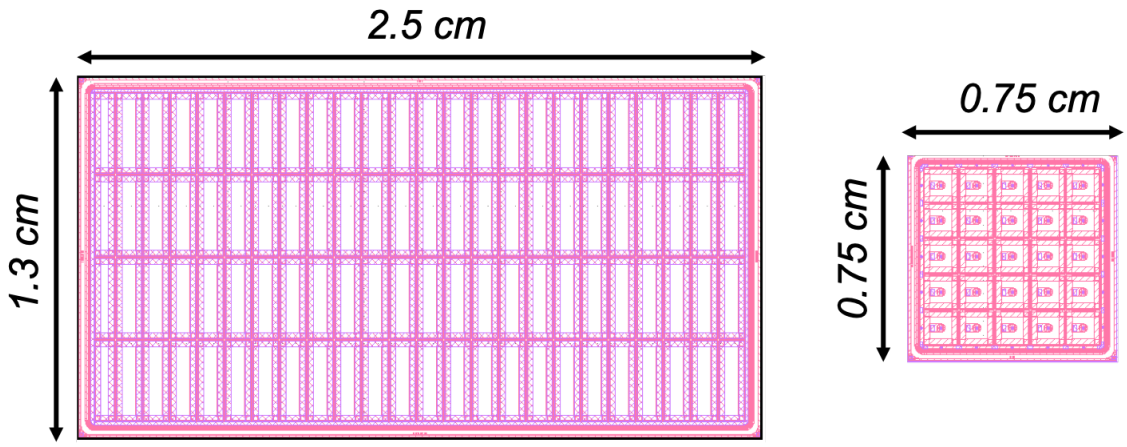

Figure 7.54: Layout of the sensors for the yield and leakage current uniformity tests on UFSD productions. On the left the array 4×24 sensor, on the right the array 5×5 sensor.

setup.

The following guidelines have been adopted to classify the pads under test:

- **Good pad**: leakage current within 10 times the mode current;
- **Warm pad**: leakage current above 10 times the mode current;
- **Bad pad**: the pad does not reach the required minimum voltage, sending the device in early BD.

A total of 152 FBK array 4×24 sensors (14592 pads), 23 FBK array 5×5 sensors (575 pads), and 16 HPK array 4×24 sensors (1536 pads) have been tested. Table 7.12 reports the results of these tests.

Foundries	Sensor type	# Sensors tested	# Warm pads	# Bad pads	Analysis bias $[V]$
FBK	array 4×24	152	14(0.1%)	0	100
FBK	array 5×5	23	4(0.7%)	0	300
HPK	array 4×24	15	20(1.3%)	0	230

Table 7.12: Summary of yield and leakage current uniformity studies on FBK and HPK production.

The FBK array 4×24 sensors have a percentage of warm pads of $\sim 0.1\%$. upon further inspection into the origin of these warm pads, it can be seen that they are

concentrated in the wafers with higher gain, namely not carbon enriched and with carbon dose A. Considering only this group of wafers, a warm pads percentage of about $\sim 0.2\%$ was found. A complete report on this testing campaign is reported in [70]. The FBK 5×5 arrays have a warm pad rate of $\sim 0.7\%$. The HPK 4×24 arrays have a warm pads percentage of $\sim 1.3\%$; of the 16 HPK devices tested, one was not considered in this analysis because it has a bad pad and it cannot be biased above 30 V. Optical inspection of this device revealed an obvious surface defect near the bad pad. Figure 7.55 shows an example of leakage current uniformity measurement performed on a HPK 4×24 array sensor biased at 230 V, just below the BD voltage (~ 250 V): this measurement shows low leakage current in the entire sensor and one warm pad with a larger but still acceptable current.

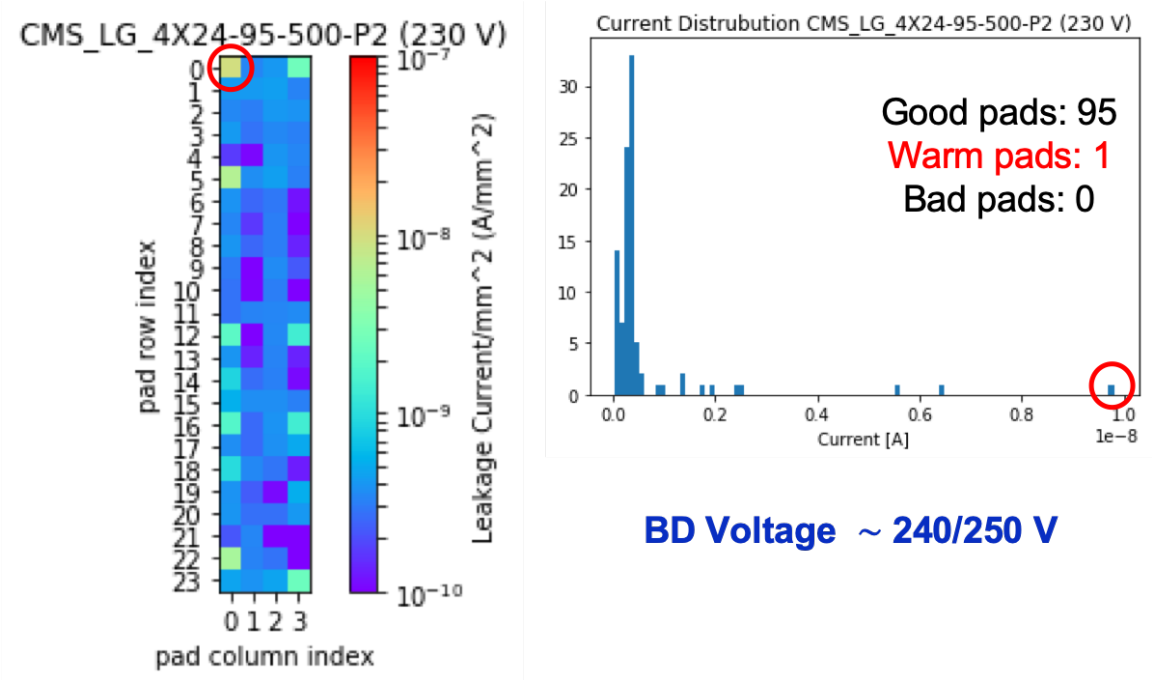

Figure 7.55: Left: map of the leakage current in a HPK array 4×24 sensor, biased at 230 V (BD voltage ~ 250 V). Right: leakage current distribution, where it is possible to notice the presence of a warm pad (red circle).

