Università degli Studi Roma Tre

Dipartimento di Matematica e Fisica

Doctoral program in Physics – XXXIV cycle

A thesis presented for the degree of
Doctor of Philosophy in Physics

Analysis of the $(W)H \rightarrow (W)WW^*$ channel and validation of resistive Micromegas detectors for the ATLAS experiment

Author:
Maria Teresa Camerlingo

Local Supervisors:
Dr. **Michela Biglietti**
Dr. **Mauro Iodice**

CERN Supervisor:
Dr. **Paolo Iengo**

PhD Program Coordinator:
Prof. **Giuseppe Degrassi**

a. y. 2021/2022

Contents

Introduction

The CERN ATLAS (A Toroidal LHC ApparatuS) [1] experiment is one of the major general-purpose experiments of High Energy Physics, operating at LHC (Large Hadronic Collider) [2]. Its collaboration searches to prove the Standard Model (SM) predictions and observe phenomena of new physics Beyond the Standard Model (BSM). Both research studies require continuous upgrades of the ATLAS detectors.
One of the main ATLAS goals is the search for the mass origin. A significant achievement has been the discovery of the Higgs boson by ATLAS [3] and CMS (Compact Muon Solenoid) [4], the other general-purpose LHC experiment that scrutinises the ATLAS results (and vice versa).
The first topic of this thesis, "Analysis of the $W(H) \rightarrow W(WW^*)$ channel", is part of the ongoing cross-section measurements of the associated Higgs boson production with a weak vector boson (VH) and subsequent $H \rightarrow WW^*$ decay using the full ATLAS Run 2 dataset of the LHC proton-proton collisions at 13 TeV centre-of-mass energy.
The previous measurements [5] have estimated the VH cross-section times the $H \rightarrow WW^*$ branching ratio in agreement with the SM without a channel observation due to the partial integrated luminosity of 36.1 fb^{-1} exploited at that time.
Thanks to the full ATLAS Run 2 integrated luminosity of 139 fb^{-1} and the inclusion of new channels, a significant improvement in the measurement precision is foreseen, and events of this channel could be observed for the first time. Indeed the expected significance of the analysis is 5.13σ. Moreover, the full ATLAS Run 2 dataset statistic also allows analysing the $VH \rightarrow VWW^*$ channel in the *Simplified Template Cross-Section* ($STXS$), a measurement framework in common with all the LHC experiments.
My personal and original contributions to the inclusive and $STXS$ analyses have consisted of implementing artificial neural networks. Their outputs have been used to define multivariate variables able to discriminate the signal among the background processes in the $WH \rightarrow W(WW^*)$ sub-channel with three charged leptons (τ excluded) in the final state where at least two leptons have the same flavour and opposite charge sign.
The second topic of this thesis, "Validation of Resistive Micromegas Detectors", is carried out at CERN in the scope of the "CERN Doctoral Student programme". It is related to the validation of the resistive Micromegas (MM) detectors in some upgraded Muon Spectrometer detector subsystems, called New Small Wheels. These detectors have replaced the old detector-equipped structures in the ATLAS innermost End-Cap regions. They are formed by combined small-strip Thin Gap Chambers (sTGC) and resistive Micromegas detectors providing efficient performances in muon triggering and tracking compatible with the expected Run 3 instantaneous and integrated luminosities up to $5{\times}10^{34}$ cm^{-2} s^{-1} and 3000 fb^{-1}, respectively.
My contribution to the MM detector validation tests has concerned different stages of the upgrade project, from the participation in construction operations and tests on the single detector components to studies on the final detector performances.
The thesis has the following structure. The ATLAS apparatus of LHC Run 2 and the status of the

ATLAS analyses on the cross-sections of the Higgs boson production processes are described in Chapter 1 to introduce the subjects of the analysis in Chapter 2.
Chapter 2 is dedicated to a general overview of the $VH \rightarrow VWW^*$ analysis and my contribution to the multivariate analysis of the final state with three leptons ($WH \rightarrow WWW^*$ process) mainly affected by the Z boson backgrounds.
A description of the Micromegas technology developed for the ATLAS detectors is reported in Chapter 3, together with the results of the studies about: the specifications of the ATLAS readout MM boards, on the high voltage stability of the MM detectors and their possible correlations.
The thesis also has five appendices that support the main chapters, except for Appendix E. It concerns characterisation studies performed on new resistive Micromegas detectors operating in an O(10 MHz/cm^2) rate environment. They are part of an INFN R&D project synergic to ATLAS Micromegas detectors.

1 ATLAS Experiment

ATLAS is a general-purpose experiment, potentially sensible to physics beyond Standard Model (BSM). It operates at Large Hadronic Collider (LHC) [2]. From the start of the ATLAS data acquisition, significant progress has been made in the search for the mass origin, measurements of the properties of the top quark, B-hadrons, Flavour Physics and supersymmetry search [6]. With the increase of the data statistics, other research items become accessible, including direct dark matter research, exotic processes and studies on interaction with heavy nuclei.
The ATLAS apparatus the principal production and decay modes of the Higgs boson are presented in this chapter.

1.1 Large Hadronic Collider

Large Hadronic Collider at CERN is the current most powerful collider, with its centre-of-mass energy of 13 TeV and 27 km of circumference. It is located underground at the Swiss-French border. Two proton or ion beams circulate inside LHC in opposite ways to collide at four points of the machine where the apparatuses of the major LHC experiments are situated. They are ATLAS (A Toroidal LHC ApparatuS) [1], CMS [7], ALICE [8] and LHCb [9]. Together with the ATLAS experiment, CMS is a general-purpose experiment, sharing the same goals. The ATLAS and CMS experiments use different technical solutions and different magnet-system designs to investigate Standard Model (SM) and Beyond Standard Model (BSM) phenomena of Particle Physics. One of their main achievements is the Higgs boson discovery [3] [4].
The ALICE experiment is dedicated to heavy-ion physics at LHC. It is designed to study the physics of strongly interacting matter at extreme energy densities, where quark-gluon plasma forms.
Finally, the LHCb experiment investigates the differences between matter and antimatter by studying b quark systems.
The detectors and the collider have fully operated since 2008, with two long shutdown periods for maintenance operations in 2013-2014 and from 2019 up to March 2022.
During Run 2 proton-proton collisions, the centre-of-mass energy has reached the 13 TeV value with a maximum instantaneous luminosity of $2.1 \cdot 10^{34}$ cm^{-2} s^{-1}. All the experiments have achieved important discoveries in the High Energy Physics, regarding the mass origin at the electroweak scale, in B-hadron Physics, in top quark Physics, in supersymmetry and in dark matter searches.
LHC still offers a broad range of physics opportunities. For the following years (see Figure 1.1), precision studies at higher intensity are planned to fully exploit the LHC potentials.

Figure 1.1: LHC time schedule for the next decades, reporting the energy of the collisions and the integrated luminosity [10]. In January 2022, the schedule has been updated with long shutdown 3 (LS3) to start in 2026 and to last for 3 years.

1.2 ATLAS experimental apparatus

To cover the wide variety of physical processes, the apparatus provides redundant and complementary signatures using electron, muon, photon, jet and missing transverse energy measurements, thanks to a complex structure with sub-detectors of different technologies. In some analyses, more complex signatures that include τ and heavy flavour tags are also used. The ATLAS sub-detectors in Figure 1.2 can be grouped into three principal systems: the Inner Detector, the Calorimetry System and the Muon Spectrometer. Starting from the interaction point (at the centre of the ATLAS apparatus), the pixel detector, the semiconductor tracker and the transition radiation tracker form the Inner Detector that provides tracking and triggering measurements (described in Section 1.2.3). Progressively distant from the interaction point (IP), there are the Liquid Argon (LAr) electromagnetic calorimeter, the Tile calorimeter, the LAr hadronic End-Cap and forward calorimeters that constitute the Calorimetry System, measuring the energy loss of the particles (Section 1.2.4). Finally, the muon Barrel and End-Cap stations form the Muon Spectrometer (described in Section 1.2.5), as shown in Figure 1.2.

1.2.1 Coordinate system

The ATLAS global right-handed Cartesian coordinate system is individuated by the Z-axis along the LHC beam pipe direction and by the X-axis and Y-axis orientated and originated as shown on the left side of Figure 1.3. Cylindrical (r, ϕ, Z) and spherical (r, ϕ, θ) coordinate systems are alternatively used due to the ATLAS apparatus symmetry, and they are represented on the right side of Figure 1.3.

In the spherical system, the coordinates are defined as the following: $r = \sqrt{X^2 + Y^2}$ is the radial distance from the interaction point, ϕ (or second coordinate) is the azimuthal angle between r and

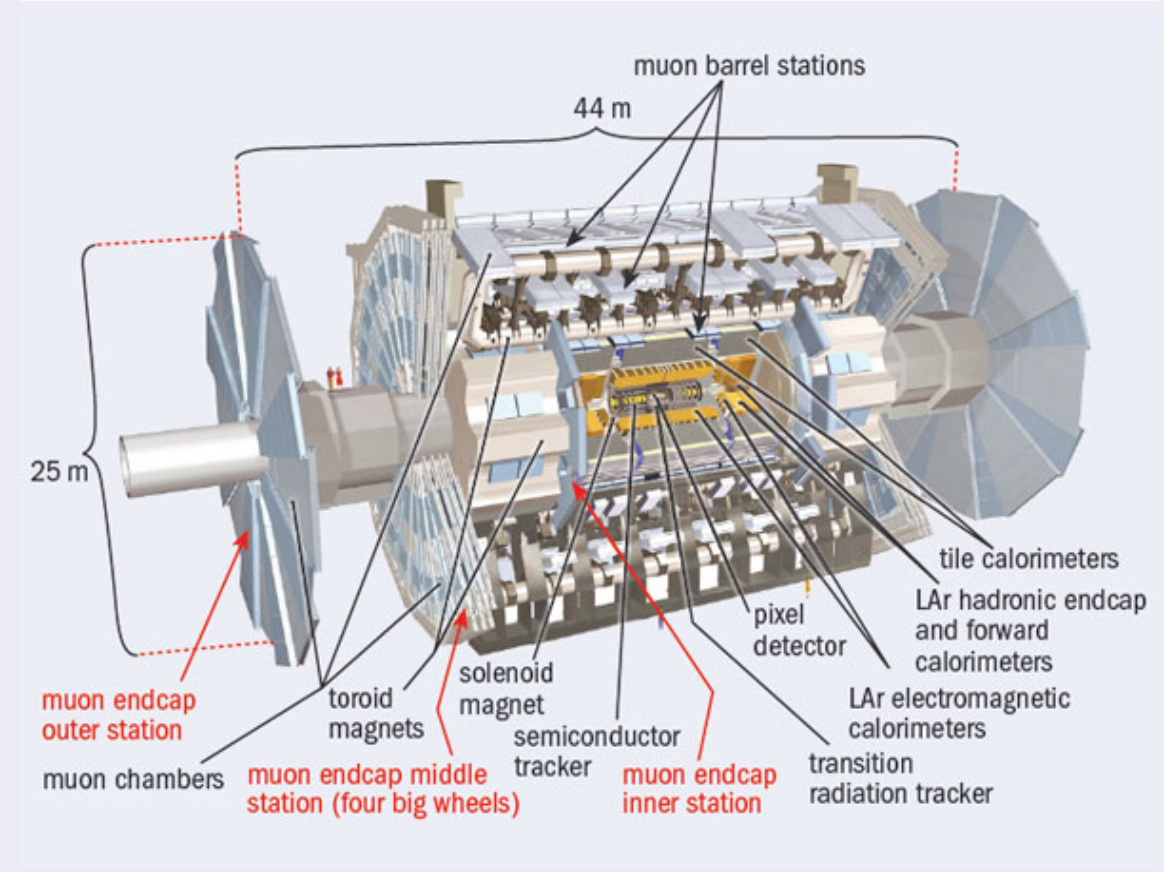


Figure 1.2: Sketch of ATLAS detector apparatus, pointing out the component positions of the three sub-detector systems: the Inner Detector, the Calorimetry System and the Muon Spectrometer (marked by red arrows the three stations of muon detectors on the End-Cap regions) [1].

the X-axis, and θ is the polar angle between the positive Z-axis and r. It is preferred to use the pseudorapidity η (or precision coordinate) defined as:

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right], \tag{1.1}$$

instead of the bare angle θ since the particle flux is uniform as a function of η.
The ATLAS Barrel region extends up to the range $|\eta| < 0.9$, while the End-Caps cover the pseudorapidity range $0.9 < |\eta| < 2.5$.
Beyond its geometrical definition, η can be interpreted as the relativistic limit of the rapidity y:

$$y = \frac{1}{2}\ln\left[\frac{(E + p_z)}{(E - p_z)}\right] \tag{1.2}$$

where E is the energy of a particle with momentum $\vec{p}$.
The energy and momentum of the particle are often projected in the transverse plane because the conservation laws can be applied in this plane. The transverse momentum p_T is defined as

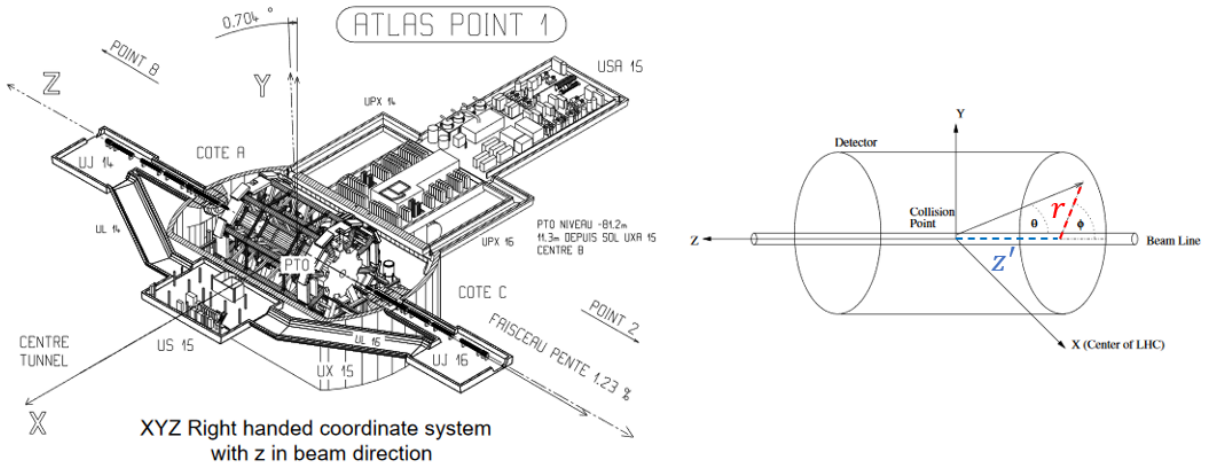


Figure 1.3: Sketch of ATLAS apparatus with the global coordinate system (left). The nominal interaction point individuates the origin of the coordinate system. The positive X-axis points from the interaction point to the centre of the LHC ring, and the positive Y-axis points upwards. Illustration of the ATLAS (r, ϕ, θ) coordinate system (right).

$\sqrt{p_X^2 + p_Y^2} = |p| \sin(\theta)$ and the transverse energy as $E_T = E \sin(\theta)$. The angular distance between two tracks ΔR is defined as $\sqrt{\Delta\eta^2 + \Delta\phi^2}$ in the pseudorapidity-azimuthal angle space.

1.2.2 Magnet system

The ATLAS magnet system gives the name to the experiment, and it distinguishes the ATLAS layout from the CMS experiment.
The magnet system provides the necessary bending power[1] on the particle path for the momentum measurements of charged particles. It includes a solenoid, Central Solenoid, and three large-scale air toroids, Barrel Toroid and End-Cap Toroids. Its total dimensions are 26 m along the beam direction and 22 m of the diameter in the transverse plane.
The Central Solenoid (CS) provides the magnetic field for the Inner Detector (ID), placed inside it (Figure 1.2). Its length is 5.8 m, and the inner and outer diameters are 2.46 m and 2.56 m.
The intensity of the CS field is 2 T at the interaction point, which decreases from ~ 1.8 T to ~ 0.9 T at the end of the Inner Detector cavity. Its material thickness is kept as low as possible (about 0.66 radiation lengths at normal incidence) to minimise the resulting energy loss of the particles at the entrance of the calorimetry system.
The Barrel Toroid (BT) consists of 8 coils (Figure 1.4). It extends over 25.3 m along the beam direction,

[1] The bending power is defined as

$$\int d\vec{l} \times \vec{B}$$

along the trajectory of the particle

with inner and outer diameters of 9.4 m and 20.1 m, respectively. It leads the overall dimensions of the ATLAS apparatus, and it envelops the calorimetry system, the CS and the Inner Detector.
According to the radial and azimuthal regions, the field intensity varies between 0.15 T and 2.5 T. The average toroidal field intensity is 0.5 T in the central regions of ATLAS.
The End-Cap Toroids (ECTs) are inserted at the BT ends to optimise the bending power in the End-Cap regions of the Muon Spectrometer (MS), as shown in Figure 1.4. Their length in beam direction is 5 m, and their inner and external diameters are 1.65 m and 10.7 m, respectively [1].
In the End-Cap regions, the average field intensity is 1 T, and the range of possible values goes from 0.15 T to 3.5 T.
In the ID, the uncertainties on its relative alignment dominate the bending power uncertainties rather

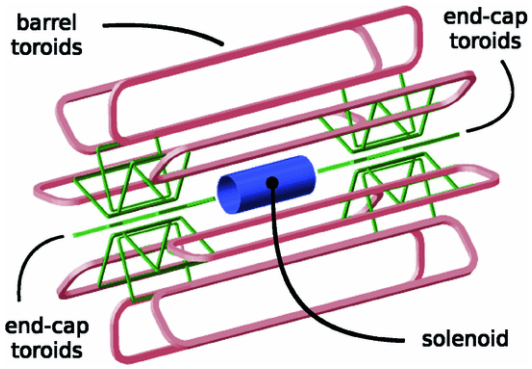


Figure 1.4: Sketch of the ATLAS Magnet System. The Central Solenoid is represented in blue, the 8 coils of the Barrel Toroid in red and the End-Cap Toroids are in green.

