UNIVERSITÀ
DEGLI STUDI
FIRENZE

PhD in
Physics and Astronomy
Cycle XXXIV

Development of Radiation Resistant Pixel Detectors for the Luminosity Frontier and Measurement of the Higgs Boson Production via Vector Boson Fusion with the CMS Experiment at the LHC

Tutors:
Prof. Piergiulio Lenzi
Dr. Marco Meschini

Coordinator:
Prof. Raffaello D'Alessandro

Candidate:
Dr. Rudy Ceccarelli

Years:
2018-2021

Aime la vérité, mais pardonne l'erreur.

Voltaire

Contents

Introduction **VII**

1 Silicon Pixel Sensors **1**
1.1 Semiconductors 1
1.2 Charge Generation 4
1.3 The *p*-*n* Junction 6
1.3.1 Properties of the Reverse Bias 8
1.3.2 Signal Formation 10
1.4 Radiation Damage 11
1.5 Pixel Sensors 16
1.5.1 Technologies 18
1.5.2 3D Pixel Sensors 21

2 LHC and the CMS Experiment **23**
2.1 The Large Hadron Collider 23
2.2 The CMS Experiment 26
2.2.1 The Tracker 28
2.2.2 The Electromagnetic Calorimeter 31
2.2.3 The Hadronic Calorimeter 33
2.2.4 The Muon Chambers 34
2.2.5 The Trigger System 36
2.3 High Luminosity LHC 38
2.3.1 The CMS Upgrade 40
2.4 The CMS Tracker Upgrade 41
2.4.1 The Inner Tracker 47
2.5 FBK Pixel Sensors Productions 52
2.5.1 Planar Batches 53
2.5.2 3D Batches 57

3 Pixel Readout Electronics **61**
3.1 The RD53A ROC . . . 61
3.1.1 The Linear AFE . . . 65
3.1.2 Single Chip Cards . . . 66
3.1.3 DAQ Systems . . . 69
3.2 Calibration Procedures . . . 71
3.3 Cross-Talk Studies . . . 76

4 Test Beam Measurements **83**
4.1 Tracking Concepts . . . 83
4.1.1 Cluster Reconstruction . . . 83
4.1.2 Trajectory Reconstruction . . . 84
4.1.3 Alignment Procedure . . . 85
4.2 DESY Test Beam Facility . . . 86
4.2.1 Beam Generation . . . 87
4.2.2 The DATURA Telescope . . . 88
4.2.3 The Experimental Setup . . . 89
4.2.4 Data Acquisition . . . 92
4.2.5 Offline Analysis . . . 93
4.3 Resolution Estimation . . . 98
4.3.1 Telescope Resolution . . . 99

5 Analysis of Pixel Detectors **103**
5.1 Overview . . . 103
5.2 Planar Pixel Detectors . . . 105
5.2.1 Irradiated 50 µm Pitch Detectors . . . 106
5.2.2 Fresh 25 µm Pitch Detectors . . . 113
5.2.3 Irradiated 25 µm Pitch Detectors . . . 122
5.3 3D Pixel Detectors . . . 129
5.3.1 Fresh 50 µm Pitch Detectors . . . 130
5.3.2 Irradiated 50 µm Pitch Detectors . . . 139
5.3.3 Fresh 25 µm Pitch Detectors . . . 146
5.4 Outlook . . . 154

6 The Higgs Boson at the LHC **155**
6.1 Electroweak Interactions . . . 155
6.1.1 The Brout-Englert-Higgs Mechanism . . . 157
6.2 Strong Interactions . . . 159
6.2.1 Jets Measurement . . . 160
6.2.2 Proton-Proton Collisions . . . 161
6.3 The Higgs Boson Phenomenology . . . 163

6.4 The WW Decay Channel . . . 166
6.4.1 The VBF Higgs Production Mechanism . . . 171

7 Analysis Strategy **175**
7.1 Analysis Overview . . . 175
7.2 Event Reconstruction . . . 175
7.3 Data and Simulated Samples . . . 177
7.3.1 Simulation Corrections . . . 180
7.4 Event Categorisation . . . 183
7.4.1 The Different-Flavour VBF Channel . . . 184
7.4.2 The Different-Flavour ggH Channel . . . 187
7.4.3 The Same-Flavour VBF and ggH Channels . . . 188
7.4.4 The VH Channels . . . 189
7.4.5 The STXS Measurement . . . 192
7.5 Deep Neural Networks . . . 196
7.5.1 The VBF-DNN . . . 200
7.6 Background Estimation . . . 209
7.7 Systematic Uncertainties . . . 211

8 Experimental Results **217**
8.1 Statistical Methodology . . . 217
8.1.1 Nuisance Parameters . . . 218
8.1.2 Statistical Significance . . . 220
8.2 Standard Analysis Results . . . 222
8.2.1 Complete Analysis . . . 228
8.3 STXS Analysis Results . . . 230

A *VBF*-DNN Input Variables **233**

Conclusions **IX**

Bibliography **XI**

Introduction

The Large Hadron Collider (LHC) is the largest circular accelerator ever build, allowing collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV, at a nominal instantaneous luminosity $\mathcal{L} = 1.0 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.
The Phase-2 of the LHC, known as High Luminosity LHC (HL-LHC), is going to start in 2027, aiming to reach an instantaneous peak luminosity up to $7.5 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. With HL-LHC, the Compact Muon Solenoid (CMS) experiment will gather an integrated luminosity up to $4000\,\mathrm{fb}^{-1}$ in 10 years, making it possible to study rare processes of the Standard Model (SM) or to search for processes beyond it. The CMS experiment will be upgraded between 2025 and 2027 to cope with the higher luminosity: especially in the regions near the collision point, unprecedented requirements in terms of radiation resistance and granularity need to be met.
The first part of this Thesis focuses on the upgrade of the CMS silicon tracker, whose inner section will be made of pixel detectors. The characteristics of the new tracker will be extremely important in the future analysis to be carried out in CMS during Phase-2.
For the new Phase, pixel sensors of new conception have been considered, in which the electrodes (p^+ and n^+) penetrate deep into the silicon from the same side of the sensor: these new pixels are referred as '3D' for their characteristic of having columnar implants as deep as the active thickness of the sensor, while the more conventional planar '2D' pixels have superficial implants of small thickness. Thanks to this structure, 3D sensors can have excellent performance even with high radiation damage, making them suitable for the use in the inner layers of the future CMS tracker. Due to the cutting edge technology needed to produce these sensors, their use for a large scale experiment has only recently become feasible. However, the production processes are more complex than those of planar sensors, and this affects costs and production efficiency. Therefore, 3D sensors have been taken into consideration only for the inner layers of the pixel tracker, while planar sensors will be used in the other layers.
In this Thesis, a complete characterisation of 3D and planar pixel detectors

is presented. The studies are performed at the INFN and CERN laboratories and in several test beam experiments at DESY. My work was crucial for the characterisation of the detectors both before and after irradiation, to verify that both the sensor and the readout chip are able to resist the high fluences expected at the HL-LHC with a minimum loss of performance.
The second part of this Thesis focuses on the measurement of the Vector Boson Fusion (VBF) Higgs production mechanism in the $H \to WW$ decay channel. A particle consistent with the SM Higgs boson was observed in 2012 by the CMS and ATLAS collaborations at the LHC. After the discovery, precision on the measurement of this new particle properties and interactions has progressed as more data were collected. Currently, all production processes have been observed in one or more decay channels or via combination of several decay channels, with no significant deviations with respect to the SM prediction. However, the VBF mechanism, being at the heart of the electroweak symmetry breaking, needs to be studied with ever-improving analysis techniques while waiting for additional data to reduce the statistical uncertainty.
The cross section for the VBF mechanism in proton-proton collisions at a center of mass energy of 13 TeV has been measured by CMS in several Higgs decay channels. The $H \to WW$ decay, thanks to its large branching ratio, is ideal for the observation of this production process. The most recent CMS analysis in the $H \to WW$ decay channel, however, was mainly focused on the measurement of the global production cross section: the analysis was not optimized with respect to the VBF production mode.
In this Thesis, a multivariate analysis was implemented in order to enhance the sensitivity of the measurement of the VBF mechanism in the $H \to WW$ decay channel. In particular, a Deep Neural Network (DNN) was developed in order to isolate the signal events from the background, which is mainly composed by top quarks events, non-resonant WW and gluon fusion Higgs boson production mechanism. The DNN yields four scores for each event, corresponding to the degree of compatibility either with the signal or with one of the main backgrounds. These scores are then combined and used in the fitting procedure. This innovative approach was necessary because one of the main backgrounds of this analysis is another Higgs production process, therefore making it difficult to tackle this analysis in a simple signal versus background paradigm. This study is based on the whole Run-2 dataset, collected from 2016 to 2018 with the CMS experiment.

Chapter 1

Silicon Pixel Sensors

1.1 Semiconductors

The periodic lattice of a crystal defines energy bands for electrons. If the bands are fully filled no current can flow. The difference between insulators and conductors is that in the fundamental state insulators have fully filled bands, while conductors have at least one partially filled band. This classification is only valid for $T = 0$ K: at higher temperatures the electrons can jump from the last filled band (valence band) to the first empty band (conduction band). The energy difference between the minimum of the conduction band and the maximum of the valence band is known as bandgap[1] E_g.
Electrons follow the Fermi-Dirac statistics: the fraction of electrons that can jump from one band to the other is proportional to $e^{-E_g/2k_BT}$: if E_g is small with respect to k_B, a conductive behaviour is observed. Crystals with $E_g < 5\,\text{eV}$ are referred to as semiconductors. The typical semiconductor resistivities are $10^{-5}\,\Omega\text{m} < \rho < 10^{3}\,\Omega\text{m}$ (for conductors $\rho < 10^{-5}\,\Omega\text{m}$ while for insulators $\rho > 10^{3}\,\Omega\text{m}$). Silicon and germanium are the most common semiconductors (both tetravalent).
When an electron jumps from the valence band to the conduction band a hole is left in the lattice: holes can be treated as a positive charge carriers following the Fermi-Dirac statistics. In semiconductors electrons and holes continuously generate and recombine. In a pure (intrinsic) semiconductor the

[1] The energy bands also depend on the crystal momentum k. If the minimum of the conduction band and the maximum of the valence band correspond to the same k value, the semiconductor is referred to as direct, otherwise as indirect. A transition between the two bands can be induced by an incident photon of energy higher than the bandgap. In an indirect semiconductor a phonon is also necessary in order to conserve the crystal momentum.

concentrations of electrons n (the number of electrons per unit volume) and holes p are always equal: this value is referred to as intrinsic concentration n_i. In particular:

$$n_i^2 = n^2 = p^2 = CT^3 e^{-E_g/k_B T} \tag{1.1}$$

where C is a constant of the material.
When an electron-hole pair is generated, both carriers move by diffusion. The charge distribution varies with time, and the transverse section can be described by a gaussian distribution with:

$$\begin{aligned} \sigma_e &= \sqrt{2D_e t} \\ \sigma_h &= \sqrt{2D_h t} \end{aligned} \tag{1.2}$$

for electrons and holes respectively. The time passed from the pair generation is t and D is the diffusion coefficient.
If an electric field is applied to the semiconductor, the charge carriers move in opposite directions. The drift velocities are proportional to the applied field:

$$\begin{aligned} \boldsymbol{v}_e &= -\mu_e \boldsymbol{E} \\ \boldsymbol{v}_h &= \mu_h \boldsymbol{E} \end{aligned} \tag{1.3}$$

where μ_e and μ_h are the mobilities of electrons are holes respectively. By increasing the electric field, the drift velocity increases as well but, as the mean free path is the same, the number of collisions with the lattice atoms also increases. Eventually, a saturation velocity is reached (of the order of 10^7 cm/s). The current density is given by:

$$\boldsymbol{J} = e(n\mu_e + p\mu_h)\boldsymbol{E} \tag{1.4}$$

Therefore, the resistivity is:

$$\rho = \frac{1}{e(n\mu_e + p\mu_h)} \tag{1.5}$$

where e is the absolute value of the electron charge.
The diffusion coefficient can be expressed in terms of the mobility with the Einstein relations:

$$\begin{aligned} D_e &= \mu_e \frac{k_B T}{e} = \mu_e V_T \\ D_h &= \mu_h \frac{k_B T}{e} = \mu_h V_T \end{aligned} \tag{1.6}$$

Property	Value
Atomic Number (Z)	14
Atomic Mass (A)	28.1
Density (ζ)	2.33 g/cm^3
Relative Dielectric Constant (ϵ_r)	12
Bandgap (E_G)	1.1 eV
Intrinsic Concentration (n_i)	$1.5 \cdot 10^{10}$ cm^{-3}
Resistivity (ρ)	2300 Ωm
Electrons Mobility (μ_e)	0.14 m^{2}V^{-1}s^{-1}
Holes Mobility (μ_h)	0.05 m^{2}V^{-1}s^{-1}
Electrons Diffusion Coefficient (D_e)	$3.4 \cdot 10^{-3}$ m^2/s
Hole Diffusion Coefficient (D_h)	$1.3 \cdot 10^{-3}$ m^2/s

Table 1.1: Most important properties of silicon at room temperature (300 K).

where $V_T \simeq 26\,\text{mV}$ at room temperature. In presence of an electric field, Equations 1.2 can be combined with Equations 1.3 and Equations 1.6:

$$\sigma = \sqrt{\frac{2V_T x}{E}} \tag{1.7}$$

where x is the drift distance. Table 1.1 reports the most important properties of silicon at room temperature.
Impurities can be added to a semiconductor in order to reach the desired doping. Atoms of trivalent elements (e.g. boron, gallium, indium), referred to as acceptors, or atoms of pentavalent elements (e.g. antimony, phosphorous, arsenic), referred to as donors, are substituted to silicon atoms in the lattice: in the first case the doping is of type n, while in the second case it is of type p. The fraction of substituted atoms is of the order of $10^{-3}-10^{-9}$, and the semiconductor is referred to as extrinsic. In the n doping the additional electron energy level is close to the conduction band: at room temperatures the impurities are fully ionised. Similarly, in the p doping the hole is close to the valence band: an electron from another atom moves to the impurity, leaving a hole in the lattice.
Increasing the concentration of one charge carrier decreases the concentration of the other one. Indeed, the electron-hole recombination probability is proportional to $n \cdot p$, while the pair generation by thermal excitation is doping independent. This equilibrium is expressed by the mass action law:

$$n \cdot p = n_i^2 \tag{1.8}$$

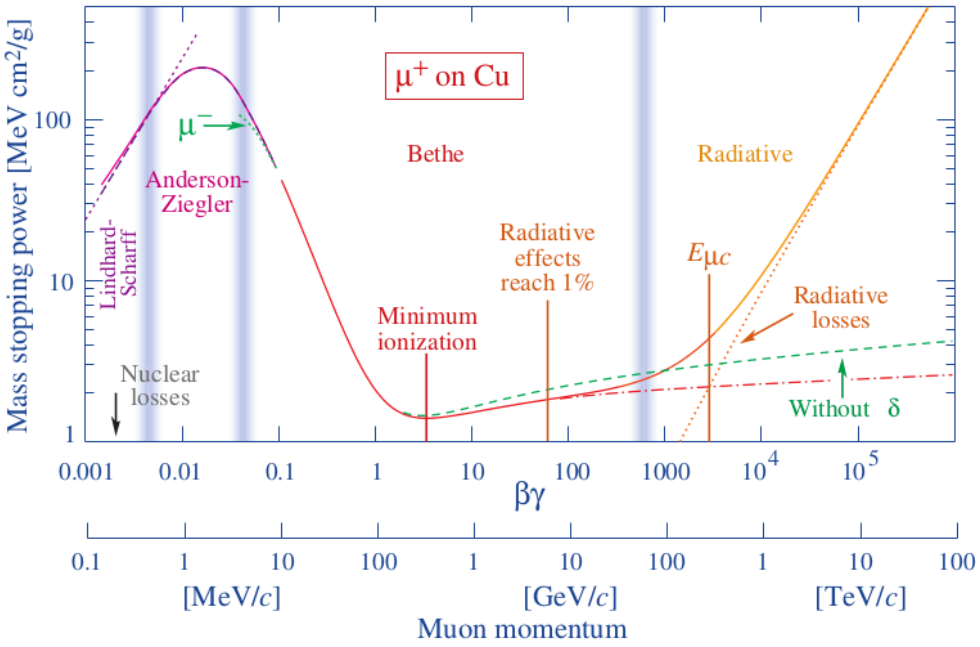


Figure 1.1: Stopping power for a positive muon in copper, as a function of $\beta\gamma$. The continuous lines indicate the total stopping power and the red curve corresponds to the Bethe-Bloch formula: the vertical lines indicate the validity limits [1].

For a n doping, if N_D is the concentration of donor atoms, it follows that:

$$N_D \gg n_i, \quad n_n \simeq N_D, \quad p_n \simeq n_i^2/N_D \tag{1.9}$$

where the subscript indicates the doping type. In this case, electrons are referred to as the majority carriers. Similarly, for a p doping, if N_A is the concentration of acceptor atoms:

$$N_A \ll n_i, \quad p_p \simeq N_A, \quad n_p \simeq n_i^2/N_A \tag{1.10}$$

and holes are referred to as the majority carriers. Typical values of N_A and N_D are $10^{15} - 10^{18}\,\mathrm{cm}^{-3}$.

1.2 Charge Generation

When a charged particle passes through a semiconductor, electron-hole pairs are generated along its trajectory (this process does not depend on the doping). The number of produced pairs $\mathcal{N}$ is proportional to the charged particle

energy loss E_{loss}:

$$\mathcal{N} = \frac{E_{loss}}{\varepsilon} \tag{1.11}$$

where ε is the average energy for generating an electron-hole pair. In silicon[2] $\varepsilon = 3.6$ eV. Delta rays can be produced during the interaction, that is high energy electrons that can generate further electron-hole pairs, even at long distances from the particle trajectory.
The mean energy loss for a charged particle traversing a medium is the stopping power S, measured in MeVcm^2/g. The stopping power be can calculated with the Bethe-Bloch formula [1]:

$$S = 4\pi N_A r_e^2 m_e c^2 z^2 \frac{Z}{A}\frac{1}{\beta^2}\left(\frac{1}{2}\ln\frac{2m_e c^2\beta^2\gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta}{2}\right) \tag{1.12}$$

where N_A is the Avogadro number; r_e the classic electron radius; m_e the electron mass; c the speed of light, z the atomic number of the interacting charged particle; Z the atomic number of the medium; A the mass number of the medium; T_{max} the maximum energy that can be transferred to an electron in a collision; I the average ionisation energy; β, γ are referred to the interacting charged particle; $\delta = \delta(\beta\gamma)$ is a density correction for high energies, due to the polarisation of the medium.
The Bethe-Bloch formula is valid in the interval $0.1 < \beta\gamma < 1000$. At low energies ($\beta\gamma \simeq 0.1$) the first term dominates, therefore S decreases with increasing energy. A minimum is reached for $\beta\gamma \simeq 3$: particles in this region are referred to as MIPs. In high energy physics experiments particles measured with tracking devices can be considered as MIPs. At higher energies the logarithmic term slowly increases S, while the density correction mitigates the growth. At very high energies ($\beta\gamma > 1000$) radiative effects, not considered in the formula, start to be important. Figure 1.1 shows the stopping power for a positive muon in copper, as a function of $\beta\gamma = p/(m_\mu c)$.
The charge generation process in the active volume of a detector has statistical fluctuations: due to the presence of delta rays, the collected charge can be described by a Landau distribution[3]. Usually, the Landau Most Probable Value (MPV) is used to estimate the energy loss (the MPV is about 30% less with respect to the mean value). The stopping power for a MIP in silicon is:

$$S_{min} = 1.66\,\text{MeVcm}^2/\text{g} = 3.87\,\text{MeV/cm} \tag{1.13}$$

[2]The silicon bandgap is 1.1 eV, but being an indirect semiconductor, a higher energy is necessary to generate an electron-hole pair.

[3]This is the case for MIPs, that fully traverse the detector, without being stopped.

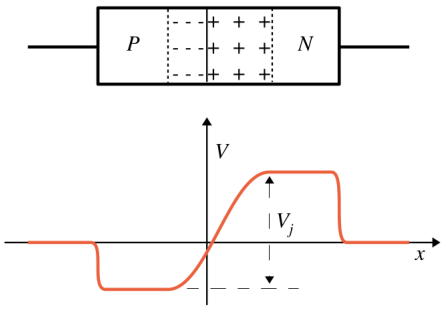


Figure 1.2: Electric potential of a p-n junction [2].

The average electron-hole pairs generated in $d = 1$ µm is $S_{min} \cdot d/\varepsilon = 108$ while the MPV is 76.
By considering an intrinsic silicon volume of area $A = 1\,\mathrm{cm}^2$ and thickness $d = 300$ µm, about $108 \cdot 300 = 3.2 \cdot 10^4$ pairs are generated by the passage of a MIP. However, the number of free carriers is four times higher: $n_i \cdot A \cdot d = 4.5 \cdot 10^8$. The signal to noise ratio is therefore too small: in order to decrease the number of free charge carriers, the p-n junction is employed for building silicon detectors.

1.3 The *p-n* Junction

A p-n junction is formed by contacting two extrinsic semiconductor regions, one of type p and one of type n. Diffusion currents will flow in the junction, due the large concentration gradients: the holes move to the n side and the electrons to the p side where they recombine to electrons and holes respectively.
A charged region, referred to as depletion region, is formed in correspondence to the junction: it is formed by the lattice atoms without the majority carriers, that went to the opposite side of the junction. In the p side the depletion region is negatively charged, while in the n side it is positively charged. The resulting electric field opposes the diffusion currents, until an equilibrium is reached. The resulting junction potential V_j has typical values between 300 mV and 800 mV.
If the n and p dopings have the same concentration, the depletion region will be symmetric. Otherwise, if for instance the n region has a higher concentration of dopants, the electrons will need to travel further away in the p region before recombining: in this case the depletion region is more extended in the p side.
The depletion region is suitable for being used as a particle detector: due

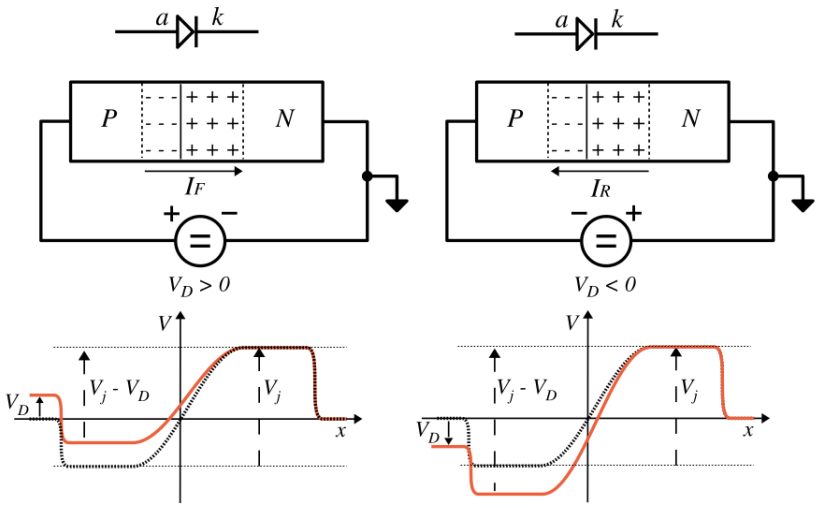


Figure 1.3: Direct bias (left) and reverse bias (right) for a p-n junction [2]. The electric field is represented in red.

to the electric field, the electron-hole pairs generated in this region drift to the n side (electrons) or to the p side (holes), where they can be respectively collected. Only the lattice fixed charges are present in the depletion region (that do not contribute to the conductivity), therefore the resistivity is high (with respect to undepleted p and n regions).
A diode is formed by a p-n junction in which both sides are connected to an ohmic contact. The p side contact is referred to as anode a while the n side contact as cathode k. Figure 1.2 shows the electric potential through the junction. By applying a potential V_D to the diode, two possibilities arise, as shown in Figure 1.3. By applying a positive V_D to the anode, the diode is forward biased: the potential barrier between the p and n regions is forced to a lower value, and the majority carriers in each region start diffusing to the other one. The resulting current I_F exponentially increases with V_D.
By applying a negative V_D the diode is reverse biased: the potential barrier is forced to a higher value and the majority carriers from each region further move away from the junction, increasing the depletion region. The reverse bias configuration is ideal for detecting particles, due the enlargement of the depletion region. The minority carriers are responsible for the leakage current I_R, usually $10^6 - 10^8$ times smaller than the forward bias current. Moreover, I_R slowly increases with $|V_D|$ due to the enlargement of the depletion region. By increasing $|V_D|$, the breakdown voltage V_{BD} is eventually reached: the leakage current abruptly increases, potentially destroying the junction for thermal effects. Two different mechanisms are responsible for this behaviour.

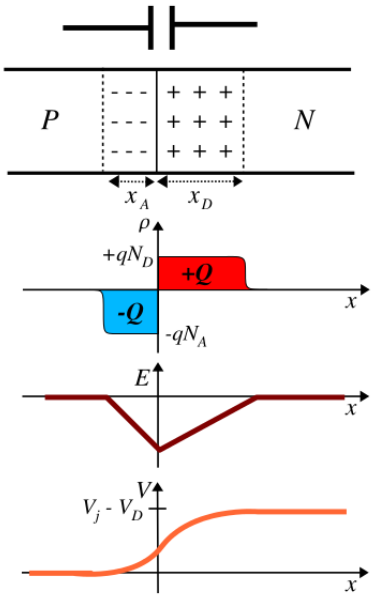


Figure 1.4: From top to bottom the charge distribution, the electric field and the electric potential for a reversed biased p-n junction [2].

The first one is the avalanche carrier multiplication: the electric field is so strong that carriers are able to further generate electron-hole pairs in an avalanche process. The second one is the Zener effect: in this case the electric field directly causes the generation of electron-hole pairs.

1.3.1 Properties of the Reverse Bias

Some properties of the reverse biased p-n junction can be derived from the one-dimensional Poisson equation (the system is reported in Figure 1.4):

$$\frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2} = -\frac{\varrho(x)}{\epsilon} = \begin{cases} +\frac{eN_A}{\epsilon} & -x_A < x < 0 \\ -\frac{eN_D}{\epsilon} & 0 < x < x_D \end{cases} \tag{1.14}$$

where the origin of the reference frame is located in correspondence to the junction. Moreover, x_A and x_D are the extensions of the depletion regions in the p and n sides respectively, $\varrho(x)$ is the charge density and ϵ is the dielectric constant of the medium. The electric field is obtained by integrating Equation 1.14 and applying the boundary conditions. The electric field must vanish at both edges of the charge distribution:

$$-\frac{\mathrm{d}\varphi}{\mathrm{d}x} = -\begin{cases} +\frac{eN_A}{\epsilon}(x + x_A) & -x_A < x < 0 \\ -\frac{eN_D}{\epsilon}(x - x_D) & 0 < x < x_D \end{cases} \tag{1.15}$$

The electric potential is obtained by integrating again:

$$\varphi(x) = \begin{cases} +\frac{eN_A}{2\epsilon}(x + x_A)^2 & -x_A < x < 0 \\ -\frac{eN_D}{2\epsilon}(x - x_D)^2 - V_D & 0 < x < x_D \end{cases} \tag{1.16}$$

where the boundary conditions are $\varphi(-x_A) = 0$ and $\varphi(x_D) = -V_D$ (neglecting the small V_j contribution). Since the solutions for either side of the junction must match at $x = 0$ and $x_A N_A = x_D N_D$ (the junction is overall neutral):

$$(x_A + x_D)x_D = \frac{2\epsilon|V_D|}{eN_D} \tag{1.17}$$

For instance, if $N_A \gg N_D$ then $x_D \gg x_A$, therefore the extension of the depletion region can be written as:

$$\ell = \sqrt{\frac{2\epsilon|V_D|}{eN}} = \sqrt{2\epsilon\mu\rho|V_D|} \tag{1.18}$$

where N, ρ and μ are referred to the junction side (p or n) with the lower doping.
The bias voltage needed to fully deplete a junction of thickness d is given by:

$$|V_{dep}| = \frac{eN}{2\epsilon}d^2 \tag{1.19}$$

Due to fixed charges on both sides of the junction, the depletion region exhibits some properties of a charged capacitor. The capacitance per unit area is given by:

$$C = \frac{\epsilon}{\ell} = \sqrt{\frac{e\epsilon N}{2|V_D|}} \tag{1.20}$$

The capacitance decreases with increasing $|V_D|$, since the depletion regions grows thicker.
Semiconductor detectors usually operate in full depletion, in order to maximise the active volume and minimise the junction capacitance. The active volume (the substrate) of a detector is made of high purity silicon (ν for a very low doping of donors or π for a very low doping of acceptors), so that the the full depletion is achieved at low bias voltages. The junction is created by depositing highly doped silicon of the opposite type (p^+ or n^+) on the substrate. This is often called the rectifying contact. Because of its high doping level, it also serves as an excellent blocking contact in which the minority carrier concentration is very low. In the nearly pure substrate, however, the minority carriers are not highly suppressed and an additional

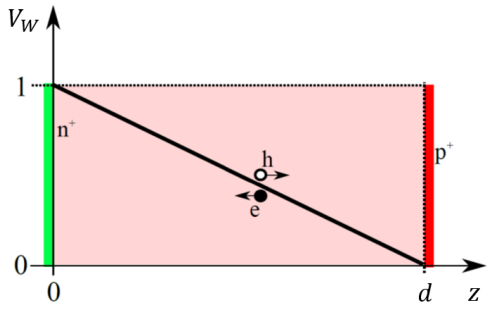


Figure 1.5: Weighting field as a function of the distance from the electrode z for a detector with electrode dimensions much larger than the detector thickness.

blocking contact is normally provided at the opposite side. If the high purity silicon is of type ν, a thin n^+ layer is applied to the backside. A metallic layer is deposited on both sides, forming the ohmic contacts.

1.3.2 Signal Formation

When a charged particle traverses a silicon detector, the generated electrons and holes start drifting to the electrodes (n^+ and p^+ respectively). The charge induced in the the electrodes can be calculated with the Shockley–Ramo theorem. The theorem states that the instantaneous current induced to an electrode by a particle of charge q (electron or hole) and drift velocity $\boldsymbol{v}$ is:

$$i = q\boldsymbol{v} \cdot \boldsymbol{E}_w \tag{1.21}$$

where $\boldsymbol{E}_w$ is the weighting field (obtained by placing the collecting electrode to a unitary potential and the other electrode to ground). The charge induced in the time interval (t_1, t_2) (during which the charge carrier moved from position $\boldsymbol{x}_1$ to position $\boldsymbol{x}_2$) is found by integrating Equation 1.21:

$$Q = \int_{t_1}^{t_2} i(t)\mathrm{d}t = q(V_w(\boldsymbol{x}_1) - V_w(\boldsymbol{x}_2)) = q\Delta V_w \tag{1.22}$$

If the electrode dimensions are larger than the detector thickness, the weighting potential is a linear function of the distance from the electrode, and goes from one to zero (the weighting field is constant), as shown in Figure 1.5. When all charge carriers reached the respective electrodes, the total charge collected by the electrode is:

$$Q_{tot} = -n_q e\left(\frac{z}{d} - 1\right) + n_q e\left(\frac{z}{d} - 0\right) = n_q e \tag{1.23}$$

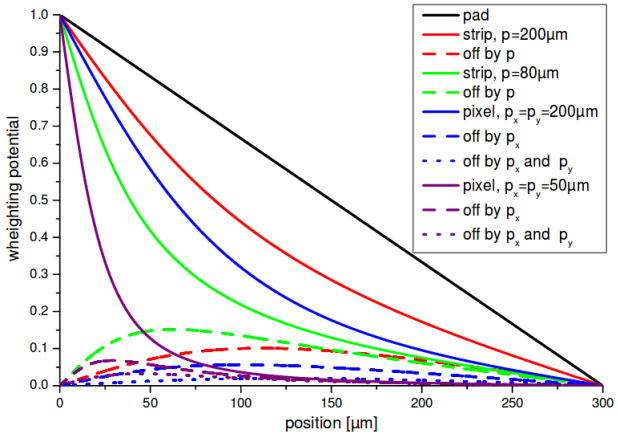


Figure 1.6: Weighting potential for different geometries of pixel and strip detectors with 300 µm thickness: p_x and p_y are the pitches of the pixels in x and y directions while p is the pitch of the strips. The continuous lines indicate the potential evaluated at the center of the electrode while the dashed lines indicate the potential evaluated between the electrodes [3].

where n_q is the number of generated electron-hole pairs, d the substrate thickness and z the point where an electron-hole pair is generated.
In the case of pixel detectors this approximation does not hold, since the electrode dimensions are lower than the detector thickness: in this case the weighting potential rapidly goes to zero by moving away from the electrode, as shown in Figure 1.6. Therefore, the majority of the charge is induced in proximity of the electrode, and the contribution of the charge carrier moving to the opposite electrode is negligible. In any case, when all charges reach the respective electrodes, the collected charge is still $n_q e$.

1.4 Radiation Damage

Silicon sensors employed to detect particles can get damaged after a long exposure to the source of radiation. The bulk damage produced in silicon by hadrons or high energy leptons is caused by displacing a Primary Knock on Atom (PKA) out of its lattice site. As a result of this displacement a vacancy is left in the crystal lattice and the recoiling atom moves to an interstitial lattice position: this is known as Frenkel pair. The PKA can only be displaced if the imparted energy is higher than about 25 eV. These are point defects. However, if the transferred energy is higher than about 2 keV, the PKA can displace other atoms, forming a dense agglomeration

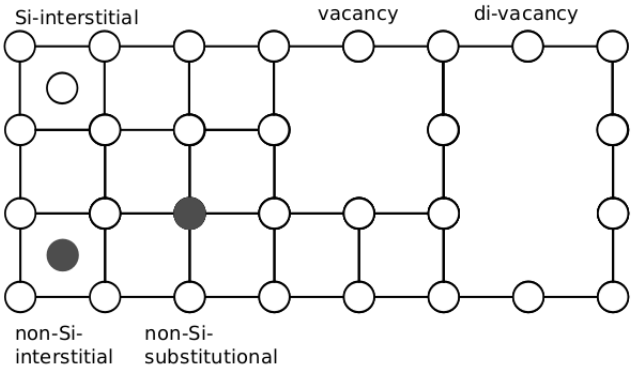


Figure 1.7: Representation of radiation damage defects of the silicon lattice [4].

of defects (or cluster). Figure 1.7 shows some of the possible defects of the silicon lattice.
The defects have a high mobility, therefore they diffuse in a process known as annealing, which is divided into the beneficial annealing and the reverse annealing. In the first case the Frenkel pairs recombine, thus reducing the damage effects. In the second case the defects, moving in the lattice, can cause further displacements. Both processes, being diffusive, strongly depend on temperature.
Charged hadrons interact with silicon primarily due to Coulomb interaction and ionization of lattice atoms whereas neutrons interact only with the silicon nuclei. In principle, is not not trivial to describe the radiation damage caused by different particles and different processes. In the Non Ionizing Energy Loss (NIEL) hypothesis, the atom displacement damage scales linearly with the energy released by the collision, regardless of the particle or the interaction type. The displacement damage can be described by the cross section $D(E)$ [5]:

$$D(E) = \sum_i \sigma_i(E) \int_0^{E_R^{max}} f_i(E, E_R) P(E_R) \mathrm{d}E_R \tag{1.24}$$

The index i indicates all possible interactions between the incoming particle of energy E and the silicon atoms in the crystal; σ_i is the cross section of process i; $f_i(E_{kin}, E_R)$ is the probability for the generation of a PKA with recoil energy E_R by a particle with energy E undergoing an interaction i; E_R^{max} is the maximum recoil energy. Finally, $P(E_R)$ is the Lindhard partition function, that describes the portion of the recoil energy deposited in form of displacement damage. Figure 1.8 shows the displacement damage cross sections for neutrons, protons, pions and electrons up to an energy of 10 GeV.

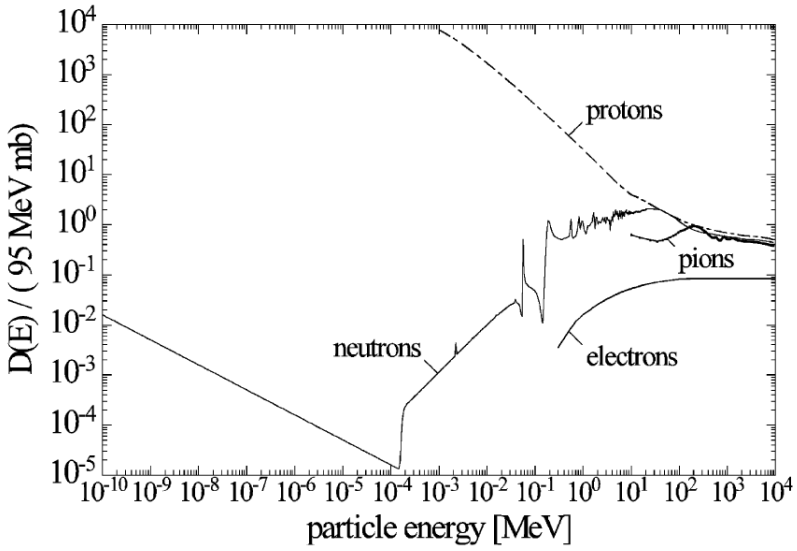


Figure 1.8: Estimation of the displacement damage cross sections for neutrons, protons, pions and electrons up to an energy of 10 GeV. [6].

The hardness factor k is used for comparing radiation damages from a source with energy spectrum per unit area $\phi(E)$ in terms of the radiation damage from 1 MeV neutrons (neutron equivalents n_{eq}):

$$\kappa = \frac{\int D(E)\phi(E)\mathrm{d}E}{D(E_n = 1\,\mathrm{MeV}) \int \phi(E)\mathrm{d}E} = \frac{\Phi_{eq}}{\Phi} \tag{1.25}$$

where Φ are the fluences (Φ_{eq} is the fluence expressed in $\mathrm{n_{eq}/cm^2}$) and $D(E_n = 1\,\mathrm{MeV}) = 95\,\mathrm{MeVmb}$.
Irradiated silicon sensors are kept at low temperatures (below 0 °C) in order to minimise the macroscopic effects of the radiation damage, described in the following. However, it can be useful to expose to room (or higher temperature) aiming for the beneficial annealing. The macroscopic effects of radiation damage can be summarised as follows:

- **Leakage Current:** an irradiated silicon sensor shows a higher leakage current, since the lattice defects add additional energy levels in the bandgap. The current variation ΔI in the silicon active volume V is proportional to the irradiation fluence:

$$\frac{\Delta I}{V} = \alpha\Phi \tag{1.26}$$

  where α is the current-related damage rate, which is independent of the initial resistivity of the silicon, as shown in Figure 1.9. The leakage

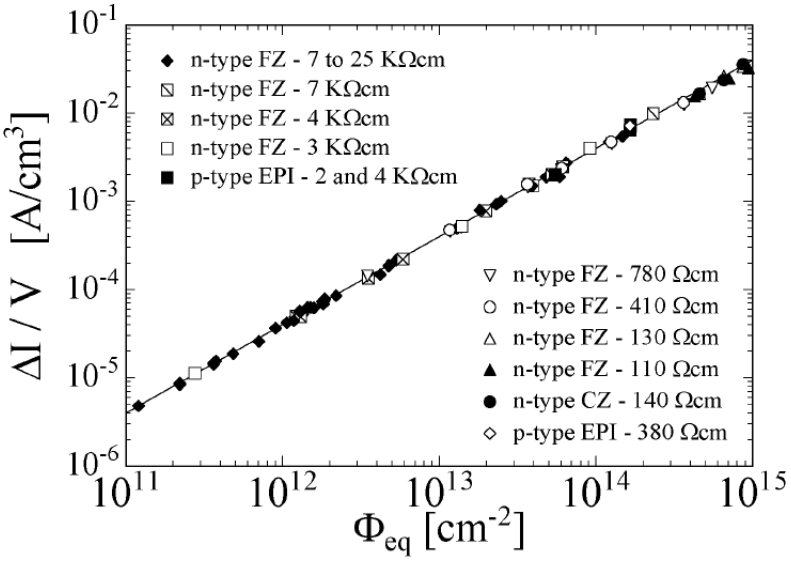


Figure 1.9: Leakage current as a function of the received fluence, for different types of silicon sensors [5].

current strongly depends on temperature:

$$I \propto T^2 e^{-E_g/2k_B T} \tag{1.27}$$

For this reason irradiated sensors are always operated at low temperatures, in order to maintain the leakage current to a reasonable level for experimental necessities.

- **Doping:** after irradiating a silicon sensor, a donor removal is observed while the bulk defects behave as acceptors. The effective doping variation ($N_{eff} = N_D - N_A$) can be expressed by:

$$N_{eff} = N_D e^{-c\Phi} - N_A - b\Phi \tag{1.28}$$

where c is the donors removal constant and b the acceptor creation rate. If the substrate is of type n, by increasing the fluence the acceptors can compensate the donors, leading to a bulk type inversion. If this happens, the junction moves from the p^+ side to the n^+ side. Due to the doping level variation, the bias voltage needed to deplete the sensor decreases with increasing fluence, reaches a minimum, and increases again after inversion, as shown in Figure 1.10. The inversion point depends on the initial doping: if the initial donor level is low (the resistivity is high), a lower fluence in necessary for inversion with respect to a substrate with a lower resistivity.

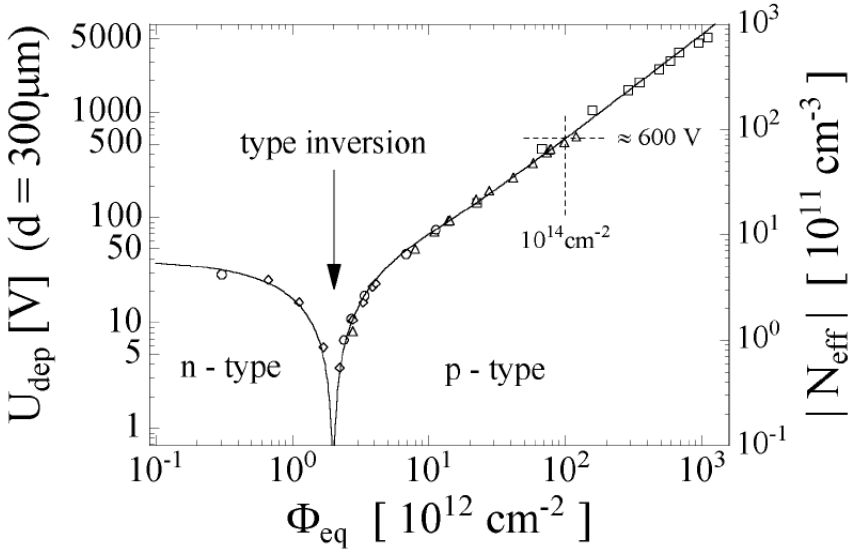


Figure 1.10: Depletion voltage as a function of the received fluence,for different types of silicon sensors [5].

- **Traps:** the energy levels in the bandgap can trap the charge carriers produced by the a traversing particle. If the charge trapping time is higher than the collection time, their contribution to the signal is lost. The Charge Collection Efficiency (CCE) is defined as the ratio of the collected charge and the generated charge. When an electron-hole pair is generated, the carriers move toward their respective electrodes: the sum of the drift distances is the active thickness d of the sensor. The effective distance after irradiation, referred to as Charge Collection Distance (CCD), is defined by:

$$d_C = \mu \tau E \tag{1.29}$$

where E is the electric field, μ the sums of electrons and holes mobilities and τ is given by:

$$\frac{1}{\tau} = \frac{1}{\tau_e} + \frac{1}{\tau_h} \tag{1.30}$$

where τ_e and τ_h are the average trapping times for electrons and holes respectively. These parameters are fixed by the substrate characteristics and by the irradiation. The collected charge is given by:

$$Q_{coll} = Q_0 \frac{d_C}{d} \tag{1.31}$$

where Q_0 is the charge generated by the interaction of a traversing particle. The CCE is therefore Q_{coll}/Q_0. By defining the generated charge per unit length $q_p = Q_0/d$, the CCD can be written as:

$$d_C = \frac{Q_{coll}}{q_p} \tag{1.32}$$

The trapping effect can be mitigated by increasing the electric field, so that d_C increases as well. For this reason irradiated sensors are always operated at a bias voltage higher than the depletion voltage (over-depletion). The CCE increases with the bias voltage, until the breakdown is reached.

The radiation damages described so far are relative to the bulk of the sensor, but the sensor surface is damaged as well. The silicon sensor is usually covered by a thin dielectric layer. Since the lattice structure is very irregular, it is not affected by displacements. However, the dielectric ionisation is more problematic, as it is not fully reversible: a positive charge is formed in the oxide, that modifies the electric field in the substrate. Various strategies are adopted to mitigate this effect, as explained in Section 1.5.1.

1.5 Pixel Sensors

In order to obtain a two-dimensional measurement of the impact point of a particle, the electrodes need to be segmented into two directions. If the pitch of the electrodes is lower than about 500 µm for both directions, they are usually referred to as pixels, and the sensor is usually referred to as pixel sensor.
The construction procedure starts from a high purity (ν or π) silicon wafer [7]. After cleaning the surface, a thin oxide layer is created for protection and insulation. Usually SiO_2 is used, since it is stable and chemically inert, while also being a good dielectric. Windows are subsequently opened in the oxide layer, with a photo-lithographic technique, in order to access the underlying silicon for doping. To perform this operation a photoresist is applied on the wafer surface, then illuminated by UV light with the pattern of the windows to be opened. The light chemically alters the photoresist, so that it can be removed only in the illuminated areas. The selected oxide areas are removed with acid etching. Finally, after the etching, the remaining photoresist is removed.
The following step is the silicon doping, performed with ion implantation: acceptor or donor atoms are implanted using an accelerator, for p^+ or n^+

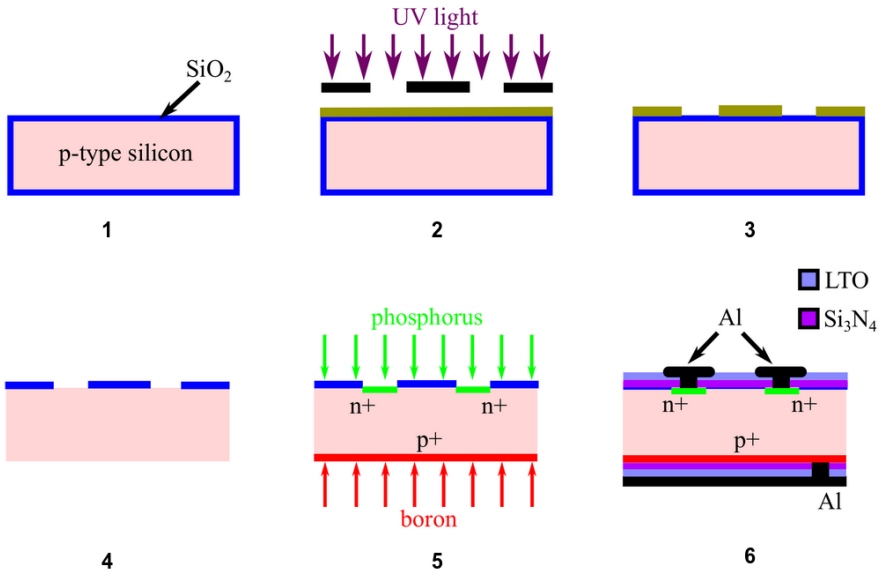


Figure 1.11: Fabrication steps of a silicon pixel sensor.

implants respectively. The doping process can damage the silicon, therefore the wafer usually undergoes an annealing procedure at high temperatures ($800 - 1000$ °C). Afterwards, aluminium is evaporated and patterned by photo-lithography to provide thin ohmic electrical contacts at the front and rear surfaces. Moreover, a passivation layer (usually polyimide or low temperature SiO_2) is applied to isolate and protect the sensor surface. Openings of about 12 µm diameter are made in the passivation layer above the electrodes: they are referred to as bump (bonding) pads. The production steps are summarised in Figure 1.11.
After fabrication, the wafer is cut, and every pixel sensor is connected to a ReadOut Chip (ROC) with a compatible pixel pitch, with the bump bonding technique. The ROC collects, amplifies and digitalises the charge produced by a particle interacting with the sensor. First, an additional metallization is applied on the bump pads (called Under Bump Metallisation, UBM) and on the ROC. Afterwards, a metallic layer (usually indium or a tin-silver alloy) gets deposited on the bump pads with a photo-lithographic technique. By heating the sensor, the metal takes a spherical shape, thus forming the so-called bumps above all pixels. The ROC is flipped and applied on the sensor: the assembly is pressed and heated in order to form the connections. This procedure is also known as flip-chipping. The sensor plus ROC assembly is referred to as pixel detector in the following.

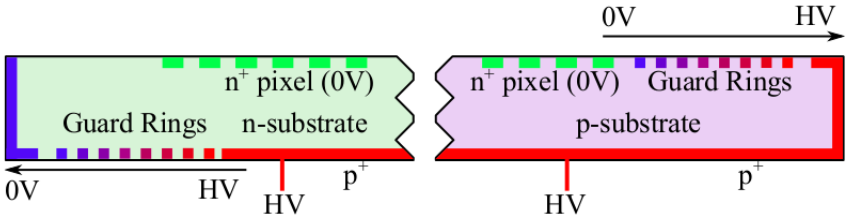


Figure 1.12: Disposition of guard rings in the n-in-n (left) e n-in-p (right) technologies [4].

1.5.1 Technologies

Various combinations are available for the construction of pixel sensors, for instance the doping of the substrate, ν or π, and the doping of the collecting electrodes (i.e. the pixels), p^+ or n^+. When using p^+ implants, the signal is mainly generated by the drift of holes. Since the mobility of holes is three times lower with respect to electrons, sensors using p^+ implants are more influenced by radiation damage after receiving high fluences. For this reason, in high energy physics n^+ implants are used as collecting electrodes, since the radiation damage is often significant. The substrate can be either be ν or π: in the first case sensors are referred to as n-in-n while in the second case they are referred to as n-in-p sensors.
In the case of n-in-n sensors, the junction is on the backside (that is the opposite face with respect to the pixels, with a uniform p^+ doping). Additional production steps are necessary in order to protect the p^+ implant, which receives the bias voltage. Moreover, n-in-n sensors can only be used if fully depleted, since the collecting electrodes are placed to the opposite side of the sensor with respect to the junction. Finally, the weighting field is maximum near the collecting electrode, but the electric field is minimum. For this reason these sensors are usually operated in over-depletion, in order to increase the electric field in this region. However, since the substrate is of type n, it undergoes an inversion after receiving high fluences. The junction moves to the pixel side and the sensors can be used even if not fully-depleted.
In the case of n-in-p sensors, the junction is on the same side of the pixels, therefore only one side of the wafer is processed, making this technology cheaper. Moreover, the substrate does not undergo inversion. In both types of sensors, to reverse bias the junction a negative voltage needs to be applied to the backside. The pixels are at zero potential through the ROC.
All the pixel sensors presented in this Thesis are n-in-p. Additional structures, presented in the following, are often used [4]:

- **Guard Rings:** the edges of a pixel sensor are conductive due to damages from the wafer cutting procedure. Therefore, the bias voltage applied to the backside of the sensor reaches the sensor edges. Since the pixels are kept at zero potential though the ROC, additional structures are necessary to control the electric field at the borders of the active region. Moreover, if the depletion region reaches the edges, the resulting current would destroy the sensor.
  For these reasons the so-called guard rings are employed, that is ring implantations placed between the active region (i.e. the pixels) and the sensor edges. Guard rings will be at an intermediate potential between zero and the bias voltage, acting as control structures. Figure 1.12 shows the disposition of guard rings in n-in-n and n-in-p technologies: in the first case they are on the backside, while in the second case they are on the pixel side. The number and shape of guard rings are often varied in the development of new pixel sensors, in order to find an optimal solution. With the n-in-p technology, additional protections are necessary in order to avoid sparks between the sensors edges and the ROC.

- **Polarisation Structures:** in applications where a large number of sensors needs to be produced (for example the construction of a tracker) it is important to have a high production yield. For this reason, electrical tests on pixel sensors needs to be performed of wafer, before cutting and bump bonding procedures. For instance, it is important to verify that the breakdown voltage is sufficiently high. Additional polarisation structures are added for this purpose.
  The Punch-Through (PT) is a common solution: an additional n^+ implant, the bias dot, is placed at the intersection of four pixels. All the bias dots are connected, with a metal route, to a common ring outside the active region (but before the guard rings). This ring is referred to as bias ring. When sensors are tested on wafer, before the bump bonding, the bias ring is kept at zero potential: through the bias dots, all pixels are kept at zero potential as well. By applying the bias voltage to the backside of the wafer, the sensors can be characterised before cutting.
  After cutting the wafer and bump bonding the sensors to the ROCs, the pixels will be at zero potential though the ROC: also the bias dots and bias ring will be at the same potential. However, since the bias dots are not connected to a readout channel, charge produced near these implants is lost. This leads to a hit detection efficiency loss, which is more severe when radiation damage is present.

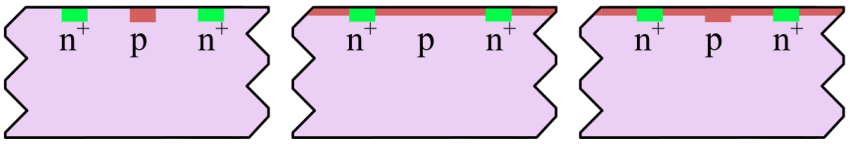


Figure 1.13: Different types of isolation structures: p-stop (left), p-spray (center), moderated p-spray (right) [4].

An alternative method to the PT structure is the temporary metal solution: all the pixels are connected together with a metallic layer so that they can be kept at zero potential. The metal layer is only used when testing the sensors on the wafer: it can subsequently be removed, with a metal etching procedure. This method has no disadvantages in terms of sensor performance, but it is more expensive.

- **Isolation Structures:** when a pixel sensor surface is damaged by radiation, due to the positive charges in the oxide layer, negative mirror charges accumulate near the silicon surface, leading to a short of neighbour pixels. Three different solutions are used to mitigate this problem, as illustrated in Figure 1.13.
  The p-stop solution consists in a p implantation between the pixels: the implant is not connected to any reference potential. This solution requires an additional photo-lithographic step in the production.
  The p-spray solution consists in a p implantation on the whole sensor surface. The dose is chosen to avoid high electric fields close to the n^+ implants. This solution does not require and additional photo-lithographic step, and it is more suitable for pixels of small pitch.
  The third solution, the moderated p-spray, combines the previous ones: an uniform p-spray is deposited on the sensor with a higher dose between the pixels.

A novel technique in silicon manufacture is the Silicon-Silicon (Si-Si) Direct Wafer Bonding (DWB). The DWB allows to produce silicon wafers with a thin high resistivity π layer (the substrate) bonded through molecular interactions to a low resistivity p layer. The latter has two functions: it is the mechanical support for the sensor and the blocking contact of the junction. This technique allows the production of sensors with a small active thickness ($100 - 150$ µm), while keeping the wafer robust enough for production. Sensors with a small active thickness are less influenced by radiation damage, because the drift patch of the charge is reduced. Moreover, the bias voltage needed to deplete them is low, as is the leakage current. On the other hand,

the charge produced by a MIP traversing a thin sensor is low, therefore the ROCs need to have low thresholds (below 2000 electrons).
The low resistivity wafer (also referred to as handle wafer) can be thinned after production. After thinning, a metal layer is deposited to ease the sensor bias.

1.5.2 3D Pixel Sensors

The sensors described so far are usually referred to as planar pixel sensors, because the implants are superficial. Another type of pixel sensors, usually referred to as 3D pixel sensors, was recently conceived. While in planar sensors the electrodes are thin superficial implants on the surface of the substrate, in 3D sensors the electrodes are cylindrical and penetrate deeply into the substrate. This configuration allows for several advantages.
In first place, the electric field is parallel to the sensor surface and it extends between the electrodes. The drift of the charge carriers (generated by the passage of a particle) is parallel to the sensor surface. The distance between electrodes is smaller than the active thickness, therefore the bias voltage needed to deplete a 3D sensor is lower with respect to a planar sensor, as is the drift time. However, the collected charge only depends on the active thickness, which is the same for 3D and planar pixel sensors. For these reasons, 3D sensors can have good performance even with high radiation damages. However, the production processes are more complicated, with consequences on costs and production yield.
The columnar electrodes are usually created with the Deep Reactive Ion Etching (DRIE) process. Two steps are iterated: an ion plasm etches the wafer with vertical direction and a passivation layer is deposited to protect the borders of the holes. The holes diameter slightly decreases while penetrating into the silicon: the shape of the columns is more properly a truncated cone. After the dopants have been deposited, the holes are filled with polycrystalline silicon with the Low Pressure Chemical Vapour Deposition (LPCVD) technique. The silicon interacts with the doping molecules (usually boron and phosphorus oxides), producing acceptors and donor atoms, that are diffused to the surrounding monocrystalline silicon by annealing.
The first 3D pixel sensors were produced with a double sided process [8]: the n^+ columns are excavated from one side of the wafer and p^+ columns from the other side. However this process has several disadvantages. For instance the limited accuracy (a few µm) of the alignment between the front and the backside of the wafers, that is critical for small pitch pixels. Moreover, the wafer is extremely fragile, due to the presence of holes on the whole volume. In the past years, the DWB technique allowed for a single sided process: p^+

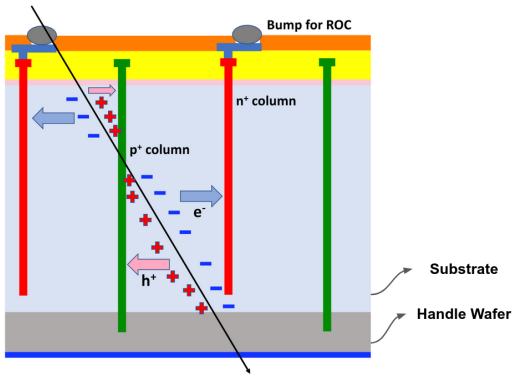


Figure 1.14: Schematic view of a 3D pixel sensor fabricated with the DWB double-sided technique. The drift path of the charge carriers is also highlighted.

columns are etched deep enough to reach the low resistivity handle wafer, so that a good ohmic contact is achieved on the sensor backside. The n^+ columns are stopped a short distance from the handle wafer (about 20 µm), in order to prevent from an early breakdown. Moreover, the handle wafer assures robustness before and after production. Figure 1.14 shows a 3D sensor produced with the DWB double sided technique.
The pitch of 3D pixel is determined by p^+ column, while one n^+ column is usually placed in the middle of the pixel cell. Aluminium is deposited on top of n^+ columns (the collecting electrodes).

Chapter 2

LHC and the CMS Experiment

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest circular hadron collider ever built [9], with a length of about 27 km. It is installed underground (at a depth varying between 50 m to 175 m) at the *Conseil Européen pour la Recherche Nucléaire* (CERN) near Geneva. Its main goal is to study high-energy proton-proton collisions[1].
The LHC is the last stage of a complex accelerator chain. Protons are obtained through ionization of Hydrogen, and are injected in the first linear collider, LINAC2, where they are accelerated to an energy of 50 MeV. Protons are then injected in three subsequent synchrotrons: the Proton Synchrotron Booster (PSB), which accelerates them to an energy of 1.4 GeV, the Proton Synchrotron (PS) which accelerates them to an energy of 25 GeV and groups them into bunches with a frequency of 40 MHz and finally the Super Proton Synchrotron (SPS), which accelerates the protons to an energy of 450 GeV. Protons are subsequently injected into the LHC in opposite directions in two parallel pipes and accelerated[2] up to 7 TeV. Inside the pipes an extreme vacuum condition is kept (about 10^{-10} Torr), in order to avoid spurious interactions between the beam and the gas remnants.
Proton beams collide at four interaction points, in which the four LHC main experiments are placed: Compact Muon Solenoid (CMS) [10], A Toroidal LHC ApparatuS (ATLAS) [11], LHC beauty (LHCb) [12], A Large Ion Collider Experiment (ALICE) [13]. More in details:

[1] Heavy ion collisions are also studied with LHC.

[2] A radio frequency acceleration system, consisting of 16 superconducting radio-frequency resonant cavities, is used to increase the proton energy by 0.5 MeV with each beam revolution.

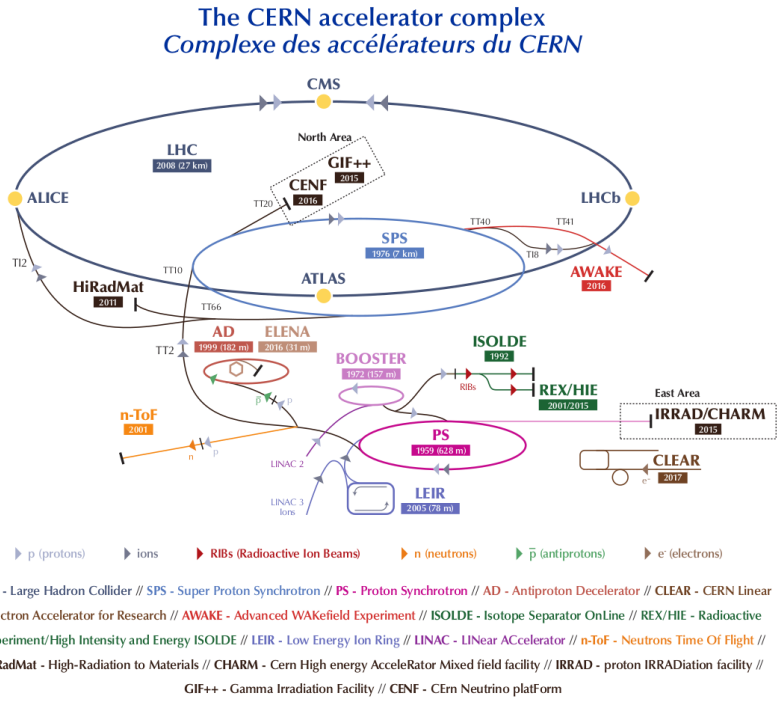


Figure 2.1: The LHC accelerator complex.

- ALICE studies heavy ion collisions and the production of the quark-gluon plasma, a new state of matter.

- ATLAS and CMS are two general purpose experiments designed to investigate a large variety of physics processes. In particular, they were conceived to find the Higgs boson and to look for evidence of new physics. The two experiments are independent, in order to cross-check any new discovery.

- LHCb is mainly designed to study CP violation (combined charge conjugation and parity symmetries) in electroweak processes and to study asymmetries between matter and antimatter through the analysis of the decay of hadrons containing b quarks.

Two other smaller experiments are also present: TOTEM (TOTal Elastic and diffractive cross section Measurement) [14] and LHCf (LHC forward) [15]. TOTEM measures proton-proton interaction cross section and monitors the luminosity of the LHC using detectors positioned at each side of the CMS experiment. LHCf is made by two detectors along the LHC beamline, at 140 m from each side of the ATLAS collision point. Its goal is to simulate

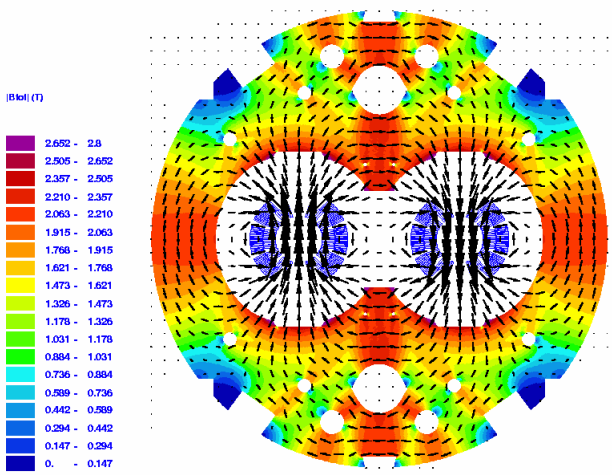


Figure 2.2: Cross section of an LHC dipole magnet with the magnetic field lines. The magnetic field is opposite in the two beam pipes.

the interaction of very high energy (between 10^{17} TeV and 10^{20} TeV) cosmic rays with the atmosphere, by using neutral particles thrown in the forward direction by LHC collisions. In Figure 2.1 a schematic view of the LHC accelerator complex and the main experiments is reported.
The LHC ring is divided into octancts, with eight curve regions, referred to as arcs, separated by rectilinear regions. In the arcs 1232 magnetic dipoles bend the beams, according to the relation [1]:

$$p[\text{GeV}] = 0.3 \cdot B[\text{T}] \cdot R[\text{m}] \tag{2.1}$$

where p is the particle momentum, R the radius of curvature and B the magnetic field. To keep 7 TeV beams within the LHC circumference a magnetic field of about 8 T is required. This is achieved by using superconductive magnets, which need to be cooled at about 2 K. In Figure 2.2 a cross section of an LHC dipole magnet is reported. In the straight LHC regions quadrupoles, sestupoles and octupoles are used for focusing the beams and to make corrections to the beam direction.
Each proton beam consists of 2808 bunches, each containing about 10^{11} protons. The beam transverse length is about 15 µm. The revolution frequency of the bunches is 11 kHz, while the bunch crossing frequency is 40 MHz (e.g. the time interval between two consecutive bunch collisions is 25 ns). The maximum achievable energy in the center of mass frame of reference with the LHC is $\sqrt{s} = 14$ TeV.
An important parameter in accelerators is the instantaneous luminosity $\mathcal{L}_{inst}$,

defined by:

$$\mathcal{L}_{inst} = f n_b \frac{n_1 n_2}{4\pi\sigma_x\sigma_y} [\mathrm{cm}^{-2}\mathrm{s}^{-1}] \tag{2.2}$$

where n_1 and n_2 are the number of particles in the bunches of each beam, n_b is the number of bunches per beam, f is the revolution frequency, σ_x and σ_y are the transversal dimensions of the beams. The LHC was designed to reach an instantaneous luminosity $\mathcal{L}_{inst} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The rate of events R is given by:

$$R = \mathcal{L}_{inst}\sigma \tag{2.3}$$

where σ is the cross section of the given event. Finally, the integrated luminosity[3] $\mathcal{L}$ is obtained by integrating the instantaneous luminosity over the the acquisition time:

$$\mathcal{L} = \int \mathcal{L}_{inst} dt \tag{2.4}$$

Natural units $\hbar = c = 1$ are used throughout this Thesis.

2.2 The CMS Experiment

The CMS experiment is optimised for proton-proton collisions in order to study a broad spectrum of physics processes, such as the properties of the Higgs boson and the research of new particles. CMS is made by several detectors that cover a solid angle of almost 4π and can distinguish between the several types of particles that are produced in the final states of the collisions [10].
The cross section of proton-proton collisions at $\sqrt{s} = 14$ TeV is about 100 mb. Therefore, using Equation (2.3), CMS observes 10^9 events per second, at the design instantaneous luminosity. For each bunch crossing the average number of proton-proton interactions is ~ 40. The main goal is the identification and the reconstruction of a single collision of interest, while the other collisions are background events due to processes occurring with high probability. These additional collisions are known as pile-up[4]. In order to compensate the pile-up the CMS detectors feature high granularity and fast readout. Moreover, all the components of the detectors (the active volume and the readout electronics) are radiation resistant, especially near the collision point, due to the higher track density.
The CMS experiment uses a right-handed reference system: the x axis is

[3] Measured in inverse femtobarn, fb^{-1}.

[4] This is the in-time pile-up. An out-of-time pile-up contribution is also present, due to collisions in previous bunch-crossings.

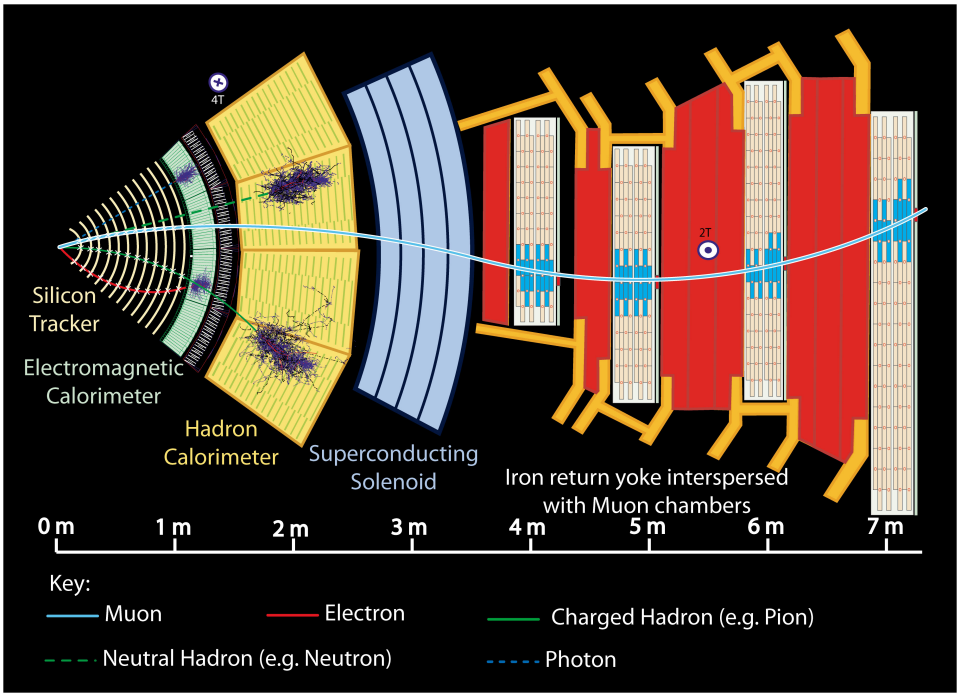


Figure 2.3: Schematic view of the CMS detectors, with the trajectories of the particles produced in the collisions.

oriented to the center of the LHC ring, the y axis upwards and the z axis is tangent to the beams counter-clockwise. Three coordinates are usually defined: the azimuthal angle ϕ (measured from the x axis in the $x-y$ plane), the polar angle θ (measured from the z axis) and the radial distance from the beam direction r.
The most commonly utilised observables in the CMS experiment are the transverse momentum p_T and the pseudorapidity η. The transverse momentum is the projection of the momentum on the $x-y$ plane:

$$p_T = \sqrt{p_x^2 + p_y^2} \tag{2.5}$$

The pseudorapidity[5] is given by:

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) \tag{2.6}$$

[5]The rapidity of a particle is given by $\mathsf{y} = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right)$. In the ultra-relativistic regimes, rapidity can be approximated by pseudorapidity: $\mathsf{y} \to \eta$

A common observable used to describe the angular separation between particles produced in the collisions is ΔR, defined as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{2.7}$$

where $\Delta\eta$ and $\Delta\phi$ are the separations in η and ϕ coordinates respectively. Another important variable used in CMS is the missing transverse energy E_T^{miss}, defined as the modulus of the missing transverse momentum $\boldsymbol{p}_T^{miss}$, which is the negative sum of transverse momenta of all reconstructed particles:

$$E_T^{miss} = |\boldsymbol{p}_T^{miss}| = \left| -\sum_i \boldsymbol{p}_{T,i} \right| \tag{2.8}$$

The transverse momentum of the initial state is null, therefore distributions of E_T^{miss} are used to study processes in which weakly interactive particles, such as neutrinos, escape from the detectors and are not reconstructed.
The CMS experiment features a cylindrical geometry coaxial with the beam pipe, and is made of a central part, the barrel, and two extremities, the endcaps. CMS has a length of 21.6 m, a diameter of 15 m and a total mass of 1.25×10^7 kg.
A super-conductive solenoidal electromagnet produces a magnetic field of about 3.8 T in the internal bore, which has a diameter of 6 m and a length of 12.5 m. The energy stored in the magnet is about 2.7 GJ at the full current (about 20 kA). The solenoid is made of four Niobium-Titanium layers[6]. Outside the solenoid is the return yoke, composed of three sections along the z axis: the field in the yoke is about 1.8 T.
The presence of the solenoid is essential for the detection of charged particles since they are deflected by the magnetic field. Indeed, it is possible to measure the transverse momentum of these particles by measuring the curvature of their trajectories through the tracking system,
Figure 2.3 shows a schematic view of the CMS detectors, with the trajectories of the particles produced in the collisions. In the following, a summary of the detectors is reported.

2.2.1 The Tracker

The tracker is designed to reconstruct the trajectories of charged particles originating from the interaction point with high resolution, and to identify the secondary vertices produced by particles with a short mean life time (like

[6]To reach the superconductivity phase, the solenoid is kept at a temperature of 4 K through a liquid Helium cooling system.

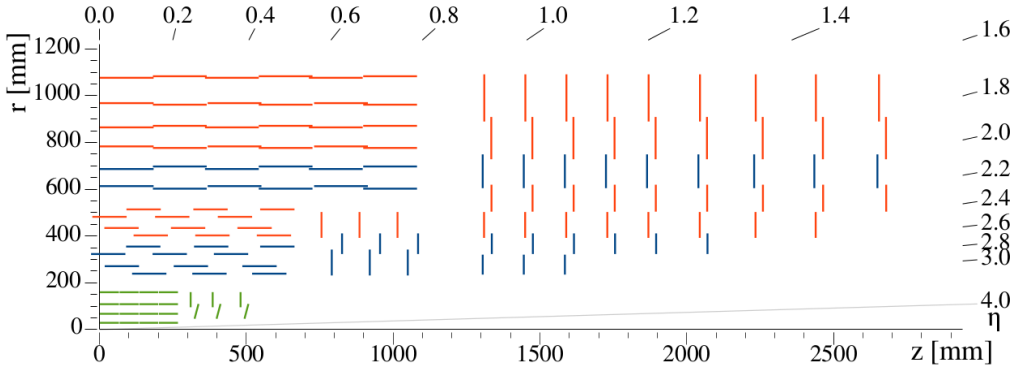


Figure 2.4: Schematic view of a quarter of the tracker in the $r - z$ plane. Pixel modules are shown in green, single strip modules in red and double strip modules in blue [20].

hadrons containing the b quark, that decays after travelling a few hundreds of µm). Charged hadrons are about two-thirds of the particles produced in proton-proton collisions, therefore a correct reconstruction is essential [16].
The scientific goals of CMS require good resolutions for transverse momenta up to 1 TeV (for searches of high mass resonances decaying in charged lepton pairs) and a high efficiency in the reconstruction of tracks with a low transverse momentum (about 100 MeV), in order to achieve an optimal jet[7] energy resolution. Moreover, it is necessary to have a high efficiency in track reconstruction, while keeping the number of ghost tracks[8] low, even in the typical high pile-up conditions of CMS. Finally, the system has to be fast, to send information to the trigger, and has to be radiation resistant.
The CMS tracker is completely made of silicon, with a decreasing granularity from the inside to the outside, and features a cylindrical symmetry with the axis aligned with the beam line [17, 18, 19]. It has a length of 5.8 m, a diameter of 2.5 m and covers the region $|\eta| < 2.5$. The tracker comprises of a large silicon strip detector with a small pixel detector inside it[9].
The pixel tracker is made of four cylindrical layers and is placed in the barrel, near the collision point. Outside the pixel tracker is the silicon trip tracker, made of ten cylindrical layers up to a radius of 110 cm. In the endcaps the pixel tracker is completed by three disks, while the strip tracker by twelve

[7]Collimated hadron clusters generated by hadronisation processes.

[8]Tracks are reconstructed from the coordinates of the hits on the detectors. A ghost track is a combination of non-correlated hits, arising from different particles, or a track of a particle badly reconstructed due to spurious hits.

[9]The first version of the CMS pixel tracker was used until 2016. In 2017 a new, improved, version was installed in CMS [19].

System	Layers	Pitch	Region
BPIX	4 cylinders	100×150 µm^2	$3.0\,\text{cm} < r < 16.0\,\text{cm}$
TIB	4 cylinders	$(80 - 120$ µm $)\times 20$ cm	$20\,\text{cm} < r < 55\,\text{cm}$
TOB	6 cylinders	$(122 - 183$ µm $)\times 25$ cm	$55\,\text{cm} < r < 116\,\text{cm}$
FPIX	3 disks	100×150 µm^2	$29.1\,\text{cm} < \lvert z\rvert < 51.6\,\text{cm}$
TID	3 disks	$(100 - 141$ µm $)\times 20$ cm	$65\,\text{cm} < \lvert z\rvert < 124\,\text{cm}$
TEC	9 disks	$(97 - 184$ µm $)\times 25$ cm	$124\,\text{cm} < \lvert z\rvert < 282\,\text{cm}$

Table 2.1: Summary of the characteristics of the tracker detectors.

disks. A schematic view of the tracker is shown in Figure 2.4.
The detectors are kept at -15 ℃ to reduce the leakage current. Aside the detectors, the tracker also hosts supporting materials, power supply cables and the cooling systems: these materials cause multiple scattering effects, that worsen the spatial resolution.
In Table 2.1 a summary of the detectors and their characteristics is reported.