It is important to notice that the analysis of the two types of FBK devices has been performed at different bias voltages, as reported in the last column of table

7.12. The 4×24 devices suffer of earlier BD caused by the fact that many pads were not connected during testing. Since the test system used (probe card and switching matrix) has a limited number of channels (up to 40 channels). These floating pads cause a lowering of BD voltage. In the FBK devices, the decrease of BD voltage is very significant and it is of about 100 V, while in HPK sensors it is only a few dozen Volts, therefore less problematic. This effect is thought to be a consequence of the pad termination layout typology. Figure 7.56 shows this effect on a FBK 5×5 array, starting from a configuration with all pads connected to ground, the pads have been progressively floated: the device initially has a BD voltage of 350 V while this value drops disconnecting even one pad. The BD voltage stabilizes at about 150 V in a condition with only 4 pads connected out of 25. This effect could be dramatic for the ETL modules, as even a single pad not connected to the readout electronic lower significantly the BD voltage. For this reason, this anomalous behaviour has been studied in an internal FBK production, not object of this Ph.D. thesis.

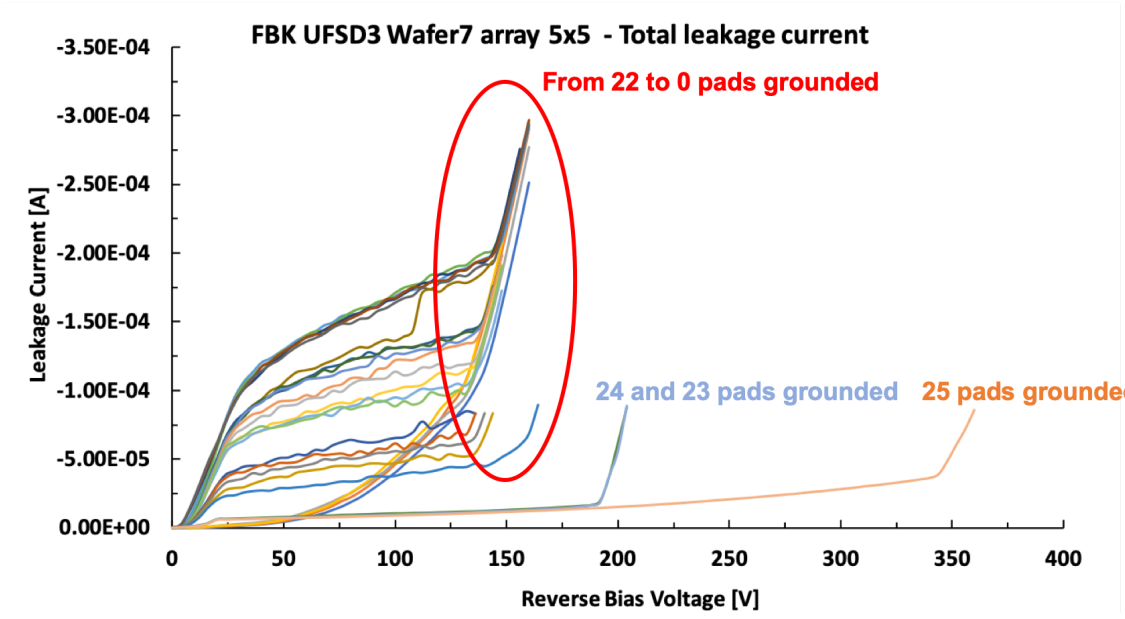

Figure 7.56: Total leakage currents on a FBK array 5×5 sensors of the wafer 7-UFSD3, in a condition of progressive disconnection of the pads from the ground. The drop in the BD voltage is remarkable even by disconnecting a single pad.

In conclusion to this study, it is possible to claim that large area UFSDs produced by the two foundries FBK and HPK have a very good overall leakage current uniformity and yield.

7.5.2 Gain uniformity study

Another important requirement for the UFSD productions of ATLAS and CMS is a good uniformity of gain, to have devices with behaviour as similar as possible to each other. Not uniformity of the gain layer implant generates not uniformity of device performances. The information of the gain layer implant uniformity can be obtained from CV measurements: the spread of the V_{GL} gives a direct measurement of the variation of the implanted doses in the gain layers (see equation 6.3). Using this measurement methodology, in this section the not uniformity of the gain layers implanted by FBK and HPK will be evaluated. The uniformity studies have been performed on wafers 3 and 4 of FKB-UFSD3 production and on wafers 2 and 4 of HPK-EXX28995 one, these pairs of wafers are identical. The sensors characterized have been selected on an area covering about half of the total wafer area, assuring a good coverage. In addition, the measurements performed on each wafer have been compared with those obtained on the twin one, in order to study the not uniformities of the productions. In figure 7.58, the regions on FBK and HPK wafers covered by these studies are surrounded by a red frame: on the left the wafer layout of the FBK production, while on the right the wafer layout of the HPK one.

A total number of 13 2×2 array sensors (one pad for each sensor) have been characterized on each FBK wafers, while a total one of 48 pads distributed on single pad sensors, 2×2, and 4×24 array have been characterized on HPK wafers. Figure 7.58 reports all the CV curves performed on the HPK wafer 4 (top) and on the FBK wafer 3 (bottom), from which it is possible to appreciate the spread of the V_{GL}. The results of the analysis are summarized in table 7.13. The FBK wafers 3 and 4

have a percentage of not uniformity of the gain layer depletion voltage ($(V_{GLmax} - V_{GLmin})/V_{GLmean}$) of $\sim$ 1.9% and $\sim$ 2.2%, respectively; the HPK wafers 2 and 4 show a not uniformity of $\sim$ 2.4% and $\sim$ 3.2%. In addition, for both foundries it is possible to estimate the not uniformity of the UFSD productions by comparing the CV measurements performed on both pairs of wafers: in the HPK production, a total not uniformity of $\sim$ 3.7% has been measured; in the FBK production a total one of $\sim$ 2.5%.