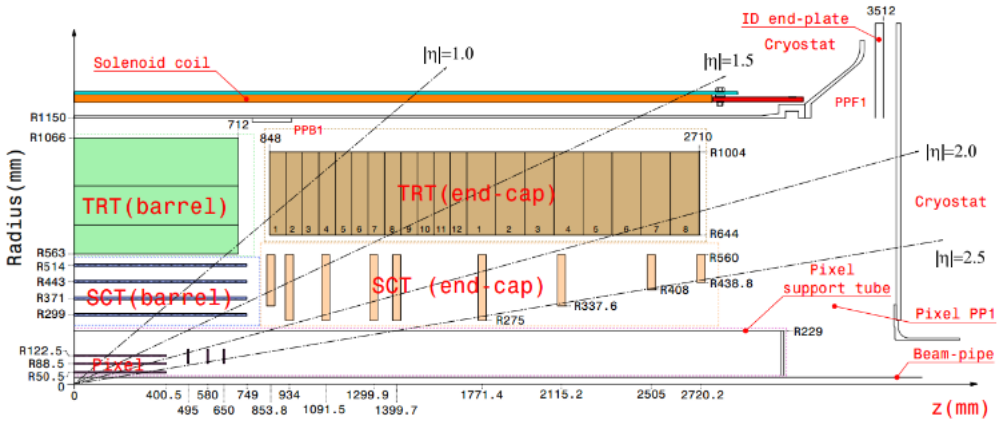
than the magnetic field uncertainties. The requirement of $\sim 5 \times 10^{-4}$ for relative uncertainty on bending power is reached by demanding a systematic alignment uncertainty of $< 0.1\%$ level of the sagitta for the most high-precision measurements in the ID [1]. While, in the Muon Spectrometer, the local uncertainties on the bending power introduce O(1%) fluctuations of the momentum resolution scale due to the non-uniform field (1 mT/mm of maximum field gradient). The single uncertainty contributions are comparable (O(1%)), and they are: the field measurement errors, the uncertainty on the distance between the muon chambers and the magnet coils and the trajectory measurement errors. Then, the total uncertainties on the measured muon curvature and the bending power establish that the maximum degradation of the overall momentum resolution has to be less than 5% anywhere in the MS.

1.2.3 Inner Detector

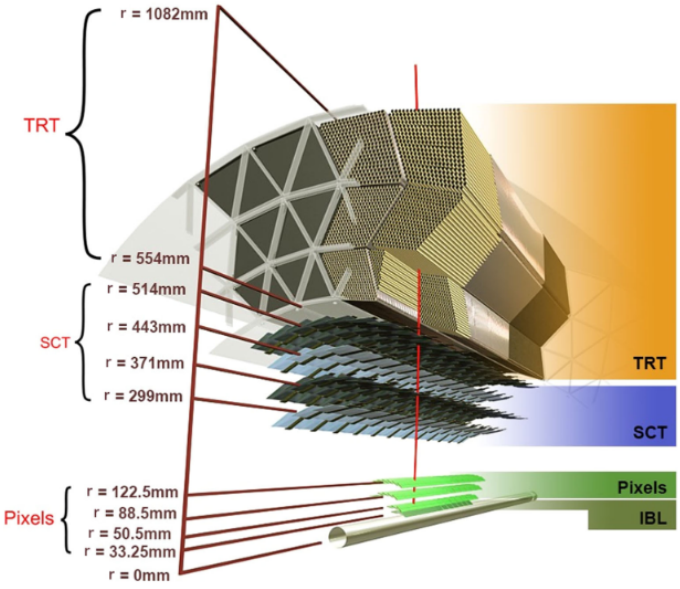
The ID provides particle identification and precise tracking of electric charged particles, from which the primary and secondary vertices are reconstructed. The ID covers the $\eta < 2.5$ range. Its overall length and radius are respectively ~ 7 m and ~ 1.08 m, as shown in Figure 1.5 (a).
For LHC Run 2, an additional tracking layer, insertable B-Layer (IBL), has been placed in the closest position to the pipe beam to improve the precision and robustness of the track reconstruction, as shown in Figure 1.5 (b). It consists of 14 staves with a total of 280 silicon pixel modules that provide a fully ϕ hermeticity for $p_T > 1$ GeV and an η coverage up to 3 [11] [12].
The (silicon) pixel detector consists of three Barrel layers and four disks on each basis of the cylinder, completing the angular coverage. The pitches of the $140{\times}10^6$ detector active elements (arranged in 1774 modules) are 50 μm in the $r\phi$ and 300 μm in the beam (z) direction, making possible the application of pattern recognition techniques in the LHC environment. The pixel detector provides three of the precision measurements over the full acceptance, and it determines the impact parameter resolution and the ability to track short-lived particles in ATLAS. Together with other detector parameters, the typical spatial resolutions over a flat η distribution are reported in Table 1.1 for the pixel detector and the other ID detector systems.

Table 1.1: The position, spatial resolution, number of readout channel and η coverage of the different systems of ID [13].

System	Position	Resolution (μm)	Channels (10^6)	η coverage
IBL	IP closest Barrel layer	$r\phi =< 10$, $z = 72$	12	± 3
Pixel	1 removable Barrel layer	$r\phi = 12$, $z = 66$	16	± 2.5
	2 Barrel layers	$r\phi = 12$, $z = 66$	81	± 1.7
	4 End-Cap disks on each side	$r\phi = 12$, $r = 77$	43	$1.7 \div 2.5$
Semiconductor	4 Barrel layers	$r\phi = 16$, $z = 580$	3.2	± 1.4
	9 End-Cap wheels on each side	$r\phi = 16$, $r = 580$	3.0	$1.4 \div 2.5$
TRT	Axial Barrel straws	170 per straw	0.1	± 0.7
	Radial Barrel straws	170 per straw	0.32	$0.7 \div 2.5$



(a)



(b)

Figure 1.5: A quarter-section of the ID layout in the z-r plane where z is the beam direction, showing the major detectors with their active dimension during LHC Run 1 (a). Zoom of the ID Barrel region, showing the radial coordinates of ID detector systems during LHC Run 2 (b).

In the intermediate radial range, four additional barrel layers and nine disks on each side of the silicon micro-strip detectors form the ATLAS semiconductor tracker (see Figure 1.5 (b)). They improve the precision of the $r\phi$ and z coordinates, contributing to the measurements of the momentum, the impact parameter and vertex position.
At the most external part of the ID, a straw-tube transition radiation tracker (TRT) adds 36 hits per track to the 7 space points of the pixel and silicon micro-strip detectors. Its measurements improve the pattern recognition. The TRT allows the implementation of an easy and fast Level-2 track trigger, too. The choice of a gas mixture with Xenon makes the ATLAS TRT able to detect transition radiation photons produced by the electrons in the radiator between the straws and, thus, to identify the electrons from the other charged particles. The p_T resolution is quite constant in the range $|\eta| < 1.4(1.7)$ while it degrades in the forward region due to the finite dimensions of the solenoid (CS), as shown in Figure 1.6.

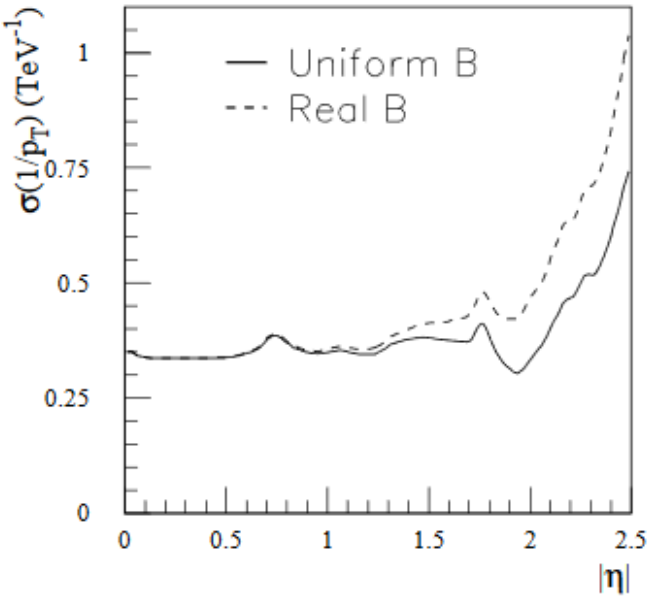


Figure 1.6: The Inner Detector transverse momentum resolution with beam constraint, for tracks with p_T = 500 GeV, for the real solenoidal field (dot line) compared to a uniform 2 T field (full line) [13].

1.2.4 Calorimeter system

The Calorimeter system (in Figure 1.7) covers the range $|\eta| < 4.9$ and the entire azimuthal angle, using different techniques to detect the wide variety of physics processes that occur after the proton collisions. It is placed outside the solenoid to avoid the shower widening due to its magnetic field.

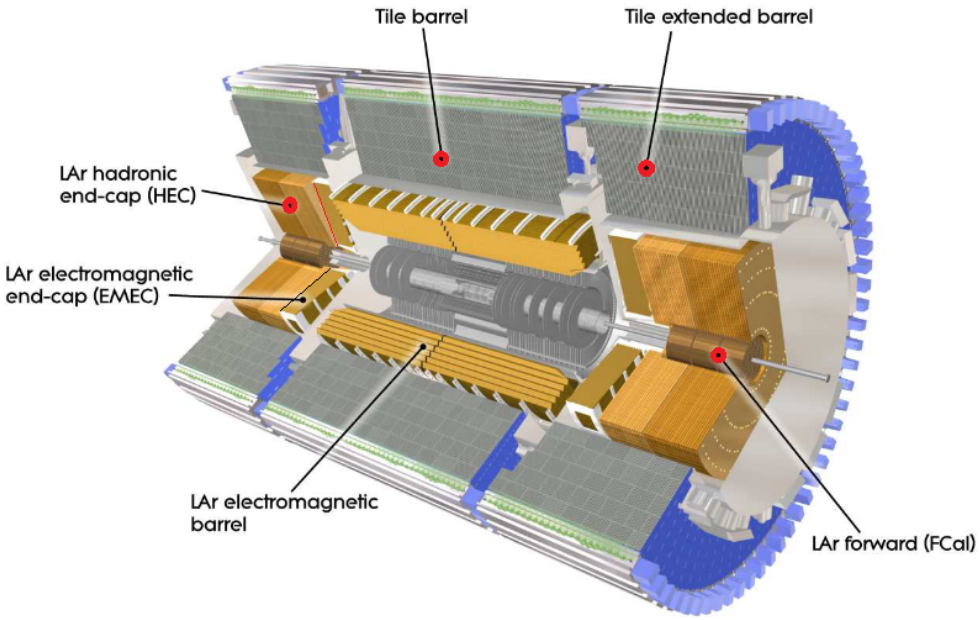


Figure 1.7: Sketch of the Calorimeter System, the black (red) marks indicate the electromagnetic (hadronic) components of the Calorimeter System.

It is formed by an electromagnetic (EMC) and a hadronic (HCal) calorimeter.
The EMC identifies electrons, positrons and photons and measures their energy. It is a sampling LAr calorimeter with lead as the passive absorber. The accordion geometry (visible in Figure 1.8) of its electrodes provides complete ϕ symmetry without azimuthal cracks. As also shown in Figure 1.8, the calorimeter layers have different $\Delta\eta \times \Delta\phi$ granularities that are optimised to assure the required precision on the electron and photon position for reconstructing the electromagnetic showers. For example, the EMC is divided into three layers with different segmentations in the Barrel region. The innermost layer (Layer 1 in Figure 1.8) has the finest segmentation along the η coordinate, except for the edge zones of the Barrel and End-Cap regions, as visible in Table 1.2. More distant from IP, the second layer (Layer 2) measures the largest fraction of the energy of the electromagnetic shower, and the third layer (Layer 3) collects only the tail of the electromagnetic shower and is, then, less segmented. The physical requirements for the cell size are less demanding in the End-Cap electromagnetic and hadronic calorimeters.

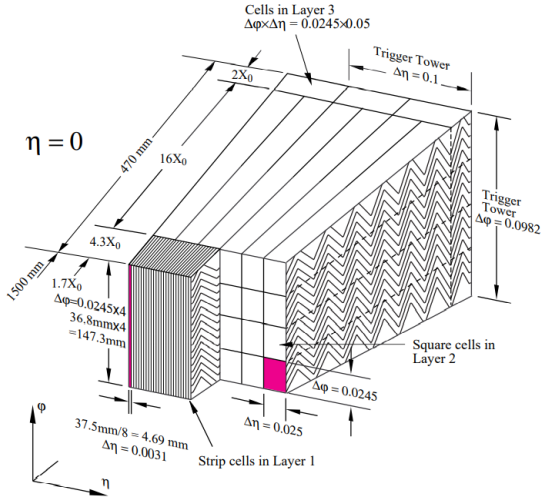


Figure 1.8: Sketch of a Barrel module where the accordion geometry in ϕ and the different layers are clearly visible. The granularity in η and ϕ of the cells of each of the three layers (see also Table 1.2) and of the trigger towers is also shown.

Table 1.2: Granularities of ATLAS EMC and HCal calorimeter cells [1]. The Barrel and End-Cap EMC calorimeters overlaps in the $1.375 < |\eta| < 1.425$ range.

Calorimeter Layer	η coverage	$\Delta\eta \times \Delta\phi$ Granularity
Barrel EMC Layer 1	$\|\eta\| < 1.40$	$0.025/8 \times 0.1$
Edge zones the Barrel EMC Layer 1	$1.40 < \|\eta\| < 1.475$	0.025×0.025
Barrel EMC Layer 2	$\|\eta\| < 1.40$	0.025×0.025
Edge zones of Barrel EMC Layer 2	$1.40 < \|\eta\| < 1.475$	0.075×0.025
Barrel EMCr Layer 3	$\|\eta\| < 1.35$	0.050×0.025
End-Cap EMC Layer 1 (overlap)	$1.375 < \|\eta\| < 1.425$	0.050×0.1
End-Cap EMC Layer 1 (overlap)	$1.425 < \|\eta\| < 1.5$	0.050×0.025
End-Cap EMC Layer 1	$1.50 < \|\eta\| < 1.80$	$0.025/8 \times 0.1$
	$1.80 < \|\eta\| < 2.00$	$0.025/6 \times 0.1$
	$2.00 < \|\eta\| < 2.40$	$0.025/4 \times 0.1$
	$2.40 < \|\eta\| < 2.50$	0.025×0.1
	$2.50 < \|\eta\| < 3.20$	0.1×0.1
End-Cap EMC Layer 2 (overlap)	$1.375 < \|\eta\| < 1.425$	0.050×0.025
End-Cap EMC Layer 2	$1.425 < \|\eta\| < 2.50$	0.025×0.025
	$2.50 < \|\eta\| < 3.20$	0.1×0.1
End-Cap EMC Layer 3	$1.50 < \|\eta\| < 2.5$	0.05×0.025
Barrel HCal Layer 1 and 2	$\|\eta\| < 1.0$	0.1×0.1
Extended Barrel HCal Layer 1 and 2	$0.8 < \|\eta\| < 1.7$	0.1×0.1
Barrel HCal Layer 3	$\|\eta\| < 1.0$	0.2×0.1
Extended Barrel HCal Layer 3	$0.8 < \|\eta\| < 1.7$	0.1×0.1
End-Cap HCal Layer 1,2,3 and 4	$1.50 < \|\eta\| < 2.5$	0.1×0.1
	$2.50 < \|\eta\| < 3.2$	0.2×0.2

The EMC thickness is > 22 radiation lengths in the Barrel and > 24 in the End-Cap regions [1], making a sufficient containment of the electromagnetic showers of the electrons (photons) with the energy of 7 (10 GeV in the case of photons) GeV up to a few TeV. The lead thickness has been optimised as a function of $|\eta|$ to reach the best performance in terms of energy resolution. The fractional energy resolution ($\frac{\sigma(\mathrm{E})}{E}$) of the EMC LAr is given by the relation:

$$\frac{\sigma\,(\mathrm{E})}{E} = \frac{10\% \cdot \sqrt{GeV}}{\sqrt{E}} \oplus 0.17\% \tag{1.3}$$

It is observed that the reconstructed energy response is linear within ± 0.1%, in the energy range of 15-180 GeV. At the lowest energy, a non-linearity of 0.8% is measured [1].
In the $|\eta| < 1.8$ region, an active LAr layer is placed as a presampler detector to correct for the energy lost by electrons/positrons and photons upstream of the calorimeter.
The HCal design provides hermetic jet and missing transverse energy measurements. The overall thickness is equal to 9.7 interaction lengths (λ_0) at $|\eta| = 0$ in the Barrel region and 10 λ_0 in the End-Cap regions. The outer support in the Barrel region adds other 1.3 λ_0 at $|\eta| = 0$ to the 9.7 λ_0 upstream the Muon Spectrometer, minimising the (no muon) particle leakage in MS.
The HCal is formed by the Tile calorimeter in the Barrel region and the LAr hadronic calorimeters in the End-Cap regions, as shown in Figure 1.7.
Covering the $|\eta| < 1.7$ range, the Tile calorimeter has scintillating tiles, as active layers, and steel, as the passive absorber. Its stand-alone fractional energy resolution has been evaluated on the 12% percentage of the Tile calorimeter modules as a function of the pion beam energy and impact angle [1]. At $\eta = 0.35$, it is:

$$\frac{\sigma\,(\mathrm{E})}{E} = \frac{56\% \cdot \sqrt{GeV}}{\sqrt{E}} \oplus 5.5\% \tag{1.4}$$

The LAr hadronic End-Cap calorimeters (EC-HCal) are next to the End-Caps of the EMC, and it covers the $1.5 < |\eta| < 1.8$. The passive absorbers of the EC-HCal are copper flat plates.
The fractional energy resolution is:

$$\frac{\sigma\,(\mathrm{E})}{E} = \frac{71\% \cdot \sqrt{GeV}}{\sqrt{E}} \oplus 5.8\% \tag{1.5}$$

It has been measured on a 25% percentage of the EC-HCal modules with pions with energies up to 200 GeV.
Further LAr forward calorimeters (FCal) are located in the same cryostats as the HCal, with copper in the innermost layer and tungsten in the outer layers as passive absorbers. The FCal cells cover the $3.1 < |\eta| < 4.9$ range, in which the particle flux is quite high. To keep the occupancy low and to avoid ion build-up problems, the LAr active gaps (close in tubes) are smaller than the usual 2 mm gap of the EMC in the Barrel region. Thanks to its fine transverse segmentation, the FCal fractional energy resolution with a pion beam with an energy of 200 GeV [1] is:

$$\frac{\sigma\,(\mathrm{E})}{E} = \frac{70\% \cdot \sqrt{GeV}}{\sqrt{E}} \oplus 3.0\% \tag{1.6}$$

1.2.5 Muon Spectrometer

The Muon Spectrometer is the most external part of ATLAS. It covers the range $|\eta| < 2.7$.
It measures the magnetic deflection (sagitta) of muons (and antimuons) with a p_T in the range from 10 to 3000 GeV that have escaped the calorimeter system, with a fractional transverse momentum resolution ($\frac{\sigma_{p_T}}{p_T}$) of 10 % at 1 TeV. The MS is used for pattern recognition in stand-alone or in combination with ID and for Level-1 trigger.
The three large scale air toroids (Barrel Toroid and End-Cap Toroids) provide the magnetic fields in the Muon Spectrometer (see Section 1.2.2 for the field intensity).
In particular, the Barrel Toroid provide the magnetic bending over the range $|\eta| < 1.4$. In the range $1.6 < |\eta| < 2.7$, muon tracks are subject to the fields of the End-Cap Toroids. In the range $1.4 < |\eta| < 1.6$, referred to as the transition regions, a combination of Barrel and End-Cap fields bends the muon tracks. This magnet configuration provides a field that is mostly orthogonal to the muon trajectories, minimising at the same time the degradation of resolution due to multiple scattering [1].
The MS is equipped with gaseous detectors (or chambers) that are placed according to Figure 1.9.
The detectors are grouped in sixteen Sectors in the r-ϕ plane, as shown on the right side of Figure 1.9.