The Pixel Tracker

The barrel layers of the pixel tracker are located at $r = 3.0\,\text{cm}$, $r = 6.8\,\text{cm}$, $r = 10.2\,\text{cm}$ and $r = 16.0\,\text{cm}$ while disks in the endcaps are located at $z = \pm 29.1$ cm, $z = \pm 39.6$ cm and $z = \pm 51.6$ cm. The region in the barrel is known as BPIX, while the region in the endcaps is known as FPIX.
The pixel tracker allows for a three-dimensional measurement of the position of a particle interacting with the silicon. In BPIX the z and $r\phi$ coordinates are measured with pixels, while r is given by the layer position. In FPIX, r and $r\phi$ coordinates are measured with pixels, while z is given by the disk position.
The silicon used in the detectors is 285 µm thick, and the pixel pitch is $100 \times 150\,\mu\text{m}^2$. The detectors are organised in modules made of 16 ROCs, each with 4160 channels (52 columns $\times$ 80 rows): each readout channel is connected to a single pixel of a common (n-in-n) planar pixel sensor.
There are 1856 modules corresponding to 124 millions pixels and 1.92 m^2 of silicon. The spatial resolutions for single hits varies from 10 µm to 40 µm, depending on the region and on the number of activated pixels.

The Strip Tracker

The strip tracker is made of 15148 modules with 9.3 millions strips, that cover and area of 198 m^2. It is made of four section. The Tracker Inner Barrel (TIB) and the Tracker Inner Disks cover the regions 20 cm $< r <$ 55 cm and 68 cm $< |z| <$ 124 cm respectively, and are made of four layers in the barrel and three disks in the endcaps.
The Tracker Outer Barrel (TOB) covers the region 55 cm $< r <$ 116 cm and is made of six layers. Finally, the Tracker EndCaps (TEC) cover the region 124 cm $< |z| <$ 282 cm. In each endcap nine disks are installed, each equipped with up to seven concentric rings of strip detectors.
The pitch of the sensors varies depending on the tracker region, and goes from 80 µm to 184 µm. Modules in TIB, TID and in the four inner rings of TEC are made of 320 µm thick silicon, while in TOB and in the three outer rings of TEC modules are made of 500 µm thick silicon.
Strips are oriented along the z axis in the barrel and along r in the endcaps. The modules located in the two inner layers of TIB and TOB, as well as the ring 1 and 2 of TID and the rings 1,2 and 5 of TEC feature a second strip detector on the back of the first, but rotated by an angle of 100 mrad. In this way a module in TIB, for instance, can combine the information of the two detectors to measure the position along z, thus performing a three-dimensional measurement.
The resolution of the strip tracker for single hits varies from 10 µm to 50 µm, depending on the region and on the number of activated strips.

2.2.2 The Electromagnetic Calorimeter

Outside the tracker is the Electromagnetic CALorimeter (ECAL), which identifies and measures the energy of electrons and photons [21, 22]. ECAL is a hermetic homogeneous calorimeter, made of scintillating crystals of lead tungstate ($PbWO_4$) with a truncated pyramidal shape. It covers the region $|\eta| < 3$. It is made of two parts, the ECAL Barrel (EB), which contains 61200 crystals, and the two ECAL Endcaps (EE), which contain 7324 crystals.
Aiming to contain the electromagnetic showers in a reduced space, $PbWO_4$ crystals were chosen due to their high density (8.3 g/cm^3), short radiation length[10] X_0 (0.89 cm) and small Molière radius[11] R_M (2.2 cm). Moreover,

[10] Characteristic of a material, related to the energy loss of high energy particles electromagnetically interacting with it.

[11] Characteristic of a material related to the transverse development of electromagnetic showers in a calorimeter. On average, 90% of the energy deposited by a shower is contained inside a cylinder with radius R_M.

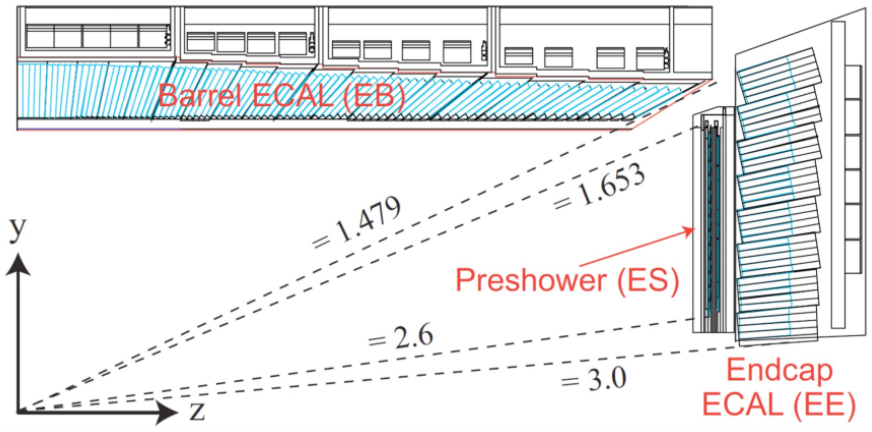


Figure 2.5: Schematic view of the CMS electromagnetic calorimeter.

the crystals are radiation hard and feature fast scintillation time (10 ns), which allows to collect 80% of the produced light between two consecutive bunch crossings. On the other hand, the light yield is low (about 10 photoelectrons/MeV) and strongly depends on temperature: for this reason the crystals are kept at a stabilised temperature (18 ℃).
The crystals are grouped into 5×5 matrices called towers. The EB has an inner radius of 129 cm, a length of 630 cm and covers the region $|\eta| < 1.479$. Each crystal in the EB has a length of 23 cm (corresponding to $25.8\,X_0$), a front face area of $22 \times 22\,\text{cm}^2$ and a rear face area of $26 \times 26\,\text{cm}^2$. The crystals are mounted tilted by 3° in ϕ and η with respect to the interaction point: in this way the empty spaces between the crystals are not aligned with the interaction point.
The two sections of EE cover the region $1.479 < |\eta| < 3$ and are formed by two semicircular aluminium halves called dees, in which crystals are arranged according to a $\eta - \phi$ symmetry. Each crystal in the EE has a length of 22 cm, a front face area of $28.6 \times 28.6\,\text{cm}^2$ and a rear face area of $30 \times 30\,\text{cm}^2$.
In the EB the scintillation light is readout using Avalanche PhotoDiodes (APDs), while in the EE Vacuum PhotoTriodes (VPTs) are used instead, due to their higher radiation resistance and the higher particles flux in this region.
Finally, a pre-shower system is also mounted in front of the endcaps, in order to separate the photons produced in the collision from those produced by forward-emitted π^0 mesons. This additional detector is a sampling calorimeter, made of two lead disks and silicon strips after each disk. In Figure 2.5 a schematic view of ECAL is reported.

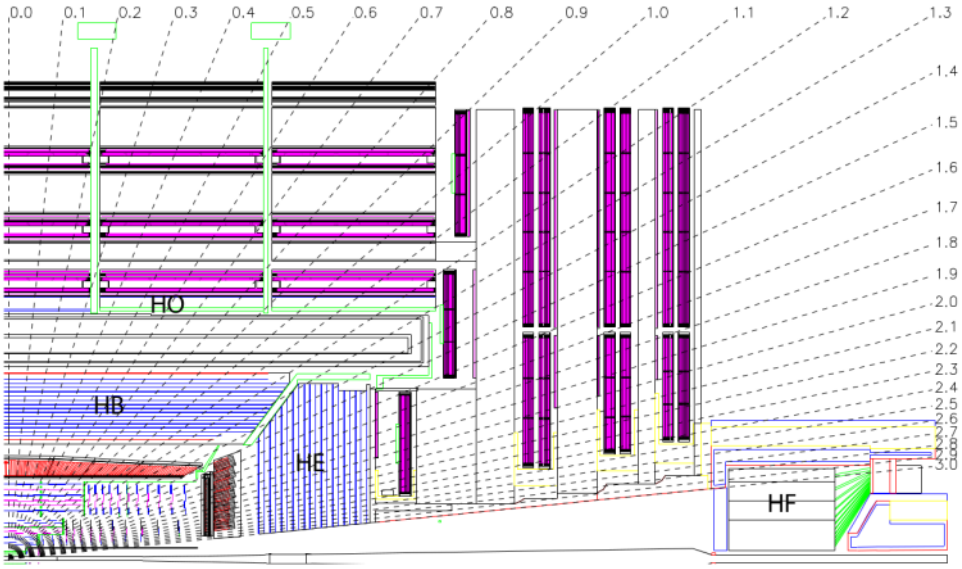


Figure 2.6: Schematic view of the CMS hadronic calorimeter.

The energy resolution, for energy below 500 GeV, is well described by:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b \oplus \frac{c}{E} \tag{2.9}$$

where a is stochastic term, that accounts for all the fluctuations in the interaction process, b takes into account the non-uniformity of the detectors and c includes the contributions from the electronic noise. In the EB these parameters were measured to be: $a = 2.8\%\,\mathrm{GeV}^{1/2}$, $b = 0.3\%$ and $c = 12\%\,\mathrm{GeV}$. Above 500 GeV showers are no longer fully contained (energy starts leaking from the rear of the crystals) and the resolution worsen.

2.2.3 The Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) is placed outside ECAL and measures the energy of hadrons produced in collisions [23]. Moreover, together with ECAL, it allows to measure energy and direction of jets. HCAL is a hermetic sampling calorimeter, covering the region $|\eta| < 5$. It is made of 5 cm layers of brass, the absorber material, interleaved with 3.7 mm plastic scintillators. The scintillation light is converted with WaveLength Shifter (WLS) fibres and readout by Hybrid PhotoDiodes (HPDs) or Silicon PhotoMultipliers (SiPMs).
It is made of four sections: HCAL Barrel (HB), located in the barrel region

inside the solenoid (covering $|\eta| < 1.4$), HCAL Endcap (HE), located in the endcaps inside the solenoid (covering $1.3 < |\eta| < 3$), HCAL Outer (HO), placed just outside the magnet (the return yoke of the solenoid is used as additional absorber material) and finally HCAL Forward (HF), made of quartz fibres between iron absorbers, is placed in the forward region $3 < |\eta| < 5$. In HF the quartz fibres produce Cherenkov light which is readout by radiation resistant Photo-MulTipliers (PMTs). The granularity of HCAL is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ for $|\eta| < 1.6$ and $\Delta\eta \times \Delta\phi \simeq 0.17 \times 0.17$ for $|\eta| > 1.6$. The calorimeter thickness varies between 7 and 11 interaction lengths[12] λ_I, depending on the η value.
In Figure 2.6 a schematic view of HCAL is reported. The resolution of HCAL can be parametrised by a stochastic term and a constant one:

$$\begin{aligned} \frac{\sigma_E}{E} &= \frac{90\%\,\mathrm{GeV}^{1/2}}{\sqrt{E}} \oplus 4.5\% \qquad \text{in HB, HE, HO} \\ \frac{\sigma_E}{E} &= \frac{172\%\,\mathrm{GeV}^{1/2}}{\sqrt{E}} \oplus 9\% \qquad \text{in HF} \end{aligned} \tag{2.10}$$

2.2.4 The Muon Chambers

Muons penetrate through all the previous detectors with a minimum loss of energy. Muon chambers, aimed to identify and measure high p_T muons in combination with the tracker, are placed outside the magnetic coil, embedded in the return yoke [24]. Three independent gaseous detectors are used:

- **Drift Tubes (DTs)** are placed in the barrel and cover the region $|\eta| < 1.2$, where the magnetic field is weak and homogeneous and the occupancy relatively low ($< 10\,\mathrm{Hz/cm^2}$).
  The barrel is divided into five wheels along the z axis and it is composed by four concentric stations in between the layers of the iron return yoke. Each station is formed by 12 DT chambers. The basic element of the DT chamber is a rectangular drift tube, with a transverse size of $13 \times 42\,\mathrm{mm^2}$ and a length varying between 2 m and 4 m. These tubes are filled with a gas mixture of Ar (85%) and CO_2 (15%), and are grouped to form detection layers: a group of four layers is called superlayer.
  In each chamber there are two superlayers with the anode wire parallel to the z axis and one superlayer perpendicular to it: in this way a chamber provides two measurements of the $r\phi$ coordinate and one

[12]The interaction length is a parameter that characterizes the transversal as well as the longitudinal shape of the hadronic shower.

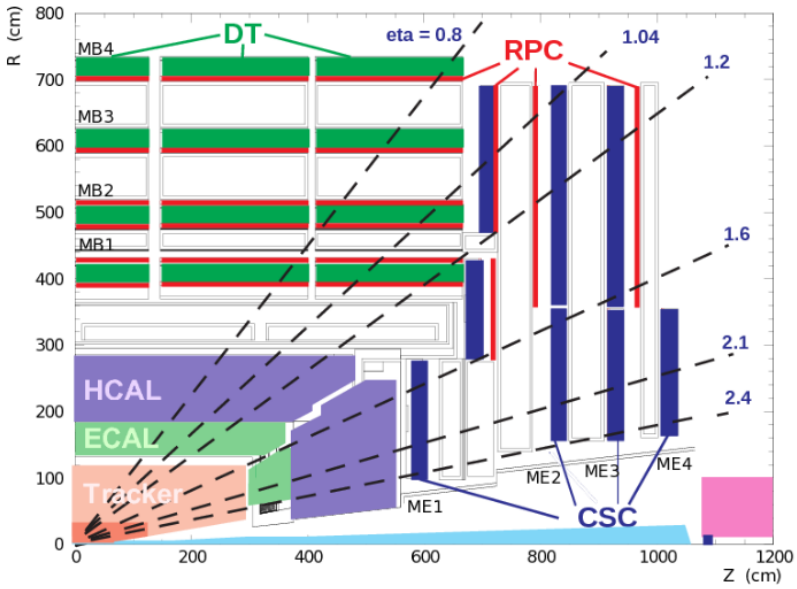


Figure 2.7: Schematic view of the muon chambers.

measurement of the z coordinate, with a resolution of about 100 µm in both cases.
In Figure 2.8 a sketch of a DT cell is reported: a stainless steel anode is placed between two parallel aluminium plates and two cathodes.

- **Cathode Strip Chambers (CSCs)** are placed in the endcaps, and cover the region $0.8 < |\eta| < 2.4$ where the occupancy is higher (greater than $100\,\mathrm{Hz/cm^2}$).
  Due to the high and non-uniform magnetic field and the high occupancy DTs can not be used. CSCs are multi-wire proportional chambers filled with a gas mixture of Ar (40%), CO_2 (50%) and CF_4 (10%). The cathodes are segmented into strips oriented radially while the anode wires are place transversally, as shown in Figure 2.9. In this way, a measurement of r and $r\phi$ is possible, with a spatial resolution of about 80 µm.

- **Resistive Plate Chambers (RPCs)** are placed both in the barrel and in the endcaps and cover the region $|\eta| < 2.1$. They are meant to complement DT and CSC detectors, by adding redundancy. RPCs are gaseous detectors that, albeit characterised by a coarse resolution, are able to perform precise time measurements, useful for the trigger system.

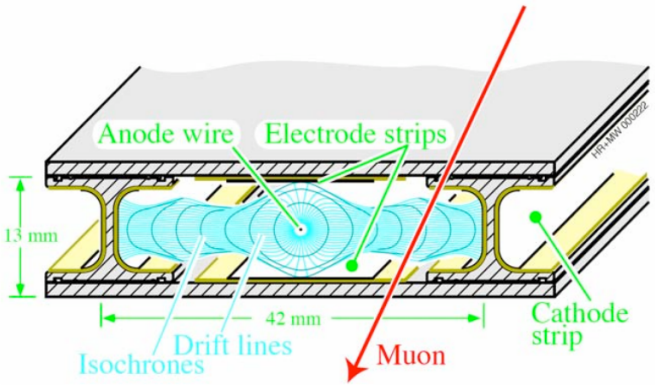


Figure 2.8: Sketch of a DT cell showing drift lines.

These detectors are made of four bakelite planes separated by 2 mm of $C_2H_2F_4$ (94.5%) and Isobutane. In the central part of the chambers insulated aluminium strips are placed to collect the signals generated by crossing particles.

2.2.5 The Trigger System

The collision rate of the LHC is 40 MHz, therefore it is necessary to reduce the amount of data to store and process. Moreover, the vast majority of events produced at the LHC is not interesting for physics analysis, as it involves events with low transverse momentum (these are called "minimum bias" events). The first step of event selection is performed by the CMS trigger system, composed of the Level-1 Trigger (L1) and the High Level Trigger (HLT) [25].
L1 uses dedicated processors and uses coarse informations from the calorimeters and the muon chambers. A decision has to be taken for each bunch crossing within 3.2 µs. The event rate is reduced from 40 MHz to 100 kHz.
HLT runs on a farm of commercial processor and uses the full granularity information of all the detectors. It reduces the event rate to a maximum of 1 kHz within 50 ms. At this point, data are saved on disk and can be used for analysis.

Level-1 (L1) Trigger

The L1 Trigger receives informations on the towers, which are clusters of energy measured by ECAL and HCAL. The readout electronics of the calorimeters is able to identify electron, photon and jet candidates, and to provide

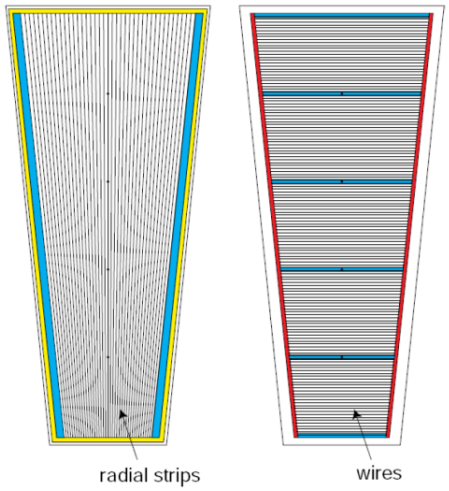
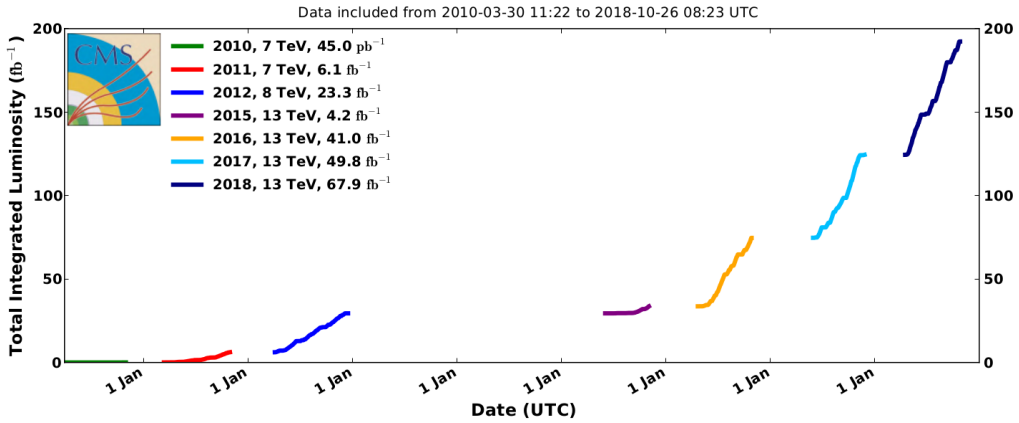


Figure 2.9: Schematic representation of CSC cathode (left) and anode (right) panels.

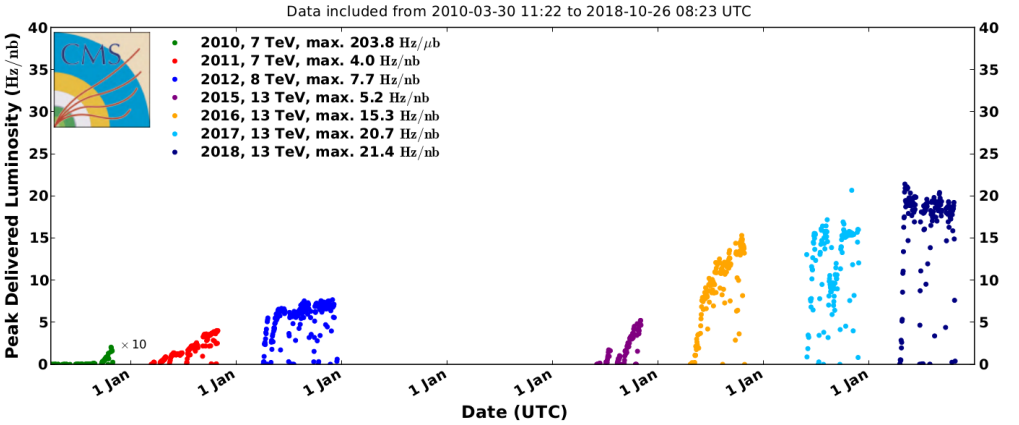
these informations to L1.
The muon system provides the track segments and a first estimate of their transverse momentum. A quality flag is assigned to each of these segments and the four with the highest quality are sent to L1.
The L1 global trigger combines these informations and decides whether or not to keep the event, by using a set of pre-programmed thresholds called L1 trigger table. If the event is flagged as interesting, all the detector subsystems are readout and sent to the HLT.

High Level Trigger (HLT)

HLT uses the informations of all the detectors. Data coming from the detectors are assembled by a builder unit and are sent to a switching network that dispatches them to the processor farm. The algorithms reconstructing the events are implemented via software and proceed by steps: at first only informations from calorimeters and muon chambers are used, for a coarse selection (Level-2), then informations from the tracker are added to refine the reconstruction (Level-2.5) and finally all the detectors informations are used to make the final decision (Level-3).
Each step reduces the number of events to be processed in the following step. The most computationally expensive tasks are executed in the Level-3. Algorithms such as track reconstruction are time consuming, therefore they are only executed in the region of interest, and on a limited set of hits, since the ultimate resolution is not required at this level.



(a) Integrated luminosity



(b) Instantaneous peak luminosity

Figure 2.10: Integrated luminosity and peak instantaneous luminosity as a function of the year, as measured by the CMS experiment.

2.3 High Luminosity LHC

The first physics run of the LHC was in 2010: the period between 2010 and 2012 is known as Run-1. In 2010 and 2011 the center of mass energy was $\sqrt{s} = 7$ TeV and CMS collected an integrated luminosity of $45\,\text{pb}^{-1}$ and $6\,\text{fb}^{-1}$ respectively. In 2012 the LHC operated at $\sqrt{s} = 8\,\text{TeV}$ and CMS collected $\mathcal{L} = 23\,\text{fb}^{-1}$.
Run-1 was followed by a two-year long shutdown, known as Long Shutdown 1 (LS1), during which the accelerator and the experiments were improved. Run-2 started in 2015 at $\sqrt{s} = 13\,\text{TeV}$. Operations in 2015 were focused on the commissioning of the LHC at the new energy, and the instan-

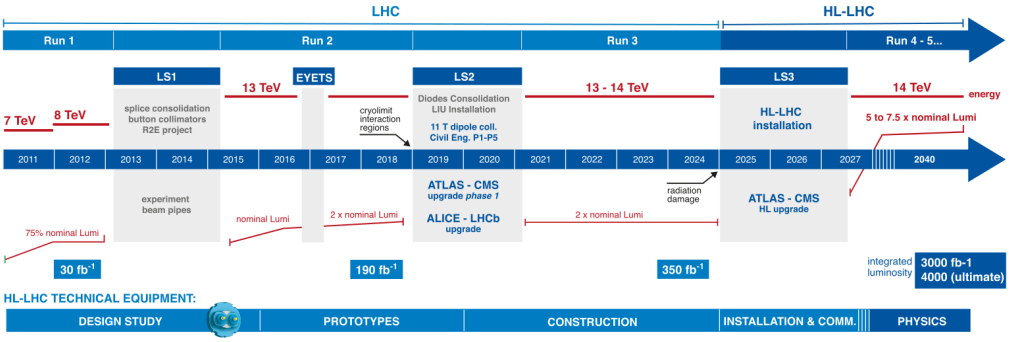


Figure 2.11: Current schedule of LHC and HL-LHC operations.

taneous luminosity was kept below 5.0×10^{33} cm^{-2}s^{-1}. In 2016 the instantaneous luminosity reached 1.5×10^{34} cm^{-2}s^{-1}, beyond the design specifications (1.0×10^{34} cm^{-2}s^{-1}) and CMS collected 38 fb^{-1}. In the following years the operations proceeded smoothly and CMS collected 45 fb^{-1} in 2017 and 63 fb^{-1} in 2018 (these numbers refer to the integrated luminosity validated for use in physics analyses). In October 2018, with the last proton-proton collisions, Run-2 ended. In Figure 2.10 the integrated and instantaneous luminosities measured by CMS are reported, as a function of the year.
Since 2018 the Long Shutdown 2 (LS2) started, in which the machine is being upgraded for future high luminosity phases. Run-3 will start in 2021 at $\sqrt{s} = 14$ TeV with a peak instantaneous luminosity twice the design value. In this way CMS is expected to collect up to 300 fb^{-1} in three years. In 2025 the Long Shutdown 3 (LS3) will begin, ending the LHC Phase-1. Between 2025 and 2027 the experiments and the machine will be upgraded, for the Phase-2 of LHC, known as High Luminosity LHC (HL-LHC).
In this new phase the instantaneous peak luminosity will reach a value of 5×10^{34} cm^{-2}s^{-1} or even 7.5×10^{34} cm^{-2}s^{-1} in the ultimate performance scenario. CMS will be able to collect 300 fb^{-1} per year for a total of 3000 fb^{-1} in ten years (or 4000 fb^{-1} in the ultimate performance scenario). In Figure 2.11 the current schedule of LHC and HL-LHC operations is reported.
The increase of the instantaneous luminosity will be achieved by a variety of methods [26]. New Nb_3Sn quadrupole magnets will be installed at the CMS and ATLAS interaction points to focus the beams. These magnets are more radiation tolerant and they are able to generate a higher magnetic field, allowing for a luminosity increase. Moreover superconducting radiofrequency crab cavities will be installed in the interaction regions, allowing a rotation of the beams before collision so that the bunches collide head on.
The pile-up in LHC reached 53 at the highest instantaneous luminosity, in

a special run in 2016: in HL-LHC the pile-up will reach 140 or 200 in the ultimate performance scenario, making it necessary to upgrade detectors significantly. Moreover, the radiation levels will be unprecedented: after ten years of operations the fluence in the innermost layer of the future CMS tracker is expected to reach $2.3 \times 10^{16}\,\mathrm{n_{eq}cm^{-2}}$, corresponding to a Total Ionising Dose of 12 MGy.
The HL-LHC goal is to significantly expand the physics potential of the LHC, both for rare Standard Model and Beyond Standard Model processes [27,28].

2.3.1 The CMS Upgrade

During the LS3 the CMS experiment will be significantly upgraded for Phase-2, in order to operate in the conditions of HL-LHC [27]. Due to the higher radiation levels, the radiation hardness of the detectors has to be improved. Moreover, the higher pile-up and the consequent increased track density require higher detector granularity, to reduce the occupancy, and increased bandwidth, to readout the higher data rates. The trigger rate will be increased at 750 kHz at L1 and at 7.5 kHz at the HLT. The trigger latency at L1 will be 12.5 µs.
The muon chambers are expected to resist the high demands of Phase-2. However the readout electronics of DTs and CSCs will be completely replaced with improved versions. In the endcaps RPCs will be improved and new chambers based on Gas Electron Multiplier (GEM) technology will be installed, in order to add redundancy, increase the acceptance and improve trigger capabilities [29].
The crystals in ECAL will be operated at a lower temperature with respect to Phase-1, and the readout electronics will be improved [30]. Moreover, thanks to new front-end boards, the L1 trigger will be able to use informations coming from a single crystal (in Phase-1 L1 uses the global information of towers). The scintillation layers of HCAL closer to the beam will also be replaced.
Both ECAL and HCAL endcaps will be replaced with a new combined electromagnetic and hadronic sampling calorimeter, HGCAL [31]. The electromagnetic compartment (CE-E) consists of 28 sampling layers with a total thickness of 34 cm and a depth of approximately $26\,X_0$ and $1.7\,\lambda_0$. The rear hadronic compartment (CE-H) consists of 25 layers with a depth of $8.5\,\lambda_I$. Detector planes are equipped with either silicon sensors or scintillator tiles readout by SiPMs. Copper, tungsten, and lead are used as absorber materials in the CE-E, while stainless steel is used in the CE-H. This detector will provide high granularity (both transverse and longitudinal) which will help in mitigating the pile-up and in the reconstruction of the physics objects.

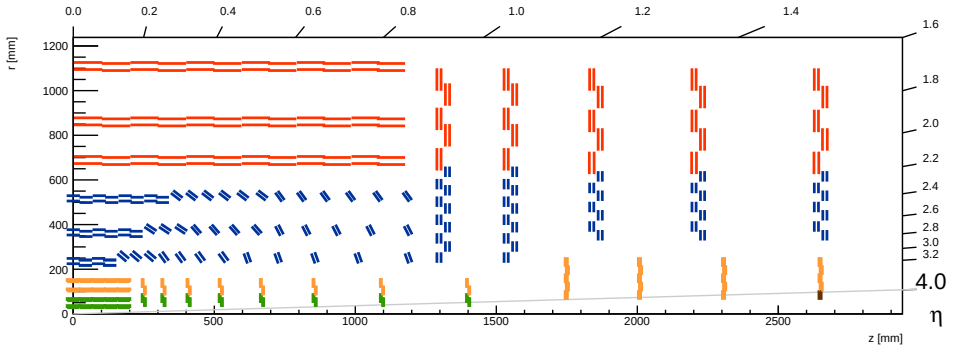


Figure 2.12: Schematic view of a quarter of the tracker in the $r-z$ plane. The Inner Tracker is depicted by green and yellow lines, corresponding to modules with two and four ROCs respectively. The Outer Tracker is depicted by blue and red lines, corresponding to modules with a strip detector and a macro-pixel detector, and modules with two strip detectors, respectively [20].

A timing detector sensitive to Minimum Ionizing Particles (MIPs), with a time resolution of the order of 30 ps, will be placed between the tracker and ECAL [32]. This additional MIP Timing Detector (MTD) will be specialized to provide timing for the individual tracks crossing it, while photon and hadron timing will be partly provided by the upgraded CMS calorimeters. The added value of a timing detector is expressly quantified in terms of improved track and vertex reconstruction abilities, lepton efficiencies, diphoton vertex location, and MET resolution. Two technologies are considered for the MTD: scintillating crystals readout by SiPMs for lower radiation areas and silicon Low Gain Avalanche Detectors (LGADs) for higher radiation areas.
The CMS silicon tracker will be completely replaced, since the Phase-1 Tracker would suffer from significant damage and performance degradation during Phase-2, while also not being compatible with the requirements of HL-LHC [33]. In the following an overview of the future tracker is presented.

2.4 The CMS Tracker Upgrade

The future CMS tracker will be made of two subsystems, the Inner Tracker (IT), made of pixel detectors, and the Outer Tracker (OT), made of strip and macro-pixel detectors. The Inner Tracker will be made of four layers in the barrel and twelve disks per side in the endcaps. It will cover the region closer to the beam pipe: $r < 20$ cm with $|z| < 160$ cm and $r < 30$ cm with $160\,\text{cm} < |z| < 270$ cm. The innermost layer will be at only 30 mm from

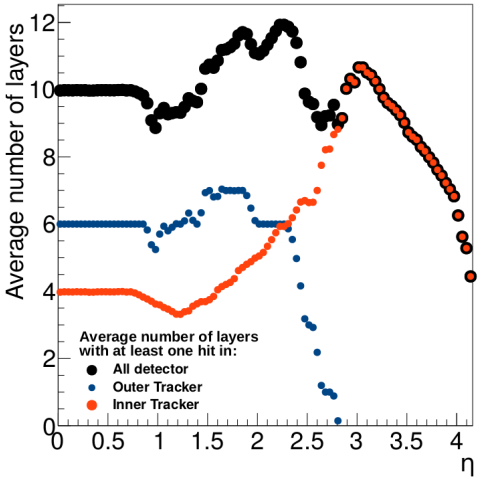


Figure 2.13: Average number of module layers traversed by particles as a function of η, including the Inner Tracker (red), the Outer Tracker (blue) and the complete tracker (black) [20].

the beam line. The Outer Tracker will be made of six layers in the barrel and five disks per side in the endcaps and covers the region $r < 120$ cm and $|z| < 270$ cm.
In Figure 2.12 a schematic view of the new tracker is reported. The number of layers was optimised to ensure robust tracking, so that the performances are unaffected when one detecting layer is lost in some parts of the rapidity acceptance. Figure 2.13 shows the simulated average number of active layers that are traversed by particle as a function of η, for the complete tracker as well as for the Inner Tracker and the Outer Tracker separately.
The requirements of the new tracker can be summarised as follows:

- **Radiation tolerance** for an integrated luminosity up to 4000 fb^{-1}. For the Outer Tracker this requirement needs to be fulfilled without any maintenance, since the detectors will not be easily accessible. However, it will be possible to extract the Inner Tracker and possibly replace modules that sustained heavy radiation damage. Detailed FLUKA [34, 35] simulations have been performed to estimate the expected fluences in the different regions of the tracker, as shown in Figure 2.14. The particle fluence depends primarily on r, while the variation with z is very moderate.

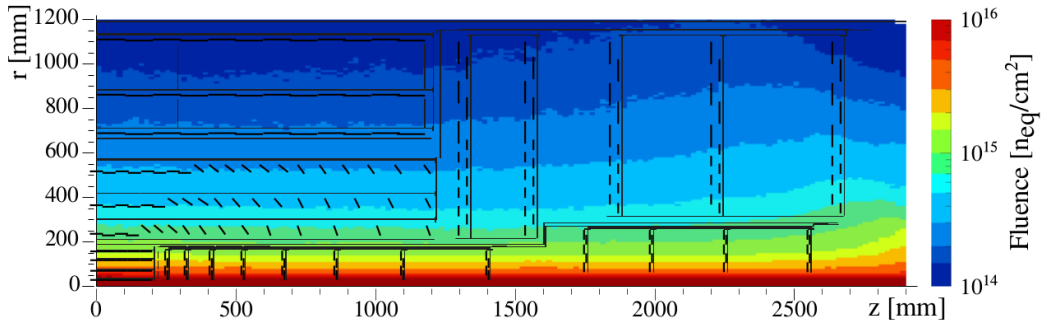


Figure 2.14: Schematic view of a quarter of the tracker in the $r - z$ plan, showing the expected fluence in Phase-2 (in $n_{eq}cm^{-2}$ units), considering $\sqrt{s} =$ 14 TeV collisions and an integrated luminosity $\mathcal{L} = 3000\,fb^{-1}$ [20].

- **Increased granularity**, in order to keep the occupancy[13] of the order of 10^{-3} in the IT and 10^{-2} in the OT. This is necessary to ensure efficient tracking performance with the high pile-up expected during Phase-2.

- **Improved two-track separation**: the present tracker has limited performance in track finding of highly energetic jets, due to hit merging in the pixel detector. To exploit the increased number of collisions in HL-LHC, an improvement of this parameter is necessary.

- **Reduced material in the tracker**: the performance of the present tracker is limited by the amount of material, which also impacts the calorimeters. Figure 2.15 shows the expected material reduction.

- **Robust pattern recognition**, so that track finding can be fast and efficient even under the expected high pile-up conditions.

- **Contribution to the L1 trigger**: the selection of interesting physics events becomes challenging under high pile-up conditions, due to rate increase and inefficiencies in the selection. Tracking information will be used by the L1 trigger, anticipating part of the reconstruction performed by the HLT.

- **Extended tracking acceptance**, in order to have efficient tracking up to $|\eta| = 4$. The physics capabilities of CMS will greatly benefit from the extended acceptance, both in tracker and in the calorimeters.

[13] Fraction of the activated channels in each bunch crossing.

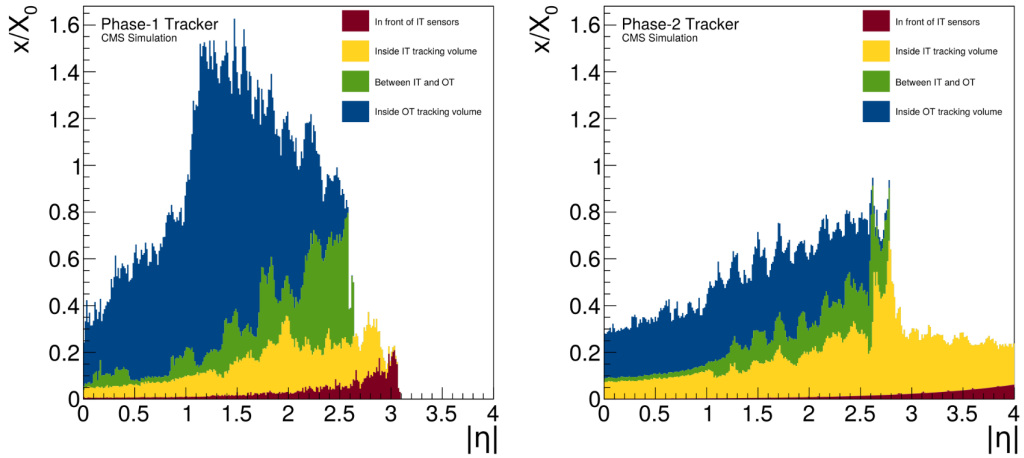


Figure 2.15: Comparison of the total thickness of the present CMS tracker and Phase-2 tracker, in X_0 units and as a function of η [20].

In order to limit the data to send to the L1 trigger at every bunch crossing, the tracker will make a self selection of interesting events. This is achieved by using detectors that are capable of rejecting signals from particles below a certain p_T threshold, referred to as "p_T modules". As shown in Figure 2.16, p_T modules are made of two closely-spaced silicon strip detectors readout by a common set of front-end Application Specific Integrated Circuits (ASICs) that correlate the signals in the two sensors and select the hit pairs, referred to as "stubs". Tracks from charged particles are bent in the transverse plane by the 3.8 T field of the CMS magnet, with the bending angle depending on the p_T of the particle. Stubs that are geometrically compatible with tracks above a certain p_T threshold ($2-3$ GeV) are selected and sent to L1 at every bunch crossing. All other hits are stored in the pipelines waiting for the L1 trigger response. The resolution of p_T modules depends on the distance from the interaction point, therefore they are only used in the Outer Tracker.
Two versions of p_T modules were conceived: modules with two strip sensors (2S modules) and modules with a strip and a macro-pixel sensor (PS modules). The strips in the 2S modules have a length of about 5 cm, while those in the PS modules are about 2.4 cm long. Both module types have a pitch of about 100 µm. In PS modules one of the two sensors is segmented into macro-pixels with a length of about 1.5 mm, providing the z (r) coordinate measurement in the barrel (endcaps).
The PS modules will be installed in the first three layers of the Outer Tracker ($200 < r < 600$ mm) while the 2S modules will be installed in the outermost three layers ($r > 600$ mm). In the endcaps the modules are arranged in rings

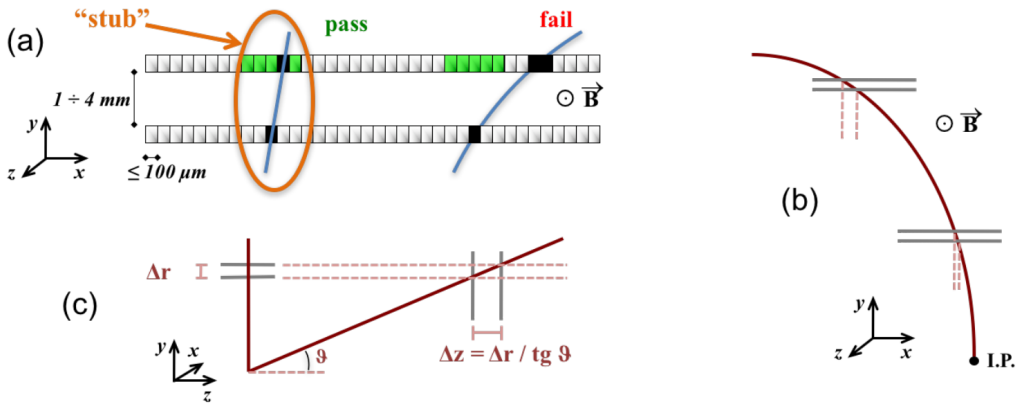


Figure 2.16: Illustration of the p_T module concept. Correlation of signals in closely-spaced sensors enables rejection of low p_T particles. The channels shown in green represent the selection window to define an accepted stub (a). The same p_T corresponds to a larger distance between the two signals at large radii for a given sensor spacing (b). For the endcap disks, a larger spacing between the sensors is needed to achieve the same discriminating power as in the barrel at the same radius (c) [20].

on disk-like structures, with the rings at low radii ($r < 700$ mm), equipped with PS modules, while 2S modules are used at larger radii.
The readout electronics require Low Voltage (LV), in the range of $1 - 2$ V, and high current powering, provided by Power Supplies Units (PSU) installed outside the CMS experiment. A direct powering scheme requires too much material in the tracker, therefore a powering scheme based on a Point of Load (PoL) conversion was chosen for the Outer Tracker. On each module a DC-DC converter generates the necessary voltage from an intermediate voltage ($10 - 12$ V) provided from the PSUs. Considering also the High Voltage (HV) necessary to polarise the silicon sensors, which follows a direct powering scheme instead, the Outer Tracker requires a power of roughly 90 kW. Thanks to the PoL conversion, the passive material associated with the power supply cables is greatly reduced. The power dissipated from the Phase-2 tracker is removed by a CO_2 cooling system. The powering scheme of the Inner Tracker is discussed in Section 2.4.1.
In Figure 2.17 the tracking efficiency in jet cores is shown as a function of the distance ΔR between a simulated track and its nearest neighbour, for the Phase-1 and the Phase-2 trackers. Although the reconstruction algorithms were not optimised for Phase-2, a significant improvement can already be seen for small values of ΔR, thanks to the higher granularity of the new

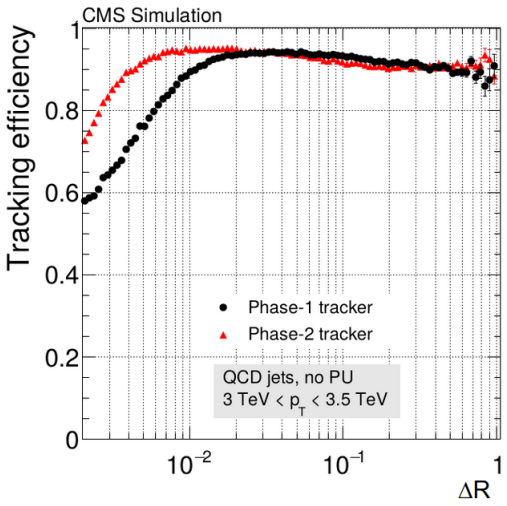


Figure 2.17: Tracking efficiency in the core of jets with $3\,\text{TeV} < p_T < 3.5\,\text{TeV}$ as a function of the distance ΔR between a simulated track and its nearest neighbour, for the Phase-1 (black) and the Phase-2 (red) trackers [20].

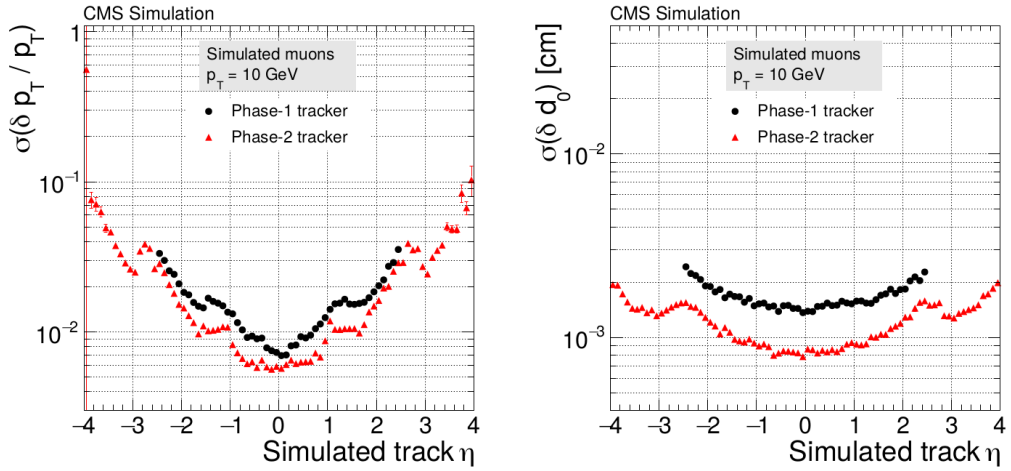


Figure 2.18: Relative p_T resolution (left) and resolution of the transverse impact parameter (right) as a function of η for the Phase-1 (black) and the Phase-2 (red) trackers, using single isolated muons with $p_T = 10$ GeV [20].

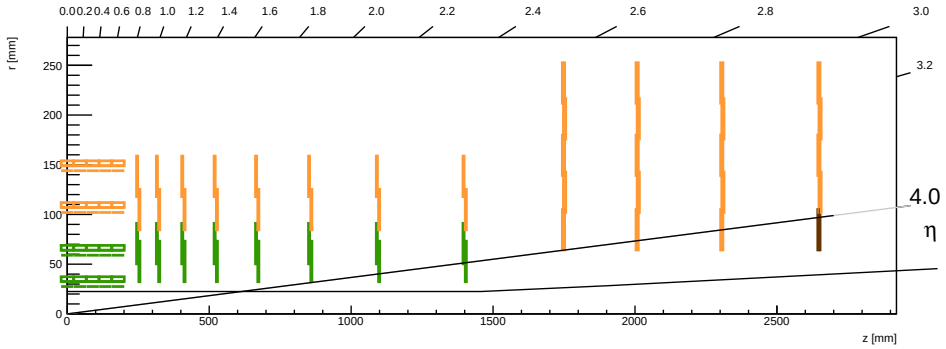


Figure 2.19: Schematic view of a quarter of the tracker in the $r - z$ plane. Modules made of two ROCs are depicted in green, while modules with four ROCs are depicted in orange.

detector. Figure 2.18 shows the p_T resolution and the transverse impact parameter[14] resolution for single muons with $p_T = 10$ GeV as a function of η for both the current and the future trackers.

2.4.1 The Inner Tracker

The Inner Tracker will be made of a barrel part with four layers (Tracker Barrel Pixel Detector, TBPX), eight small double-disks per side (Tracker Forward Pixel Detector, TFPX) and four large double-disks per side (Endcap Pixel Detector, TEPX), as shown in Figure 2.19 and in Figure 2.20.
In the TBPX the pixel modules will be arranged in "ladders" along the z axis. In each layer, neighbouring ladders will be mounted staggered in radius, so that an overlap in $r\phi$ is achieved. Modules on a ladder will not overlap in z. To avoid a projective gap at $\eta = 0$, an odd number of modules will be mounted on the ladders and the barrel mechanics will be split in two asymmetric halves. In TFPX and TEPX the modules will be arranged in concentric rings. Each double-disk is physically made of two disks, so that modules can be mounted onto four planes, with overlaps in r and $r\phi$.
In Figure 2.21 the three subsystems are visible inside the supporting structure, which is referred to as service cylinder since it hosts the services necessary to the detectors (power supply cables, readout connections, cooling pipes). The Inner Tracker will have a total pixel surface of approximately 4.9 m^2 and 2×10^9 channels, from almost 4000 pixel modules.
In order to increase the granularity with respect to Phase-1, the pixel pitch

[14]Minimum distance in the transverse plane between the reconstructed track and the primary interaction vertex.

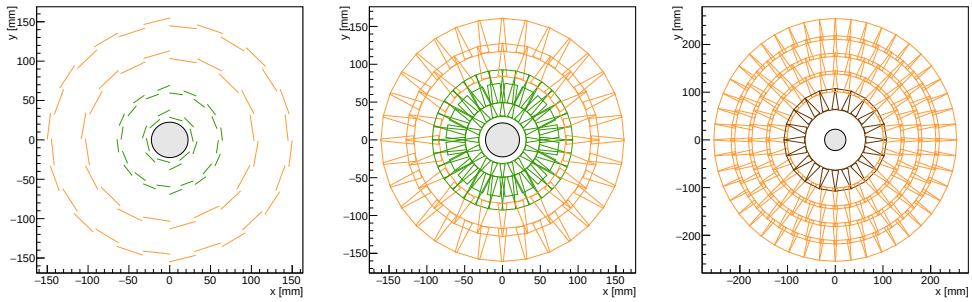


Figure 2.20: Schematic view of the geometry of FBPX (left), TFPX (center), TEPX (right) in the $x - y$ plane. Modules made of two ROCs are depicted in green, while modules with four ROCs are depicted in orange. The central circle represents the beam pipe.

will be 25×100 µm^2 (with the long side pointing along z in the barrel and along r in the endcaps) or 50×50 µm^2, while the present tracker is made of 100×150 µm^2 pixels. The performances of these two cell types are still under study, and depend on the region in which they will be installed. Moreover, the active thickness will be reduced from 285 µm (Phase-1) to $100 - 150$ µm (Phase-2), and the silicon sensors will be n-in-p instead of n-in-n. Finally, two pixel technologies are being considered: the standard planar pixel sensors and 3D pixel sensors. More details on the future pixel sensors are reported in Section 2.5.

The Phase-2 ROC has an active area of 16.4×22.0 mm^2, corresponding to a pixel matrix of 328 cells with a 50×50 µm^2 pitch. This ROC is developed by RD53, a joint ATLAS-CMS collaboration [36], using 65 nm CMOS technology. More details on the Phase-2 ROC are reported in Section 3.1.

A pixel module is made of a pixel sensor, several ROCs, and a thin, high density flex circuit, referred to as High Density Interconnect (HDI). Sensors are bump bonded to the ROCs and glued to the HDI. The ROCs are then wire bonded[15] to the HDI. The HDI provides clock trigger and control signals, as well as the power distribution, while also shipping the data out.

Two types of modules are foreseen, differing only on the pixel surface and the number of ROCs: modules with two ROCs (1×2 modules) and with four ROCs (2×2 modules). The 1×2 modules will be placed in the innermost regions, while the 2×2 modules will be placed in the outer layers, as show in Figure 2.19 and in Figure 2.20. Figure 2.22 shows a sketch of the two types of

[15]Interconnection between the integrated circuit and its packaging, usually made with aluminium wires.

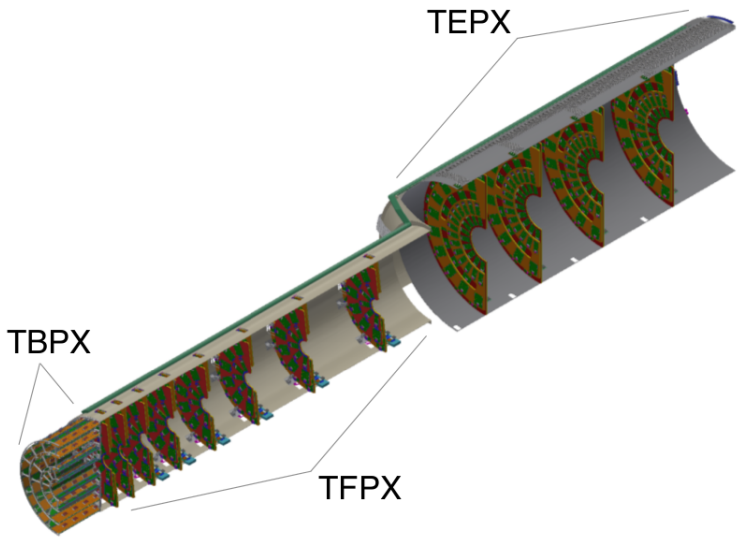


Figure 2.21: Prospective view of a quarter of the Inner Tracker, showing TBPX, TFPX, TEPX and the service cylinder [20].

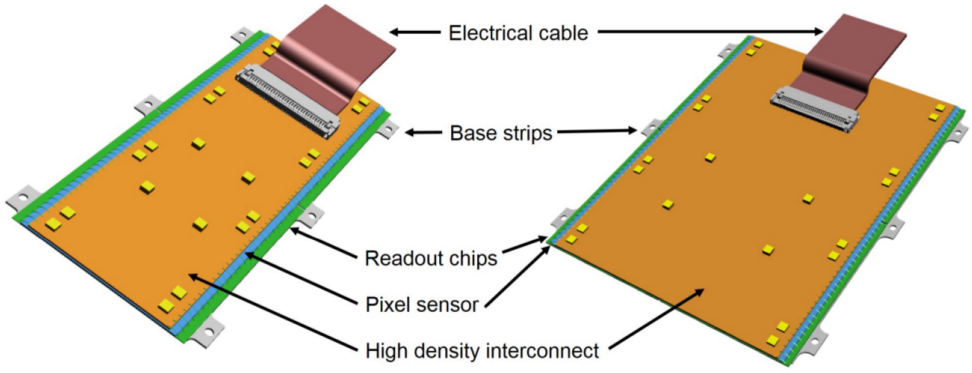


Figure 2.22: Schematic sketch of the 1×2 pixel module (left) and the 2×2 pixel modules (right) [20].

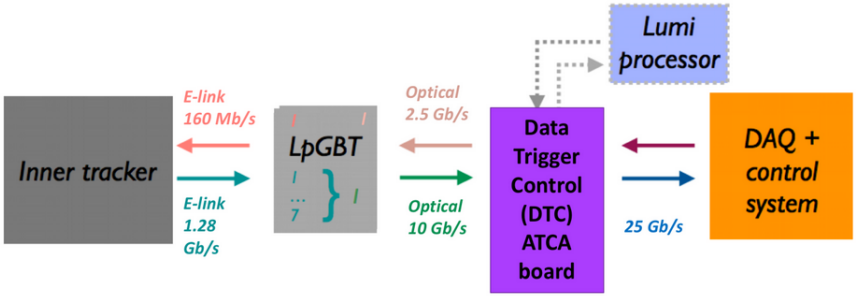


Figure 2.23: Scheme of the readout chain of the Inner Tracker.

modules. The cooling system will keep the modules at a stable temperature of about -20 ℃.
Hits data are stored by ROCs during the 12.5 µs trigger latency. Triggered data events are then extracted from the ROCs, compressed, and sent to differential digital lines (E-links), at 1.28 Gb/s. The number of E-links per module is configurable $(1-6)$ in order to have sufficient bandwidth in the innermost layers. In the outer layers, where the hit rates are lower, event data from all chips of a module are merged into a single E-link. The Data AcQuisition (DAQ) system is placed too far from the detectors to be compliant with an electrical transmission protocol. Therefore, opto-conversion modules[16], based on the LpGBT (Low-power Gigabit Transceiver) chip set, merge data from up to seven E-links into 10 Gb/s optical links for transmission to the off-detector DAQ system. Clock, trigger, real-time commands, and configuration data are received by the LpGBTs on the opto-conversion modules via 2.5 Gb/s optical down links and sent to the pixel chips via one 160 Mb/s E-link per pixel module.
The DAQ interface of the Inner Tracker consists of a modular DTC (Data, Trigger, Control), which communicates with the on-detector electronics via the LpGBT optical links. Figure 2.23 shows a scheme of the readout chain of the Inner Tracker. The DTC is planned to accommodate 72 pairs of LpGBT optical up and down links. A crate with six pixel DTC modules can accommodate the readout and control of a quarter of the Inner Tracker: there will be a total of four crates with 24 DTC modules and 1728 available optical link pairs.
The Phase-2 ROC necessitates a modern low density CMOS technology with low supply voltage (about 1.2 V), that however requires a high current level

[16]The opto-conversion modules are placed in the service cylinder, where the radiation levels are tolerable.

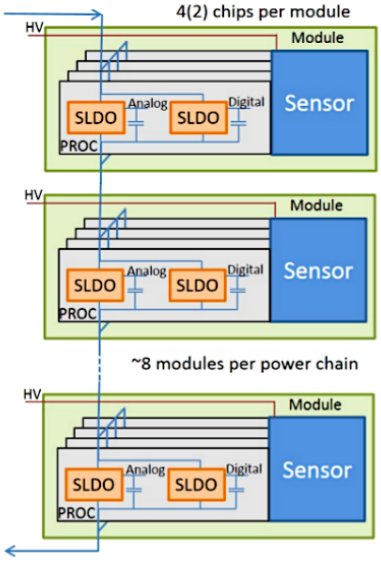


Figure 2.24: Serial power distribution with the ROCs in parallel on each module and modules connected in series [20].

(about 2 A), corresponding to a power consumption of about $0.5\,\mathrm{W/cm^2}$. A direct powering scheme would require power cables of large section, which would dramatically increase the passive material in the tracker. A possible approach, based on the present pixel tracker, is based on DC-DC converters. However, these are affected by two problems. In the first place, they are not radiation resistant enough for the Inner Tracker environment. In the second place, they are large and heavy objects that are difficult to place inside the tracker, while also adding significant passive material near the collision point. While this scheme is indeed used in the Outer Tracker, since it is far less affected by these problems, the Inner Tracker has opted for a direct powering scheme instead.
The Inner Tracker is organised in chains of up to five modules, with the ROCs of each module connected in parallel, as shown in Figure 2.24. The ROC includes a highly specialized circuit that combines the functionality of a current shunt and a Low-DropOut (LDO) regulator, thus referred to as Shunt-LDO (SLDO)[17]. The SLDO ensures that power and current consumption are kept constant, independent of hit and trigger rates. Moreover, the SLDO is designed to ensure appropriate current sharing between multiple chips, powered in parallel. Thanks to this scheme, the serial chain presents

[17]Two SLDOs are present on the ROC: one for the digital domain and one for the analog domain.

itself as a constant load to the power supplies: the SLDO manages the ROC power consumption variations.
One downside of the serial powering scheme is that it is more susceptible to failures, since a malfunction in one module propagates to all the chain. Therefore a careful study of the possible failure scenarios is necessary.
Each module in a chain works with a different local ground, since the output of a module is the input of the following one. Since the HV (needed to polarise the sensors) will be provided with a direct powering scheme, with the modules in a chain connected in parallel, the applied bias voltage on the silicon sensors will be different through the chain. With a maximum voltage drop of $\simeq 2.5\,\mathrm{V}$ between two modules in a chain, and considering five modules in the chain, the maximum bias voltage difference between the first and the last module in a chain is about 10 V. This could lead to performance differences along the chain for some types of pixel sensors, as will be shown in Chapter 5.

2.5 FBK Pixel Sensors Productions

For the development of the pixel sensors to be installed in the future CMS Inner Tracker, the CMS collaboration works with several companies producing silicon sensors. All the pixel sensors analysed in this Thesis (both planar and 3D) are developed and fabricated in the framework of the Istituto Nazionale di Fisica Nucleare (INFN) and Fondazione Bruno Kessler (FBK) collaboration activity [8, 37]. A total of four FBK batches were produced since 2016, two for planar sensors and two for 3D sensors. A batch consists in a production of several silicon wafers, with each wafer containing several (n-in-p) pixel sensors. The wafers processed by FBK have a diameter of 150 mm.
The wafers were manufactured with the Si-Si DWB technique[18], using a 500 µm thick handle wafer made of Czochralski (CZ) silicon (with less than $1\,\Omega$ cm resistivity) and a high resistivity ($> 3000\,\Omega\,\mathrm{cm}$) Float Zone (FZ) silicon wafer. The thickness of the FZ wafer (and hence the active thickness of the sensors) was 100 µm, 130 µm or 150 µm depending on the production. The CZ wafer is thinned after production, in order to achieve a total silicon thickness between 200 µm and 250 µm. Sensors on wafers were tested at FBK premises using the temporary metal technique. FBK sensors are always treated with the p-spray isolation technique.
As explained in Section 2.4.1, two pixel pitches are considered for the future CMS Inner Tracker (25×100 µm^2 and 50×50 µm^2): both pitches were

[18] Some wafers, not studied for this Thesis, featured the Silicon-On-Insulator (SOI) technology.

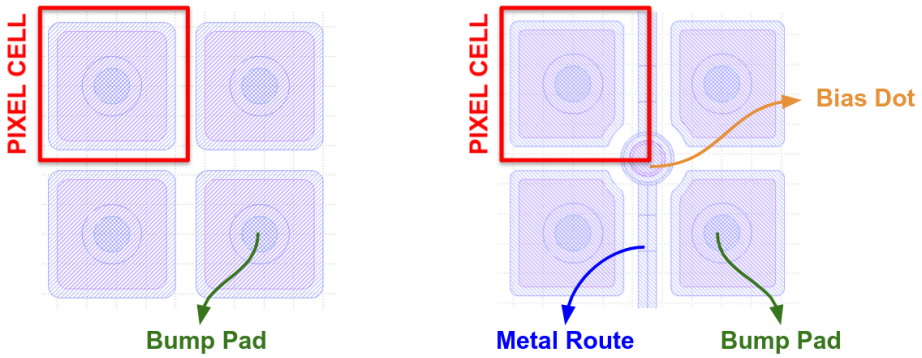


Figure 2.25: Schematic drawings of 50×50 µm^2 pixels (for a 2×2 pixel grid) for two sensors of the P-1 batch: without the PT structure (left) and with the PT structure (right). The metal route connects the bias dot to the bias grid. A 50×50 µm^2 pixel cell is highlighted in red.

studied with FBK batches. However, since these sensors need to be bump bonded to a readout chip (presented in Section 3.1) with a pixel pitch of 50×50 µm^2, the 25×100 µm^2 sensors had to be adapted to the different pitch with a particular arrangement of the bump bonding pads.
In the following, an overview of the FBK batches is reported, highlighting the different pixel designs which are studied in this Thesis.

2.5.1 Planar Batches

The two planar batches will be referred to as P-1 and P-2 in the following. The P-1 batch was produced in 2016, while the P-2 batch in 2019. Sensors are protected by a minimum of one and a maximum of three guard rings, depending on the batch and on the sensor type.
Figure 2.25 shows the designs of 50×50 µm^2 pixels (in a 2×2 pixel grid) for two sensors from the P-1 batch: one features the PT structure[19] while the other does not. The size of the n^+ implant is 32×36 µm^2 and 37×37 µm^2 respectively: with the PT structure, the n^+ implants is reduced in order to allow the metal route to connect the bias dot to the bias ring at the periphery of the sensor. Section 5.2.1 reports the effects of the PT structure on the sensor performance. Figure 2.26 shows a photograph of a 50×50 µm^2 pixel sensor (without the PT structure) on wafer: the temporary metal is still present and connects all the pixels in the matrix.

[19]While the sensors were tested on wafer with the temporary metal solution, some sensors were designed with the PT structure in order to evaluate hit detection efficiency losses on a particle beam.

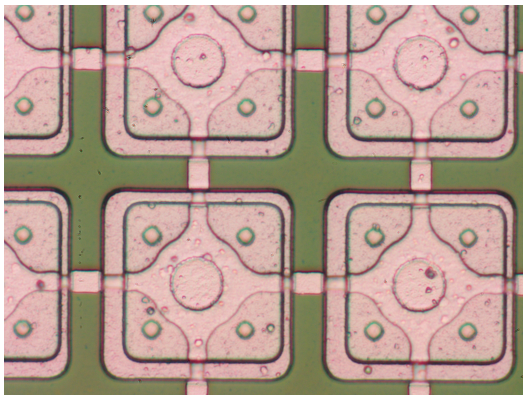


Figure 2.26: Photograph of a 50×50 µm^2 pixel sensor from the P-1 batch, taken on wafer. The temporary metal is still present and connects all the pixels in the matrix.

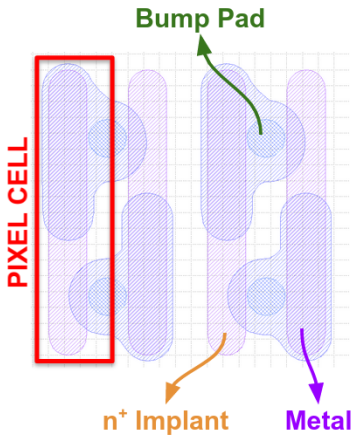


Figure 2.27: Schematic drawing of 25×100 µm^2 pixels (in a 4×1 pixel grid) for a sensor of the P-1 batch. A 25×100 µm^2 pixel cell is highlighted in red.

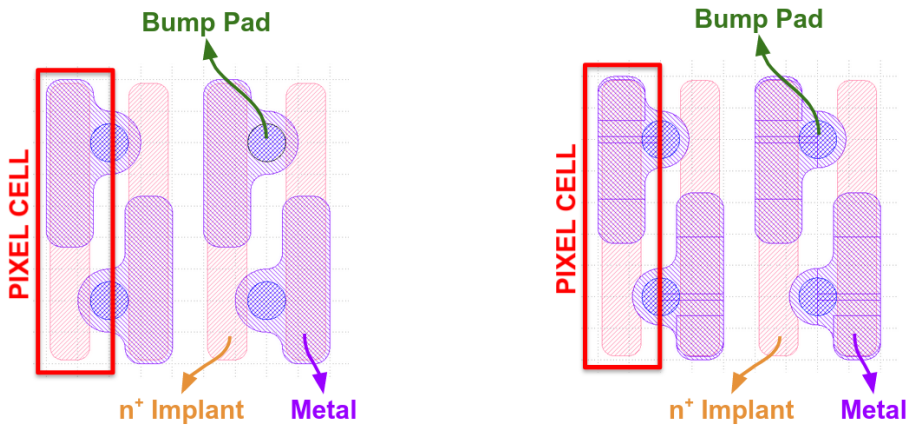


Figure 2.28: Schematic drawings of 25×100 µm^2 pixels (in a 4×1 pixel grid) for two sensors of the P-2 batch: standard (left) and BT (right). The metal only partially covers the n^+ implant. A 25×100 µm^2 pixel cell is highlighted in red.

Figure 2.27 shows the design of 25×100 µm^2 pixels (in a 4×1 pixel grid) for a sensor from the P-1 batch. The size of the n^+ implant is 12×90 µm^2. The bump bonding pads partially overlap the neighbour pixel: this was necessary to adapt the 25×100 µm^2 pixels of the sensor to the 50×50 µm^2 pixels of the ROC. This overlap may induce cross-talk between neighbour coupled pixels, as detailed in Section 3.3. In order to reduce the cross-talk effect, the P-2 batch contained variants of this design.
In the P-2 batch the dimension of the n^+ implant was slightly modified in 12.5×87.5 µm^2, and the metal around the bump pad was reduced. Moreover, four alternative designs for this pitch were included in the batch. Figure 2.28 shows the designs of 25×100 µm^2 pixels (in a 4×1 pixel grid) for two sensors of the P-2 batch: one has a standard n^+ implant while the other one has a peculiar structure and is referred to as BiTen (BT). In the BT sensor, the n^+ implant is slightly reduced in correspondence to the bump pad and the metal layer of the neighbour pixel (bitten design). This particular design is expected to reduce the cross-talk effect, since the bump pad and the implant of the neighbour pixel do not overlap. Figure 2.29 shows a photograph of the BT sensor on wafer. The n^+ implant (in light green) is bitten in correspondence to the bump pad of the neighbour pixel.
Figure 2.30 shows two designs of 25×100 µm^2 pixels (in a 4×1 pixel grid) for other two sensors of the P-2 batch: they are referred to as BiTen Field Plate (BTFP) and Extended BiTen Field Plate (EBTFP). In the BTFP sensor, the n^+ implant features the same bitten design of the BT sensor, however in this

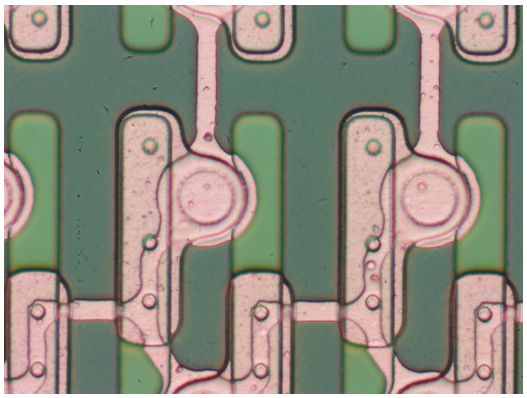


Figure 2.29: Photograph of a BT sensor from the P-2 batch on wafer. The temporary metal is still present and connects all the pixels in the matrix.

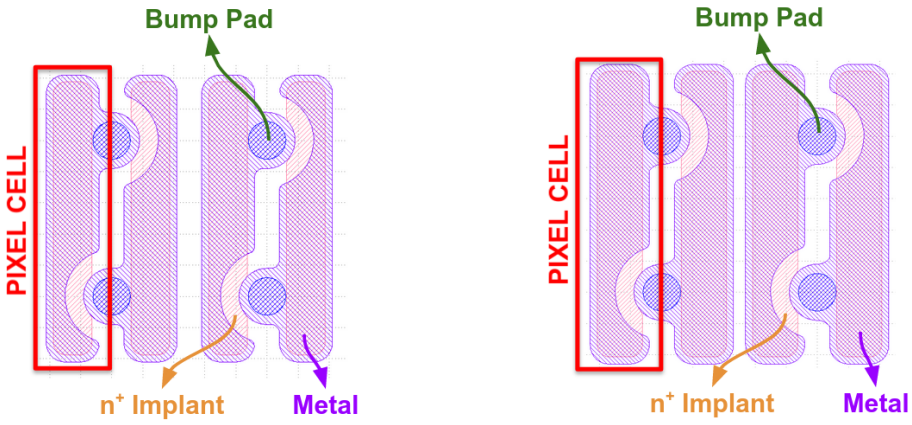


Figure 2.30: Schematic drawings of 25×100 µm^2 pixels (in a 4×1 pixel grid) for two sensors of the P-2 batch: BTFP (left) and EBTFP (right). The metal fully covers the n^+ implant, except in proximity of the bump pad of the neighbour pixel. A 25×100 µm^2 pixel cell is highlighted in red.

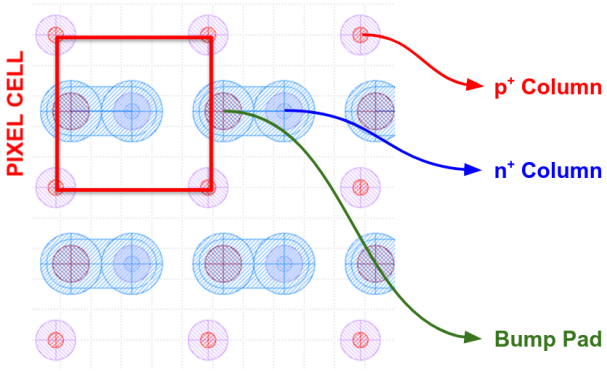


Figure 2.31: Schematic drawings of 50×50 µm^2 pixels (in a 2×2 pixel grid) for a 3D sensor. A 50×50 µm^2 pixel cell is highlighted in red.

case the overlying metal covers the whole implant (field plate design). This design aims to reach a more uniform electric field in the inter implant regions, in order to increase the breakdown voltage. However, the increased metal layer might increase the coupling between neighbour pixels and hence the cross-talk effect. In order to reduce this effect as much as possible, the metal layer is kept as far as possible from the the bump pad of the neighbour pixel. The EBTFP sensor shares the same characteristics of the BTFP sensor, however the n^+ implant is slightly larger (14.5×91.5 µm^2). Also in this case, the larger implant was made in order to possibly increase the breakdown voltage, while also decreasing the inter-pixel regions that might have a lower hit detection efficiency.
The current-voltage curves for all sensors were tested on wafer using the temporary metal. The breakdown voltage was found to be higher than 300 V for most of the sensors. No significant difference between the various pixel designs has been found in terms of the breakdown voltage.
Section 3.3 reports the cross-talk studies for all the different 25×100 µm^2 pixel designs, while Section 5.2 reports a performance comparison.

2.5.2 3D Batches

The two 3D batches will be referred to as 3D-1 and 3D-2 in the following. The 3D-1 batch was produced in 2017 using the the Mask-Aligner photo-lithographic technology, while the 3D-2 batch was produced in 2019 using the Stepper-and-Repeat photo-lithographic technology (with a single sided process in both cases). The latter has a higher, sub-µm, precision but it is less flexible with respect to the Mask-Aligner. When there is the need to

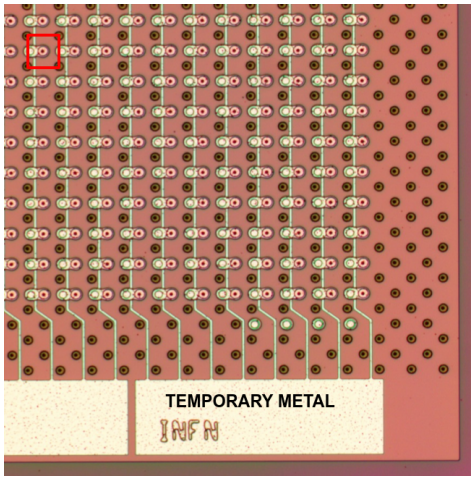


Figure 2.32: Photograph of a 50×50 µm^2 3D pixel sensor on wafer: the temporary metal is still present. A 50×50 µm^2 pixel cell is highlighted in red.

have different cell and sensor designs on the same wafer layout, they can be patterned in one go on the full wafer by the Mask Aligner. The same is not possible with the Stepper.
The diameter of n^+ and p^+ columns of 3D sensors fabricated by FBK is 5 µm. Aside from the different production technique, both batches featured the same design for 50×50 µm^2 and 25×100 µm^2 3D pixel sensors.
Figure 2.31 shows the layout of 50×50 µm^2 pixels (in a 2×2 pixel grid) for a 3D sensor. The pixel cell is delimited by the p^+ columns, while the n^+ column is placed in the center. The bump pad is placed to the left of the n^+ column, with a metal route for electrical contact. The bump pad is not placed on top of the n^+ columns since bump bonding could be more problematic in this configuration.
Figure 2.32 shows a photograph of a 50×50 µm^2 3D pixel sensor on wafer. At the periphery of the pixel matrix, four lines of p^+ columns are placed to control the electric field in the border regions (functioning as a guard ring for 3D sensors).
Figure 2.33 shows the layout of 25×100 µm^2 pixels (in a 4×1 pixel grid) for a 3D sensor. Also in this case, the pixel cell is delimited by the p^+ columns, while the n^+ column is placed in the center. The bump pads are placed in diagonal with respect to the n^+ column, in order to match the 50×50 µm^2 pixel pitch of the readout chip. As shown in Section 3.3, the cross-talk induced by this configuration is negligible.
Finally, Figure 2.34 shows a photograph of a wafer from the 3D-1 batch.

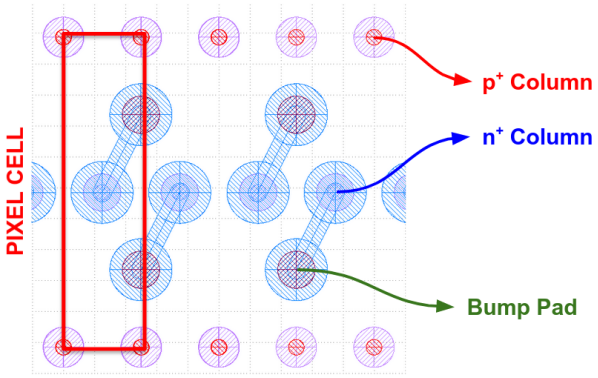


Figure 2.33: Schematic drawings of 25×100 µm^2 pixels (in a 2×2 pixel grid) for a 3D sensor. A 25×100 µm^2 pixel cell is highlighted in red.