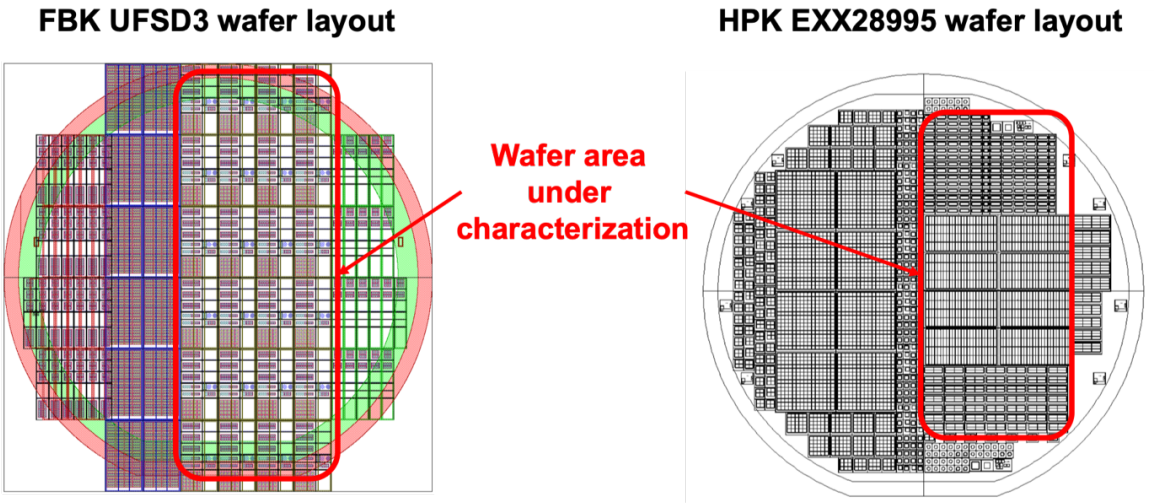

Figure 7.57: Layouts of wafers for gain uniformity study on UFSD productions FBK-UFSD3 (on the left) and HPK-EXX28995-Type3.1 (on the right); surrounded in red the wafer areas under characterization.

Foundries	Production	Wafer	$V_{GLmax} - V_{GLmin}$ [V]	$\triangle V_{GL}/V_{GLmean}$ [%]
FBK	UFSD3	3	0.48	$\sim$ 2.2
		4	0.42	$\sim$ 1.9
HPK	EXX28995-Type3.1	2	0.98	$\sim$ 2.4
		4	1.30	$\sim$ 3.2

Table 7.13: Summary of gain layer uniformity studies in FBK and HPK production.

Upon a further inspection into the CV curves in figure 7.58 a correlation between the position of the sensor on the wafer and its V_{GL} has been found. For FBK wafers, the not uniformity slightly increases from the centre to the top periphery of the wafer, where the depletion voltage of the gain layer is larger. Instead about the HPK

wafers, the not uniformity increases from top to bottom of the wafer and vice versa: the depletion voltage of the gain layer is always higher for the sensors located at the bottom of the wafer than those in the top.

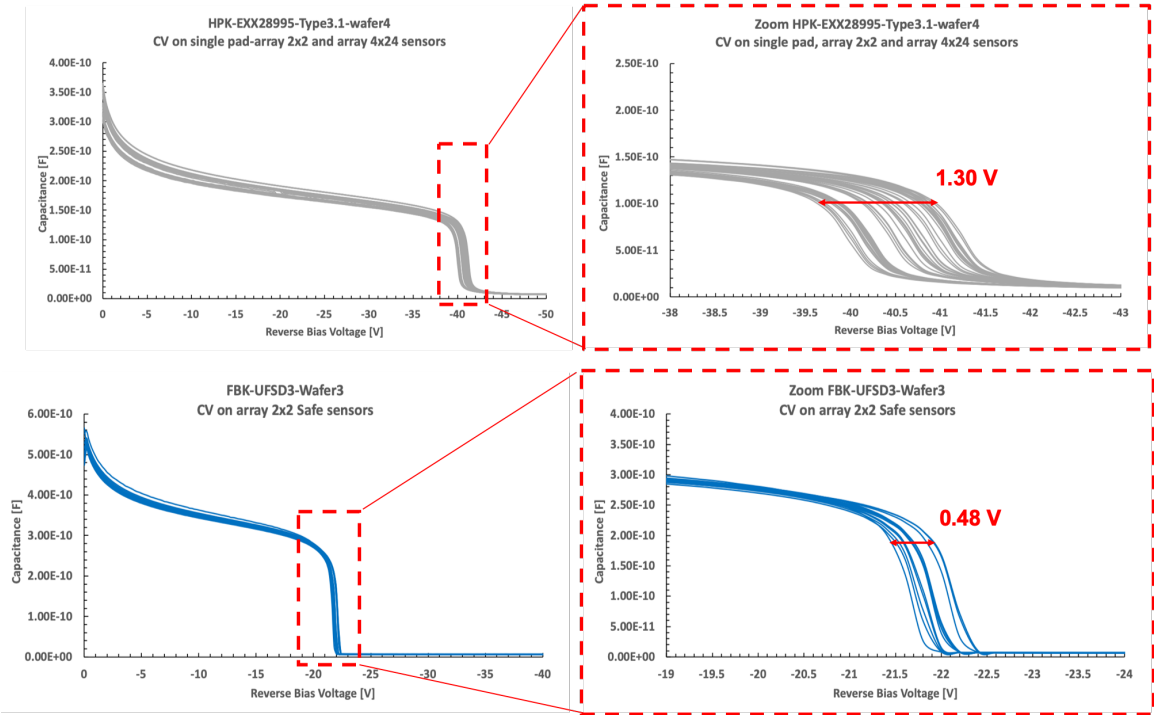

Figure 7.58: CV measurements on $1 \times 3\ mm^2$ pads from HPK-EXX28995-Type 3.1 wafer 4 (top) and FBK-UFSD3 wafer 3 (bottom). On the left, two zooms on CVs knee show the uniformity of the gain layer implant.