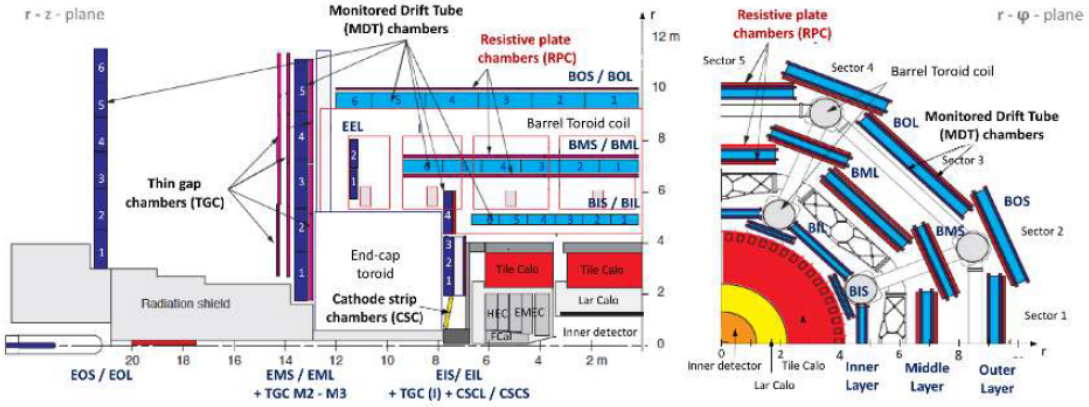


Figure 1.9: Longitudinal (left) and transverse (right) sections of Muon Spectrometer respect to the beam direction.

The odd (even) Sectors are referred to as Large (Small) Sectors.
In the Barrel region, the chambers are installed on three cylindrical (B) Layers (Inner, Middle and Outer on the left side of Figure 1.9) as the minimum multiplicity is three planes to measure sagittas. While, in the transition and End-Cap regions, the detectors are placed in three disks (per each side) perpendicular to the beam axis, referred to as muon (E) Inner, Middle and Outer Layers or stations (see Figure 1.2). The MS chambers are dedicated to precision tracking or trigger operations, as reported in Table 1.3.

An ATLAS MDT (Monitored Detector Tube) chamber contains two groups of 3 or 4 tube layers. Using the gaseous mixture Ar:CO_2(93%:7%) at 3 bar pressure, the average resolution is 80 μm at the level of a single tube and 35 μm at the level of a chamber (also reported in Table 1.3), respectively. The tube wires have been positioned in a configuration that allows measuring the precision coordinate

Table 1.3: The position, function, spatial or time resolution and η coverage of the different systems of MS [1]. The spatial resolution does not include chamber-alignment uncertainties. The reported time resolutions are the intrinsic chamber time resolutions.

Chamber	Position	Function	Resolution	$\lvert\eta\rvert$ coverage
MDT	Barrel	Precision Tracking	$z = 35\mu$m	2
	End-Cap	Precision Tracking	$r = 35\mu$m	$2 \div 2.7$
CSC	End-Cap	Precision Tracking	$r = 40\mu$m, $r\phi = 5$ mm	$2 \div 2.7$
RPC	Barrel	Level-1 Trigger,	< 2 ns,	± 1.05
		second coordinate	$r\phi = 10$ mm, $z = 10$ mm	1.05
TGC	End-Cap	Level-1 Trigger,	4 ns	$1.05 \div 2.4$
		second coordinate	$r\phi$ =(3-7) mm, r =(2-6) mm	$1.05 \div 2.7$

η in the Barrel region and the coordinate r in the End-Cap regions.
The CSC (Cathode-Strip Chamber) is a multiwire proportional chamber with strip-segmented cathodes. The strips have been orientated in orthogonal directions of the tracking plane, providing simultaneous measurements of the track coordinates in the two directions. Operating with the gas mixture Ar:CO_2(80%:20%), the spatial resolution of the CSC in the bending plane varies between 40 μm and 60 μm , and the spatial resolution in the transverse plane is 5 mm, according to the different pitches in the readout planes. The stand-alone CSC intrinsic time resolution is 7 ns.
As reported in Table 1.3, the trigger-dedicated detectors also provide measurements of the second coordinate ϕ with an angular resolution of 1-3 mrad.
The RPC (Resistive Plate Chamber) is a gaseous parallel electrode-plate detector, with two orthogonal series of capacitively coupled strips. The η-strips are parallel to the Barrel MDT wires and the ϕ-strips are orthogonal to the MDT wires providing the second-coordinate measurement.
The RPCs operate in avalanche mode using the gas mixture[2] $C_2H_2F_4$:iC_4H_{10}:SF_6 (94.7%:5%:0.3%). They have a time resolution below 2 ns [15], a detection efficiency per layer beyond 98.5% with a local rate capability of $\sim$1 kHz/cm^2.
The TGC (Thin Gap Chamber) is a multiwire proportional chamber with planar geometry. The CO_2:n-pentane(55%:45%) gas mixture prevents the occurrence of streamers in all operating conditions, and it assures good resolutions and efficiency for the trigger and tracking functions of the TGCs.
Groups of anode wires provide redundant measurements of the bending coordinate (η). The wire plane is among two cathode planes with different readout geometries. One of the cathodes has readout strips that run perpendicular to the wires. The strips measure the azimuthal coordinate ϕ with a spatial resolution of $3 \div 7$ mm corresponding to an azimuthal granularity of $2 \div 3$ mrad. The other cathode has readout pads that are involved in L1 trigger operations. The TGCs have a time resolution of 4 ns. Even if they have a worse time resolution than RPCs, the TGCs provides better efficiency at a very high hit rate, increasing the robustness of the trigger system at large values of rapidity.

[2] A small amount (8000-12000 ppm) of water vapour is added to reach a mixture with a relative humidity between 35% and 45% and, thus, maintain a constant resistivity of the Bakelite electrodes [14]. Also, the gas mixture of the MDT detectors is humidified (about 300 ppm) to improve the high voltage stability [1].

In the innermost region of the End-Cap Inner stations, the CSCs and the MDTs are arranged together with thin gap chambers (TGCs) in detector structures called Small Wheels. While the Big Wheels are equipped with only TGCs and MDTs, and they are in the End-Cap Middle stations.

During the ATLAS subsystem installation, the initial positioning accuracy has been approximately established. The MS precision tracking detectors are installed with a precision of about 5 mm and 2 mrad with respect to their nominal positions, together with their alignment objects. However, a $\leq 30\,\mu$m overall accuracy on the tracking chamber locations along the muon trajectory is required to achieve a 45 μm sagitta resolution. To measure the tracking chamber alignment, the detectors are equipped with optical sensors. The tracking chambers in three Barrel Layers form towers in a projective optical system, shown in Figure 1.10 (a).
In the End-Cap regions, the projective system can not be adopted because the cryostat vessels of the toroid magnets would obstruct the light path between the inner and the middle layer. Then, a polar optical system is defined in the three End-Cap Layers per side in which the tracking chamber position is determined in the grid shown in Figure 1.10 (b). Combining the optical sensor systems with

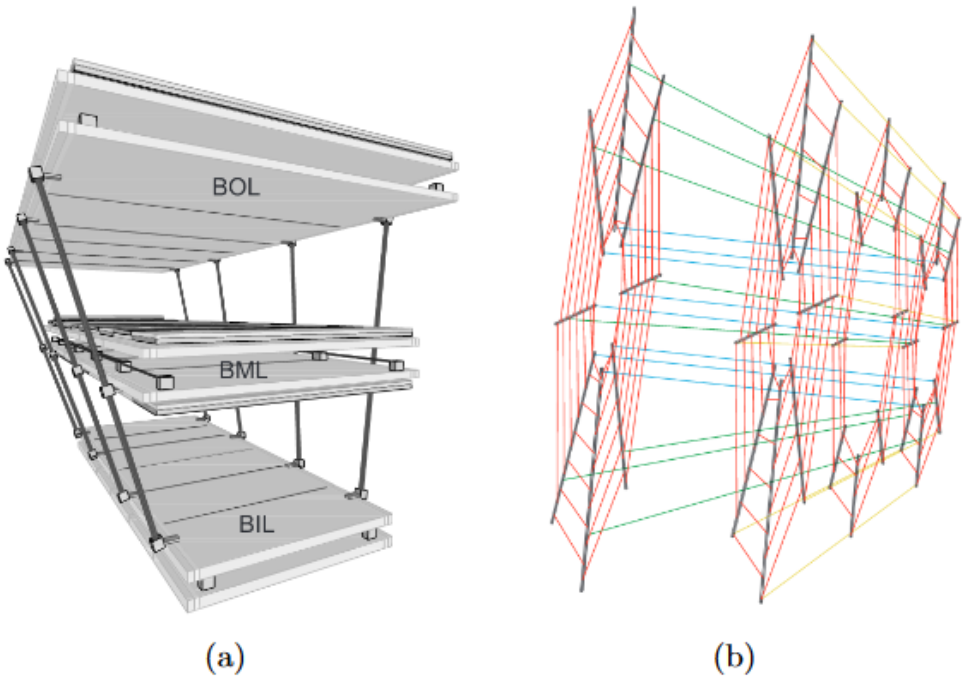


Figure 1.10: Three dimensional view of the projective alignment system of one half octant where a tower is defined by the chambers on the three MS Barrel Layers that cover part of the solid angle with origin in the interaction point (a). Polar optical system grid of used in an MS End-Cap region where alignment bars are sketched in black and light rays in red (b).

track-based alignment algorithms, the projective optical system finally detects the relative variations in the tracking chamber position with 20 μm accuracy in the Barrel region. It can also determine the absolute chamber position with a few hundred microns. The final accuracies on the relative and absolute positions are 40 μm and 300 μm in the End-Cap regions [1]. Considering the average

between the Barrel and End-Cap regions, the overall 30 μm requirement is satisfied. Then, the relative alignment of the Barrel and forward regions of the muon spectrometer, calorimeters and inner detector is determined by using high-momentum muon trajectories.

1.3 Trigger system

The ATLAS multi-level trigger system has a pipelined front-end readout and a hierarchical data acquisition architecture. It is divided into three levels: Level-1 (L1), Level-2 (L2) and the event filter. The trigger system after L1 event selection is also referred to as High-Level Trigger (HLT).
The L1 trigger has < 2.5 μs latency time with a maximum counting rate of 100 kHz. Based on customised electronics, it selects events among the colliding LHC events that occur every 25 ns (at 40 MHz). The L1 searches the signatures of high-p_T muons, electrons, photons, jets and τ-leptons decaying hadronically [1]. It also selects events with large total transverse energy and missing transverse energy.
For the case of the muons, the muon L1 hardware trigger is based on hit coincidences between different RPC or TGC detector layers inside programmed geometrical windows (roads) with reduced granularity, which define the muon transverse momentum p_T. The full calorimeter system is exploited for the electrons, photons, τ-leptons, jets and missing transverse energy.
Especially for Higgs boson and B-hadrons physics processes, an addition L1 trigger, named topological trigger (L1Topo), has been implemented to improve the background rejection and signal event acceptance. It is based on FPGAs that take the topological decision to accept or reject events using different criteria, such as isolation requirements, angular relations, missing transverse energy and invariant masses. Its latency is ~200 ns.
Different from the L1 trigger that uses reduced-granularity information, the HLT algorithms use the full granularity and precision of calorimeter and muon chamber data, as well as the data from the inner detector, to refine the trigger selections. Indeed, the L2 trigger processes the information coming from the Regions-of-Interest (RoIs) [16], limiting the amount of data that must be transferred from the detector readout. The RoI is a region of the detector where the L1 trigger has identified possible trigger objects within the event. The L2 trigger reduces the rate to 3.5 kHz with a maximum latency time of 40 ms.
The event filter uses offline analysis on full-reconstructed events to select the event that will be stored. It reduces the event rate to approximately 200 Hz, with an average event processing time of order 4 s.

1.4 Higgs boson Physics

The Standard Model (SM) is currently the accepted theory to describe three of the fundamental interactions in a unified way. Excluding the gravitational force, it includes the electromagnetic, weak and strong interactions inside a quantum field model. It provides observables and predictions that can be compared with the results observed by the experiments at colliders. Among the SM input parameters, there is the mass of the Higgs boson [17] [18]. The first measurement of the Higgs boson mass ($m_H = 125.09 \pm 0.24$ GeV) has been made by the ATLAS [3] and CMS [4] experiments, and it

is compatible with the predicted limits of the SM model in the Figure 1.11.

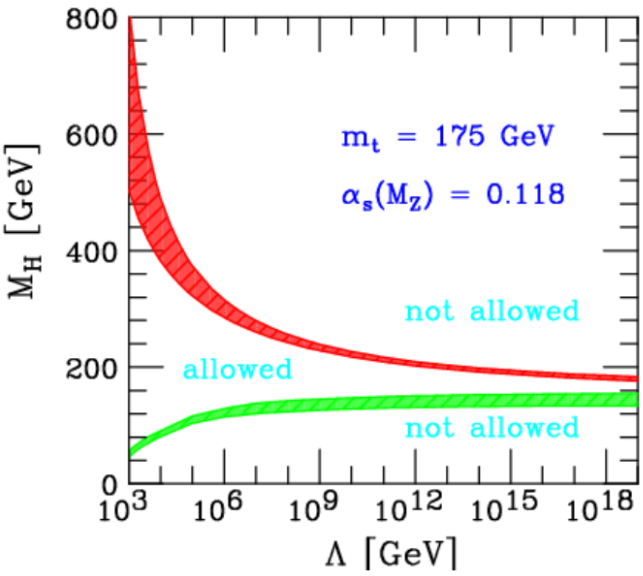


Figure 1.11: Theoretical upper and lower limits on the Higgs mass in the SM valid up to cut-off scale Λ. m_t is the mass of the top quark, α_S is the QCD coupling and M_Z = 91.187 GeV is the mass of Z boson [19].

As visible in Figure 1.12, the most dominant production mechanism of the Higgs boson at LHC is the gluon-gluon fusion. The relative Feynman diagram of leading order is shown in Figure 1.13 (a) and it includes a loop on virtual quarks. As the Yukawa coupling between the Higgs boson and the quarks depends on the square of quark masses, the only sensitive diagrams are with virtual bottom and top quarks at LHC energy. The gluon-gluon fusion (ggF) production mechanism has the highest sensitivity when the Higgs boson decays in leptons since the decay leptons are exploited in the trigger selection to discriminate ggF events with respect to QCD background.
The second most probable production mechanism is the vector boson fusion (VBF), where two quarks radiate virtual V bosons that fuse and produce the Higgs boson in the final state of the Feynman diagram in Figure 1.13 (b). Independently on the Higgs boson decay mode, the final state of VBF production processes has a distinct signature represented by the two jets coming from the quarks that have radiated the vector bosons.
Other production processes are the associated production with a vector boson (VH) or with a $t\bar{t}$ top pair ($t\bar{t}H$). The leading-order Feynman diagrams are reported in Figures 1.13 (c) and 1.13 (d). These two last production mechanisms are also interesting for the direct couplings between the Higgs boson and a weak vector boson (Figure 1.13 (c)) or top quark (Figure 1.13 (d)).
The branching ratios of the Higgs boson decays are shown on the left side of Figure 1.14 as a function of the Higgs boson mass. For the measured Higgs boson mass [21], the principal decay channels are $H \to b\bar{b}$, $H \to WW^*$ together with the processes where the Higgs boson decays in a pair of gluons, Z bosons and τ leptons (ZZ and $\tau\tau$). Their values are listed in the table on the right side of Figure 1.14.
The decay mode $H \to b\bar{b}$ has the largest branching ratio (58%). Its final state is marked by the presence of two reconstructed b-jets. In the gluon-gluon fusion, this decay channel of the Higgs boson

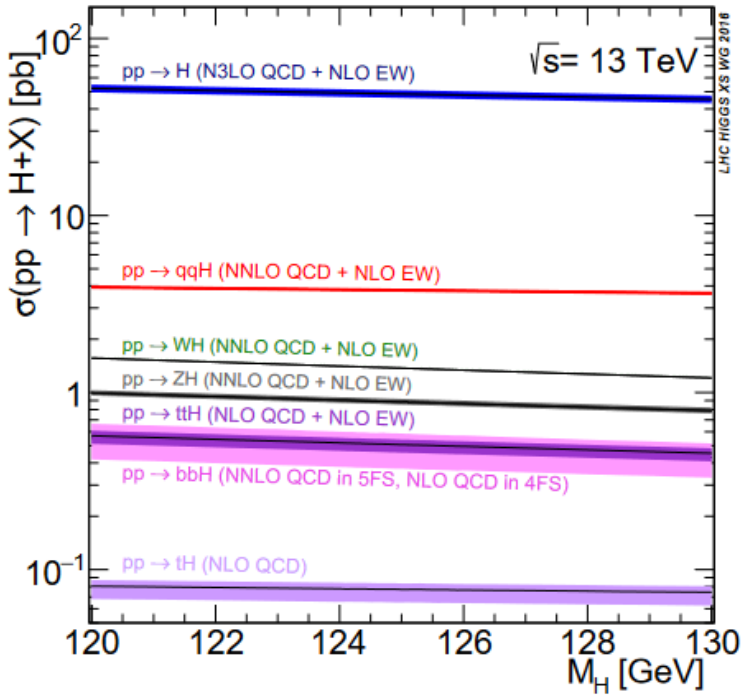
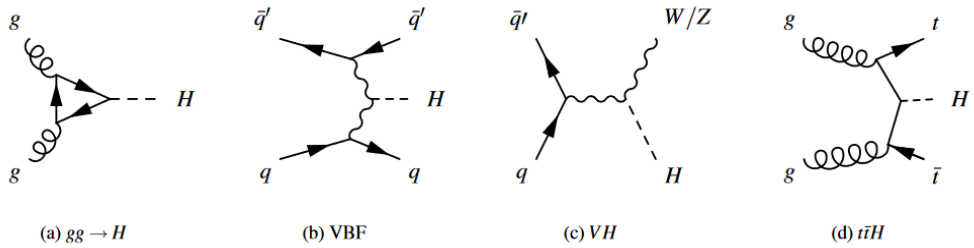


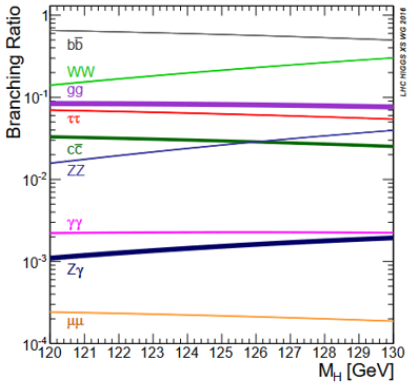
Figure 1.12: Standard Model Higgs production cross-sections. The gluon-gluon fusion is marked in blue, the vector boson fusion in red, the associated production with W boson in green, the associated production with Z boson in grey and the associated production with $t\bar{t}$ in purple [20].