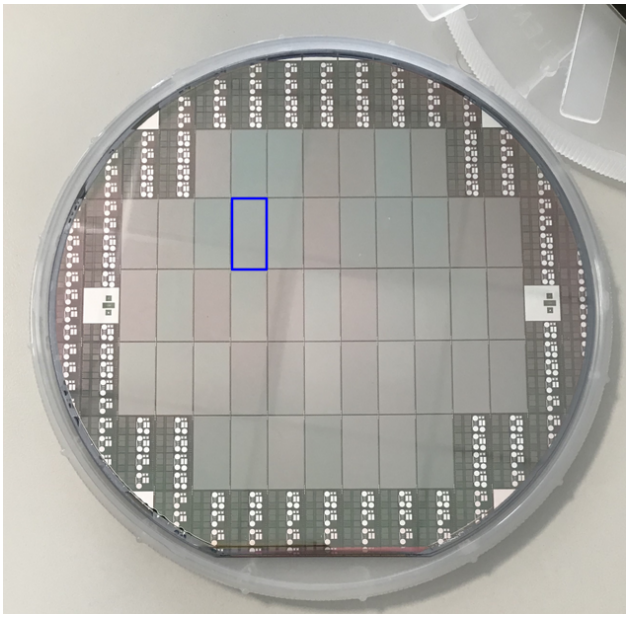


Figure 2.34: Photograph of a wafer from the 3D-2 batch. A pixel sensor to be diced and bump bonded to the readout chip is highlighted in blue.

Chapter 3

Pixel Readout Electronics

3.1 The RD53A ROC

The pixel ROC is one of the most critical component of the future CMS Inner Tracker. Among the requirements are an extreme radiation tolerance, low thresholds and noise (in order to detect signals from thin sensors) and a high channel density. The RD53 collaboration is a joint project supported by CMS and ATLAS collaborations for the development of the ROC to be used by both experiments in the HL-LHC. Developed in 65 nm CMOS technology, the project started in 2016. Every pixel from the silicon sensor is DC coupled to the relative readout channel via bump bondings. The collected charge is amplified, shaped, and digitized using the Time Over Threshold (TOT) method: the time during which the analogue pulse exceeds a certain threshold is digitized and taken as a measure of the deposited charge. Hits are stored for at most 12.5 µs (the L1 trigger latency) in the pixel array within multi-pixel regions.
RD53A is the first prototype of the future ROC, and has been available to the CMS and ATLAS collaborations since late 2017 [38]. The pixel matrix is organised in 192 rows and 400 columns, with a pixel area of 50×50 µm^2 and a total active area of 20×9.6 mm^2 (about half the size of the final CMS ROC). Figure 3.1 shows the geometry of RD53A. The pixel matrix is organised in 8×8 pixel cores: inside each core the 64 pre-amplification circuits are placed as 16 analog "islands" with four circuits each. The digital circuitry is placed in the"sea" surrounding the analog islands, as shown in Figure 3.2.
The chip periphery, placed at the bottom of the ROC (below the pixel matrix), contains the wire bonding pads as well as the Analog Chip Bottom (ACB) and the Digital Chip Bottom (DCB) macro blocks. The ACB groups the analog building blocks, such as Analog to Digital Converters (ADCs),

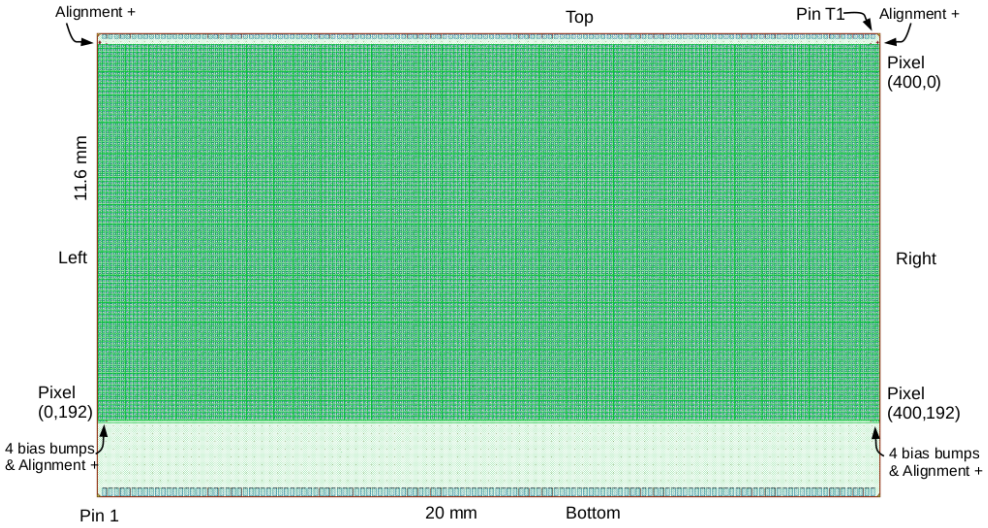


Figure 3.1: The RD53A geometry: the ROC dimensions are 20 mm × 11.8 mm, with a pixel matrix of 400 × 192 pixels, corresponding to an active area of 20 mm × 9.6 mm [38].

Digital to Analog Converters (DACs), temperature sensors, etc. The DCB implements the input, output and configuration digital logic. The wire bonding pads are organized as a single row and are separated from the first row of bump bondings by 1.7 mm, in order to allow for wire bonding when a sensor is bump bonded to the ROC.
In the future CMS Inner Tracker the ROC will be powered using the Shunt-LDOs described in Section 2.4.1. The two SLDOs present on the ROC separately provide VDDD (the internal voltage for the digital domain) and VDDA (the internal voltage for the analog domain). This powering mode requires a current generator: the current that is not used by the ROC is dissipated by the shunts. The power dissipation is therefore critical and a cooling system is always necessary.
In the RD53A ROC, being a prototype, other powering options are available. For instance, it is possible to bypass the shunts and only use the LDOs. This is the usual operational condition in laboratory, since the required power is far lower (the LDOs are usually powered in parallel with typical voltage and current values of ∼1.8 V and ∼600 mA respectively).
It is also possible to bypass the LDOs and directly provide VDDD and VDDA (the nominal values are 1.3 V and 1.2 V respectively): this is however not advised. Indeed, the LDOs function as a protection, so that input voltage drops are not transmitted to the ROC. Moreover, the LDOs are usually op-

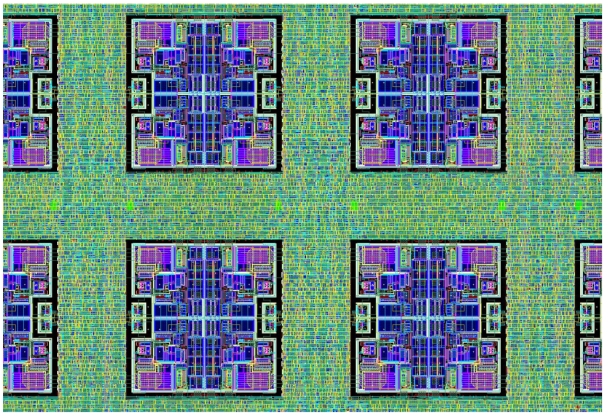


Figure 3.2: Layout of four analog islands (corresponding to 16 pixels) surrounded by the digital sea [38].

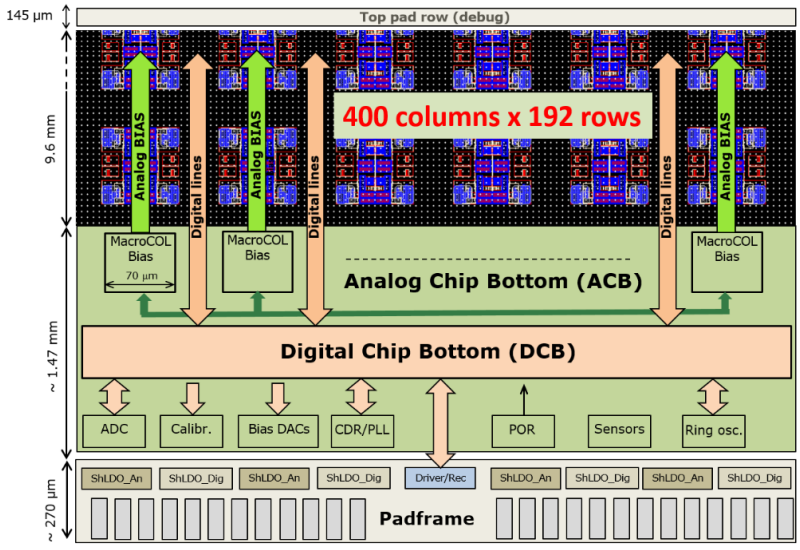


Figure 3.3: The RD53A ROC functional view. In the bottom the wire bonding pads are visible [38].

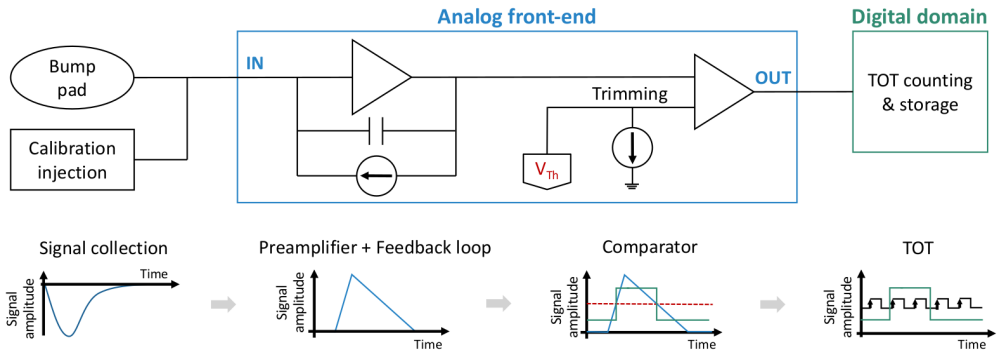


Figure 3.4: Signal processing steps for a generic AFE, from charge collection to digitalisation [39].

erated at 1.8 V, but can accept up to 2 V, whereas the ROC can be damaged if VDDD or VDDA exceed 1.35 V.
The analog to digital conversion is performed in each pixel channel by the Analog Front-Ends (AFEs), as shown in Figure 3.4. An AFE contains a charge sensitive pre-amplifier, a threshold discriminator, a threshold trimming circuit to address pixel to pixel variation of the threshold voltage, and the TOT counting of the input signal amplitude. In the RD53A ROC the TOT digitalisation (with 4 bit resolution) is done with respect to rising edges of the 40 MHz LHC clock. Therefore, one TOT unit corresponds to 25 ns. A feedback circuit takes care of the signal return to the baseline and compensates the leakage current from the sensor.
The chip also features a circuit for the generation of internal calibration charge injection signals, as visible in Figure 3.4. The circuit, connected to the input of the pre-amplifier of each pixel channel, enables the injection of a well-defined and programmable calibration charge to test the front-end functionalities and calibrate the chip response. The calibration injection circuit uses the difference of two voltages: this allows a precision differential voltage that will be independent of local ground drops in the chip. This differential voltage ΔV_{cal} is applied on an injection capacitance. The conversion between V_{cal} units and electrons was measured to be:

$$Q[e^-] = Q[\Delta V_{cal}] \cdot (10.40 \pm 0.10)[e^-/\Delta V_{cal}] + (180 \pm 60)[e^-] \qquad (3.1)$$

The RD53A ROC can be configured with dedicated DAC registers. A global reference current, called I_{ref}, is responsible for scaling the currents that are produced by all the bias current DACs. The reference voltage V_{ref} is used as reference for the DACs used for injection pulses. The nominal values of I_{ref} and V_{ref} are 4 µA and 900 mV, respectively.

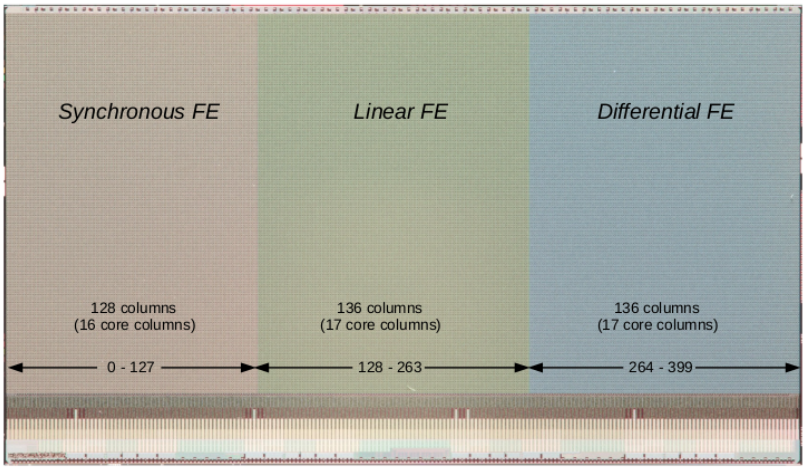


Figure 3.5: Disposition of the three AFEs with respect to the pixel matrix [38].

Three different types of AFEs are present on RD53A, as shown in Figure 3.5 The chip is divided vertically into three sections, each one having one type of AFE. The three designs are referred to as Differential, Linear and Synchronous: they are not small variants of a common design, but substantially different projects. They were placed in the same chip to allow a detailed performance comparison.
After extensive measurements, the CMS collaboration chose the Linear AFE for its final ROC [39]. Various improvements with respect to the Linear AFE of RD53A were implemented by the CMS collaboration in the next iteration of RD53A, named CROC (CMS ROC). The CROC chip went in production in mid 2021, therefore it not object of this Thesis. All the results reported in this Thesis are relative to the Linear AFE of RD53A.

3.1.1 The Linear AFE

The Linear AFE implements a linear pulse amplification in front of the discriminator. The pre-amplifier of the Linear AFE features a Krummenacher feedback, which ensures both the sensor leakage current compensation and the discharge of the feedback capacitor. The Krummenacher current I_{Krum} can be adjusted with the relative DAC register (called KRUM_CURR): increasing I_{Krum} results in a faster pre-amplifier return to baseline and a reduced TOT as schematically shown in Figure 3.6.
An important consideration for a highly efficient particle detector is the event loss due to the pixel channels dead time, especially at high luminosity and

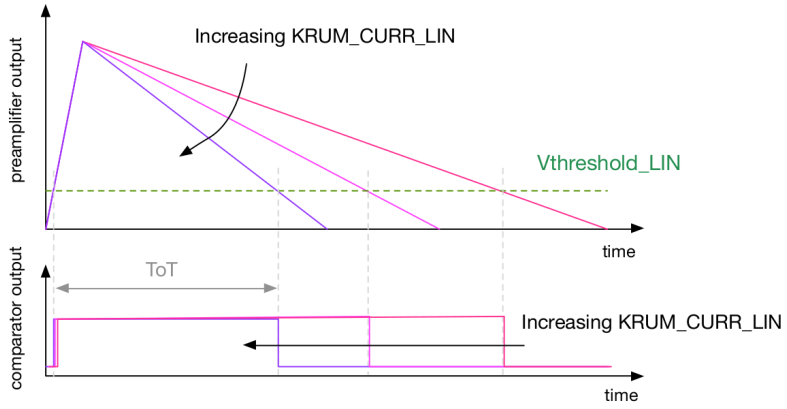


Figure 3.6: Sketch of the effects of the KRUM CURR parameter on the pre-amplifier output waveform and corresponding TOT.

high pile-up conditions. In the innermost layer of the future CMS tracker, a maximum dead time of 1% is required. The dead time depends on the TOT response, since every TOT unit correspond to 25 ns. Therefore, I_{Krum} needs to be regulated depending on the tracker layer requirements: a faster pre-amplifier discharge leads to a reduced dead time.
The signal from the pre-amplifier is fed to a low power threshold discriminator, which compares the signal with the threshold. The V_THR DAC register sets the global threshold of the Linear AFE, corresponding to the threshold voltage applied to the discriminator input. Increasing V_THR results in an increased global threshold. In each pixel channel, a 4-bit Trimming DAC (TDAC), allows to reduce threshold dispersion across the pixel matrix.
Finally, the PA_BIAS register regulates the current of the pre-amplifier input branch. This current represents the main contribution to the Linear AFE current consumption and determines the speed of rise of the signal.

3.1.2 Single Chip Cards

In order to test pixel sensors bump bonded to the RD53A ROC, dedicated Printed Circuit Boards (PCBs) have been developed by the CMS collaboration. They are usually referred to as Single Chip Cards (SCCs) since they only host one ROC (while in the future CMS Inner Tracker a single, large, pixel sensor will be bump bonded to two or four ROCs, as explained in Section 2.4.1). The SCCs are useful for testing both the ROCs and the sensors. Two main types of SCCs have been employed by the CMS collaboration, which will be referred to as Type-A or Type-B.
The SCC Type-A, shown in Figure 3.8, was designed by the CMS group of

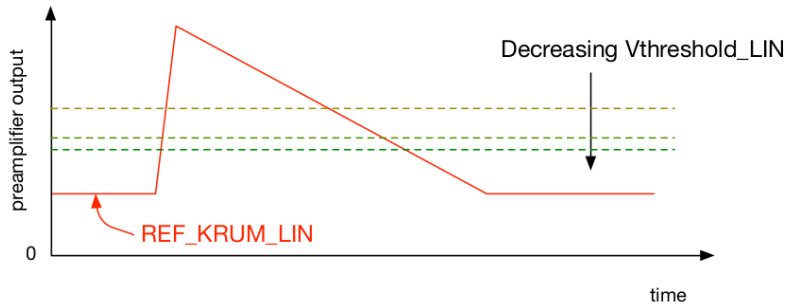


Figure 3.7: Sketch of the effects of V_THR on the global threshold.

Zurich University (UZH). The pixel detector (i.e. the assembly of the ROC and the sensor) is glued to a an aluminium frame which is then glued to the PCB. The PCB presents a hole in correspondence to the aluminium frame: this permits to cool the ROC with a refrigerator system. Moreover, the limited amount of material under the detector minimizes the activation under irradiation[1] and multiple scattering effects if the pixel detector is tested on a particle beam.
Wire bondings connect the ROC to the PCB. Many additional elements are welded on the PCB: powering connectors, monitoring terminals, DisplayPort connectors for sending and receiving data, configuration jumpers etc., allowing for a full control of the chip. The metallised backside of the sensor (which is the visible surface of the detector placed on the SCC) is connected to the bias voltage with wire bondings and a dedicated circuit (visible in the upper left side of Figure 3.8). All the wire bondings are usually encapsulated with silicone elastomer to be protected.
The Type-B SCC, designed by the CMS group of Rice University (Houston, Texas), is shown in Figure 3.9: it is a cheap and lightweight PCB. The card region below the pixel detector is free from material. The Type-B SCC has to be connected to its own adapter card in order to allow for readout by the DAQ system in laboratory. The Type-B SCC is particularly suitable for irradiations, since all the additional components are placed on the adapter card (which is not irradiated).
All the pixel sensors tested and characterised for this Thesis were bump bonded to the RD53A ROC and glued and wire bonded to one of the two types of SCCs, depending on availability. In particular, pixel sensor wafers

[1]Both the ROC and sensor need to be characterised after receiving fluences similar to those expected in the future CMS Inner Tracker. The RD53A ROC is certified to resist to a Total Ionising Dose of 5 MGy.

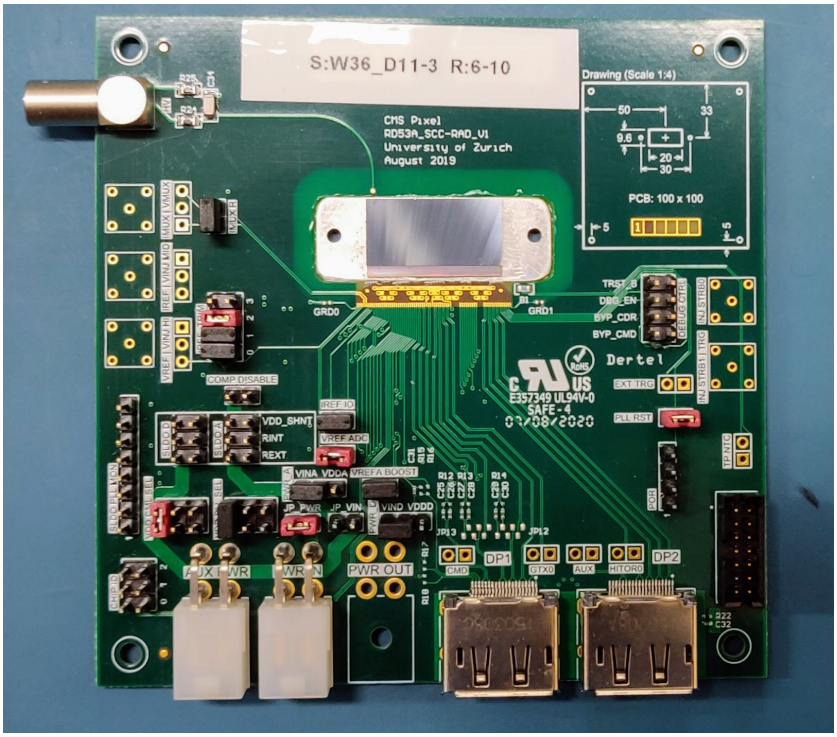


Figure 3.8: Pixel sensor and RD53A ROC assembly glued and wire bonded onto a Type-A SCC. The light-gray front surface is the metallised sensor back-side.

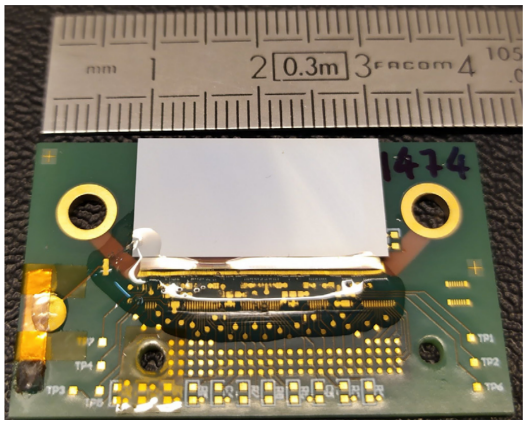


Figure 3.9: Pixel sensor and RD53A ROC assembly glued and wire bonded onto a Type-B SCC. Wire bonds are encapsulated with silicone elastomer to be protected. The light-gray front surface is the metallised sensor back-side, where high voltage bias wire bondings are connected in the lower left corner.

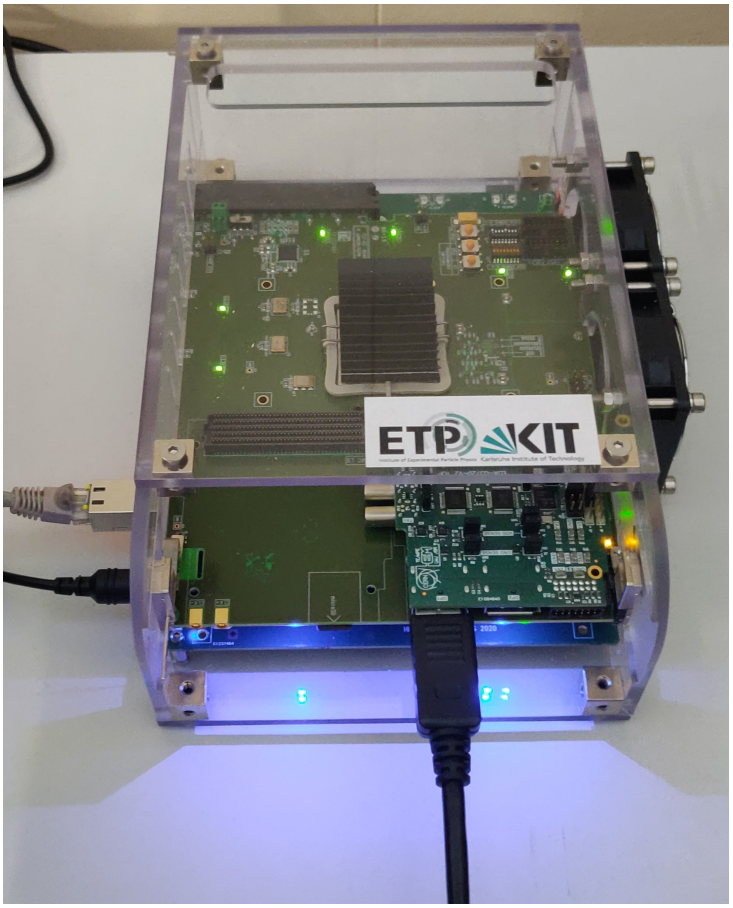


Figure 3.10: The FC7 board operated on the nano-crate. An custom FMC card allows the electrical communication with the ROC through a DisplayPort cable.

(both planar and 3D) were processed for UBM (Under Bump Metalization), thinned, diced and bump bonded to the RD53A ROC at the Fraunhofer-Institut IZM (Berlin), while the wire bondings were performed at the INFN laboratories in Florence.

3.1.3 DAQ Systems

Two different DAQ systems to readout the RD53A ROC were used for this Thesis: BDAQ53 and Ph2ACF. Both systems allow an extensive configuration and provide a large variety of tests and calibrations to characterise the pixel detectors. Most of the calibrations are similar for both systems, and will be presented in the Section 3.2.

BDAQ53 was the first available DAQ system when the RD53A ROC was available for testing in late 2017 [40]. It was developed by Bonn University and consists of a firmware and a software.
The BDAQ53 firmware, written in Verilog, is compatible with the Kintex7 Field Programmable Gate Array (FPGA) [41]. The CMS collaboration used the commercially available Xilinx KC705 development board, which houses this FPGA. A custom FMC[2] card was developed by CERN to connect the board to a SCC with electric links through DisplayPort cables. The FPGA communicates with the PC via an Ethernet link.
The BDAQ53 software[3] consists in a Python based modular framework. Data received from the ROC are written to the PC in the HDF5 format, a data format for efficient storage of large amounts of scientific data.
The Ph2ACF (Phase2 Acquisition and Control Framework) DAQ system is developed by the CMS collaboration. The hardware consists on a custom board, named FC7, based on a Kintex7 FPGA. The FC7 is designed to be an AMC[4] compatible with the µTCA standard. It can be operated both on a crate or on a specifically designed nano-crate for desktop usage (this is the typical case in laboratory), with an Ethernet link to communicate with the PC. The FC7 presents two FCM connectors that can host custom FMC cards for electrical communication with the ROC through DisplayPort cables, as shown in Figure 3.10. Other custom FMCs can provide optical communication via LpGBT and external triggering and clocking.
The firmware developed for the FC7 is called IT-µDTC (Inner Tracker Data Trigger and Control), entirely written in the VHDL language and based on the IPBus protocol, developed at CERN. The firmware can handle both electrical and optical links to SCCs or pixel modules made by two or four ROCs (described in Section 2.4.1). The IT-µDTC will be the foundation of the final DTC system that will be used in the Phase-2 Inner Tracker system of the CMS experiment.
The Ph2ACF software[5] is written in C++ and uses the ROOT software [42], developed at CERN, for data analysis. It allows to configure the ROC DAC registers with XML files and provides a full set of calibration procedures.

[2]FPGA Mezzanine Card (FMC) is an industrial standard that defines I/O mezzanine modules with connections to an FPGA.

[3]`https://gitlab.cern.ch/silab/bdaq53`

[4]Advanced Mezzanine Card (AMC) is an industrial standard for PCBs.

[5]`https://gitlab.cern.ch/cms_tk_ph2/Ph2_ACF`

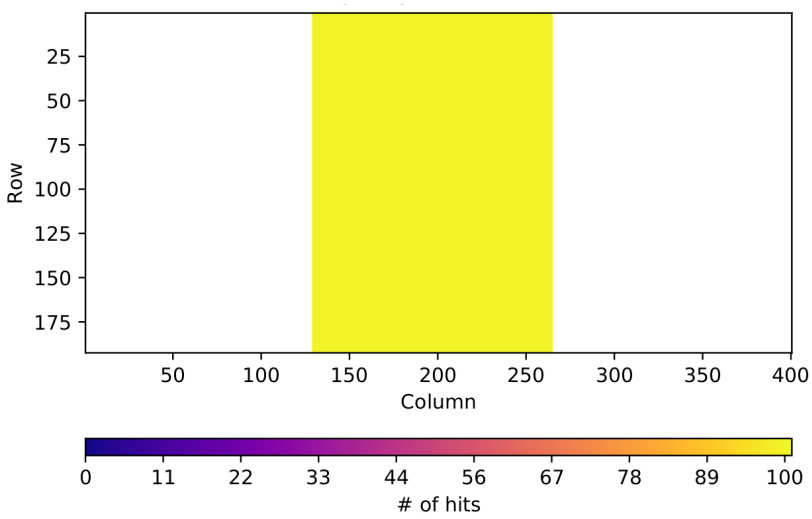


Figure 3.11: Result of the *PixelAlive* test performed on a pixel detector. All the Linear AFE pixels responded with 100% efficiency to the charge injections. The Synchronous AFE and the Differential AFE, placed on the left and right the Linear AFE respectively, were not tested.

3.2 Calibration Procedures

The RD53A ROC needs to be tested and calibrated in laboratory: this is important both for characterising the chip and for testing the pixel detector on a particle beam. Over the course of my PhD I have worked extensively with the RD53A ROC and I tested and calibrated many pixel detectors, both at the Florence INFN laboratories and at CERN.
When calibrating a pixel detector, it is important to bias and fully deplete the silicon sensor, since this is the operative condition (the different capacitance of an undepleted silicon sensor with respect to a fully depleted sensor can modify the ROC response, and hence the calibrations).
In the following, a typical calibration procedure for a pixel detector on a SCC is reported. While using different algorithms, both DAQs presented in the previous Section share the same procedures and lead to comparable results. The ROC is usually operated in LDO mode (i.e. bypassing the shunt), and only the Linear AFE is tested.

- The first step consists in regulating VDDD, VDDA as well as V_{ref} and I_{ref} to their nominal values. This is done by modifying dedicated chip DAC registers or with specific jumpers on the SCC (this is the case for I_{ref}). This sets the RD53A ROC to its working condition.

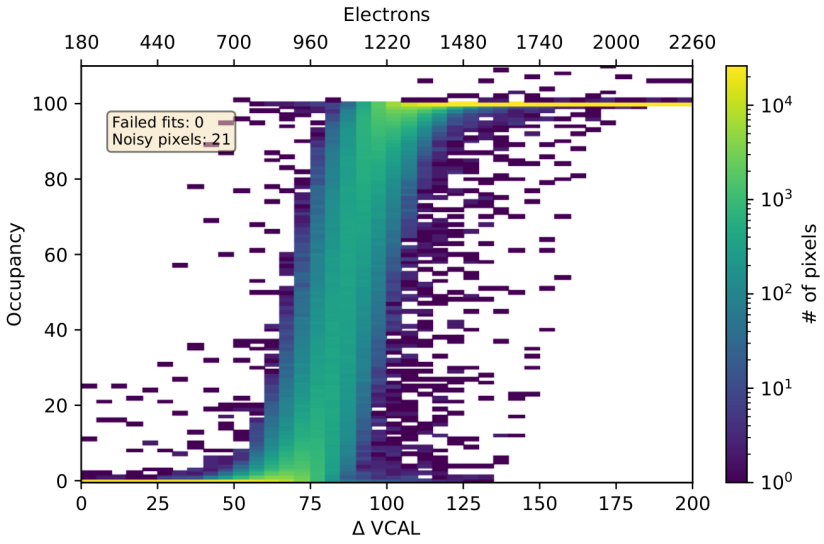


Figure 3.12: Result of the *SCurve* test performed on a pixel detector. The pixel channels are superimposed on the z axis (represented with a coloured scale).

- Some pixel channels in the ROC might be noisy, especially in irradiated detectors. The *Noise* scan readouts all pixels 10^7 times without any injection: pixels with an occupancy higher than the target value (usually 10^{-6}) are masked in subsequent tests.

- In order to detect non-functioning pixel channels, a calibration charge is injected in all pixels (usually 100 times): this test is usually referred to as *PixelAlive*. If the pixel is functioning (and the calibration charge is sufficiently higher than the pixel threshold), it should respond to the injection. Figure 3.11 shows the *PixelAlive* test performed on a pixel detector: in this case all the pixels of the Linear AFE correctly responded to all 100 injections. If the test finds pixels with a low occupancy, they are probably defective, and are masked in subsequent tests.

- In order to determine threshold and noise values for each pixel the so called *SCurves* test is performed. These quantities are evaluated by measuring the pixel channel efficiency as a function of the injected charge ΔV_{cal}. Every ΔV_{cal} value is injected 100 times to measure the efficiency. If the injected charge is close to the pixel threshold, the pixel channel might be activated or not depending on the noise, that sums to the injected charge. For high ΔV_{cal} values, the efficiency is always

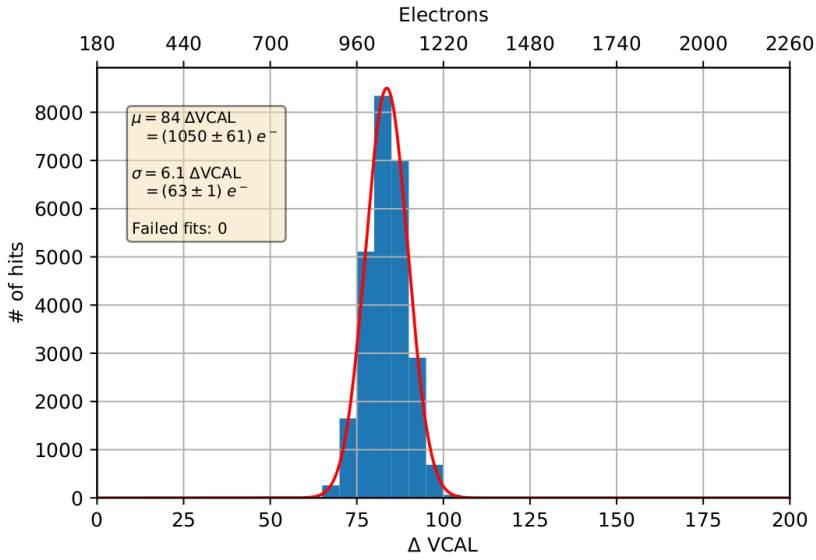


Figure 3.13: Threshold distribution from the *SCurves* test for a pixel detector. The red line represents a fit to a gaussian distribution.

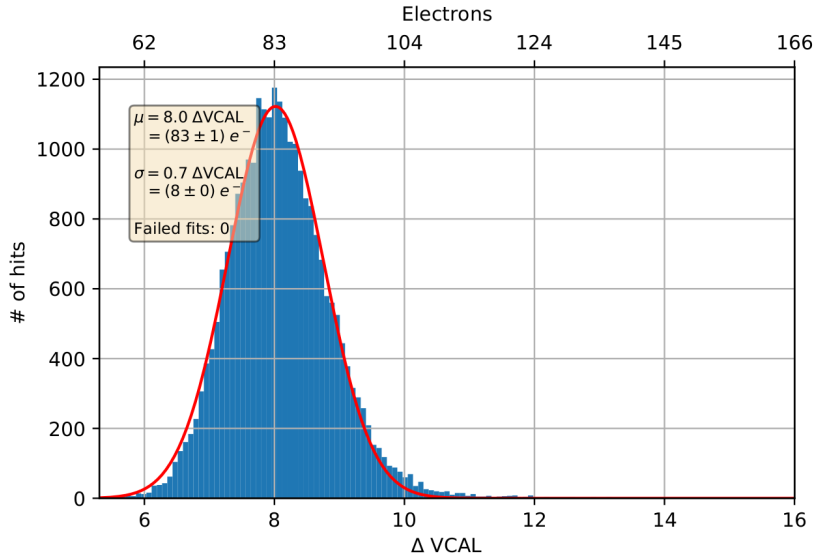


Figure 3.14: Noise distribution from the *SCurves* test for a pixel detector. The red line represents a fit to a gaussian distribution.

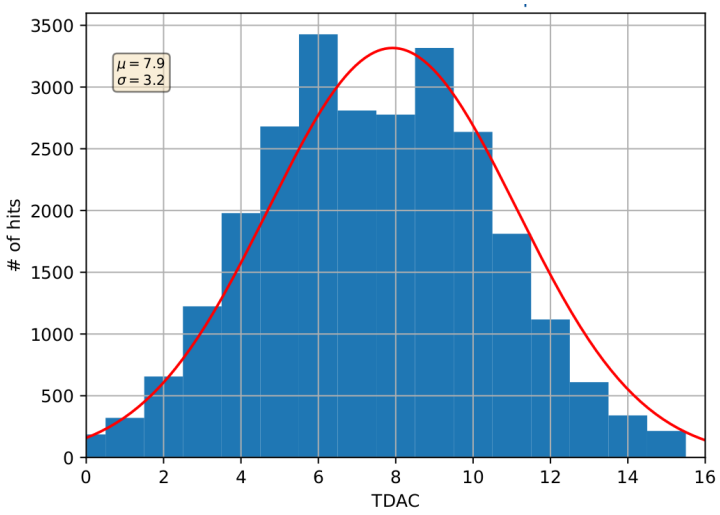


Figure 3.15: TDAC distribution for a pixel detector after the *ThresholdAdjustment* test. The red line represents a fit to a gaussian distribution.

100%. Figure 3.12 shows the *SCurves* test performed on a pixel detector: the pixel channels are superimposed on the z axis (represented with a coloured scale). The injected charge is both in ΔV_{cal} units and in electrons (using Equation 3.1 for the conversion). The efficiency curves have a characteristic "S" shaped profile, hence they are usually referred to as SCurves. The function that better characterise this behaviour is:

$$y = \frac{1}{2}\left(1 + \text{erf}\left(\frac{q-\mu}{\sqrt{2}\sigma}\right)\right) \tag{3.2}$$

where $\text{erf}(q)$ is the error function and q the injected charge. The μ and σ parameters represent threshold and noise respectively. For each pixel channel, μ and σ are derived by fitting the corresponding SCurve with this function [43].
Figure 3.13 and Figure 3.14 show the threshold an noise distributions respectively, for all the pixel channels. The red line represents a fit to a gaussian distribution: the average pixel threshold is about 1050 electrons while the average noise is 83 electrons.

- In order to reach the low average threshold of Figure 3.13, the V_THR chip register (presented in Section 3.1.1) needs to be significantly lowered with respect to the default value. The *ThresholdAdjustment* test lowers the V_THR register while performing *SCurves* tests, in an iterative procedure, until the target threshold is met. Lowering too much

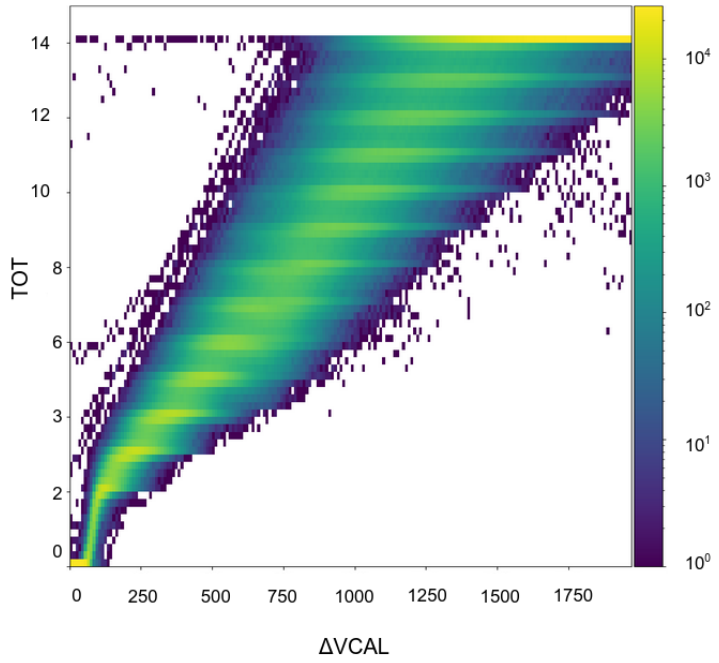


Figure 3.16: Result of the *GainScan* test performed on a pixel detector. The pixel channels are superimposed on the z axis (represented with a coloured scale).

the threshold might drastically increase the noisy pixels: for this reason the *Noise* scan needs to be repeated every time the threshold is lowered. If the the noisy pixels reach 1% of the total pixels, the thresholds should not be further lowered, in order to maintain a reasonable amount of masked pixels.

- The threshold distribution reported in Figure 3.13 is very narrow, meaning the pixel responses are very close to each others. This is obtained with the *ThresholdEqualisation* scan, that regulates the TDAC (presented in Section 3.1.1) for each pixel channel, so that the pixel thresholds are as uniform as possible. Figure 3.15 shows the TDAC distribution for all pixel channels: the distribution is centred around the middle register value $(7-8)$, which is an indication of the goodness of the equalisation. If the distribution is not symmetric, the calibration might not be optimal. The *ThresholdEqualisation* and the *ThresholdAdjustement* tests need to be performed iteratively, progressively aiming to lower thresholds.

- The *GainScan* test measures the TOT response for every pixel channel as a function of the injected charge ΔV_{cal}, as shown in Figure 3.16. The pixel channels are superimposed on the z axis (represented with a coloured scale). Since the TOT is a 4-bit register, the TOT range goes from 0 to 14 (where 14 is the saturation[6]). During the *GainScan* test, every ΔV_{cal} value is injected 100 times in a given pixel channel: the average TOT is calculated and reported on the y axis. Since a pixel channel will not always respond with the same TOT to a given ΔV_{cal} injection, the average TOT values reported in Figure 3.16 are not integers.
  The TOT response varies drastically across the pixel matrix, especially for high ΔV_{cal} values: this variation can not be corrected and can only be accounted for. Indeed, a lookup table can be built by evaluating the most probable ΔV_{cal} value for each (integer) TOT value. This table contains the TOT $\rightarrow \Delta V_{cal}$ conversion for every pixel channel. A reliable charge calibration, that is a TOT $\rightarrow$ electrons conversion for each pixel channel, can be performed by combining this lookup table with Equation 3.1.

- The slope of the distribution reported in Figure 3.16 can be optimised by regulating I_{Krum}, as explained in Section 3.1.1. The *GainOptimisation* test aims to a target TOT value for a specific ΔV_{cal} injection, while scanning the KRUM_CURR register. If the calibration was successful, the average pixel response to the specific ΔV_{cal} injection should be the target TOT (albeit with a considerable dispersion across the pixel matrix). This effectively changes the slope of the distribution of Figure 3.16.
  As mentioned in Section 3.1.1, this scan is required for optimising the dead times of the pixel detectors. The requirements are different for the various layers of the future CMS Inner Tracker. The scan is also useful in other applications, such as studies with radioactive sources in laboratory, in order to optimise the TOT range to the measurements to be performed.

3.3 Cross-Talk Studies

As explained in Section 2.5, 25×100 µm^2 pixel sensors were adapted in order to be readout by the RD53A ROC, which has a pixel pitch of 50×50 µm^2 However, the arrangement of the bump pads on 25×100 µm^2 pixel sensors

[6] A TOT value of 15 is associated with a readout error.

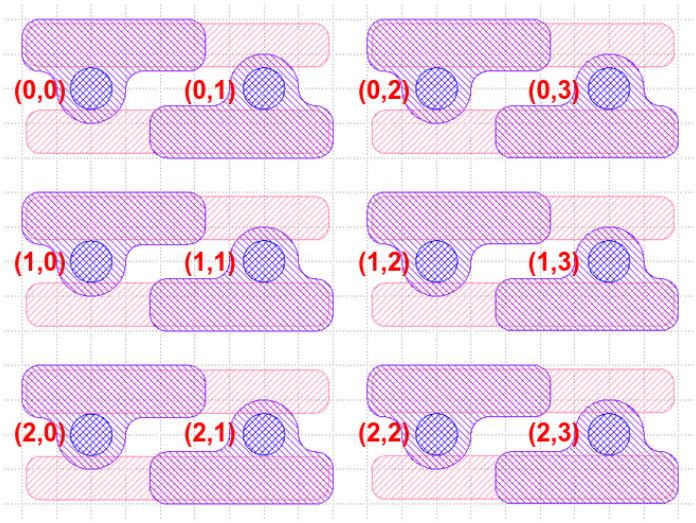


Figure 3.17: 2×6 pixel matrix for a 25×100 µm^2 planar pixel sensor. The bump pads are enumerated with the corresponding RD53A readout channel.

causes a cross-talk effect between adjacent pixels. I characterised the various designs of FBK 25×100 µm^2 pixel sensors, both 3D and planar, in order to quantify the cross-talk effect for the various pixel designs. The measurement of the cross-talk is performed by injecting a large amount of charge in one pixel: a fraction of this charge (referred to as $X[\%]$) will be induced to the neighbour coupled pixel.
In order to make a clean measurement of this effect, a particular injection pattern was adopted. Figure 3.17 shows a 2×6 pixel matrix for a 25×100 µm^2 planar pixel sensor. The bump pads are enumerated with the corresponding RD53A readout channel. Focusing on the first column of 25×100 µm^2 pixels, pixel (0,0) is injected and both pixels (0,0) and (0,1) are readout: a cross-talk effect is expected between the two. In the following, the (0,0) pixel is referred to as the primary pixel, while (0,1) as the secondary pixel (i.e. the pixel with an expected cross-talk effect). The same applies for pixels (2,0) and (2,1), while pixels (1,0) and (1,1) are not injected nor readout[7]. The same injection pattern is applied to the other columns of 25×100 µm^2 pixels[8].
The analysed pixel detectors were calibrated to an average pixel threshold of about 1000 electrons. An *SCurve* test was performed with the aforementioned injection pattern, and up to very high injected charge values (about

[7]This choice was made in order to avoid second order cross-talk effects, for instance between pixels (0,1) and (1,0).
[8]Cross-talk between 25×100 µm^2 pixel columns is expected to be negligible.

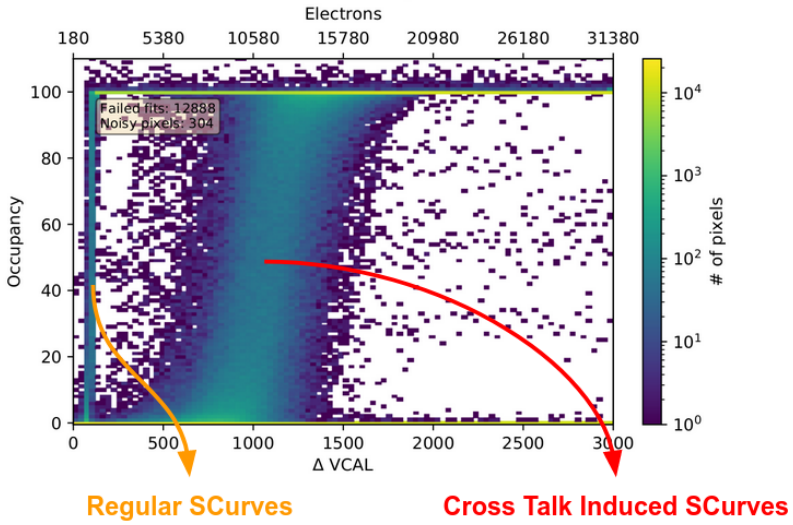


Figure 3.18: *SCurve* test performed by injecting up to about 30000 electrons on a 25×100 µm^2 planar pixel sensor. Two SCurve distributions are visible: the first one is related to the injected pixels, while the second one is related to the cross-talk coupled pixels.

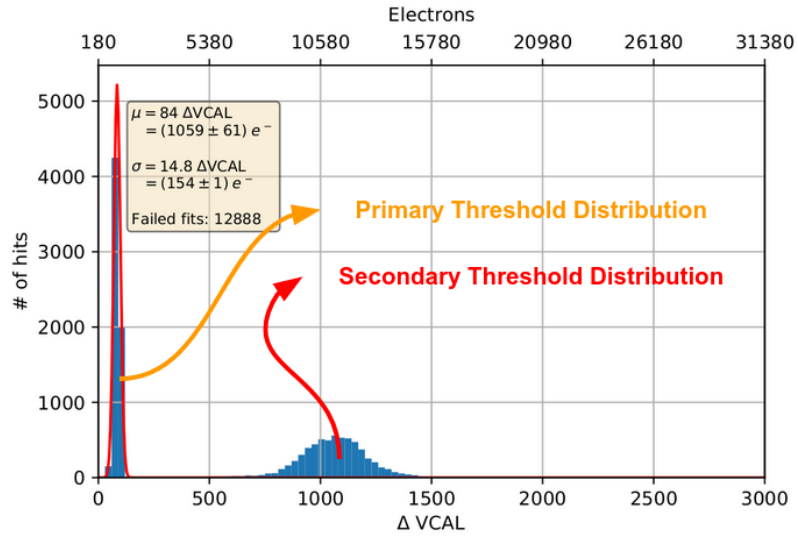


Figure 3.19: Threshold distribution for 25×100 µm^2 planar pixel sensor. The first distribution is related to the injected pixels, while the second one is related to the cross-talk coupled pixels.

30000 electrons). As shown in Figure 3.18, two different SCurves are visible: the first one is the usual SCurve (i.e. relative to the primary pixel), while the second one is due to the cross-talk effect (i.e. relative to the secondary pixel). For high values of the injected charge, for which the secondary pixel activates, the primary pixel turns on 100% of the times. Figure 3.19 shows the corresponding threshold distributions for the two sets of SCurves: one is the primary threshold distribution, while the other one is the secondary threshold distribution caused by the cross-talk effect.
A gaussian fit is performed on both distributions: μ_1, σ_1 and μ_2, σ_2 are the fit parameters for the primary and secondary threshold distributions respectively. Both μ_1 and μ_2 are influenced by cross-talk. Indeed, if μ_1 is injected in the primary pixel, only a fraction $(100 - X[\%])$ is actually readout. In other words, the nominal threshold of the primary pixel is $\mu_1(100 - X[\%])$. Since the primary and the secondary pixels were tuned together, they are expected to have the same nominal threshold, that is:

$$\mu_2 \cdot X[\%] = \mu_1 \cdot (100 - X[\%]) \tag{3.3}$$

By defining r as the ratio of μ_1 and μ_2 the cross-talk value is given by:

$$X[\%] = \frac{r}{1 + r} \tag{3.4}$$

The dispersion of the secondary threshold distribution can be large, due to the non-uniform response of the pixel channels. Therefore, errors on the $X[\%]$ are estimated by evaluating the ratio $r_{\pm}$ between μ_1 and $\mu_2 \pm \sigma_2$ and substituting in Equation 3.4.
Table 3.1 reports the measured cross-talk for the various designs of planar pixel sensors (described in Section 2.5.1) as well as for a 3D pixel sensor. All the pixel detectors were tuned to average threshold of about 1000 electrons and the sensors biased with at least 10 V above full depletion. The standard (i.e. without the bitten design) 25×100 µm^2 planar pixel detector has the higher cross-talk value. The three variations of the bitten design show comparable measured cross-talk values, which are systematically lower with respect to the standard (non-bitten) pixel design, proving the effectiveness of the bitten design. The cross-talk on 3D pixel detector is lower with respect to all planar pixel detectors. This is expected due to the peculiar configuration of the implants on 3D sensors.
It should be noted that small differences in $X[\%]$ determine whether or not the cross-talk effect is actually visible if testing a pixel detector on a particle beam. Considering a pixel sensor with active thickness of 150 µm, and MIPs traversing it, the Landau distribution of the collected charge has an expected

Pixel Sensor Type	**X**[%]
Planar Standard	$11.5^{+3.0}_{-2.0}$
Planar BT	$8.5^{+2.0}_{-2.0}$
Planar BTFP	$9.0^{+2.0}_{-1.5}$
Planar EBTFP	$8.5^{+2.0}_{-1.5}$
3D	$5.5^{+0.5}_{-0.5}$

Table 3.1: Cross-talk measured on various types of 25×100 µm^2 pixel sensors (both planar and 3D). BT stand for BiTten, BTFP for BiTten Field Plate, EBTFP for Extended BiTen Field Plate. The standard PA_BIAS register value (350) was used for these results.

Pixel Sensor Type	**X**[%]
Planar Standard	$9.5^{+3.0}_{-2.0}$
Planar BT	$7.0^{+2.0}_{-1.5}$
Planar BTFP	$7.5^{+2.0}_{-1.5}$
Planar EBTFP	$7.0^{+1.5}_{-1.5}$
3D	$5.0^{+0.5}_{-0.5}$

Table 3.2: Cross-talk measured on various types of 25×100 µm^2 pixel sensors (both planar and 3D). BT stand for BiTten, BTFP for BiTten Field Plate, EBTFP for Extended BiTen Field Plate. A slightly higher PA_BIAS register value (410) was used for these results.

MPV of about 11400 electrons. If $X[\%] = 11\%$, 1250 electrons are induced in the coupled pixel, while if $X[\%] = 6\%$, only 680 electrons are induced. If the thresholds were tuned to an average value of 1000 electrons, in the first case the cross-talk effect is visible, while in the second case it is not[9]. For the same reason, it should be noted that while $X[\%]$ does not not depend on the average pixel threshold, if the chip is tuned to a high threshold the cross-talk effect is not visible (i.e. the coupled pixels do not activate).
The Linear AFE registers can be slightly adjusted in order to further reduce the cross-talk effect. In particular, increasing the pre-amplifier bias current is very effective, since it changes the input impedance of the readout channels. However, increasing this current significantly increases the power consumption[10]. The compromise was found for a slightly larger value for the PA_BIAS register (presented in Section 3.1.1), which has an acceptable impact on power consumption.
Table 3.2 shows the measured $X[\%]$ with the slightly larger value for the PA_BIAS register. The cross-talk is significantly reduced for all pixel designs. For this reason, the larger pre-amplifier bias current is now used as default by the CMS collaboration.

[9]This is the case by only considering the Landau MPV. If a larger charge is produced,the cross-talk effect is visible with both thresholds.

[10]Increasing too much the power consumption could be problematic for the overall powering scheme of the future CMS Inner Tracker.

Chapter 4

Test Beam Measurements

4.1 Tracking Concepts

The main goal of a tracker in high energy physics is to reconstruct the trajectories of the charged particles passing through it. The first step consists in the reconstruction of the local hits on the tracker layers. These informations are then combined in tracks, that are the reconstructions of the particles trajectories. In the following, a brief summary of track reconstruction, starting from the local hits, is reported.

4.1.1 Cluster Reconstruction

When a particle passes through a pixel[1] detector, charge is deposited below one or more pixels: charge sharing is due to diffusion, by Lorentz drift if the detector is in a magnetic field, or by the incident particle traversing the detector with a certain angle. The cluster reconstruction is the procedure of grouping these pixels (provided that the readout electronics channels are above threshold). In the following, cluster size indicates the number of pixels in a cluster.
From this collection of points, the particle impact point on the detector can be reconstructed. Since pixel detectors can usually measure the collected charge, weighted position calculations are possible, which significantly improve the position resolution. Many algorithms have been studied and can be more or less effective, depending of the cluster type [44].
The most common algorithm in cluster reconstruction is the Center of Grav-

[1] Many of the following concepts also apply to strip detectors.

ity (CoG). The cluster center (in one dimension) is simply given by:

$$x_{cog} = \frac{\Sigma_i Q_i x_i}{\Sigma_i Q_i} \tag{4.1}$$

where Q_i is the charge collected by pixel i, which has a coordinate x_i. In this case the algorithm can be used for all cluster sizes.
The Head-Tail algorithm is used in case of very long clusters. The central pixels in such clusters do not contribute significantly to the information provided by the first and the last pixels in the cluster. The position is therefore calculated as:

$$x_{ht} = \frac{x_{head} - x_{tail}}{2} + \frac{Q_{tail} - Q_{head}}{2Q_{average}} \cdot p \tag{4.2}$$

where x_{head} and x_{tail} are the positions of the first and the last pixels respectively, Q_{head} and Q_{tail} the respective collected charge, $Q_{average}$ is the average pixel charge of the cluster and p is the pixel pitch. The first term only uses the positions of the first and the last pixels, while the second term takes into account charge information.
The CMS experiment uses more advanced algorithms, such as the Template Matching [45]. The first step is a detailed simulation of the detector response based on the particle incidence angle and the radiation damage. Then, the reconstruction algorithm compare the full distributions of the observed cluster charge to the expected distributions from the simulations at different hit positions.
After a cluster has been reconstructed, its coordinates are usually converted from the local coordinates to a global frame of reference, so that the track from the traversing particle can be reconstructed. These cluster three dimensional space positions are referred to as hits.

4.1.2 Trajectory Reconstruction

Different types of algorithms have been developed to reconstruct the particle trajectories from the hit positions in the global frame of reference. If no magnetic field is present, a simple straight line is often sufficient. However, depending on the experimental setup, multiple scattering effects needs to be considered. If a magnetic field is present, as in the CMS experiment, the particle trajectory is described by a helix with curvature $k = 1/r$, where r is determined by the p_T of the particle and the magnetic field strength. Some of the most common reconstruction algorithms are reported in the following.
The triplet method is used when three hits from three different layers of a tracker are available. First, a straight line fit is performed using two of the

hits. The resulting track is interpolated (or extrapolated) to the position of the third hit. The residual is then defined by:

$$\Delta x = x_{track} - x_{meas} \tag{4.3}$$

which is the distance between the measured hit and the interpolation of the reconstructed track. The residual can be used to select good tracks: if the residual is above a certain threshold Δx_{thr}, the track is rejected, because it was likely reconstructed out of uncorrelated hits.
The General Broken Lines (GBL) is a more sophisticated algorithm that takes into account Coulomb scattering that may randomly change the particle trajectory [46, 47]. While the mean deflection is zero, its variance depends on the traversed material and on the particle energy. The variance of the deflection angle θ of a single particle with momentum p and velocity β is given by [1]:

$$V[\theta] = \left(\frac{13.6\,\text{MeV}}{\beta p} \right)^2 \cdot t \cdot [1 + 0.038 \cdot \ln(t)]^2 \tag{4.4}$$

where t is the scatterer thickness in terms of its radiation length X_0. The scattering materials are represented by zero-thickness scatterers between the hits. The GBL algorithm is able to include the scattering material in the re-fit of an already selected track, referred to as seed. The trajectory is then built from the seed, the hits, and the scatterers: the fit parameters are determined using a χ^2 minimisation of these contributions.
The GBL is mathematically equivalent to the Kalman filter [48], which adds hits and scatterers one-by-one while updating the fit. The advantage of the GBL is that it provides the global covariance matrix for each track, which is needed in alignment methods.

4.1.3 Alignment Procedure

The resolution of modern pixel detectors is far better than the mechanical precision of the the detector itself. Therefore, misalignments have to be corrected during data analysis. Various algorithms have been developed: while the approaches are different, all the algorithms use the detector data to perform a minimisation of the deviation between the track fit and the measurements.
The CMS experiment uses the MILLEPEDE-II (MP-II) algorithm [49, 50], which implements a global minimisation approach: it can fit $\simeq 10^6$ parameters and millions of tracks simultaneously, making it optimal for complex

tracking detectors. The alignment is performed by minimising the following equation:

$$\chi^2(\boldsymbol{p}, \boldsymbol{q}) = \sum_i \sum_j \boldsymbol{r}_{ij}^T(\boldsymbol{p}, \boldsymbol{q}_j) V_{ij}^{-1} \boldsymbol{r}_{ij}(\boldsymbol{p}, \boldsymbol{q}_j) \tag{4.5}$$

where $\boldsymbol{p}$ are the alignment parameters, $\boldsymbol{q}$ the track parameters, and $\boldsymbol{r}_{ij}$ is the track residual for a given hit i on track j. Finally, V_{ij} is the covariance matrix of each measurement. The inversion of this matrix is resource intensive and memory consuming. MP-II uses the MINRES algorithm [51], which iteratively inverts the matrix, an only requires non-zero matrix elements as input: this permits to speed up the process.
A common problem for alignment algorithms are the weak modes, that is parameters to which the alignment is not sensitive. If, for example, a tracker made by parallel pixel detectors is traversed by an orthogonal beam, the alignment is not sensitive to translations along the beam axis, as the residuals would be unchanged. Additional constraints are therefore necessary (such as requiring the mean shift of the three planes to be zero).

4.2 DESY Test Beam Facility

Test beams refer to experiments made on particle beams (usually protons or electrons), which are essential to characterize pixel detectors (or other particle physics detectors) and complement the studies made in laboratory. In test beams, pixel detectors are operated in conditions as close as possible to the final experiment. With test beam experiments, new pixel designs and front-end electronics can be studied by performing various measurements (for instance, hit detection efficiency and resolution).
In order to perform such studies, unbiased reference measurements are necessary. The so-called beam telescope detectors are used to provide the reference tracks: they consist of planes of well-known and characterized silicon pixel or strip detectors with good spatial resolution. They are mounted around the Device Under Test (DUT), that is the pixel detector to be studied. Usually, a telescope is made by two parts: the upstream arm, placed before the DUT with respect to the beam line), and the downstream arm, placed after the DUT. In this way it is possible to compare the telescope reconstructed tracks with the DUT measurements.
The work presented in this Thesis is based on test beam experiments at the DESY (Deutsches Elektronen-Synchrotron) laboratories in Hamburg, in particular at the test beam area 21 [52]. Therefore, an overview of this test beam facility is presented on the following.

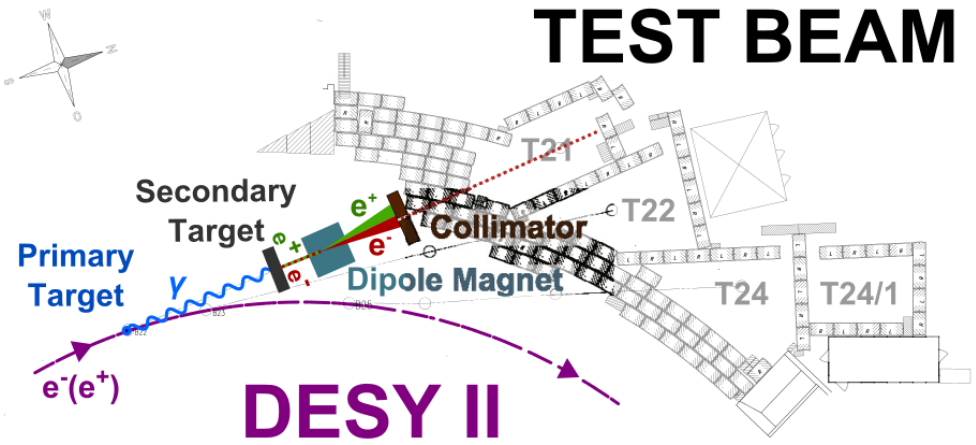


Figure 4.1: Schematic view of the test beam generation at the DESY test beam facility, for the test beam area 21.

4.2.1 Beam Generation

The DESY test beam facility provides electrons or positrons beams up to an energy of 6 GeV. The beam is provided by the DESY-II synchrotron: with a circumference of about 293 m, it is mainly used as the injector of the PETRA-III storage ring. The DESY-II dipole magnets operate in a sinusoidal ramping mode with a frequency of 12.5 Hz. The length of one acceleration is therefore about 80 ms, while the bunch length is about 30 ps. The revolution frequency is 1 MHz, and each bunch contains about 10^{10} electrons or positrons.

The primary beam is not directly used in the test beam facility: a twofold conversion in used to reduce the particle rate. Carbon fibres are positioned in the beam, so that bremsstrahlung photons are created, escaping the beam line tangentially. The photons are then converted to electron and positron pairs on a secondary metal target. The energy distribution of the electrons/positrons reaches 6 GeV. A spectrometer dipole magnet is then used to select electrons or positrons, as well as their momentum. The resulting beam is then collimated and sent to the test beam facility. The particles rate is about 10 kHz to 100 kHz. The beam divergence is about 0.5 mrad, and the energy spread approximately 5%.

The particles rate has a maximum around an energy of 3 GeV. However, at this energy multiple scattering effects significantly deteriorate test beam measurements. Therefore, the spectrometer is usually operated to produce a beam with an energy between 5 GeV and 5.6 GeV. This is a good compromise between rate and resolution of the measurements. A total of three indepen-

dent test beam areas are available for users. Figure 4.1 shows a schematic view of the beam generation.

4.2.2 The DATURA Telescope

The beam telescope used in test beam area in 21 is called DATURA [53], and is part of a few identical telescopes built within the EUDET and AIDA programs. DATURA is made of six planes of monolithic pixel detectors, featuring the MIMOSA-26 (M-26) sensors, manufactured with the 350 nm CMOS technology. The M-26 sensors consist of 1152 columns and 576 rows of pixels with 18.4×18.4 µm^2 pitch, covering an active area of 21.2×10.6 mm^2. The sensor thickness is about 50 µm. The M-26 sensors are readout with a rolling-shutter method: the columns are readout in parallel while each row is read in 16 cycles of a 80 MHz clock. At this clock frequency, the M-26 integration time equals 115.2 µs: about 8680 frames to be readout per second. This time is necessary to perform a Correlated Double Sampling (CDS) to remove charge offsets, and to perform zero suppression via the discriminators in the digital circuitry placed outside the active pixel array. The discriminator thresholds can be configured as multiples of the noise Root Mean Square (RMS). The threshold is usually kept at level six (that is six times the noise RMS): in this configuration the noise occupancy per pixel is about 6×10^{-5}. Due to the limited buffer size, the maximum rate of particles that can be detected per unit area is about 1 MHz/cm^2.
Each pixel sensor is mounted in an aluminium jig, and the jigs are placed on rails. Lightproof Kapton foils of 25 µm thickness protect the sensors on each side. The three upstream jigs (the upstream arm) and the three downstream jigs (the downstream arm) can be moved independently along the beam direction, in order to facilitate the installation of the DUT in the middle of the telescope. The jigs composing an arm are usually kept at a 150 mm distance from each other. Moreover, the jigs are cooled through the use of chiller with water as coolant, so that the sensors are kept at a stable temperature of 18°C.
The beam traverses only 300 µm of silicon and 300 µm of Kapton foils. The material budget of a telescope plane is defined by: $\epsilon = \Sigma_i x_i / X_{0,i}$ where x_i is the physical material thickness of the scatterer and $X_{0,i}$ the corresponding radiation length. By also accounting for air, the total telescope material budget is 4.8×10^{-3}.

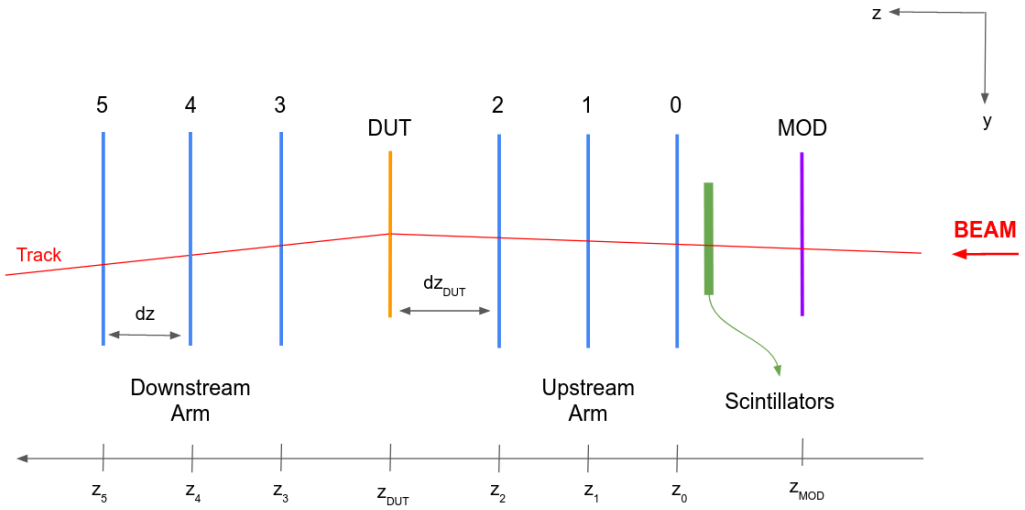


Figure 4.2: Sketch of the DESY test beam experimental setup in the $y-z$ plane. The telescope planes are represented in blue. See text for details.

4.2.3 The Experimental Setup

A sketch of the experimental setup is shown in Figure 4.2. In the chosen right-handed reference frame, the y direction points vertically down and the z direction along the beam direction. The z positions of the telescope planes are denoted to as z_0 to z_5, while the distance between two nearby planes as dz (usually dz=150 mm). The origin is placed on $z_0 = 0$.
The DUT[2] is installed between the upstream and the downstream arms, and the z position is denoted to as z_{DUT}. To limit the material budget, only one DUT at a time is tested. The DUT is placed on a mechanical support that can be controlled and moved with micrometer precision along the x and y axes. This is used to optimise the orientation with respect to the beam. Moreover, the DUT is mounted on a hinge that can be rotated: this is used to test the DUT at different rotation angles around the y axis. The DUT[3] is connected to a KC705 board and readout with the BDAQ53 software presented in Section 3.1.3. Figure 4.3 shows a photo of the test beam setup with a mounted DUT.
If the DUT is irradiated[4], an aluminium cooling box (covered by armaflex for

[2]Both Type-A and Type-B SCCs, presented in Section 3.1.2 can be tested on beam.
[3]All the the pixel sensors tested for this Thesis were bump bonded to the RD53A ROC, and only the Linear AFE was tested on beam.
[4]As explained in Chapter 5, some pixel detectors were tested after being irradiated to fluences similar to those expected in the future CMS Inner Tracker.

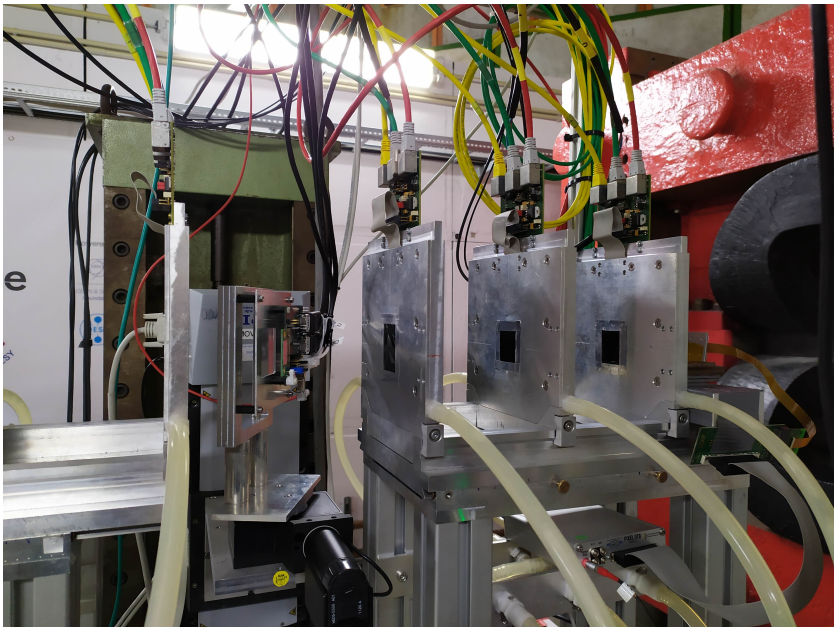


Figure 4.3: Test beam setup with a non-irradiated DUT. The upstream arm and the first jig of the downstream arm are visible. The DUT is rotated around the y axis. The beam is coming from the right.

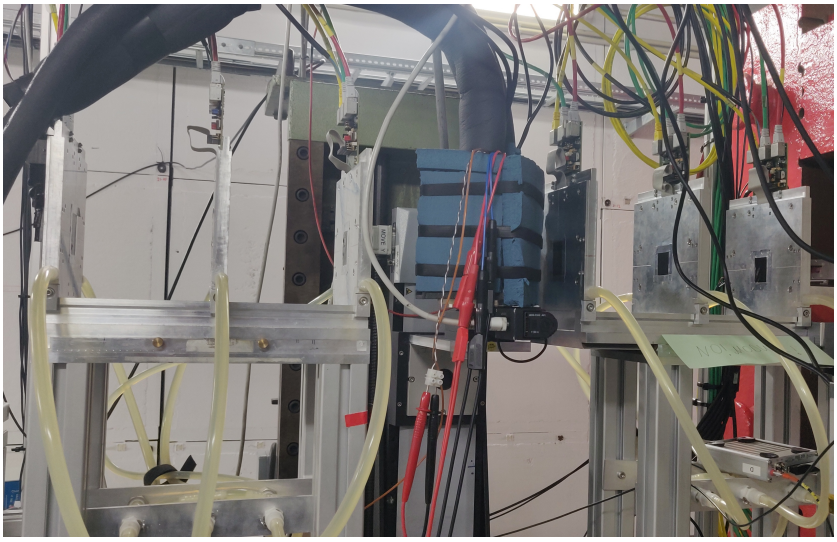


Figure 4.4: Test beam setup with an irradiated DUT. All the telescope planes are visible, as well as the cooling box with the DUT inside it. The beam is coming from the right.

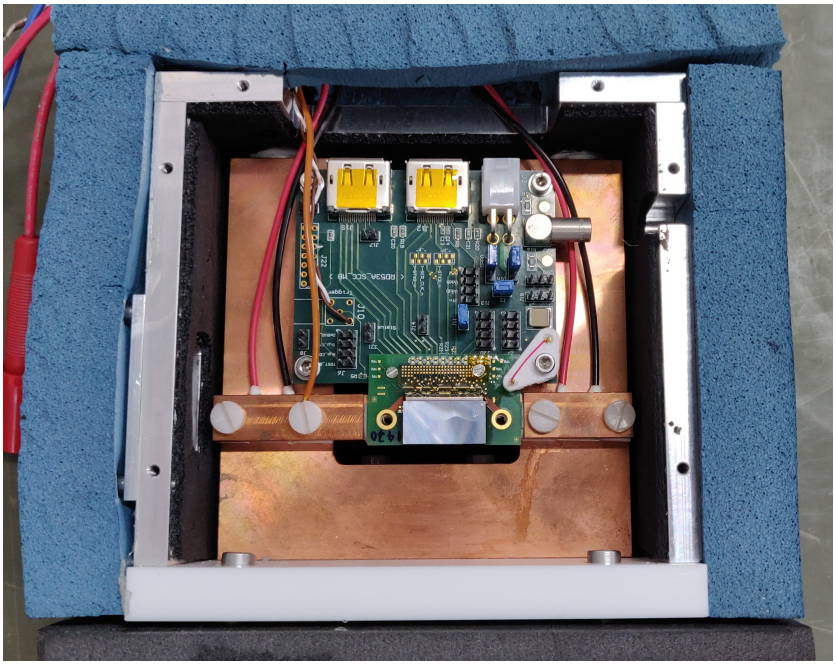


Figure 4.5: The DUT (hosted on the Type-B SCC) mounted inside the cooling box. The copper bar used for cooling the DUT is visible.

thermal insulation), is placed between the telescope arms. The DUT is kept inside, in order to keep it at a low and constant temperature (hence reducing the radiation damage effects). Inside the box, the downstream face of the DUT is in contact, through a thermally conductive paste, to a thin copper bar, which is kept cold (at about -27 °C) by two Peltier cells, cooled by an external ethanol-based chiller. Moreover, dry air is flushed inside the box to avoid condensation. The box features two cut-outs at the beam impact points, that are covered by thin Kapton foils. Figure 4.4 shows a photo of the the test beam setup with the cooling box, while Figure 4.5 shows a photo of the DUT inside the cooling box.
Due to the presence of a copper bar near the DUT, in this configuration the downstream arm can not be used due to severe beam degradation. Therefore, only the upstream arm of the telescope is used in track reconstruction in the case of an irradiated DUT.
With a beam energy of about 5 GeV, the track multiplicity per event in the DATURA telescope is about 2.5. Such high multiplicity is due to the very long time integration of the M-26 sensors: further particles are detected after the trigger has been sent. Therefore, an external timing reference is necessary, since the M-26 sensors are not capable of giving time informations of

individual pixels. A CMS Phase-1 module is placed upstream the telescope (at a position z_{MOD}) to function as a reference module (MOD in the following). The telescope tracks are first matched with the MOD, in order to select particles with the correct time stamp. The selected tracks can be compared with the hits on the DUT so that, for instance, a reliable hit detection efficiency measurement can be performed. In order to maximize charge sharing, and therefore the spatial resolution, the MOD is mounted rotated around the x and y axes.
Two plastic scintillators, connected to PMTs through light guides, are used to generate trigger signals for particles traversing the setup. The scintillators are placed upstream the telescope, but downstream the MOD. They are crossed with respect to each other, defining a rectangular acceptance window of 20 mm $\times$ 10 mm (hence matching the M-26 active area). The Trigger Logic Unit (TLU) is based on a commercial Spartan 3 board and features a coincidence unit with discriminator boards accepting the PMT input signals. Additional custom made PCBs allow for an easy integration of the DAQ systems. The TLU takes a trigger decision based on its input channels, with a programmable logic. Moreover, the TLU implements a handshake to handle busy signals. For instance, during the slow M-26 integration time, the DATURA telescope vetoes new triggers.
The 40 MHz clock (necessary for the RD53A ROC) is not synchronous with the beam (which will be the final deployment situation at the LHC). This is due to the fact that the 1 MHz beam frequency is constantly re-synchronised to the 50 Hz power grid before injecting a new bunch.
Due to the slow DATURA integration time, the DUTs can not be tested at high trigger rates at the DESY test beam facility. The CMS collaboration has been developing a dedicated telescope, CHROMIE (CMS High Rate telescOpe MachInE) [54], to test the new pixel detectors with the 750 kHz trigger rate of the Phase-2 CMS experiment. The first tests with this telescope will be made at CERN after the Long Shutdown 2. Since the spatial resolution of CHROMIE is worse than that of EUDET, the measurements made with the two telescopes will be complementary.