To complete these studies of uniformity, it is interesting to quantify the effect of the not uniformity of the gain layer implant on the gain value. This analysis has been performed on four single pad devices $1 \times 3\ mm^2$ from wafer 1 and 8 of the HPK-EXX28995-Type 3.1 production. The CV measurements of these devices, in figure 7.59 (top), show the spread in the depletion voltage of the gain layer, indicating a gain layer doping concentration spread of $\sim 2.5\%$. On the bottom of figure 7.59, the measurements of the collected charge performed with the TCT technique are shown. The collected charge is higher for the device with higher depletion voltage of the gain layer, as expected. A bias increment of $\sim 30\ V$ is required to collect the same amount of charge in the UFSD with a lower gain dose compared to the higher one.

This indicates that $\simeq 12V$ is needed to compensate 1% in doping variation.

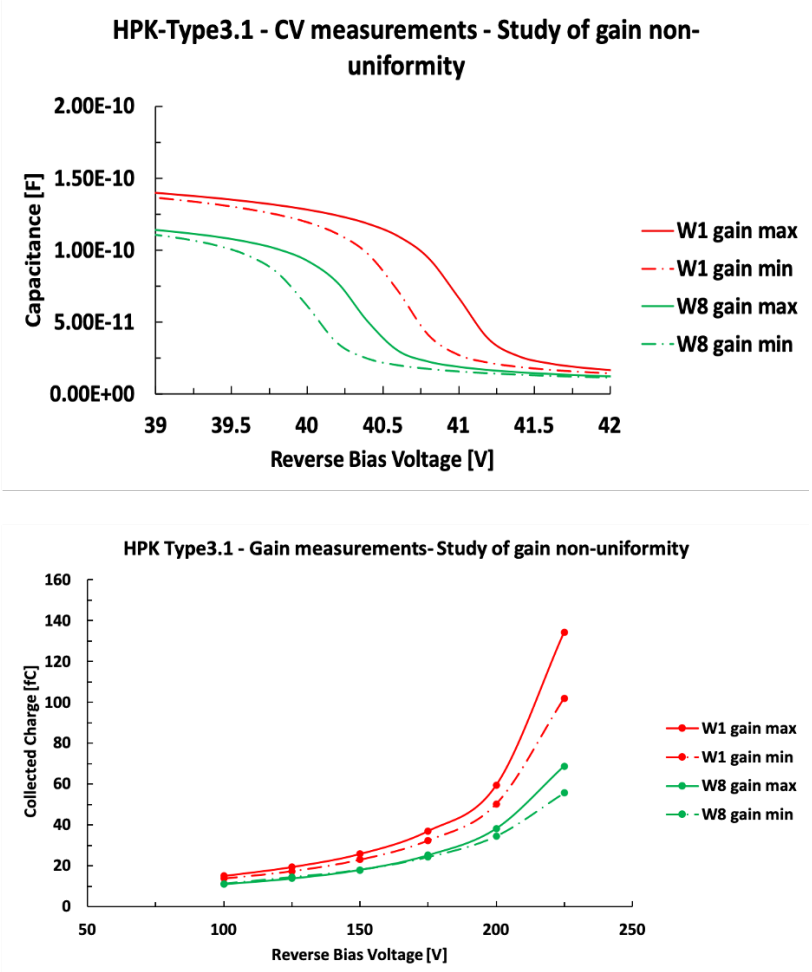

Figure 7.59: Evaluation of the not uniformity gain layer implant from CV measurements on the amount collected charge. On the top, the CV curves on $1 \times 3\ mm^2$ single pad devices under test, from wafer 1 and 8 of HPK-EXX28995-Type 3.1 production. On the bottom, the respective collected charge measurements as a function of external bias voltage.

From the collected charge measurements just discussed, it is possible to extract the fraction of the collected charge (CC/CC_{ref} at fixed bias) as a function of the fraction of gain layer doping (V_{GL}/V_{GLref}). This analysis, in figure 7.60 shows that, in this case of study, at a fixed voltage bias value of 200 V a variation in gain layer doping concentration of 1% causes a variation in the collected charge of $\sim 20\%$; similar results have been obtained at other bias voltage values. The data in figure 7.60 have

been compared with WF2 simulations of a UFSDs with an active thickness of 45 μm and gain layer types HPK-type 3.2 and FBK-B LD. The simulation of both types of devices has been performed at a fixed bias voltage of 150 V and normalized to a collected charge value of 5.5 fC. Figure 7.61 shows a very good agreement between simulation results and experimental data.

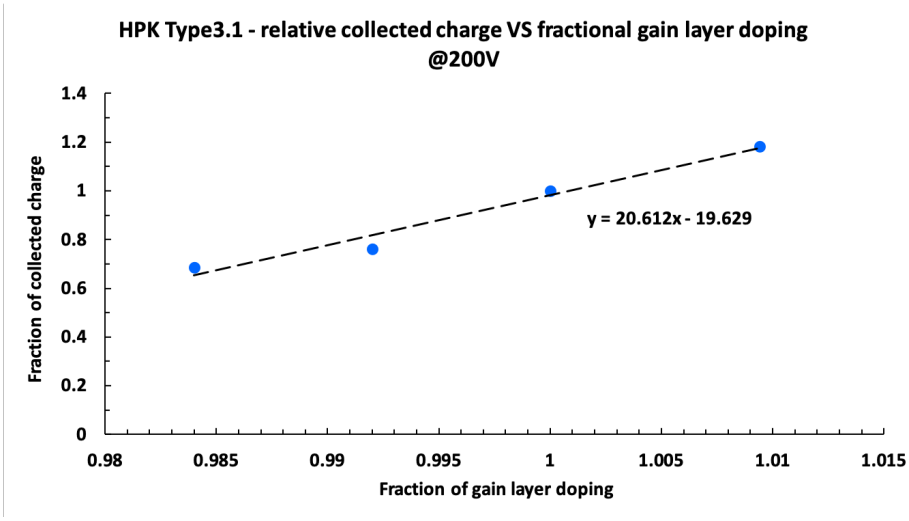

Figure 7.60: Impact of the fraction of the gain layer doping on the fraction of collected charge in HPK devices with gain layer Type 3.1 and bulk thickness of $\sim$ 45 μm, at 200 V of external bias.