(a) $gg \to H$ (b) VBF (c) VH (d) $t\bar{t}H$

Figure 1.13: Leading-order Feynman diagrams for the gluon-gluon fusion (a), vector boson fusion (b), associated production with a vector boson (c) and $t\bar{t}$ top pair (d) mechanisms of the Higgs boson production.

suffers from a huge background from direct b-quark production. For this reason, the study of $H \to b\bar{b}$ is combined with the production modes associated with a vector boson V or a $t\bar{t}$ top pair, as well as with the VBF production mode.
In the case of the VH production mode, the selected final states have at least a charged lepton (i.e. $ZH \to e^+e^-b\bar{b}$, $ZH \to \mu^+\mu^-b\bar{b}$ and $W^\pm H \to e^\pm \nu b\bar{b}$, $W^\pm H \to \mu^\pm \nu b\bar{b}$) that provides a significant reduction of the multijet background. After the event selection, the dominant background processes are $V + jets$, $t\bar{t}$, single top and diboson processes [23].
The decay mode $H \to WW^*$ is the second most probable (from the table on the right side of Figure



Decay channel	Branching fraction
$H \to b\bar{b}$	5.81×10^{-1}
$H \to WW^*$	2.15×10^{-1}
$H \to gg$	8.18×10^{-2}
$H \to \tau^+\tau^-$	6.26×10^{-2}
$H \to ZZ^*$	2.64×10^{-2}
$H \to \gamma\gamma$	2.27×10^{-3}
$H \to Z\gamma$	1.54×10^{-3}
$H \to \mu^+\mu^-$	2.17×10^{-4}

Figure 1.14: The SM Higgs boson branching ratios as a function of Higgs boson mass M_H (left) [20]. The values of the SM Branching Ratios at $M_H = 125.09$ GeV (right) [22].

1.14), and it is studied using its leptonic final states. The sensitivity relative to the leptonic final states of $H \to WW^*$ is better than $H \to b\bar{b}$, but it is not as favourable as $H \to ZZ \to 4l$ due to the presence of neutrinos that do not allow the kinematic closure of the process. The relevant backgrounds after the event selection are $t\bar{t}$, single top and diboson processes.
Less probable but with clean signatures, the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ are the discovery channels of the Higgs boson in 2012 [21].
The process $H \to \gamma\gamma$ corresponds to a narrow peak in the diphoton invariant mass distribution over the irreducible backgrounds ($qq \to \gamma\gamma$, $qg \to q\gamma\gamma$ and $gg \to \gamma\gamma$ processes) and the reducible backgrounds consisting of the inclusive prompt photon (γ-jet) and multijet production processes with one or more jet(s) misidentified as a photon(s).
The $H \to ZZ^* \to 4l$ with $l = e, \mu$ has the cleanest signature [24]. A narrow four lepton invariant mass peak has been observed on top of a smooth background. The main irreducible background is $ZZ \to 4l$ process. The reducible backgrounds from $Z + jets$ and $t\bar{t}$ add to the irreducible background. Their rejection is provided by tight lepton isolation cuts.

To add more accurate constraints to the Standard Model or to probe new physics, all the ATLAS analyses are aiming to provide newly updated measurements exploiting the full integrated luminosity (~ 139 fb^{-1}) statistics of data collected by ATLAS during Run 2 proton-proton collisions at 13 TeV centre-of-mass energy.
Specific to the cross-section measurements of the Higgs boson production mechanisms, all the related analyses are releasing the differential results to update the total Higgs production cross-section.
The expectations for the new measurements using the entire ATLAS Run 2 dataset are pretty high since the analyses on partial Run 2 data statistics have already provided new observations compatible with SM, other than confirming the results of Run 1.
Limiting the discussion on the possible combinations of the production and decay modes of Higgs boson, cited before, most of them have already provided results from ATLAS Run 2 pp data corresponding to luminosities higher than 36 fb^{-1}, and $VH \to V(WW^*)$ and $t\bar{t}H$ groups are finalising their analysis with higher statistics. First observations of the $H \to b\bar{b}$[23] and $H \to WW^*$ decays

[25] have been possible with observed significance greater than 5 σ, by exploiting the most sensitive combinations of the production mechanisms.
Direct observation of the Yukawa coupling between the Higgs boson and the top quark has also been made by studying the $t\bar{t}H$ production mechanism [26] in the analyses of the single decay channels: $H \to b\bar{b}$ in Reference [27], multilepton final states (WW, ZZ, $\tau\tau$) and their combination in Reference [28] and with $H \to \gamma\gamma$ in Reference [29].
As well as the combined $V(H) \to V(b\bar{b})$ results on the 79.8 fb^{-1} ATLAS Run 2 dataset with the $V(H) \to V(ZZ^* \to 4l)$ and $V(H) \to V(\gamma\gamma)$ results on 36.1 fb^{-1} ATLAS Run 2 dataset have given the first VH production mechanism observation with an observed (expected) significance of 5.3 (4.8) σ [23].
The most recent combination of the ggF, VBF, VH and $t\bar{t}H$ production cross-section measurements is shown in Figure 1.15. The combined measurement leads to an observed significance above 5 σ

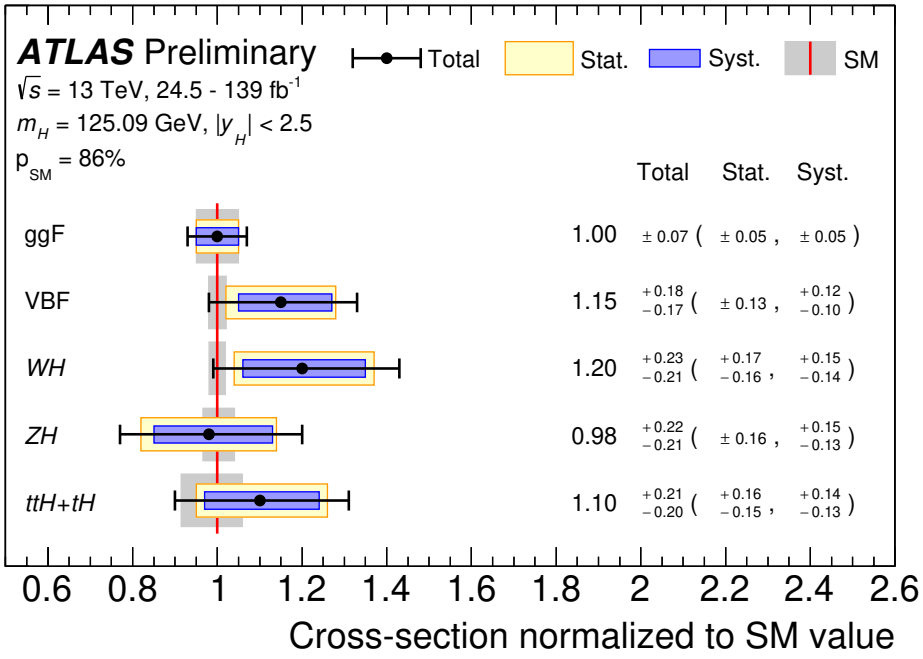


Figure 1.15: The ggF, VBF, WH, ZH and $t\bar{t}H$+tH production cross-sections normalised to their SM predictions. They are measured by assuming SM values for the decay branching fractions. The level of compatibility between the measurement and the SM prediction corresponds to a *p-value* $p_{SM} = 86\%$ [30].

for all the Higgs boson production mechanisms considered in the fit. It is compatible with the SM predictions like the observed global signal strength $\mu = 1.06 \pm 0.07$, defined as the measured Higgs boson signal yield normalised to its SM prediction.
The integrated luminosities corresponding to the most recent analyses of the Higgs boson production mechanisms studied with the discussed decay modes are summarised in Table 1.4.

Table 1.4: The decay channels, target production modes and integrated luminosity used for each analysis of the combination between decay channel and production modes. The $t\bar{t}H(4l)$ and $t\bar{t}H(\tau_{had}\tau_{had})$ analyses do not have overlapping event selection criteria with the $t\bar{t}H$ analysis in Reference [28].

Analysis decay channel	Target Production Model	Integrated Luminosity [fb^{-1}]
[31]	VBF	126
$H \to b\bar{b}$ [32]	VH	139
[33]	$t\bar{t}H$	36.1
[34]	ggF, VBF	139
$H \to WW^*$ [35]	VH,	36.1
[28]	$t\bar{t}H$	36.1
$H \to \tau\tau$ [36]	$ggF, VBF, VH, t\bar{t}H(\tau_{had}\tau_{had})$	139
[28]	exclusive $t\bar{t}H$	36.1
$H \to \gamma\gamma$ [37]	$ggF, VBF, VH, t\bar{t}H$	139
$H \to ZZ^*$ [38]	$ggF, VBF, VH, t\bar{t}H(4l)$	139
[28]	exclusive $t\bar{t}H$	36.1

2 Analysis of $WH \rightarrow WWW^* \rightarrow l\nu l\nu l\nu$

The analysis of the associated Higgs boson production with a vector boson, VH, with $H \rightarrow WW^*$ decay is performed using events with two ($2l$), three ($3l$) or four ($4l$) charged leptons (l are only electrons and/or muons) in the final state to select the WH and ZH processes.
Up to Section 2.4, the sections concern all the channels of $VH \rightarrow VWW^*$ analysis. The other sections of this chapter only deal with the analysis of the $WH \rightarrow WWW^* \rightarrow l\nu l\nu l\nu$ channel with at least a pair of leptons with Same Flavour (electronic or muonic) and Opposite electric charge Sign ($SFOS$), which is the subject of this thesis.

2.1 VH production with subsequent $H \rightarrow WW^*$

The process of the associated production of the Higgs boson with a vector boson is deeply studied by the ATLAS collaboration to provide more precise measurements of the quantities that describe the Higgs boson like mass, coupling constants and cross-sections, as discussed in general in Section 1.4. The present work is part of the analysis that studies the VH production mechanism in combination with the $H \rightarrow WW^*$ decay. The overall analysis aims to update the VH production cross-section measurements by using the data acquired by ATLAS during the Run 2 pp collisions at 13 TeV centre-of-mass energy and corresponding to 139 fb^{-1}. In particular, this thesis focuses on updating the signal strength parameter μ, defined as the ratio between the measured signal yields and the predicted signal yields by the SM, for a specific sub-channel of the Higgs boson production associated with the W boson. From the previous analysis of the Run 2 proton-proton data acquired by ATLAS apparatus at LHC with a centre-of-mass energy of 13 TeV and corresponding to the integrated luminosity of 36 fb^{-1} [35], it has been measured:

$$\mu_{WH} = 2.3^{1.1}_{-0.9}(\text{stat})^{+0.49}_{-0.36}(\text{exp syst})^{+0.41}_{-0.3}(\text{theo syst}) = 2.3^{+1.2}_{-1.0} \tag{2.1}$$

for the $(W)H \rightarrow (W)WW^*$ channel with the leptonic final state.
It has also been combined with the strength parameter μ_{ZH} of the $(Z)H \rightarrow (Z)WW^* \rightarrow lll\nu l\nu$ process to evaluate the combined measurement,

$$\mu_{VH} = 2.5^{+0.8}_{0.7}(\text{stat})^{+0.30}_{0.23}(\text{exp syst})^{+0.37}_{0.26}(\text{theo syst}) = 2.5^{+0.9}_{-0.8}. \tag{2.2}$$

The tree-level Feynman diagrams of the $(W)H \rightarrow (W)WW^* \rightarrow l\nu l\nu l\nu$ and $(Z)H \rightarrow (Z)WW^* \rightarrow lll\nu l\nu$ processes are reported in Figure 2.1.
To update the measurements μ_{WH} and μ_{ZH}, the semileptonic final state of $(V)H \rightarrow (V)WW^*$ with two charged leptons is also considered. In the analysed processes with two charged leptons in the final

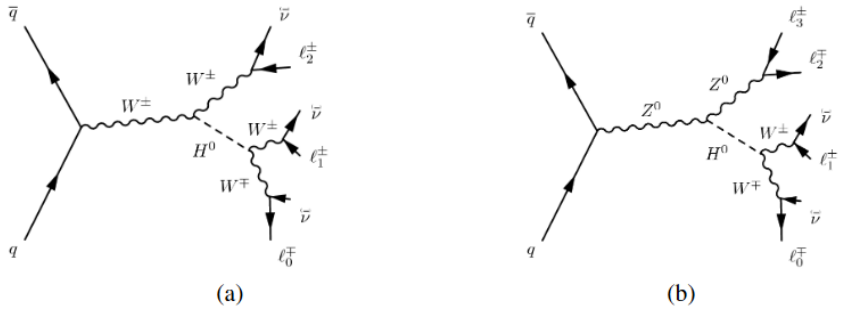


Figure 2.1: Tree-level diagrams for $VH \rightarrow V(WW^*)$. $3l$ WH channel (a) and $4l$ ZH channel (b).

state, one of the involved vector bosons decays hadronically giving two jets in the final state. The possible tree-level Feynman diagrams that contribute to this final state are reported in Figure 2.2.
For the case of the Higgs boson production associated with W boson (WH), there are two contributions: the diagram of Figure 2.2 (a) in which the two quark jets come from the decay of the Higgs boson, and the diagram of Figure 2.2 (b) in which the two quark jets come from the associated W boson.
The final state with two charged leptons of the Higgs boson production associated with Z boson (ZH) has only one corresponding tree-level diagram in which the two quark jets come from the associated Z boson (in Figure 2.2 (b)).

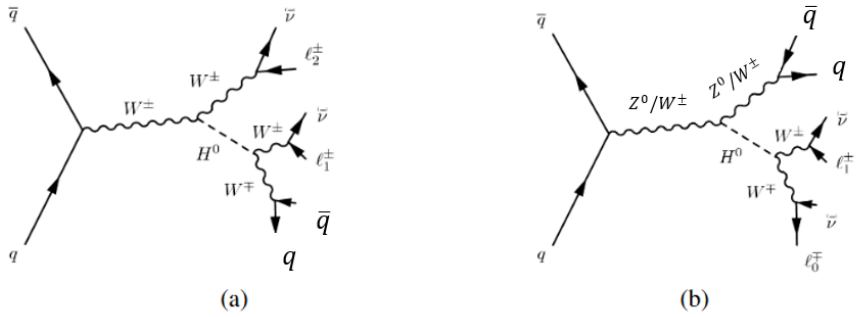
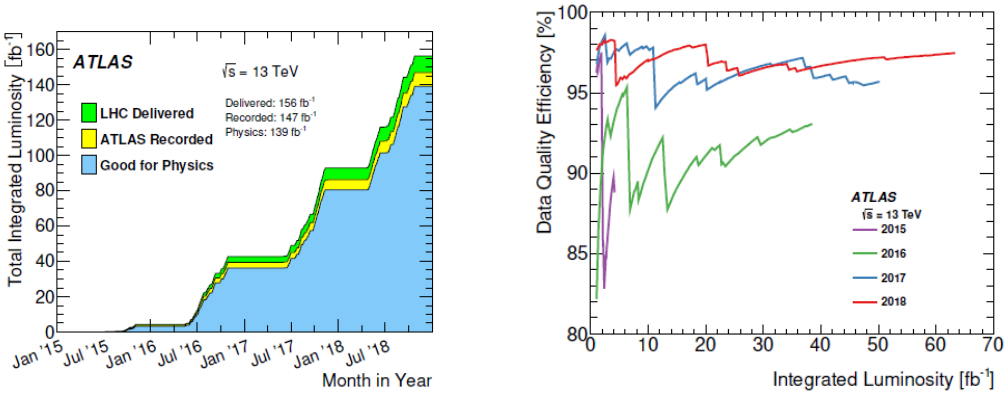


Figure 2.2: Tree-level diagrams for $VH \rightarrow V(WW^*)$ with semileptonic final state with two charged leptons (muons/antimuons and/or electrons/positrons). The diagram (a) represents the contribution in which W boson, from the decay of the Higgs boson, hadronically decays and the diagrams (b) represent the contributions in which the associated vector W or Z boson hadronically decays.

2.2 Data sample for the analysis of the $WH \rightarrow W(WW^*)$ channel

During LHC proton-proton Run 2 at 13 TeV centre-of-mass energy, the ATLAS apparatus has collected data corresponding to a final integrated luminosity of 139 fb^{-1} (as shown in Figure 2.3 (a)) with an uncertainty of 1.7% [39]. The data have been acquired in the "Standard" Run 2 machine configuration, consisting of 25 ns bunch-spacing proton beams.

The data quality efficiency in Figure 2.3 (b) is expressed in terms of a luminosity-weighted fraction of the good quality data recorded during stable beam periods (dedicated to the physics analysis). It has improved during the Run 2, with 88.8% in 2015, 93.1% in 2016, 95.7% in 2017 and 97.5% in 2018. Combining these values, the overall efficiency for the full Run 2 is 95.6% [39].



(a) The integrated Luminosity of ATLAS as a function of the Run 2 time period

(b) The ATLAS data quality efficiency as a function of the integrated Luminosity of ATLAS

Figure 2.3: The integrated Luminosity of ATLAS as a function of time (a) and the ATLAS data quality efficiency as a function of the integrated Luminosity of ATLAS (b) for the LHC Run 2 [39].