4.2.4 Data Acquisition

The DAQ of the test beam has to record data from all the different detectors, merge them and store them on disk for later, offline analysis. The DAQ provides the interface between the detectors readout electronics and the PCs, as well as central components for controlling all the subsystems.
The EUDAQ framework is the data taking software used for the EUDET type beam telescopes. It consists of independent modules executed on different

machines and communicating over Ethernet network via TCP/IP. EUDAQ requires one event per trigger per every connected detector: the trigger rate is hence limited by the slowest device. The Run Control is the central operational point for users, features a Graphical User Interface (GUI), and provides all the necessary controls.
Data from the detectors are delivered to EUDAQ via Producers, which function as links between the framework and the connected systems (the telescope, the DUT, the TLU, etc.). Producers provide a set of commands to be called by the Run Control. Data readout from each subsystem by the respective Producers are sent to the Data Collector, which performs the event building, that is the correlation of events from all subsystems in single, global events with all data belonging to one trigger.
Aside basic sanity checks, the data quality is verified via the Online Monitor, that connects to the Data Collector to request a fraction of the recorded events, to decode them and to produce basic plots such as hit maps. Data from each subsystem is decoded using Data Converter plug-ins.
The EUDAQ version used at DESY has been patched to work with the BDAQ53 software.

4.2.5 Offline Analysis

During DESY test beam experiments, data is usually acquired in Runs of about 500000 events. After each run a raw file is produced, to be analysed offline. A custom software[5], based on EUDAQ, has been developed by DESY specifically for the analysis of the RD53A chip in EUDET type telescopes. The software is made of two parts: *tele*, for track reconstruction and telescope alignment, and *scope*53, for the DUT and MOD alignments.
The *tele* software takes the raw file as inputs, as well a geometry file with rough estimates of the telescope planes positions. The first step in the track reconstruction performed by *tele* is the masking of the telescope noisy pixels. Pixels exceeding a frequency threshold of 1% are marked as noisy and are not used in the track reconstruction. Usually, noisy pixels are between 50 and 100 per M-26 sensor (less than 0.2 ‰). After this step, clusters are reconstructed by connecting adjoining pixels. Since the M-26 sensors do not give charge information (i.e. the readout is binary), a simple geometrical interpolation of the cluster center is performed.
After the cluster reconstruction, tracks are searched: a track candidate is required to have one hit in each of the six planes. The initial track candidate is built from triplets, built independently is the upstream and downstream

[5] `https://stash.desy.de/projects/RDA/repos/tele-scope/browse`

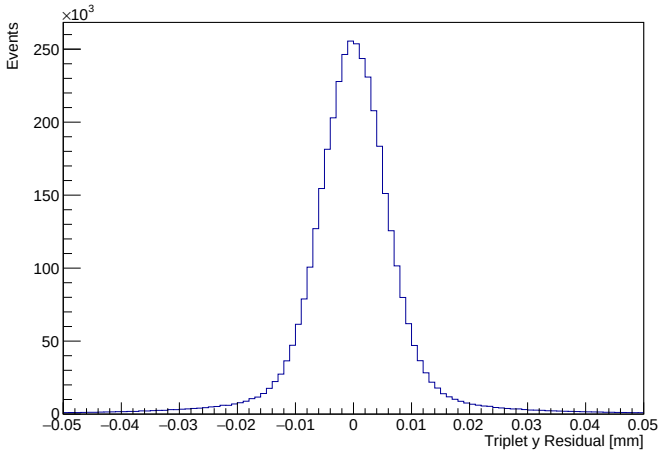


Figure 4.6: Upstream triplet y residual evaluated on telescope plane 1 after alignment.

arms. A triplet is found by searching for hits in planes 0 and 2 (3 and 5), and by looking for a corresponding hit on plane 1 (4). The two triplets are then extrapolated to the DUT position in the center of the telescope: if an intersection point is found within 100 µm, the track candidate is accepted. It should be noted that, due to multiple scattering effect on the DUT, the triplet slopes can be different.
The telescope alignment consists of an iterative procedure. Usually, only the first 100000 events of a Run are used for telescope alignment. In the first iteration over selected events, for each telescope plane the x and y residual distributions are fitted to gaussian distributions: the respective mean values are subtracted to the plane positions (so that the residual mean values are zero). The residual is simply defined as the difference between the reconstructed track and the hit on the telescope plane to be aligned. Subsequent iterations progressively correct others degrees of freedom as well: the z position and rotations around the z axis. Other rotations are kept fixed, considering the tight constraints imposed by the telescope mechanical structure. In order to constrain the weak modes, the second telescope plane is fixed at its initial position. About ten iterations are necessary to complete the alignment with sufficient precision. The upstream triplet y residual evaluated on telescope plane 1 after alignment is reported in Figure 4.6: the distribution is narrow and centred at zero. Telescope alignment is performed every time the mechanical structure is altered.
After the telescope alignment procedure has been completed, the DUT and MOD alignments are performed using the *scope*53 software. Rough estimates

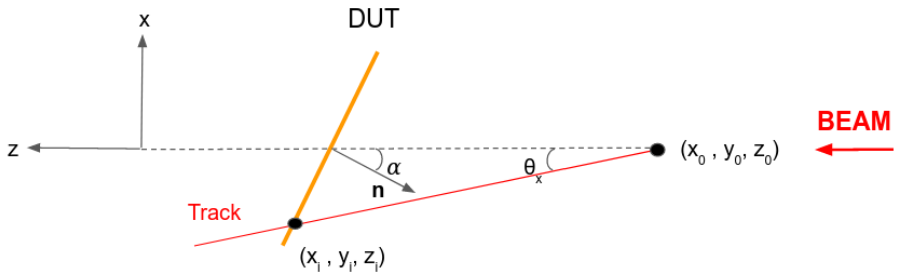


Figure 4.7: Sketch of the DUT and a reconstructed track in the $x-z$ plane.

of the positions and rotation angles around the x, y and z axes are provided by the user. The first step of the alignment procedure is the cluster reconstruction. Starting from a seed pixel, adjacent pixels are added to the cluster and removed from the list of un-clustered pixels. When no more pixels can be found, the cluster is marked as complete, and the procedure starts again with a new seed. The position of the cluster center if found using the CoG algorithm, described in Section 4.1.1.
The reconstructed telescope tracks have to be transformed in the DUT (or MOD) reference frame, in order to facilitate the comparison of the sensor features. By defining (x_0, y_0, z_0) the origin of the reference system, and θ_x and θ_y the track slopes, the inclined track can be expressed by:

$$\begin{aligned} x &= x_0 + (z - z_0)\tan(\theta_x) \\ y &= y_0 + (z - z_0)\tan(\theta_y) \end{aligned} \tag{4.6}$$

If the DUT is placed at $\boldsymbol{r_p} = (0, 0, z_p)$ and its normal vector is $\boldsymbol{n} = (n_x, n_y, n_z)$ the DUT plane can be parametrised as:

$$\boldsymbol{n} \cdot (\boldsymbol{r} - \boldsymbol{r_p}) = 0 \tag{4.7}$$

The intersection point between the track and the DUT, $\boldsymbol{r_i} = (x_i, y_i, z_i)$, can then be calculated by inserting Equations 4.6 in Equation 4.7:

$$z_i - z_0 = \frac{n_z\,(z_p - z_0) - n_y\,y_0 - n_x\,x_0}{n_x\tan(\theta_x) + n_y\tan(\theta_y) + n_z} \tag{4.8}$$

By inserting z_i in Equations 4.6, x_i and y_i can be calculated as well. The DUT plane orientation can be parametrised using $\boldsymbol{n}$ and the Euler angles (ω, α, ϕ). If $\mathcal{R}_x(\omega)$, $\mathcal{R}_y(\alpha)$, $\mathcal{R}_z(\phi)$ are the corresponding rotation matrices,

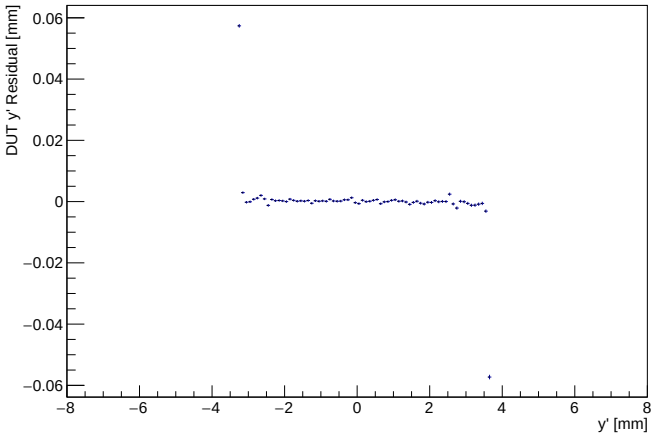


Figure 4.8: DUT (planar 25×100 µm^2 planar pixel detector) y' residual versus y' after alignment.

the normal vector can be expressed as:

$$\begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} = \mathcal{R}_x(\omega)\mathcal{R}_y(\alpha)\mathcal{R}_z(\phi) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{4.9}$$

It is important to note that rotations do not commute. Finally, the intersection point $\boldsymbol{r_i}$ expressed in the DUT frame of reference is given by:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathcal{R}_z(-\phi)\mathcal{R}_y(-\alpha)\mathcal{R}_x(-\omega) \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} + \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} \tag{4.10}$$

where $z' \equiv 0$ and a_x, a_y, a_z are the alignment shifts. A sketch of the DUT (rotated by an angle α around the y axis) and the reconstructed track in the $x - z$ plane is reported in Figure 4.7. The procedure is analogous for the MOD.

The alignment procedure is similar to the one described for the telescope. Also in this case, the alignment is performed by selecting the first 100000 events of the Run. The triplets for the upstream and downstream arms are extrapolated to the DUT position and averaged. The intersection point is obtained using Equation 4.6 and converted to the DUT reference frame: residuals between the track impact and the DUT hit are thus calculated. In the first alignment iteration, the residual distributions for x' and y' are fitted to gaussian distributions: the mean values correspond to the shifts a_x and a_y,

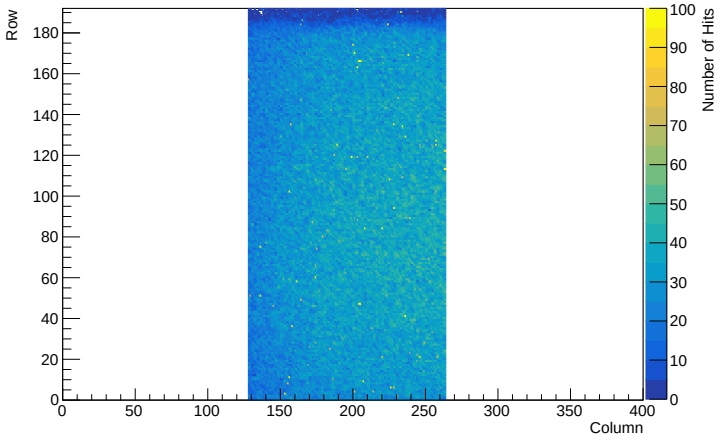


Figure 4.9: Hits on the DUT during a Run. Only the Linear AFE of the RD53A ROC has been tested on beam.

to be added or subtracted to the DUT position (so that the residual mean values are zero). Subsequent iterations also take the three Euler angles and the z coordinate into account. For instance, rotations around the x axis (that is a non-zero ω angle) are corrected by considering the y' residual versus the DUT pixel position along y'. In case the slope of this distribution appears to be different from zero, the rotation angle is modified. Figure 4.8 shows the DUT y' residual distribution versus y' after alignment: the distribution does not show a dependence that would indicate a remaining misalignment. In order to maximise the alignment precision, the DUT is re-aligned at every Run. The procedure is analogous for the MOD.
DUT analysis is performed inside *scope*53, so that users can define histograms that are produced in output ROOT [42] files. One important DUT parameter is the hit detection efficiency, defined as the number of reconstructed telescope track with an associated hit in both DUT and MOD, divided by the number of reconstructed telescope tracks with an associated hit in the MOD:

$$\epsilon = \frac{\text{Tracks linked to MOD and DUT}}{\text{Tracks linked to MOD}} \tag{4.11}$$

Moreover, a fiducial cut is made to select only tracks intersecting the DUT. As already mentioned, the MOD link is necessary for selecting only in-time tracks. Since the MOD can only reduce the number of reconstructed tracks, no bias in efficiency measurement is induced.
The binomial error on the efficiency is always $< 1\%$ in a single Run, due to the high number of reconstructed tracks. A conservative error of 1% has however

been considered in this Thesis, to account for eventual misalignments or to the presence of undetected noisy pixels[6] (which would artificially increase the efficiency).
Finally, Figure 4.9 shows the number of hits on the DUT during a Run. The coloured scale indicates the number of registered hits in every pixel: the beam profile can be seen. The two white bands at the borders correspond to the Synchronous and Differential front-ends, which were not tested on Beam.

4.3 Resolution Estimation

The segmentation of the pixels determines the attainable precision of the measurement of the particle position. Due to charge sharing effects between adjacent pixels, the measured spatial resolution is almost always better than the digital resolution:

$$\sigma_{dig} = \frac{\text{pixel pitch}}{\sqrt{12}} \tag{4.12}$$

It should be noted that the limitation of hybrid pixel detectors is the achievable density of the electronic needed to amplify, discriminate, and process the hit information in the ROC. The 65 nm technology used for RD53A (and the future CROC) allows for a pixel cell area of 50×50 µm^2, which is already six time smaller than the ROC used in the present CMS tracker.
The spatial resolution of tracking detectors is defined as the width of the residual distributions. In the test beam, the measured DUT residual width σ_{meas} is influenced by the telescope resolution σ_{tel} as well, so that:

$$\sigma^2_{meas} = \sigma^2_{int} + \sigma^2_{tel} \tag{4.13}$$

where σ_{int} is the intrinsic DUT resolution to be measured. Therefore, the telescope resolution needs to be precisely known in order to perform resolution measurements.
One major problem in determining the width of the residual distributions is that they do not follow the shape of a gaussian distribution. Indeed, delta rays add non-gaussian tails and clustering algorithms further distort the shape significantly. One approach is to calculate the RMS of the distribution, preferably excluding large residuals values. However, this method usually leads to overestimated width values.
In certain cases, more suitable fit functions can be used. The Student's

[6]Pixels exceeding a frequency threshold of 1% are masked offline, during data analysis.

t-distribution is typically used at the DESY test beam. It is defined as:

$$f(x) = a + \frac{b}{\sigma\sqrt{\pi\nu}} \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)} \left[1 + \frac{1}{\nu}\left(\frac{x-\mu}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}} \tag{4.14}$$

where $\Gamma(x)$ is the gamma function and a, b, μ, σ, ν are the free fit parameters. Being a convolution between a gaussian and a Breit-Wigner distribution, it accounts for non-gaussian tails. The σ parameter describes the width of the distribution, and is used to estimate the spatial resolution in this Thesis.

4.3.1 Telescope Resolution

To evaluate the telescope resolution when using all the six planes (e.g. when the DUT is not irradiated, and the cooling box is not on the beam line), the following strategy was adopted. The sixtuplet x residual is defined as the difference between the upstream and downstream triplets extrapolated on the DUT:

$$\Delta x_{six} = x_{up} - x_{down} \tag{4.15}$$

The width of this residual distribution can be written as:

$$\sigma^2_{six} = \sigma^2_{up} + \sigma^2_{down} = 2\sigma^2_{tri} \tag{4.16}$$

where $\sigma_{tri} = \sigma_{up} = \sigma_{down}$ is the triplet resolution on the DUT plane (it is assumed that the resolution is the same for the upstream and the downstream arms). The x' DUT residual ($\Delta x'_{DUT}$) can be written as (the procedure is the same for $\Delta y'_{DUT}$):

$$\Delta x'_{DUT} = x'_{DUT} - \frac{x'_{up} + x'_{down}}{2} \tag{4.17}$$

The width of the DUT residual distribution σ_{meas} is therefore:

$$\sigma^2_{meas} = \sigma^2_{int} + \frac{\sigma^2_{tri}}{2} = \sigma^2_{int} + \frac{\sigma^2_{six}}{4} \tag{4.18}$$

By inverting this equation, the intrinsic DUT resolution σ_{int} is given by:

$$\sigma_{int} = \sqrt{\sigma^2_{meas} - \left(\frac{\sigma_{six}}{2}\right)^2} \tag{4.19}$$

The telescope resolution is therefore estimated with $\sigma_{tele} = \sigma_{six}/2$. This is correct if the DUT is parallel to the telescope planes (e.g. the Euler angles

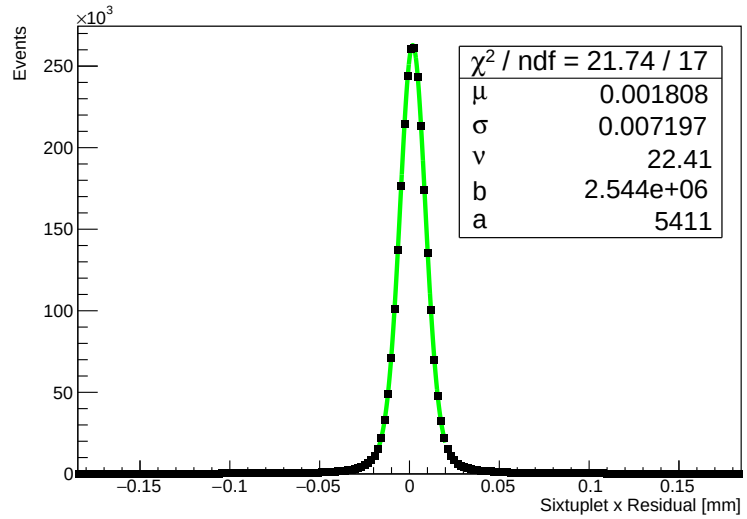


Figure 4.10: Sixtuplet x residual distribution, fitted with Equation 4.14.

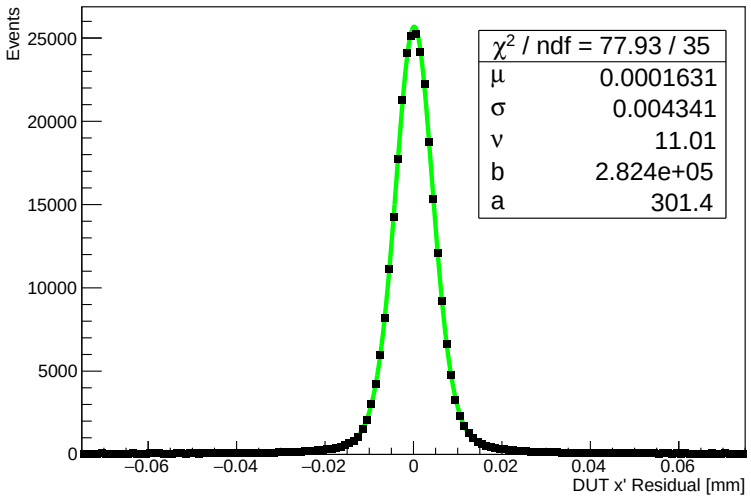


Figure 4.11: DUT (planar 25×100 µm^2 planar pixel detector) x' residual distribution, fitted with Equation 4.14. The DUT was rotated by an angle $\alpha = 8°$ around the y axis.

are zero). If the DUT is rotated by an angle α around the y axis[7], as shown in Figure 4.7, a projection in the DUT plane is necessary:

$$\sigma_{tele} = \frac{\sigma_{six}}{2\cos(\alpha)} \tag{4.20}$$

In Figure 4.10 the sixtuplet residual distribution fitted with Equation 4.14 is reported. The σ of the fit is 7.2 µm: the telescope resolution (with $\alpha = 0$) is therefore $\sigma_{tele} = 3.6$ µm. This number depends on the distance between the upstream and downstream arms. For this reason, the telescope resolution is evaluated every time the telescope mechanics are moved at different distances/positions.
To rotate the DUT around the y axis, the telescope arms need be moved away for mechanical constraints (Figure 4.3 shows the typical setup), resulting in a worse telescope resolution. In any case, to optimise resolution measurements, the telescope arms need to be placed as close as possible with respect to each others and to the DUT (i.e. the distance $\mathrm{d}z_{DUT}$ from Figure 4.2 needs to be minimised). This is particularly important when the DUT resolution is close to the telescope resolution. Figure 4.11 shows the DUT x' residual distribution, fitted with Equation 4.14. In this case $\sigma_{meas} = 4.3$ µm (the pixel pitch is 25 µm and the DUT is rotated by $\alpha = 8°$.) With Equation 4.19, the DUT intrinsic resolution can be evaluated.
In order to estimate errors on σ_{meas} and σ_{six}, the alignment procedure was repeated from scratch for different Runs with the same telescope and DUT configurations. The fits of the residual distributions were compared: the maximum variation of the fitted σ was found to be about ± 0.1 µm for both σ_{meas} and σ_{six}, which was taken as the measurement error for these quantities. The error on σ_{int} ($\Delta\sigma_{int}$) is found by propagating Equation 4.19: the typical case is $\Delta\sigma_{int} = 0.2$ µm. This result was validated by comparing σ_{int} estimated from Runs with the same DUT configuration, but different telescope configurations: the measured resolutions were always found to be within $\Delta\sigma_{int}$.
In the case of an irradiated DUT, the cooling box placed in the middle of the telescope prevents the use of the downstream arm. Therefore, in this case, the resolution is evaluated using the triplets from the upstream arm only. Evaluating the telescope resolution in this case is non-trivial. Therefore, a simulation[8] has been performed at DESY, based on the GBL algorithm. The simulation takes into account the material budget of the telescope and the beam energy (usually 5.2 GeV or 5.6 GeV). Figure 4.12 shows the estimated

[7] This is the typical case at DESY, since the DUT can be rotated around the y axis to study efficiency and resolution for inclined tracks.

[8] `https://github.com/simonspa/resolution-simulator`

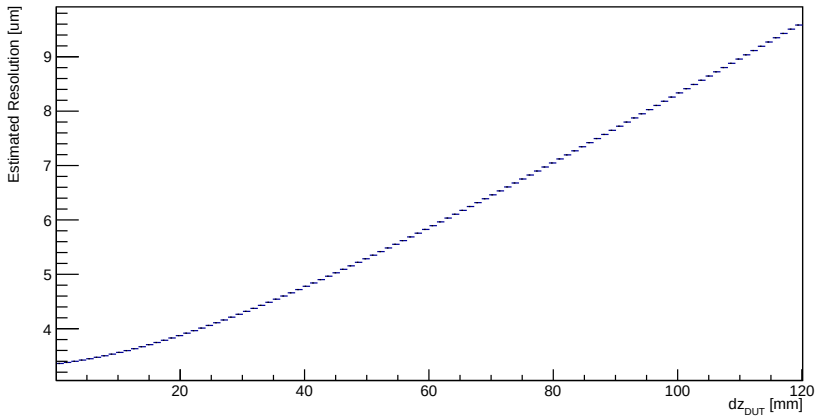


Figure 4.12: Estimated telescope resolution as a function of $\mathrm{d}z_{DUT}$ for a 5.2 GeV electrons beam.

resolution versus the distance $\mathrm{d}z_{DUT}$ between the DUT and the last upstream telescope plane for a 5.2 GeV electrons beam. The intrinsic DUT resolution in this case is given by:

$$\sigma_{int} = \sqrt{\sigma^2_{meas} - \sigma^2_{sim}} \tag{4.21}$$

where σ_{sim} is the telescope resolution evaluated from the simulation and σ_{meas} the width of the DUT residual distribution.
Since this resolution strongly depends on $\mathrm{d}z_{DUT}$, small misalignments in this direction can cause overestimations or underestimations on the telescope resolution. Moreover, the alignment procedure is not very sensitive to z, and misalignments of about 1-2 µm are common. The error on the simulated telescope resolution was estimated to be ± 0.2 µm. By propagating Equation 4.21, $\Delta\sigma_{int}$ is calculated. Also in this case, the result was validated by comparing σ_{int} estimated from Runs with the same DUT configuration, but different telescope configurations: the measured resolutions were always found to be within $\Delta\sigma_{int}$.
All the resolution studies reported in this Thesis are performed without making any cut on the collected charge. In principle, a selection on the collected charge could be used to reject contributions from delta rays, hence enhancing the measured resolution.

Chapter 5

Analysis of Pixel Detectors

5.1 Overview

Over the three years of my PhD I participated to several test beam experiments at DESY[1], in which different types of 3D and planar pixel sensors, bump bonded to the RD53A ROC, were thoroughly analysed. Many different pixel designs were studied over the course of the years. The aim of this Chapter is to give an overview of the results, in sight of the pixel design choice that the CMS collaboration will make for the Phase-2 Inner Tracker. The results I obtained are based on FBK sensors.
It is important to test the pixel detectors after receiving irradiation fluences similar to those expected in the future CMS Inner Tracker. Some of the pixel detectors presented in this Chapter were characterised after receiving an irradiation[2] to fluences from $5 \times 10^{15}\,\mathrm{n_{eq}/cm^2}$ to $24 \times 10^{15}\,\mathrm{n_{eq}/cm^2}$.
It should be noted that the corresponding received dose varies dramatically depending of the irradiation facility. While this is not a problem for the pixel sensor, the ROC can suffer significant damage, since electronics is sensitive to the received dose. For instance, the majority of the pixel detectors presented in this Chapter were irradiated at the Karlsruher Institut für Technologie (KIT) irradiation facility [55], using a 23 MeV proton beam. In this facility, a fluence of $1 \times 10^{16}\,\mathrm{n_{eq}/cm^2}$ corresponds to a dose[3] of 14.4 MGy, while

[1] In presence or remotely, due to the COVID-19 pandemic.

[2] Supported by the H2020 project AIDA-2020, GA no. 654168.

[3] The k factor at KIT is $\simeq 2$, therefore a neutron equivalent fluence of $1 \times 10^{16}\,\mathrm{cm^{-2}}$ corresponds to an effective fluence $\Phi = 0.5 \times 10^{16}\,\mathrm{cm^{-2}}$. The Total Ionising Dose (TID) is therefore equal to:

$$\mathrm{TID} = \frac{\Phi(-dE/dx)A\Delta x}{\rho A \Delta x} = -\Phi\frac{-dE}{\rho dx} = \Phi S$$

the RD53A ROC is certified to resist up to 5 MGy. However, the ROC has demonstrated to resist well above the specifications, and up to $\sim 30\,\mathrm{MGy}$, as will be shown in this Chapter. It should be emphasised that these radiation levels are unprecedented in high energy physics.
One 3D detector was irradiated at the CERN PS IRRAD facility [56], using 24 GeV protons. Aside from this detector (and a few others, not reported in this Thesis), this irradiation facility could not be exploited after the end of 2019 due to the Long Shutdown 2 of the LHC.
While planar pixel detectors show hit detection efficiencies greater than 99% at low bias voltages ($\simeq$ 70 V) when they are new, far higher bias voltages (hundreds of volts) are necessary to reach similar efficiencies after being exposed to these extreme radiation fluences. Above $\simeq$ 800 V however, sparks[4] between the borders of the sensor and the ROC can potentially destroy the detector, while also being infeasible from a power supply point of view. As presented in this Chapter, 3D detectors can be operated at far lower bias voltages, about 150 V, to reach high efficiencies after being exposed to the same fluence. A careful evaluation of the hit detection efficiency as a function of the received fluence and the applied bias voltage is thus necessary for the construction of the future CMS tracker.
In this Chapter a characterisation of the resolution of the 50×50 µm^2 and 25×100 µm^2 pixel cells is presented, for both planar and 3D sensors. Only the resolution along the 25 µm pitch has been studied, being of particular interest for the CMS experiment. Indeed, in the future CMS tracker, the pixel modules will be mounted in such a way that the long 100 µm pitch will be along the z axis: the short 25 µm pitch will therefore determine the resolution of the $r\phi$ coordinate.
Radiation damage reduces the collected charge and therefore the charge sharing (since it is easier for a pixel cell to be under-threshold). Moreover, the ROC is tuned to higher thresholds in order to keep the number of noisy pixels to a reasonable level. For these reasons, resolution measurements are performed also after irradiation.

where $S = 18\ \mathrm{MeVcm^2g^{-1}}$ is the stopping power for silicon with 23 MeV protons. Therefore:

$$\mathrm{TID} = 9 \cdot 10^{22}\,\mathrm{eV/g} = 9 \cdot 10^{22} \cdot 1.6 \cdot 10^{-19} \cdot 10^{3}\,\mathrm{J/Kg} = 14.4\,\mathrm{MGy}$$

[4]Coating materials, such as parylene, have been studied by the CMS collaboration to prevent sparks.

5.2 Planar Pixel Detectors

The complete list of the tested planar detectors is reported in Table 5.1. The DUT letter (A to G) is used to label the detectors. Sensors from the P-1 batch have an active thickness of 100 µm and a total thickness of 200 µm, while sensors from the P-2 batch have an active thickness of 150 µm and a total thickness of 250 µm. While thinner sensors are less affected by radiation damage, due to the shorter drift path, they also collect less charge, requiring the ROC to be tuned at low threshold values. For this reasons, both configurations are interesting to study, especially after high fluences.
CMS recently chose 150 µm active thickness sensors as baseline, since their performance is on par with 100 µm sensors after being exposed to the same fluence, while also allowing for a higher collected charge.
Additional details on the pixel design are reported under Notes. Section 2.5 presents a description of the various structures.

DUT	**Batch**	**Pixel Pitch [µm²]**	**Active Thickness [µm]**	**Notes**	**Irradiation Fluence [n_{eq}/cm^2]**	**On Beam**
A	P-1	50×50	100	-	5×10^{15} (K)	02/19
B	P-1	50×50	100	PT	5×10^{15} (K)	02/19
C	P-1	25×100	100	-	7×10^{15} (K)	07/20
D	P-2	25×100	150	BTFP	-	06/20
E	P-2	25×100	150	BT	-	07/20
					11×10^{15} (K)	12/20
F	P-2	25×100	150	-	18×10^{15} (K)	03/21
G	P-2	25×100	150	BT	24×10^{15} (K)	03/21

Table 5.1: Summary of the planar pixel detectors which were tested at DESY Test Beam. PT stands for "Punch-Through", BT for "BiTten" and BTFP for "BiTten Field Plate". (K) indicates that the detector was irradiated at the KIT irradiation facility. The DUT letter will be used to refer to these detectors in the following.

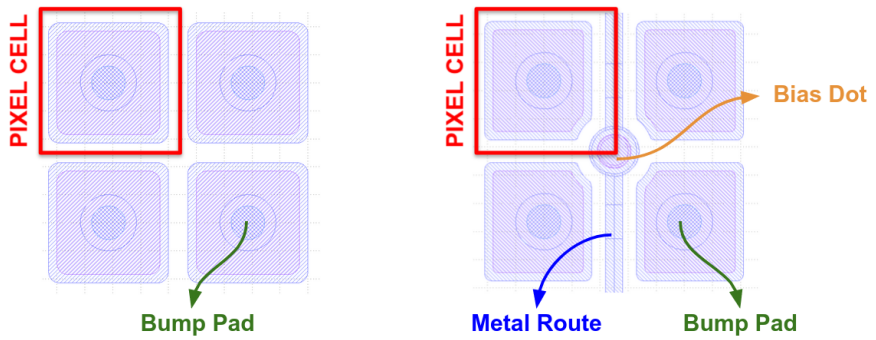


Figure 5.1: Schematic drawing of a 2×2 pixel grid of DUT A (left) and DUT B (right). The 50×50 µm^2 pixel cells are highlighted in red.

5.2.1 Irradiated 50 µm Pitch Detectors

Two irradiated planar 50×50 µm^2 pixel detectors (referred to as DUT A and DUT B) were tested on beam, both from the P-1 batch and with an active thickness of 100 µm. Both were irradiated at KIT at a fluence $\phi = 0.5 \times 10^{16}\,\mathrm{n_{eq}/cm^2}$. The DUT B features the punch-through structure, while the DUT A does not. In Figure 5.1 schematic drawings of a 2×2 pixel grid are shown for both modules.
Being irradiated, the detectors were tested inside the cooling box described in Section 4.2.3. Thresholds of both modules were tuned inside the cooling box, to an average pixel threshold of 1400 electrons for DUT A and 1200 electrons for DUT B, with a dispersion of about 50 electrons. The tuning was made targeting low thresholds and noise, having at most 1% noisy pixel channels.
Both DUTs were kept at a temperature of about -27 °C. The leakage current for DUT A was $\sim$ 540 µA at a bias voltage[5] $V_{bias} = 295$ V, which is higher than expected and probably caused by a bad cooling contact between the copper bar and the DUT. The leakage current for DUT B was $\sim$ 50 µA at $V_{bias} = 390$ V.
In Figure 5.2, the global hit detection efficiency versus the applied bias voltage is shown, for perpendicular incident tracks. DUT A reaches a hit detection efficiency greater than 99% at a bias voltage of about 210 V. The hit efficiency for the DUT B starts saturating at around 200 V, and reaches a maximum value of 95.8% at 390 V.
In Figure 5.3 the DUT A hit efficiency maps for a 2×2 pixel grid are shown,

[5]All the bias voltages shown in this Chapter are effective voltages, evaluated after having taken into account the voltage drop on the limiting resistors in series to the high voltage bias circuit.

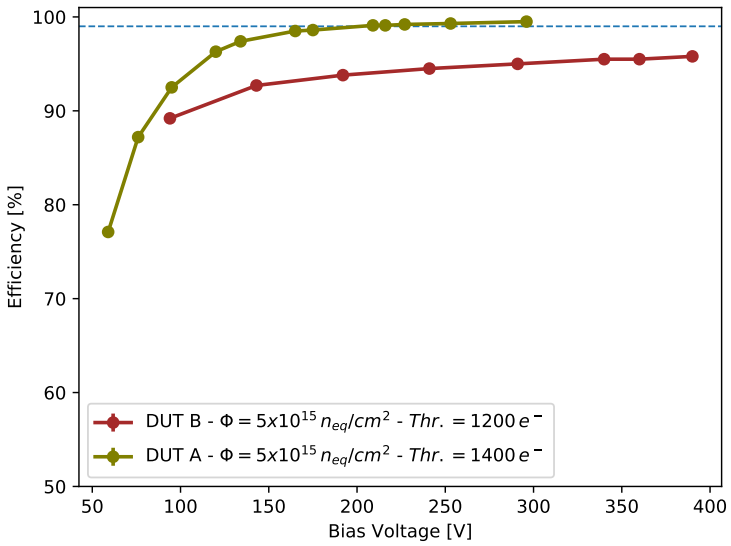
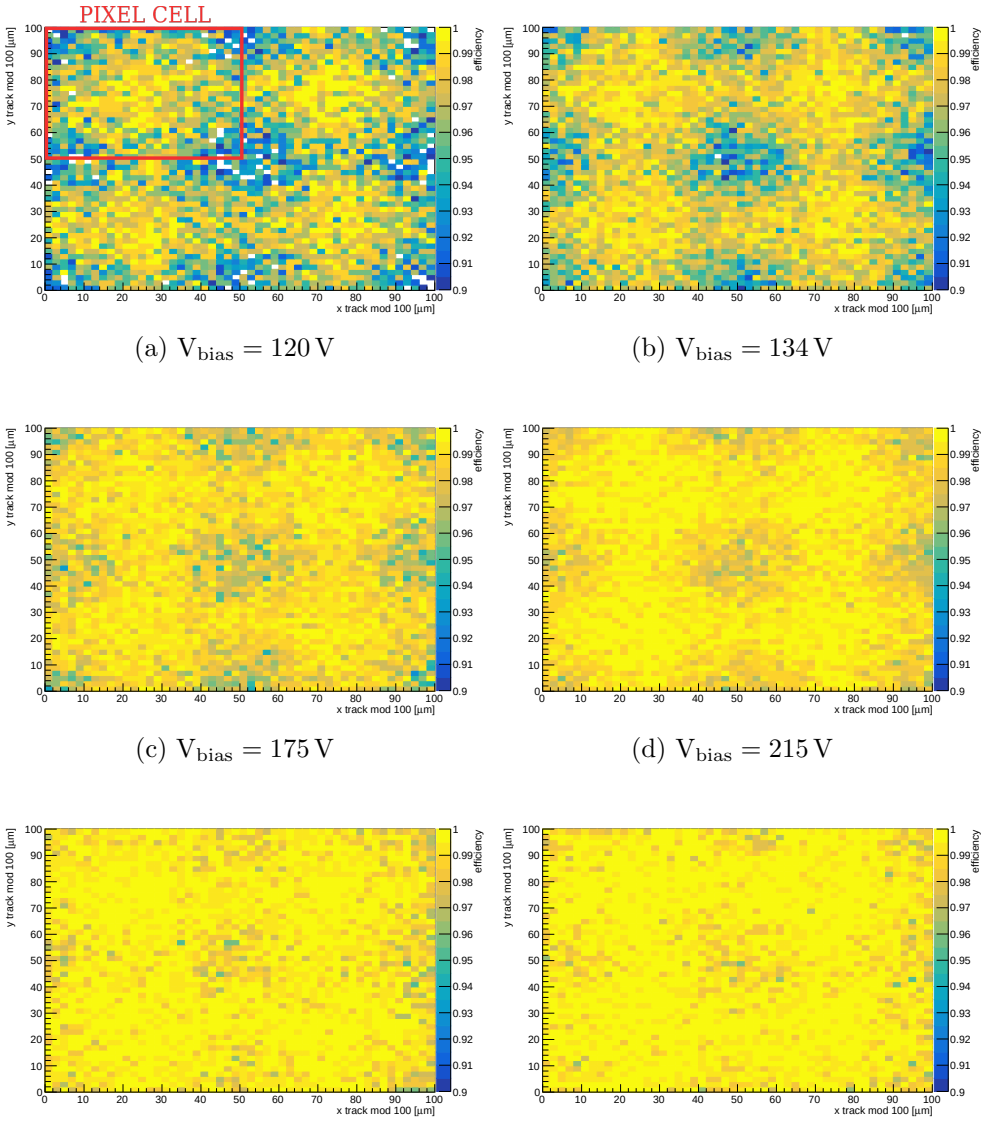


Figure 5.2: DUT A and DUT B hit detection efficiency versus the applied bias voltage with orthogonal beam incidence. The dashed line represents 99% efficiency.

for perpendicular tracks and for different bias voltages: the coloured scale indicates the hit detection efficiency. These maps are evaluated by calculating the hit efficiency inside all pixel cells, with a 2 µm binning. The result is then averaged for all 2×2 pixel grids in the sensor. The efficiency grows and gets more uniform by increasing the bias voltage. For low bias voltages, a small efficiency drop is visible at the intersection of four pixels. This effect is due to charge sharing between the four pixels and the charge trapping due to the irradiated silicon: hits in this region are more likely to be under-threshold and therefore not detected by the readout chip. By increasing the bias voltage, and hence the electric field, this effect is considerably mitigated.
In Figure 5.4 the DUT B hit efficiency maps for a 2×2 pixel grid are shown, for perpendicular tracks and for different bias voltages. In the bias dot implant region, the efficiency goes down to $\simeq 80\%$, but it recovers by increasing the bias voltage.
Rotations around the y axis (as shown in Figure 4.7) are also extremely effective, as can be seen in the hit efficiency maps in Figure 5.5. The mechanics of the DUT rotation procedure were explained in Section 4.2.3. With a rotation of only 12°, the hit detection efficiency in the punch-through implant region is greater than 85%, while at 26° the efficiency is almost uniform across the

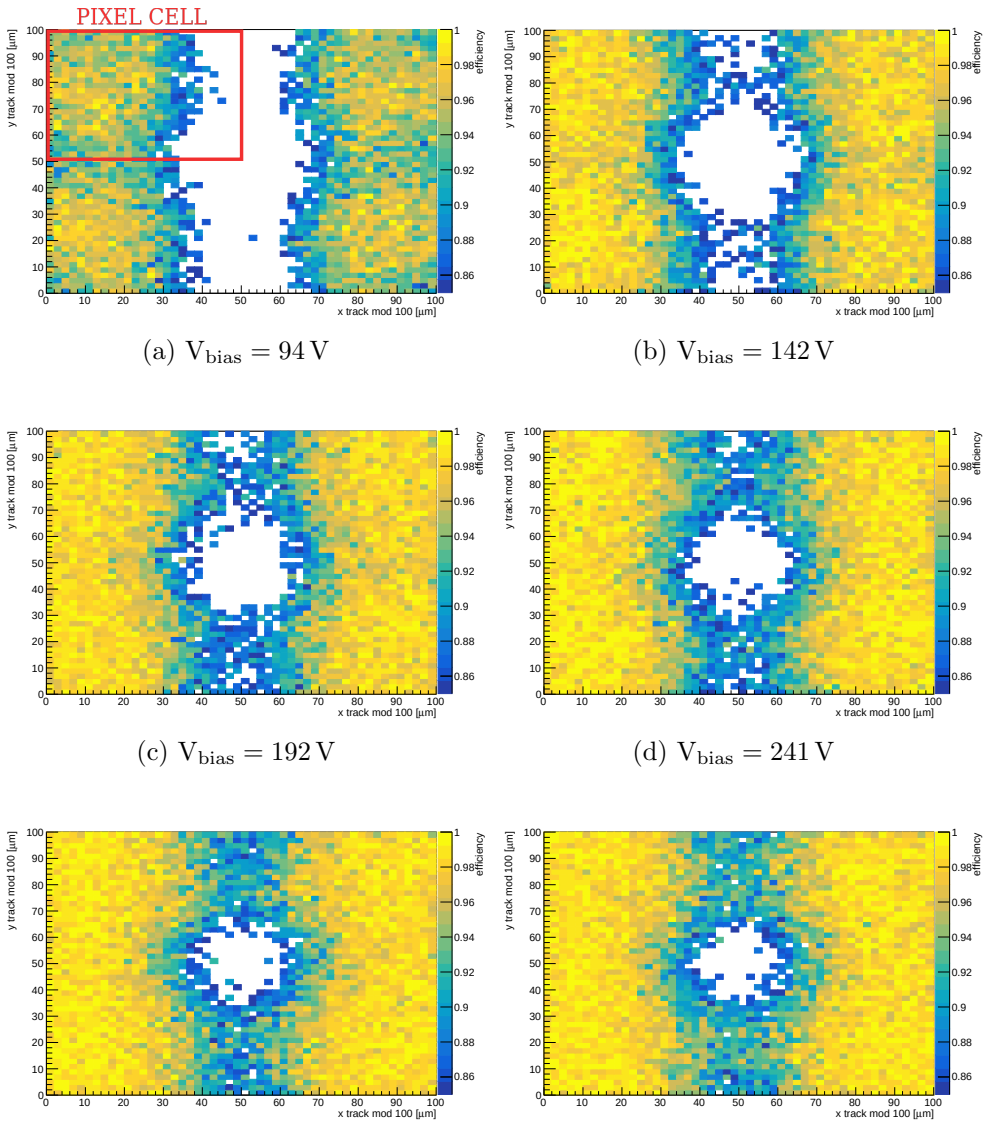


(a) $V_{bias} = 120\,V$ (b) $V_{bias} = 134\,V$

(c) $V_{bias} = 175\,V$ (d) $V_{bias} = 215\,V$

(e) $V_{bias} = 253\,V$ (f) $V_{bias} = 295\,V$

Figure 5.3: DUT A hit detection efficiency maps for a 2×2 pixel grid, with orthogonal beam incidence and for increasing values of the applied bias voltage. The efficiency colour scales are the same, and start at 90% efficiency.

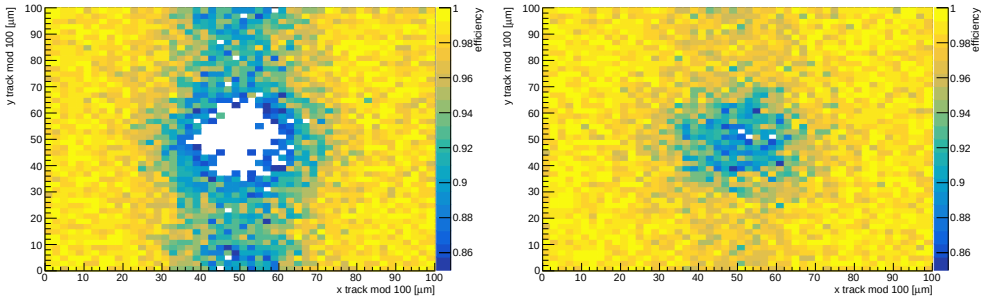


(a) $V_{bias} = 94\,V$ (b) $V_{bias} = 142\,V$

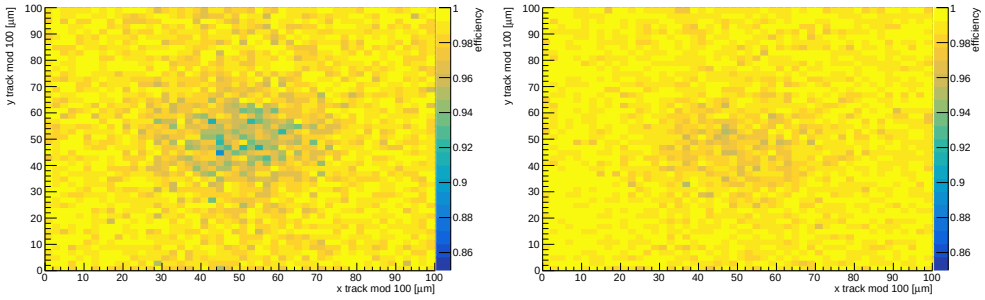
(c) $V_{bias} = 192\,V$ (d) $V_{bias} = 241\,V$

(e) $V_{bias} = 340\,V$ (f) $V_{bias} = 390\,V$

Figure 5.4: DUT B hit detection efficiency maps for a 2×2 pixel grid, with orthogonal beam incidence and for increasing values of the applied bias voltage. The efficiency colour scales are the same, and start at 85% efficiency.



(a) $\alpha = 0^\circ$ (b) $\alpha = 12^\circ$



(c) $\alpha = 20^\circ$ (d) $\alpha = 26^\circ$

Figure 5.5: DUT B hit detection efficiency maps for a 2×2 pixel grid, for different rotation angles around the y axis, with $V_{bias} = 390\,V$. The efficiency colour scales are the same, and start at 85% efficiency.

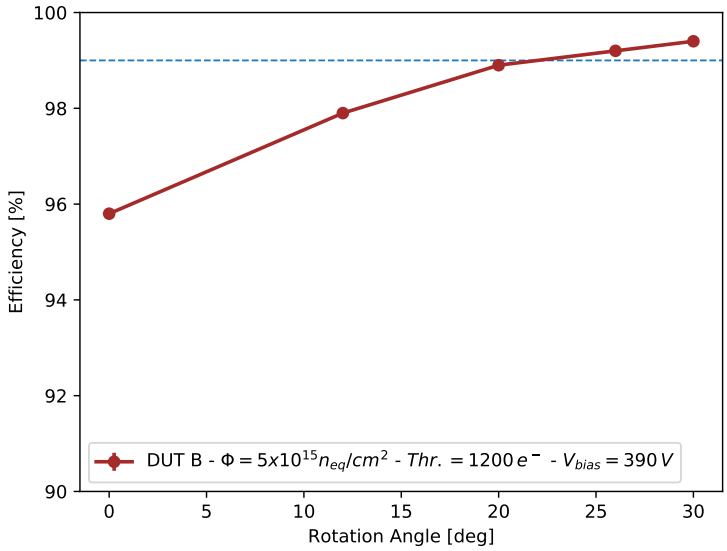


Figure 5.6: DUT B hit detection efficiency versus the rotation angle around the y axis, with $V_{bias} = 390\,V$. The dashed line represents 99% efficiency.

four pixels. The global efficiency is greater than 99% with a rotation of 20°, as shown in Figure 5.6.
In Figure 5.7 the average cluster size map for a 2×2 pixel grid are shown, for both DUTs at the same bias voltage, and perpendicular tracks. Similarly to the hit efficiency maps, the coloured scale indicates the average cluster size. The cluster size increases at the borders of the pixels, as expected due to charge sharing by diffusion. In DUT B an asymmetry due to the bias dot implant can be observed. Indeed, the distance between two adjacent pixels in a row (22 µm) is slightly larger than the distance between two adjacent pixels in a column (14 µm), due to the the presence of the metal route that connects the bias dot structure to the bias ring at the periphery of the sensor. In Figure 5.8 the average cluster size versus the bias voltage is reported. At high bias voltages, the average cluster size is about 1.2 for both DUTs, despite the lower charge sharing effect in DUT B: this is due to the lower average threshold achieved with the DUT B with respect to the DUT A.
Sensors without the punch-through structure are now the baseline for CMS, since the inefficiencies in the bias dot region are especially pronounced at vertical beam incidence, which is important for tracking in the forward regions of the future CMS Inner Tracker.

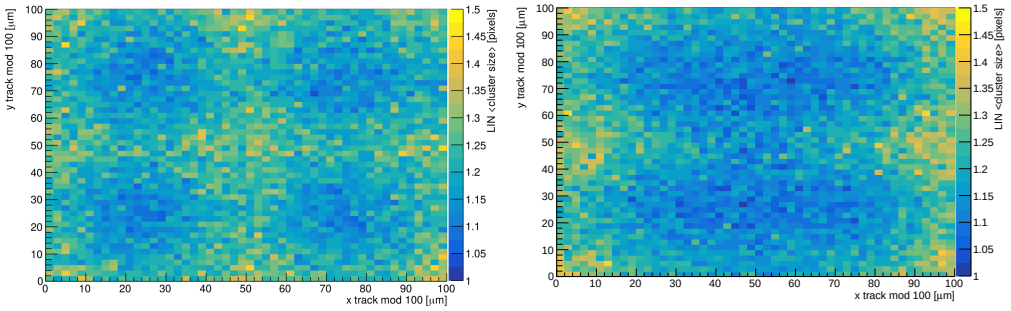


Figure 5.7: Cluster size for a 2×2 pixel grid for DUT A (left) and the DUT B (right), with $V_{bias} = 295\,V$ for both DUTs and with orthogonal beam incidence.

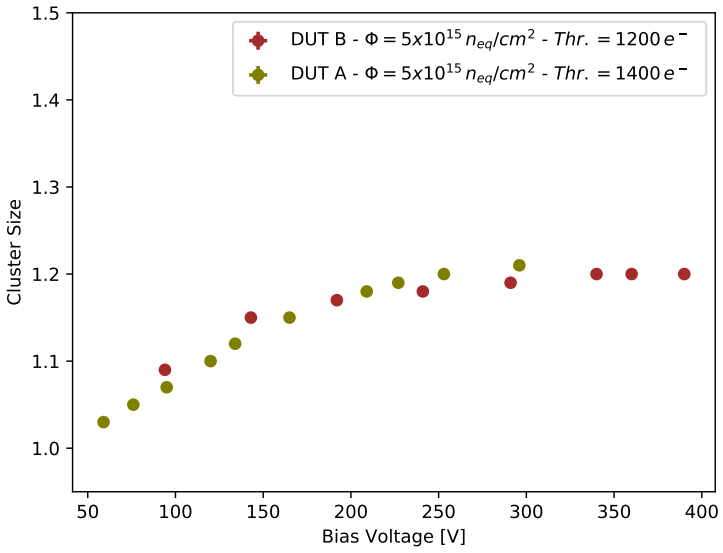


Figure 5.8: Average cluster size versus the applied bias voltage for DUT A and DUT B, with orthogonal beam incidence.

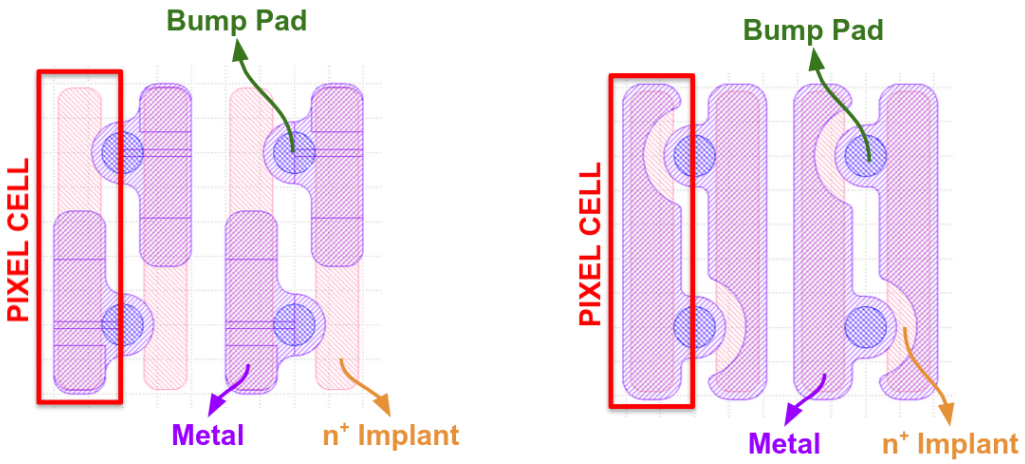
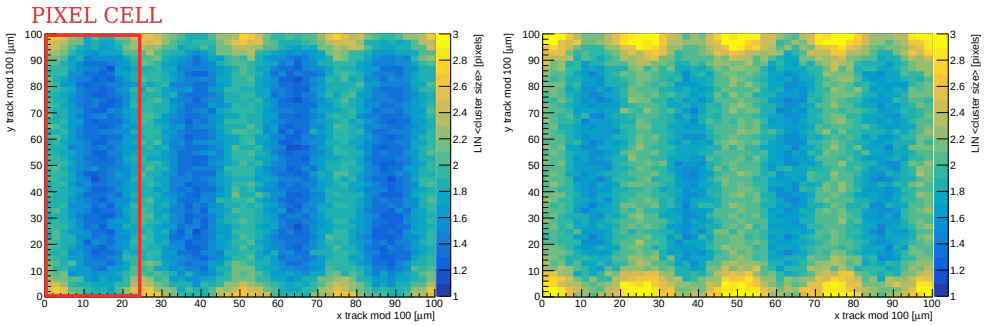


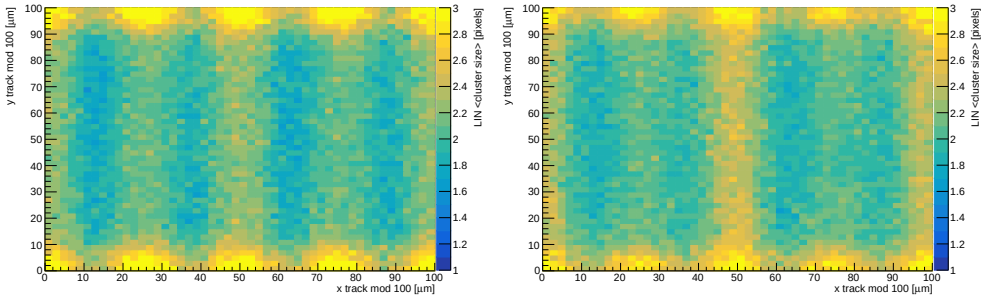
Figure 5.9: Schematic drawing of a 1×4 pixel grid of the DUT E (left) and the DUT D (right). The 25×100 µm^2 pixel cells are highlighted in red.

5.2.2 Fresh 25 µm Pitch Detectors

Two not irradiated (fresh) 25×100 µm^2 planar pixel detectors, both from the P-2 batch and referred to a DUT D and DUT E, were tested on beam. Both DUTs feature a bitten implant, and in DUT D the metal covers almost all the pixel implant (field plate design). Module E has also been tested after irradiation, as reported in the following. In Figure 5.9 drawings of a 4×1 pixel grid are shown for both DUTs.
The thresholds were tuned to an average pixel threshold of 800 electrons for both DUTs, with a dispersion of about 50 electrons. The leakage current was about 0.5 µA at a bias voltage of 70 V for both DUTs. The hit detection efficiency is greater than 99% at a bias voltage of 10 V for both DUTs.
In Figure 5.10 the DUT E cluster size maps for a 4×1 pixel grid are shown, for increasing bias voltages. As expected, the cluster size is higher at the intersection of two pixels, due to charge sharing. The cluster size initially increases with the bias voltage: this effect can be connected with cross-talk. At low bias voltages the sensors in under-depleted, therefore the collected charge is low. While it is enough to be above threshold and register a hit (the efficiency is $> 99\%$ even with $V_{bias} = 10$ V), it is not enough to activate the cross-talk coupled pixels. Increasing the bias voltage, the collected charge increases as well, and the cross-talk starts to be visible. This behaviour is in competition with another effect: increasing the bias voltages leads to an increase of the the electric field, hence reducing the charge sharing by diffusion. This effect tends to decrease the cluster size with increasing bias voltages.
While at low bias voltage the charge sharing is symmetric for the four pixel



(a) $V_{bias} = 10\,V$ (b) $V_{bias} = 20\,V$



(c) $V_{bias} = 30\,V$ (d) $V_{bias} = 70\,V$

Figure 5.10: DUT E cluster size maps for a 4×1 pixel grid, with orthogonal beam incidence and for increasing values of the applied bias voltage.

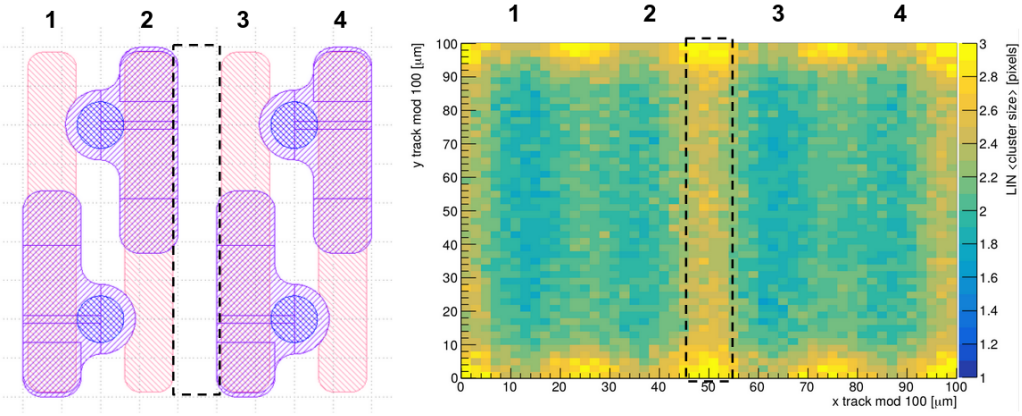
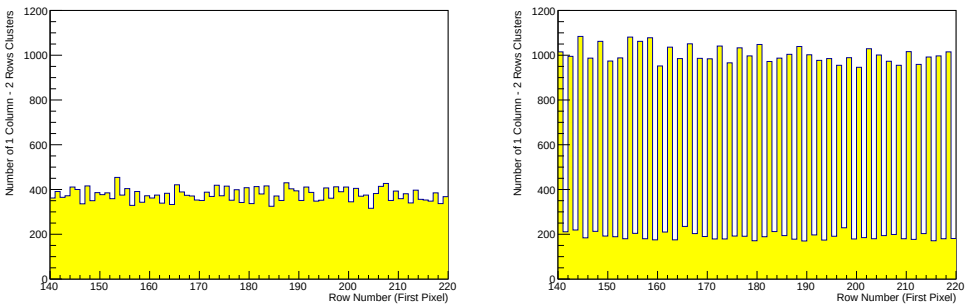
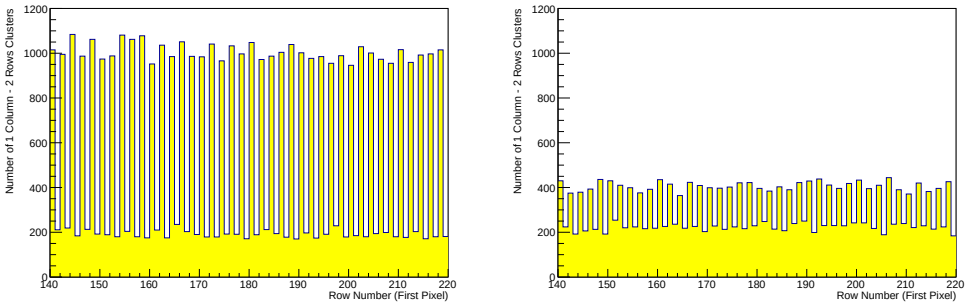


Figure 5.11: Schematic drawing of a 1×4 pixel grid of the DUT E (left) and the corresponding cluster size map, with $V_{bias} = 70\,V$ and perpendicular tracks (right). See text for details.



(a) $V_{bias} = 10\,V$ (b) $V_{bias} = 70\,V$

Figure 5.12: DUT E row number of the first pixel in a cluster of one column and two rows, for different values of the applied bias voltage and with orthogonal beam incidence.



(a) Threshold $= 800\,e^-$ (b) Threshold $= 1450\,e^-$

Figure 5.13: DUT E row number of the first pixel in a cluster of one column and two rows, for different pixel thresholds, with $V_{bias} = 70\,V$ and orthogonal beam incidence.

matrix, at high bias voltages the cluster size is higher at the intersection of non-coupled pixels, that is pixels between which cross-talk is not present. When a particle passes through this area, as shown in Figure 5.11, there is regular charge sharing between pixel 2 and 3, but also cross-talk induced from pixel 2 to pixel 1 (and from pixel 3 to pixel 4). This behaviour artificially increases the cluster size in this area. Figure 5.12 further confirms the cross-talk effect: it shows the row number for the first pixel in a cluster of one column and two rows. With $V_{bias} = 70, V$, a large asymmetry can be observed between even and odd rows, due to sharing between coupled pixels. At a lower bias voltage, however, this effect disappears, confirming the interpretation of Figure 5.10. Figure 5.13 shows the threshold effect on cross-talk: by increasing the average pixel threshold from 800 e^- to 1450 e^- the effect is significantly reduced.
The same behaviour observed in DUT E is also present in DUT D, due to the same cross-talk effect. In Figure 5.14 the average cluster size versus the bias voltage is reported, for DUT D and DUT E. Around $V_{bias} = 30$ V the cluster size reaches a maximum of about 2.1 in both cases. Slight differences between the two DUTs can be explained by slightly different threshold distributions. It should be noted that cross-talk measurements at a test beam can only be qualitative, due to the additional presence of regular charge sharing between adjacent pixel. Section 3.3 presents systematic cross-talk studies performed in laboratory.
In order to evaluate the pixel spatial resolution as a function of the rotation, measurements at different DUT rotation angles around the y axis (along the 25 µm pitch) were performed. In Figure 5.15 the average cluster size as a function of the rotation angle is reported. As expected, the cluster size increases with the angle due to the higher charge sharing, as the incoming tracks pass through more pixels. The increase in charge sharing is expected to improve the resolution as well.
In Figure 5.16 the resolution of both DUTs versus the rotation angle around the y axis is reported, with $V_{bias} = 70$V. The resolution of both DUTs is well below the digital resolution (Equation 4.12). The results of the two DUTs are compatible within the errors, as expected from the similar pixel layout. The error on the resolution has been propagated as explained in Section 4.3. The rotation angle has been derived from the DUT alignment procedure, with an estimated error of $\pm 0.5°$.
The minimum resolution, close to 2 µm, is achieved around 8-9°. This is expected from the geometry of the pixel sensor, because tracks passing through the sensor with this angle always release charge in at least two pixels, which leads to a better resolution. A sketch of this configuration is reported in Figure 5.17. These sensors are 150 µm thick, and considering the 25 µm pitch,

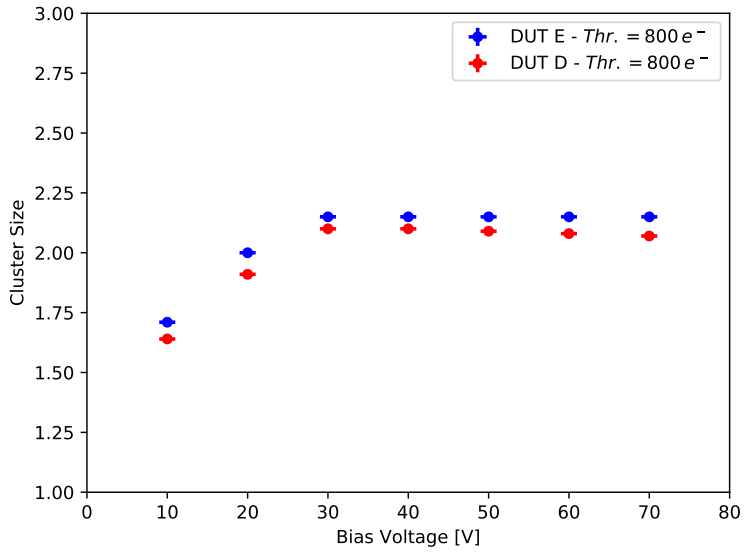


Figure 5.14: DUT D and DUT E cluster size versus the applied bias voltage, for perpendicular tracks.

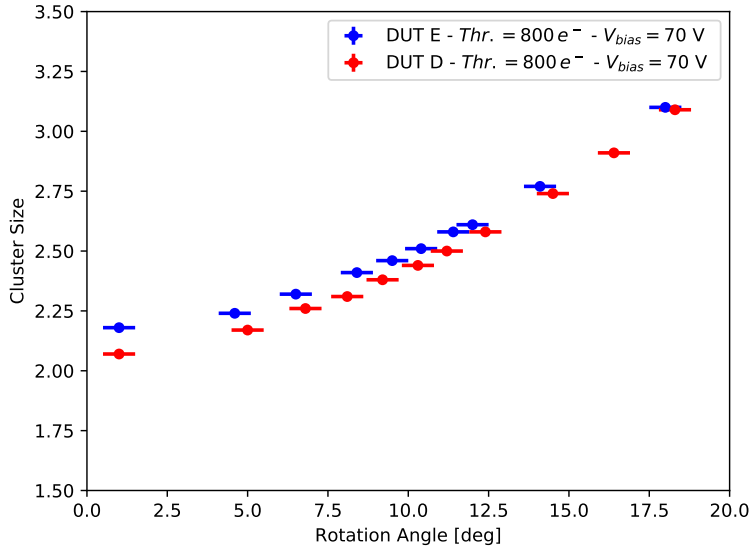


Figure 5.15: DUT D and DUT E cluster size versus the rotation angle around the y axis (along the 25 µm pitch), with $V_{bias} = 70\,V$.

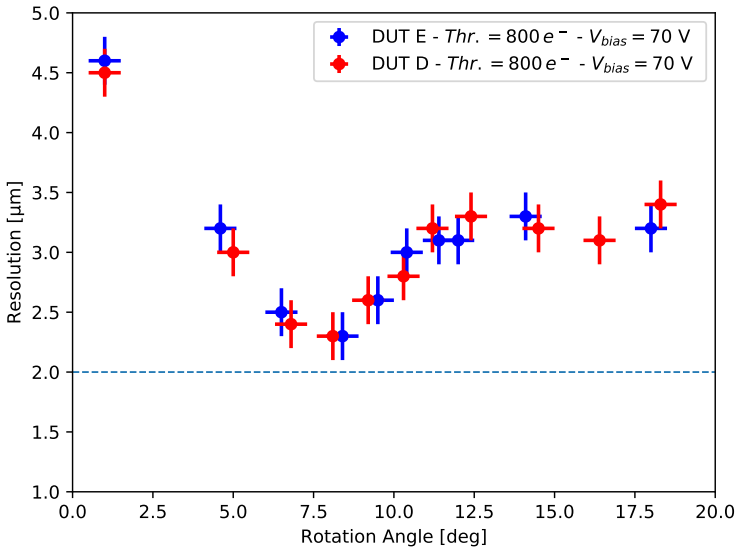


Figure 5.16: DUT D and E resolutions versus the rotation angle around the y axis (along the 25 µm pitch), with $V_{bias} = 70$ V. The horizontal line marks 2 µm resolution.

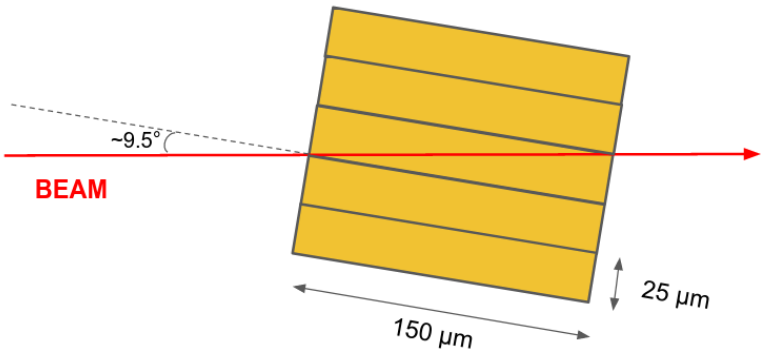


Figure 5.17: Sketch of the DUT pixel position with respect to the beam line. This is the optimal rotation angle for resolution. See text for details.

the optimal resolution angle is expected to be:

$$\alpha_{opt} = \tan^{-1}\left(\frac{\text{Pixel Pitch}}{\text{Active Thickness}}\right) \tag{5.1}$$

which is about 9.5° in this case. The measured resolution is compatible within the errors. Around this minimum, steep slopes are present, deteriorating the position resolution for angles deviating only slightly from the optimal incidence angle. This effect is related to the Center of Gravity algorithm used in cluster reconstruction. This algorithm has a strong dependence on the track incidence angle and only yields the best possible resolution at the optimal charge sharing angle for two-pixel clusters. Corrections of this behaviour have been explored in literature, but are not object of this Thesis [57].
Another local minimum is expected around 18°, which is the geometrical configuration for which cluster are at least of size three. Indeed, the resolution pattern hints for a minimum around this angle. Higher rotation angles have not been studied because the alignment software was not optimised for high rotation angle studies.
Figure 5.18 shows the resolution as a function of the bias voltage. At higher bias voltages the resolution increases due to the cross-talk effect, that artificially increases the cluster size. Indeed, Figure 5.19 shows the residual distributions with $V_{bias} = 70$ V calculated on the indicated sub-zones. Between pixels 2 and 3, for instance, cross-talk can turn on pixels 1 or 4: this results in the "horned" residual distribution, which slightly spoils the spatial resolution. The residual distribution evaluated between pixels 1 and 2 (or 3 and 4) is regular, because in this case the cross-talk contributes to regular charge sharing. The cross-talk effect on residuals is dominant only for perpendicular tracks: even with small rotation angles the resolution drops to really low values, as seen in Figure 5.16, and consistent with resolution measured in 3D sensors (as shown in Section 5.3), in which the cross-talk effect is negligible.
For DUT D charge calibrations were made, making it possible to measure the collected charge. Figure 5.20 shows the charge distribution for DUT D, with $V_{bias} = 70$ V, superimposed with a fit to a Landau distribution convoluted with a gaussian distribution. The gaussian part is due to the pixel calibrations (as shown in Section 3.2, the dispersion can be significant) and electronic noise. Only cluster sizes one, two, three and four have been considered for the collected charge distribution.
The MPV of the Landau has been reported in Figure 5.21, as a function of the applied bias voltage. At low bias voltages the sensor in under-depleted,

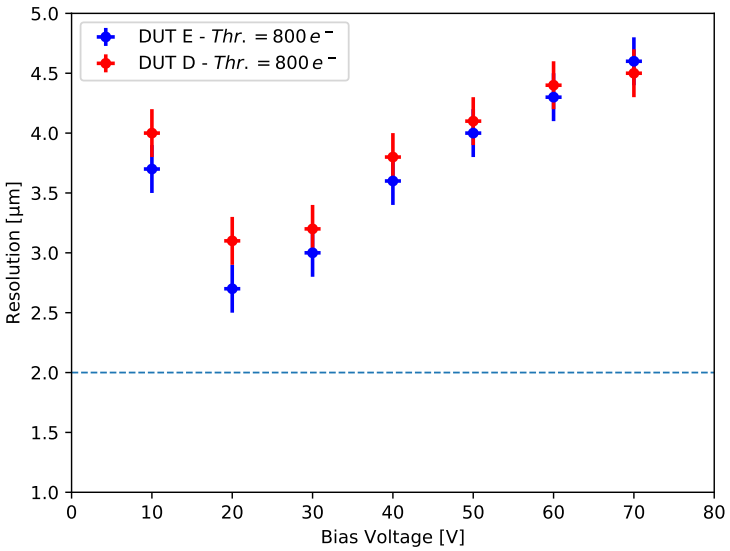


Figure 5.18: DUT D and DUT E resolutions versus the applied bias voltage, with orthogonal beam incidence. The horizontal line marks 2 µm resolution.

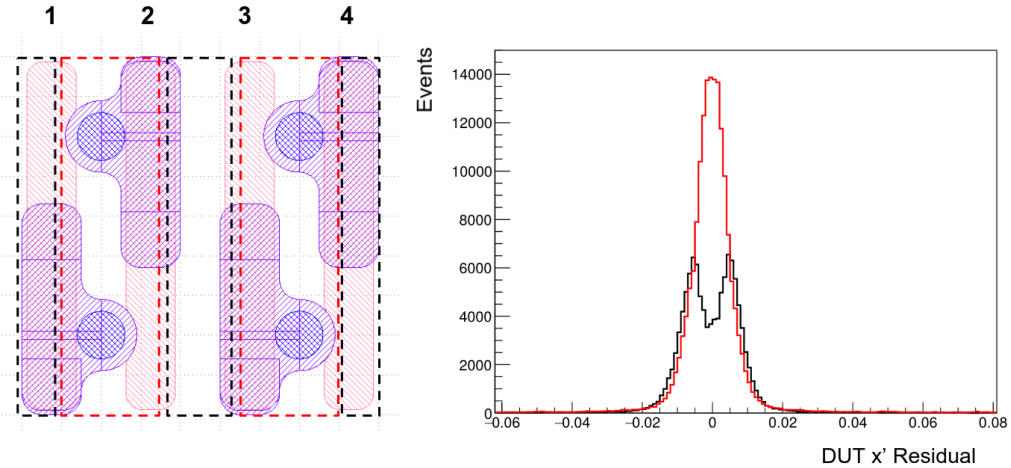


Figure 5.19: Schematic drawing of a 1×4 pixel grid of the DUT E (left) and the residual distributions evaluated on the corresponding areas of the same color (right), with $V_{bias} = 70\,V$. See text for details.

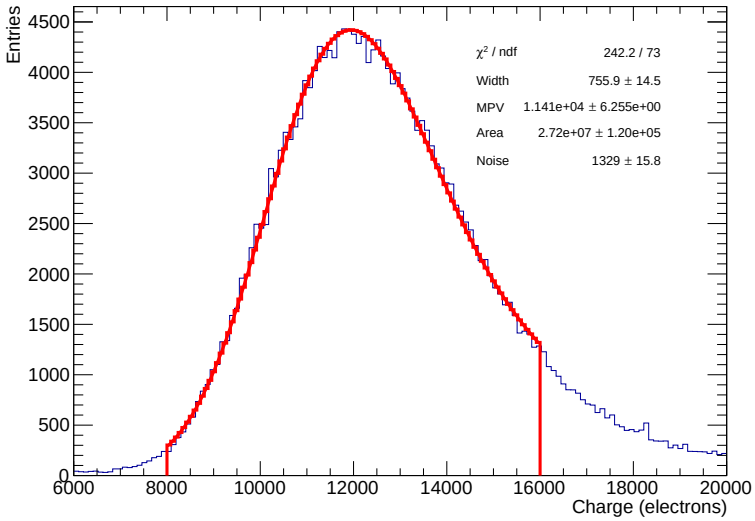


Figure 5.20: DUT D collected charge distribution with $V_{bias} = 70\,V$. The red line represents a fit to a Landau distribution convoluted with a gaussian distribution.