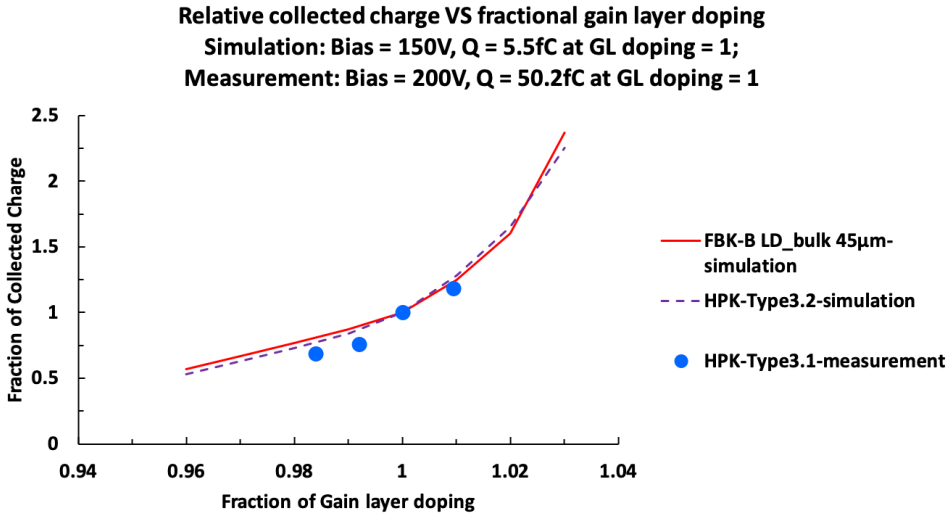

Figure 7.61: Fraction of collected charge variation, comparison between measurements on HPK-Type 3.1 and WF2 simulations on FBK-B LD and HPK-Type 3.2 45 μm thick sensors.

Chapter 8

Conclusions

This Ph.D. thesis is focused on the study and characterization of UFSD sensors. The UFSDs are innovative silicon sensors, based on LGAD technology, optimized to achieve a time resolution of 30 ps and a spatial resolution of $10s$ μm exploiting a moderate internal gain (~ 20) and an active thickness of ~ 50 μm. Moreover, recent radiation hardness studies proved the possibility to use these sensors up to irradiation fluence of $3 \cdot 10^{15}$ n_{eq}/cm^2.

The UFSD technology, thanks to its excellent time resolution and radiation hardness, has been chosen to instrument the endcap regions of the Minimum Ionizing Particle Timing Detector (MTD) of the CMS experiment at High Luminosity LHC. The CMS experiment, thanks to the MTD, will be able to maintain its current resolution, reconstruction efficiency, and background rejection also in the HL-LHC environment.

In this thesis, several studies performed on UFSDs produced by Fondazione Bruno Kessler (FBK) and Hamamatsu Photonics (HPK) have been reported. In particular, this work investigates the sensor characteristics required by CMS experiment, such as the time resolution of 30 ps up to an irradiation fluence of $1.5 \cdot 10^{15}$ n_{eq}/cm^2, the narrow inactive area, and uniformity of large area sensors. The UFSDs characterizations have been performed in the Laboratory of Innovative Silicon Sensors in the Physics Department of the University and INFN of Turin. The measurements have been performed with two setups: the probe station for the DC electrical tests, and the Transient Current Technique Setup (TCT) which exploits the signal induced in the device by a pulsed laser. Below, the key results obtained during this Ph.D. thesis are summarized.

- **Gain layer design:** different gain layer designs, of the UFSD productions by

FBK and HPK, have been characterized. The acceptor dopants used by FBK for their gain layers are boron or gallium, enriched or not with carbon; the gain layer of type gallium is wider than boron one, due to the higher diffusivity of the gallium in the silicon lattice. The gain layer with boron dopant has been processed at two different thermal loads, the one processed at lower thermal load has a narrower and higher profile and higher gain at the same bias voltage condition. The HPK gain layers are boron type and are deeper than FBK one; due to the different depth of implant, the HPK gain layers work in lower electric field regions than FBK ones. The relationship of gain value with gain layer dose and working temperature have been studied: a 1% variation in gain layer dose can be compensated with a bias variation of $\sim 12\ V$; while, a $1\ ^{\circ}C$ variation can be compensated with a bias variation of $\sim 1\ V$ and of $\sim 2\ V$ for HPK and FBK sensors, respectively.

- **Carbon implantation:** the effect of the co-implantation of carbon in the gain layer has been studied for FBK UFSD2 and UFSD3 productions, in which five different carbon doses have been implanted. The carbon reduces the active fraction of the gain layer (carbon-boron capture), reducing the internal gain of the device. Experimental measurements showed that carbon-boron capture is a threshold mechanism, which becomes more important for high doses of carbon.

- **Pad termination:** FBK-UFDS3 devices with termination layout Aggressive, Intermediate, and Safe show early breakdown; this phenomenon has been investigated mapping the devices with TCT setup, at a value of bias voltage close to the BD. These maps show an increase of the collected charge in the corners of the pads, in a region where there are the termination implants. These results have been confirmed by measurements performed with a CCD camera. early BD is probably due to a high electric field between JTE and p-stop termina-

tion. This phenomenon has not been observed in Super-Safe devices and in HPK sensors.