2.2.1 Pile-up

The LHC beams are organised in bunches. For LHC proton-proton Run 2, the maximum number of bunches is 2556, with a bunch intensity of 1.2×10^{11} protons per bunch and a stored intensity in the machine of 3×10^{14} protons [40]. The bunch spacing is 25 ns at the regime. When the bunches cross in the ATLAS interactions point, several collisions are acquired per bunch crossing. During LHC proton-proton Run 2, the ATLAS apparatus has measured a maximum instantaneous luminosity of 2.1×10^{34} cm^{-2} s^{-1}.

The undesired interactions are referred to as pile-up interactions. There are five contributions to the pile-up in ATLAS. They are: *in-time pile-up*, *out-of-time pile-up*, *cavern background*, *beam halo events* and *beam gas events*.

The *in-time* and the *out-of-time* pile-up are the dominant contributions. The *in-time* pile-up is due to the additional collisions that occur in the same bunch crossing of the collision of interest. The *out-of-time* contribution is from collision before or after the bunch crossing of the collision of interest. Its relevance depends on the detector technology and the integration time of the readout electronics. The other contributions come from the LHC background of neutrons and photons (*cavern background*), from the interactions of the beam protons with the collimators (*beam halo events*) and from the interactions between the beam protons and the residual gas in the beam pipe (*beam gas events*).

During the overall Run 2, the mean number of total ATLAS pile-up interactions is ~34, as shown in Figure 2.4. The average number of the interactions per bunch crossing (μ) and the number

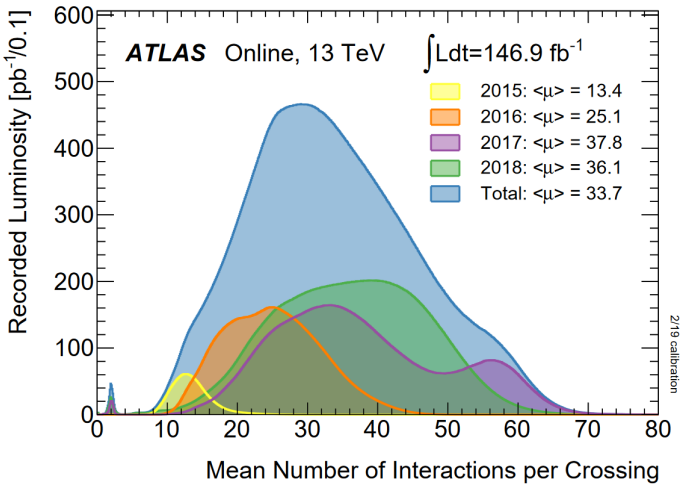


Figure 2.4: ATLAS pile-up distribution during LHC Run 2 [41].

of primary vertices (N_{PV}) are used to investigate the pile-up conditions because N_{PV} reflects the amount of the *in-time* pile-up and $< \mu >$ is proportional to the total inelastic proton-proton cross-section.

2.3 Object selection

The selection of the candidate signal events is based on offline requirements on several objects, like vertices, electrons, muons and jets, that have been reconstructed in an ATLAS event. This section is organised in subsections, in which the specific requirements of $VH \rightarrow VWW^*$ analysis are reported for each object.

2.3.1 Primary vertex

The primary vertices are the reconstructed locations of the inelastic proton-proton interactions from which all the other objects come. An object coming from a primary vertex is labelled as a prompt object. Conversely, a non-prompt object is from secondary vertices.
A vertex is extrapolated by a set of reconstructed tracks in an event. The vertex reconstruction essentially comprises two components: the association of the tracks to a particular vertex seed (vertex finding) and the vertex fitting.
The $VH \rightarrow VWW^*$ analysis selects events with at least one primary vertex with at least two associated tracks, each with transverse momentum $p_T > 500$ MeV.
In an event with more than one reconstructed primary vertex, the vertex seed with the largest $\sum {p_T}^2$ over all the tracks is included in the final vertex fitting procedure.

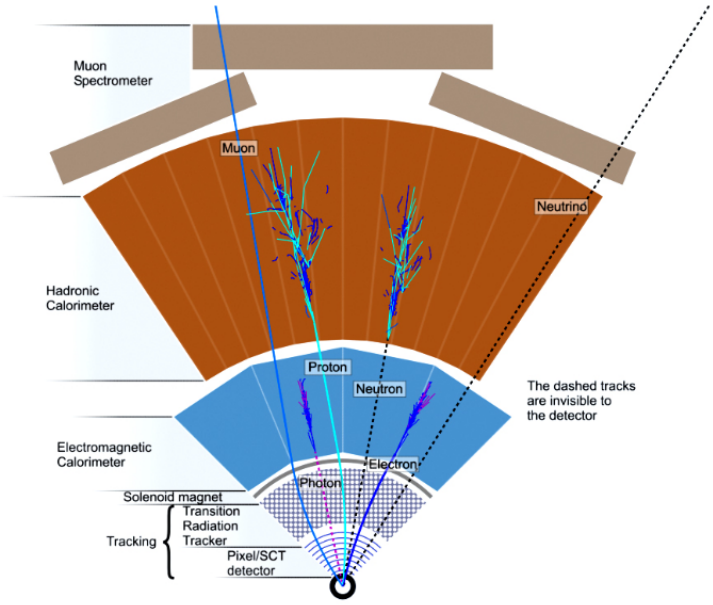


Figure 2.5: Sketch of the ATLAS particle identification according to the combination of the sub-detectors that receive a signal.

2.3.2 Leptons

The objects that give a track in the Inner Detector associable with an energy deposition in the Electromagnetic Calorimeter are identified as electrons, as shown in Figure 2.5. For the electrons, further and more rigorous identification criteria are based on the calorimeter electromagnetic shower shape, the quality of the match between the track and the cluster and the amount of transition radiation in the TRT of the ID. While muons are associated with combined tracks having hits in the Inner Detector and Muon Spectrometer (see Figure 2.5). Different algorithms are used to combine the information from the ID, calorimeters and MS, taking into account the different acceptances of ATLAS detector systems. There are four possible types of reconstructed muon: "Combined (CB)","Segment-tagged (ST)", "Calorimeter-tagged (CT)" or "Extrapolated Muon (EM)"[42].
Different levels of identification efficiencies can be selected for the leptons. An electron can be "Loose", "Medium", or "Tight", corresponding to $\sim 96\%$, $\sim 94\%$,$\sim 88\%$ of identification efficiency, measured for an electron with a transverse energy of 100 GeV.
For muons, four different identification levels are possible, as reported in Table 2.1.
In the $VH \rightarrow V(WW^*)$ analysis, all sub-channels use the “Tight” identification level for electrons with the exception of the $ZH \rightarrow Z(WW^*) \rightarrow lll\nu l\nu$ channel. In this channel, the electrons with $p_T > 25$ GeV pass the “Medium”, while the “Tight” level is kept on the electrons with $p_T < 25$ GeV. Muon candidates have to pass the "Middle" identification level in all sub-channels.
Other than selecting the identification level, the $VH \rightarrow V(WW^*)$ sub-channels impose requirements on the leptons to individuate well-isolated charged leptons coming from a primary vertex.
To improve the event purity, the lepton pseudorapidity η and transverse momentum p_T are limited in selected ranges. For electrons, the selected η range is $|\eta| < 2.47$ (coverage of ID), excluding the

Table 2.1: Efficiency of prompt muons from W boson decays for low ($4 < p_T < 20$ GeV) and high ($20 < p_T < 100$ GeV) momentum muons for candidates with $|\eta| < 2.5$. The statistical uncertainties are negligible.

	$4 < p_T < 20$ GeV	$20 < p_T < 100$ GeV
Loose	96.7%	98.1%
Medium	95.5%	96.1%
Tight	89.9%	91.8%
High-p_T	78.1%	80.4%

transition region $1.37 < |\eta| < 1.52$ between the Barrel and End-Cap electromagnetic calorimeters, and p_T has to be larger than 15 GeV. For the muons, their pseudorapidity has to belong to the $|\eta| < 2.7$ range and their $p_T > 15$ GeV.
Requirements on the absolute longitudinal impact parameter (z_0) and the transverse impact parameter significance $\left(\frac{|d_0|}{\sigma_{d_0}}\right)$ maximise the association with a primary vertex, and they are reported below:

- $\left|z_0 \sin(\theta)\right| < 0.5$ mm;
- $\frac{|d_0|}{\sigma_{d_0}} < 5(3)$ for electrons (muons).

To reject leptons originating in decays of the *c/b*-hadrons (secondary vertices), specific isolation criteria are applied to the leptons in the final states of interest since these hadrons are generally embedded within particle jets with a large number of tracks and calorimeter energy deposits close to the leptons not originated from a primary vertex (non-prompt leptons).
Except for the $ZH \rightarrow Z(WW^*) \rightarrow lll\nu l\nu$ channel, the others adopt the newly developed multivariate isolation method "PromptLeptonImprovedVeto" (PLIV). Three separate Boosted Decision Tree (BDT) classifiers have been trained to discriminate prompt and non-prompt leptons. Two BDTs are reserved for prompt and non-prompt electrons in Barrel and End-Cap regions. Third BDT is for the prompt and non-prompt muon discrimination.
The excluded ZH channel uses looser isolation criteria.
On the electrons, an additional veto is applied to reject the electrons from gamma conversions.
In the following sections, the leptons that pass the identification and isolation criteria are also indicated as *id* leptons. Conversely, the leptons that satisfy looser isolation criteria with respect to the *id* lepton are referred to as *anti-id* leptons. The *anti-id* leptons are only exploited to estimate the misidentified lepton background in Section 2.6. In the rest of the cases, the leptons are *id* objects.

2.3.3 Jets

A jet is a collimated spray of particles arising from the fragmentation and hadronisation of partons after an inelastic collision. The jet reconstruction can be performed using algorithms that admit as inputs electromagnetic topological (EMTopo) or Particle Flow (PFlow) objects [43]. The EMTopo jet reconstruction only exploits calorimeter-based energy information. While the PFLow algorithms also include the ID information in the jet reconstruction. Specifically, the PFlow objects pass a cell-based energy subtraction algorithm that removes overlaps between the momentum and energy measurements

made in the inner detector and calorimeters, respectively. After that, the different pieces of information are combined according to the $|\eta|$ range to reconstruct hadronic jets and soft activity. In the $|\eta| < 2.5$ range that is covered by ID, the inputs are the tracks that match with a primary hard-scattering or pile-up vertex and the calorimeter topological clusters[1]. In the $|\eta| > 2.5$ forward regions, the inputs of the algorithm are the topological clusters with significant energy depositions. After the buildings of jets, some corrections are applied to take into account: the calorimeter non-compensation, the energy loss in the non-instrumental regions, the signal losses due to residual electronics noise and pile-up contributions, together with a residual correction based on *in-situ* measurements.
Compared with the EMTopo techniques, the PFlow approach improves the energy and angular resolutions of jets in the central region of the detector, as reported in Reference [43]. Including the vertex information in the jet reconstruction also improves the suppression of jets due to pile-up interactions.
For these reasons, the $VH \rightarrow VWW^*$ analysis uses the PFlow objects in the jet reconstruction. The adopted jet reconstruction algorithm is the anti-k_t sequential algorithm [44]. All the sequential clustering algorithms proceed by identifying the smallest between the distances d_{ij} and d_{iB}.
The first distance d_{ij} is the minimum distance between the particle j with transverse momentum p_T^j and the building-up cluster i with p_T^i, defined as:

$$d_{ij} = min\left(p_{Ti}^a, p_{Tj}^a\right) \times \frac{R_{ij}^2}{R}, \tag{2.3}$$

where $R_{ij} = \sqrt{\left(\eta_i - \eta_j\right)^2 + \left(\phi_i - \phi_j\right)^2}$ is the angular distance between the entities i and j, and the parameter R is related to the jet radius. While d_{iB} is the distance between the building-up cluster i and the beam axis.
If the smallest is d_{ij}, the algorithm recombines entities i and j. While, if the smallest is d_{iB}, it recognises the entity i as a jet and removes i from the list of entities.
The anti-k_t algorithm fixes the exponential parameter $a = -2$ in the definition d_{ij} in Equation 2.3. The $VH \rightarrow V(WW^*)$ analysis uses the anti-k_t algorithm with parameter $R = 0.4$ to reconstruct jets from PFlow objects.
It also requires that the jet pseudorapidity belongs to $|\eta| < 4.5$ and the jet transverse momentum is greater than 30 GeV.
The analyses use dedicated tagging algorithms for the jet and vertex association and for the identification and reconstruction of jets originating by long lifetime hadrons containing b quarks (b-jets).
In the present analysis, the jets and vertices are associated by using the multivariate discriminator Jet Vertex Tagger (JVT) [45] in the central region and the optimised version, forward JVT (fJVT), in the forward regions [46]. Specific to the jets with $p_T < 60$ GeV and $|\eta| < 2.4$, a tight level requirement is applied to the JVT discriminant to suppress the pile-up contribution and assure a good purity. The forward fJVT is applied with a loose working point to jets with $p_T < 60$ GeV.
The b-jet tagging algorithms are based on the b-jet topology that includes displaced tracks originating from a secondary vertex, as reported in Figure 2.6.
In particular, the DL1r algorithm [47] is used in the $VH \rightarrow V(WW^*)$ analysis. It is based on a neural

[1] A topological cluster is formed by the signals of topologically connected cells of the calorimeters that are combined on the basis of the energy significance $\frac{E_{cell}}{\sigma_{cell}}$ at the High Level Trigger stage to improve the rejection of electronics noise and other contributions like pile-up.

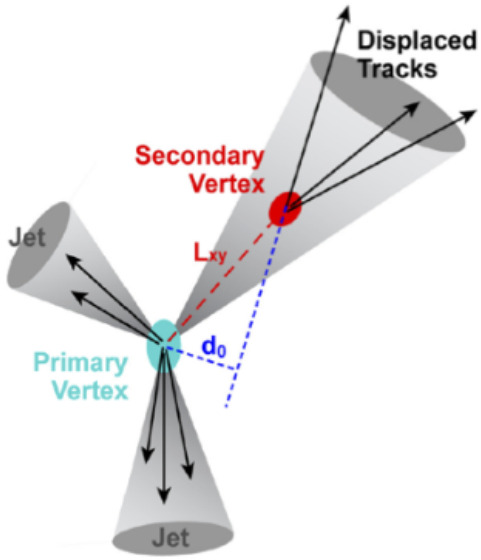


Figure 2.6: Sketch of a b-jet signature with displaced tracks from a secondary vertex (in red). The impact parameter d_0 and transverse decay length L_{xy} are reported in blue dot line and red dot line, respectively.

network multi-classifier to discriminate b-jets, c-jets or $light$-flavour jets. The multi-classifier input variables are taggers themselves that have been optimised to target the specific b-jet topology and kinematics.
For example, the peculiarity of the DL1r algorithm is to include the RNNIP [47] tagger that discriminates the large b quark transverse momentum fraction, the high multiplicity of the b quark decay and the large impact parameters due to the long lifetime of the $b \rightarrow c$ oscillation.
The other input taggers of the DL1r algorithm consider the properties of the secondary vertices and c-hadron flight path.
In the $VH \rightarrow V(WW^*)$ analysis, the reconstruction of b-jets is limited in the $|\eta| < 2.5$ range. Using simulated $t\bar{t}$ events, the b-tagging efficiency is determined, and it has to be 85%. The jets with $20 < p_T < 30$ GeV are only included in the b-jet counting, whereas they are considered under-threshold in the global counting of the jets as the $VH \rightarrow V(WW^*)$ sub-channels require that jet $p_T > 30$ GeV.

2.3.4 Missing Transverse Energy

The ATLAS apparatus indirectly detects the neutrinos. The neutrinos produce a momentum imbalance in the transverse plane.
This imbalance is evaluated as the negative vectorial sum of the transverse momenta of all calibrated selected objects, representing the hard terms, and tracks compatible with the primary vertex and not matched to any objects, which are the soft term [48].
Then, the components of the transverse momentum imbalance are:

$$E^{miss}_{x(y)} = -\left(\sum p^{electrons}_{x(y)} + \sum p^{photons}_{x(y)} + \sum p^{muons}_{x(y)} + \sum p^{jets}_{x(y)} + \sum p^{soft}_{x(y)}.\right) \tag{2.4}$$

The absolute value of the momentum imbalance in the transverse plane is referred to as missing transverse energy,

$$E_T^{miss} = \sqrt{\left(E_x^{miss}\right)^2 + E_y^{miss}}^2 . \tag{2.5}$$

Its azimuthal angle can be expressed as:

$$\phi^{miss} = \arctan \frac{E_y^{miss}}{E_x^{miss}} \tag{2.6}$$

In the E_T^{miss} definition, the objects are reconstructed from mutually exclusive detector signals with an ordered sequence: electrons, photons, hadronically decaying τ leptons, jet, while the muons are not involved in the overlapping because their reconstructions are principally based on the combinations between ID and MS tracks. The demand for calibrated objects in the hard term and the compatibility with primary vertices in the soft term strongly suppresses pile-up contributions in E_T^{miss}.

2.4 Monte Carlo sample

2.4.1 General overview

Monte Carlo generators reproduce detailed aspects of the final states in individual scattering events. In Quantum Mechanics, the differential cross-section of a scattering process can be factorised in a phase space term, which takes into account the normalisation factor and the quadrimomentum conservation, and a Lorentz-invariant term, invariant Matrix Element (ME), that includes the dynamics. The ME can be expanded in terms of Feynman diagrams, and it can be numerically integrated by Monte Carlo methods.
The ME of an inelastic scattering at LHC is usually evaluated up to the next-leading order (NLO).
In a hard scattering between partons, like the scattering processes in LHC, the involved partons can radiate virtual gluons that can themselves emit further gluons or produce quark-antiquark pairs. This radiation chain leads to the formation of parton showers that connect the coloured scale of the initial hard scattering with the hadronisation scale.
Then, the Monte Carlo methods are also used to model the parton showers, describing their QCD evolution as a chain of processes in which a parton can or can not split into two partons and simulating the initial and final radiation.
The Underlying Events, i.e. particle productions that are not associated with the hard parton processes, are also described in the parton shower models because they may have not a negligible impact on the experiment. They are the multiple scattering and exchange of multiple partons between the protons of the initial state.
To complete the simulation, a *hadronisation* or *jet fragmentation* model is applied, which makes the quark and gluon from the hard scattering (at LHC), the parton shower and multiple scattering processes into colour-neutral final states that can be compared with experimental data.
Different Monte Carlo (MC) generators are used to simulate the processes during the analysis up to the next-leading, NLO, or next-next-to-leading (NNLO) order of the perturbative expansion, by exploiting the individual generator characteristics.
Among the adopted generators in the analysis, Pythia8 [49] can model high-energy collisions, predicting the evolution from a few-body hard process to multi-hadronic final states. Powheg (Positive Weight Hard-est Emission Generator) [50] is a method to interface parton-shower generators with NLO QCD computations. The Sherpa (Simulation of High-Energy Reactions of PArticles) [51] generator can treat intermediate unstable resonances (cascade decays) with high-multiplicity final states. It also includes NLO QCD corrections to a given scattering process. MadGraph5_aMC@NLO [52] permits to compute any arbitrary observables in an arbitrary process at the tree-level or NLO precision, with or without the matching to parton showers, based on a set of process-independent building blocks. The EvtGen [53] generator is used for the simulation of b-quark and c-quark decays in the present analysis.