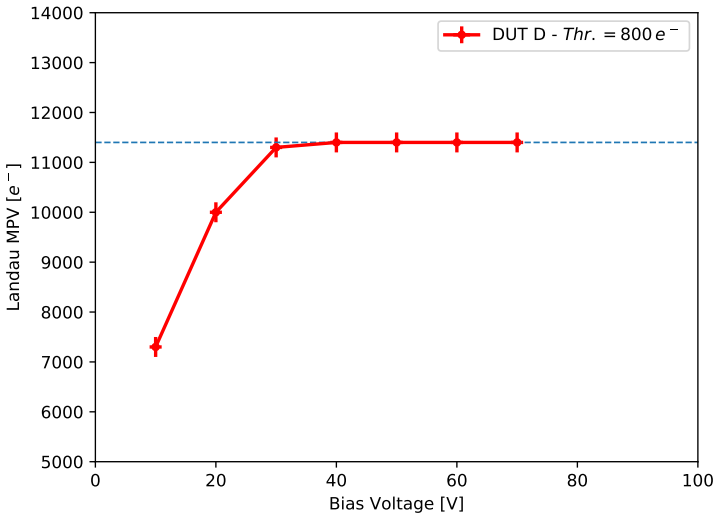


Figure 5.21: DUT D collected charge versus the applied bias voltage. The charge values refer to the MPVs from the fit of a Landau distribution convoluted with a gaussian distribution for each bias voltage. The dashed line represents the expected Landau MPV for a MIP traversing 150 µm thick silicon.

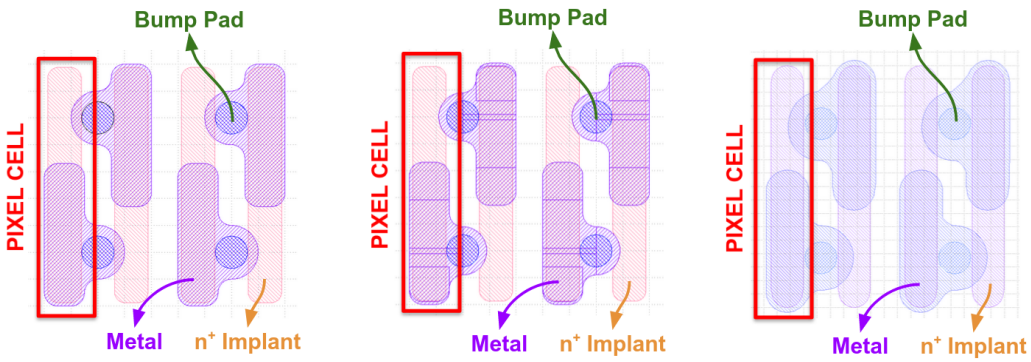


Figure 5.22: Schematic drawing of a 1 × 4 pixel grid of DUTs G and E (left), DUT F (center) and DUT C (right). The 25 × 100 µm^2 pixel cells are highlighted in red.

therefore less charge is collected. Around $V_{bias} = 30\,V$ the sensor is fully depleted and the collected charge saturates around 11400 electrons, which is close to the theoretical expectation[6].
Errors on the Landau MPV have been estimated to be $\pm 200\,e^-$ from Equation 3.1, using the Landau MPV in ΔV_{cal} units. While this is valid only for cluster size one, restricting the collected charge distribution to clusters of size one do not affect the measured Landau MPV and hence the associated error.

5.2.3 Irradiated 25 µm Pitch Detectors

Four irradiated planar 25 × 100 µm^2 pixel detectors (referred to as DUT C, DUT E, DUT F and DUT G) were tested on beam. DUT C is from the the P-1 batch, and has an active thickness of 130 µm. DUTs E, F and G are from the P-2 batch, and have an active thickness of 150 µm. In Figure 5.22 schematic drawings of a 1 × 4 pixel grid are shown for all the DUTs. DUTs C and F have standard, non-bitten implants, while DUTs E and G have the bitten implant. All the DUTs were irradiated at KIT at increasing fluences: DUT C to a fluence of $7.5 \times 10^{15}\,n_{eq}/cm^2$, DUT E to $11 \times 10^{15}\,n_{eq}/cm^2$, DUT F to $18 \times 10^{15}\,n_{eq}/cm^2$ and DUT G to $24 \times 10^{15}\,n_{eq}/cm^2$, which is higher than the expected fluence in the innermost tracker layer after ten years of operations at the HL-LHC.
The thresholds of the DUTs were tuned inside the cooling box, to an average pixel threshold of 1150 electrons (DUTs C and G), 1250 electrons (DUT E)

[6]For a MIP, the electrons-holes number MPV is 76 $e^-\mu m^{-1}$: with an active thickness of 150 µm, the MPV is = 11400 e^-.

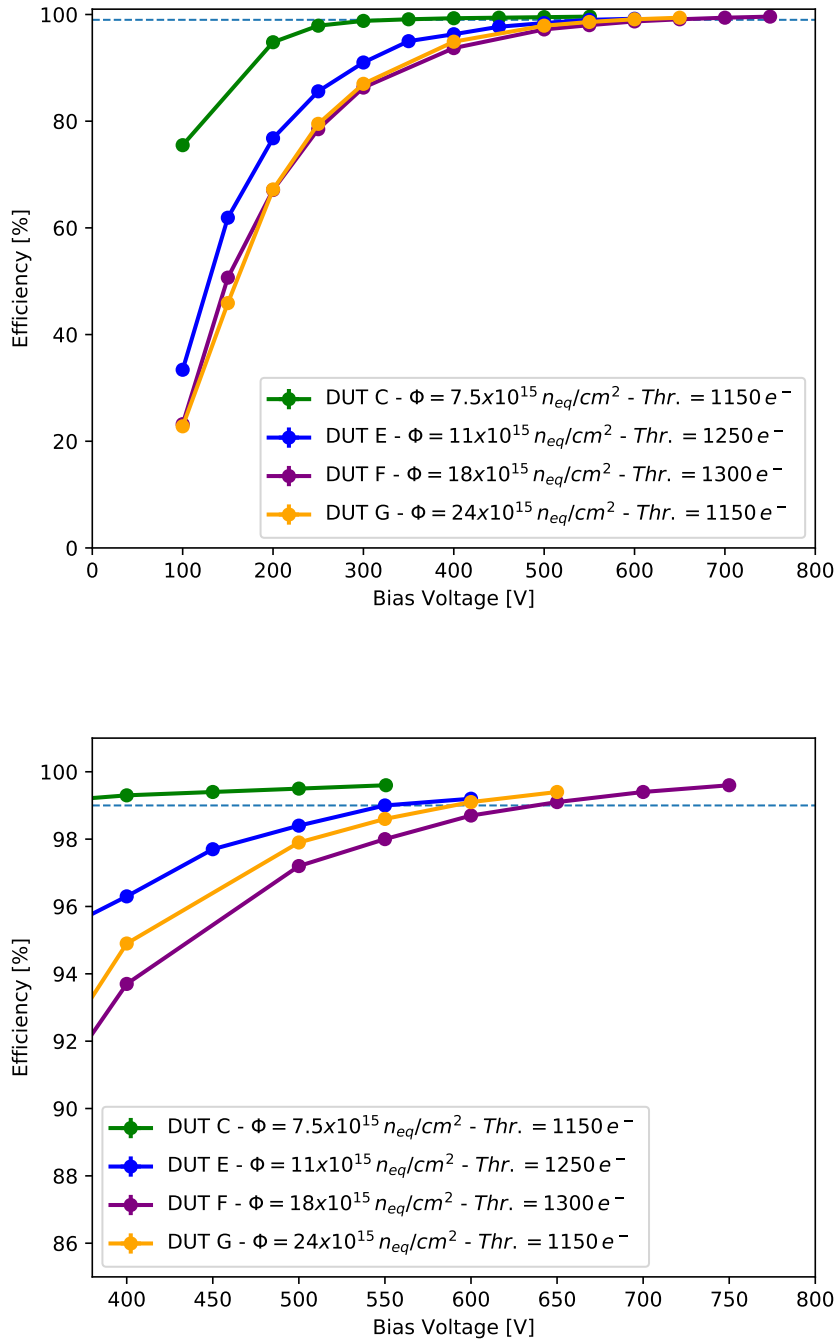
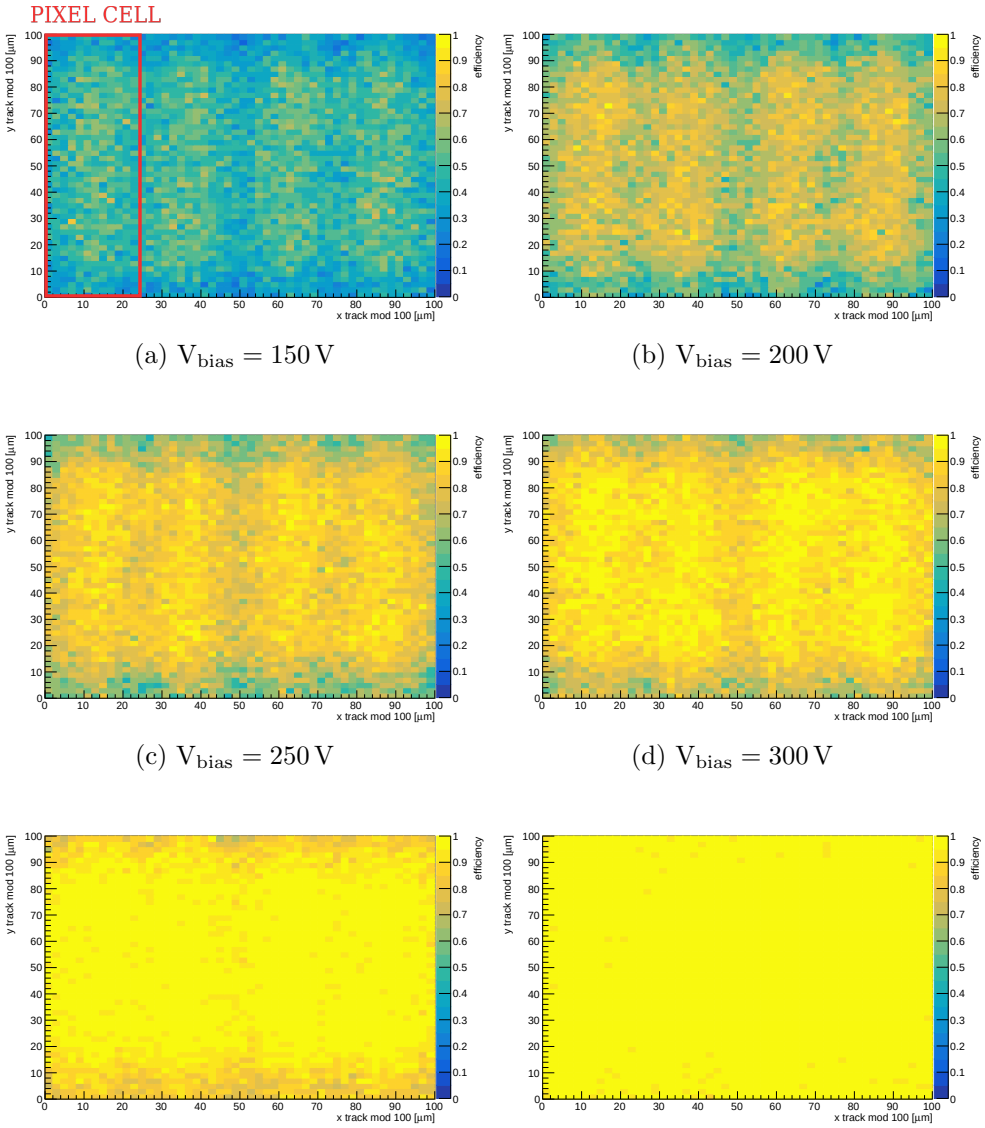


Figure 5.23: Hit detection efficiency for DUTs C, E, F and G versus the applied bias voltage with orthogonal beam incidence. The dashed line represents 99% efficiency.



(a) $V_{bias} = 150\,V$ (b) $V_{bias} = 200\,V$

(c) $V_{bias} = 250\,V$ (d) $V_{bias} = 300\,V$

(e) $V_{bias} = 400\,V$ (f) $V_{bias} = 650\,V$

Figure 5.24: DUT G hit detection efficiency maps for a 4×1 pixel grid, with orthogonal beam incidence and for increasing values of the applied bias voltage.

and 1300 electrons (DUT F) with a dispersion of about 50 electrons. For the resolution measurements DUT F was tuned again to a lower threshold of 1200 electrons. The tuning was made targeting low thresholds and noise, having at most 1% noisy pixel channels. Due to higher noise in irradiated detectors, thresholds are higher with respect to fresh DUTs.
The DUTs were kept at a temperature of about -27 ℃. At this temperature, the leakage currents were $\simeq$ 50 µA at 500 V for DUT C, $\simeq$ 140 µA at 500 V for DUT E, $\simeq$ 200 µA at 600 V for DUT F and $\simeq$ 620 µA at 600 V for DUT G. DUTs F and G where coated with parylene before being irradiated, in order to prevent spark issues after irradiation, at very high bias voltages. Both DUTs were tested in laboratory up to an unprecedented value of 1200 V, without any sign of sparks, proving the effectiveness of parylene coating.
Figure 5.23 shows the hit detection efficiency for the four irradiated DUTs, as a function of the bias voltage. The efficiency of DUT C, irradiated at a fluence of $7.5 \times 10^{15}\,\mathrm{n_{eq}cm^{-2}}$ is greater than 99% around 300 V, while for DUTs F and G, irradiated at $18 \times 10^{15}\,\mathrm{n_{eq}/cm^2}$ and $24 \times 10^{15}\,\mathrm{n_{eq}/cm^2}$ respectively, bias voltages greater than 600 V are necessary. While DUT G was irradiated to a higher fluence with respect to DUT F, the hit detection efficiency is higher for DUT G (with the same bias voltage), due to the lower average pixel threshold.
Planar pixel sensors were not expected to survive the extreme fluences of DUTs F and G. In particular, the irradiation fluence of module G corresponds to the expected fluence after ten years of operation in the innermost layer of the Phase-2 CMS tracker. It was originally planned to substitute the planar detectors in the innermost tracker layer[7] after five years of operations of HL-LHC (due to the extreme radiation damages): however, in light of these studies, it might not be necessary. Moreover, the RD53A ROC worked without significant problems after receiving a dose of about 30 MGy, six times the certification. The obtained results are therefore very encouraging towards the construction of the Phase-2 CMS tracker.
In Figure 5.24 DUT G hit detection efficiency maps for a 4×1 pixel grid are reported, with orthogonal beam incidence and increasing values of the bias voltage. The efficiency increases with the bias voltage as the sensor gets more depleted and more charge is collected. At low bias voltage values, the efficiency is lower at the intersection of adjacent pixels due to charge sharing between the pixels and charge trapping in the highly irradiated silicon. It is more probable for a hit to be under-threshold if the particle passes though the intersection of two or four pixels. By increasing the bias voltage, and consequently the electric field, the efficiency becomes more uniform across

[7] The alternative consists in using 3D pixel detectors in the innermost tracker layer.

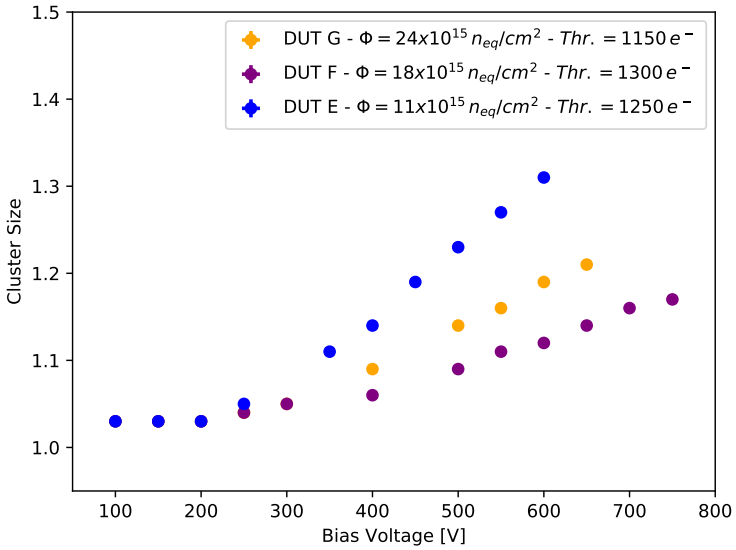


Figure 5.25: Average cluster size versus the applied bias voltage for DUTs E, F and G, with orthogonal beam incidence.

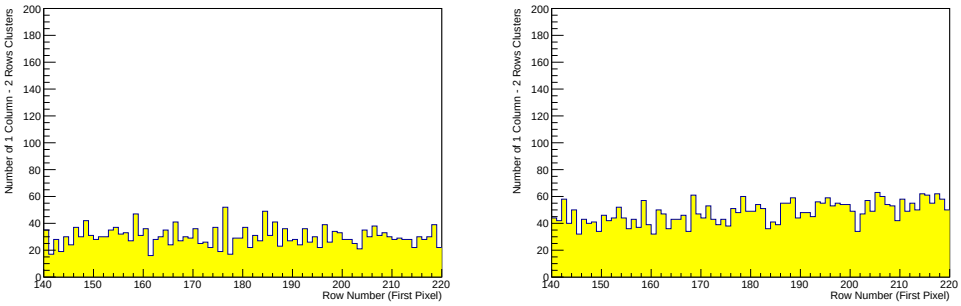


Figure 5.26: Row number of the first pixel in a cluster of one column and two rows for DUT F (left) and DUT G (right), with $V_{bias} = 650\,V$ in both cases and orthogonal beam incidence.

the pixel cells.
The average cluster size versus the applied bias voltage for DUTs E, F and G is reported in Figure 5.25. While the cluster size is close to one for low bias voltages, it starts increasing at about 300 V for all three DUTs. The cluster size of DUT E, the less irradiated, is higher than the other two, due to less charge trapping. Also in this case, DUT G shows a higher average cluster size with respect to DUT F, despite a higher irradiation fluence, due to the lower average threshold.
It should be noted that the cluster size is significantly lower with respect to fresh DUTs. Due to charge trapping, two effects sum up: the charge sharing between adjacent pixels is reduced as is the collected charge. The latter suggests that the cross-talk is absent in this highly irradiated detectors. Indeed, Figure 5.26 shows the row number of the first pixel in a cluster of one column and two rows for DUTs F and G: no asymmetries between even and odd rows can be observed, confirming the absence of cross-talk. The higher thresholds with respect to fresh DUTs also contribute to the cross-talk suppression.
Also in the case of irradiated DUTs, spatial resolution measurements were performed. Figure 5.27 shows the average cluster size as a function of the DUT rotation angle around the y axis, along the 25 µm pitch. For these measurements, the thresholds of DUT F were tuned to a lower average value. The average cluster size increases with the rotation angle, and is higher for the less irradiated DUTs. Due to the lower threshold of DUT F, the measurements for the three DUTs scale correctly with respect to the received irradiation fluence.
The spatial resolution versus the rotation angle around the y axis is reported in Figure 5.28. As explained in section 4.3, only three planes of the telescope are used in the case of irradiated DUTs. Therefore, the telescope resolution has been evaluated with a simulation. The spatial resolution is close to the digital value for low rotation angles, but starts decreasing for higher angles. The measurements are compatible for the three DUTs. The resolution deterioration with respect to fresh DUTs is due to the lower cluster size due to charge trapping and due to the higher thresholds. Despite the deterioration, resolutions close to 6 µm can be achieved, even after extreme irradiation fluences.

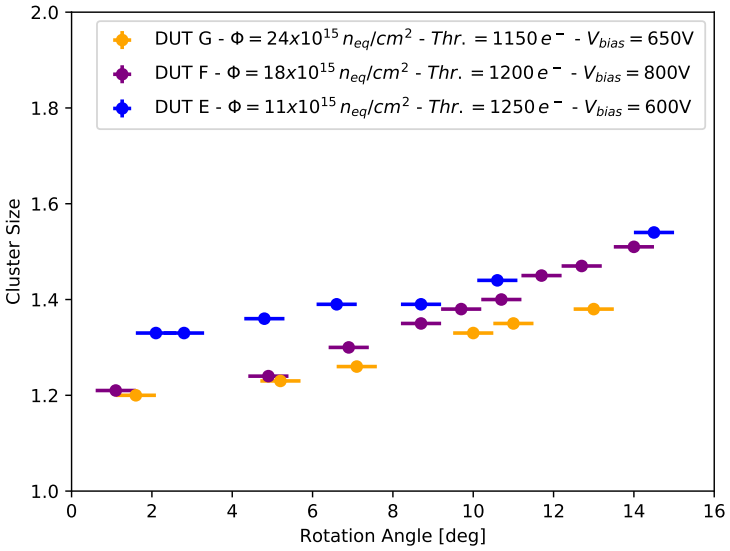


Figure 5.27: Average cluster size versus the rotation angle around the y axis (along the 25 µm pitch) for DUTs E (with $V_{bias} = 600\,V$), F ($V_{bias} = 800\,V$) and G ($V_{bias} = 600\,V$).

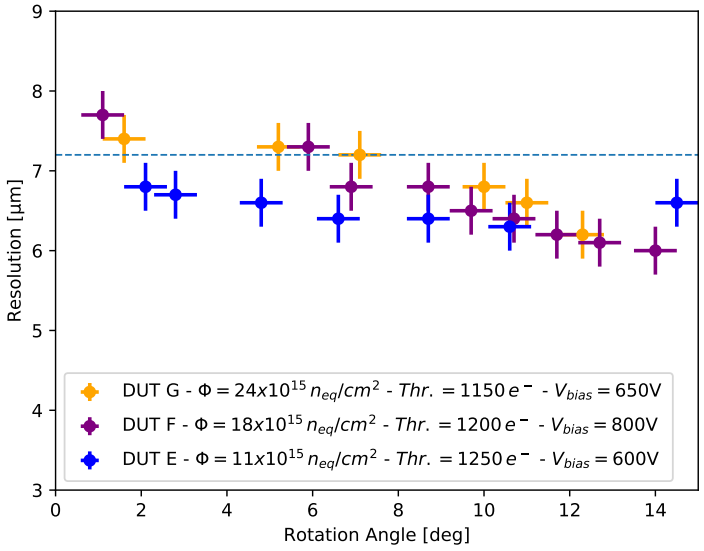


Figure 5.28: Resolution versus the rotation angle around the y axis (along the 25 µm pitch) for DUTs E (with $V_{bias} = 600\,V$), F ($V_{bias} = 800\,V$) and G ($V_{bias} = 600\,V$). The horizontal line marks the digital 7.2 µm resolution.

5.3 3D Pixel Detectors

The complete list of the tested 3D detectors is reported in Table 5.2. The DUT letter (J to Y) is used to reference the detectors. Sensors from the 3D-1 batch have an active thickness of 130 µm and a total thickness of 200 µm, while sensors from the 3D-2 batch have an active thickness of 150 µm and a total thickness of 250 µm.
Only one 3D detector was tested after irradiation, to a fluence of $10 \times 10^{15}\,n_{eq}/cm^2$. Other 3D detectors were irradiated, but due to sudden mechanical failures of the single chip cards used to host the detectors, they could not be tested. Irradiation of other 3D detectors are ongoing and they will hopefully be tested in the near future. In any case, the only tested irradiated 3D detector is still significant for comparisons with planar detectors.

DUT	**Batch**	**Pixel Pitch [µm²]**	**Active Thickness [µm]**	**Irradiation Fluence [n_{eq}/cm^2]**	**On Beam**
J	3D-1	50×50	130	10×10^{15} (PS)	04/19
K	3D-1	25×100	130	-	11/19
W	3D-1	50×50	130	-	11/19
X	3D-2	50×50	150	-	11/19
Y	3D-2	25×100	150	-	06/20

Table 5.2: Summary of the 3D pixel detectors which were tested at DESY Test Beam. (PS) indicates that the detector was irradiated at the CERN PS IRRAD facility. The DUT letter will be used to refer to these modules in the following.

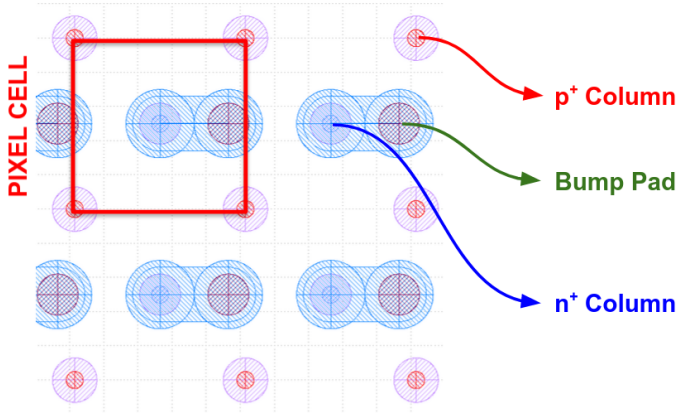


Figure 5.29: Schematic drawing of a 2×2 pixel grid of DUTs W and X. The 50×50 µm^2 pixel cell is highlighted in red.

5.3.1 Fresh 50 µm Pitch Detectors

Two fresh 50×50 µm^2 3D pixel detectors, referred to as DUT W and DUT X were tested on beam. DUT W is from the 3D-1 batch and has an active thickness of 130 µm, while DUT X is from the 3D-2 batch and has an active thickness of 150 µm.
The thresholds were tuned to an average pixel threshold of 950 electrons for DUT W and 900 electrons for DUT X, with a dispersion of about 50 electrons in both cases. The leakage current was about 0.2 µA at a bias voltage of 30 V for DUT W and 16 µA at a bias voltage of 30 V for DUT X. DUT X had a bias current much higher than it was when measured on sensor wafer (before dicing and flip-chipping). However even with the higher bias current the DUT was working without any impact on performance even in long overnight runs of over 25 millions collected triggers. No increase in noise figure or thermal drift due to the bias current was observed. The reason for taking data with high statistics and high bias voltage was to prove that 3D sensors can be safely operated with a large margin with respect to the depletion voltage (which is about 10 V for these sensors). This could eventually be needed because the future CMS Inner Tracker will be powered according to the serial powering scheme described in Section 2.4.1. In a serial powering chain of five modules, the high voltage bias drop between the first and the last module in the chain is about 10 V. Recently, fresh 3D detectors were safely operated up to 60 V in laboratory studies. Therefore, in serial

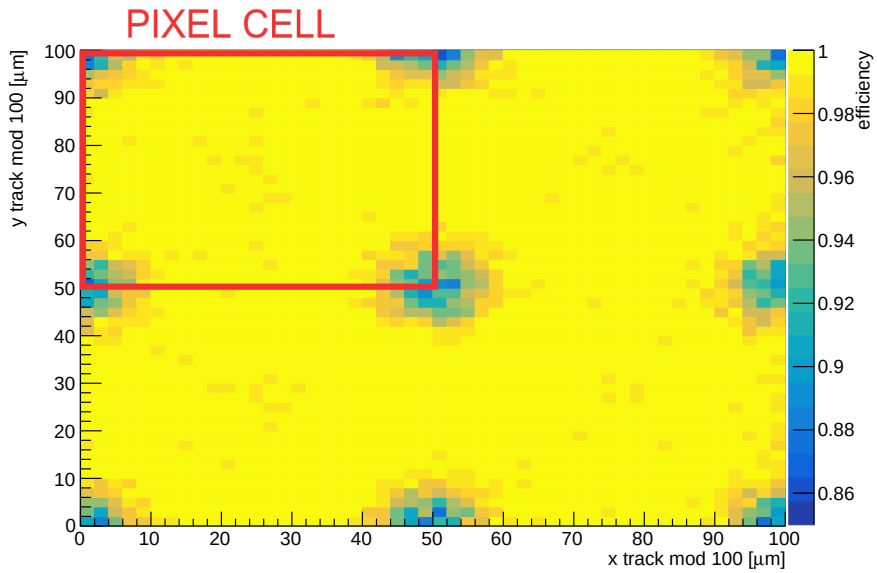


Figure 5.30: DUT X hit detection efficiency map for a 2×2 pixel grid, with $V_{bias} = 30\,V$ and orthogonal beam incidence.

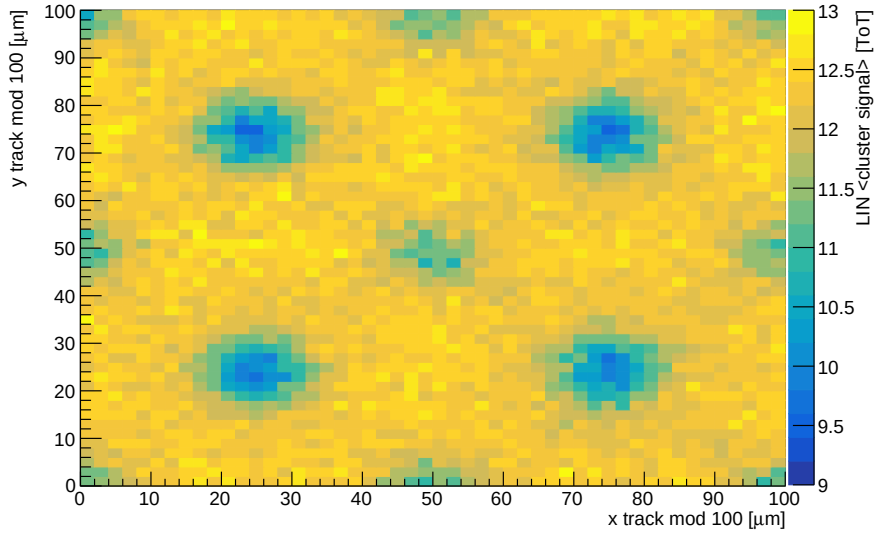
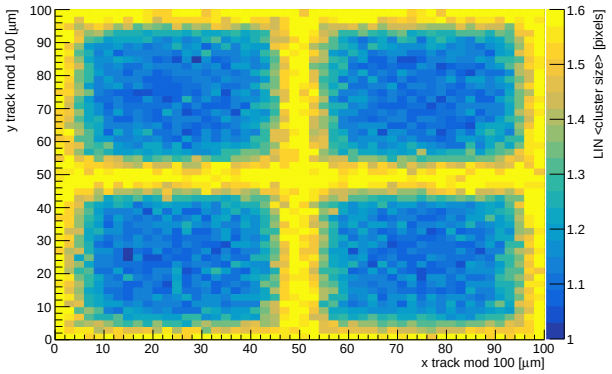
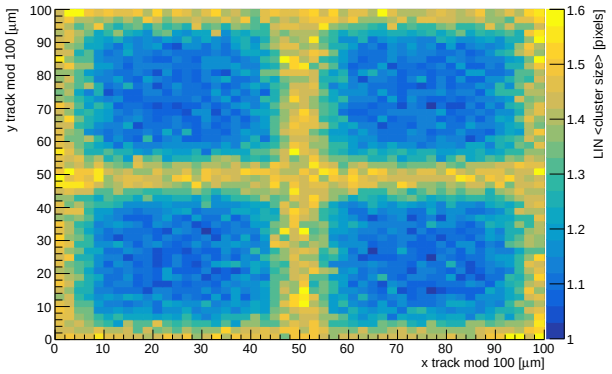


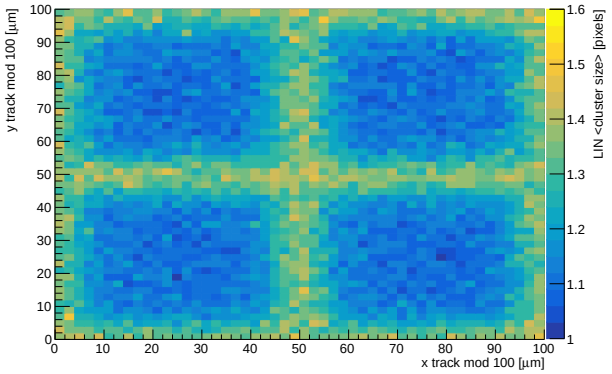
Figure 5.31: DUT X TOT map for a 2×2 pixel grid, with $V_{bias} = 30\,V$ and orthogonal beam incidence.



(a) $V_{bias} = 5\,V$



(b) $V_{bias} = 10\,V$



(c) $V_{bias} = 30\,V$

Figure 5.32: DUT X average cluster size maps for a 2×2 pixel grid, for increasing values of the applied bias voltage and with orthogonal beam incidence.

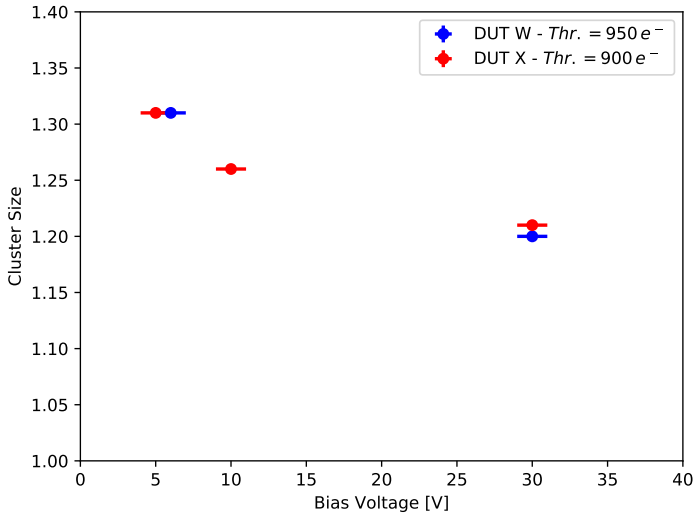
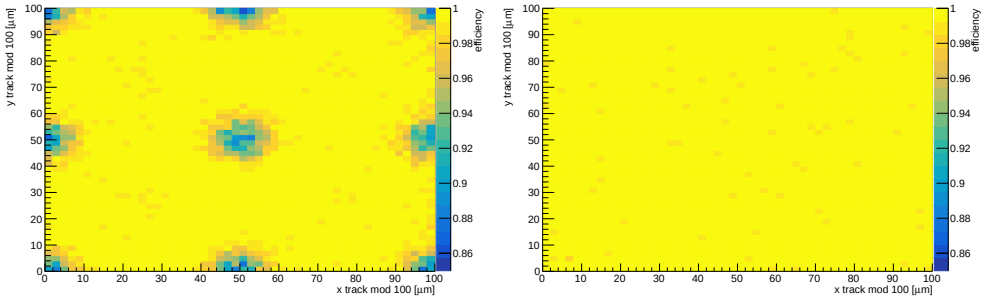


Figure 5.33: DUTs W and X average cluster size versus the applied bias voltage with orthogonal beam incidence.

powering chains, fresh 3D modules[8] could be operated from 20 V to 40 V. The hit detection efficiency is greater than 99% at a bias voltage of 5 V for both DUTs, with orthogonal beam incidence. Figure 5.30 shows the DUT X hit detection efficiency map for a 2×2 pixel grid and perpendicular tracks. Efficiency drops can be observed in correspondence to p^+ columns. Indeed, since columns are partially empty, when a particle passes through them, charge is not collected. However, since the beam is not perfectly perpendicular to the DUT, the incident particles can escape the columns: if enough charge is generated in the active silicon, the hit is registered. For this reason[9] the efficiency in correspondence to the p^+ columns is about 85%. The efficiency in correspondence to the n^+ columns is close to 100% despite being made of passive material as well. This is because the n^+ columns are the collecting electrodes, therefore it is highly probable that the (low) generated charge gets collected. Instead, if low charge is produced near the p^+ columns, it is split between adjacent pixels, making it likely for the involved pixels to be under-threshold. This effect in confirmed in Figure 5.31, which shows the DUT X average TOT map for a 2×2 pixel grid and perpendicular

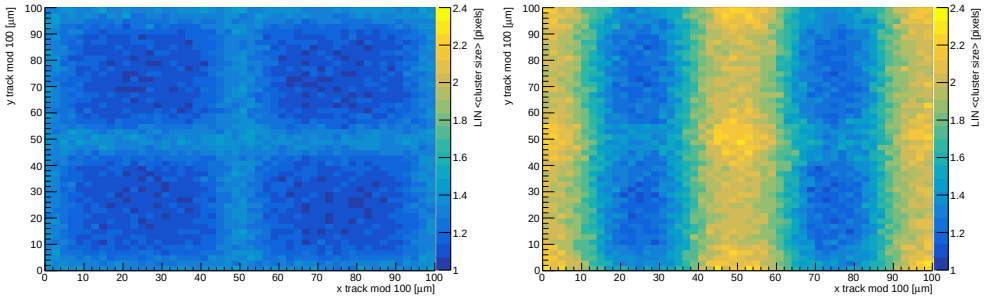
[8]Planar pixel detectors are usually operated around 70 V, twice the full depletion voltage. For this reason, planar pixel modules are less affected by a $\sim$ 10 V voltage drop along the chain.

[9]The telescope resolution, being higher than the column diameter (5 µm), further smears this effect.



(a) $\alpha = 0^\circ$ (b) $\alpha = 12^\circ$

Figure 5.34: DUT X hit detection efficiency maps for a 2×2 pixel grid, for different rotation angles around the y axis, with $V_{bias} = 30\,V$. The efficiency colour scales are the same, and start at 85% efficiency.



(a) $\alpha = 0^\circ$ (b) $\alpha = 12^\circ$

Figure 5.35: DUT X average cluster size maps for a 2×2 pixel grid, for different rotation angles around the y axis, with $V_{bias} = 30\,V$.

tracks. Indeed, hits in correspondence to n^+ columns are associated with a lower TOT (and therefore a lower collected charge) with respect to the rest of the pixel cell. Moreover, since hits in correspondence to p^+ columns need a higher collected charge to be over-threshold, hits in these regions feature a higher average TOT with respect to hits near n^+ columns.
Figure 5.32 shows DUT X average cluster size maps for a 2×2 pixel grid for increasing values of the bias voltage and perpendicular tracks. The cluster size is higher at the borders of a pixel cell due to charge sharing. Increasing the bias voltage reduces the cluster size in these regions: this is because the higher electric field reduces charge sharing by diffusion. Figure 5.33 shows the average cluster size as a function of the bias voltage, for DUTs W and X. As expected, the average cluster size decreases with the bias voltage for both DUTs.
The efficiency drops in correspondence to p^+ columns are visible only with orthogonal beam incidence. Figure 5.34 shows the DUT X hit detection efficiency map for a 2×2 pixel grid with orthogonal beam incidence and with a DUT rotation of 12° with around the y axis. In the latter case, no efficiency drops can be observed, because the incoming beam particles always escape from the column passive material and enough charge is collected for the involved pixels to be over-threshold.
Figure 5.35 shows the DUT X average cluster size map for a 2×2 pixel grid with orthogonal beam incidence and with a DUT rotation of 12° around the y axis. In the latter case, charge sharing is increased in one direction: this is because the incoming beam particles are likely to traverse two (or more) pixels since the DUT is rotated.
In order to perform resolution measurements, the DUTs were tested at various rotation angles around the y axis. Figure 5.36 shows the DUTs W and X average cluster size as a function of the rotation angle and with $V_{bias} = 30\,V$. For DUT W some data were taken also with $V_{bias} = 5\,V$. No significant difference between the two DUTs can be observed. The points with $V_{bias} = 5\,V$ have a slightly higher cluster size due to the higher charge sharing by diffusion.
In Figure 5.37 the resolution of DUTs W and X is shown, as a function of the rotation angle around the y axis. Both DUTs reach a minimum resolution of 5 µm at about 18°-19°. This result is compatible with the expected optimal angle (Equation 5.1), which should be 18° for DUT X (150 µm active thickness) and 21° for DUT W (130 µm active thickness). A second minimum, expected above 30°, is hinted by the resolution pattern.
Resolution for DUT W was also measured with $V_{bias} = 5\,V$: no significant differences in resolution can be observed between the two tested bias voltages. This is extremely important for the serial powering scheme, since it shows

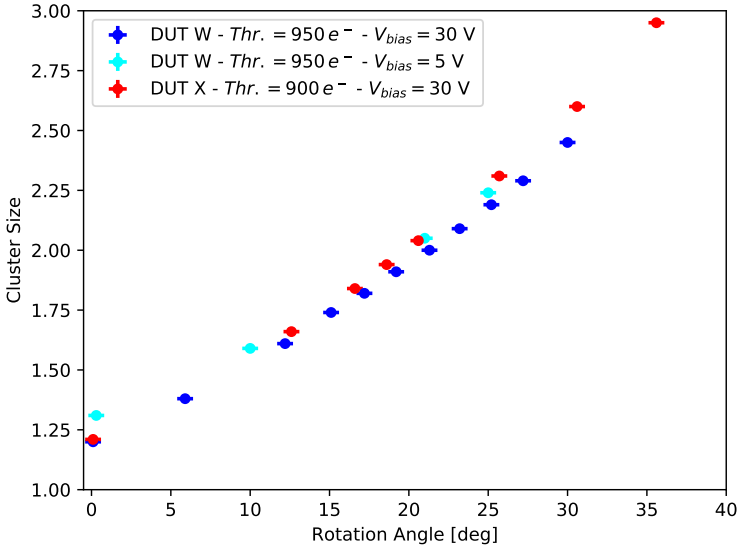


Figure 5.36: DUTs W and X average cluster size versus the rotation angle around the y axis, with $V_{bias} = 30$ V for both DUTs (for DUT W measurements with $V_{bias} = 5$ V are also reported).

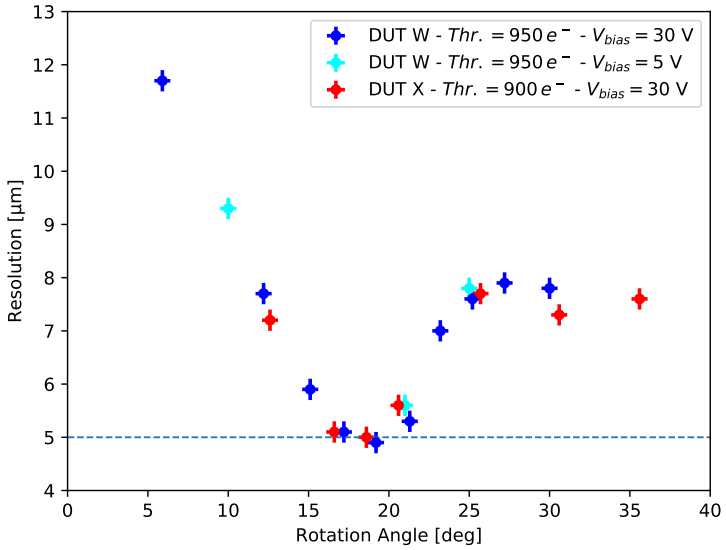


Figure 5.37: DUTs W and X resolution versus the rotation angle around the y axis with $V_{bias} = 30$ V for both DUTs (for DUT W measurements with $V_{bias} = 5$ V are also reported). The horizontal line marks 5 µm resolution.

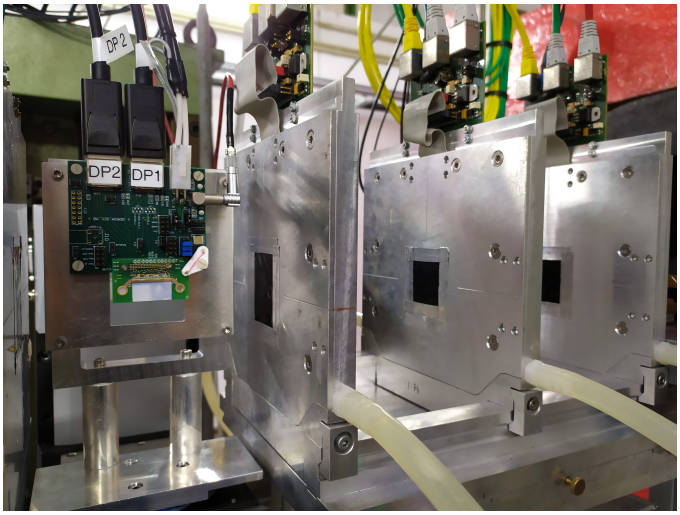


Figure 5.38: Test beam setup for high angle measurements. The beam is coming from the right.

that no significant performance gap along a 3D modules chain is expected, at least with rotation angle greater than 10°. However, it should be noted that $V_{bias} = 5\,V$ is an extreme case, since 3D modules could be operated between 20 V and 40 V, as mentioned before.
Finally, DUT W was tested with rotation angles of 83° and 90° around the y axis, which means that the beam entered sideways the silicon sensor. The setup is shown in Figure 5.38. In the future CMS Inner Tracker, in the innermost barrel layer will be at only 30 mm from the beam line. Therefore, at the extremities of the barrel particles from the collision point can enter the tracker at very high angles, up to 84°. Since a particle entering a pixel sensor at these high angles can traverse many pixels before exiting, one or more of the involved pixels may be under-threshold: if many pixels are under under-threshold, the cluster reconstruction could become problematic. Indeed, the traversing particle can only interact with 50 µm of silicon per pixel (or 100 µm in the case of $25 \times 100\ \mu m^2$ sensors), instead of the $130 - 150\,\mu m$ in the case with orthogonal incidence.
This issues is different between planar and 3D sensors, since the respective depletion regions extends in perpendicular directions. Moreover, the issue is more pronounced with the $50 \times 50\ \mu m^2$ sensors, due to lower traversed silicon. Figure 5.39 shows the DUT W activated pixels for an event, with a DUT rotation angle of 83°: for each of the activated pixels the corresponding TOT is indicated with the colour scale. In the reported event, 33 pixels in row activated, without any missing pixel. Since the alignment software was

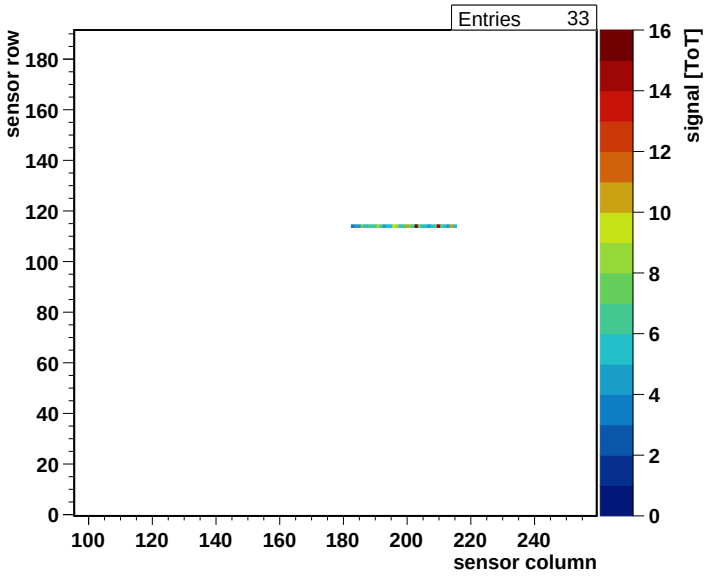


Figure 5.39: Reconstructed cluster on DUT W, with a DUT rotation angle of 83° around the y axis. For each of the activated pixels, the corresponding TOT is indicated with the colour scale.

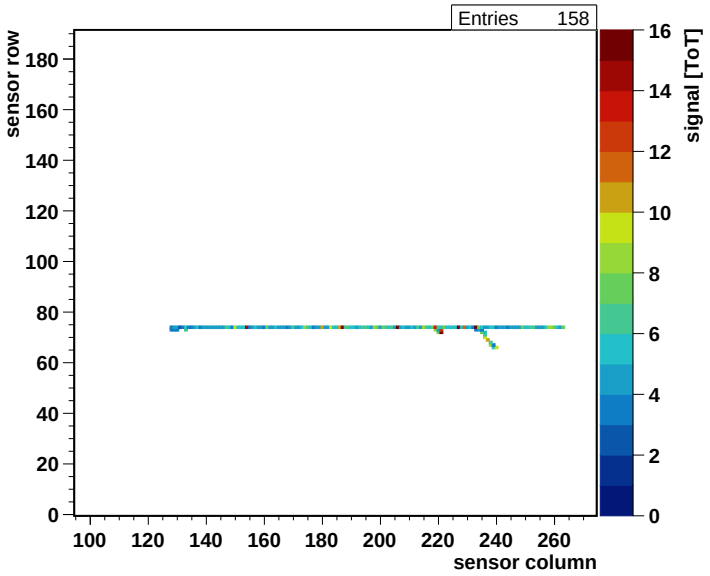


Figure 5.40: Reconstructed cluster on DUT W, with a DUT rotation angle of 90° around the y axis. For each of the activated pixels, the corresponding TOT is indicated with the colour scale.

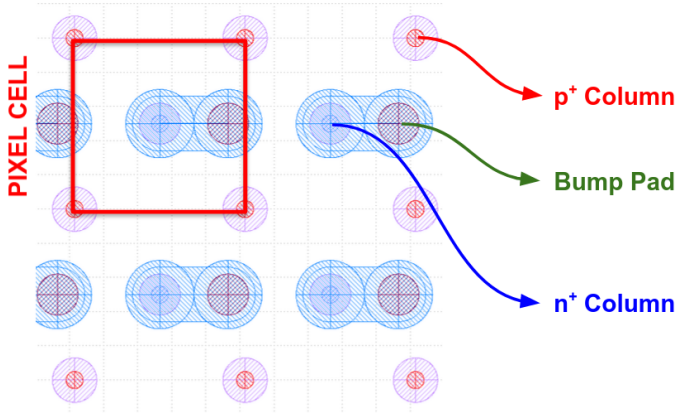


Figure 5.41: Schematic drawing of a 2 × 2 pixel grid of DUT J. The 50 × 50 µm^2 pixel cell is highlighted in red.

not optimised for these kind of studies, more systematic measurements could not be performed. In any case, no signs of cluster breaking were found in the examined events. Figure 5.40 shows the DUT W activated pixels for an event, with a DUT rotation angle of 90°: in this case the cluster is much longer (it traverses the whole Linear front-end) and a delta ray can be observed exiting from the particle trajectory. No signs of cluster breaking have been found in this extreme condition.

5.3.2 Irradiated 50 µm Pitch Detectors

One irradiated 50 × 50 µm^2 3D pixel detector, referred to as DUT J, was tested on beam. The DUT was irradiated at the CERN PS IRRAD facility in a high intensity 24 GeV proton beam (which has size of about 12×12 mm^2), to an estimated fluence of 1×10^{16} n_{eq}/cm^2. The DUT was tilted with respect to the beam at an angle of 55° in order to achieve a more uniform irradiation. In Figure 5.41 a schematic drawing of a 2 × 2 pixel grid of DUT J is shown. The DUT was tuned inside the cooling box, to an average pixel threshold of 1150 electrons, with a dispersion of about 50 electrons. Also in this case, the tuning was made targeting low thresholds and noise, having at most 1% noisy pixel channels. The leakage current was 100 µA at a bias voltage of 110 V.
In Figure 5.42, the global hit detection efficiency versus the applied bias voltage is shown, for perpendicular tracks. The hit efficiency starts saturating around 110 V, and reaches a maximum value of 98.8% at 146 V. A DUT

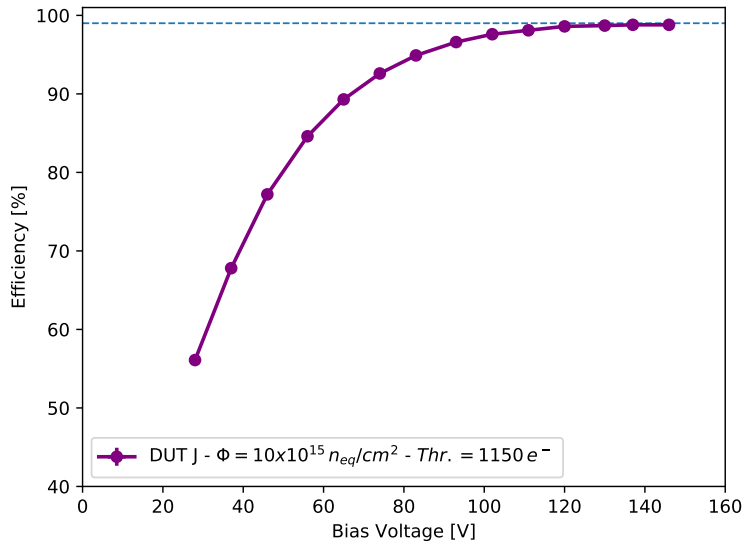
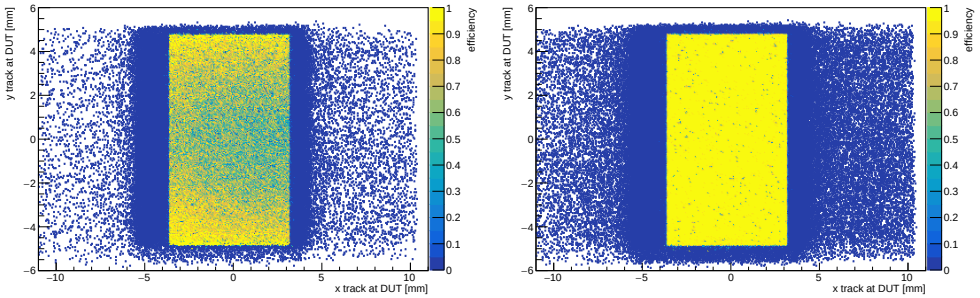


Figure 5.42: DUT J hit detection efficiency versus the applied bias voltage with orthogonal beam incidence. The dashed line represents 99% efficiency.

rotation of 6° around the y axis is enough to reach a hit efficiency greater than 99%.
This DUT was irradiated at the CERN PS: due to the limited beam dimensions, the irradiation was not uniform across the DUT. In Figure 5.43 the hit efficiency map for the entire tested area[10] of the DUT is shown, for two different applied bias voltages: at 37 V the sensor is under-depleted and in the center of the tested area a large efficiency drop can be seen, corresponding to the most irradiated zone. With a bias voltage of 146 V the sensor is over-depleted, and the efficiency is uniform across the tested area, therefore the inefficiency in correspondence to the most irradiated area appears to be recovered. In order to verify this assumption, the tested area was divided into six smaller zones, as shown in Figure 5.44, and the hit efficiency was evaluated on each of these smaller subsets of pixels. Figure 5.45 shows the hit detection efficiency evaluated on all these sub-zones as a function of the bias voltage. At low bias voltages some of the sub-zones present a lower hit efficiency due to the higher irradiation. However, by increasing the bias voltage, the hit efficiencies of all the sub-zones reach comparable values. Therefore the hit detection efficiency is uniform across the DUT tested area,

[10]The DUT tested area corresponds to the Linear front-end, since the Synchronous and Differential front-ends were not tested on beam.



(a) $V_{bias} = 37V$ (b) $V_{bias} = 146V$

Figure 5.43: DUT J hit detection efficiency maps for the whole Linear AFE, for different values of the applied bias voltage. The blue halo corresponds to the tracks reconstructed outside the active area of the DUT: the Linear AFE area is smaller than the telescope and trigger acceptances.

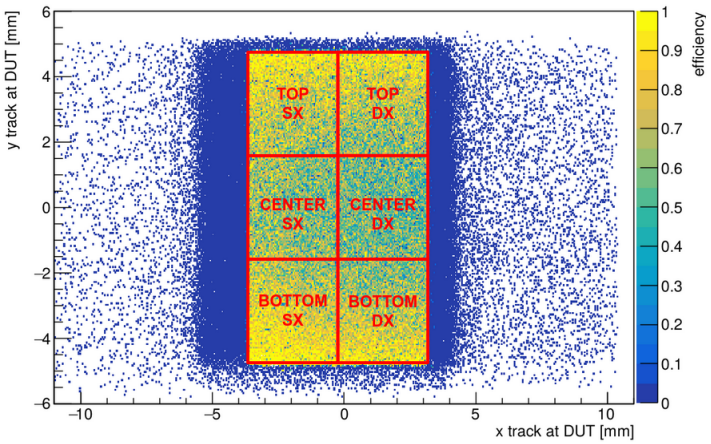


Figure 5.44: DUT J hit detection efficiency maps for the whole tested area with $V_{bias} = 56\,V$. The tested area has been divided into the indicated subzones. The blue halo corresponds to the tracks reconstructed outside the active area of the DUT: the Linear AFE area is smaller than the telescope and trigger acceptances.

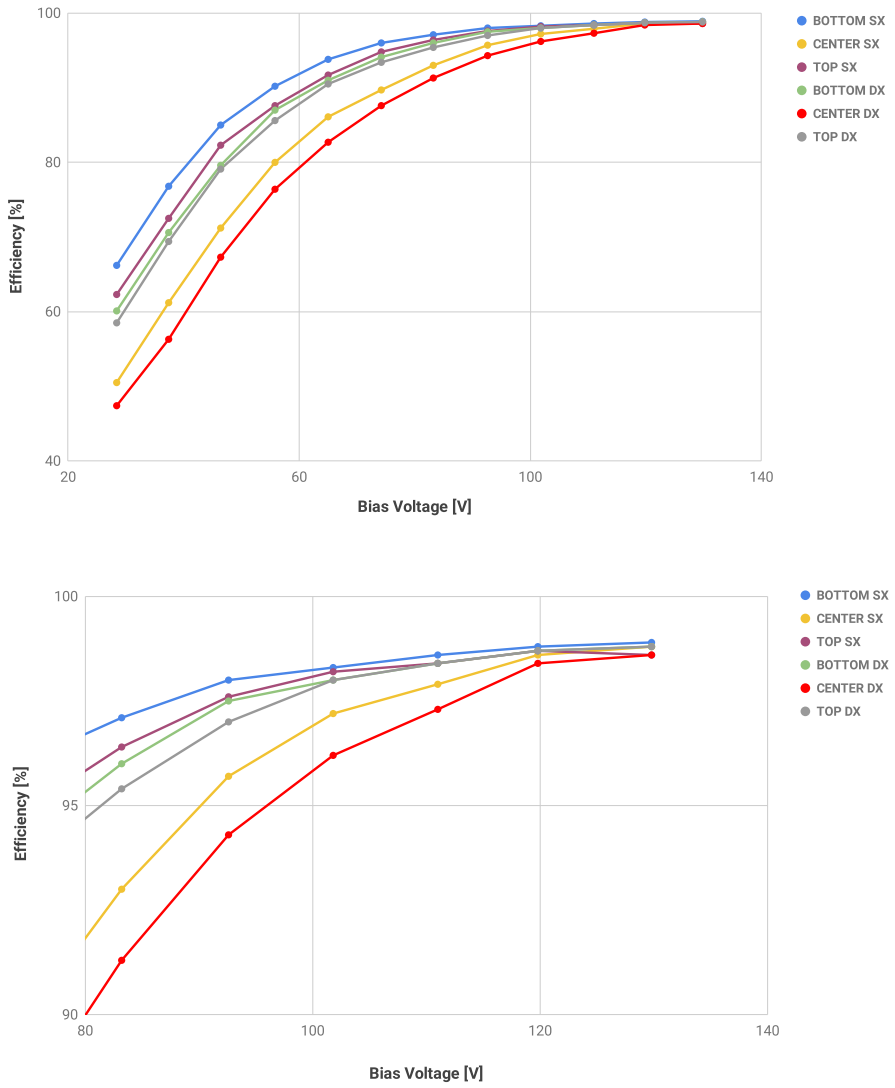
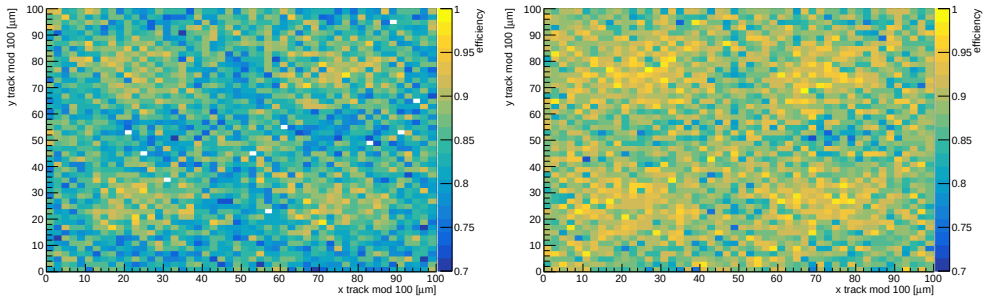


Figure 5.45: DUT J hit detection efficiency versus the applied bias voltage, for each of the sub-zones in which the tested area has been divided into and with orthogonal beam incidence.



(a) $V_{bias} = 56V$ (b) $V_{bias} = 65V$

Figure 5.46: DUT J hit detection efficiency maps for a 2×2 pixel grid, for different values of the applied bias voltage. The efficiency colour scales are the same, and start at 70% efficiency.

at least at high bias voltages.
In Figure 5.46 the DUT J hit efficiency maps for a 2×2 pixel grid are shown, for perpendicular tracks and two different applied bias voltages. At low bias voltages the efficiency is higher for hits near the n^+ columns, which are the collecting electrodes, because the drift path of the charge carriers is shorter, and they have a lower probability to get trapped in the silicon defects produced by irradiation. The efficiency becomes more uniform across the pixels by increasing the applied bias voltage. Since the DUT was tested inside the cooling box, the downstream arm of the telescope could not be used. Therefore, the telescope resolution was worse with respect to fresh DUTs, and inefficiencies due to the columns passive material are smeared in these efficiency maps.
Figure 5.47 shows the DUT J average cluster size map for a 2×2 pixel grid and perpendicular tracks, with $V_{bias} = 146\,V$. As expected, the cluster size is higher at the intersection of adjacent pixels. Figure 5.48 shows the DUT J average cluster size versus the bias voltage, with orthogonal beam incidence. Contrary with what happens with fresh, not irradiated detectors, the cluster size increases with the bias voltage. Indeed, increasing the bias voltage increases the electric field hence reducing the effects of charge trapping.
In order to evaluate the resolution on DUT J, measurements were taken with DUT rotations around the y axis. Figure 5.49 shows the average cluster size as a function of the rotation angle. As expected, the cluster size increases with the angle. The DUT J resolution as a function of the rotation angle is reported in Figure 5.50. The resolution was evaluated as explained in Section 4.3, using a simulation to determine the telescope resolution. The measured

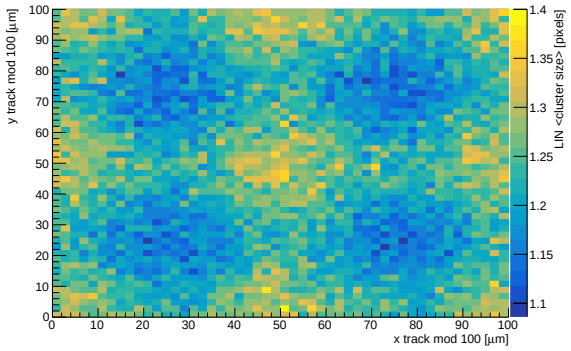


Figure 5.47: DUT J cluster size for a 2×2 pixel grid, with $V_{bias} = 146\,V$ and orthogonal beam incidence.

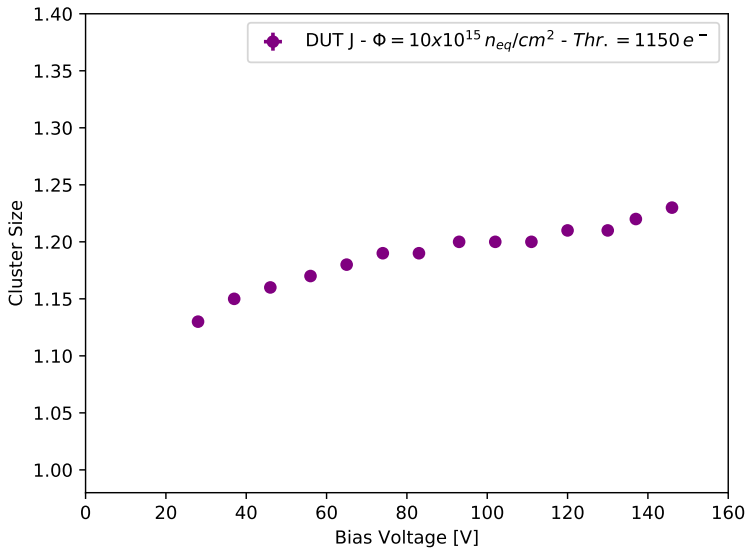


Figure 5.48: DUT J average cluster size versus the applied bias voltage, with orthogonal beam incidence.

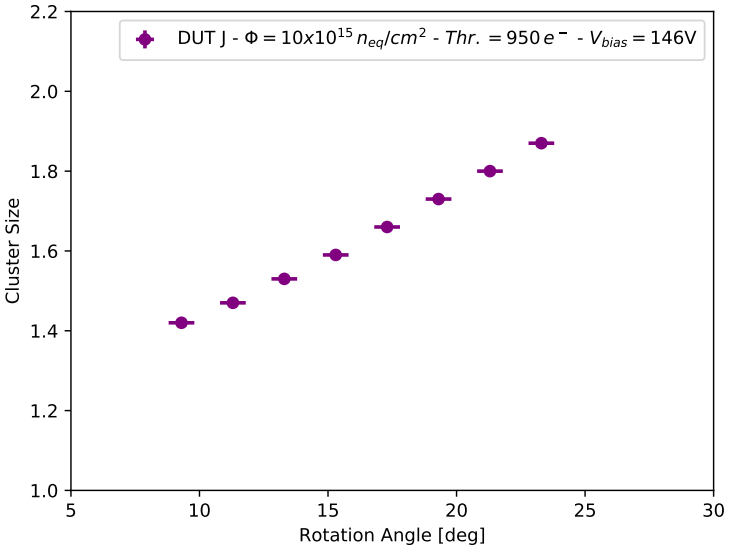


Figure 5.49: DUT J average cluster size versus the rotation angle around the y axis, with $V_{bias} = 146\,V$.

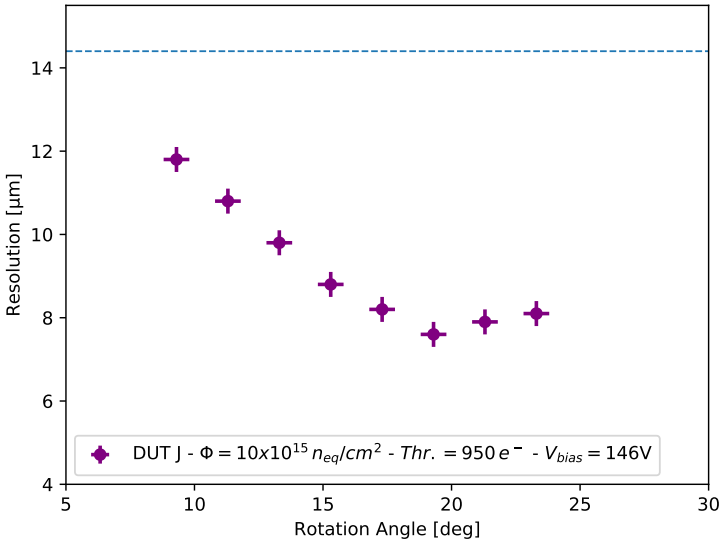


Figure 5.50: DUT J resolution versus the rotation angle around the y axis, with $V_{bias} = 146\,V$. The horizontal line marks the digital 14.4 µm resolution.

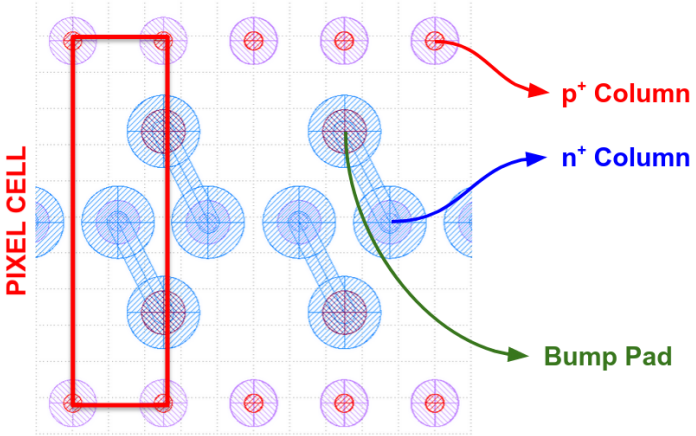


Figure 5.51: Schematic drawing of a 2×2 pixel grid of DUTs K and Y. The 25×100 µm^2 pixel cell is highlighted in red.

resolution is higher with respect to fresh $50 \times 50\,\mu m^2$ 3D pixel detectors, but is significantly lower than the digital resolution (14.4 µm). Moreover, a minimum resolution of 7.6 µm is reached at an angle of about 20°. As the sensor has an active thickness of 130 µm, the expected optimal angle from Equation 5.1 is 21°, compatible with the measurement within errors.
The only irradiated 3D pixel detector showed remarkable efficiency and resolution. DUT E (planar pixel detector) was irradiated at a similar fluence, but had to be biased at more than 500 V before reaching $\sim 99\%$ efficiency, while only 110 V are required for DUT J.

5.3.3 Fresh 25 µm Pitch Detectors

Two fresh 25×100 µm^2 3D pixel detectors, referred to as DUT K and DUT Y were tested on beam. DUT K is from the 3D-1 batch and has an active thickness of 130 µm, while DUT Y is from the 3D-2 batch and has an active thickness of 150 µm.
The thresholds were tuned to an average pixel threshold of 900 electrons for DUT K and 800 electrons for DUT Y, with a dispersion of about 50 electrons in both cases. The leakage current was about 0.2 µA at a bias voltage of 30 V for DUT W and 16 µA at a bias voltage of 30 V for DUT X.
The hit detection efficiency is greater than 99% at a bias voltage of 5 V for both DUTs, with orthogonal beam incidence. Figure 5.52 shows the DUT K hit detection efficiency map for a 4×1 pixel grid and perpendicular tracks. As seen seen with fresh 50×50 µm^2 3D pixel detectors, efficiency drops can

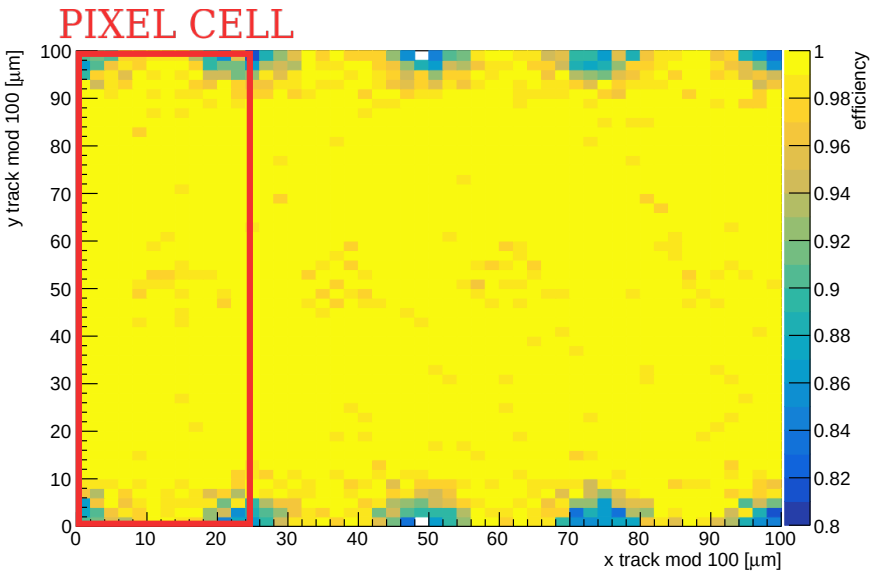


Figure 5.52: DUT K hit detection efficiency map for a 2×2 pixel grid, with $V_{bias} = 30\,V$ and orthogonal beam incidence.

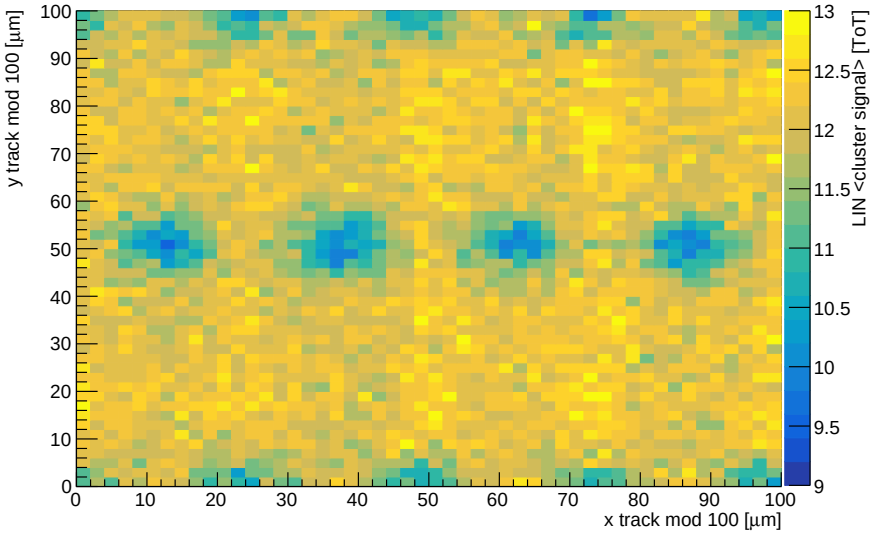
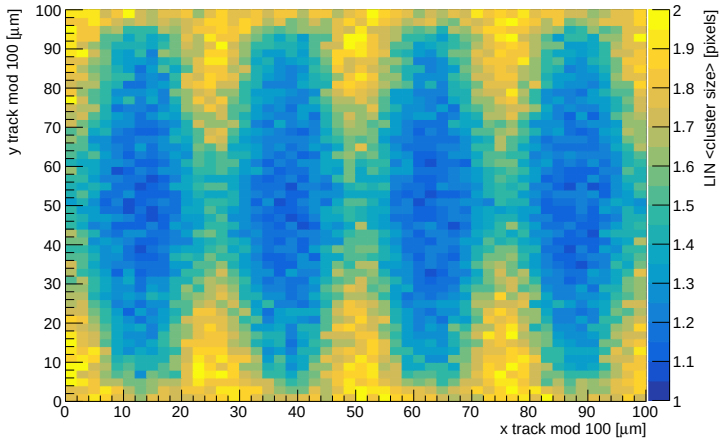
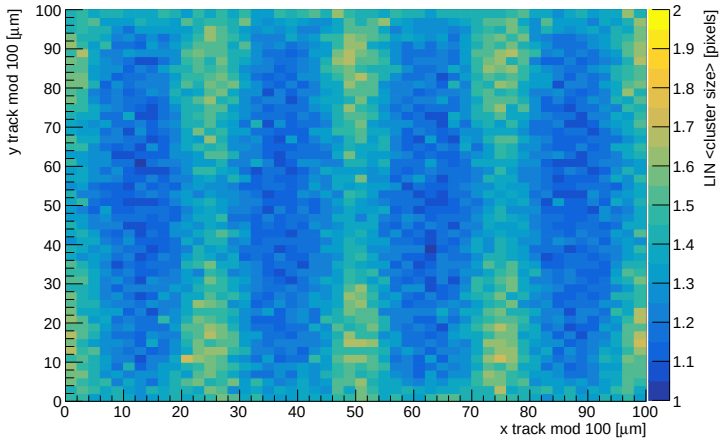


Figure 5.53: DUT K TOT map for a 2×2 pixel grid, with $V_{bias} = 30\,V$ and orthogonal beam incidence.



(a) $V_{bias} = 5\,V$



(b) $V_{bias} = 30\,V$

Figure 5.54: DUT K average cluster size maps for a 2×2 pixel grid, for increasing values of the applied bias voltage and with orthogonal beam incidence.

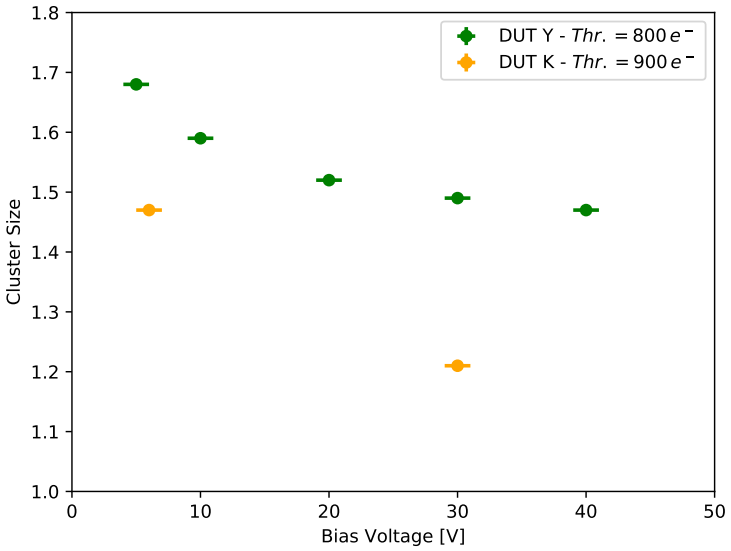
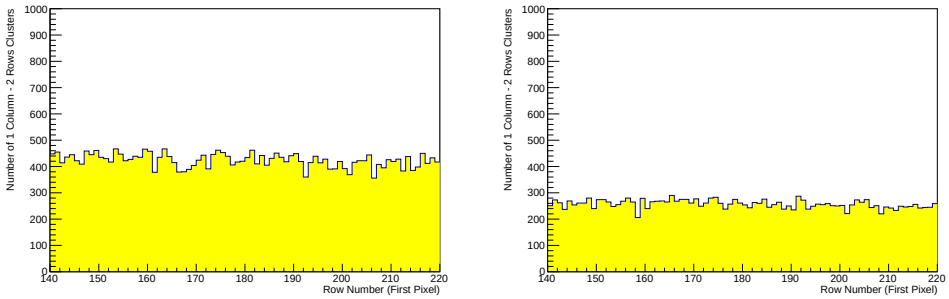


Figure 5.55: DUTs K and Y average cluster size versus the applied bias voltage, with orthogonal beam incidence.