- **Pop-corn noise:** an unexpected effect called Pop-corn noise has been observed in FBK-UFSD3 sensors. This effect consists of microdischarges that occur above a certain value of bias, before the BD. This phenomenon could be caused by the p-stop too doped and sharp, which generates a p-n junction with the free charges in the oxide covering the sensor surface. In order to investigate this phenomenon and the premature BD, an internal FBK production called UFSD3.1 has been designed.

- **Inter-pad extension:** measurements of the extension of the inter-pad region have been performed on Aggressive, Intermediate, Safe and Super-Safe FBK UFSD3 sensors, and on HPK devices with four different termination layouts (layout factor 95, 70, 50 and 30). These measurements have been performed with the TCT setup. For FBK sensors have been measured inter-pad distances of 38.3 μm for the Super-Safe layout, 30.4 μm for the Safe, 16.7 μm for the Intermediate and 16.4 μm for the Aggressive. The results are in good agreement with the nominal distance between two neighboring gain layers except for the Aggressive layout, for which the measurements have been performed a too low value of bias voltage due to the premature BD. Inter-pad width measurements have been performed on HPK sensors with four different termination layouts; inter-pad widths of 130 μm, 114 μm, 92 μm, and 72 μm have been measured for the layout factors 95, 70, 50 and 30, respectively.

- **Trench Isolated-LGADs:** the requirement of a fine segmentation in Ultra Fast Silicon Detectors, for 4D particle tracking applications, led Fondazione Bruno Kessler to the development of the innovative Trench Isolated-LGADs, based on trench isolation technology as an alternative to the standard one

based on Junction Termination Extension and p-stop implants. Preliminary characterizations on Trench Isolated-LGADs show excellent electrical isolation between adjacent pads, an inter-pad distance of $\sim$ 7 μm has been measured reducing more than a factor of five the inter-pad extension compared to the Safe design in UFSD3 and a factor of two compared to the Intermediate and Aggressive ones.

- **Acceptor removal:** several irradiation campaigns with neutrons and protons with different energy (23 MeV/c, 70 MeV/c and 24 GeV/c) have been performed, aiming at studying the acceptor removal mechanism in the gain layer. From the comparison of CV measurements between new and irradiated sensors has been possible to estimate the acceptor removal rate for different types of gain layers, identifying the most radiation hard design. It has been measured that the gain layer enriched with carbon dose A doubles its radiation resistance. The B LD+ C-A is the most radiation resistant gain layer, which maintains the 80% of the active fraction of gain at a fluence of $1.5 \cdot 10^{15}$ n_{eq}/cm^2. An acceptor removal parametrization as a function of the initial acceptor density has been proposed, with good agreement with the experimental data. The comparison between data and parametrization shows that: boron and gallium dopants have the same radiation resistance; the gain layers enriched with carbon doses higher than dose A worsens the radiation resistance, and the carbon dose A is currently the dose that maximizes the gain layer radiation hardness. For sensors irradiated with protons, the acceptor removal rate has been measured as a function of the energy of the protons: the acceptor removal caused by protons of energy above GeV/c is similar to the same one caused by 1 MeV neutron equivalent, while the rate is higher for protons at lower energy.

- **Acceptor creation and charge collection efficiency:** acceptor creation and

charge collection efficiency measurements have been performed on PiN diodes irradiated with neutrons up to fluence $1 \cdot 10^{16}$ n_{eq}/cm^2. Experimental data, for both effect, show a saturation trend above a fluence of $3 \cdot 10^{15}$ n_{eq}/cm^2.

- **Time resolution:** several beam tests have been performed at CERN and Fermilab, aiming at measuring the time resolution of new and irradiated UFSDs. A time resolution of $\sim$ 30 ps has been measured on sensors irradiated up to a fluence of $1.5 \cdot 10^{15}$ n_{eq}/cm^2, while a time resolution of $\sim$ 50 ps has been measured on sensors irradiated to a fluence of $3 \cdot 10^{15}$ n_{eq}/cm^2. These results meet the CMS time resolution request for ETL sensors.

- **Production yield:** measurements of yield and leakage current uniformity on UFSD3 production by FBK and on EXX28995 production by HPK have been performed. For UFSD3 production, the 4×24 array devices have been characterized at FKB, while the 5×5 array devices have been characterized in Torino. On the 4×24 arrays, maps of leakage current of each sensor have been acquired selecting the leakage current for each pad at a bias of 100 V; from these maps a percentage of $\sim$ 0.1% of warm pads, that is pads with a leakage current above 10 times the mode current, has been calculated. Most of these warm pads are concentrated in wafers with high gain, where the percentage is $\sim$ 0.2%. On the 5×5 arrays, a percentage of warm pads of $\sim$ 0.7% at a bias voltage of 300 V has been calculated. For EXX28995 production, the yield measurements have been performed on the 4×24 array devices showing a percentage of warm pads of $\sim$ 1.3% at a bias voltage of 230 V.

- **Gain uniformity:** studies of the uniformity of the gain layer implant have been performed on the UFSD3 and EXX28995 productions. These studies have been performed on single pad devices, 2×2, and 4×24 arrays. UFSD3 production shown an overall not uniformity of $\sim$ 2.5%, while EXX28995 production shown

an overall not uniformity of $\sim$ 3.7%. Concluding these uniformity studies, the effect of not uniformity of gain layer implant on the gain value has been quantified: a 1% variation in gain layer doping causes a variation of the 20% of the collected charge. To compensate a 1% variation in the gain layer doping a bias variation of $\sim 12\ V$ is needed.

References

[1] L. Rossi and et al., "High luminosity large hadron collider hl-lhc," *CERN Yellow Report 2015-005*, 2017.

[2] B. Schmidt, "The high-luminosity upgrade of the lhc: Physics and technology challenges for the accelerator and the exsperiments," *Journal of Physics: Conference Series*, vol. 706, 2 2016.