2.4.2 Monte Carlo samples for $VH \rightarrow V(WW^*)$

In Table 2.2, the used generators are reported per process together with the corresponding cross-section times $H \rightarrow WW^*$ branching ratio and the precision of the inclusive cross-sections applied to the

sample normalisation.

Table 2.2: MC generators used to model the signal and background processes, and corresponding cross-sections (given for $m_H = 125$GeV in the case of the Higgs boson production processes) times branching ratio (Br). When the leptonic decay filter is applied on W/Z bosons the quoted cross-sections include the branching ratios and are summed over lepton flavors. The column "Precision $\sigma_{\text{incl.}}$" gives the precision of the inclusive cross-section applied to the sample.

Process	Generator	$\sigma \cdot$ Br (pb)	Precision $\sigma_{\text{incl.}}$
$WH\,H \to WW^*$	Powheg +Pythia8 (MinLO)	0.293	NNLO
$ZH\,H \to WW^*$	Powheg +Pythia8 (MinLO)	0.189	NNLO
ggF $H \to WW^*$	Powheg +Pythia8 NNLOPS	10.4	N3LO+NNLL
VBF $H \to WW^*$	Powheg +Pythia8	0.808	NNLO
Inclusive $Z/\gamma^{(*)} \to \ell\ell$ ($40 \geq m_{\ell\ell} \geq 10$ GeV)	Sherpa 2.2.1	6.80×10^3	NNLO
Inclusive $Z/\gamma^{(*)} \to \ell\ell$ ($m_{\ell\ell} \geq 40$ GeV)	Sherpa 2.2.1	2.107×10^3	NNLO
$(W \to \ell\nu)\gamma$ ($p_{\text{T}}^{\gamma} > 7$ GeV)	Sherpa 2.2.8	1071	NLO
$(Z \to \ell\ell)\gamma$ ($p_{\text{T}}^{\gamma} > 7$ GeV)	Sherpa 2.2.8	353	NLO
$t\bar{t}$ di-leptonic (e, μ, τ)	Powheg +Pythia8	87.6	NNLO+NNLL
Wt leptonic	Powheg +Pythia8	7.55	NLO
$t\bar{t}V$ leptonic	MadGraph5_aMC@NLO+Pythia8	0.67	NLO
tZ leptonic	MadGraph5_aMC@NLO+Pythia8	0.21	NLO
tWZ leptonic	MadGraph5_aMC@NLO+Pythia8	0.016	NLO
$q\bar{q}/g \to WW \to \ell\nu\ell\nu$	Sherpa 2.2.2	12.7	NNLO
$q\bar{q} \to WWq\bar{q} \to \ell\nu\ell\nu jj$	Sherpa 2.2.2	0.18	NNLO
$gg \to WW \to \ell\nu\ell\nu$	Sherpa 2.2.2	0.60	NLO
$q\bar{q}/g \to WZ/W\gamma^{(*)} \to \ell\nu\ell\ell$	Sherpa 2.2.2	7.54	NLO
$q\bar{q}/g \to ZZ/Z\gamma^{(*)} \to \ell\ell\ell\ell$	Sherpa 2.2.2	2.71	NLO
VVV	Sherpa 2.2.2	0.11	NLO

Following the order in Table 2.2, the Higgs boson production processes with successive $H \to WW^*$ decay are simulated with Powheg interfaced with Pythia8. They are generated assuming 125 GeV as Higgs boson mass. The expected cross-sections and branching ratios of the LHC Higgs Working group [54] are adopted in the sample normalisation.
In particular, the Powheg MinLO interfaced with Pythia8 [55] gives predictions at the NLO in the cases of the Higgs boson production in association with a vector boson with 0 and 1 jet. NNLO QCD and NLO EW corrections are applied to the cross-sections of these events. The predictions for gluon-gluon fusion production mode are generated at the NNLO precision for 0-jet events, at the NLO precision for 1-jet events and at the LO precision for 2-jets events, with Powheg NNLOPS interfaced with Pythia8 [56].
The ggF has the most precise inclusive normalisation with an N^{3}LO+NNLLO (Next-Next-to-Leading Logarithmic Order) precision [57].
The vector boson fusion production is simulated by means of Powheg interfaced with Pythia8. Its cross-section used in the sample normalisation includes EW corrections up to NLO precision and QCD corrections up to an approximated NNLO precision.
Only the associated Higgs boson production with a W boson is considered as signal in the $WH \to W(W^-W^+)$ analysis and in this thesis.
Sherpa 2.2.1 is used to simulate the $Z/\gamma^{(*)}$ process in two intervals of the invariant mass m_{ll} of the leptons from Z boson or γ conversion. The samples are normalised using cross-sections calculated

with an NNLO precision.
In $VH \rightarrow V(WW^*)$ analysis, the $Z/\gamma^{(*)}$ process is only simulated in the $2l$ sub-channels with two charged leptons from the W boson decay in the final state. As well as, the $V\gamma$ events are generated with the NLO precision for events with 0 or 1 jet by means of Sherpa 2.2.8 in the sub-channels where γ is not a source of misidentified electrons. The photons are required to have transverse momenta $p_T > 7$ GeV and angular distances[2], ΔR, from V boson greater than 0.1. An additional requirement is applied to the mass m_{ll} of the leptons that has to be greater than 2 GeV.
In the $WH \rightarrow W(WW^*)$ analysis with three leptons in the final state, the Z+jet and $V\gamma$ processes are estimated using data-driven methods.
Powheg interfaced with Pythia8 is also used for the $t\bar{t}$ production and the single top production Wt events. The $t\bar{t}$ samples include a filter to require that the W bosons decay leptonically, and they are normalised using cross-sections calculated with a NNLO+NNLL precision. For the Wt process, at least two charged leptons are required in the final state.
The $t\bar{t}V$, tZ and tWZ processes are generated using MadGraph5_aMC@NLO interfaced with Pythia8. For the tZ process, it is required to have three charged leptons in the final state and the invariant mass of the two leptons from the Z boson decay greater than 10 GeV to include the off-shell Z boson case.
The diboson processes (WZ, WW and ZZ) are simulated using the Sherpa 2.2.2 generator. The events with 0 or 1 jet are generated with an NLO precision while events with greater jet multiplicity with a LO precision. For these samples, the cross-sections are normalised with an NLO precision.
The Sherpa 2.2.2 generator is also used for the triboson processes. The events with 0 jet are simulated up to the NLO precision and events with greater jet multiplicity up to the LO precision.
Different Parton Distribution Functions (PDF) and tunes are used according to the MC generators and considered process.
The CT10 and NNPDF 3.0 parton distribution function sets [58] are used for the hard scattering processes that are simulated through Powheg interfaced with Pythia8, with the exception of the Higgs boson production processes for which the PDF4LHC15 PDF set [59] is used.
Finally, NNPDF 3.0 NNLO (with dedicated sets of tuned parameters for the parton shower [60]) and NNPDF3NLO [61] PDF sets are respectively used for Sherpa and MadGraph5_aMC@NLO (version 2.3.3).
The AZNLO [62] tune is used for the diboson and signal processes, and the A14 tune [63] is for the other processes.
All the simulated MC samples include the effect of pile-up (defined in Section 2.2.1). To model the pile-up, the Monte Carlo samples are reweighted with the data scale factor μ to match the distribution of the number of pile-up interactions observed in the data. For the analysis of the $VH \rightarrow V(WW^*)$ channel, the applied data scale μ is 1/1.03.
To avoid a statistical dilution of samples, two different strategies are used for the MC reweighting of the 2015-2016 dataset and the 2018-2019 dataset. For the 2015-2016 dataset, the distribution of the inelastic collisions over all LHC bunches is considered in the MC/data matching. While the distribution of the inelastic collisions of each LHC proton bunch is considered in the MC/data matching for the 2017-2018 dataset.
The MC events are processed through the Geant-4 [64] based ATLAS detector simulations, which reproduce the interactions with detector material, electromagnetic fields and the electronic contributions

[2] The angular distance $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ is defined in Section 1.2.1.

to translate them into realistic signals. Finally, they are reconstructed by the ATLAS reconstruction software, such as the real data.

Until this section, the cited object selections and Monte Carlo analysis are in common among all $VH \rightarrow V(WW^*)$ sub-channels. From the following section, the discussion will focus on one of the two signal regions of $WH \rightarrow W(WW^*)$ channel with three leptons in the final state, as anticipated at the beginning of this chapter.

2.5 Reconstruction of WH events in $WWW^* \rightarrow l\nu l\nu l\nu$

The entire reconstruction of the $WH \rightarrow WWW^* \rightarrow l\nu l\nu l\nu$ channel (Figure 2.1 (a)) with three charged *id* leptons in the final state is presented in this section, referring to the common requirements in Section 2.3.2.
The selected WH events in $WWW^* \rightarrow l\nu l\nu l\nu$ decay are required to have at least a primary vertex, three isolated high transverse momentum ($p_T > 15$ GeV) leptons $l^{\pm}l^{\mp}l^{\mp}$ plus missing transverse energy E_T^{miss}. If an event has more than one primary vertex, the vertex with the largest associated track $\sum p_T^2$ is selected in the analysis. The charged leptons ($l^{\pm}l^{\mp}l^{\mp}$) can be muons or electrons, and they are ordered in the following way:

- l_0: the lepton with the unique electric charge;
- l_1: the lepton with the smallest angular distance from l_0;
- l_2: the remaining lepton.

According to elicity and charge conservation laws, l_0 and l_1 are the candidates from Higgs boson decay.
The common preselection requirements (already discussed in Section 2.3.2) on lepton association with a primary vertex, identification and isolation are summarised in Table 2.3 for *id* and *anti-id* leptons. Loose identification and isolation criteria are applied to the *anti-id* leptons, according to the *anti-id* definition in Section 2.3.2.
In particular, the FCLoose isolation working point foresees fixed cuts[3] on the electron isolation variables $E^{topocone20}$ and $p_T^{varcone20}$. The events with 3 charged leptons are divided into two signal regions (SR) to better separate the signal from the competitive backgrounds. The first signal region does not accept events with leptons with Same Flavour and Opposite electric charge Sign ($SFOS$), called $3l$-"Z-depleted" SR. The other signal region collects the events with one or two combinations of $SFOS$ leptons. Accepting this type of events, the second SR includes the 3/4 of signal events with the disadvantageous inclusion of irreducible background processes involving Z boson. Indeed, this

[3] The cuts are $E_T^{topocone20}/p_T < 0.2$ and $p_T^{varcone20}/p_T < 0.15$ where $E_T^{topocone20}$ is defined as the transverse energy sum of the topological clusters in a cone of $= 0.2$ size around the electron, after subtracting the contribution from the energy deposit of the electron itself and correcting for pile-up effects. While $p_T^{varcone20}/p_T < 0.15$ is defined as the scalar sum of the transverse momenta of the tracks with $p_T > 1$ GeV in a cone of size $=$ min (10 GeV/p_T^e, 0.2) around the electron of transverse momentum p_T^e, excluding the electron track itself. The cone size is chosen to be p_T-dependent to improve the performance for electrons produced in the decay of particles with large transverse momenta.

Table 2.3: Lepton preselection for the $WH \rightarrow WWW^* \rightarrow l\nu l\nu l\nu$ analysis, where Author = 1 is the veto to reject the electrons from gamma conversions.

	Muons		Electrons	
	id	Anti id	id	Anti id
p_T(GeV)	> 15	> 15	> 15	> 15
$\|z_0 \sin\theta]\|$ (mm)	< 0.5	< 0.5	< 0.5	< 0.5
$\|d_0\|/\sigma_{d_0}$	< 3	< 15	< 5	< 5
Identification	Medium	Medium	LHTight	LHLoose
Isolation	PLIVTight	-	PLIVTight	FCLoose
Author	-	-	1	-
Veto id	-	yes	-	yes

SR is nominated $3l$-"Z-dominated".

2.6 Misidentified lepton background

Relevant backgrounds consist of processes with misidentified leptons from photons, jets and b-jets. The sources of misidentified leptons depend on the lepton type.
Misidentified electrons can arise from jets, photon conversions, or semileptonic heavy-flavour decays. For the $3l$-"Z-dominated" SR channel, the dominant sources of misidentified electrons are the jets from the V+jet processes, the photon conversion from the $V\gamma$ events and the not prompt electrons from $t\bar{t}$ and Wt processes. Thanks to the isolation requirements (in Table 2.3) on electrons, the contribution of W+jet and the $W\gamma$ events are negligible in the $3l$-"Z-dominated" SR because they include more than one misidentified electron.
The misidentified muons are real non-prompt muons from semileptonic heavy-flavour decays, mainly from b-hadron jets.
Beyond the requirements in Table 2.3 that suppress most of the misidentified leptons, a cut on the minimum angular distance ($\Delta R > 0.1$) between lepton pairs improves the rejection of these objects without any loss of WH signal in the $3l$-"Z-dominated" SR selection. The contribution of the misidentified lepton background inside the SR is evaluated using data driven techniques. Specific transfer factors are extrapolated from the dedicated control regions that are enriched by misidentified leptons and associated with other control regions with *anti-id* leptons. In the following sections, the transfer factors are distinguished into Scale Factors (SF) for misidentified electrons and Fake Factors (FF) for misidentified muons.

2.6.1 Estimation of the misidentified electron background

Misidentified electrons are estimated by scaling the MC yields in the SR with Scale Factors measured from data in dedicated control regions (CRs). Differently from the other processes in the SR, the MC processes with a misidentified electron are simulated by selecting an *anti-id* electron.

As reported above, the main sources of misidentified electrons are: $t\bar{t}$ and Wt (heavy flavour jets), Z+jet (light flavour jets) and $Z\gamma$ (photon conversion). Using a truth classifier based on the ΔR matching between truth and reconstructed objects, the MC events are firstly divided according to the source of the misidentified electrons. Then, the MC events are normalised to data simultaneously in the dedicated control regions (CRs) to measure the SFs.
Table 2.4 reports the CR definitions together with their target processes and the sources of misidentified electrons.

The control region $CR^{e^{\pm}\mu^{\mp}e^{\pm}}_{heavyflav}$ is enriched by the misidentified electrons from b-jets thanks to

Table 2.4: Definition of the CRs for electron scale factors.

Control Regions	Selections	Target source of misidentified electron
$CR^{e^{\pm}\mu^{\mp}e^{\pm}}_{heavyflav}$	require at least one *id* lepton to match single lepton triggers $e^{\pm}\mu^{\mp}e^{\pm}$, lepton p_{T} > 15 GeV $N_{bjet} = 1$	b-jets
Common selections for $CR^{\mu^{\pm}\mu^{\mp}e}_{\gamma\ conv.}$ and $CR^{\mu^{\pm}\mu^{\mp}e}_{lightflav}$	require at least one *id* lepton to match single lepton triggers $\mu^{\pm}\mu^{\mp}e$, lepton p_{T} > 15 GeV $N_{bjet} = 0$ m^{min}_{SFOS} > 12 GeV E^{miss}_{T} < 30 GeV, m_T < 40 GeV	common requirements for γ conversion and Z+jets process
$CR^{\mu^{\pm}\mu^{\mp}e}_{\gamma\ conv.}$	40 GeV < $m_{\mu\mu}$ <80 GeV, $\lvert m_{lll} - m_Z \rvert$ < 10 GeV	γ conversion
$CR^{\mu^{\pm}\mu^{\mp}e}_{lightflav}$	80 GeV < $m_{\mu\mu}$ <100 GeV, m_{lll} < 200 GeV	Z+jets process
id CR	3 *id* leptons	
anti-id CR	2 *id* leptons + 1 *anti-id* electron ($\not{e}$)	

$e^{\pm}\mu^{\mp}e^{\pm}$ event topology selection and the number of b-jets = 1 cut, assuring the orthogonality to the other CRs and the $3l$-"Z-dominated" SR, too. By requiring two electrons with the same electric charge sign, this CR specifically selects the $t\bar{t} \rightarrow W^+W^-b\bar{b}$ process with the final state in which $e^{\pm}\mu^{\mp}$ are from the W boson decays and the other $e^{\pm}$ from a heavy flavour jet. The electron with lower p_T is identified as the misidentified electron since a b-tagged jet is expected to have less energy than a prompt electron from W boson decay due to the larger W boson mass than the b quark mass and the electron from a b decay typically carries only a small fraction of the original momentum of the b quark.
The control regions $CR^{\mu^{\pm}\mu^{\mp}e}_{\gamma\ conv.}$ and $CR^{\mu^{\pm}\mu^{\mp}e}_{lightflav}$ have in common the selection cuts: the $E^{T}_{miss} < 30$ GeV cut and $\mu^{\pm}\mu^{\mp}e$ event topology selection, which make them orthogonal to the other signal regions. They respectively select the γ conversion and Z+jets processes. The requirements on $m_{\mu\mu}$ and m_{lll} cuts are specific to the single process, improving its purity without severe statistics losses and assuring the orthogonality between $CR^{\mu^{\pm}\mu^{\mp}e}_{\gamma\ conv.}$ and $CR^{\mu^{\pm}\mu^{\mp}e}_{lightflav}$.
Every CR is split into *id* CRs (with 3 *id* leptons) and *anti-id* CRs (with 1 *anti-id* and 2 *id* leptons), as

shown in the last rows of Table 2.4.
Events in the *id* control regions are binned in terms of the misidentified electron p_T. In detail, events in the *id* $CR^{\mu^\pm\mu^\mp e}_{\gamma conv}$ and *id* $CR^{\mu^\pm\mu^\mp e}_{light flav}$ are split into the p_T bins of the electrons while $CR^{e^\pm\mu^\mp e^\pm}_{heavy flav}$ is divided into p_T bins of the electron with lower p_T between the two same sign electrons. The events in all three *anti-id* CRs are binned by p_T of the *anti-id* electrons. Therefore, the scale factors are evaluated by fitting MC yields of the *anti-id* CRs to data yields in *id* CRs, for each source of misidentified *anti-id* electrons and their p_T bin. The fit also constrains the normalisation of prompt processes as $WZ, ZZ, VVV, ttV, VH, ttH, tZ$. The fit results are reported in Table 2.5, together with the contributions of their uncertainties. The σ^{fit}_{stat} are the statistical uncertainties from fit, their values

Table 2.5: Data driven electron scale factors with their uncertainties.