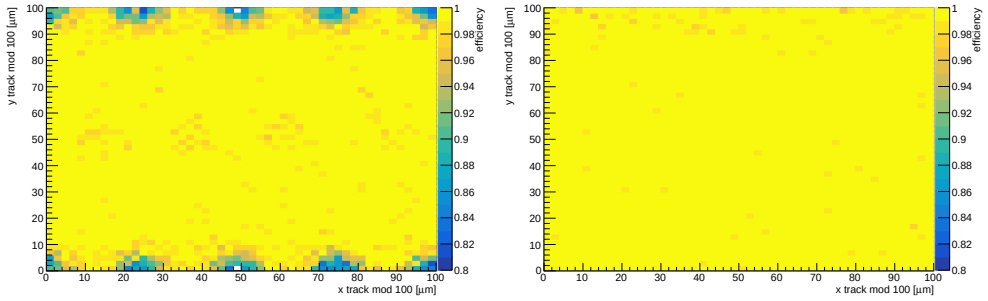
be observed in correspondence to the p^+ columns, as they are made by passive material. The efficiency near the n^+ columns is near 100% since they are collecting electrodes. Figure 5.53 shows the DUT K average TOT map for a 4×1 pixel grid and perpendicular tracks. Again, the behaviour is the same as observed before: hits in correspondence to n^+ and p^+ columns are associated with a lower TOT with respect to the rest of the cell.
Figure 5.54 shows the DUT K average cluster size map for a 4×1 pixel grid for increasing values of the bias voltage and perpendicular tracks. As expected, the cluster size is higher at the intersection of pixels, while it is lower towards the center of the pixel cells: this is because the n^+ column is placed in the center of the cell, therefore the electric field is stronger in this region, hence reducing the charge sharing by diffusion. By increasing the bias voltage, the cluster size decreases as the higher electric field further suppresses the charge sharing by diffusion.
The average cluster size as a function of the bias voltage for DUTs K and Y is reported in Figure 5.55, with orthogonal beam incidence. As expected, the cluster size decreases with the bias voltage due to reduced charge sharing by diffusion. The large discrepancy between the two DUTs is given by the different average thresholds and (in minor part) by the different active thicknesses. As DUT K was tuned to a higher threshold, the cluster size is



(a) $V_{bias} = 5\,V$ (b) $V_{bias} = 30\,V$

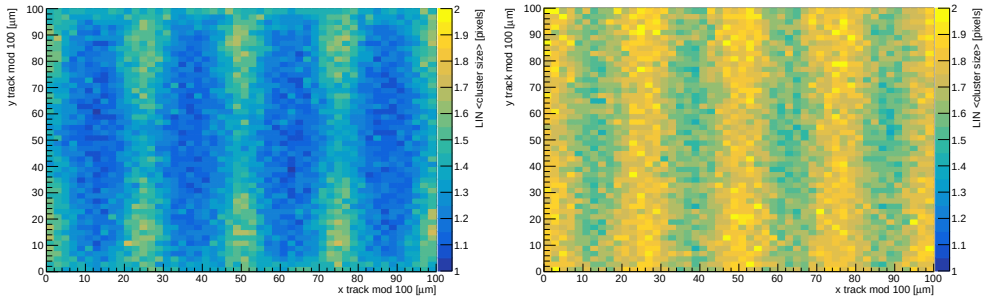
Figure 5.56: DUT K row number of the first pixel in a cluster of one column and two rows, for different values of the applied bias voltage and with orthogonal beam incidence.

systematically lower with respect to DUT Y. Since the average thresholds are very low, small variations can cause significant differences in performance. It should be noted that the cluster sizes for both DUTs are significantly lower with respect to 25×100 µm^2 planar pixel detectors, shown in Section 5.2. This is because the cross-talk is negligible in 3D pixel detectors, as shown in Section 3.3. Indeed, Figure 5.56 shows the row number of the first pixel in a cluster of one column and two rows for DUTs J and K. No asymmetries can be observed between even and odd rows, confirming that cross-talk is negligible with 3D pixel detectors.
By rotating the DUT, the inefficiencies in correspondence to the p^+ columns disappear: in Figure 5.57 the DUT K hit detection efficiency map for a 4×1 pixel grid is reported, with orthogonal beam incidence and with a DUT rotation of 12° around the y axis. Moreover, Figure 5.58 shows the DUT K cluster size map for a 4×1 pixel grid, again with orthogonal beam incidence and with a DUT rotation of 12°. The cluster size is increased in one direction since beam particles are likely to traverse two (or more) pixels.
In order to evaluate the resolution of the detectors, measurements were taken with different DUT rotation angles around to the y axis. Figure 5.59 shows the average cluster size as a function of the rotation angle, for DUTs K an Y, both tested with $V_{bias} = 30\,V$. The cluster size increases with the rotation angle as expected. Again, the differences between the two DUTs are due to different thresholds and different active thicknesses.
Figure 5.60 shows the measured resolution versus the DUT rotation angle, for DUTs K and Y, both tested with $V_{bias} = 30\,V$. DUT Y reaches a resolution of 2 µm at an angle of about 8°, consistent with the optimal angle from



(a) $\alpha = 0^{\circ}$ (b) $\alpha = 12^{\circ}$

Figure 5.57: DUT K hit detection efficiency maps for a 2×2 pixel grid, for different rotation angles around the y axis (along the 25 µm pitch), with $V_{bias} = 30\,V$. The efficiency colour scales are the same, and start at 80% efficiency.



(a) $\alpha = 0^{\circ}$ (b) $\alpha = 12^{\circ}$

Figure 5.58: DUT K average cluster size maps for a 2×2 pixel grid for different rotation angles around the y axis (along the 25 µm pitch), with $V_{bias} = 30\,V$.

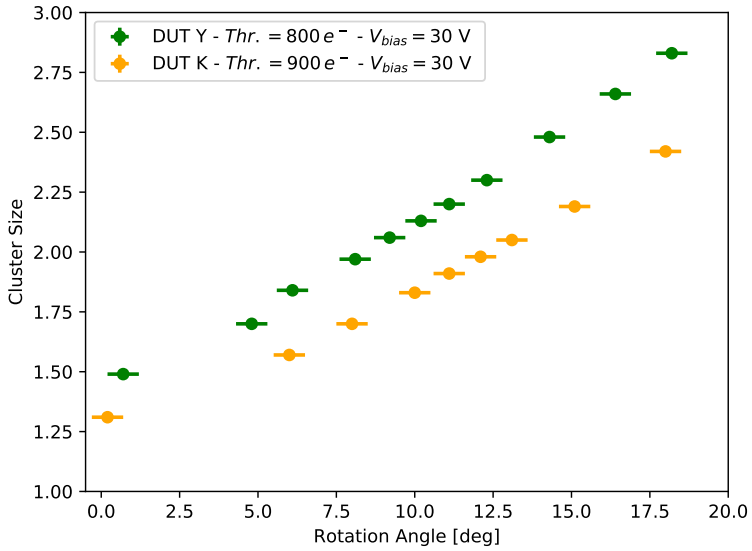


Figure 5.59: Average cluster size versus the rotation angle around the y axis (along the 25 µm pitch) for DUTs K and Y, with $V_{bias} = 30\,V$ for both DUTs.

Equation 5.1. DUT K reaches a resolution of 3 µm at angle of about 11°, again consistent with the optimal angle from Equation 5.1 (DUT K has a lower active thickness with respect to DUT Y, and therefore a higher optimal angle). The higher average threshold of DUT K degrades the resolution by about 1 µm with respect to DUT Y.
The resolution was also evaluated as a function of the bias voltage with orthogonal beam incidence, as shown in Figure 5.61. This was done in order to further test the viability of a serial powering chain of 3D modules. The resolution is better at lower bias voltage, due to increased charge sharing by diffusion. However, the resolution saturates at about 20 V. Therefore, if the 3D modules in serial powering chain are operated between 20 V and 40 V of bias voltage (as mentioned before), no significant performance gap in the chain is expected. This is the extreme case of orthogonally incident particles: with small DUT rotations the effect is further reduced, as shown in Figure 5.37 with the 50×50 µm^2 3D pixel detectors.

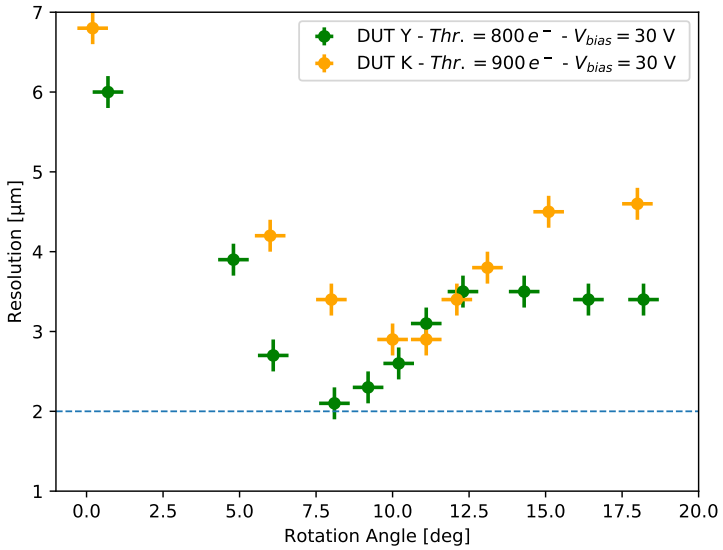


Figure 5.60: DUTs K and Y resolution versus the rotation angle around the y axis (along the 25 µm pitch) with $V_{bias} = 30$ V for both modules. The horizontal line marks 2 µm resolution.

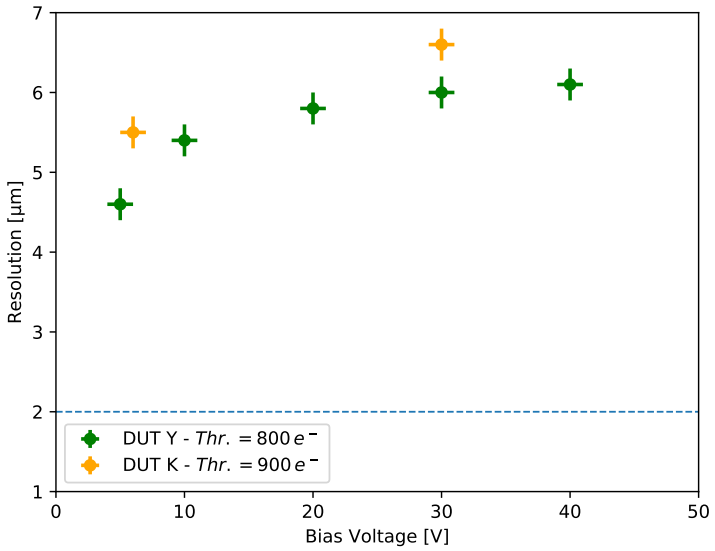


Figure 5.61: DUTs K and Y resolution versus the applied bias voltage, with orthogonal beam incidence. The horizontal line marks 2 µm resolution.

5.4 Outlook

The results presented in this Chapter will contribute to the choice of the pixel sensor technology for the future CMS Inner Tracker. Planar pixel detectors have shown remarkable efficiencies and resolutions when irradiated to fluences up to $24 \times 10^{15}\,\mathrm{n_{eq}cm^{-2}}$. However, they require very high bias voltages (up to 600 V) in order to become fully efficient. The $25 \times 100\,\mu\mathrm{m}^2$ pixel cell design was found to be susceptible to cross-talk effects when non-irradiated. As shown in Section 3.3, the bitten implant design moderately reduces this effect.
These issues could be solved using 3D pixel detectors. In particular, far lower bias voltages are necessary in order to reach good performances: this would ease the power consumption requirements of the future CMS Inner Tracker. Moreover, as presented in Section 3.3, 3D pixel detectors are less affected by cross-talk effects.
During my PhD, only one irradiated (to a fluence of $10 \times 10^{15}\,\mathrm{n_{eq}cm^{-2}}$) 3D pixel detector could be tested on beam, and it showed excellent performances. Further test beams are envisioned in the near future, with 3D detectors irradiated to fluences greater than $20 \times 10^{15}\,\mathrm{n_{eq}cm^{-2}}$. Afterwards, a comprehensive comparison will be performed in order to optimise the usage of 3D and planar pixel detectors in the future CMS Inner Tracker.
Resolution studies presented in this Thesis will also be crucial in order to optimise the usage of the different pixel cells ($50 \times 50\ \mu\mathrm{m}^2$ and $25 \times 100\ \mu\mathrm{m}^2$) in different regions of the future tracker.

Chapter 6

The Higgs Boson at the LHC

6.1 Electroweak Interactions

The Standard Model (SM) describes all fundamental constituents of matter and their interactions. The SM is a renormalisable quantum field theory based on the local gauge symmetry $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ and it is capable to provide a quantitative description of three of the four interactions in nature: electromagnetism, weak and strong interactions. The field content of the SM consists of three generations of leptons, three generations of quarks (all of them are fermions, with spin 1/2), and the gauge bosons.
The $SU(2)_L \otimes U(1)_Y$ gauge group describes the electroweak interaction. The generator of the $SU(2)_L$ group is the weak isospin $\boldsymbol{T}_L$, while the generator of $U(1)_Y$ is the weak hypercharge Y. They are connected to the charge Q by the following equation:

$$Q = T_{3L} + \frac{Y}{2} \tag{6.1}$$

where T_{3L} is the component of $\boldsymbol{T}_L$ along the third axis, arbitrarily chosen as the quantization axis. Leptons are described by Dirac fields whose left-chiral components form $SU(2)_L \otimes U(1)_Y$ doublets:

$$\begin{pmatrix} \nu \\ \ell \end{pmatrix}_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L , \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L , \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \tag{6.2}$$

The T_{3L} component of weak isospin is equal to $+1/2$ for up components and $-1/2$ for down components of the doublets, while $Y = -1$ for all.
Only the down components have right-chiral equivalents, which behave as $SU(2)_L \otimes U(1)_Y$ singlets with $Y = -2$:

$$\ell_R = e_R \, , \, \mu_R \, , \, \tau_R \tag{6.3}$$

Together, these fields describe the electron, the muon, the tau, and their respective neutrinos. Since their $Q = 0$, neutrinos are only sensitive to weak interactions, making them the least interactive particles in the SM.
Similarly, quarks are described by Dirac fields organised in $SU(2)_L \otimes U(1)_Y$ doublets:

$$\begin{pmatrix} q_u \\ q_d \end{pmatrix}_L = \begin{pmatrix} u \\ d \end{pmatrix}_L , \begin{pmatrix} c \\ s \end{pmatrix}_L , \begin{pmatrix} t \\ b \end{pmatrix}_L \tag{6.4}$$

where T_{3L} is equal to $+1/2$ for up components and $-1/2$ for down components of the doublets, while $Y = 1/3$ for all. Contrary to leptons, both up and down components have right-chiral equivalents behaving as $SU(2)_L \otimes U(1)_Y$ singlets:

$$q_{u,R} = u_R \, , \, c_R \, , \, t_R \tag{6.5}$$

with $Y = 4/3$ and $Q = 2/3$, and:

$$q_{d,R} = d_R \, , \, s_R \, , \, b_R \tag{6.6}$$

with $Y = -2/3$ and $Q = -1/3$.
Finally, for every particle q an antiparticle $\bar{q}$ exists, obtained by applying the charge conjugation transformation to the field. Such particles have all of the same properties as regular particles, save for changing the sign of all additive quantum numbers.
Interactions between fermions occur through the exchange of spin 1 bosons: $SU(2)_L \otimes U(1)_Y$ has four gauge fields, combinations of which give W^+, W^- and Z bosons, the mediators of the weak interaction, and the photon γ, mediator of the electromagnetic force.
Weak charged currents, mediated by vector bosons $W^\pm$, only involve left-handed particles or right-handed anti-particles. This is in accordance with the fact that parity symmetry is maximally violated for weak charged current interactions. In the case of leptons, charged currents can only connect two particles within the same generation, for example the electron and the electron neutrino, while a mixing of different generations may occur in the case of quarks, according to the Cabibbo-Kobayashi-Maskawa matrix (CKM).
Another possible interaction in the weak sector is the neutral current interaction, mediated by the neutral Z boson. In the vertex of this interaction the identity of the interacting leptons does not change, resembling in this matter the electromagnetic current. Concerning the quark sector, the weak neutral currents involving different quark flavours, i.e. flavour changing neutral currents, are strictly suppressed by the Glashow-Iliopoulos-Maiani (GIM) mechanism.
The $SU(2)_L$ local gauge symmetry requires the gauge bosons to be massless, which is unproblematic for photons and gluons, but in drastic contrast to the

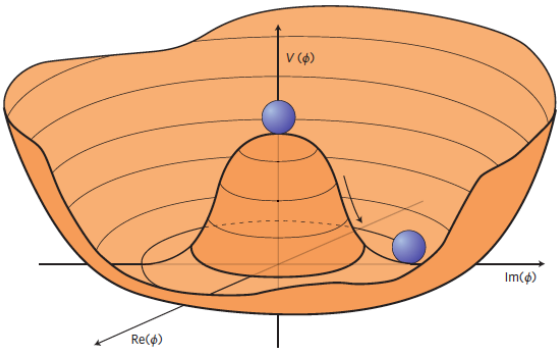


Figure 6.1: $V(\phi)$ potential for $\lambda > 0$ and $\mu^2 < 0$.

known masses of the Z and $W^{\pm}$ bosons, which are $m_Z = 91.1876 \pm 0.0021$ GeV and $m_W = 80.385 \pm 0.015$ GeV, respectively. Also the fermion fields should be massless for gauge invariance, while it is experimentally established that all fermion fields carry mass (except possibly for one neutrino[1]).

6.1.1 The Brout-Englert-Higgs Mechanism

The proposed solution to the mass problem is the mechanism of spontaneous symmetry breaking, also called the Brout-Englert-Higgs mechanism. The gauge symmetry is still intrinsic to the Lagrangian density of the theory, but not manifest in its energy ground state, which in this case is the quantum vacuum. The spontaneous symmetry breaking requires the introduction of a self-interacting complex scalar field, which is an isospin doublet:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} (\phi_1 + i\phi_2)/\sqrt{2} \\ (\phi_3 + i\phi_4)/\sqrt{2} \end{pmatrix} \tag{6.7}$$

The most simple lagrangian density involving this term is given by:

$$\mathcal{L}_H = D^\mu \phi^\dagger D_\mu \phi - V(\phi) \tag{6.8}$$

where:

$$D^\mu = \partial^\mu + ig\boldsymbol{A}_\mu \cdot \boldsymbol{T}_L^\phi + ig' B_\mu Y^\phi \tag{6.9}$$

is the covariant derivative. $\boldsymbol{A}_\mu$ and B_μ are the gauge fields, $\boldsymbol{T}_L^\phi$ and Y^ϕ are the $SU(2)_L \otimes U(1)_Y$ generators for a doublet of complex scalar fields and g, g' represent the coupling constants for the gauge fields. The potential term:

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2 \tag{6.10}$$

[1]Neutrinos in the SM are massless. However, due to observed flavour oscillations, the masses of the three neutrinos need to be different, and only one is allowed to be zero.

depends on two parameters, μ and λ, with $\lambda > 0$ in order to have vacuum stability. If μ is chosen so that $\mu^2 < 0$, the minimum of the potential field may assume infinite values, since $V(\phi)$ is rotationally invariant in the ϕ space:

$$\phi^\dagger\phi = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2} \tag{6.11}$$

where v is the Vacuum Expectation Value (VEV). Figure 6.1 shows the potential "mexican hat" shape. In the SM, the lagrangian density is expanded around the potential minimum, which needs to be chosen among the infinite set of Equation 6.11. When the choice is performed, the symmetry is broken.
However, since the photon has null mass, the Brout-Englert-Higgs mechanism has to preserve the electromagnetic symmetry, that is the lagrangian density needs to be invariant under $U(1)_Q$ transformations. Therefore, the ground state has to be electrically neutral and can be written as:

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v\end{pmatrix} \tag{6.12}$$

The field can be expanded at the first order around the ground state:

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v+H\end{pmatrix} \tag{6.13}$$

where H is the field associated to the scalar Higgs boson which remains after the symmetry breaking. Inserting this field in Equation 6.8, the bosonic fields acquire mass:

$$m_W = \frac{v}{2}g \quad , \quad m_Z = \frac{v}{2}\sqrt{g^2+g'^2} \tag{6.14}$$

Furthermore, the mass of the Higgs boson also arises:

$$m_H = v\sqrt{2\lambda} \tag{6.15}$$

whose value can not be predicted by theory since the λ is unknown. On the other hand, v can be obtained from the relation between m_W and the Fermi constant G_F ($v = 247$ GeV).
The mass of fermions is achieved without breaking the gauge symmetry of the Lagrangian by introducing a coupling term, known as Yukawa coupling, between the fermion doublets and the Higgs field. Thus, all the massive fields within the SM interact with the Higgs boson through a term that is found to be proportional to the particle mass itself.

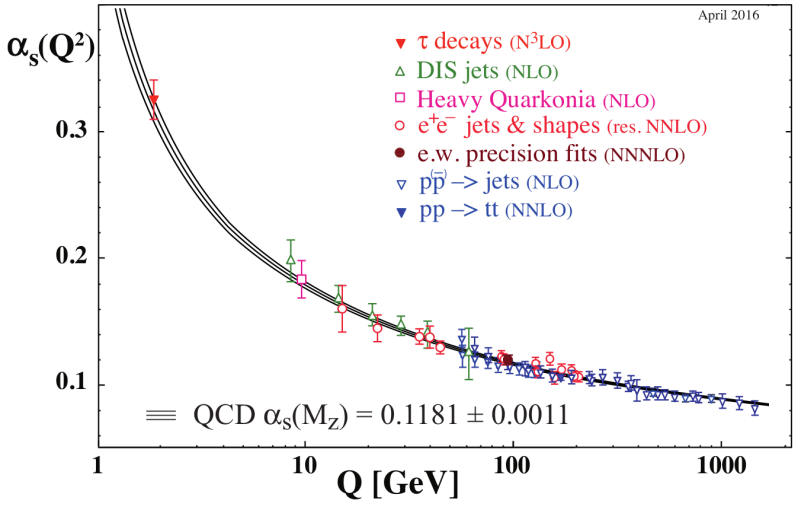


Figure 6.2: Measurements of α_s as a function of the energy scale Q [1].

6.2 Strong Interactions

Quarks also interact via the strong force, responsible of their confinement within hadrons. In fact, free quarks are not observed in nature, but they bind together forming two main categories of hadrons: mesons, bound states of a quark q and an anti-quark $\bar{q}$, and baryons, bound states of three quarks. The strong interaction is described by Quantum ChromoDynamics (QCD), a non-abelian gauge theory based on the $SU(3)_C$ symmetry group. The mediators of the interaction are eight massless gauge bosons, the gluons. The symmetry is assumed to be exact: a physical quantity called colour is preserved. Colour can assume three different values for quarks (blue, green and red) and anti-quarks (anti-blue, anti-green and anti-red).
The physical vertices in QCD include the gluon-quark-antiquark vertex, analogous to the Quantum ElectroDynamics (QED) photon-fermion-antifermion coupling, but also the three-gluon and four-gluon vertices. Indeed, gluon themselves carry colour charge, making it possible for these particles to autointeract: this aspect has no analogue in an abelian theory like QED. Quarks and gluons are the only particles that interact through the strong interaction. The non-abelian nature of the theory leads to two important characteristics:

- **Colour Confinement:** the QCD coupling constant $\alpha_s = g_s^2/4\pi$ is a decreasing function of the transferred momentum Q, as shown in Figure 6.2. At low energies (corresponding to distances of the order of the fb), α_s is large and a perturbative approach is not applicable. Increasing the distance between a quark-antiquark pair, the intensity of the force

increases up to the point that the creation of a new quark-antiquark pair from the vacuum is more energetically favourable than further increasing the interaction strength. This is the reasons why quarks are not observed isolated, but they form colourless bound states, the hardons.

- **Asymptotic Freedom:** at large Q scales (i.e. at small distances) α_s decreases approaching zero, meaning that quarks can be asymptotically considered as free particles. The small value of the coupling constant at large scales justifies the usage of a perturbative approach in this regime, the pQCD (Perturbative QCD).

6.2.1 Jets Measurement

Colour confinement is also the origin of the hadronisation process which causes the formation of jets. Quarks and gluons produced after strong scattering processes recombine with each others, forming collimated hadron clusters, the jets.
The energy of hadrons composing a jet may be down to values for which the pQCD fails. In particular, a fixed order cross section calculation can diverge if a coloured particle is emitted at either low momentum (infrared divergence) or at low angle (collinear divergence). In particle physics experiments is therefore necessary to define observables insensitive to configurations that are difficult to compute theoretically. Any observable Ω, function of N measured momenta, must therefore satisfy two requirements:

- Infrared safety: $\Omega(p_1, ..., p_N) = \Omega(p_1, ..., p_N, \epsilon)$, where ϵ is the arbitrary small energy of another emitted particle.
- Collinear safety: $\Omega(p_1, ..., p_i, ..., p_N) = \Omega(p_1, ..., p_{i1}, p_{i2}, ..., p_N)$ where $p_i = p_{i1} + p_{i2}$ and p_{i1}, p_{i2} are the momenta of particles emitted at low angle.

The CMS experiments uses the $anti{-}k_T$ recursive algorithm to group charged hadron candidates into jets [58]. Given a set of particles with transverse momentum $p_{T,i}$, two kind of distances are calculated:

$$d_{i,B} = p_{T,i}^{-2} \tag{6.16}$$

which is the distance between object i and the beam, and:

$$d_{i,j} = \mathrm{Min}\left(p_{T,i}^{-2}, p_{T,j}^{-2}\right)\left(\frac{\Delta R_{ij}^2}{R^2}\right) \tag{6.17}$$

which is the distance between objects i and j, where ΔR was defined in Equation 2.7. The distance parameter R is arbitrary, and broadly speaking defines the size of the jet in the $\phi - \eta$ plane. Both CMS and ATLAS use 0.4 as a standard value for R. The algorithm proceeds as follows.
First, the distance $d_{i,B}$ is calculated for all particles and $d_{i,j}$ is calculated for all particle pairs. The smallest distance is found: if it is d_{ij}, objects i and j are combined into a single object summing their four-momenta. The distances are then calculated again. The iterative procedure ends when the smallest distance is found to be of type $d_{i,B}$: in this case the object i is defined as jet and is removed from the input particles. The algorithm restarts, until no particles are left.
The $anti - k_T$ algorithm is both infrared safe (if a cut on the minimum jet p_T is applied) and collinear safe: if a soft hadron is added to the initial set of particles, the same jet is found, and the same is true if a particle is replaced by two collinear ones. Moreover, the $anti - k_T$ algorithm starts the clustering procedure from the highest p_T objects, resulting in jets with a smoother shape when compared to other algorithms.

6.2.2 Proton-Proton Collisions

In proton-proton collisions at the LHC the internal structure of the protons is relevant for describing scattering processes. Indeed, protons are composed by three valence quarks (u,u,d) and a sea of virtual quarks and gluons as an effect of their strong interaction. The collision can be described by the scattering of the proton constituents, the partons, namely quarks and gluons. Each parton carries a fraction x of the proton four-momentum:

$$p^{\mu}_{parton} = x p^{\mu}_{proton} \tag{6.18}$$

where x is the Bjorken variable. The actual energy of the elementary scattering at the LHC is therefore a fraction of the nominal 13 TeV.
The x distribution for a given parton is the Probability Density Function (PDF) $f(x)$. The normalisation condition is given by the fact that the sum of the four-momenta of all partons must correspond to the proton four-momentum:

$$\sum_i \int_0^1 x f_i(x) \mathrm{d}x = 1 \tag{6.19}$$

where the index i runs over all partons. Figure 6.3 show the partons PDFs as a function of x: gluons are dominant at low x values. Moreover, the PDFs shapes depend on the energy scale at which they are measured, the factorisation scale μ. If α_s is small, which is true in the high energy regime,

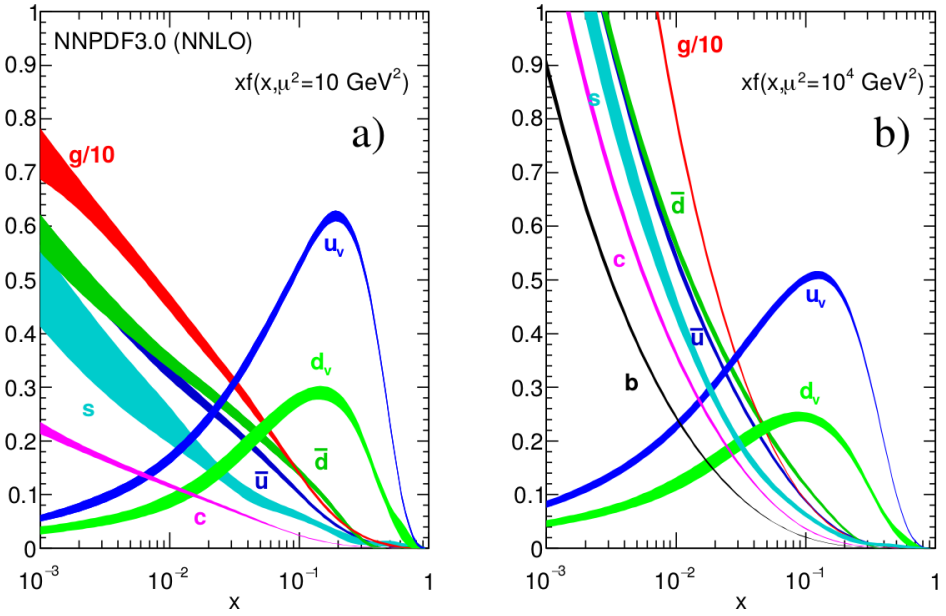
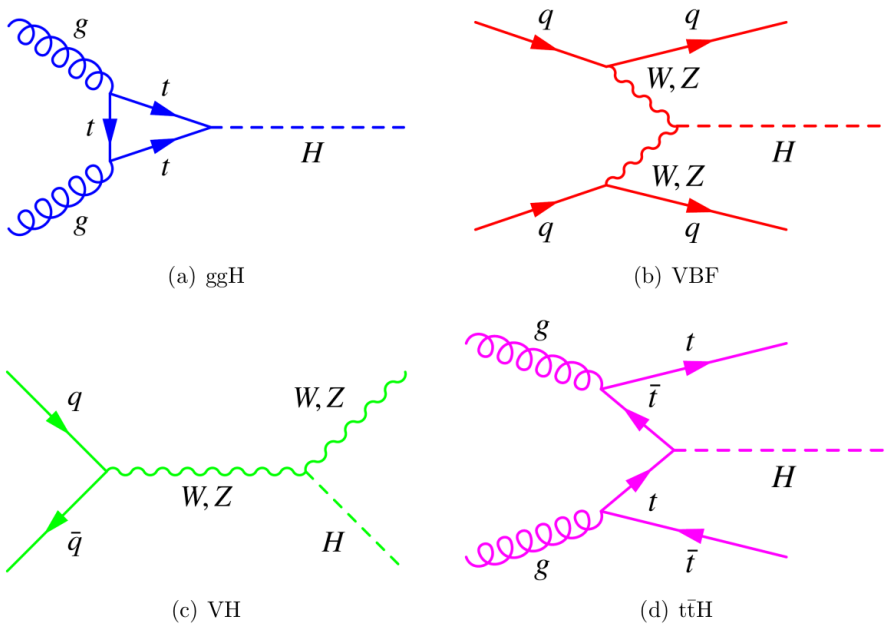


Figure 6.3: Parton Distribution Functions (multiplied by x) evaluated at a factorisation scale $\mu = 100$ GeV (a) and $\mu = 10$ TeV (b), which is the LHC energy scale [1].

a perturbative approach is valid, and the evolution of the PDFs with respect to μ can be inferred by using the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [59].
Within a proton-proton collision, several partons can interact. Only hard scatterings are interesting for particle physics experiments, that is interactions between partons with a large transferred momentum, which allow for the production of a variety of particles. The QCD factorisation theorem allows to factorise these interactions into hard scatterings at parton level, convoluted with the PDFs [60]:

$$\sigma_{p_1+p_2\to F+X} = \sum_{i,j} \int \mathrm{d}x_1 \mathrm{d}x_2 f_i(x_1,\mu^2) f_j(x_2,\mu^2) \hat{\sigma}_{ij\to F}(x_1 x_2 s, \mu_R^2, \mu^2) \quad (6.20)$$

where p_1 and p_2 are the interacting protons, F is the final state of interest, X is any other product of the scattering process, i (j) is a parton type in proton p_1 (p_2), f_i (f_j) is the PDF for parton i (j) and $\hat{\sigma}_{ij\to F}$ is the parton level cross section of the process. The sum runs over all partons i and j that can initiate the process of interest, while the integrals are over the Bjorken variables x_1 and x_2 for protons p_1 and p_2 respectively. Finally, μ_R is the renormalisation scale, an additional parameter introduced in pQCD to



(a) ggH (b) VBF

(c) VH (d) tt̄H

Figure 6.4: Higgs boson main production modes: Gluon Fusion (a), Vector Boson Fusion (b), Vector Boson Associated Production (c), Top Quark Associated Production (d).

treat the ultraviolet divergences. Other partons in the two colliding protons undergo soft scattering processes: these are known as the underlying event.

6.3 The Higgs Boson Phenomenology

A neutral, scalar resonance compatible with the SM Higgs boson H was observed in 2012 by the ATLAS and CMS experiments at the LHC, at a mass of 125.09 ± 0.24 GeV [61]. The Higgs boson couples to every massive particle in the SM, therefore its phenomenology is particularly rich.
The main Higgs production processes at the LHC (with proton-proton collisions at $\sqrt{s} = 13$ TeV) are represented by the Feynman diagrams shown in Figure 6.4. In order of decreasing cross section the main production modes are:

- **Gluon Fusion (*ggH*):** two incoming gluons give rise to the Higgs boson through a heavy quark loop. The quark flowing in the loop is mainly the top quark, since Higgs-couplings with heavy particles are preferred. The ggH production cross section is estimated to be 48.6 pb at $\sqrt{s} = 13$ TeV [62].

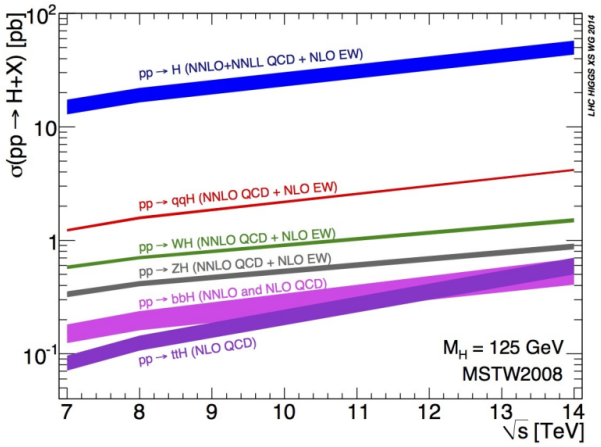


Figure 6.5: Cross sections for the main Higgs production modes in proton-proton collisions as a function of the center of mass energy $\sqrt{s}$.

- **Vector Boson Fusion (*VBF*):** two incoming quarks emit two vector bosons (W^+W^- or ZZ) which further interact to produce the Higgs boson. The *VBF* production cross section is estimated to be 3.8 pb at $\sqrt{s} = 13$ TeV [62], about one order of magnitude lower than *ggH*.

- **Vector Boson Associated Production (*VH*):** also known as Higgsstrahlung, this process is characterized by the emission of a Higgs boson from a W or a Z boson produced by two incoming quarks. Cross sections are estimated to be about 1.4 pb and 0.9 pb for *WH* and *ZH* production modes respectively at $\sqrt{s} = 13$ TeV [62].

- **Top Quark Associated Production (*ttH*):** a pair of top quarks, originated from the splitting of two incoming gluons, interacts to give rise to a Higgs boson. The cross section of this process is estimated to be 0.5 pb at $\sqrt{s} = 13$ TeV [62]. Another analogous production mechanism, with a similar cross section, is the b quark associated production.

The cross sections for the main Higgs production modes in proton-proton collisions are reported in Figure 6.5, as a function of the center of mass energy $\sqrt{s}$. The experimental signature of each Higgs production mechanism is also determined by the Higgs decay channel. Indeed, the predicted Higgs Boson mean lifetime is $\sim 10^{-22}$ s [1], therefore its properties are measured at LHC through the study of its decay products. Figure 6.6 reports the Branching Ratio (BR) of the decay channels. The most probable decay of the Higgs boson is $H \to b\bar{b}$, with a BR higher than 50%. However, analyses

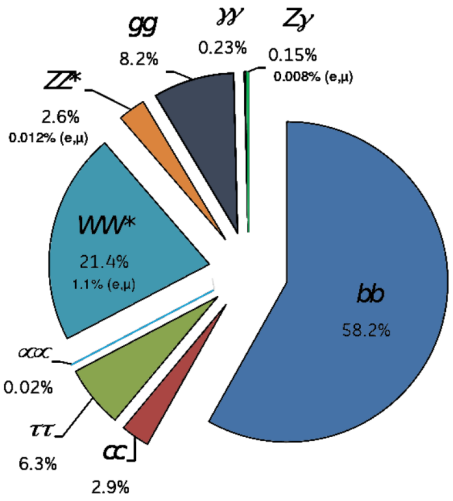


Figure 6.6: Branching ratios for the Higgs bosons decay modes.

looking for this decay are limited by the overwhelming $pp \to b\bar{b}$ background, and usually target the Higgs boson production via VBF, VH or ttH, where additional jets or leptons can be used to select events [63]. The $H \to gg$ decay is problematic due to the high contamination from other hadronic processes. The leptonic decays $H \to e^+e^-$ and $H \to \mu^+\mu^-$ [64] have a low probability to occur whereas the $H \to \tau^+\tau^-$ decay is more difficult to reconstruct due to the subsequent decay of the τ leptons [65].
Bosonic decays are particularly interesting to study, thanks to their large BR and the possibility to tag fully leptonic or semi-leptonic final states. For instance, the $H \to ZZ \to 4\ell$ has a very clean signature but a relatively low cross section: the BR is only $\simeq 0.01\%$ if ℓ is an electron or a muon [66]. The $H \to \gamma\gamma$ decay features a larger BR ($\simeq 0.23\%$), but is contaminated by the QCD backgrounds (both for real photons or misidentified jets) [67]. The channel taken into consideration for this Thesis is the $H \to W^+W^-$ decay, which is the second most probable.

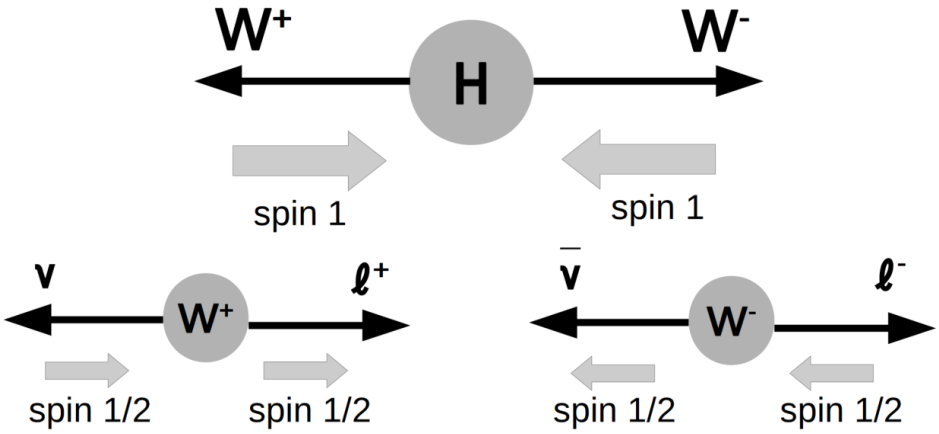


Figure 6.7: Spin correlation of the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay channel.

6.4 The WW Decay Channel

In the $H \to W^+W^-$ decay channel, since $m_H < 2m_W$, one of the two W bosons is produced off-shell[2]. However, since the W mean lifetime is $\sim 10^{-25}$, both bosons are observed through their decay products. Each W boson has a 67.4% probability to decay into hadrons. However, in this case, the final state is strongly contaminated by multijet backgrounds. Instead, the BR for a W boson to decay into leptons is 10.9% per each lepton family. The fully leptonic final state:

$$H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu} \tag{6.21}$$

despite the relatively low BR ($\sim 1\%$), is characterized by a cleaner signature and is affected by much less background contributions.
The signature of this final state is the presence of two leptons (electrons or muons[3]) with opposite charge and a moderate amount of missing transverse energy, due to the presence of two neutrinos in the final state. The two leptons are characterized by high p_T values (with the lepton from the on-shell W boson having a higher p_T value). Because of the presence of neutrinos in the final state, the kinematics is not closed and an invariant mass peak can not be reconstructed in this channel: other methods must be used to

[2]A particle is on-shell if its four-momentum satisfies the relation $p_\mu p^\mu = m^2$. Otherwise, it is off-shell (or virtual).

[3]The τ leptons have a short lifetime ($\sim 10^{-13}$ s), therefore they decay inside the detector. Hadronic decays of the τ (BR $\sim 65\%$) are not considered as signal in $H \to WW$ analyses, while electrons and muons coming from τ decays are included.

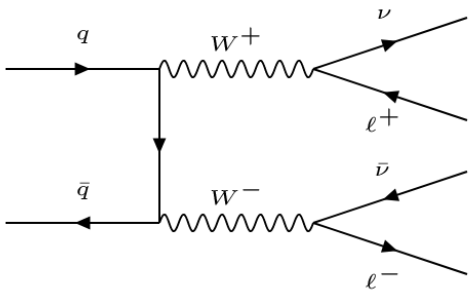


Figure 6.8: Leading order Feynman diagram for the non-resonant WW process, with the same final state of the signal.

distinguish between signal and background contributions.
Since the Higgs boson is a scalar particle, according to the angular momentum conservation law, fermions spins must sum up to zero. Moreover, each lepton-neutrino pair is generated by the decay of a W boson, which is a spin 1 system. The overall decay chain forces the two leptons (and neutrinos) spatial momenta to be preferably collinear, and the spin correlation, shown in Figure 6.7, is exploited to tag this final state. The main background contributions to this channel, which have the same final state of Equation 6.21, are reported in the following.

Non-Resonant WW Background

One of the most important background sources is the production of a W^+W^- pair that does not involve a Higgs boson, which is referred to as non-resonant WW production. If both W bosons decay leptonically, the final state has the same experimental signature as the signal. The leading order Feynman diagram is shown in Figure 6.8. In this case the background is referred to as irreducible, since only kinematic selections can mitigate its contribution.
For instance, the direction of the resulting lepton momenta can be used to discriminate the signal and the background. In the resonant production (i.e. when a Higgs boson is produced), the invariant mass of the lepton system:

$$m_{\ell\ell} = \sqrt{2p^{\ell_1}p^{\ell_2}(1-\cos\Delta\alpha)} \tag{6.22}$$

is smaller, since the separation angle $\Delta\alpha$ tends to be smaller. The lepton candidates with the highest and second highest p_T values are referred to as ℓ_1 and ℓ_2 respectively. The $m_{\ell\ell}$ variable tends to be larger in the non-resonant case also because both W bosons are produced on-shell. Therefore, this background can be distinguished through the shape of the $m_{\ell\ell}$ distribution.

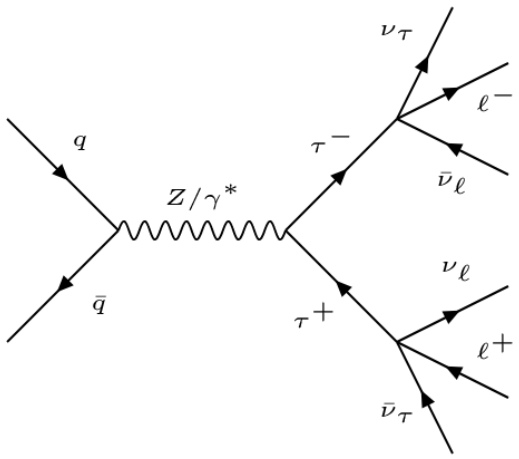


Figure 6.9: Leading order Feynman diagram for $DY \to \tau\tau$ processes, with the same final state of the signal.

Drell-Yan Background

The Drell-Yan (DY) processes occur when a $q\bar{q}$ pair annihilate in a photon or Z boson, which subsequently splits into two fermions. If two leptons are produced together with instrumental E_T^{miss} (defined in Equation 2.8), due to detector or instrumental inefficiencies, the DY process can mimic the signal final state. By requiring two different flavour leptons (i.e. one electron and one muon), this background can be suppressed. However, the contribution from $DY \to \tau\tau$ processes can not be completely eliminated, due to lepton decays of the τ: $\tau \to \ell\nu_\ell\nu_\tau$ (where ℓ is an electron or a muon). This process can lead to the same signal final state, as shown in Figure 6.9. However, leptons from τ decays are less energetic than those originating from W boson, since the Higgs mass is larger with respect to the Z boson and because only two particles emerge from a W decay, whereas the τ decays into three particles. This background can be reduced by requiring a high di-lepton momentum:

$$p_T^{\ell\ell} = |\boldsymbol{p}_T^{\ell_1} + \boldsymbol{p}_T^{\ell_2}| \tag{6.23}$$

Another important variable to discriminate this background is the Higgs boson transverse mass:

$$m_T^H = \sqrt{2p_T^{\ell\ell} E_T^{miss}[1 - \cos\Delta\phi(\boldsymbol{p}_T^{\ell\ell}, \boldsymbol{p}_T^{miss})]} \tag{6.24}$$

where $\Delta\phi(\boldsymbol{p}_T^{\ell\ell}, \boldsymbol{p}_T^{miss})$ is the opening angle between $\boldsymbol{p}_T^{\ell\ell}$ and $\boldsymbol{p}_T^{miss}$.

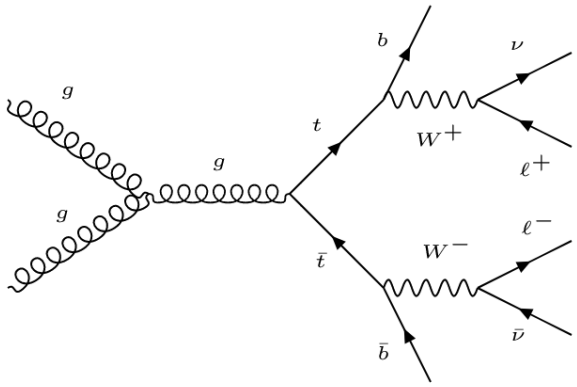


Figure 6.10: Leading order Feynman diagram for a $t\bar{t}$ pair production process, with the same final state of the signal.

Top Quark Background

The top quark is responsible for another major background. Top quarks undergo the weak decay $t \to bW^+$ before being able to form a bound state (the BR is close to 100% [1]). Processes with a $t\bar{t}$ pair production, or with a tW associated production can mimic the signal final state if the W bosons decay leptonically. An example of a $t\bar{t}$ pair production process is shown in Figure 6.10.
These processes are characterised by the presence of b quarks in the final state: jets originating from these quarks contain B hadrons, which are characterised by a long lifetime ($\sim 10^{-12}$ s) and a high mass (~ 5 GeV). Since B hadrons usually decay inside the tracker, it is possible to distinguish the secondary decay vertexes. Dedicated b-tagging algorithms allow for the identification of jets likely originating from b quarks with respect to jets originating from lighter quarks or gluons [69]. Therefore, the top quark background can be reduced by imposing the absence of b-tagged jets in the final state. This requirement is not sufficient on its own because b-tagging algorithms are not fully efficient in recognising b-jets.
Since the two W bosons are not produced by a Higgs resonance, the kinematic selections used to distinguish non-resonant WW events help suppressing this background as well.

Non-Prompt Background

Another important background is due to non-prompt leptons i.e. not originating from the hard scattering process. The typical scenario is given by events in which a W boson is produced in association with jets: one of the jets can be misidentified as a lepton, hence populating the signal region. In this case the lepton is referred to as non-prompt, being either a lepton generated from a secondary vertex in a b-jet or a pion misidentified as an electron. The latter can happen if a charged pion undergoes a charge exchange reaction with the detector material, for instance $\pi^- + p \to \pi^0 + n$. The resulting π^0 decays into photons (with a BR close to 99%), giving rise to a cluster in ECAL.
This background can be suppressed by applying isolation criteria on electrons and muons, since non-prompt leptons are characterised by the presence of many nearby particles. The isolation criteria must also account for pile-up effects, since prompt leptons can be surrounded by charged hadrons from pile up vertices. The relative isolation variable for muons is defined as:

$$I^{rel}_{muon} = \left[\sum_{ChH} p_T + \mathrm{Max}\left(0, \sum_{NH} p_T + \sum_{Ph} p_T - 0.5 \sum_{ChHPU} p_T\right)\right] \Big/ p_T^{muon} \quad (6.25)$$

where the subscripts ChH, NH, Ph and ChHPU refer to charged hadrons, neutral hadrons, photons and charged hadrons from pileup vertexes, respectively. The sum is calculated in a cone of radius $\Delta R < 0.4$ around the muon direction. The subtractive term is actually a correction for neutral hadrons coming from pile-up, as their contribution is estimated to be roughly half (hence the 0.5 factor) of that of charged hadrons. This correction is motivated by the fact that neutral hadrons do not leave any track in the tracker, therefore they can not be assigned a vertex with sufficient precision. The applied cut on the isolation variable is usually $I^{rel}_{muon} < 0.15$.
The relative isolation variable for electrons is defined as:

$$I^{rel}_{electron} = \left[\sum_{ChH} p_T + \mathrm{Max}\left(0, \sum_{NH} p_T + \sum_{Ph} p_T - \rho A\right)\right] \Big/ p_T^{electron} \quad (6.26)$$

where ρ is the energy density due to pile-up events and A is an effective area [68]. The applied cut on the isolation variable is usually $I^{rel}_{electron} < 0.06$.

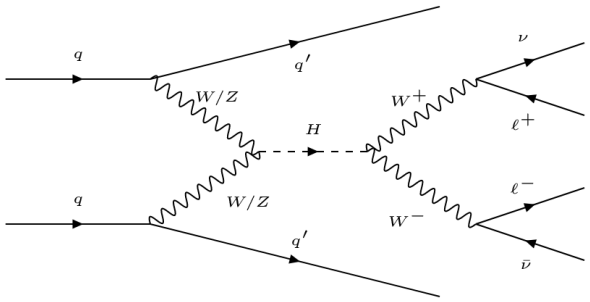


Figure 6.11: Feynman diagram for the *VBF* Higgs production mechanisms in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay channel.

Other Backgrounds

Other minor backgrounds include:

- The WZ process mimics the signal final state if the Z boson decays leptonically or if both W and Z bosons decay leptonically and one of the three charged leptons is not reconstructed.
- The $W\gamma^*$ background[4] is similar to the previous one, but the two leptons from the γ^* decay have a lower invariant mass.
- The $W\gamma$ process can contribute to the backgrounds if the photon is converted to electrons by interacting with the detector material.
- ZZ or VVV (with $V = W, Z$) events can enter the selection due to inefficiencies in lepton reconstruction.

6.4.1 The *VBF* Higgs Production Mechanism

Currently, all production processes have been observed in one or more decay channels or via combination of several decay channels, with no significant deviations with respect to the SM prediction. However, an accurate study of the *VBF* mechanism is still crucial to test the properties of the new particle. Such a rare process is sensitive to new physics phenomena and allows to set stringent constraints on the compatibility of the Higgs boson with the SM. The WW decay channel, thanks to its large BR, is ideal for the observation of this production process.
The aim of this Thesis is the measurement of the Vector Boson Fusion Higgs

[4]The * superscript indicates an off-shell particle

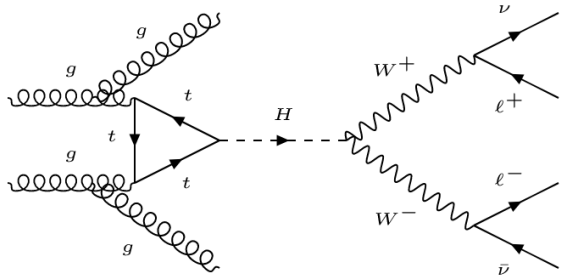


Figure 6.12: Feynman diagram for the *ggH* Higgs production mechanisms in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay channel, with jets coming from radiative processes.

production mechanism in the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay channel with the full LHC Run-2 dataset. The leading order Feyman diagram for the *VBF* production in this decay channel is reported in Figure 6.11. The experimental signature of this production mechanism is the presence of two jets emerging from the outgoing partons. All background processes presented in the previous section can feature jets in the final state due to radiative processes[5]. However, the *VBF* jet topology can be used to distinguish the signal from the backgrounds.

The quark pair produces jets with a large pseudorapidity gap $|\Delta\eta_{jj}|$, with the one coming from the q quark typically emitted in the forward region. This is because q, a valence parton, can carry a significant fraction of the proton momentum, as shown in Figure 6.3. The pseudorapidity gap is caused by the fact that no coloured particle is exchanged between the interacting quarks [70]. This results in high values of the invariant mass of the two jets, m_{jj}. Both $|\Delta\eta_{jj}|$ and m_{jj} can be used to separate the *VBF* production mode from the backgrounds.

The measurement of the *VBF* mechanism is complicated by the fact that the *ggH* mechanism is a very challenging background as well, since it shares the same final state including the presence of the Higgs boson, as shown in Figure 6.12. Since jets in the *ggH* final state are due to radiative processes, they tend to have a lower pseudorapidity gap. Therefore, high thresholds on $|\Delta\eta_{jj}|$ and m_{jj} can help to distinguish the two production modes. Figure 6.13 shows distributions of these two variables for simulated samples of *VBF* and

[5]When two partons take part in the hard scattering, accelerated colour charges are present, thus emission of radiation may occur. This effect is called Initial State Radiation (ISR). Also the final state partons can produce further radiation, called Final State Radiation (FSR).

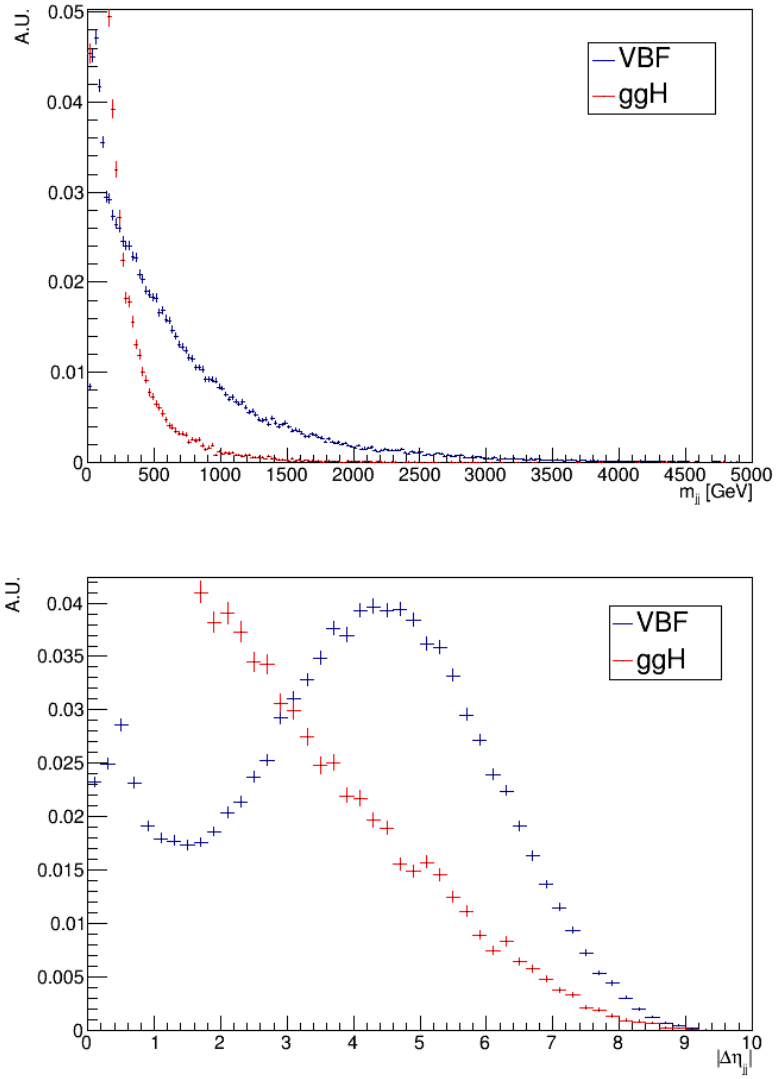


Figure 6.13: Normalised m_{jj} (top) and $|\Delta\eta_{jj}|$ (bottom) distributions for VBF and ggH simulated samples.

ggH mechanisms.
One of the most recent CMS analysis in the WW decay channel [71] was centred on the measurement of the global production cross section, and it was not optimised with respect to the *VBF* production mode. The *VBF* signal extraction was carried out through a fit on the $m_{\ell\ell}$ distribution in two different phase space regions, one for low m_{jj} values and one high m_{jj} values (higher than 700 GeV). Indeed, $m_{\ell\ell}$ is a good discriminating variable and the split in m_{jj} helps increasing the sensitivity to the signal, since the high m_{jj} region is enriched with *VBF* events. However, the analysis was barely sensitive to the *VBF* signal.
In order to improve the sensitivity to the *VBF* process, I implemented a Multi-Variate Analysis (MVA) based on a multi-classification Deep Neural Network (DNN). This approach was motivated by the fact that one of the main backgrounds of this analysis is another Higgs production process. Therefore, disentangling the signal from the backgrounds is particularly challenging, and the DNN was employed for this scope.
The new analysis performed by the CMS collaboration, to which I contributed, aims to measure the Higgs boson properties in the $H \to W^+W^-$ decay channel targeting not only the *VBF* production mechanism, but also *ggH* and *VH*, considering final states with at least two charged leptons arising either from the associated vector boson or from the products of the Higgs boson decay. The Higgs boson properties are probed by measuring the inclusive cross sections. The analysis is based on proton-proton collision data produced at the LHC with $\sqrt{s} = 13$ TeV and collected by the CMS detector during the LHC Run-2, for a total integrated luminosity of about $137\,\text{fb}^{-1}$.

Chapter 7

Analysis Strategy

7.1 Analysis Overview

In this Chapter, the analysis of the Higgs signal in the WW decay channel with full Run-2 dataset is presented. The analysis is composed by several Higgs production channels, and I mainly contributed to the *VBF* measurement. Therefore, this channel is described in details, while also giving the salient information on the other channels that contribute to the measurement. The description of the other channels was included in order to properly describe the analysis results, that are presented in Chapter 8. Indeed, in order to give a comprehensive view of the contribution of my channel to the complete analysis, the results obtained by combining all the analysis channels are reported together with a set of results limited to the channel I studied.

7.2 Event Reconstruction

The event reconstruction in the CMS experiment is performed by the Particle Flow (PF) algorithm [72], developed by the Collaboration and used in all analyses. The PF technique allows for the identification of different physics objects, as well as the measurement of their properties. This is achieved by combining informations from all the CMS detectors. The PF candidates are identified in the following order:

- **Muons** are identified and their momenta are measured in the pseudorapidity interval $|\eta| < 2.5$, as tracks in the central tracker consistent with either a track or several hits in the muon system. The matching can be done both inside-out (starting form the tracker) as well as outside-in (starting from hits in the muon system). The efficiency to

reconstruct and identify muons is greater than 96%, while the relative transverse momentum resolution for muons with p_T up to 100 GeV is 1% in the barrel and 3% in the endcaps. The p_T resolution in the barrel is better than 7% for muons with p_T up to 1 TeV.

- **Electrons** are identified and their momenta are measured in the pseudorapidity interval $|\eta| < 2.5$ by combining tracks in the central tracker with spatially compatible energy deposits in ECAL. Electrons typically emit bremsstrahlung radiation when traversing the tracker, resulting in additional ECAL deposits spread along ϕ due to bremsstrahlung photons. These deposits are grouped together in a super-cluster, which is then matched with a track from the central tracker to define a candidate electron.
  The efficiency to reconstruct and identify electrons ranges between 60% and 80% depending on the lepton p_T and identification and isolation criteria. The p_T resolution for electrons with $p_T \simeq 45\,\text{GeV}$ from $Z \to ee$ decays ranges from 1.7% to 4.5% depending on the η region. The resolution is generally better in the barrel than in the endcaps and also depends on the bremsstrahlung energy emitted by the electron.

- **Charged Hadrons** are identified from remaining tracks associated to energy clusters in ECAL and HCAL. The energy of charged hadrons is determined from a combination of their momenta (measured with the tracker) and the matching energy deposits of ECAL and HCAL.

- **Photons and Neutral Hadrons** are identified from clusters in ECAL and HCAL (respectively) that do not match any track.

Multiple proton-proton interaction vertices are identified from tracking information by using the Adaptive Vertex Fitting algorithm [73]. The primary proton-proton interaction vertex is defined as the candidate vertex with the largest value of summed physics object p_T^2. The physics objects include those returned by a jet-finding algorithm applied to all charged tracks assigned to the vertex, and the associated missing transverse momentum, defined as the negative vector sum of the p_T of those objects.
Jets are reconstructed by clustering PF candidates using the $anti-kT$ algorithm (described in Section 6.2.1), with a distance parameter $R = 0.4$. The additional proton-proton interactions (i.e. the pile-up effect) can be responsible for additional tracks and calorimetric energy depositions in the reconstructed jet. To mitigate this effect, charged particles identified to be originating from pile-up vertices are discarded and an offset correction is applied to correct for remaining contributions.

The measured jet energy does not correspond to its true value, primarily due to the non-uniform and non-linear response of HCAL. Jets are corrected with Jet Energy Correction (JEC) factors in bins of jet p_T and η, to reproduce on average the true jet four-momentum [74]. Moreover, measurements show that the Jet Energy Resolution (JER) in data is worse than in simulation, therefore simulated jets are smeared to match the measured resolution. The jet energy resolution typically is 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV.
Additional selection criteria are applied in order to remove jets potentially dominated by anomalous contributions from various detector components or reconstruction failures. For instance, while all jets measured in the range $|\eta| < 4.7$ were considered in the analysis of the data collected in 2016 and 2018, for the analysis of the data recorded in 2017 all jets in the range $2.5 < |\eta| < 3.0$ were excluded due to anomalous detector noise in that region.
Various b-tagging algorithms have been developed by the CMS collaboration, making use of Boosted Decision Tree (BDT) and DNN machine learning techniques. For each jet in the event, a score is calculated through a multivariate combination of different jet properties. Jets are considered b-tagged if their associated score exceeds a threshold, tuned to achieve a certain tagging efficiency in top quarks events (characterised by b-jets in the final state).
The missing transverse momentum vector $\boldsymbol{p}_T^{miss}$ (described in Section 2.2) is computed as the negative vector sum of the transverse momenta of all the PF candidates in an event [75]. The Pile-Up Per Particle Identification (PUPPI) algorithm is applied to reduce the pile-up dependence of the $\boldsymbol{p}_T^{miss}$ observable [76].

7.3 Data and Simulated Samples

The datasets used in the analysis were collected by the CMS experiment in 2016, 2017 and 2018, corresponding to integrated luminosities of 35.9 fb^{-1}, 41.5 fb^{-1} and 59.7 fb^{-1} respectively.
Events are triggered by requiring the presence of either one or two high-p_T electrons or muons. For the 2016 dataset, the single electron trigger requires a p_T threshold of 25 GeV for electrons with $|\eta| < 2.1$ and 27 GeV with $2.1 < |\eta| < 2.5$. For the single muon trigger the p_T threshold is 24 GeV for muons with $|\eta| < 2.4$. In the di-electron (di-muon) trigger the p_T thresholds of the leading[1] and trailing electron (muon) are respectively of 23 GeV

[1] All physics objects are listed in decreasing p_T order: the first and the second are called leading and trailing respectively.

(17 GeV) and 12 GeV (8 GeV). In the di-lepton ($e\mu$) trigger, the p_T thresholds are 23 GeV and 12 GeV for the leading and trailing lepton respectively. For the first part of data taking in 2016, a lower p_T threshold of 8 GeV for the trailing muon was used.
In the 2017 dataset the p_T thresholds of the single electron and single muon triggers were raised respectively to 35 GeV and 27 GeV, while they were set to 32 GeV and 24 GeV in the 2018 dataset. For both 2017 and 2018 datasets, the p_T thresholds of the di-lepton triggers are kept the same as the last part of the 2016 dataset.
Monte Carlo (MC) event generators are employed in the analysis to model signal and background processes. Three independent sets of simulated events (corresponding to the 2016, 2017, and 2018 detector conditions) are used for each process of interest, in order to take into account year dependent effects in the CMS detector, data taking and event reconstruction. Despite different Matrix Element (ME)[2] generators are used for different processes, all simulated events corresponding to a given dataset share the same set of PDFs, the same description of underlying events, and the same Parton Shower (PS)[3] configuration. In particular, the ME event generators are interfaced with PYTHIA8 [77] for the PS evaluation: it allows for the simulation of initial and final state radiations, hadronisation and underlying events. The PDF set used is NNPDF [78].
Simulated events are used in the analysis to model the Higgs boson production through *ggH*, *VBF*, *VH* and *ttH*[4] using POWHEG v2 [79, 80, 81] event generator, which describes the Higgs boson production at Next-to-Leading Order (NLO) accuracy in QCD. The Z boson associated production (ZH) is simulated including both gluon and quark induced contributions. The MiNLO HVJ extension [82] of POWHEG v2 is used for the simulation of the W boson associated production (WH) and the quark induced ZH production, providing a description of VH processes with zero and one jet processes in the final state with NLO accuracy.
For the ggH production, the simulated events are reweighted to match the

[2]The exact matrix element (hence the cross-section) calculation of the Feynman diagrams of the hard process of interest. The calculation can be performed at various perturbative orders.

[3]The ME calculation can suffer from infrared and collinear divergences of QCD, that would require higher order calculations to be corrected. Instead, PS algorithms offer an alternative and simple way to eliminate these divergences and to handle the complexity of several successive branchings. However, since the PS machinery relies on a collinear approximation of the ME, it is supposed to perform well in the description of the evolution of jets, but not to provide a precise description of configurations with well separated partons. The ME and PS methods are therefore complementary.

[4]This process has a negligible contribution in the phase space of this analysis.

Next-to-Next-to-LO PS (NNLOPS)[5] prediction [83, 84] in the hadronic jet multiplicity (i.e. the number of jets in the final state, n_{jet}) and Higgs boson transverse momentum distributions. However, in the $n_{jet} \geq 2$ phase space, only a LO accuracy is achieved. In order to have a better description, in the $n_{jet} \geq 2$ phase space the MiNLO HJJ generator is used instead of the reweighting: this guarantees NLO accuracy for $n_{jet} \geq 2$ and LO for $n_{jet} \geq 3$. The simulated samples are normalised to the cross sections recommended by the LHC Higgs Cross Section Working Group [62]. For instance, the Next-to-Next-to-Next-to-LO (N3LO) cross section is used to normalise the *ggH* sample.

The Higgs boson mass in the event generation is assumed to be 125 GeV, and the decay to a pair of W bosons and subsequently to leptons is performed using the JHUGen generator [85] for *ggH*, *VBF* and quark-induced *ZH* samples. For the other signal simulations, the Higgs boson and W boson decays are performed using PYTHIA 8. Finally, for the *ggH*, *VBF* and *VH* production mechanisms, additional Higgs boson simulations are produced using the POWHEG v2, where the Higgs boson decays to a pair of τ leptons.

The background processes are simulated using several event generators. The quark-initiated non-resonant WW process is simulated using POWHEG v2 with NLO accuracy [86]. The MCFM v7 generator [87, 88] is used for the simulation of gluon-induced non-resonant WW production at LO accuracy, and the normalisation is chosen to match the NLO cross section [89]. The non-resonant electroweak production of W^+W^- pairs with two additional jets, that is the Vector Boson Scattering (*VBS*) topology, is simulated at LO accuracy with with MadGraph5_aMC@NLO [90].

Top pair production ($t\bar{t}$) as well as single top processes (such as the tW associated production) are simulated with POWHEG v2 [91, 92, 93]. The Drell-Yan production of a charged lepton pair is simulated at NLO accuracy with MadGraph5_aMC@NLO.

For all processes, the detector response is simulated using a detailed description of the CMS detector, based on the Geant4 package [94].

[5]Matching of the NNLO ME with the PS event generator to avoid double counting effects.

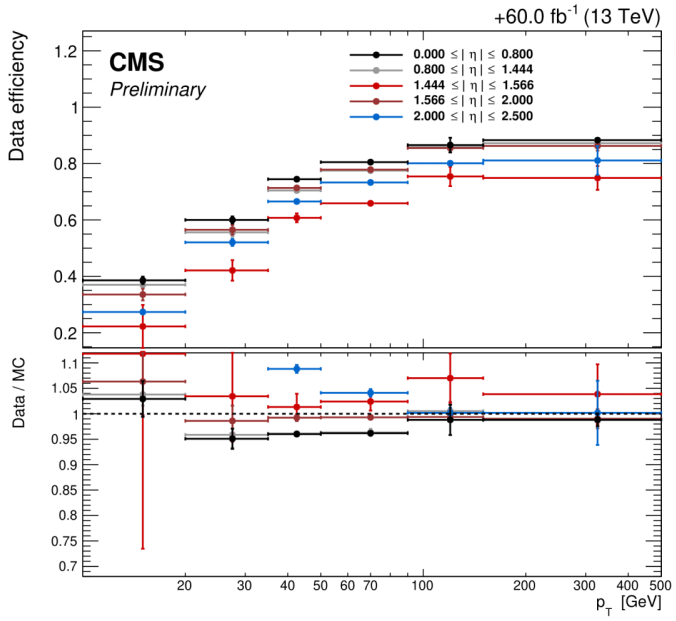


Figure 7.1: Data isolation and identification efficiencies (top) and corresponding scale factors (bottom) as a function of p_T and η for electrons in the 2018 dataset.

7.3.1 Simulation Corrections

Discrepancies between data and simulation in various distributions have multiple causes, such as the difference in the pile-up distribution and the imperfect modelling of the detector, leading to offsets in event reconstruction efficiency. In order to correct these discrepancies, each Monte Carlo event is reweighted with several Scale Factors (SFs), aiming to correct any disagreement with data at the best possible level. The main SFs are reported in the following.

- **Lepton Identification and Isolation SF:** the efficiencies related to the identification and isolation selections applied to muons and electrons can show differences between data and Monte Carlo simulations. The isolation selections are given by Equation 6.25 and Equation 6.26 for muons and electrons respectively, while the identification selection is performed by dedicated algorithms aiming to select genuine leptons with a known efficiency while rejecting fake ones. In order to estimate these efficiencies, both in data and simulations, the Tag and Probe method is employed.
  First, events with two electrons (muons) with an invariant mass close

to that of the Z boson are selected: one electron (muon) is required to be very well reconstructed (the "tag" lepton), while the other electron (muon) is only required to pass a loose selection (the "probe" lepton). After that, the probe is tested against the (electron or muon) selections. Non-Z events are subtracted with a fit to the invariant mass of the di-lepton mass allowing for an exponential or polynomial background. The selection efficiency is then evaluated with:

$$\epsilon(p_T, \eta) = \frac{N_{pass}^{probe}}{N_{pass}^{probe} + N_{fail}^{probe}} \tag{7.1}$$

The efficiencies are estimated separately for electrons and muons, and as a function of the probe lepton p_T and η.
These efficiencies are computed both in data (ϵ_{data}) and simulation (ϵ_{MC}). The scale factor is then calculated by taking the ratio of the data and simulation efficiencies ($\epsilon_{data}/\epsilon_{MC}$), which is then applied to reweight simulated events. Figure 7.1 shows data identification and isolation efficiencies and the corresponding scale factors as a function of p_T for electrons in the 2018 dataset, for different $|\eta|$ intervals.

- **Pile-Up SF:** the distribution of the number of pile-up interactions is different between data and simulation, because the instantaneous luminosity can vary during data taking. A dedicated scale factor to address this issue is calculated as the ratio between the measured and the estimated number of pile-up interactions. The average number of pile-up interactions was 23 (32) in 2016 (2017 and 2018).

- **Heavy Flavour Tagging Efficiency SF:** for each jet in the event where the b-tagging discriminator is evaluated, a specific scale factor depending on the jet p_T, η, flavour, and the b-tagging discriminator value is applied to the event. The scale factors are designed to make the distribution of the discriminator value in the simulation close to that found in collision data.

- **L1 Trigger Pre-Firing SF:** in 2016 and 2017 datasets an issue caused highly energetic readouts in the ECAL endcaps (from jets, photons and electrons) to be assigned by the L1 Trigger to the previous bunch-crossing. This is known as "pre-firing", as the Trigger is fired to the previous bunch-crossing, rather than the bunch-crossing where the highly energetic object came from. The event with interesting physics is therefore lost in data, while the issue is not reproduced in MC events. In order to correct this difference, a weight is applied to all events in MC,

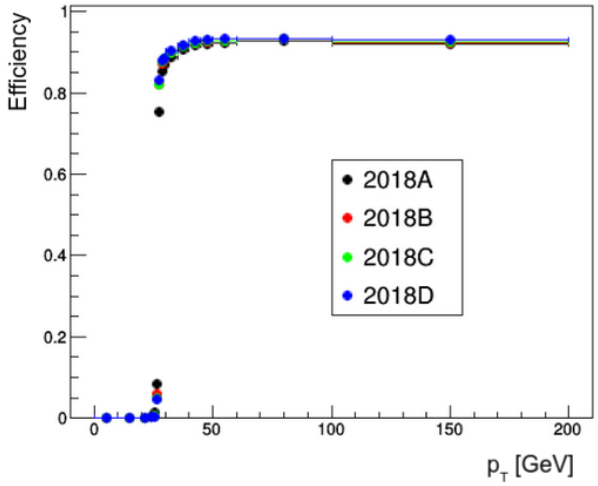


Figure 7.2: Trigger efficiency as a function of p_T for a single muon trigger in the 2018 dataset. Events are separated per periods of data taking.

which is essentially $(1 - x)$, where x is the probability of an event to be pre-fired, based on all objects that may induce the pre-firing.

- **Jet Pile-Up SF:** additional scale factors related to the identification of pile-up jets are applied, for possible discrepancies between data and MC events in terms of mis-tag rate or efficiency.

The high level lepton triggers are characterized by p_T thresholds above which the trigger efficiency is very high (plateau region). Nevertheless, the trigger efficiency as a function of the lepton p_T is not a step function, but is characterized by a steep increase of the efficiency around the p_T threshold. This turn-on region is due to the fact that the HLT performs a rougher estimate of the lepton transverse momentum with respect to the full reconstruction.
The trigger selection is not applied to MC samples: instead they are weighted with the trigger efficiency, measured from data as a function of p_T and η of the leptons.
The trigger efficiencies are also evaluated with the Tag and Probe method. Figure 7.2 shows the trigger efficiency as a function of p_T for a single muon trigger in the 2018 dataset. Since a mixture of single and double triggers is used in the analysis, the combined efficiency has to be computed and applied to simulated samples as an event weight.

Channel	N_{lepton}	N_{jet}
(DF, SF) ggH	2	-
(DF, SF) VBF	2	2
(DF, SF) $VH2j$	2	2
$WHSS$	2	> 0
$WH3\ell$	3	0
$ZH3\ell$	3	> 0
$ZH4\ell$	4	0

Table 7.1: Overview of the selection defining the analysis channels. N_{jet} indicates the number of jets with $p_T > 30$ GeV required in the analysis, while N_{lepton} the number of required leptons (the requirements on the leptons momenta depend on the channel and on the dataset).