[3] CMS-Collaboration, "A mip timing detector for the cms phase-2 upgrade," CERN, Tech. Rep., 2019.

[4] A. M. Sirunyan and et al., "Particle-flow reconstruction and global event description with the cms detector," *Journal of Instrumentation*, vol. 12, P10003–P10003, 10 2017.

[5] H. F.-W. Sadrozinski, A. Seiden, and N. Cartiglia, "4d tracking with ultra-fast silicon detectors," *Reports on Progress in Physics*, vol. 81, p. 026 101, 2 2017.

[6] H. F.-W. Sadrozinski and et al., "Ultra-fast silicon detectors (ufsd)," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 831, pp. 18–23, 2016.

[7] S. M. Sze, *Physics of Semiconductor Devices*. John Wiley and Sons, 1981.

[8] M. Moll, "The title of the work," PhD thesis, University of Hamburg, 1999.

[9] H. Bethe and W. Heitler, "On the stopping of fast particles and on the creation of positive electrons," *Proc. Royal Soc. London*, vol. A 146, pp. 83–112, 856 1934.

[10] S. Meroli, D. Passeri, and L. Servoli, "Energy loss measurement for charged particles in very thin silicon layers," *Journal of Instrumentation*, vol. 6, P06013–P06013, 06 2011.

[11] W. Shockley, "Current to conductors induced by a moving point charge," *J. Appl. Phys.*, vol. 9, pp. 635–636, 1938.

[12] S. Ramo, "Currents induced by electron motion," *Proceedings of the IRE*, vol. 27, pp. 584–585, 9 1939.

[13] M. S. Lazo, D. M. Woodall, and P. J. McDaniel, "Silicon and silicon dioxide neutron damage functions," *In Proc. Fast Brut React. Workshop, 1986*, vol. 1, pp. 85–103, 1987.

[14] G. Lindstrom, "Radiation damage in silicon detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 512, pp. 30–43, 1-2 2003.

[15] A. Affolder, P. Alport, and G. Casse, "Collected charge of planar silicon detectors after pion and proton irradiations up to 2.2×10^{16} $n_{eq}cm^{-2}$," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 623, pp. 177–179, 1 2010.

[16] G. Kramberger and et al., "Annealing studies of effective trapping times in silicon detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 571, pp. 608–611, 3 2007.

[17] N. Cartiglia and et al., "Design optimization of ultra-fast silicon detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 796, pp. 141–148, 2015.

[18] N Cartiglia and et al., "The 4d pixel challenge," *Journal of Instrumentation*, vol. 11, pp. C12016–C12016, 12 2016.

[19] N. Cartiglia and et al., "Tracking in 4 dimensions," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 845, pp. 47–51, 2017.

[20] N Cartiglia and et al., "Tracking in 4 dimensions," *Proceedings of science*, vol. 266, 2017.

[21] V. Sola and et al., "Ultra-fast silicon detectors for 4d tracking," *Journal of Instrumentation*, vol. 12, pp. C02072–C02072, 02 2017.

[22] C. Piemonte and et al., "Performance of nuv-hd silicon photomultiplier technology," *IEEE Transactions on Electron Device*, vol. 63, pp. 1780–1783, 03 2016.

[23] A. G. Chynoweth, "Ionization rates for electrons and holes in silicon," *Phys. Rev.*, vol. 109, pp. 1537–1540, 5 1958.

[24] M. Mandurrino and et al., "Tcad simulation of silicon detectors: A validation tool for the development of lgad," in *30th RD50 Workshop*, `https://indico.cern.ch/event/637212/contributions/2608617/`, 2017.

[25] F. Cenna, "Design and Test of Sensors and Front-End Electronics for Fast Timing in High Energy Physics," PhD thesis, Università degli Studi di Torino, 2018.

[26] M. Tornago, "Development of Ultra-Fast Silicon Detectors for 4D tracking at High Luminosity LHC: laboratory measurements and numerical simulations," Master's thesis, Università degli Studi di Torino, 2019.

[27] *Sentaurus device user guide*, N-2017.09, Synopsys, 2017.

[28] R. Van Overstaeten and H. De Man, "Measurement of the iuonization rates in diffused silicon p-n junctions," *Solid-State Elect.*, vol. 13, pp. 583–608, 5 1970.

[29] D. J. Massey, J. P. R. David, and G. J. Rees, "Temperature dependence of impact ionization in submicrometer silicon devices," *IEEE Trans. on Elect. Dev.*, vol. 53, pp. 2328–2334, 9 2006.

[30] Y. Okuto and C. R. Crowell, "Threshold energy effect on avalanche breakdown voltage in semiconductor junctions," *Solid-State Elect.*, vol. 18, pp. 161–168, 2 1975.

[31] A. Rivetti and et al., "Fast front-end electronics for semiconductor tracking detectors: Trends and perspectives," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 765, pp. 202–208, 2014.

[32] F. Cenna and et al., "Weightfield2: A fast simulator for silicon and diamond solid state detector," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 796, pp. 149–153, 2015.

[33] N. Cartiglia and et al., "Performance of ultra-fast silicon detectors," *Journal of Instrumentation*, vol. 9, pp. C02001–C020011, 02 2014.

[34] M. Mota and J. Christiansen, "A high-resolution time interpolator based on a delay locked loop and an rc delay line," *IEEE Journal of Solid State Circuits*, vol. 34, pp. 1360–1366, 10 1999.

[35] G. F. Dalla Betta and et al., "Design and tcad simulation on double-sided pixelated low gain avalanche detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 796, pp. 154–157, 2015.

[36] A. Rivetti, *CMOS: Front-End Electronics for Radiation Sensors*. CRC Press, 2015.

[37] M. Mandurrino and et al., "Analysis and numerical design of resistive ac-coupled silicon detectors (rsd) for 4d particle tracking," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 959, 2020.

[38] G. Paternoster and et al., "Trench-isolated low gain avalanche diodes (ti-lgads)," *IEEE Electron Device Letters*, 2020.