	$SF_{\gamma\ conv}$	σ^{fit}_{stat}	$\sigma^{fit}_{stat+xs}$	$\sigma_{non\text{-}closure}$	σ_{total}
15 GeV < p_T < 20 GeV	0.071	$^{+5.8\%}_{-5.8\%}$	$^{+5.8\%}_{-5.8\%}$	35.5%	36.0%
20 GeV < p_T < 30 GeV	0.111	$^{+5.6\%}_{-5.6\%}$	$^{+6.6\%}_{-5.6\%}$	19.2%	20.2%
p_T > 30 GeV	0.176	$^{+8.9\%}_{-9.5\%}$	$^{+9.5\%}_{-9.5\%}$	19.0%	21.3%
	SF_{heavy}	σ^{fit}_{stat}	$\sigma^{fit}_{stat+xs}$	$\sigma_{non\text{-}closure}$	σ_{total}
15 GeV < p_T < 20 GeV	0.012	$^{+118.8\%}_{-93.8\%}$	$^{+118.8\%}_{-93.8\%}$	35.5%	112.0%
20 GeV < p_T < 30 GeV	0.020	$^{+100.0\%}_{-100.0\%}$	$^{+100.0\%}_{-100.0\%}$	19.2%	101.8%
p_T > 30 GeV	0.007	$^{+100.0\%}_{-100.0\%}$	$^{+100.0\%}_{-100.0\%}$	19.0%	101.8%
	SF_{light}	σ^{fit}_{stat}	$\sigma^{fit}_{stat+xs}$	$\sigma_{non\text{-}closure}$	σ_{total}
15 GeV < p_T < 20 GeV	0.057	$^{+19.6\%}_{-19.6\%}$	$^{+19.6\%}_{-19.6\%}$	35.5%	40.6%
20 GeV < p_T < 30 GeV	0.014	$^{+75.0\%}_{-75.0\%}$	$^{+75.0\%}_{-81.2\%}$	19.2%	80.4%
p_T > 30 GeV	0.012	$^{+100.0\%}_{-100.0\%}$	$^{+106.2\%}_{-106.2\%}$	19.0%	107.9%

are driven by the limited misidentified electron yields in *anti-id* CRs and data yields in *id* CRs. The $\sigma^{fit}_{stat+xs}$ are the uncertainties from fit including the normalisation factors for the prompt lepton MC processes in CRs.
A conservative approach is applied to the statistical uncertainties on the SF_{heavy} and for the high p_T bins (100% uncertainties) when the fit returns a value smaller than the physical lower limit. In these cases, the upper statistical uncertainties $+\sigma^{fit}_{stat}$ is used both as the central value (SF) and as the uncertainty of the final scale factors.
The $\sigma_{non-closure}$ are the uncertainties that take into account the difference between the yields of misidentified electrons MC in an *id* CR and the scaled misidentified *anti-id* electrons MC. The scaled misidentified *anti-id* electrons MC yields are the *anti-id* electron events from the *anti-id* CR normalised by the MC-driven scale factors resulting from the fit using Asimov dataset [4] in *id* CRs. For all the CRs, the $\sigma_{non-closure}$ uncertainties are evaluated in the $CR^{\mu^\pm\mu^\mp e}_{light flav}$ since its selection is the closest to the SR definitions of the $2l$ WH with a lepton from the Higgs boson decay chain (Feynman diagram in 2.2 (a)) and $3l$ WH channels. The total uncertainties σ_{total} on scale factors combine $\sigma_{non-closure}$ and the symmetrised $\sigma^{fit}_{stat+xs}$.
In the end, the SFs in Table 2.5 are applied to MC yields that have been selected with an *anti-id* electron

[4] The Asimov dataset is generated without statistical fluctuations where the observed values are equal to the expected ones.

in the SR according to the source and *anti-id* electron p_T to estimate the misidentified electrons in $3l$-"Z-dominated" SR.

2.6.2 Estimation of misidentified muon background

A different method, called Fake Factor method, is used for estimating the misidentified muons where the fit is applied to data in both the *id* CR and in the *anti-id* CR[5].
The CR, enriched by misidentified muons from b-jets, is defined in Table 2.6.
The $CR_{t\bar{t}}^{\mu^{\pm}e^{\mp}\mu^{\pm}}$ selects events with two muons with same sign and an electron with opposite sign.

Table 2.6: Definition of the CRs for muon fake factors.

Control Regions	Selections
$CR_{t\bar{t}}^{\mu^{\pm}e^{\mp}\mu^{\pm}}$	require at least one *id* lepton to match single lepton triggers $\mu^{\pm}e^{\mp}\mu^{\pm}$, sum of charge ± 1, lepton p_{T} > 15 GeV $N_{bjet} = 1$
id $CR_{t\bar{t}}^{\mu^{\pm}e^{\mp}\mu^{\pm}}$	3 *id* leptons
anti-id $CR_{t\bar{t}}^{\mu^{\pm}e^{\mp}\not{\mu}^{\pm}}$	2 *id* leptons + 1 *anti-id* muon ($\not{\mu}$)

Similar to $CR_{heavyflav}^{e^{\pm}\mu^{\mp}e^{\pm}}$ described in Section 2.6.1, the event topology selection and the cut on the number of b-jets ($N_{bjet} = 1$) are optimised for the $t\bar{t}$ and tW processes with a misidentified muon and make sure that the CR is orthogonal to the other kinematic regions in the analysis.
Based on the same consideration in Section 2.6.1 made on the $CR_{heavyflav}^{e^{\pm}\mu^{\mp}e^{\pm}}$, the muon with lower p_T is associated with the b-jet. Dedicated studies at the truth MC level have shown that this assumption is corrected in 75% of the cases. The remaining cases are $t\bar{t}$ events where the less energetic muon is from W bosons decay while the more energetic muon is from b-hadrons in the CR, leading to a contamination of prompt muons in the misidentified muons. The contamination fraction is within the statistical uncertainty on the fake factors. Other contamination sources of prompt muons are the $t\bar{t}Z$ and $t\bar{t}H$ processes but with negligible contributions with respect to the $t\bar{t}W$ process.
Then, the CR is split into *id*, with three *id* leptons, and *anti-id* CR, which includes events with two *id* leptons and *anti-id* muon.
The different isolation requirements on the *id* and *anti-id* muons introduce p_T spectrum differences between the underlying jet that originates the *id* misidentified muon and the *anti-id* muon. The p_T of *anti-id* muons is corrected by adding the variable $ptvarcone30$[6], $p'_T = p_T + ptvarcone30$, before defining the p_T bins. The fake factors are defined as the ratios between the yields of the misidentified muons ($N_{misindentified}^{id,p_T}$) that satisfy the *id* requirements and the yields ($N_{misindentified}^{anti-id,p'_T}$) of the *anti-id*

[5] In Section 2.6.1 the estimation of the SF includes MC yields in *anti-id* CRs.

[6] $ptvarcone30$ is a track-based isolation variable defined as the scalar sum of the transverse momentum of selected tracks within a cone of $\Delta R_{conesize} = \min(\frac{10GeV}{p_T^{muon}[GeV]}, 0.3)$ centred around the muon track direction, excluding the muon track itself.

muons,

$$FF = \frac{N^{id,p_T}_{misindentified}}{N^{anti-id,p'_T}_{misindentified}} = \frac{N^{id,p_T}_{data} - N^{id,p_T}_{prompt}}{N^{anti-id,p'_T}_{data} - N^{anti-id,p'_T}_{prompt}}. \quad (2.7)$$

These yields can be expressed in terms of data ($N^{id\,or\,anti-id,p_T}_{data}$) and yields of prompt process ($N^{id\,or\,anti-id,p_T}_{prompt}$). The FFs are estimated by fitting the data in both the *id* and *anti-id* CR. The fit constraints the normalisation of the prompt backgrounds (i.e. $WZ, ZZ, VVV, ttV, VH, ttH, tZ$ processes) in this case, too.
The measured fake factors are reported in Table 2.7 for each corrected *anti-id* p_T bin. The uncertainties

Table 2.7: Data driven muon fake factors with uncertainties where σ^{fit}_{stat} is the statistical uncertainty from fit, due to limited statistics in the *id CR*s. The $\sigma^{fit}_{stat+xs}$ is the uncertainty from fit with including the inclusive cross section uncertainty for the prompt lepton MC processes using the nuisance parameters. The $\sigma_{non\text{-}closure}$ is the uncertainty due to MC non-closure in the 3*l*-"Z-dominated" SR. The σ_{total} is the total uncertainty of the fake factors.

	FF^{DD}_{μ}	σ^{fit}_{stat}	$\sigma^{fit}_{stat+xs}$	$\sigma_{non\text{-}closure}$	σ_{total}
15 GeV < p_{T} < 20 GeV	0.023	20.0%	20.0%	37%	42.0%
20 GeV < p_{T} < 30 GeV	0.0039	60.5%	63.2%	24%	66.6%
p_{T} > 30 GeV	0.016	27.5%	28.8%	23%	37.1%

in Table 2.7 are the same ones considered on the misidentified electrons. The *non-closure* uncertainty is evaluated by considering the misidentified muon MC yields in the 3*l*-"Z-dominated" SR and the scaled misidentified *anti-id* muon MC yields by MC-driven muon fake factors that have been measured by replacing the data with the Asimov data.
The misidentified lepton background in the 3*l*-"Z-dominated" SR is, then, estimated by applying the FF in Table 2.7 to the *anti-id* muon MC yields in the SR.

2.6.3 Validation of misidentified lepton background

The estimation of the misidentified lepton background is validated using dedicated control samples. The misidentified leptons predicted directly through MC are compared with the misidentified leptons estimated from a validation *anti-id* SR by using the transfer factors.
The *anti-id* SRs ($SR_{ll\not{e}}$, $SR_{ll\not{\mu}}$) are associated with 3*l*-"Z-dominated" and 3*l*-"Z-depleted" SRs that deal with the misidentified muons and electrons separately. A third region, $SR_{lll_{misid}}$, is used to validate the misidentified lepton background regardless of the l_{misid} flavour.
The results of the validation test are reported in Table 2.8 for the corresponding SRs. The misidentified direct MC predictions and the scaled *anti-id* estimations of misidentified leptons show a satisfying agreement.

Table 2.8: Ratio between the MC predicted yields and the MC driven estimated yields in SRs.

Z-dominated	$SR_{lle_{misid}}$	$SR_{ll\mu_{misid}}$	$SR_{lll_{misid}}$
predicted/estimated	1.07 ± 0.41	0.70 ± 0.31	0.99 ± 0.38
Z-depleted	$SR_{lle_{misid}}$	$SR_{ll\mu_{misid}}$	$SR_{lll_{misid}}$
predicted/estimated	1.12 ± 0.57	0.84 ± 0.37	0.97 ± 0.45

2.7 Definition and background composition of the 3l-"Z-dominated" Signal Region

The 3l-"Z-dominated" SR includes the events with 3 charged leptons forming at least a $SFOS$[7] and missing transverse energy.
Processes such as $WZ/W\gamma^*$, VVV (especially WWW) and ZZ^* with a not detected lepton represent the irreducible backgrounds in 3l-"Z-dominated" SR. They show the same signature of the WH process, with 3 prompt leptons that pass the isolation and identification criteria (in Table 2.3).
To better module the $WZ/W\gamma^*$ process in SR, the $WZ/W\gamma^*$ process is divided into two bins of the number of jet variable on which the corresponding normalisation factors (NFs) are applied. The NFs are estimated by the profile likelihood fit (in Section 2.11) performed on data of dedicated control regions (in Section 2.9) for the $WZ/W\gamma^*$ process without jets ($WZ/W\gamma^*_{w0jet}$) and with at least a jet ($WZ/W\gamma^*_{wjets}$).
Background processes that include a misidentified lepton from $Z + jets$, $Z\gamma$, $t\bar{t}$ and Wt processes are estimated by data-driven techniques, as already described in Section 2.6.
The 3l-"Z-dominated" SR searches for the $WH \rightarrow WWW^* \rightarrow l\nu l\nu l\nu$ events that satisfy the selection

Table 2.9: Event selection criteria used to define the signal region in the 3l-"Z-dominated" analysis. The symbols are defined in the text.

Quantities	Cuts and vetos	Actions of Cuts and vetos
Preselection Number of SFOS	3 isolated leptons ($p_T > 15$ GeV) total lepton charge ± 1 2 or 1	SR defintion
E_T^{miss} [GeV]	> 30	Suppression of Z+jets and ZZ^* processes
Number of b-jets	0	Suppression of processes including top quark
$\lvert m_{\ell\ell}^{SFOS} - m_Z \rvert$ [GeV]	> 25 (SFOS)	Suppression of $WZ/W\gamma^*$, ZZ^*, Z+jets and $Z\gamma$
$m_{\ell\ell}$ [GeV]	> 12 (min. SFOS)	low mass resonance events from heavy-flavour quarkonia

in Table 2.9. Its first row selects the SR. While the successive cuts and vetos optimise the rejection of specific background processes. The cut on E_T^{miss} (> 30GeV) reduces Z+jets and ZZ^* processes, in addition to selecting final states with neutrinos.
The prompt $t\bar{t}V$ and tZ processes contribute to the irreducible background, and they are suppressed thanks to the veto on the b-jets number. The requirements on isolation and veto on the b-jets number also suppress the $t\bar{t}$ and Wt processes including top quark with an additional non-prompt lepton from b-hadron decay.
The Z-veto is defined by the requirement $\lvert m_{\ell\ell} - m_Z \rvert > 25$ GeV, It is applied to suppress in part the dominant irreducible background including Z boson, as visible in $m_{l_0 l_1}$ and $m_{l_0 l_1}$ MC distributions in next Section 2.8.2. In order to suppress background events from heavy-flavour quarkonia, the smallest invariant mass of $SFOS$ pairs is required to be greater than 12 GeV.
Processes with more than a not prompt lepton are considered negligible.

[7] The $SFOS$ is acronym of Same Flavour Opposite Sign, as defined in Section 2.5.

Table 2.10: Cutflow for the event selection in the 3l-"Z-dominated" SR. Yields for the WH signal process and total background. Only MC statistical uncertainties are shown. Data driven methods are used to estimate misidentified leptons, e-fakes and μ-fakes, from top quark, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective WZ CRs (Section 2.11) are applied to the $WZ/W\gamma^*$ processes.

$\sqrt{s} = 13TeV$, $\mathcal{L} = 139fb^{-1}$ (Full Run 2)	WH	Total Bkg	Significance
Scale factors		NFs Applied	
Preselection	58.00 ± 0.10	29002.43 ± 37.87	0.33 ± 0.00
1 or 2 SFOS	44.05 ± 0.08	28777.31 ± 37.83	0.25 ± 0.00
$E_T^{miss} > 30$ GeV	36.58 ± 0.07	18138.09 ± 27.34	0.27 ± 0.00
bjet-veto	34.77 ± 0.07	15676.91 ± 26.55	0.27 ± 0.00
Z-veto	25.16 ± 0.06	1241.97 ± 8.50	0.69 ± 0.00
$m_{\ell,\ell;min} > 12$ GeV	24.27 ± 0.06	1149.43 ± 8.25	0.69 ± 0.00

The MC yields of the WH signal process and total background in 3l-"Z-dominated" SR are presented in Table 2.10 after each cited cut. The quoted expected approximate Poisson significance is defined as:

$$\sqrt{2((s+b)\ln(1+\frac{s}{b})-s)}, \tag{2.8}$$

in Table 2.11. It reports the yields of every background component after each 3l-"Z-dominated" SR cut, showing how the cuts act on the different background processes. The $WZ/W\gamma^*$ process turns to be the dominant background.

Table 2.11: Cutflow for the event selection in the 3l-"Z-dominated" SR. Yields for the ZH process (not included in the total background column) and every single background component. Only MC statistical uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top quark, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective WZ CRs (Section 2.9) are applied to the $WZ/W\gamma^*$ processes.

(Full Run 2)	ZH	$WZ/W\gamma^*_{w0jet}$	$WZ/W\gamma^*_{wjets}$	ZZ^*	VVV	tV+ttV	μ-fakes	e-fakes	other Higgs
Scale factors		NF = 1.04	NF = 0.90						
Preselection	94.17 ± 0.47	9937.73 ± 24.99	10655.06 ± 16.77	3999.93 ± 15.19	120.60 ± 0.23	1299.63 ± 2.45	748.71 ± 3.94	2146.03 ± 16.61	94.74 ± 0.46
1 or 2 SFOS	92.72 ± 0.47	9920.26 ± 24.97	10630.37 ± 16.75	3994.57 ± 15.19	99.59 ± 0.22	1236.81 ± 2.36	697.20 ± 3.82	2121.22 ± 16.60	77.30 ± 0.43
$E_T^{miss} > 30$ GeV	63.40 ± 0.37	6939.66 ± 20.78	8123.78 ± 14.38	1153.28 ± 6.26	88.10 ± 0.21	1062.30 ± 2.20	322.10 ± 2.57	383.52 ± 7.63	65.35 ± 0.38
bjet-veto	53.83 ± 0.35	6824.69 ± 20.58	7156.66 ± 13.85	992.43 ± 5.88	77.08 ± 0.19	112.86 ± 0.64	194.79 ± 2.05	298.43 ± 7.08	19.97 ± 0.33
Z-veto	2.80 ± 0.08	406.31 ± 5.32	530.03 ± 3.65	132.55 ± 3.08	32.95 ± 0.09	14.64 ± 0.31	33.25 ± 0.84	87.57 ± 4.51	4.68 ± 0.16
$m_{\ell,\ell;min} > 12$ GeV	2.72 ± 0.08	382.32 ± 5.21	482.36 ± 3.49	122.04 ± 2.99	32.73 ± 0.09	14.29 ± 0.30	31.63 ± 0.81	79.45 ± 4.35	4.62 ± 0.16

In Figure 2.7, the MC distribution of $\Delta R_{\ell_0 \ell_1}$, the angular separation of the two Higgs boson leptons candidates, for the signal and the background processes in the $3l$-"Z-dominated" SR is reported as an example of a variable with good separation between the signal and the background processes. To further separate between signal and background processes in $3l$-"Z-dominated" SR, a discriminant is implemented using an Artificial Neural Network (ANN), as described in next Section 2.8.2.