7.4 Event Categorisation

The Higgs production and decay channels are studied through the definition of categories for the selected events.
First, in order to tag the WW decay channel of the Higgs boson, events in the analysis are selected by requiring a pair of oppositely charged leptons (electrons or muons) with high p_T , high E_T^{miss} and a varying number of jets. The Higgs boson production via ggH, VBF and VH with a hadronically decaying vector boson ($VH2j$ channel), are measured in Different-Flavour (DF) and Same-Flavour (SF) channels by selecting opposite-sign $e\mu$ and $ee/\mu\mu$ pairs respectively. The DF channels benefit from the high suppression of the DY background. Four channels target the VH production with a leptonically decaying vector boson, depending on the number of required leptons and jets: $WHSS$ (Same-Sign leptons), $WH3\ell$, $ZH3\ell$ and $ZH4\ell$. An overview of the different channels is reported in Table 7.1.
Each channel is further split into different regions, constructed through cuts on variables that show some discrimination power between signal and background processes. Signal Regions (SRs) are defined to enhance the sensitivity with respect to the process of interest, while Control Regions (CRs) are constructed to be enriched in a particular background process (and so that the number of signal events is expected to be low). Their purpose is to extract background normalisation from data, as explained in Section 7.6.
Pre-selection criteria are applied in all phase space regions. The thresholds on both leading and trailing leptons p_T are slightly above the trigger con-

ditions described in Section 7.3. In the 2016 dataset, the leading lepton in the event is required to have $p_T^{\ell 1} > 25$ GeV (20 GeV) for electrons (muons), while the trailing lepton is required to have $p_T^{\ell 2} > 13$ GeV (10 GeV) for electrons (muons). In the 2017 and 2018 datasets the thresholds for muons were raised to 25 GeV and 13 GeV for the leading and trailing leptons respectively, matching the electrons requirements. The difference across the datasets is due to the different trigger setup.
The statistical approach used to interpret the datasets for this analysis and to combine the results from the independent categories was developed by the ATLAS and CMS collaborations in the context of the LHC Higgs Combination Group [96]. The signal extraction procedure, presented in detail in Section 8.1, is performed using binned templates based on variables that allow for a good discrimination between signal and background. Thus, every channel has its own set of discriminating variables.
The strategy of the analysis has been scrutinized and approved by a selected committee of internal reviewers of the CMS collaboration before looking at the data in the signal regions. This approach ("blind" policy) prevents the analysts and the internal reviewers from being biased by the data in the developing phase of the analysis.
My work was centred on the DF *VBF* channel: in the following the event categorisation and analysis strategy for this channel is presented, as well as a brief summary of the other channels of the analysis.

7.4.1 The Different-Flavour *VBF* Channel

Events in the SR of the DF *VBF* channel are required to have no third lepton with $p_T > 10$ GeV in order to suppress processes with three or more leptons in the final state. Moreover, exactly two jets with $p_T > 30$ GeV are required ($N_{jet} = 2$), with an invariant mass $m_{jj} > 120$ GeV. Contributions arising from top quark production are reduced by vetoing events containing any jet with $p_T > 20$ GeV that was flagged as b-jet by the dedicated b-tagging algorithm. The di-lepton invariant mass $m_{\ell\ell}$ is required to be higher than 12 GeV to suppress low mass resonances and QCD events with non-prompt leptons. Events with no genuine E_T^{miss} (arising due to the presence of neutrinos in signal events), as well as $DY \to \tau\tau$ events, are suppressed by requiring $E_T^{miss} > 20$ GeV. The latter are further reduced by requiring $p_T^{\ell\ell} > 30$ GeV (as explained in Section 6.4).
In order to further suppress the $DY \to \tau\tau$ process and W production is association with jets (with a non-prompt lepton), the following transverse

Region	**Selection**
Global Selection	$e\mu$ pair with opposite charge $p_T^{\ell_1} > 25$ GeV, $p_T^{\ell_2} > 13$ GeV, $p_T^{\ell_3} < 10$ GeV $p_T^{\ell\ell} > 30$ GeV, $m_{\ell\ell} > 12$ GeV $E_T^{miss} > 20$ GeV $N_{jet} = 2$, $m_{jj} > 120$GeV
SR	60 GeV $< m_T^H < 125$ GeV $m_T^{\ell_2, E_T^{miss}} > 30$ GeV No b-tagged jets with $p_T > 20$ GeV
Top Quark CR	$m_{\ell\ell} > 50$ GeV $m_T^{\ell_2, E_T^{miss}} > 30$ GeV Leading or trailing jets b-tagged
$DY \to \tau\tau$ CR	$m_T^H < 60$ GeV 30 GeV $< m_{\ell\ell} < 80$ GeV No b-tagged jets with $p_T > 20$ GeV

Table 7.2: Global selection used in the DF *VBF* channel, and specific variable cuts employed in the signal region and in the control regions. In the 2016 datasets, the p_T thresholds for the leading and trailing leptons were 25 GeV (20 GeV) and 13 GeV (10 GeV) for electrons (muons) respectively.

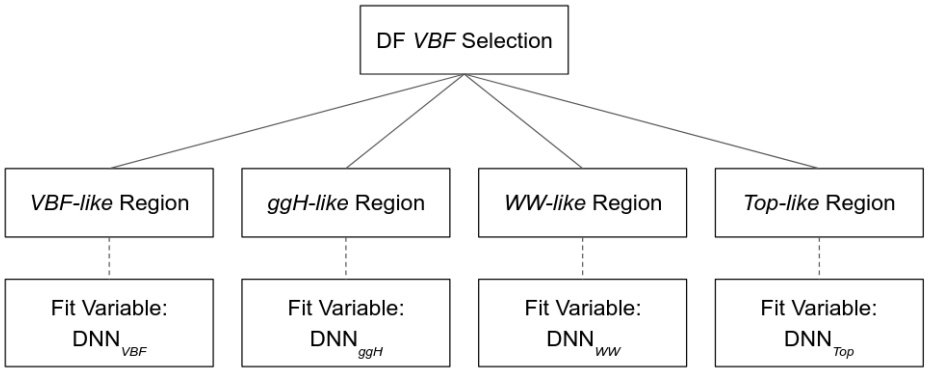


Figure 7.3: Schematic view of the SR sub-categories in the DF *VBF* channel and the respective variable used in the signal extraction fit.

mass is employed:

$$m_T^{\ell 2, E_T^{miss}} = \sqrt{2 p_T^{\ell 2} E_T^{miss} [1 - \cos \Delta\phi(\boldsymbol{p}_T^{\ell 2}, \boldsymbol{p}_T^{miss})]} \tag{7.2}$$

which is required to be above 30 GeV. Finally, m_T^H (described in Section 6.4) is required to be higher than 60 GeV.
Two control regions are also defined, enriched in top quark and $DY \to \tau\tau$ events respectively. The top quark CR is defined with the same selection of the SR, but requiring either the leading or the trailing jets to be b-tagged. Moreover, the threshold on $m_{\ell\ell}$ is set to 50 GeV in order to reduce contributions coming from processes that involve the production of a Higgs boson.
The $DY \to \tau\tau$ CR is defined by inverting the SR cut on m_T^H. Moreover, it is also required that 30 GeV $< m_{\ell\ell} <$ 80 GeV, since the DY phase space is populated at this values. An overview of the selections in the DF *VBF* channel is reported in Table 7.2: the global selection defines the DF *VBF* phase space, while specific selections define the SR and the two CRs.
In order to increase the *VBF* sensitivity in the analysis, an approach based on a multi-classification DNN was followed. Unlike binary classification that distinguishes between two classes of events (usually one signal process and all background processes combined), multi-classification has the advantage that different physics processes can be treated with equal importance. This is especially desirable when trying to isolate the *VBF* production, since one of the main backgrounds is another Higgs production mechanism (ggH): the kinematic selection employed to reduce other backgrounds is not necessarily adequate to suppress the ggH production. Technical details on the DNN I

developed (and referred to as *VBF*-DNN in the following) are reported in Section 7.5.
The *VBF*-DNN is constructed to perform a multi-classification of an event as either signal (*VBF*) or one of the three main background processes, namely non-resonant WW, top quark and *ggH*. Drell-Yan has not been taken into account for the training since it is suppressed by pre-selection criteria. The input variables are various kinematic properties of the particles in in the final state. As a result, an event is attributed a vector of four components $\boldsymbol{\mathcal{O}}$. The numerical value of each of these components represents the degree of agreement of such event with the signal and the three background processes: these outputs can be interpreted as a probability, since they are normalised to one. Therefore, for a given event, the process j with the highest output $\mathcal{O}_j$ is interpreted as the most probable process.
The four outputs are referred to as classifiers: DNN_{VBF}, DNN_{ggH}, DNN_{WW} and DNN_{Top}. It should be emphasised that the first two classifiers target two different Higgs production mechanisms. In the SR, four orthogonal sub-categories are made using the classifiers. This strategy aims to define a signal region sub-category, the *VBF-like*, with a high signal purity. Indeed, most of the background events are expected to fall in the other three sub-categories. The signal extraction fit is performed on binned templates of the classifiers: in the *j-like* sub-category the DNN_j classifier is used as a fit variable. The bin boundaries are chosen with an algorithm that aims to split the bins as finely as possible, while also requiring at least five expected signal events and ten expected signal plus background events in each bin as well as a maximum relative statistical uncertainty on the background of 20%. Figure 7.3 shows a schematic view of the sub-categories and the respective fit variables.

7.4.2 The Different-Flavour *ggH* Channel

Events in the SR of the DF *ggH* channel are required to pass a very similar selection to the DF *VBF* channel, the only difference being the number of jets in the final state. In order to exploit the peculiar kinematics of the target final state, events with $N_{jet} = 0$, $N_{jet} = 1$ and $N_{jet} > 1$ are separated into distinct categories.
In order to better constraint the W production in association with jets (with a non-prompt lepton), the $N_{jet} = 0$ and $N_{jet} = 1$ categories are subdivided into two sub-categories each, according to the charge of the leading lepton. This subdivision is motivated by the fact that while the signal is charge symmetric, in W+jets events W^+ bosons are more abundant than W^- bosons. Finally, these two sub-categories are further divided according to whether the p_T of the trailing lepton is above or below 20 GeV. This is done in order to

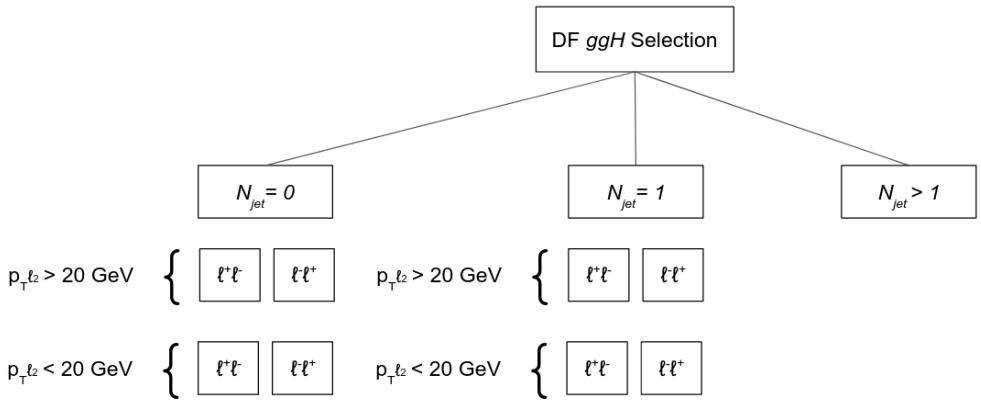


Figure 7.4: Schematic view of the SR sub-categories in the DF ggH channel and the respective variable used in the signal extraction fit.

further constrain the non-prompt background, since non-prompt leptons are more likely at low p_T values. This results in a four-fold partitioning of the $N_{jet} = 0$ and $N_{jet} = 1$ sub-categories.
The $N_{jet} > 1$ category is not further sub-categorised. The orthogonality with the DF VBF channel is assured by a different selection on the m_{jj} variable: in this case it is required to be lower than 120 GeV. Indeed, as shown in Section 6.4.1, the ggH production in association with jets populates lower m_{jj} values with respect to the VBF production. Moreover, the region $65\,\text{GeV} < m_{jj} < 105\,\text{GeV}$ is excluded since it is used in the $VH2j$ channel.
In order to disentangle the signal from the backgrounds, both $m_{\ell\ell}$ and m_T^H are employed. Indeed, it is found that signal and background events populate different regions of the $(m_{\ell\ell}, m_T^H)$ plane. The signal extraction fit is therefore performed on a two-dimensional $(m_{\ell\ell}, m_T^H)$ binned template, allowing for good signal to background discrimination. Figure 7.4 shows a schematic view of the SR sub-categories.
As with the DF VBF channel, top quark and $DY \to \tau\tau$ CRs are defined in order to optimise the background subtraction in the SR. Two CRs are defined for each N_{jet} category, for a total of six control regions.

7.4.3 The Same-Flavour *VBF* and *ggH* Channels

The major difference between the DF and the SF channels is that in the latter the DY background is dominant, since the two leading leptons in the event are required to form an oppositely signed ee or $\mu\mu$ pair. In order to reduce this background, a veto is placed on events in which the di-lepton

invariant mass is within 15 GeV from the nominal mass of the Z boson.

- The SR of the SF *VBF* channel is defined by requiring a selection similar to the respective SR of the DF channel. Therefore, exactly two jets with $p_T > 30$ GeV are required also in this case, but with a higher invariant mass: $m_{jj} > 350$ GeV. In order to constraint the *DY* background, a dedicated multivariate discriminant based on a DNN, referred to as DYMVA [71], is employed. A selection on the DYMVA discriminant is used to define a phase space enriched in Higgs boson events. The signal is extracted with a fit on the number of events in this phase space. Two CRs are defined, one for the top quark processes, and one for the non-resonant WW production.

- The SR of the SF *ggH* channel is defined by requiring a similar selection to the respective SR of the DF channel. Also in this case the DYMVA discriminant is employed to suppress the *DY* background. Events in the SR are then divided in sub-categories based on the number of jets with $p_T > 30\,\text{GeV}$: $N_{jet} = 0$, $N_{jet} = 0$ and $N_{jet} > 1$. Further selections on m_T^H, $m_{\ell\ell}$ and the azimuthal angle between the two leading leptons $\Delta\phi_{\ell\ell}$ are applied depending on the sub-category. The signal is extracted with a simultaneous fit on the number of events in each sub-category. For each N_{jet} category, two CRs are defined, one for the top quark process and one for the non-resonant WW production.

7.4.4 The *VH* Channels

The channels targeting the VH Higgs production mechanism are defined depending on the decay of the V boson. The $VH2j$ channel, targeting events in which the vector boson decays into two resolved jets, is split in DF and SF channels. Each of the four VH channels with a leptonically decaying vector boson ($WHSS$, $WH3\ell$, $ZH3\ell$ and $ZH4\ell$) follow specific strategies.

- **DF *VH2j***: the final state, and therefore the SR selection, is analogous to the DF *ggH* channel with $N_{jet} = 2$, with the added requirement that the invariant mass of the two jets has to be close to that of the W and Z bosons: $65\,\text{GeV} < m_{jj} < 105\,\text{GeV}$. Figure 7.5 shows the m_{jj} distribution for *VBF*, *ggH* and *VH* productions: the *VH* production is peaked in the selected m_{jj} region.
  The signal extraction fit is performed on a template shape of the dilepton invariant mass $m_{\ell\ell}$, which has a different profile for the signal and the non-resonant WW background. Two CRs are defined, one for the top quark process, and one for the *DY* background.

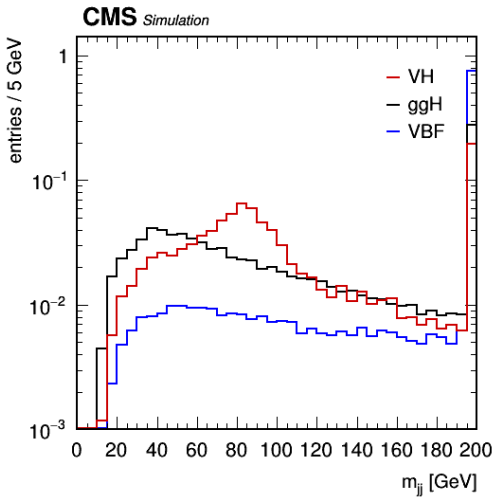


Figure 7.5: Normalised distributions of m_{jj} for *ggH*, *VBF* and *VH* Higgs production mechanisms.

- **SF $VH2j$**: the SR selection is analogous to the SF *ggH* channel with $N_{jet} = 2$. Also in this case, the invariant mass of the two jets is required to be close to that of the W and Z bosons. Moreover, the DYMVA tagger is employed to reduce the *DY* background. The signal is extracted with a fit on the number of events in this phase space. Finally, two CRs are defined, one for the top quark process, and one for the non-resonant WW production.

- **$WHSS$**: this channel targets the $WH \to 2\ell 2\nu qq$ final state, where the two leptons are required to have the Same-Sign (SS) in order to reduce Drell-Yan background. The final state therefore contains two same-sign leptons, E_T^{miss}, and at least one jet.
  Events in the SR are categorized based on the number of jets with $p_T >$ 30 GeV: $N_{jet} = 1$ and $N_{jet} > 1$. For events containing more than two jets, only the leading and trailing jets are considered for the analysis. These jets are required to have an invariant mass $m_{jj} < 100\,\text{GeV}$.
  The N_{jet} categories are further divided into $e\mu$ and $\mu\mu$ sub-categories. Events with two electrons are not considered, as this flavour category is less sensitive to signal. To improve discrimination between signal and background, the following variable is defined:

$$\tilde{m}_H = \sqrt{(p^{jj} + 2p^{\ell})^{\mu}(p^{jj} + 2p^{\ell})_{\mu}} \tag{7.3}$$

  This variable serves as a proxy for the Higgs boson mass. It is com-

puted as the invariant mass of the jet pair and twice the lepton four-momentum closest to the jet pair. The lepton four-momentum serves as a proxy for the neutrino. If an event in the $N_{jet} = 1$ categories contains a second jet with $20\,\text{GeV} < p_T < 30\,\text{GeV}$, that jet is included in the computation of this variable. Otherwise the four-momentum of the single jet is used. Events in all categories are required to have $\tilde{m}_H > 60\,\text{GeV}$. The signal extraction fit is performed with a simultaneous fit on binned templates of the $\tilde{m}_H$ variable.
The main backgrounds in the $WHSS$ channel are WZ, $V\gamma^{(*)}$ processes and those related to non-prompt leptons. Two control regions for the WZ background are defined, one for each N_{jet} category.

- ***WH3ℓ***: this channel targets the $WH \to 3\ell 3\nu$ decay. The final state therefore contains three leptons and E_T^{miss}. The analysis selects events containing three leptons with $p_T > 25, 20, 15$ GeV respectively and total charge $Q = \pm 1$. Events with fourth lepton with $p_T > 10$ GeV or with a jet with $p_T > 30$ GeV are rejected. Events in the SR are categorised as follows: events with a Same-Sign SF lepton pair are placed in the SSSF category, while all other events are placed in the OSSF category. Leptons in the OSSF category are required to satisfy $|m_{\ell\ell} - m_Z| > 20$ GeV and $E_T^{miss} > 40$ GeV, in order to suppress events with a Z boson. The signal extraction is performed with a simultaneous fit on binned templates of a discriminant built with a BDT.
  The main backgrounds are WZ, ZZ and $V\gamma^{(*)}$ processes, as well as backgrounds containing non-prompt leptons. Two control regions are defined, one for the WZ background (with $N_{jet} = 0$) and one for the $V\gamma^{(*)}$.

- ***ZH3ℓ***: this channel targets the $ZH \to 3\ell 3\nu qq$ decay. The final state therefore contains three leptons with total charge $Q = \pm 1$. Selected events must contain an opposite-sign SF lepton pair with the requirement $|m_{\ell\ell} - m_Z| < 20$ GeV. Events are categorized based on the number of jets with $p_T > 30$ GeV in the event: $N_{jet} = 1$ and $N_{jet} > 1$. The signal extraction is performed with a simultaneous fit on binned templates of m_T^H.
  The main backgrounds of the $ZH3\ell$ channel are WZ, ZZ and $V\gamma^{(*)}$ processes, as well as backgrounds containing non-prompt leptons. This channel shares the same WZ control regions of the $WHSS$ channel, one for each N_{jet} category.

- ***ZH4ℓ***: this channel targets the $ZH \to 4\ell 4\nu$ decay. The final state therefore contains four leptons and E_T^{miss}. Selected events must con-

tain four leptons with $p_T > 25, 15, 10, 10$ GeV respectively and total charge $Q = 0$. Events containing a fifth lepton with $p_T > 10$ GeV are rejected. The opposite-sign SF lepton pair with invariant mass closest to the Z boson mass is designated as the Z boson candidate, while the remaining lepton pair, coming from the Higgs boson, is referred to as the X candidate. The Z boson candidate mass is required to be within 15 GeV of the nominal Z boson mass.
Events in the signal region are categorized based on the flavour of the lepton pair forming the X candidate. Events in the XSF category have a SF X lepton pair, while events in the XDF category have a DF X lepton pair. The signal extraction is performed with a simultaneous fit on binned templates of a discriminant built with a BDT.
Production of ZZ pairs is the main background in this channel. Two ZZ control regions are defined, one for each X lepton flavour category.

7.4.5 The STXS Measurement

The analysis described above (referred to as Standard Analysis in the following) targets the measurements of inclusive production cross sections. Another analysis, aiming to the measurement of the Simplified Template Cross Sections (STXS) framework [97], is performed.
The STXS framework aims to separate more cleanly measurement and interpretation steps in order to reduce the theory dependencies that are folded into the measurements (including the dependence on the underlying physics model). In addition, they provide more finely-grained measurements (and hence more information for theoretical interpretations). The framework is designed to be inclusive over the Higgs decays, allowing one to perform a global combination of all decay channels. Hence, the primary goals of the STXS framework are to maximize the sensitivity of the measurements while at the same time to minimize their theory dependence.
In the STXS framework the cross sections of different Higgs boson production mechanisms are measured in mutually exclusive regions of generator level phase space, that is without detector effects, and are referred to as STXS bins. These bins are designed to enhance sensitivity to possible deviations from the SM.
The measurement is done by defining a set of analysis categories that match the STXS bin definitions as closely as possible, and results are then unfolded to generator level. The full set of the STXS bins is reported in Figure 7.6 for the ggH production and in Figure 7.7 for the VBF production. The $VH2j$ channel is included in the VBF phase space, in the STXS bin with the requirement 60 GeV $< m_{jj} <$ 120 GeV.

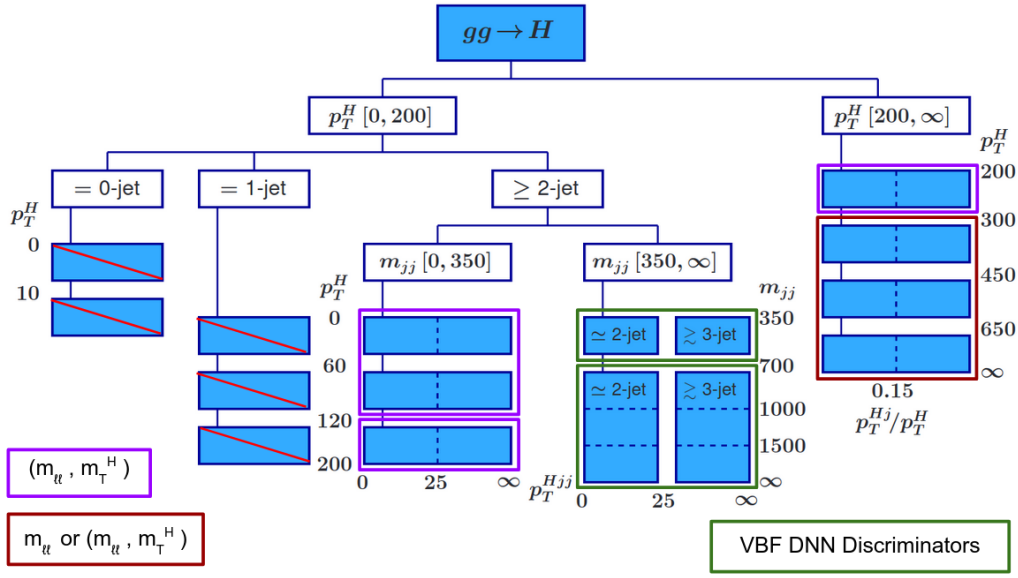


Figure 7.6: STXS bins for the *ggH* production. Groups of merged bins are highlighted with a colour scheme indicating different variables employed in the signal extraction fit. Barred STXS bins are non included in the analysis.

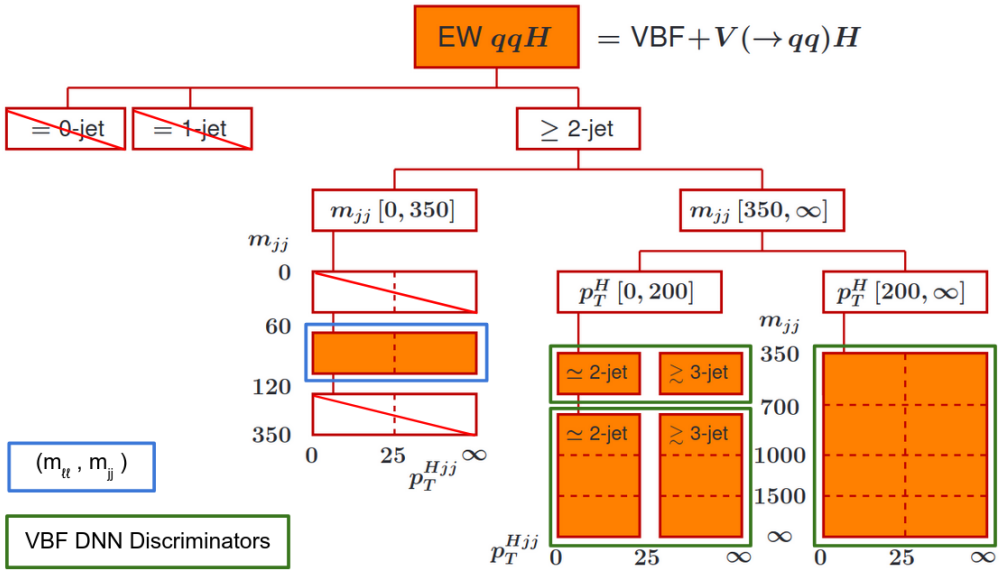


Figure 7.7: STXS bins for the *VBF* production. Groups of merged bins are highlighted with a colour scheme indicating different variables employed in the signal extraction fit. Barred STXS bins are non included in the analysis.

Analysis Category	**Target STXS Bins**
$N_{jet} > 1$, 350 GeV $< m_{jj} <$ 700 GeV, $p_T^H < 200$ GeV The *VBF*-DNN is employed to split this SR in the *j-like* sub-categories	$\sigma_{VBF}^{[350\text{ GeV}<m_{jj}<700\text{ GeV}, p_T^H<200\text{ GeV}]}$ $\sigma_{ggH}^{[350\text{ GeV}<m_{jj}<700\text{ GeV}, p_T^H<200\text{ GeV}]}$
$N_{jet} > 1$, $m_{jj} > 700$ GeV, $p_T^H < 200$ GeV The *VBF*-DNN is employed to split this SR in the *j-like* sub-categories	$\sigma_{VBF}^{[m_{jj}>700\text{ GeV}, p_T^H<200\text{ GeV}]}$ $\sigma_{ggH}^{[m_{jj}>700\text{ GeV}, p_T^H<200\text{ GeV}]}$
$N_{jet} > 1$, $m_{jj} > 350$ GeV, $p_T^H > 200$ GeV The *VBF*-DNN is employed to split this SR in the *j-like* sub-categories	$\sigma_{VBF}^{[m_{jj}>350\text{ GeV}, p_T^H>200\text{ GeV}]}$
$N_{jet} > 1$, $m_{jj} < 350$ GeV, $p_T^H < 120$ GeV (65 GeV $< m_{jj} <$ 105 GeV excluded)	$\sigma_{ggH}^{[m_{jj}<350\text{ GeV}, p_T^H<120\text{ GeV}]}$
120 GeV $< p_T^H <$ 200 GeV, $N_{jet} > 1$, $m_{jj} < 350$ GeV (65 GeV $< m_{jj} <$ 105 GeV excluded)	$\sigma_{ggH}^{[m_{jj}<350\text{ GeV}, 120\text{ GeV}<p_T^H<200\text{ GeV}]}$
$N_{jet} > 1$, 65 GeV $< m_{jj} <$ 105 GeV	$\sigma_{VH2j}^{[60\text{ GeV}<m_{jj}<120\text{ GeV}]}$

Table 7.3: Selection of the analysis categories and target STXS bins.

Analysis Category	**Target STXS Bins**
$200\text{ GeV} < p_T^H < 300\text{ GeV}$, $m_{jj} < 350\text{ GeV}$ if $N_{jet} > 1$ ($65\text{ GeV} < m_{jj} < 105\text{ GeV}$ excluded)	$\sigma_{ggH}^{[200\text{ GeV}<p_T^H<300\text{ GeV}]}$
$p_T^H > 300\text{ GeV}$, $m_{jj} < 350\text{ GeV}$ if $N_{jet} > 1$ ($65\text{ GeV} < m_{jj} < 105\text{ GeV}$ excluded)	$\sigma_{ggH}^{[p_T^H>300\text{ GeV}]}$

Table 7.4: Selection of the analysis categories and target STXS bins.

Given the statistical power of the LHC Run-2 dataset, sensitivity to some of the STXS bins is limited. Some bins are therefore measured in unison, by fixing the corresponding cross section ratios to the value predicted by the SM. This procedure is referred to as bin merging. Some STXS bins were excluded given the very low expected sensitivity (for example, for the *VBF* process only categories with at least two jets have been considered). Groups of STXS bins merged with this procedure are highlighted in Figure 7.6 and in Figure 7.7.

The selections used in the STXS measurements are similar to the ones described for the Standard Analysis, but adapted to the STXS bins. The analysis is performed only for DF leptons in the final state. Two analysis categories are in common between *VBF* and *ggH* STXS bins, in the region[6] $m_{jj} > 350\,\text{GeV}$ and $p_T^H < 200$ GeV. The *VBF*-DNN, employed for the DF *VBF* channel in the Standard Analysis, is employed for these two categories, taking advantage of the fact that the DNN can discriminate *VBF* from *ggH* productions. The signal extraction fit is performed on a two-dimensional $(m_{\ell\ell}, m_{jj})$ template in the category targeting the $VH2j$ channel, while either $m_{\ell\ell}$ or $(m_{\ell\ell}, m_T^H)$ templates are used in the remaining categories, depending on the number of expected events. Figure 7.6 and Figure 7.7 also show the variables employed in the signal extraction fit.

Table 7.3 and Table 7.4 show the selections in each of the defined analysis

[6]The reconstructed Higgs boson transverse momentum is given by the transverse momentum of the W bosons pair.

categories, and the corresponding STXS bins they target[7]. Additional cuts on m_{jj} are applied in order to guarantee the orthogonality of the defined categories. Control regions, similar to the ones described for the Standard Analysis, are also employed in the STXS Analysis.

7.5 Deep Neural Networks

The *VBF*-DNN, employed for the DF *VBF* channel of the Standard Analysis, as well as for the STXS Analysis, is built with Keras [98] using TensorFlow [99] back-end. In the following, a brief introduction to feed-forward deep neural networks is presented, while in Section 7.5.1 details regarding the *VBF*-DNN are reported.
The building block of neural networks is the perceptron, which is a simple algorithm that, given an input vector $\boldsymbol{x}$ of n values (referred to as "features"), outputs either one (yes) or zero (no):

$$f(\boldsymbol{x}) = \begin{cases} 1 & \text{if} \quad \sum_{j=1}^{n} w_j x_j + b > 0 \\ 0 & \text{otherwise} \end{cases} \tag{7.4}$$

where $\boldsymbol{w}$ is a vector of real-valued weights and b is the bias. By changing $\boldsymbol{w}$ and b parameters, the $f(\boldsymbol{x})$ output is changed as well.
Adding more layers between the inputs and the output results in a multi-layer perceptron, as shown in Figure 7.8. These additional layers are usually referred to as hidden layers. Each node (or neuron) in the first layer receives the input and fires according to the local $\boldsymbol{w}$ and b values. The output of the first layer is then passed to the second layer, the result of which is passed to the next layer and so on, until the output layer, consisting of a single neuron, is reached, The DNN built in this way is called dense, since each neuron in a layer is connected to all neurons located in the previous layer and to all the the neurons in the following layer. Layers are the building blocks of a neutral network, acting as a filter for data. More precisely, layers extract representations of data, hopefully more meaningful for the problem at hand. If the network has more than one output, the DNN performs a multi-classification: an input event can be classified in two or more output categories.
The output of each neuron is usually weighted with an activation function g: this gives the neuron the non-linear property necessary to model complex

[7]For this Thesis, a subset of the STXS bins targeted in the complete CMS analysis was considered. In the complete analysis other categories are defined, targeting ggH productions with $N_{jet} = 0$, $N_{jet} = 1$ and VH leptonic productions. SF channels are also included for *VBF*, ggH and $VH2j$ productions.

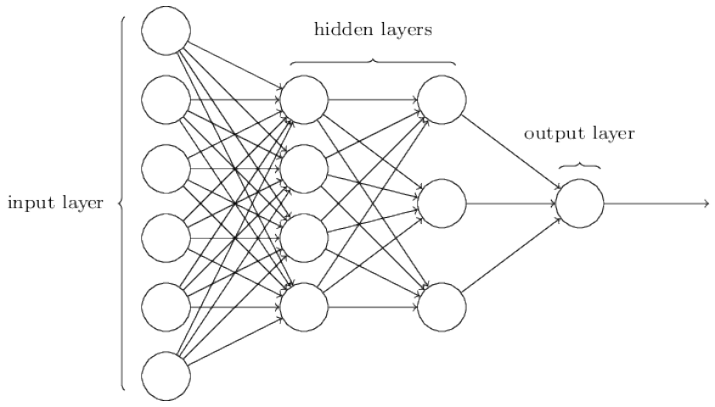


Figure 7.8: Schematic view of a multilayer perceptron.

datasets. The most common activation function employed in DNNs is the Rectifier Linear Unit (ReLU): $\mathrm{ReLU}(x) = \mathrm{Max}(x, 0)$.
The output A of each neuron in a multilayer perceptron can be written as:

$$A_i^k = g\left(\sum_j w_{ij}^k A_j^{k-1} + b_i^k\right) \tag{7.5}$$

where i indicates a neuron in a layer, k indicates a layer, j runs on the neurons of the previous layer $(k-1)$ and g is the activation function. The output nodes usually feature a different activation function. For instance, in multi-classification problems, the softmax function is employed:

$$\mathcal{O}_i = \frac{e^{\tilde{\mathcal{O}}_i}}{\sum_j e^{\tilde{\mathcal{O}}_j}} \tag{7.6}$$

where $\tilde{\mathcal{O}}_i$ are the output nodes before the transformation. The softmax function normalises the output nodes, $\sum_i \mathcal{O}_i = 1$: in this way they can be interpreted as probabilities for an input event to belong to one of the output categories.
The $\boldsymbol{w}$ and $\boldsymbol{b}$ tensors of the network layers are the trainable parameters of the DNN. Initially, the weights are filled with small random values (a step called random initialization). The network gradually adjusts these weights, based on a feedback signal: this gradual adjustment is the training process of the DNN, which gives it predictive power.
The first step of the training procedure is to prepare a batch of training events (with each event x having n features) with the corresponding targets y_i. The targets indicate to which of the classification categories the x event belongs:

with a DNN with four outputs, for instance, if the event x belongs to the category corresponding to the third output, the target will be: $\mathbf{y} = (0, 0, 1, 0)$.
The second step is to run the DNN on the training batch: this gives predictions $\mathcal{O}_i$ for every event x. Then, the loss of the network is computed: the loss function is a measure of the mismatch between between $\mathcal{O}_i$ and y_i. The most used loss function for multi-classification problems is the categorical cross-entropy:

$$E = -\sum_{i=1}^{m} \mathsf{y}_i \log(\mathcal{O}_i) \tag{7.7}$$

where $\mathcal{O}_i$ are the output nodes normalised with the softmax transformation, m is the number of output nodes and y_i is either zero or one.
The weights of the network are then updated in a way that slightly reduces the loss on the training batch. This is performed by computing the gradient of the loss function, with regard to the network weights. The weights are moved in the opposite direction from the gradient, thus decreasing the loss function. The most simple optimisation algorithm that performs this operation is the batch gradient descent:

$$\begin{aligned} \boldsymbol{w}_{n+1} &= \boldsymbol{w}_n - \xi \nabla_{\boldsymbol{w}} E \\ \boldsymbol{b}_{n+1} &= \boldsymbol{b}_n - \xi \nabla_{\boldsymbol{b}} E \end{aligned} \tag{7.8}$$

where ξ is the learning rate, i.e. how much the DNN parameters are changed at each step. The batch gradient descent algorithm updates the model parameters according to the gradient computed on the whole training batch. Another possibility is the stochastic gradient descent, that updates the model parameters according to the gradient computed from a single event: in this case each update is more accurate, but far more computationally expensive. The mini-batch gradient descent is a good compromise, using small subsets of the training batch and basing updates to parameters on the average gradient over the events in the subset.
In order to escape plateaus of the loss function, more sophisticated approaches are available. For instance, since the mini-batch approach operates on a small subset of data, information from other recently seen mini-batches can contribute to the current parameter update. This is the momentum method, in which updates on the parameters $\Delta\boldsymbol{w}$ are also influenced by the previous update $\Delta\boldsymbol{w}_{old}$:

$$\Delta\boldsymbol{w} = -\xi \nabla_{\boldsymbol{w}} E - \alpha \Delta\boldsymbol{w}_{old} \tag{7.9}$$

where $\alpha \in [0, 1]$. Another issue with the standard gradient descent algorithm is the fixed learning rate: if the learning rate is too small the convergence

might be slow, while if it is too large the algorithm might not converge on a minimum. In order to solve this issue, optimisation algorithms capable of adapting the learning rate during the training have been developed. The most used optimisation algorithm is ADAptive Momentum estimation (ADAM) [100], which employs the momentum method as well as an adaptive learning rate.
The operation of evaluating the gradient across the network layers is called back-propagation, which is based on the chain rule of calculus. Frameworks like TensorFlow [99] are capable of symbolic differentiation: given a chain of operations with a known derivative, they can compute a gradient function for the chain that maps network parameter values to gradient values.
Each iteration over the whole training batch is called epoch. Depending on the problem to be solved, hundreds of epochs are necessary in order to converge to the minimum of the loss function. The metric is a measure of how well the network is doing to solve the problem. The metric function is constantly monitored after each epoch: in multi-classification problems the categorical accuracy is usually employed. An event is considered successfully classified if the highest DNN output corresponds to the true category of the training event. The categorical accuracy is given by the ratio of successfully classified events and the total number of analysed events. It should be noted that the results from evaluating a metric are not used to train the network.
Processing the training batch for too many epochs may result in the so called over-training, that is learning representations that are specific to the training batch and do not generalize to data outside of it. In order to prevent over-training, a second batch of data is prepared, the validation batch. The loss function and the metric function are evaluated on both the training batch and the validation batch after each epoch. For instance, the loss evaluated on the training batch will continue to decrease as the training proceeds, but the loss evaluated on the validation batch at some point will start increasing: this is the onset of over-training.
To prevent this situation, the early-stopping technique is usually employed: if the validation loss keeps increasing for N epochs, the training is stopped, and the configuration for which the validation loss was minimum is saved. The choice of waiting at N epochs before stopping the training is made in order to escape local minima of the loss function. The metric function can be used with the same principle: the metric evaluated on the training batch will continue to increase as the training proceeds, but the metric evaluated on the validation batch at some point will start decreasing.
Another method of preventing the over-training is the use of drop-out layers between the the network hidden layers. A drop-out layer randomly sets input units to zero with a certain frequency at each step during training time. This

reduces the degree to which hidden units co-adapt and thus mitigates over-training.

7.5.1 The *VBF*-DNN

The *VBF*-DNN was trained on simulated events, using Monte Carlo samples[8] of *VBF*, top quark, non-resonant WW and *ggH* processes, passing the SR selection of the DF *VBF* channel. Each of the four samples contain about 15000 simulated events after the selection. Half of the WW sample is composed by *VBS* events, which features a topology similar to the VBF signal. Each of these samples is split into three sub-samples: 70% of the events are used for training, 20% are used for validation during the training, and 10% are used for testing the performance of the DNN after the training. The training, validation and testing events from the four samples are then merged together and shuffled in the training batch, validation batch and testing batch respectively.
The DNN is made by three hidden layers of 240, 120 and 60 neurons each and drop-out layers between them, in order to reduce over-training effects. The number of hidden layer and the number of neurons per layer were optimised for this analysis. The training aims to minimize the categorical cross-entropy loss function, while the categorical accuracy metric is monitored during the training.
The training is performed over a maximum of 500 epochs. Each epoch is performed on mini-batches of 500 events, and the ADAM optimiser is used. In order to minimise over-training effects, the early-stopping technique is adopted: if the the categorical cross-entropy on the validation batch keeps increasing after 100 epochs, the training is stopped.
The training used in this analysis lasts about 250 epochs, as can be seen in Figure 7.9, where both the loss function and the accuracy are reported as a function of the training epoch, for the training and validation batches. The categorical accuracy reaches a value of about 65%.
It can be noted that for the first 50 epochs the validation loss is actually lower than the training loss: this is the effect of the drop-out layer that forces the network to not learn representations which are too close to the training batch samples.
The following kinematic variables are used as input features of the DNN:

- $|\Delta\eta_{jj}|$, m_{jj} : as shown in Section 6.4.1, these two variables are among the few that allow to discriminate *VBF* from *ggH* production mecha-

[8]The Monte Carlo samples employed to train the *VBF*-DNN are independent from those used in the analysis.

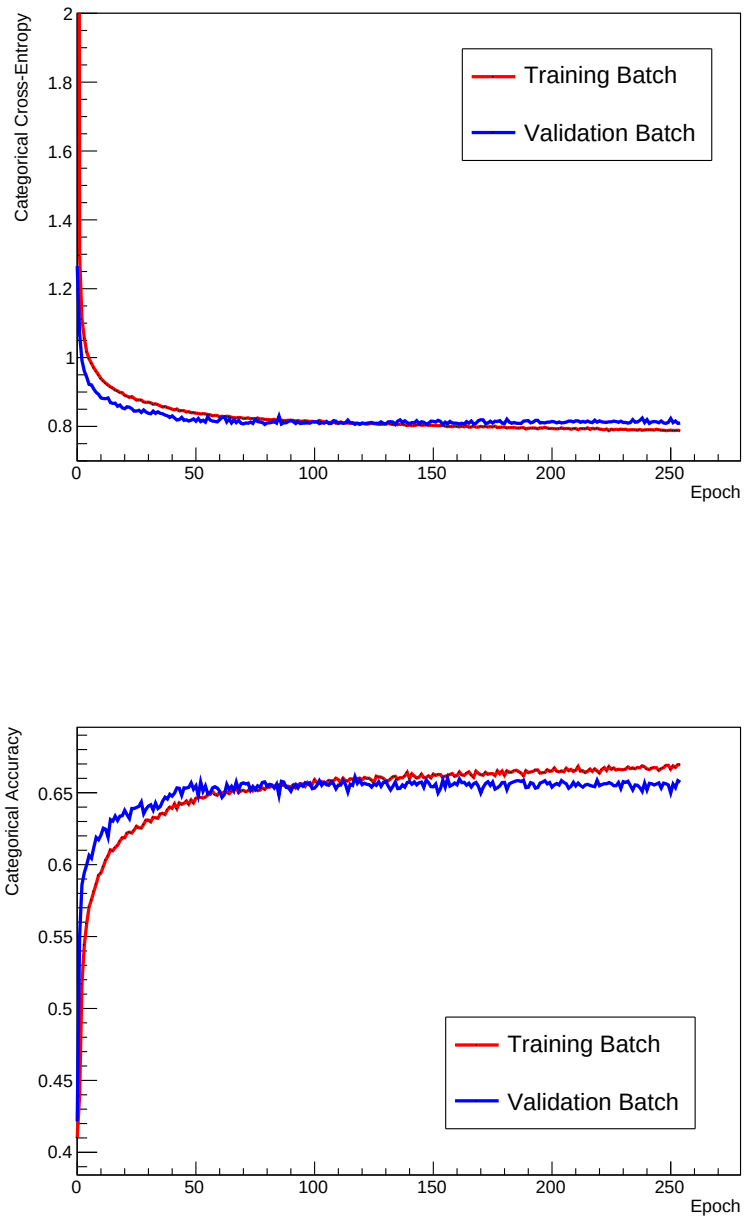


Figure 7.9: Categorical cross-entropy loss function (top) and categorical accuracy metric (bottom) as a function of the training epoch.

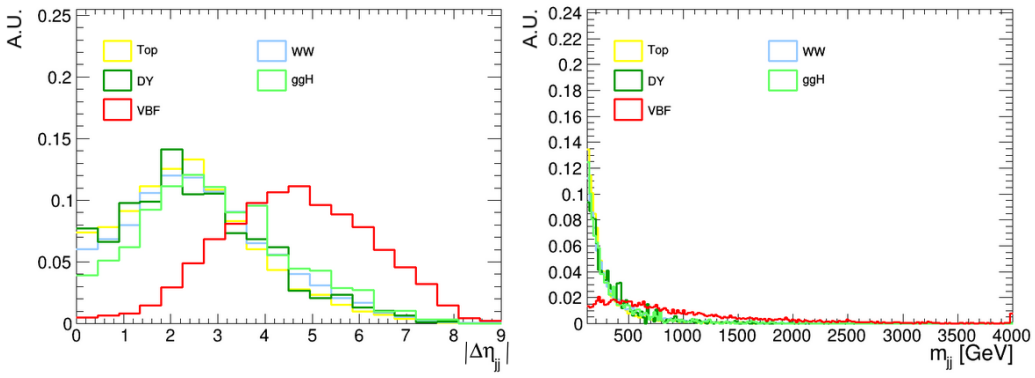


Figure 7.10: Normalised $|\Delta\eta_{jj}|$ (left) and m_{jj} (right) distributions for the *VBF* and *ggH* Higgs production mechanisms and top quark, non-resonant *WW* and $DY \to \tau\tau$ processes in the SR of the DF *VBF* channel.

nisms, being related to the topology of the two jets. For instance, the *VBF* signal shows a larger pseudorapidity gap with respect to other process. Figure 7.10 shows the normalised distributions of signal and background processes (including $DY \to \tau\tau$, although not trained by the DNN) in the SR of the DF *VBF* channel.
As a consequence of the pseudorapidity gap, also m_{jj} has good discriminating power: at high m_{jj} values the *VBF* signal is the dominant process, as shown in Figure 7.10.

- $m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$: leptons coming from a $H \to WW$ decay have nearly collinear spatial momenta, due to the spin correlation. Therefore, $m_{\ell\ell}$ is peaked at low values for *VBF* and *ggH* mechanisms and it has a broadened shape for the other backgrounds, as shown in Figure 7.11. For the same reason, processes involving the production of a Higgs boson, the angular separation in ϕ between the two leptons is limited, and it is almost flat for events produced without a resonance, as shown in Figure 7.11.

- $p_T^{\ell\ell}$, $\Delta R_{\ell\ell}$: the di-lepton momentum is not expected to have a high discriminating power, while $\Delta R_{\ell\ell}$, similarly to $\Delta\phi_{\ell\ell}$, can discriminate Higgs boson events from non-resonant backgrounds.

- $p_T^{j_1}$, $p_T^{j_2}$: the transverse momenta of leading and trailing jets are not expected to have a high discrimination power, but are added to give a complete description of the di-jet system.

- η^{j_1}, η^{j_2} : the leading jet coming from the *VBF* mechanism is usually

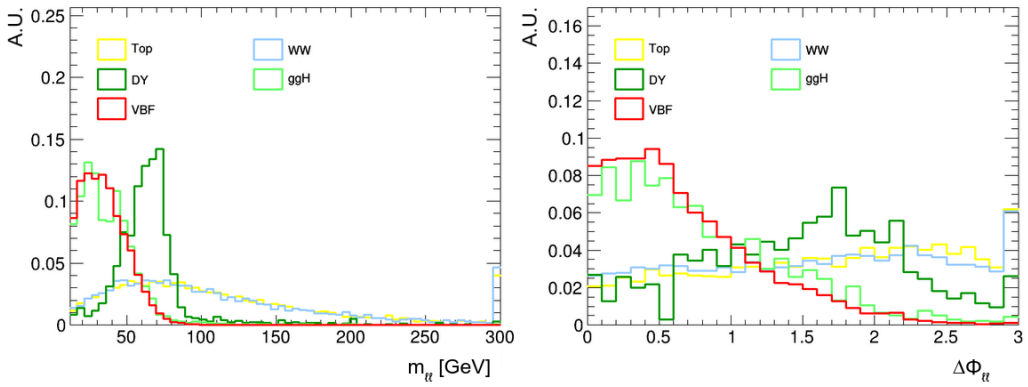


Figure 7.11: Normalised $m_{\ell\ell}$ (left) and $\Delta\phi_{\ell\ell}$ (right) distributions for the *VBF* and *ggH* Higgs production mechanisms and top quark, non-resonant *WW* and $DY \to \tau\tau$ processes in the SR of the DF *VBF* channel.

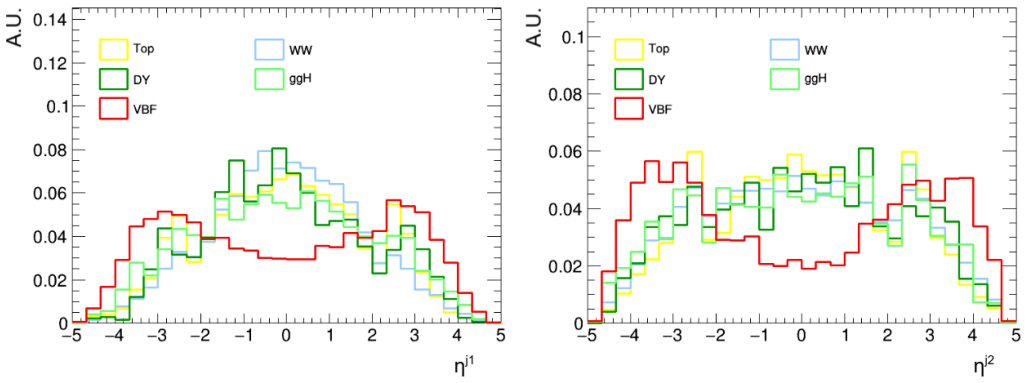


Figure 7.12: Normalised η^{j_1} (left) and η^{j_2} (right) distributions for the *VBF* and *ggH* Higgs production mechanisms and top quark, non-resonant *WW* and $DY \to \tau\tau$ processes in the SR of the DF *VBF* channel.

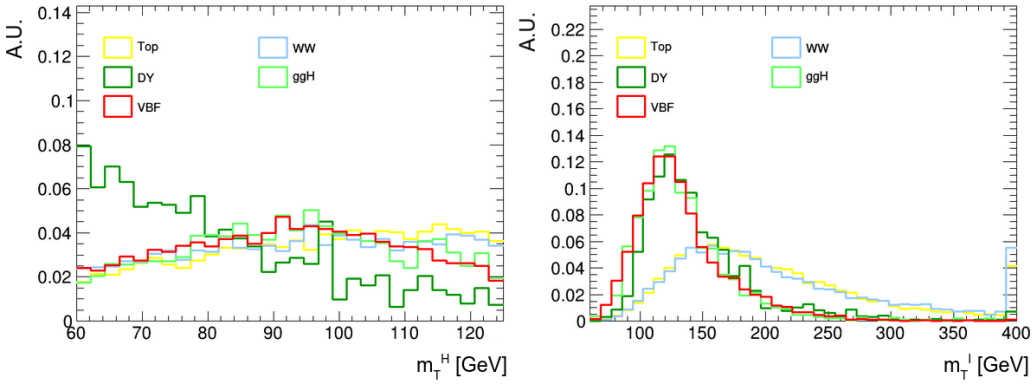


Figure 7.13: Normalised m_T^H (left) and m_T^I (right) distributions for the *VBF* and *ggH* Higgs production mechanisms and top quark, non-resonant *WW* and $DY \to \tau\tau$ processes in the SR of the DF *VBF* channel.

emitted at large $|\eta|$ values, being generated by the hadronisation of a quark that has taken a high fraction of the proton four-momenta. This is a peculiar characteristic of the signal, while backgrounds are more likely to present a leading jet emitted in the central region, as shown in Figure 7.12.
The trailing jet associated to the signal is highly correlated to the leading one, since the two jets are separated by a large pseudorapidity gap. Indeed, for the *VBF* mechanism, the η^{j_2} distribution is similar to the η^{j_1} distributions, while for background processes η^{j_2} is more homogeneous, as shown in Figure 7.12.

- $p_T^{\ell_1}$, $p_T^{\ell_2}$, η^{ℓ_1}, η^{ℓ_2} : the transverse momenta of leading and trailing lepton, as well as the respective pseudorapidities, are not expected to have a high discrimination power, but are added to give a complete description of the di-lepton system.

- $m_{\ell_1 j_1}$, $m_{\ell_2 j_1}$, $m_{\ell_1 j_2}$, $m_{\ell_2 j_2}$: the invariant mass of the system composed by the leading or trailing lepton and leading or trailing. Four combinations are possible: the discriminating power is low, but not completely negligible.

- E_T^{miss}, $\Delta\phi(p_T^{\ell\ell}, E_T^{miss})$: MET informations are also included in the training. While E_T^{miss} does not contribute significantly (since all the trained events have neutrinos in the final state), $\Delta\phi(p_T^{\ell\ell}, E_T^{miss})$ has a good discriminating power. It is also included in the the definition of m_T^H.

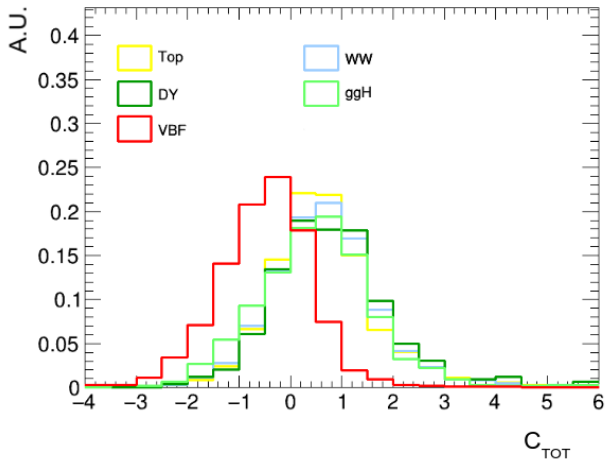


Figure 7.14: Normalised C_{TOT} distributions for the *VBF* and *ggH* Higgs production mechanisms and top quark, non-resonant WW and $DY \to \tau\tau$ processes in the SR of the DF *VBF* channel.

- m_T^H, m_T^I : both m_T^H and the so-called improved transverse mass m_T^I defined as:

$$m_T^I = \sqrt{(p^{\ell\ell} + E_T^{miss})^2 - (\boldsymbol{p}_T^{\ell\ell} + \boldsymbol{p}_T^{miss})^2} \tag{7.10}$$

  tend to have a different shape for events containing the production of an Higgs boson, as shown in Figure 7.13.

- C_{TOT}, H_T : the centrality C_{TOT} represents a measure of how much the charged leptons are emitted centrally with respect to the di-jet system. It is defined as:

$$C_{TOT} = \log\left(\sum_{\ell} |2\eta^{\ell} - \sum_{j} \eta^{j}| \,/\, |\Delta\eta_{jj}|\right) \tag{7.11}$$

  where ℓ runs on leading and trailing leptons and j on leading and trailing jets. Distributions of this variable are shown in Figure 7.14: the *VBF* signal is significantly different with respect to the backgrounds. The H_T variable is calculated as scalar sum of the transverse momenta of all jets in the event. It gives a measure of the hadronic activity of the event, as is expected to be a bit different between the trained processes.

- qgl_{j_1}, qgl_{j_2} : the so quark-gluon likelihood discriminant *qgl* [101] is constructed with a multivariate technique and it gives a measure of the

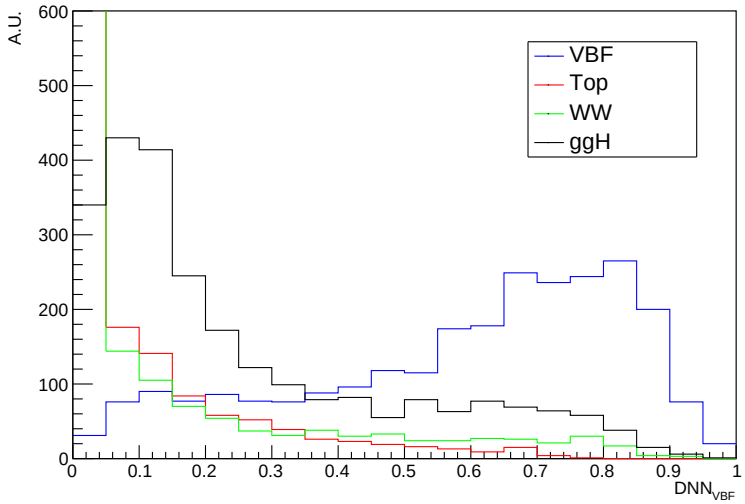


Figure 7.15: DNN_{VBF} classifier distribution on the training batch.

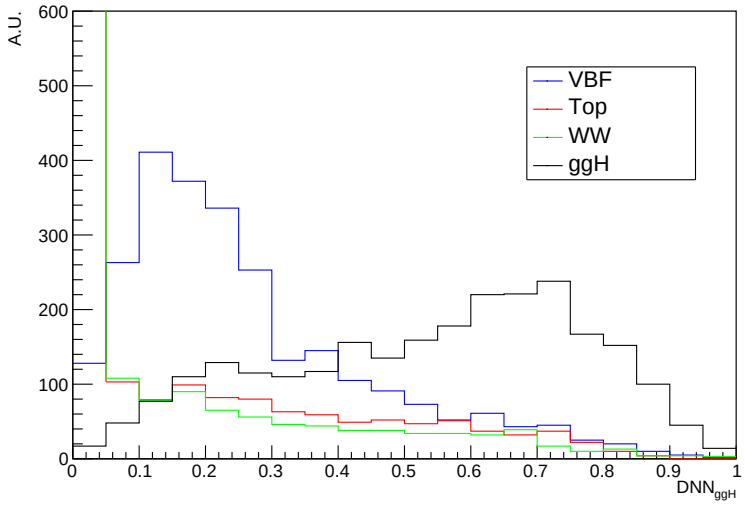


Figure 7.16: DNN_{ggH} classifier distribution on the training batch.

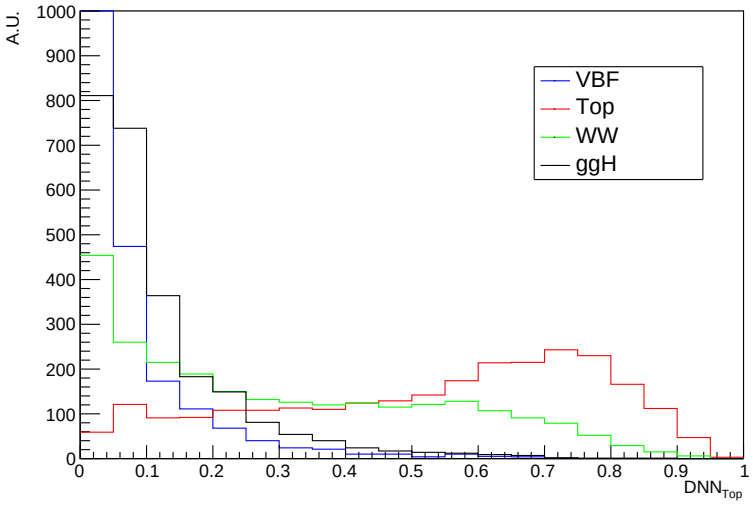


Figure 7.17: DNN_{Top} classifier distribution on the training batch.

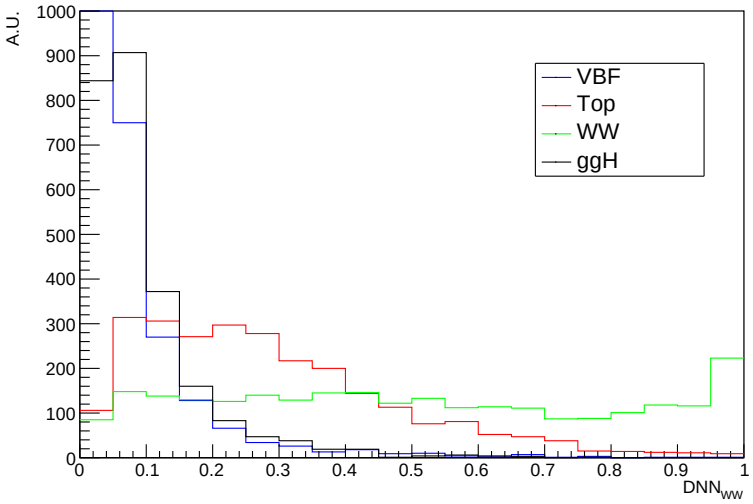


Figure 7.18: DNN_{WW} classifier distribution on the training batch.

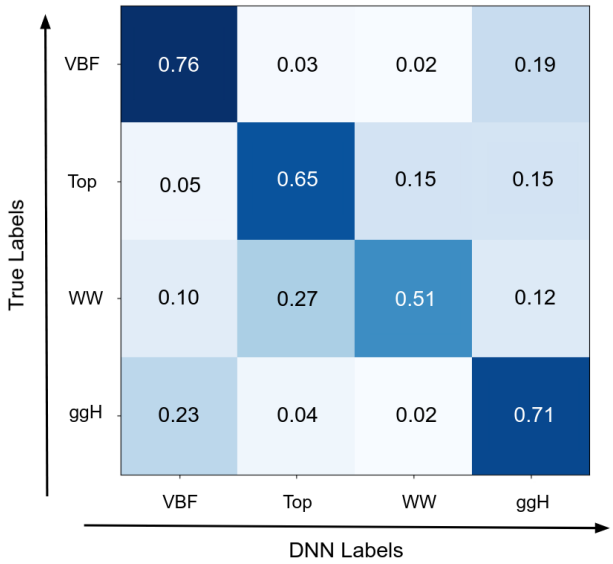


Figure 7.19: Confusion matrix of the *VBF*-DNN, evaluated on the testing batch.

probability that the jet has been emitted either by a light quark or a gluon: the discriminator tends to assume values close to one for quark-like jets and close to zero for gluon-like jets. For this reason, *qgl* is very useful to distinguish the *VBF* mechanism from other processes, since background events present jets coming from ISR or FSR, which are usually induced by gluons, while the signal has always jets induced from quarks.

Even if some of the input variable do not show evident discrimination power, the DNN is capable of finding non-trivial correlations that can help the DNN to discriminate signal from background events.
The DNN outputs (i.e. the four classifiers DNN_{VBF}, DNN_{ggH}, DNN_{WW}, DNN_{Top}) are normalised with the softmax transformation. The DNN model has been applied to the testing sample, in order check the goodness of the training. In Figures 7.15, 7.16, 7.17, 7.18 the distributions of the four classifiers DNN_{VBF}, DNN_{ggH}, DNN_{WW}, DNN_{Top} are respectively reported, for *VBF*, *ggH*, top quark and WW events. The shape of the four classifiers shows the effectiveness of the discrimination strategy. For instance, the *VBF* events are peaked for high values of the DNN_{VBF} classifier, while the three backgrounds are peaked at lower values.
Finally, Figure 7.19 shows the confusion matrix of the DNN, evaluated on the testing batch. The y axis reports the true labels, i.e. to what Monte Carlo

sample the processed event belongs. The x axis reports the DNN labels: for a given event, if the DNN_{VBF} is the highest discriminator, the event is labelled as VBF by the DNN. This is the same approach used for defining sub-categories in the analysis. For instance, VBF events are successfully classified 76% of the times, while 19% of the times are wrongly classified as ggH events.
Top quark and WW events show a higher confusion, i.e. it is easy for a top quark event to be mis-classified as a WW event. For this reason, during the development of this analysis, top quark and WW events were merged in order to obtain a single DNN classifier for both processes. However, this approach had a worse sensitivity with respect to the one in which the two processes are separated. Even if limited, the distinction between top quark and WW backgrounds, and the subsequent SR sub-categories, improves the VBF signal sensitivity.

7.6 Background Estimation

In the following, additional details regarding the estimation of the main backgrounds of the analysis are reported, with a focus on the most important backgrounds for the DF VBF channel: non-prompt leptons, top quark, non-resonant WW and $DY \to \tau\tau$ backgrounds.

- **Top Quark:** reweighting of the top and anti-top quark p_T spectrum is performed for the $t\bar{t}$ simulation in order to match the NNLO and Next-to-Next-to-Leading Logarithm (NNLL)[9] QCD predictions, including also the NLO electroweak contribution [102].
  For ggH, VBF, and $VH2j$ channels (both DF and SF), in which the contribution of top quark background is dominant, the normalisation of the simulated templates is left unconstrained in the signal extraction fit separately for each channel. The normalisations in these phase spaces are therefore measured from the observed data, by constraining the free-floating normalisation parameters in the respective top quark control regions (the statistical treatment of this method is described

[9]When calculating an observable predicted by QCD in a perturbative way, the expansion in powers of α_s contains terms of the type $\alpha_s^n L^k$ ($k < 2n$) where $L = \ln(q_{cut}/s)$, with q_{cut} the cut on resolvable emission. With low q_{cut} values, the logarithm becomes large and the perturbative series diverges. It is therefore necessary to consider the terms that have a high value of the logarithm. The study of these terms is called resummation and is done by putting the terms together in the perturbation series according to their degree of divergence, starting from the Leading Logarithm (LL) therm.

in Section 8.1.1). For the *ggH* channels, independent normalisation parameters and control regions are used for each N_{jet} sub-category.

- **Non-Resonant WW**: in the non-resonant WW background, the quark induced WW simulated events are reweighted to match the di-boson p_T spectrum computed at NNLO + NNLL QCD accuracy [103, 104]. For *ggH*, *VBF*, and $VH2j$ channels (both DF and SF) the normalisations of the quark-induced and gluon-induced WW backgrounds are measured from the observed data, keeping a different parameter for each signal phase space as done for the top quark background. In the DF channels the normalisation parameters are constrained directly in the signal regions without the need of defining control regions, as the signal regions span the high $m_{\ell\ell}$ phase space that is enriched in WW events with a negligible Higgs boson signal contribution. In the SF channels, dedicated WW control regions are defined. The *VBS* background normalisation is instead fixed to the SM cross section provided by the MC simulation.

- **Drell-Yan:** In DF *ggH*, *VBF* and $VH2j$ channels the $DY \to \tau\tau$ background is estimated with an embedding data technique. First, $Z \to \mu\mu$ with well identified muons are selected in a data sample. In each event, the selected muons are removed and replaced with simulated τ leptons, keeping the same four-momentum of the initial muons. The data embedded sample is then corrected using scale factors related to the simulation of τ leptons.
  The usage of the data embedded sample allows for a better modelling of the observables that are sensitive to the detector response and calibration, such as E_T^{miss} and other variables related to the hadronic activity in the event. Since the data embedded sample takes into account all processes with a $\tau^+\tau^-$ pair decaying to either electrons or muons, all the simulated backgrounds that contain a $\tau^+\tau^-$ pair are not considered in the analysis to avoid any double counting. To correct for any additional discrepancy associated to the different acceptance of the $H \to WW$ signal phase space, the normalisation of the data embedded samples is left unconstrained in the fit as done for top quark and WW backgrounds. The free normalisation parameters are constrained separately for DF *ggH*, *VBF* and $VH2j$ channels in the respective $DY \to \tau\tau$ CRs.
  The data embedded samples cover the events that pass the $e\mu$ triggers, which represent the vast majority of the events selected in the DF final state. The remaining $Z \to \tau^+\tau^-$ events that enter the analysis phase space thanks to the single lepton triggers (about 5% of the total) are

estimated using MC simulation.
In SF ggH, VBF and $VH2j$ channels the dominant background contribution arises from $Z \to \ell^+\ell^-$ production and it is estimated using a data driven technique [71]. Finally, in $WHSS$, $WH3\ell$, $ZH3\ell$ and $ZH4\ell$ channels the Drell-Yan represents a minor background and is estimated using MC simulations.

- **Non-Prompt:** the lepton identification and isolation criteria are not sufficient to completely remove this background. The residual contribution is hence estimated from data. A control sample is defined using events in which one lepton passes the standard lepton identification and isolation criteria and another lepton candidate fails these criteria but passes a looser selection, resulting in a sample of "pass-fail" lepton pairs. The pass-fail sample is dominated by non-prompt leptons. The efficiency (ϵ_{misID}) for a jet that satisfies this looser selection to pass the standard selection is estimated directly from data in an independent sample dominated by events with non-prompt leptons from multi-jet processes. The contamination of prompt leptons from electroweak processes in such a sample is removed using the simulation. The efficiency ϵ_{misID} is parametrised as a function of the p_T and η of the leptons, and is used to weight the events in the pass-fail sample by $\epsilon_{misID}/(1-\epsilon_{misID})$, to obtain the estimated contribution from this background in the signal region. The contamination of prompt leptons in the pass-fail sample is corrected by using their probability to pass the standard selection given that they pass the looser selection, as measured in a Drell-Yan data control sample.

7.7 Systematic Uncertainties

The signal extraction procedure must take into account systematic uncertainties, which are represented as nuisance parameters in the fit, as explained in Section 8.1.1. The effect of each source of systematic uncertainty is either a change of the normalisation of a given signal or background process, a change of the template shape, or both. Uncertainties correlated between different year datasets are represented by a single common nuisance parameter, while uncorrelated uncertainties are represented by independent parameters. The systematic uncertainties arise either from an experimental or a theoretical source.

Experimental Uncertainties

The following experimental uncertainties are included in the signal extraction fit:

- The uncertainty in the integrated luminosity amounts to 2.5%, 2.3% and 2.5% for 2016, 2017, and 2018 datasets respectively [105, 106, 107]. This uncertainty is partially correlated among the three datasets (considering the luminosity measurement scheme), and is applied to all samples that are purely based on simulations (i.e. it is not considered for processes that are measured using data).

- The uncertainties in the trigger efficiency and lepton reconstruction and identification efficiencies are measured in bins of the lepton p_T and η, independently for electrons and muons. The impacts on template normalisations from the uncertainties in the trigger efficiency are less than 1%, while the uncertainties in the reconstruction and identification efficiency cause shape and normalisation changes of about 1% for electrons and about 2% for muons. These uncertainties are dominated by the statistical fluctuations of the dataset where they are measured, and are thus kept uncorrelated among the datasets.

- The uncertainties in the determination of the lepton momentum scale and jet energy scale cause a migration of the simulated events inside or outside the analysis acceptance, as well as migrations across bins of signal and background templates. The impact of these sources in the template normalisations is $0.6 - 1.0\%$ for the electron momentum scale and 0.2% for the muon momentum scale. The main contribution to these uncertainties arise from the limited data sample used for their estimation. They are therefore treated as uncorrelated nuisance parameters among the three datasets.
  The jet energy scale uncertainty is modelled by implementing eleven independent nuisance parameters corresponding to different jet energy correction sources, six of which are correlated among the three datasets. Their effect vary in the range of $1 - 10\%$, according mainly to the jet multiplicity in the analysis phase space.

- The uncertainty in the jet energy resolution smearing applied to simulated samples causes both a normalisation and a shape change of the templates. This uncertainty has a minor impact in all the analysed channels (below 1%) and is uncorrelated among the three datasets.

- The E_T^{miss} resolution uncertainty includes the propagation of lepton and jet energy scale and resolution uncertainties to E_T^{miss}, as well as the uncertainties on the energy scales of particles that are not clustered into jets, and the uncertainty on the amount of energy coming from pile-up interactions.

- The uncertainty in the pile-up jet identification efficiency is modelled in bins of the jet p_T and η and considered for jets with $p_T < 50$ GeV, since pile-up jet identification techniques are only used for low p_T jets. This uncertainty cause both a normalisation and a shape change of the signal and background templates and is kept uncorrelated among the three datasets.

- The uncertainty in the b-tagging efficiency is modelled by implementing seventeen nuisance parameters, five of which are related to the theoretical uncertainties involved in the measurements and are therefore correlated among the three datasets. The remaining four parameters per dataset, which arise from the statistical accuracy of the efficiency measurement, are kept uncorrelated [69]. These uncertainties have an impact on both the shape of the templates and their normalisation for all the simulated samples.

- The uncertainties in the non-prompt lepton background estimation affect both the normalisation and the shape of the templates of this process. They arise from the limited size of the dataset used for the ϵ_{misID} measurement and from the difference in the flavour composition of jets mis-tagged as leptons between the pass-fail sample and the signal phase space. Both sources are implemented as uncorrelated nuisance parameters between electrons and muons, given the different mis-measurement probabilities for the two flavours, and are uncorrelated among the three datasets. Their effect vary between few percent to about 10% depending on the signal region.
  An additional normalisation uncertainty of 30% is assigned to cover any additional mis-modelling of the jet flavour composition. This uncertainty is correlated among the datasets, but uncorrelated among signal regions containing different lepton flavour combinations (since the main mechanism of non-prompt lepton production arises from different processes).

- The statistical uncertainty due to the limited number of simulated events is associated to each bin of the simulated signal and background templates.

Theoretical Uncertainties

Theoretical uncertainties have different sources, such as the employed PDF set, the α_s value and missing higher order corrections in the perturbative ME calculation. Template variations, both in shape and normalisation, associated to the aforementioned sources are treated as correlated nuisance parameters for the three datasets. The following theoretical uncertainties are included in the signal extraction fit:

- The uncertainties in the employed PDF set and the α_s choice are found to have a negligible effect in terms of variation of the shape of the simulated templates, therefore only the normalisation change is considered. These uncertainties are not considered for backgrounds whose normalisation is constrained using data in dedicated control regions. For the Higgs boson signal processes, these theoretical uncertainties are computed by the LHC Higgs Cross Section Working Group [62] for each production mechanism.

- The effect of missing higher-order corrections for the background processes is estimated by reweighting the MC simulation events to alternative event weights where the factorisation and renormalisation scales are varied up and down by a factor of two, and the envelopes of the varied templates are taken as one standard deviation variation. The extreme cases where one scale is varied by 0.5 and the other one by 2 are not considered. For backgrounds whose normalisation is constrained using data in dedicated control regions, only shape variations are considered.
  For the *ggH* signal sample the uncertainties are decomposed into several components [62]. For instance, they account for the overall cross section, and migrations of events among N_{jet} and p_T^H bins. A similar procedure is followed for the *VBF* signal sample, also including electroweak corrections to the production cross section.

- In order to estimate the uncertainty in the pile-up modelling, the total inelastic proton-proton cross section of 69.2 mb [108] is changed within a 5% uncertainty, corresponding to the uncertainty in the inelastic cross section measurement as well as the difference in the primary vertex reconstruction efficiency between data and simulation.

- Theoretical uncertainties due to the modelling of the PS and the underlying events are taken into account for all the simulated samples. The uncertainty in the PS modelling is evaluated by varying the PS weights computed by PYTHIA 8 on an event-by-event basis, keeping

the variations of the weights related to initial and final state radiation contributions uncorrelated. The uncertainty in the underlying events modelling is evaluated by shifting the nominal templates according to alternative MC simulations generated with a variation of the underlying event tune within its uncertainty.
The PS uncertainty affects the shape of the templates mainly through a migration of the events across jet multiplicity bins, while the underlying event uncertainty is found to have a negligible impact in the shape of the templates and the normalisation effect is of about 1.5%.

- Additional theoretical uncertainties on specific background processes are also taken into account. A 15% uncertainty is assigned to the relative fraction of the gluon-induced component in the non-resonant WW background process [89]. An uncertainty of 8% is assigned to the relative fraction of single top quark and $t\bar{t}$ processes.