[39] M. Mandurrino and et al., "Demonstration of 200, 100, and 50 μm pitch resistive ac-coupled silicon detectors (rsd) with 100% fill-factor for 4d particle tracking," *IEEE Electron Device Letters*, vol. 40, pp. 1780–1783, 11 2019.

[40] G. Kramberger and et al., "Radiation effects in low gain avalanche detectors after hadron irradiations," *Journal of Instrumentation*, vol. 10, P07006–P07006, 07 2015.

[41] M. Ferrero and et al., "Radiation resistant lgad design," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 919, pp. 16–26, 2019.

[42] https://rd50.web.cern.ch/rd50/.

[43] B. Baldassarri and et al., "Signal formation in irradiated silicon detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 845, pp. 20–23, 2017.

[44] http://www.imb-cnm.csic.es/index.php/en/.

[45] https://www.hamamatsu.com/eu/en/index.html.

[46] https://www.fbk.eu/en/.

[47] G. Pellegrini and et al., "Technology developments and first measurements on low gain avalanche detectors (lgad) for high energy physics applications," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 765, pp. 12–16, 2014.

[48] M. Ferrero and et al., "Study and characterization of low gain avalanche diode," in *12th "Trento" Workshop on Advanced Silicon Radiation Detectors*, `https://indico.cern.ch/event/587631/contributions/2471714/`, 2017.

[49] G. Paternoster and et al., "Developments and first measurements of ultra-fast silicon detectors produced at fbk," *Journal of Instrumentation*, vol. 12, pp. C02077–C02077, 02 2017.

[50] A. Vignati and et al., "Innovative thin silicon detectors for monitoring of therapeutic proton beams: Preliminary beam tests," *Journal of Instrumentation*, vol. 12, pp. C12056–C12056, 12 2017.

[51] M. Berretti and et al., "Test of ultra fast silicon detectors for the totem upgrade project," *Journal of Instrumentation*, vol. 12, P03024–P03024, 03 2017.

[52] R. Arcidiacono, "A new timing detector for the ct-pps project," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 845, pp. 16–19, 2017.

[53] *Keysight technologies b1505a power device analyzer/curve tracer*, Keysight Technologies.

[54] M. K. Patterson and et al., "Charge collection and capacitance–voltage analysis in irradiated n-type magnetic czochralski silicon detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 583, pp. 189–194, 1 2007.

[55] https://www.technoprobe.com.

[56] http://particulars.si.

[57] https://cividec.at.

[58] F. Siviero, "Development of ultra-fast silicon detectors for time measurements at High-Luminosoty LHC," Master's thesis, Università degli Studi di Torino, 2018.

[59] N. Cartiglia, "Working points of ufsd at hl-lhc," in *35th RD50 workshop*, `https://indico.cern.ch/event/855994/contributions/3637011/`, 2019.

[60] Y. Unno and et al., "Optimization of surface structures in n-in-p silicon sensors using tcad simulation," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 636, pp. 118–124, 1 2011.

[61] A. Tournier and et al., "Pixel-to-pixel isolation by deep trench technology: Application to cmos image sensor," 2011.

[62] G. Snoj L. Zerovnik and A. Trkov, "Computational analysis of irradiation facilities at the jsi triga reactor," *Applied Radiation and Isotopes*, vol. 70, pp. 483–488, 3 2012.

[63] G. Kramberger and et al., "Radiation hardness of thin low gain avalanche detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 891, pp. 78–77, 2018.

[64] G Kramberger and et al., "Overview of sensor radiation tolerance at hl-lhc levels," in *HSTD11*, `https://indico.cern.ch/event/577879/contributions/2743556/attachments/1572732/2484235/Kramberger-RadHard-Final.pdf`, 2017.

[65] M. del Mar Carulla Areste and et al., "Last measurements and developments on lgad detectors," in *TREDI2017*, `https://indico.cern.ch/event/587631/contributions/2471730/`, 2017.

[66] G. Kramberger and et al., "Radiation hardness of gallium doped low gain avalanche detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 898, pp. 53–59, 2018.

[67] N. Cartiglia and et al., "Beam test results of a 16 ps timing system based on ultra-fast silicon detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 850, pp. 83–88, 2017.

[68] A. Apresyan and et al., "Study of uniformity of 50 μm low-gain avalanche detectors at the fermilab test beam," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 895, pp. 158–172, 2018.

[69] C. Allaire and et al., "Beam test measurements of low gain avalanche detector single pads and arrays for the atlas high granularity timing detector," *Journal of Instrumentation*, vol. 13, P06017–P0601, 06 2018.

[70] M. Tornago, "Performances of the third ufsd production at fbk," in *33rd RD50 workshop*, `https://indico.cern.ch/event/754063/contributions/3222642/`, 2018.

Acknoledgements

We thank the Fondazione Bruno Kessler of Trento, the INFN - Gruppo V and RD50 collaboration for the development of UFSD sensors. The work was supported by the United States Department of Energy, U.S. Department of Energy grant number DE-SC0010107. Part of this work has been financed by the European Union Horizon 2020 Research and Innovation funding program, under Grant Agreement no. 654168 (AIDA-2020) and Grant Agreement no. 669529 (ERC UFSD669529), and by the Italian Ministero degli Affari Esteri and INFN Gruppo V. Thanks to Dipartimento di Eccellenza, University of Torino (ex L. 232/2016, art. 1, cc. 314, 337).

Firstly, I would like to express my gratitude to my advisor Nicolò Cartiglia for the continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

My sincere thanks also go to my colleagues Prof. M. Costa, A. Staiano, Prof. R. Arcidiacono, Prof. M.M. Obertino, Dr. V. Sola, Dr. M. Mandurrino, Dr. J. Olave, F. Siviero and M. Tornago in for the stimulating discussions, and for all the fun we have had in the last four years.

Last but not least, I would like to thank my family and especially my wife Ilenia for supporting me during the writing of this thesis and my life in general.