- For the measurement of the signal cross sections in the STXS framework, the effect of theoretical uncertainties in the template normalisations is removed for signal processes in each STXS bin being measured. In cases where two or more STXS bins are measured together, the shape effect of theoretical uncertainties causing event migrations among the merged bins is kept. In addition, residual theoretical uncertainties arising from factorisation and renormalisation scales variations are taken into account to describe the shape variations that cause an acceptance effect of the signal templates within each STXS bin. The latter uncertainties are correlated among STXS bins that share a similar phase space definition. For the measurement of leptonic VH cross sections in STXS bins, the aforementioned theoretical uncertainties are found to have a negligible impact with respect to the measurement statistical accuracy and have been neglected.

Chapter 8

Experimental Results

8.1 Statistical Methodology

In the following, a summary of the statistical methodology employed in the analysis is reported.
The method of Maximum Likelihood (ML) is the standard statistical tool employed in particle physics [109, 110]. It allows to estimate a set of parameters $\boldsymbol{\theta} = (\theta_1, ..., \theta_m)$ from a finite sample of data. Given a random variable x distributed accorded to a probability density function (p.d.f) $f(x, \theta)$, the likelihood function is defined as:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i, \boldsymbol{\theta}) \tag{8.1}$$

where x is measured n times and x_i is the i-th measurement. The ML estimators $\hat{\boldsymbol{\theta}}$ of the true parameters $\boldsymbol{\theta}$ are found by maximising the likelihood function, hence they are given by the solutions to the equations:

$$\frac{\partial L}{\partial \hat{\theta}_i} = 0 \quad \text{with} \quad i = 1, ..., n \tag{8.2}$$

It is often useful to consider the Negative Log Likelihood (NLL) instead of the likelihood $L(\boldsymbol{\theta})$:

$$NLL = -\log(L(\boldsymbol{\theta})) \tag{8.3}$$

With the NLL, the maximization of a product becomes minimization of a sum, which is computationally more manageable.
It can be proven that the ML estimator satisfies the requirements for a good estimator. For instance it is asymptotically unbiased:

$$\lim_{n \to \infty} |E[\hat{\boldsymbol{\theta}}] - \boldsymbol{\theta}| = 0 \tag{8.4}$$

where $E[\hat{\boldsymbol{\theta}}]$ is the expectation value of the estimator. Moreover, still in the large sample limit, the estimator is gaussian distributed: this property is known as asymptotic normality. Finally, it is an efficient estimator: it has the minimum variance allowed for any estimators, at least in the large sample limit. Even for finite samples, if an efficient estimator exists, the ML method will find it.
Since in most cases in particle physics the p.d.f. are not known analytically, binned fits are usually performed. The p.d.f. are replaced by histograms of the random variable x, called templates, which represent signal and background contributions obtained from Monte Carlo simulations. In a single bin histogram, if N^{obs} events are observed, and s and b are the expected number of signal and background events respectively, N^{obs} is modelled as a Poisson variable distributed around the mean value $s + b$.
In order to check if the measured number of events is consistent with the SM prediction, it is useful to introduce the signal strength modifier μ, defined as the ration between the measured cross section and the SM expectation. In this case N^{obs} is modelled as a Poisson variable distributed around the mean value $\mu s + b$, and the likelihood function is given by:

$$L(\mu) = \frac{(\mu s + b)^{N^{obs}}}{N^{obs}!} e^{(-\mu s + b)} \tag{8.5}$$

By maximising $L(\mu)$, the ML estimator $\hat{\mu}$ for the true value μ is found.
In the case of a histogram with m bins, the likelihood function can be generalised:

$$L(\mu) = \prod_{i=1}^{m} \frac{(\mu s_i + b_i)^{N_i^{obs}}}{N_i^{obs}!} e^{(-\mu s_i + b_i)} = \prod_{i=1}^{m} \mathcal{P}(N_i^{obs}; \mu s_i + b_i) \tag{8.6}$$

where s^i and b^i are the expected number of signal and background events in bin i and N^i_{obs} is the observed number of events in bin i. The estimation of μ with this binned ML method is referred to as template fit.

8.1.1 Nuisance Parameters

As explained in Section 7.7, the effect of systematic uncertainties is a variation of either the normalisation or the shape of the templates (the latter are referred to as shape uncertainties). Systematic uncertainties are therefore included in the likelihood as additional parameters, which are determined through the same fit procedure. However, since their measurement is not the purpose of the analysis, they are referred to as nuisance parameters $\boldsymbol{\nu}$, while

the signal strength is referred to as the Parameters Of Interest (POI). The likelihood function can therefore be written as:

$$L(\mu, \boldsymbol{\nu}) = \prod_{i=1}^{m} \mathcal{P}(N_i^{obs}; \mu s_i(\boldsymbol{\nu}) + b_i(\boldsymbol{\nu}))\mathcal{N}(\boldsymbol{\nu}) \tag{8.7}$$

where $\mathcal{N}(\boldsymbol{\nu})$ is a constraint of the likelihood functions determined by the nuisances $\boldsymbol{\nu}$.
For instance, uncertainties in templates normalisations are treated by introducing a nuisance parameter ν, whose p.d.f. is a log-normal distribution:

$$f(\nu; \mu_\nu, \sigma_\nu) = \frac{1}{\nu\sigma_\nu\sqrt{2\pi}} e^{-\frac{(\ln\nu - \mu_\nu)^2}{2\sigma_\nu^2}} \tag{8.8}$$

where μ_ν is the expected normalisation and σ_ν is the estimated uncertainty on the normalisation. Log-normal distributions are chosen instead of normal distributions because they vanish when ν approaches zero. Since these parameter enters the fit as multiplicative factors for the MC templates, this feature avoids negative yields.
Shape uncertainties can not be estimated by a single a-priori distribution, as they alter the expected number of events bin by bin. For each process and phase space region in which a certain shape uncertainty is applied, two additional input templates have to be provided, corresponding to a variation of one standard deviation of the considered systematic error. In general, shape uncertainties also have an effect on the overall normalisation. This feature is to be avoided if the normalisation coincides with the quantity to be measured, for instance the Higgs boson signal template.
As explained in Section 7.6, the normalisations for some backgrounds are estimated directly from data. Nuisance parameters (associated with these processes) with a flat prior distributions are included in the likelihood function. For instance, in the DF *VBF*, *ggH* and $VH2j$ channels, the number of background events in the i-th bin is given by:

$$b_i(\boldsymbol{\alpha}) = \alpha^{Top} b_i^{Top} + \alpha^{WW} b_i^{WW} + \alpha^{DY} b_i^{DY} + b_i^{Other} \tag{8.9}$$

where $\alpha^{Top}, \alpha^{WW}, \alpha^{DY}$ are the nuisance parameters. In each bin i, N_i^{obs} is therefore modelled as a Poisson variable distributed around the mean value $\mu s_i + b_i(\boldsymbol{\alpha})$: $\boldsymbol{\alpha}$ is extracted from data, together with the POI. In the DF *VBF*, *ggH* and $VH2j$ channels, normalisations of top quark and $DY \to \tau\tau$ processes are constrained by including in the likelihood function the number of events in their respective CRs, as additional bins in the templates. For the non-resonant WW events, normalisations are instead evaluated in the respective signal regions.

8.1.2 Statistical Significance

For purposes of discovering a new signal process (e.g. the Higgs boson signal), the null hypothesis H_0 is defined, describing only known processes: the backgrounds. This is to be tested against the alternative H_1, which includes the backgrounds as well as the signal. In order to verify the compatibility between data and the H_0 hypothesis, it is useful to introduce a test statistic [111]. The Neyman-Pearson lemma affirms that the Profile Likelihood (PL) ratio $\lambda(\mu)$ is the test statistic with the highest discrimination power:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\nu}}})}{L(\hat{\mu}, \hat{\boldsymbol{\nu}})} \tag{8.10}$$

where $\hat{\mu}$ and $\hat{\boldsymbol{\nu}}$ are the ML estimators for the POI and the nuisances respectively, while $\hat{\hat{\boldsymbol{\nu}}}$ is the ML estimator for the nuisances for a given signal strength μ. The assumption of the H_0 hypothesis corresponds to computing the PL ratio with $\mu = 0$. To extract the significance for a $\hat{\mu}$ obtained from the ML fit, the following test statistic is therefore employed:

$$q_0 = \begin{cases} -2 \ln \lambda_0 & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{8.11}$$

where $\lambda_0 = \lambda(\mu = 0)$. This definition reflects the fact that only upward fluctuations of data are regarded as signal evidence. In other words, data show a lack of agreement with the background-only hypothesis only if $\hat{\mu} > 0$. The Wilks' theorem [110] affirms that, assuming some regularity conditions of the likelihood function, q_0 follows a χ^2 distribution with one degree of freedom, in the large sample limit. Therefore, if q_0^{obs} is the observed value of the test statistic, the statistical significance can be computed as:

$$Z = \sqrt{q_0^{obs}} \tag{8.12}$$

Given the relationship between gaussian and χ^2 distributions, $\sqrt{q_0}$ follows a standard gaussian distribution: Z represents the number of standard deviations by which the observed test statistic deviates from zero, assuming the H_0 hypothesis.
The p-value, that is the probability of obtaining a result as compatible or less with H_0, is given by:

$$p = \int_{q_0^{obs}}^{\infty} f(q_0|H_0) \mathrm{d}q_0 = 1 - \Psi\left(\sqrt{q_0^{obs}}\right) \tag{8.13}$$

where $f(q_0|H_0)$ is the p.d.f. of q_0 (which is a χ^2 with one degree of freedom), and Ψ is the cumulative gaussian distribution.
In order to estimate the uncertainty on $\hat{\mu}$ (the ML estimator of the signal strength modifier), the test statistic:

$$q(\mu) = -2 \ln \lambda(\mu) \tag{8.14}$$

is employed, where $\lambda(\mu)$ is defined in Equation 8.10. By definition $q(\mu)$ has a minimum in zero for $\mu = \hat{\mu}$. The signal strength modifier is usually measured with a Confidence Level (CL) $\gamma = 68.3\%$. A CL γ for a given estimator has to be interpreted as the interval that includes the true parameter with a probability γ.
The CL $\gamma = 68.3\%$ for the signal strength is taken as the interval in which the test statistic increases by one, since $\Xi(1) \simeq 0.683$, where Ξ is the cumulative χ^2 distribution with one degree of freedom. Indeed, in the large sample limit, $q(\mu)$ follows the χ^2 distribution for Wilks' theorem. It should be noted that the interval so defined will be in general asymmetric with respect to $\hat{\mu}$. However, in the large sample limit, the likelihood is gaussian distributed, therefore $q(\mu)$ is a parabolic function.
The total uncertainty on the signal strength is composed by three contributions, statistical, experimental and theoretical:

$$\Delta_{tot} = \sqrt{\Delta^2_{stat} + \Delta^2_{exp} + \Delta^2_{theo}} \tag{8.15}$$

The statistical uncertainty is obtained by performing a fit keeping all nuisance parameters fixed to their best fit values. To extract the theoretical uncertainty a second fit is performed, fixing the theoretical nuisances to their nominal values, thus calculating $\Delta_{theo;\, freeze}$. The theoretical uncertainty is then calculated as:

$$\Delta_{theo} = \sqrt{\Delta^2_{tot} - \Delta^2_{theo;\, freeze}} \tag{8.16}$$

while the experimental uncertainty is given by:

$$\Delta_{exp} = \sqrt{\Delta^2_{theo;\, freeze} - \Delta^2_{stat}} \tag{8.17}$$

As explained in Section 7.4, the blind policy forbids to look into real data while developing the analysis. Therefore, the analysis is first tested on the so-called Asimov dataset [111]. This pseudo dataset is defined as the one that, when employed to evaluate the estimators for all the parameters entering the likelihood function, yields their a priori values. In other words, the fit is performed on a dataset that corresponds exactly to the expected signal

Process	**Expected Significance**	**Observed Significance**
VBF	$4.3\,\sigma$	$3.3\,\sigma$
ggH	$2.5\,\sigma$	$2.1\,\sigma$
Higgs Total	$6.0\,\sigma$	$4.7\,\sigma$

Table 8.1: Expected and observed significances for the *VBF* and *ggH* Higgs production mechanisms and for the total Higgs boson production. The ML fit is performed only in the DF *VBF* and DF *ggH* with $N_{jet} > 1$ channels.

plus background. Therefore, the signal strength modifier estimator from the ML fit is exactly equal to one by construction. Since the Asimov dataset is treated as an actual data set complete with statistical errors, it is useful to estimate the expected significance and uncertainties on the signal strength modifier, without looking at the real data.

8.2 Standard Analysis Results

In this analysis, a RooFit [112] interface called Combine [113] was adopted to estimate the physics quantities from the maximisation of the likelihood function. In particular, the minimization of the NLL is implemented via Minuit [114], a numerical minimization routine implemented in RooFit libraries. For this Thesis, the DF *VBF* channel was measured separately from the rest of the CMS analysis. In particular, it was measured together with the DF *ggH* with $N_{jet} > 1$ channel, in order to measure *VBF* and *ggH* Higgs production mechanisms in the phase space with at least two jets with $p_T > 30\,\text{GeV}$. A first measurement was performed by defining a single signal strength modifier for both Higgs production mechanisms. The signal strength modifier obtained from a simultaneous fit on the considered categories is:

$$\hat{\mu} = 0.80^{+0.21}_{-0.19} = 0.80^{+0.15}_{-0.14}(\text{stat})^{+0.09}_{-0.06}(\text{theo})^{+0.11}_{-0.10}(\text{exp}) \tag{8.18}$$

where the uncertainty is decomposed in statistical, experimental and theoretical contributions. The measured signal strength modifier is compatible with one (i.e. the SM expectation) within one standard deviation. The probability of observing a signal at least as large under the background-only

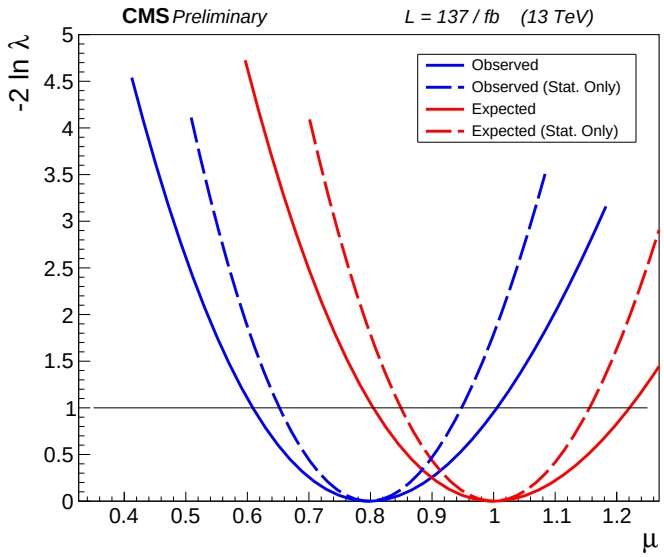


Figure 8.1: Observed and expected likelihood profiles for the global signal strength modifier μ. Dashed curves correspond to the likelihood profiles obtained including only the statistical uncertainty. The crossings with the horizontal line at $-2\ln\lambda = 1$ define the 68.3% CL interval.

hypothesis corresponds to an observed significance of 4.7 standard deviations, to be compared with the expected value of 6.0 standard deviations, obtained with the Asimov dataset.

The expected and observed likelihood profiles as a function of the signal strength modifier are shown in Figure 8.1 (in the form of $q(\mu)$ of Equation 8.14). The likelihood profiles are shown both including all uncertainties, and including only statistical uncertainties. The horizontal line marks the 68.3% CL.

The observed events as a function of the DNN_j discriminators are shown in Figures 8.2, 8.3, 8.4, 8.5 in the respective *j-like* category of the DF *VBF* channel, after the ML fit. An excellent agreement between data and Monte Carlo simulations can be observed in all distributions. Moreover, Figure 8.2 shows that a high *VBF* signal purity is reached for high values of the DNN_{VBF} classifier in the *VBF-like* category, proving the effectiveness of the DNN approach.

Figure 8.6 shows the observed events as a function of the 2D variable employed in the ML fit, $(m_{\ell\ell}, m_T^H)$, in the SR of the DF *ggH* with $N_{jet} > 1$ channel. The 2D variable is unrolled on the horizontal axis for clarity: for every m_T^H bin, the corresponding $m_{\ell\ell}$ distribution is shown. Values on the

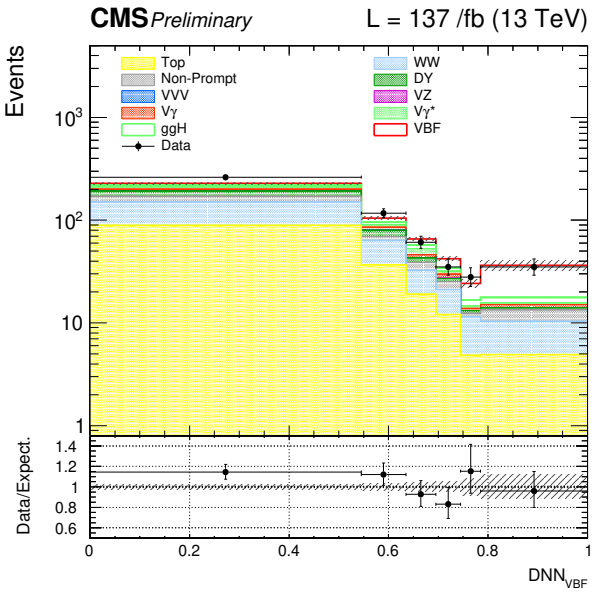


Figure 8.2: Post-fit number of events as a function of the DNN_{VBF} classifier in the *VBF-like* category of the DF *VBF* channel. The dashed gray band accounts for all systematic uncertainties on the signal and background yields.

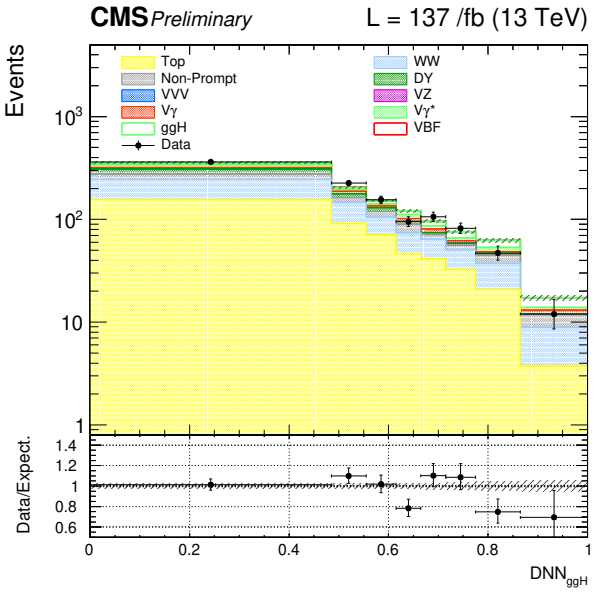


Figure 8.3: Post-fit number of events as a function of the DNN_{ggH} classifier in the *ggH-like* category of the DF *VBF* channel. The dashed gray band accounts for all systematic uncertainties on the signal and background yields.

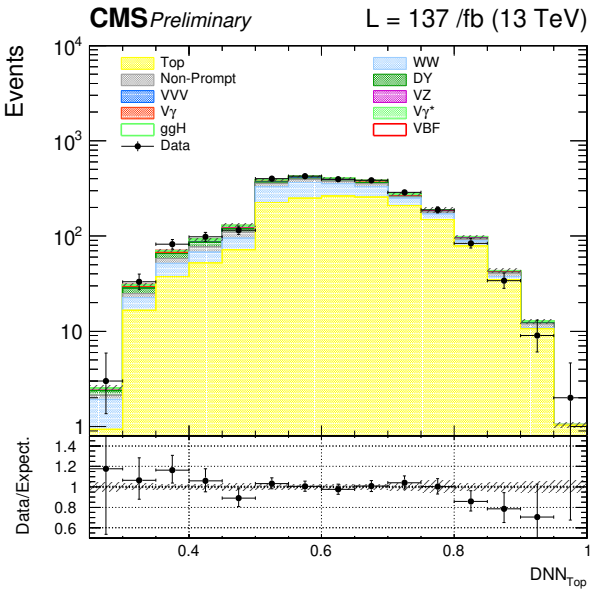


Figure 8.4: Post-fit number of events as a function of the DNN_{Top} classifier in the *Top-like* category of the DF *VBF* channel. The dashed gray band accounts for all systematic uncertainties on the signal and background yields.

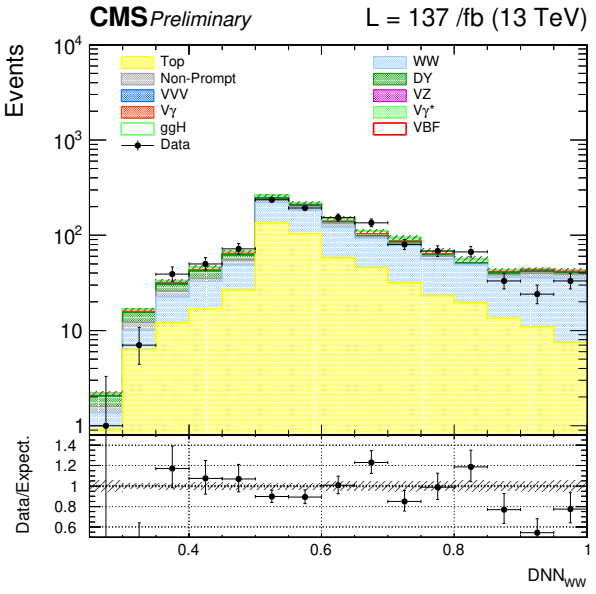


Figure 8.5: Post-fit number of events as a function of the DNN_{WW} classifier in the *WW-like* category of the DF *VBF* channel. The dashed gray band accounts for all systematic uncertainties on the signal and background yields.

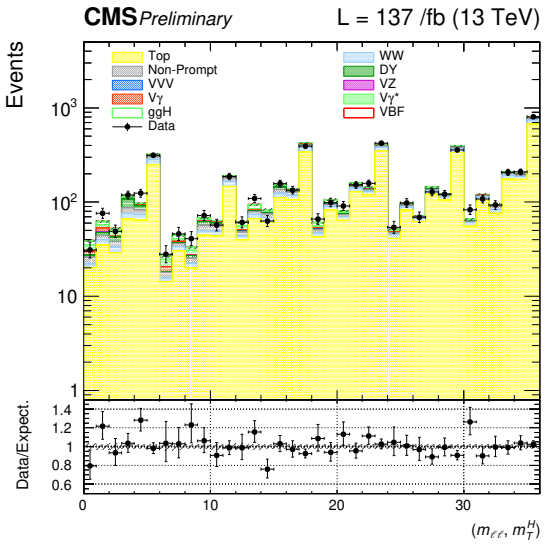


Figure 8.6: Post-fit number of events as a function of $(m_{\ell\ell}, m_T^H)$ in the SR of the DF ggH with $N_{jet} > 1$ channel: for every m_T^H bin, the corresponding $m_{\ell\ell}$ distribution is shown. Values on the horizontal axis correspond to bin numbers. The dashed gray band accounts for all systematic uncertainties on the signal and background yields.

horizontal axis correspond to bin numbers. Also in this case, a good agreement between data and Monte Carlo simulations can be observed.
Appendix A reports the observed distributions (in the 2018 dataset) of the *VBF*-DNN input variables in the SR of the DF *VBF* channel (before applying the DNN selection), after the ML fit. Also in this case, a good agreement between data and Monte Carlo simulations can be observed for all input variables, confirming the goodness of the simulations and of the ML fit.
A second measurement was performed by scaling each signal processes (*VBF* and *ggH* productions) with a different signal strength modifier:

$$\begin{aligned} \hat{\mu}_{VBF} &= 0.77^{+0.28}_{-0.25} = 0.77^{+0.22}_{-0.21}(\text{stat})^{+0.13}_{-0.07}(\text{theo})^{+0.12}_{-0.12}(\text{exp}) \\ \hat{\mu}_{ggH} &= 0.82^{+0.45}_{-0.39} = 0.82^{+0.31}_{-0.30}(\text{stat})^{+0.20}_{-0.11}(\text{theo})^{+0.25}_{-0.23}(\text{exp}) \end{aligned} \tag{8.19}$$

Both measured signal strength modifiers are compatible with one (i.e. the SM expectation) within one standard deviation. The probability of observing a *VBF* (*ggH*) signal at least as large under the background-only hypothesis corresponds to an observed significance of 3.3 (2.1) standard deviations, to be compared with the expected value of 4.3 (2.6) standard deviations. The obtained significances are reported in Table 8.1. In the selected phase space,

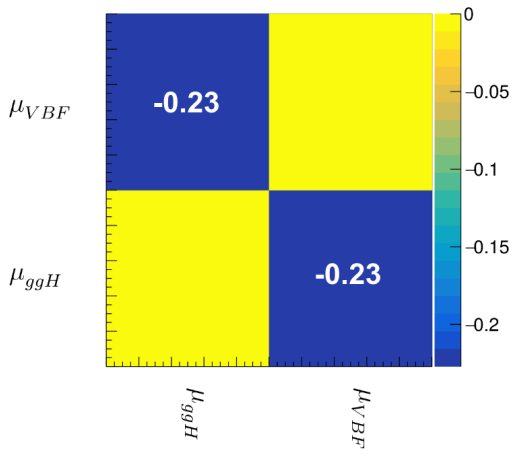


Figure 8.7: Correlation matrix of the measured signal strength modifiers.

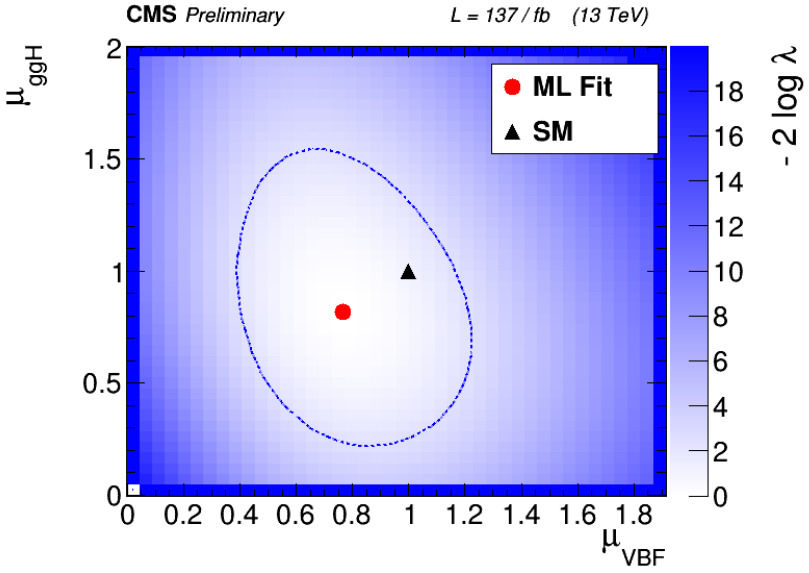


Figure 8.8: Two-dimensional likelihood profile as a function of signal strength modifiers associated with either *VBF* (μ_{VBF}) or *ggH* (μ_{ggH}) Higgs production mechanisms. The 68.3% CL contour is shown as a dashed line. The red circle represents the ML fit value, while the black triangle corresponds to the SM prediction.

where two or more jets with $p_T > 30\,\text{GeV}$ are required, the *VBF* production mechanism is measured with a higher precision, benefiting the selection performed by the *VBF*-DNN.
Figure 8.7 shows the correlation matrix of the measured signal strength modifiers. Figure 8.8 shows the two-dimensional likelihood profile as a function of the two signal strength modifiers μ_{VBF} and μ_{ggH}. The SM expectation is within the 68.3% CL contour of the ML fit value.
On both sets of measurements, it should be noted that the higher error contribution is due to the statistical uncertainty on the measured data. Therefore, the analysis could benefit from a luminosity increase. This will be possible during the LHC Run-3.

8.2.1 Complete Analysis

In the following, results obtained by considering all the analysis categories of the complete CMS analysis are reported.
By defining a global signal strength for all Higgs production mechanisms, the measured signal strength modifier is:

$$\hat{\mu}^{complete} = 0.92^{+0.10}_{-0.09} = 0.92^{+0.05}_{-0.05}(\text{stat})^{+0.06}_{-0.05}(\text{theo})^{+0.06}_{-0.05}(\text{exp}) \tag{8.20}$$

Also in this case, the measured signal strength modifier is compatible with the SM expectation within one standard deviation.
By defining a signal strength modifier for each Higgs production mechanism, the obtained results are:

$$\begin{aligned}
\hat{\mu}^{complete}_{VBF} &= 0.76^{+0.26}_{-0.23} = 0.76^{+0.20}_{-0.19}(\text{stat})^{+0.12}_{-0.07}(\text{theo})^{+0.11}_{-0.11}(\text{exp}) \\
\hat{\mu}^{complete}_{ggH} &= 0.88^{+0.11}_{-0.10} = 0.88^{+0.06}_{-0.06}(\text{stat})^{+0.08}_{-0.06}(\text{theo})^{+0.05}_{-0.05}(\text{exp}) \\
\hat{\mu}^{complete}_{VH} &= 2.08^{+0.47}_{-0.45} = 2.08^{+0.37}_{-0.36}(\text{stat})^{+0.08}_{-0.05}(\text{theo})^{+0.28}_{-0.27}(\text{exp})
\end{aligned} \tag{8.21}$$

Due to the higher cross section, the analysis is more sensitive to the *ggH* production. Moreover, the *ggH* production measurement is greatly improved (with respect to Equation 8.19) due to the inclusion of the DF $N_{jet} = 0$ and $N_{jet} = 1$ categories and the SF *ggH* channel.
The *VBF* production measurement is also slightly improved, due to a better constraint of the *ggH* production and to the inclusion of the SF *VBF* channel. In particular, the probability of observing a *VBF* signal at least as large under the background-only hypothesis corresponds to an observed significance of 3.6 standard deviations, to be compared with the expected

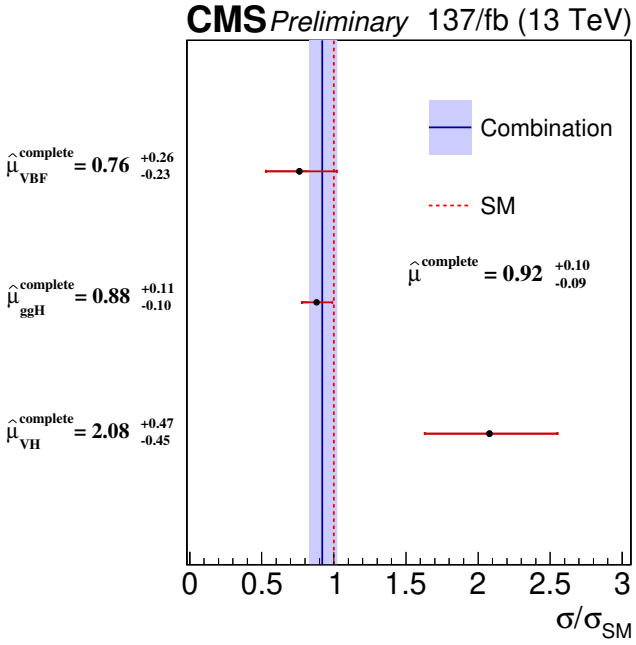


Figure 8.9: Observed signal strength modifiers corresponding to the main Higgs production mechanisms in the complete analysis. The vertical continuous line represents the measurement of the combined signal strength modifier for the three Higgs production mechanisms. The vertical dashed line corresponds to the SM expectation.

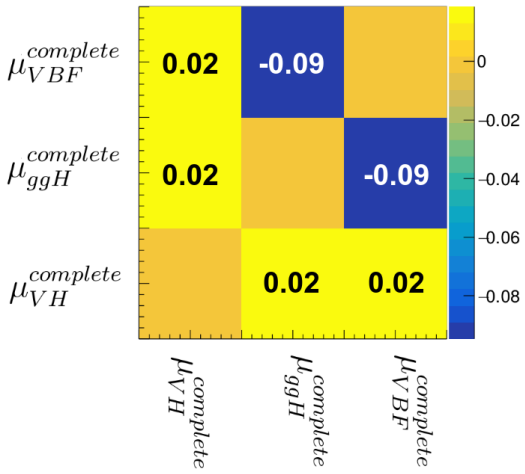


Figure 8.10: Correlation matrix of the measured signal strength modifiers in the complete analysis.

value of 4.7 standard deviations. Both values are slightly higher with respect to Table 8.1.
Even if the VH signal strength modifier is twice the expected SM value, it is compatible with the SM within three standard deviations.
Figure 8.9 shows the measured signal strength modifiers of the three production mechanisms, the combined modifier from Equation 8.20, and the SM expectation.
Finally, Figure 8.10 shows the correlation matrix of the measured signal strength modifiers. The correlation between μ_{VBF} and μ_{ggH} is lower with respect to Figure 8.7, since the ggH production is better constrained in the complete analysis.

8.3 STXS Analysis Results

The STXS bins defined in Section 7.4.5 were measured by defining one signal strength modifier for each bin. In the case of merged bins, a single signal strength modifier is defined. The measured signal strength modifers are reported in Figure 8.11: a good agreement with the SM can be observed.
The cross sections are obtained by multiplying the measured signal strength modifier to the SM cross section[1] of the STXS bin of interest. The measured STXS cross sections are:

$$\begin{aligned}
\hat{\sigma}_{VBF}^{[m_{jj}>700\text{ GeV}, p_T^H<200\text{ GeV}]} &= 0.18^{+0.06}_{-0.08}\,\text{pb} \\
\hat{\sigma}_{VBF}^{[m_{jj}>350\text{ GeV}, p_T^H>200\text{ GeV}]} &= 0.015^{+0.023}_{-0.015}\,\text{pb} \\
\hat{\sigma}_{VBF}^{[350\,\text{GeV}<m_{jj}<700\text{ GeV}, p_T^H<200\text{ GeV}]} &= 0.12^{+0.11}_{-0.11}\,\text{pb} \\
\hat{\sigma}_{ggH}^{[p_T^H>300\text{ GeV}]} &= 0.000^{+0.013}_{-0.000}\,\text{pb} \\
\hat{\sigma}_{ggH}^{[200\text{ GeV}<p_T^H<300\text{ GeV}]} &= 0.30^{+0.11}_{-0.12}\,\text{pb} \\
\hat{\sigma}_{ggH}^{[m_{jj}>700\text{ GeV}, p_T^H<200\text{ GeV}]} &= 0.00^{+0.21}_{-0.00}\,\text{pb} \\
\hat{\sigma}_{ggH}^{[350\text{ GeV}<m_{jj}<700\text{ GeV}, p_T^H<200\text{ GeV}]} &= 0.00^{+0.20}_{-0.00}\,\text{pb} \\
\hat{\sigma}_{ggH}^{[m_{jj}<350\text{ GeV}, p_T^H<120\text{ GeV}]} &= 1.78^{+0.49}_{-0.50}\,\text{pb} \\
\hat{\sigma}_{ggH}^{[m_{jj}<350\text{ GeV}, 120\text{ GeV}<p_T^H<200\text{ GeV}]} &= 0.00^{+0.25}_{-0.00}\,\text{pb} \\
\hat{\sigma}_{VH2j}^{[60\text{ GeV}<m_{jj}<120\text{ GeV}]} &= 0.00^{+0.22}_{-0.00}\,\text{pb}
\end{aligned} \tag{8.22}$$

Due to the limited statistics on some analysis categories, the measured signal strength modifier, hence the cross-section, was zero. A completely asymmet-

[1] The $H \to WW$ decay BR is included in the reported STXS cross sections.

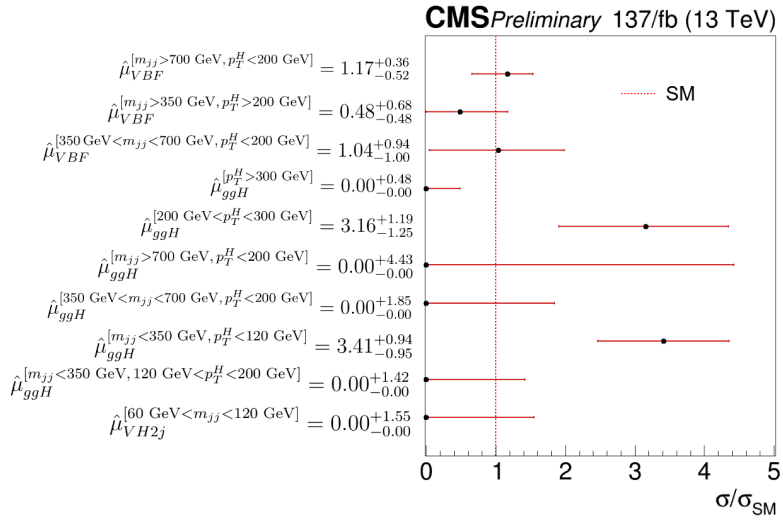


Figure 8.11: Observed signal strength modifiers corresponding to the measured STXS bins. The vertical dashed line corresponds to the SM expectation.

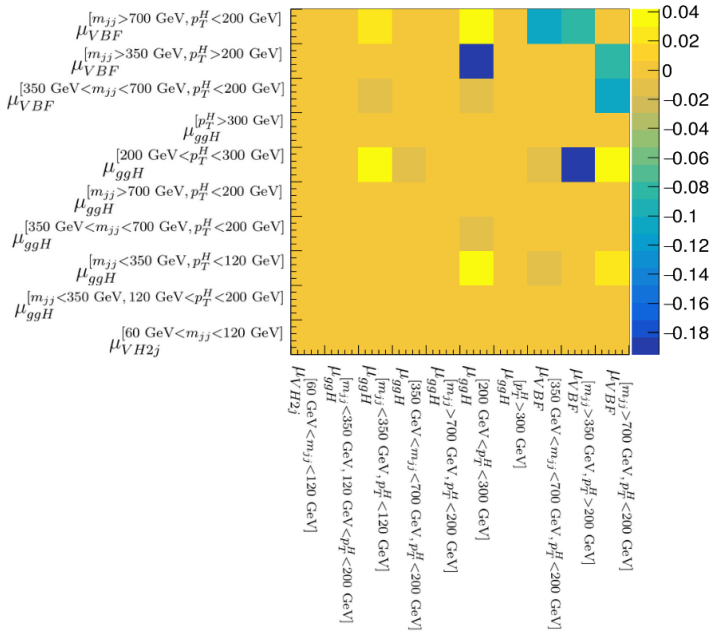
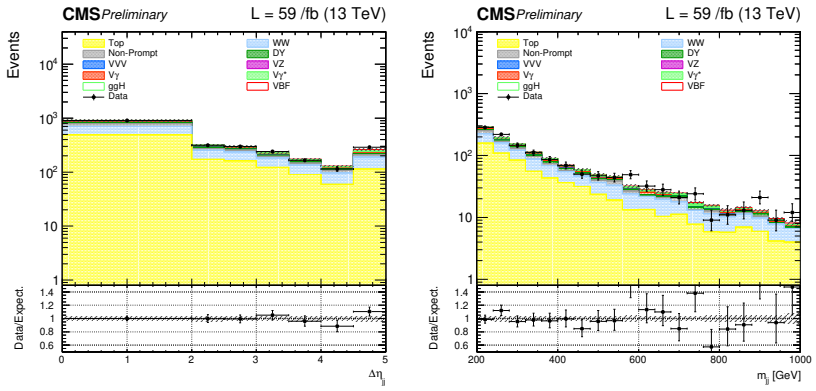


Figure 8.12: Correlation matrix of the measured STXS bins signal strength modifiers.

ric error (assuring a 68% CL) was defined in these cases.
Figure 8.12 shows the correlation matrix of the signal strength modifiers associated to the measured STXS bins: the correlation is always found to be negligible. The signal strength modifiers with a higher correlation correspond to analysis categories defined in nearby phase spaces: this can lead to signal migration from one analysis category to the other.

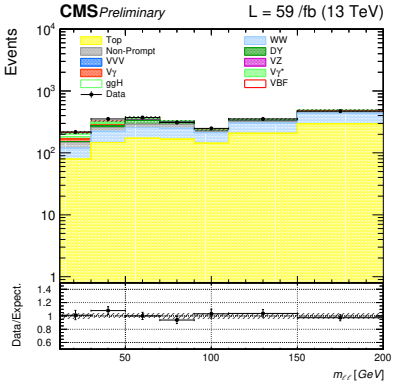
Appendix A

VBF-DNN Input Variables

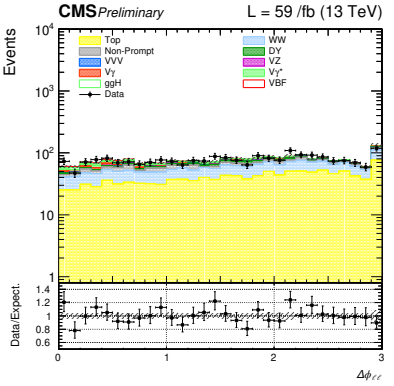


(a) $|\Delta\eta_{jj}|$ (b) m_{jj}

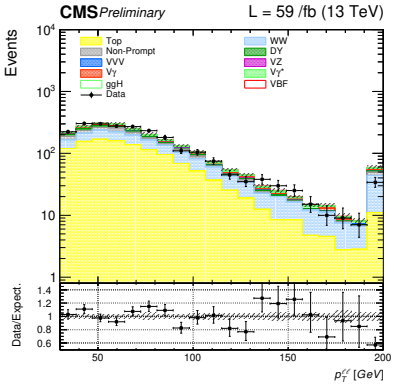
Figure A.1: Post-fit number of events as a function of the *VBF*-DNN input variables in the SR of the DF *VBF* channel for the 2018 dataset. The dashed gray band accounts for all systematic uncertainties on the signal and background yields.



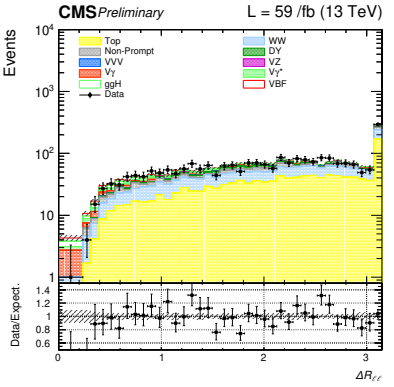
(a) $m_{\ell\ell}$



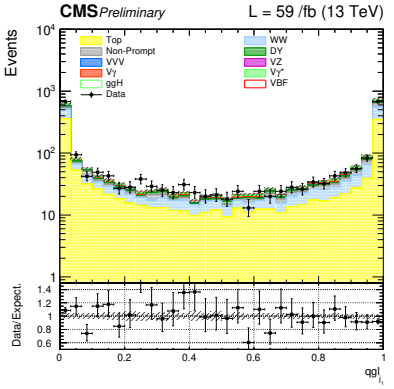
(b) $\Delta\phi_{\ell\ell}$



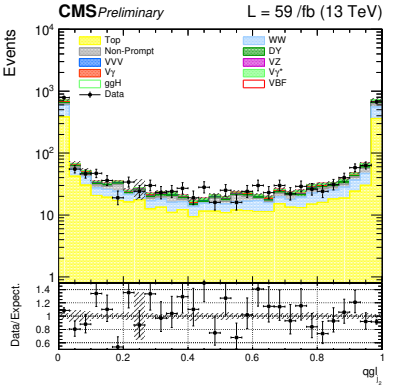
(c) $p_T^{\ell\ell}$



(d) $\Delta R_{\ell\ell}$

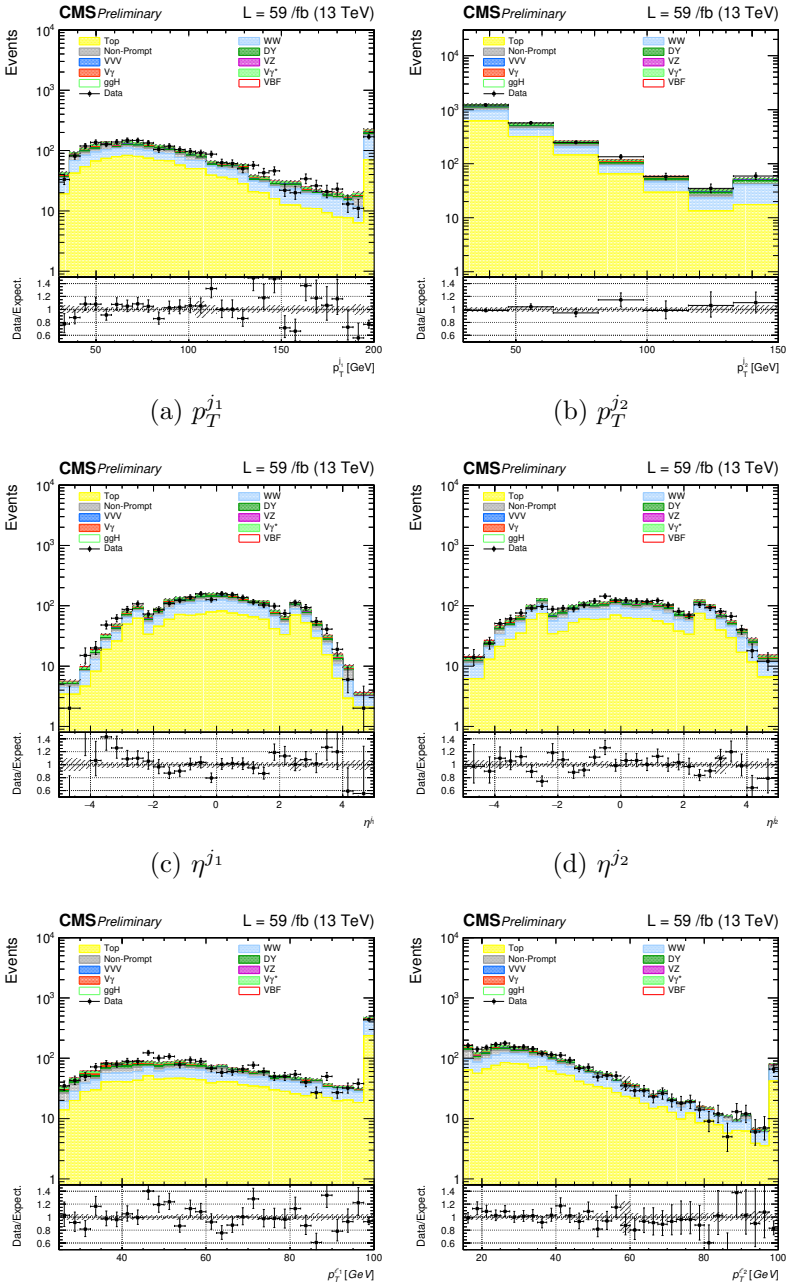


(e) qgl_{j_1}



(f) qgl_{j_2}

Figure A.2: Post-fit number of events as a function of the *VBF*-DNN input variables in the SR of the DF *VBF* channel for the 2018 dataset. The dashed gray band accounts for all systematic uncertainties on the signal and background yields.

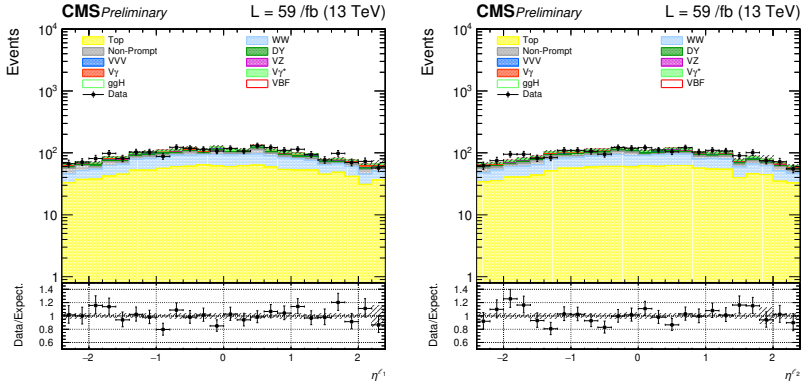


(a) $p_T^{j_1}$ (b) $p_T^{j_2}$

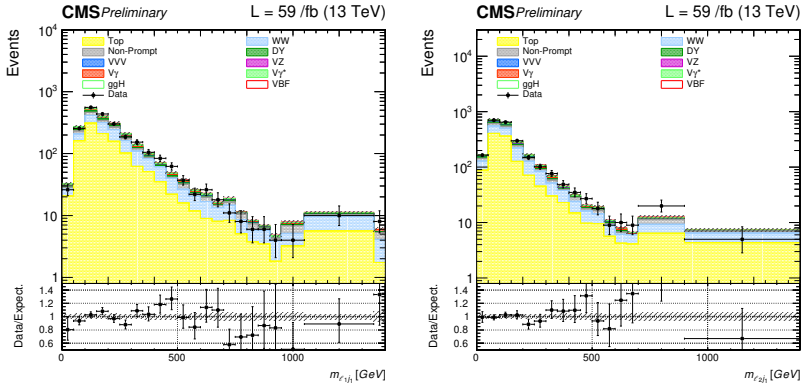
(c) η^{j_1} (d) η^{j_2}

(e) $p_T^{\ell_1}$ (f) $p_T^{\ell_2}$

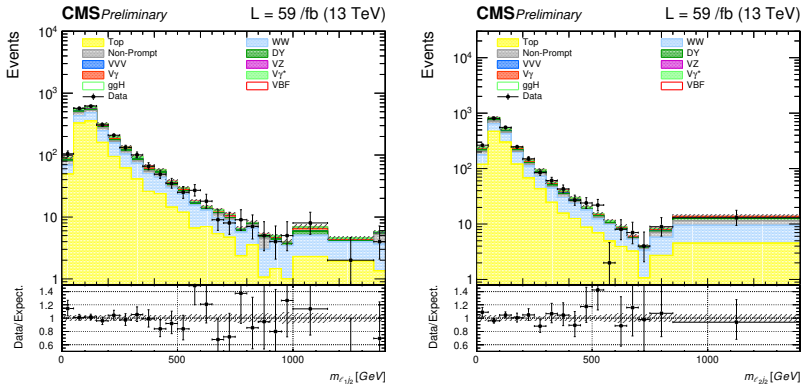
Figure A.3: Post-fit number of events as a function of the *VBF*-DNN input variables in the SR of the DF *VBF* channel for the 2018 dataset. The dashed gray band accounts for all systematic uncertainties on the signal and background yields.



(a) η^{ℓ_1} (b) η^{ℓ_2}

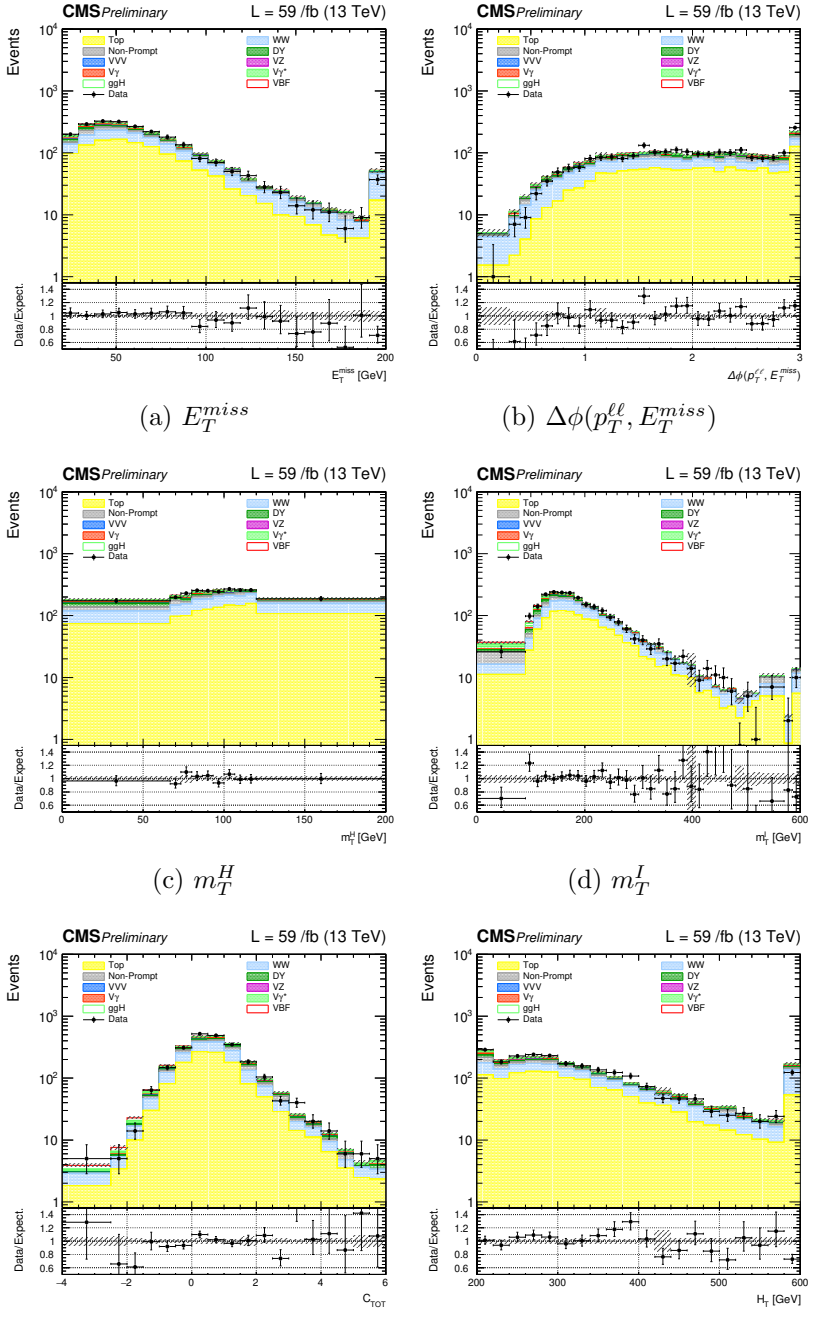


(c) $m_{\ell_1 j_1}$ (d) $m_{\ell_2 j_1}$



(e) $m_{\ell_1 j_2}$ (f) $m_{\ell_2 j_2}$

Figure A.4: Post-fit number of events as a function of the *VBF*-DNN input variables in the SR of the DF *VBF* channel for the 2018 dataset. The dashed gray band accounts for all systematic uncertainties on the signal and background yields.



(a) E_T^{miss} (b) $\Delta\phi(p_T^{\ell\ell}, E_T^{miss})$

(c) m_T^H (d) m_T^I

(e) C_{TOT} (f) H_T

Figure A.5: Post-fit number of events as a function of the *VBF*-DNN input variables in the SR of the DF *VBF* channel for the 2018 dataset. The dashed gray band accounts for all systematic uncertainties on the signal and background yields.

Conclusions

The first part of this Thesis is centred on silicon pixel detectors for the future tracker of the CMS experiment, foreseen for the High Luminosity Phase of the LHC. The 3D concept for silicon pixel sensors is innovative and presents several advantages with respect to traditional, planar, sensors. Thanks to their peculiar structure, 3D sensors are extremely resistant to radiation damage, making them suitable for use in the inner layers of the future CMS tracker. During my PhD, I have worked extensively with 3D and planar pixels detectors. I participated and contributed to test beam experiments at DESY, in order to fully characterise the detectors. The studies I made demonstrated that planar pixel detectors reach a hit detection efficiency of over 99% at a bias voltage of 600 V after an irradiation corresponding the fluence expected after ten years of operations of HL-LHC. 3D pixel detectors have not been tested to these fluences yet (new test beams in the near future will target their characterisation), but are expected to reach similar efficiencies with far lower bias voltages, around 150 V. Having high efficiencies at relatively low bias voltages leads to a lower power consumption and reduces the susceptibility to sparking issues with respect to planar sensors. Both of these features are invaluable in the inner tracker environment.
Aside from the fabrication technology (3D or planar sensors), two pixel pitches are being considered for the future CMS Inner Tracker: $25 \times 100\,\mu m^2$ and $50 \times 50\,\mu m^2$. However, since these sensors are bump bonded to a readout chip with a pixel pitch of $50 \times 50\ \mu m^2$, the $25 \times 100\ \mu m^2$ sensors needs to be adapted to the different pitch with a particular arrangement of the bump bonding pads. This arrangement causes a cross-talk effect between adjacent pixels in planar pixel sensors. Different variations of the design of the pixels have been produced in order to reduce the cross-talk effect. I characterised various designs of $25 \times 100\ \mu m^2$ pixel sensors, both 3D and planar, in order to quantify the cross-talk effect. The design variations have proven successful to reduce the cross-talk in planar pixel sensors, while in 3D sensors the cross-talk is found to be negligible.
Among the studies presented in this Thesis, the spatial resolution of 3D and

planar pixel detectors was thoroughly evaluated. Non-irradiated 3D and planar pixel detectors have shown remarkable spatial resolution, down to 2 µm for the 25 µm pitch, and down to 5 µm for the 50 µm pitch. The results presented in this Thesis will contribute significantly to the choice of the pixel sensors to be used in the future CMS Inner Tracker.
The second part of this Thesis is centred on the measurement of the Vector Boson Fusion (VBF) Higgs production mechanism: such a rare process is sensitive to new physics phenomena, and allows to put constraints on the compatibility of the Higgs boson with the Standard Model.
I studied the $H \to WW$ decay channel, and I implemented a multivariate analysis to boost the sensitivity. I developed and trained a Deep Neural Network (DNN), built with Keras with TensorFlow back-end, to isolate the signal from the main backgrounds: top quark events, non-resonant WW and gluon fusion Higgs production mechanism.
My work is part of a larger analysis performed by the CMS collaboration, which aims at measuring of the Higgs boson properties in the $H \to WW$ decay channel targeting not only the VBF, but also the other Higgs production mechanisms. The analysis is based on proton-proton collision data produced at the LHC with $\sqrt{s} = 13$ TeV and collected by the CMS detector during the LHC Run-2, for a total integrated luminosity of about 137 fb^{-1}.
The VBF Higgs production mechanism is observed with a significance of 3.6 standard deviations, resulting in the first evidence of this production mechanism in the WW decay channel with the CMS experiment. Moreover, the measured cross section is compatible with the Standard Model within one standard deviation.
The analysis is also implemented in the Simplified Template Cross Section (STXS) framework. The goal of the STXS scheme is to reduce the theoretical uncertainties that are directly folded into the measurements, while at the same time allowing for the combination of the measurements between different decay channels as well as between experiments. The DNN I developed was used to disentangle Higgs production mechanisms in the STXS analysis. This work established an analysis strategy that will be used for the LHC Run-3 and possibly beyond it.

Bibliography

[1] P.A. Zyla et al. (Particle Data Group), *Review of Particle Physics*, Progress of Theoretical and Experimental Physics, 083C01, 2020.

[2] M. Carlà, *Appunti di Elettronica per Fisici*, http://studenti.fisica.unifi.it/~carla/appunti/2016-17/.

[3] J. Becker, *Signal development in silicon sensors used for radiation detection*, University of Hamburg PhD Thesis, 2010, http://www.desy.de/~beckerj/cc/files/Thesis.pdf.

[4] P. Weigell, *Investigation of Properties of Novel Silicon Pixel Assemblies Employing Thin n-in-p Sensors and 3D-Integration*, CERN-THESIS-2012-229 & MPP-2013-5, 2013.

[5] M. Moll, *Radiation Damage in Silicon Particle Detectors*, DESY-THESIS-1999-040, 1999.

[6] G. Lindstrom et al., *Radiation hardness of silicon detectors: a challenge from high-energy physics*, Nuclear Instruments and Methods in Physics Research A Vol. 426 pp. 1-15, 1999, https://doi.org/10.1016/S0168-9002(98)01462-4.

[7] G. F. Knoll, *Radiation Detection and Measurement*, John Wiley & Sons, 2010.

[8] G.F. Dalla Betta, R. Mendicino, D. Sultan, M. Boscardin, G. Giacomini, S. Ronchin, N. Zorzi, G. Darbo, M. Meschini, A. Messineo, *Small pitch 3D devices*, PoS Vertex2016 28, 2017, https://doi.org/10.22323/1.287.0028.

[9] L. Evans et al., *LHC Machine*, JINST Vol. 3 S08001, 2008.

[10] The CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST Vol. 3 S08004, 2008.

[11] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST Vol. 3 S08003, 2008.

[12] The LHCb Collaboration, *The LHCb Detector at the LHC*, JINST Vol. 3 S08005, 2008.

[13] The ALICE Collaboration, T*he ALICE experiment at the CERN LHC*, JINST Vol. 3 S08002, 2008.

[14] TOTEM Collaboration, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST Vol. 3 S08007, 2008.

[15] LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, JINST Vol. 3 S08006, 2008.

[16] The CMS Collaboration, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, CMS-TRK-11-001 & CERN-PH-EP-2014-070, 2014.

[17] The CMS Collaboration, *CMS: The Tracker Project Technical Design Report*, CERN-LHCC-98-6 & CMS-TDR-5, 1998.

[18] The CMS Collaboration, *The CMS tracker: addendum to the Technical Design Report*, CERN-LHCC-2000-016 & CMS-TDR-5-add-1, 2000.

[19] The CMS Collaboration, *CMS Technical Design Report for the Pixel Detector Upgrade*, CERN-LHCC-2012-016 & CMS-TDR-011, 2012.

[20] The CMS Collaboration, *The Phase-2 Upgrade of the CMS Tracker*, CERN-LHCC-2017-009 & CMS-TDR-014, 2017.

[21] The CMS Collaboration, *The CMS electromagnetic calorimeter project: Technical Design Report*, CERN-LHCC-97-033 & CMS-TDR-4, 1997.

[22] The CMS Collaboration, *Changes to CMS ECAL electronics: addendum to the Technical Design Report*, CERN-LHCC-2002-027 & CMS-TDR-4-add-1, 2002.

[23] The CMS Collaboration, *The CMS hadron calorimeter project: Technical Design Report*, CERN-LHCC-97-031 & CMS-TDR-2, 1997.

[24] The CMS Collaboration, *The CMS muon project: Technical Design Report. Technical Design Report CMS*, CERN-LHCC-97-032 & CMS-TDR-3, 1997.

[25] The CMS Collaboration, *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger*, CERN-LHCC-2002-026 & CMS-TDR-6, 2002.

[26] G. Apollinari, I. Béjar Alonso, O. Brüning, M. Lamont and L Rossi, *High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report*, CERN-2015-005, 2015.

[27] D. Contardo, M. Klute, J. Mans, L. Silvestris and J. Butler, *Technical Proposal for the Phase-II Upgrade of the CMS Detector*, CERN-LHCC-2015-010 & CMS-TDR-15-02, 2015.

[28] The CMS Collaboration, *Projected Performance of an Upgraded CMS Detector at the LHC and HL-LHC: Contribution to the Snowmass Process*, 2015, `https://arxiv.org/abs/1307.7135`.

[29] The CMS Collaboration, *The Phase-2 Upgrade of the CMS Muon Detectors*, CERN-LHCC-2017-012, 2017.

[30] The CMS Collaboration, *The Phase-2 Upgrade of the CMS Barrel Calorimeters*, CERN-LHCC-2017-011, 2017.

[31] THe CMS Collaboration, *The Phase-2 Upgrade of the CMS Endcap Calorimeter*, CERN-LHCC-2017-023 & CMS-TDR-019, 2017.

[32] The CMS Collaboration, *Technical proposal for a MIP timing detector in the CMS experiment Phase 2 upgrade*, CERN-LHCC-2017-027, 2017.

[33] The CMS Collaboration, *CMS Phase II Upgrade Scope Document*, CERN-LHCC-2015-019, 2015.

[34] T. T. Böhlen, F. Cerutti, M. P. W. Chin, A. Fassò, A. Ferrari, P. G. Ortega, A. Mairani, P. R. Sala, G. Smirnov, and V. Vlachoudis, *The FLUKA Code: Developments and Challenges for High Energy and Medical Applications*, Nuclear Data Sheets 120 pp. 211-214, 2014, `https://doi.org/10.1016/j.nds.2014.07.049`.

[35] Alfredo Ferrari, Paola R. Sala, Alberto Fasso, and Johannes Ranft, *FLUKA: A multi-particle transport code*, CERN-2005-10, 2005.

[36] J. Christiansen, M. Garcia-Sciveres, *RD Collaboration proposal: Development of pixel readout integrated circuits for extreme rate and radiation*, CERN-LHCC-2013-008, 2013.

[37] M. Meschini, G.F. Dalla Betta, M. Boscardin, G. Calderini, G. Darbo, G. Giacomini, A. Messineo, S. Ronchin, *The INFN-FBK pixel R&D program for HL-LHC*, Nuclear Instrumentation and Methods in Physics Research A Vol. 831 pp. 116-121, 2016, https://doi.org/10.1016/j.nima.2016.05.009.

[38] RD53 Collaboration, *The RD53A Integrated Circuit*, CERN-RD53-PUB-17-001, 2017.

[39] N. Emriskova, *Comparative evaluation of analogue front-end designs for the CMS Inner Tracker at the High Luminosity LHC*, 2021, https://arxiv.org/abs/2105.00070v2.

[40] M. Daas et al., *BDAQ53, a versatile pixel detector readout and test system for the ATLAS and CMS HL-LHC upgrades*, 2020, https://arxiv.org/abs/2005.11225.

[41] Xilinx, *7 Series FPGAs Datasheet*, https://www.xilinx.com/support/documentation/data_sheets/ds180_7Series_Overview.pdf.

[42] ROOT (Data Analysis Framework) website, https://root.cern.ch/.

[43] Marius C. Mertens and James Ritman, *A method for fast feature extraction in threshold scans*, Nuclear Instrumentation and Methods in Physics Research A Vol. 735 pp. 615-619, 2014, https://doi.org/10.1016/j.nima.2013.10.022

[44] R. Turchetta, *Spatial resolution of silicon microstrip detectors.*, Nuclear Instrumentation and Methods in Physics Research A Vol. 335 pp. 48-58, 1993, https://doi.org/10.1016/0168-9002(93)90255-G.

[45] V. Chiochia et al., *A novel technique for the reconstruction and simulation of hits in the CMS pixel detector*, Nuclear Science Symposium Conference Record pp. 1909-1912, 2008.

[46] V. Blobel, *A new fast track-fit algorithm based on broken lines*, Nuclear Instrumentation and Methods in Physics Research A Vol. 556 pp. 14-17, 2006, https://doi.org/10.1016/j.nima.2006.05.156.

[47] C. Kleinwork, *General broken lines as advanced track fitting method*, Nuclear Instrumentation and Methods in Physics Research A Vol. 673 pp. 107-110, 2012, https://doi.org/10.1016/j.nima.2012.01.024.

[48] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, uclear Instrumentation and Methods in Physics Research A Vol. 262 pp. 444-450, 1987, https://doi.org/10.1016/0168-9002(87)90887-4.

[49] V. Blobel, *Software alignment for tracking detectors*, Nuclear Instrumentation and Methods in Physics Research A Vol. 556 pp. 5-13, 2006, https://doi.org/10.1016/j.nima.2006.05.157.

[50] V. Blobel, C. Kleinwort, and F. Meier, *Fast alignment of a complex tracking detector using advanced track models*, Computer Physics Communications Vol. 182 9 pp. 1760-1753, 2011, https://doi.org/10.1016/j.cpc.2011.03.017

[51] C. C. Paige and M. A. Saunders, *Solution of Sparse Indefinite Systems of Linear Equations*, SIAM Journal of Numerical Analysis Vol. 12 4 pp. 617-629, 1975.

[52] J. Dreyling-Eschweiler et al., *The DESY II test beam facility*, Nuclear Instrumentation and Methods in Physics Research A Vol. 922 pp. 265-286, 2019, https://doi.org/10.1016/j.nima.2018.11.133.

[53] H. Jansen et al., *Performance of the EUDET-type beam telescopes*, EPJ Techniques and Instrumentation Vol. 03 7, 2016, https://doi.org/10.1140/epjti/s40485-016-0033-2

[54] Nikkie Deelen, *Characterizing detector modules for the Upgrade of the Silicon Tracker of the Compact Muon Solenoid experiment*, CERN-THESIS-2019-275, 2019.

[55] Karlsruhe Irradiation Center website, https://www.etp.kit.edu/english/irradiation_center.php

[56] F. Ravotti et al., *A New High-Intensity Proton Irradiation Facility at the CERN PS East Area*, PoS TIPP2014 354, 2014, https://doi.org/10.22323/1.213.0354.

[57] S. Spannagel, *Test Beam Measurements for the Upgrade of the CMS Pixel Detector and Measurement of the Top Quark Mass from Differential Cross Sections*, CMS-TS-2016-010 & CERN-THESIS-2016-059, 2016.

[58] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-kt jet clustering algorithm*, JHEP Vol. 4 63, 2008.

[59] G. Altarelli, G. Parisi, *Asymptotic Freedom in Parton Language*, Nuclear Physics B Vol. 126 2, 1977, https://doi.org/10.1016/0550-3213(77)90384-4.

[60] J. M. Butterworth, G. Dissertori, G. P. Salam *Hard Processes in Proton-Proton Collisions at the Large Hadron Collider*, Annual Review of Nuclear and Particle Science Vol. 62, 2012, https://arxiv.org/abs/1202.0583.

[61] The ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, Physical Review Letters Vol. 114 191803, 2015.

[62] LHC Higgs Cross Section Working Group Collaboration, *Handbook of LHC Higgs cross sections: deciphering the nature of the Higgs sector*, 2016, https://arxiv.org/abs/1610.07922.

[63] A. M. Sirunyan et al., *Observation of Higgs boson decay to bottom quarks*, Physical Review Letters Vol. 121 121801, 2018.

[64] A. M. Sirunyan et al., *Search for the Higgs boson decaying to two muons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Physical Review Letters Vol. 122 021801, 2019.

[65] A. M. Sirunyan et al., *Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector*, Physics Letters B Vol. 779 pp. 283-316, 2018, https://doi.org/10.1016/j.physletb.2018.02.004.

[66] A. M. Sirunyan et al., *Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP Vol. 11 47, 2017.

[67] A. M. Sirunyan et al., *Measurement of inclusive and differential Higgs boson production cross sections in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP Vol. 01 183, 2019.

[68] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Physics Letters B Vol. 659 pp. 119-126, 2008, https://doi.org/10.1016/j.physletb.2007.09.077

[69] The CMS Collaboration, *Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV*, JINST Vol. 13 P05011, 2018.

[70] D. L. Rainwater, D. Zeppenfeld, *Observing $H \to W^{(*)}W^{(*)} \to e^{\pm}\mu^{\mp}p\!\!\!/_T$ in weak boson fusion with dual forward jet tagging at the CERN LHC*, Physics Review D Vol. 60 113004, 1999.

[71] A. Sirunyan et al., *Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s}$=13 TeV*, Physics Letters B Vol. 791 pp. 96-129, 2019, https://doi.org/10.1016/j.physletb.2018.12.073.

[72] CMS Collaboration, *Particle-flow reconstruction and global event description with the CMS detector*, JINST Vol. 12 P10003, 2017.

[73] W. Waltenberger, R. Frühwirth, P. Vanlaer, *Adaptive vertex fitting*, Journal of Physics G Vol. 34 12, 2017.

[74] The CMS Collaboration, *Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016-2018*, CMS-DP-2020-019, 2020.

[75] The CMS Collaboration, *Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector*, JINST Vol. 14 P07004, 2019.

[76] D. Bertolini, P. Harris, M. Low, N. Tran, *Pileup Per Particle Identification*, JHEP Vol. 10 59, 2014.

[77] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Computer Physics Communications Vol. 191 pp. 159-177, 2015, https://doi.org/10.1016/j.cpc.2015.01.024.

[78] NNPDF Collaboration, *Parton distributions from high-precision collider data*, The European Physical Journal C Vol. 77 663, 2017.

[79] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP Vol. 11 70, 2017.

[80] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini, *Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM*, JHEP Vol. 02 088, 2012.

[81] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP Vol. 02 37, 2017.

[82] G. Luisoni, P. Nason, C. Oleari, F. Tramontano, *HW/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO*, JHEP Vol. 10 83, 2013.

[83] K. Hamilton, P. Nason, E. Re, G. Zanderighi, *NNLOPS simulation of Higgs boson production*, JHEP Vol. 10 222, 2013.

[84] K. Hamilton, P. Nason, G. Zanderighi, *Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator*, JHEP Vol. 05 140, 2015.

[85] S. Bolognesi et al., *On the spin and parity of a single-produced resonance at the LHC*, Physics Review D Vol. 86 095031, 2012.

[86] P. Nason, G. Zanderighi, W^+W^-, *WZ and ZZ production in the POWHEG-BOX-V2*, The European Physical Journal C Vol. 74 2702, 2014.

[87] J. M. Campbell, R. K. Ellis, C. Williams, *Vector boson pair production at the LHC*, JHEP Vol. 07 018, 2011.

[88] J. M. Campbell, R. K. Ellis, and W. T. Giele, A multi-threaded version of MCFM, The European Physical Journal C Vol. 75 246, 2015.

[89] F. Caola et al., *QCD corrections to vector boson pair production in gluon fusion including interference effects with off-shell Higgs at the LHC*, JHEP Vol. 07 087, 2016.

[90] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP Vol. 07 79, 2014.

[91] S. Frixione, P. Nason, G. Ridolfi, *A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadrproduction*, JHEP Vol. 09 126, 2007.

[92] S. Alioli, P. Nason, C. Oleari, E. Re, *NLO single-top production matched with shower in POWHEG: s- and t-channel contributions*, JHEP Vol. 09 111, 2009.

[93] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, The European Physical Journal C Vol. 71 1547, 2011.

[94] The Geant4 Collaboration, *GEANT4 — a simulation toolkit*, Nuclear Instrumentation and Methods in Physics Research A Vol. 506 3 pp. 250-303, 2003, `https://doi.org/10.1016/S0168-9002(03)01368-8`.

[95] The CMS Collaboration, *Higgs boson production in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at* $\sqrt{s} = 13$ *TeV*, CMS-PAS-HIG-19-008, 2020.

[96] The ATLAS and CMS Collaborations, LHC Higgs Combination Group, *Procedure for the LHC Higgs boson search combination in Summer 2011*, ATL-PHYS-PUB 2011-1 & CMS NOTE 2011/005, 2011.

[97] N. Berger et al., *Simplified template cross sections - stage 1.1*, `https://arxiv.org/abs/1906.02754`, 2019.

[98] F. Chollet et al., *Keras*, 2015, `https://keras.io`.

[99] M. Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, `https://www.tensorflow.org/`.

[100] D. P. Kingma, J. Ba, *Adam: A Method for Stochastic Optimization*, 2015, `https://arxiv.org/abs/1412.6980`.

[101] The CMS Collaboration, *Performance of quark/gluon discrimination in 13 TeV data*, CMS-DP-2016-070, 2016.

[102] M. Czakon et al., *Top-pair production at the LHC through NNLO QCD and NLO EW*, JHEP Vol. 10 186, 2017.

[103] P. Meade, H. Ramani, M. Zeng, *Transverse momentum resummation effects in* W^+W^- *measurements*, Physics Review D Vol. 90 114006, 2014.

[104] P. Jaiswal, T. Okui, *Explanation of the WW excess at the LHC by jet-veto resummation*, Physics Review D Vol. 90 073009, 2014.

[105] The CMS Collaboration, *CMS luminosity measurements for the 2016 data-taking period*, CMS-PAS-LUM-17-001, 2017.

[106] The CMS Collaboration, *CMS luminosity measurement for the 2017 data-taking period at* $\sqrt{s} = 13$ *TeV*, CMS-PAS-LUM-17-004, 2017.

[107] The CMS Collaboration, *CMS luminosity measurement for the 2018 data-taking period at* $\sqrt{s} = 13$ *TeV*, CMS-PAS-LUM-18-002, 2019.

[108] The CMS Collaboration, *Measurement of the inelastic proton-proton cross section at* $\sqrt{s} = 13$ *TeV*, JHEP Vol. 07 161, 2018.

[109] G. Cowan, *Statistical Data Analysis*, Clarendon Press, 1998.

[110] L. Lista, *Statistical Methods for Data Analysis in Particle Physics*, Springer, 2016.

[111] G. Cowan, K. Cranmer, E. Gross et al., *Asymptotic formulae for likelihood-based tests of new physics*, The European Physical Journal C Vol. 71 1554, 2011.

[112] W. Verkerke, D. Kirkby, *The RooFit toolkit for data modeling*, 2003, https://arxiv.org/abs/physics/0306116.

[113] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group, *Procedure for the LHC Higgs boson search combination in Summer 2011*, CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, 2011.

[114] F. James, M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Computer Physics Communications Vol. 10 pp. 343-367, 1975, https://doi.org/10.1016/0010-4655(75)90039-9.