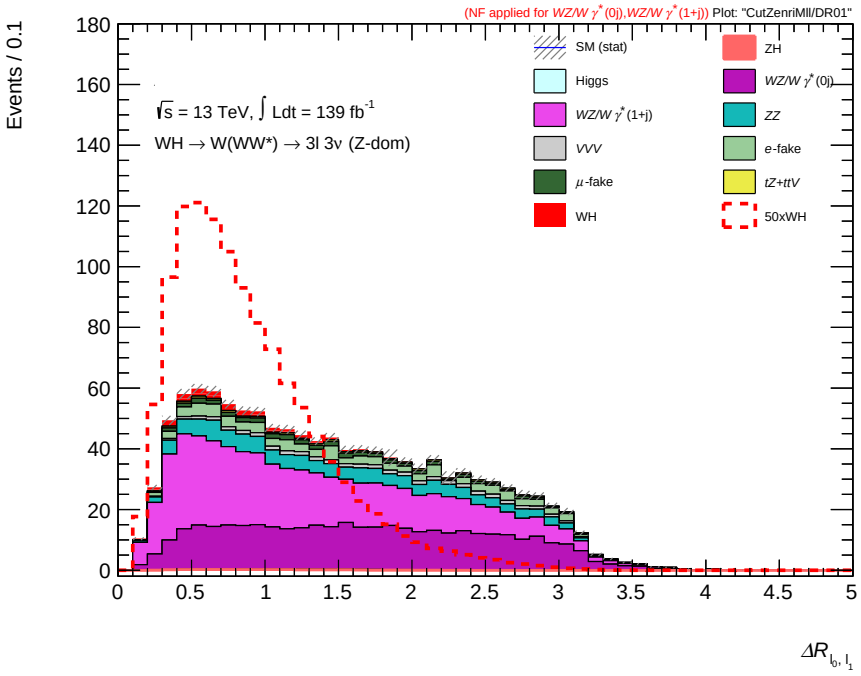


Figure 2.7: Angular distances between the l_0 and l_1 leptons. Only MC statistical uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top quark, Z+jets and $Z\gamma$ background. The WH signal is both stacked and superimposed (with a factor of 50). Normalisation factors from a fit to data in the respective WZ CRs (Section 2.9) are applied to the $WZ/W\gamma^*$ processes.

2.8 MVA analysis

In a multivariate analysis, a discriminant variable is built to maximise the sensitivity of the signal process with respect to the irreducible backgrounds. A general overview of ANNs is given in Section 2.8.1 before describing the ANN for the $3l$-"Z-dominated" channel in Section 2.8.2.
A binary Artificial Neural Network classifier is implemented to discriminate the WH signal versus the background processes in $3l$-"Z-dominated" SR in the Toolkit for MultiVariate Analysis (TMVA) [65] ROOT-integrated environment.
The TMVA software package allows interfacing object-oriented implementations in C++/ROOT for several multivariate analysis techniques with the high-level wrapper Keras [66] for machine learning frameworks with TensorFlow [67] backend.

2.8.1 Artificial Neural Network

An artificial neural network can be used in classification or regression problems. Its structure is conceived to resemble a brain. The elementary unit is called neuron, and its schematic representation is reported in Figure 2.8.

In a neuron, the input $x = (x_1, ..., x_n)$ is rearranged in a linear combination f_j with weights

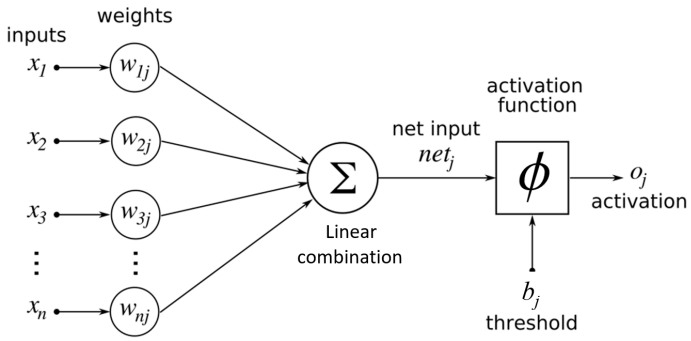


Figure 2.8: Schematic of a neuron.

$w_j = (w_{j,1}, ..., w_{j,n})$ and biases b_j, and $f_j = w_{ji}x^i + b_j$ is associated with an output o_j by an application ϕ called activation function. Then, an activation function shapes the node output, given a set of input. It determines the ability of the neural network to approximate the target function that is successively minimised in the weight computation.

Some typical activation functions are the identity (Equation 2.9), sigmoid (Equation 2.10), hyperbolic tangent function (tanh in Equation 2.11) and Rectified Linear Unit (ReLU in Equation 2.12):

$$\phi(k) = k \tag{2.9}$$

$$\phi(k) = \frac{1}{1 + \exp(-k)} \tag{2.10}$$

$$\phi(k) = \frac{\exp(+k) - \exp(-k)}{exp(+k) + \exp(-k)} \tag{2.11}$$

$$\phi(k) = \max(0, k) \tag{2.12}$$

$$\tag{2.13}$$

The output $o_j = \phi(w_{ji}x^i + b_j)$ is analogous to an axon. Its value propagates to the input of a possible next layer of neurons, forming a multilayer architecture with hidden layers, as the one reported in Figure 2.9. The neurons of a layer are linked to all the units of the next layer, but they have no link among themselves.

On the last layer, the output layer, a different activation function is used with respect to the hidden layers, depending on the type of problems.

In the regression problems, no activation function is defined in the output layer.

In the classification, the output layer of the neural network model contains a neuron per class, predicting of which class the input event belongs to. Therefore, the activation function of the output layer has to assure the sum of all output neuron predictions is equal to 1. The function softmax is one of the most

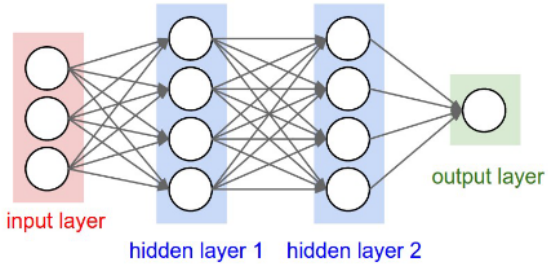


Figure 2.9: Schematic of a multilayer architecture.

used, defined as:

$$\text{softmax}(z)_i = \frac{\exp(z_i)}{\sum_j \exp(z_j)} \tag{2.14}$$

Besides the activation function, the neural network parameters are the number of layers and neurons. Set the architecture, the weights and biases of the neurons are estimated from a training sample using a supervised or unsupervised learning approach.
The supervised algorithms analyse input labelled[8] data to infer weights and biases that will be used to map new unlabelled data. In contrast, the unsupervised learning, the algorithms are not provided with any pre-assigned labels in the training data.
In practice, the supervised estimation is obtained by minimising a loss function with a gradient descent algorithm (also called optimizer).
In a classifier application, the adopted loss function is the categorical crossentropy, defined in Equation 2.15, where y_i is the true label and $\hat{y}$ is the prediction of i^{th} output neuron.

$$L(y, \hat{y}) = -\sum_{i}^{classes} y_i \log \hat{y}_i \tag{2.15}$$

Two different optimizer algorithms have been applied in the following analysis: the Stochastic Gradient Descent (SGD) and the ADAptive Momentum estimation (ADAM). Each optimizer has its own parameters. In these algorithms, the gradient computation is performed in different random and not-intersecting subsets (batches) of the training sample. The mean of the computed gradient in a batch is used to update the weights and biases. The user can set the batch size and the number of epochs. The composition inside the batches changes in every epoch. An epoch is the iterative step in which all the elements of the training sample are sent in input to the algorithm.
The SGD algorithm trains the weights and biases according to:

$$w_{jk}^{(new)} = w_{jk}^{(old)} - \eta \frac{\partial L}{\partial w_{jk}} + m\Delta w_{jk}^{(old)} \tag{2.16}$$

$$b_{j}^{(new)} = b_{j}^{(old)} - \eta \frac{\partial L}{\partial b_{j}} + m\Delta b_{j}^{(old)} \tag{2.17}$$

[8] The adjective labelled indicates that the class of every element is a priori-known in "labelled" data (samples), while the adjective unlabelled refers to data (samples) in which the classes of the elements is unknown.

In Equation 2.17, η and m are the learning rate and momentum of SGD and L is the loss function. The momentum m avoids to get trapped in local minima and it speeds up the algorithm in case of function with wide plateau regions by adding a term times m that is dependent on the previous interactions.
Beyond the weights and biases, the ADAM algorithm adapts its two momenta, m and v, after each iteration according to:

$$m^{(new)} = \beta_1 m^{(old)} + \left(1 - \beta_1\right) \nabla_w L^{(old)} \tag{2.18}$$

$$v^{(new)} = \beta_2 v^{(old)} + \left(1 - \beta_2\right) \left(\nabla_w L^{(old)}\right)^2 \tag{2.19}$$

In Equations 2.19, β_1 and β_2 are parameters that the user can set before the training. Therefore, the weights (and biases) are updated according to:

$$w_{jk}^{(new)} = w_{jk}^{(old)} - \eta \frac{\hat{m}}{\sqrt{\hat{v}} + \epsilon} \tag{2.20}$$

In Equation 2.20, the corrected expressions of the momenta $\hat{m} = \frac{m^{(new)}}{1-\beta_1^{(new)}}$ and $\hat{v} = \frac{v^{(new)}}{1-\beta_2^{(new)}}$ are considered, and ϵ is simply a smoothing term to avoid a null denominator.
To improve the regularisation of multilayer architecture, the dropout method randomly excludes a fraction (set by the user) of neurons of a layer from the learning.
To evaluate the outcome of neural network training, the following learning curves are taken into account: loss function and the accuracy as a function of epochs.
The accuracy is an evaluation metric that computes how often the ANN prediction matches the true label. It is used to select the best model, and it does not participate in the weight computation.
Generally, a portion of the training sample is reserved for test and validation samples, which do not participate in the ANN training. The test sample is used to test the final trained model on data on which the ANN learning does not rely. The validation sample is reserved to evaluate the model while tuning the ANN parameters.
From the single trends and comparison between the curves relative to the training and validation samples, it is possible to understand if the learning (or training) is underfitted, good or overfitted. The loss function trend can indicate if the network predictions converged, stacked on a local minimum or diverged with respect to the true labels.
For $3l$-"Z-dominated" SR analysis, the best model has the weights corresponding to the maximum measured validation accuracy in an epoch. When the training and validation of the model are correct, the weights corresponding to the maximum accuracy are also a point of local loss function minimum. Therefore, it is routine to simultaneously check the curves of the accuracy and loss function as a function of the epochs.
The Receiver Operating Curve (ROC), the integral of the ROC (AUC) and the significance approximation in Equation 2.8 are used to evaluate the network discriminant power between the signal and the background.
In a classification problem, the prediction of a classifier belongs to one of the cases in Figure 2.10: True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) after fixing the significance level β. Therefore, the classifier prediction successfully classifies the objects in the TP and TN cases, and it fails in the FP and FN cases with a $sensitivity$ given by $\frac{TP}{TP+FN}$ and a $specificity$ $\frac{TN}{TN+FP}$.

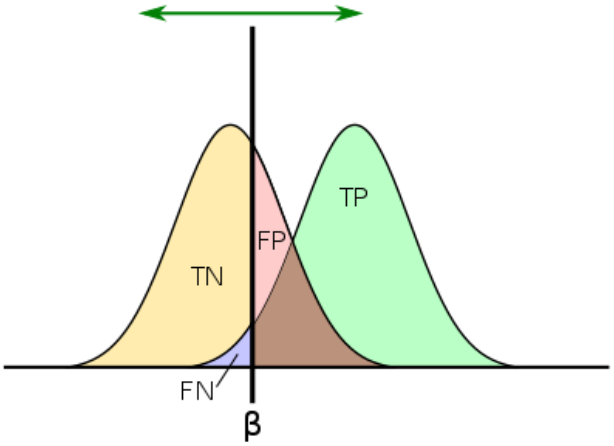


Figure 2.10: True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) integrals at fixed significance level β for a generic hypothesis test.

The ROC is the curve on the plane individuated by the $1 - specificity$ on the horizontal axis and *sensitivity* on the vertical axis. Therefore, the AUC (integral) can be interpreted as the overall probability that the classifier correctly selects the signal from the background in a binary case or a class vs the other classes in a multi-class application.

2.8.2 Training and validation of ANN

The ANN for 3l-"Z-dominated" SR analysis is trained using alternative high statistics samples for the WH signal process and the dominant irreducible background, the $WZ/W\gamma^*$ process.
For the WH signal process, the training sample is simulated through POWHEG interfaced with PYTHIA8 while the alternative sample of $WZ/W\gamma^*$ background is a customised filtered high statistic sample generated with SHERPA 2.2.1. The events are filtered with the following requirements: a cut at 15 GeV is applied to the transverse momentum p_T of the leading lepton and at 7 GeV to the p_T of the sub-leading leptons; the pseudorapidity of the leptons is restricted to $|\eta|$ <2.8; the τ leptons are forced to decay leptonically; and the events from resonant Z boson decay are removed ($|m_{\ell\ell} - m_Z| >$ 5 GeV) to mimic the Z-veto cut.
To define the input sample, the 3l-"Z-dominated" SR requirements (see Table 2.9) are applied. The resulting statistics of the training and validation test[9] samples are reported in Table 2.12.

[9] In TMVA, the validation and test samples have been merged to increase statistics [65]. The bias introduced by this on the evaluation results does not affect the analysis. It has been independently validated by comparing the training, test with the ANN output on nominal MC yields.

Table 2.12: Number of signal and background MC events used in the training and in the validation.

MC sample	training events	validation/test events
Signal WH	345.8x10^3	38.8x10^3
$WZ/W\gamma^*$	930.0x10^3	235.1x10^3

Before going into the input of the ANN, each variable has its mean subtracted, and it is scaled to its standard deviation. The input data are weighted by physics weights, consisting of the product of the predicted production cross-section of the process at LHC, the k-factor (it corrects the cross-section for the absence of higher-order terms) and filter efficiency which represents the expected fraction of events that pass the applied filters (like the case of $WZ/W\gamma^*$ background).
To select the final architecture, studies are carried out varying the network hyperparameters and input variables and comparing the AUC values and the shape of the ANN output distributions. The scans on the hyperparameters are performed, fixing the input variables at a total of 30 variables derived by the three charged leptons and E_T^{miss} and varying one or more of the network parameters. The differences in the AUC are minimal, ranging between 0.84 and 0.88. Therefore, the architecture is fixed with the one in Table 2.13 that has the best AUC.

Table 2.13: Model of the 3l-"Z-dominated" ANN.

Parameter	Values
number of input variables	15
Kernel initializer	glorot normal
number of hidden layers	6
layer activation function	ReLU
output activation function	softmax
dropout rate	0.15
number of neurons	variable
Layers x Neurons	2x128; 2x64; 2x32
loss function	binary crossentropy
optimizer	SGD (learning rate: 0.01, momentum: 0.7)
batch size	500

Keeping the same selection criteria, a different scan is performed by removing one of the 30 variables at a time.The first dismissed variables have been those with the worst separation between the signal and background and those that are strongly correlated both in signal and background samples. Finally, a set of 15 variables is selected:

- $p_{\mathrm{T}}^{\ell_0}$, the transverse momentum of the first lepton,
- $|\Sigma p_{\mathrm{T}}^{\ell_i}|$, the magnitude of the vectorial sum of lepton transverse momenta,
- $\Delta\eta_{\ell_0\ell_1}$, $\Delta\eta_{\ell_1\ell_2}$, the pseudorapidity separation between the first and the second leptons, and the second and third leptons,
- $\Delta\phi_{\ell_0\ell_2}$, the azimuthal separation between the first and third leptons,

- $\Delta R_{\ell_0\ell_1}$, $\Delta R_{\ell_0\ell_2}$, the angular distances between the first and the second leptons, and the first and third leptons,
- $m_{\ell_0\ell_1}$, $m_{\ell_0\ell_2}$, $m_{\ell_1\ell_2}$, the invariant mass of the three lepton pairs,
- $E_{\mathrm{T}}^{\mathrm{miss}}$, the missing transverse energy,
- $\Delta\phi_{\ell_i E_{\mathrm{T}}^{\mathrm{miss}}}$ $(i = 1, 2, 3)$, the azimuthal separation between leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$,
- $m_{\mathrm{T}}^{W} = \sqrt{2\, p_{\mathrm{T}}^{\ell_W} \cdot E_{\mathrm{T}}^{\mathrm{miss}} \cdot \left(1 - \cos\Delta\phi\left(\ell_W, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}\right)\right)}$, the transverse mass of the W boson, built from the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ and the lepton ℓ_W which is the lepton not belonging to the SFOS lepton pair with invariant mass closer to the Z boson mass, and could be either ℓ_1 or ℓ_2.

In the correlation matrices shown in Figure 2.11, it is visible the effort to keep at minimum the number of the correlated variables in each sample and the couples of variables that show the same correlation coefficient in the signal and background samples. The distributions of the input variables of the alternative samples used for the training are reported in Appendix A.
As visible in Figure 2.12 (a), the validation accuracy of the final optimised ANN reached 0.817 as maximum value at the 112^{th} epoch with 0.88 as AUC. The training and validation accuracy curves do not diverge as a function of the epoch, excluding overfitting.
As shown in Figure 2.12 (b), the ANN output distributions of signal and background classes for the training and test samples agree in the entire range, supporting the goodness of the ANN learning. For values greater than 0.8, the observed degradation of the agreement between the training background and test background curves is considered negligible because the statistic is poor in this range.
Figures 2.13 and 2.14 show the distributions of the ANN input variables for the nominal MC samples in the $3l$-"Z-dominated" SR. The WH signal distributions (x50) in red dotted lines are compared with all the background processes in $3l$-"Z-dominated" SR as well as the $WZ/W\gamma^*$ process. Figure 2.15 shows the ANN output distributions for the WH signal process and the background contributions. The ANN output also shows a considerable rejection power against the background processes other than the $WZ/W\gamma^*$ process. Their distributions peak in the $0.2 \div 0.3$ interval of ANN output as the $WZ/W\gamma^*$ process.

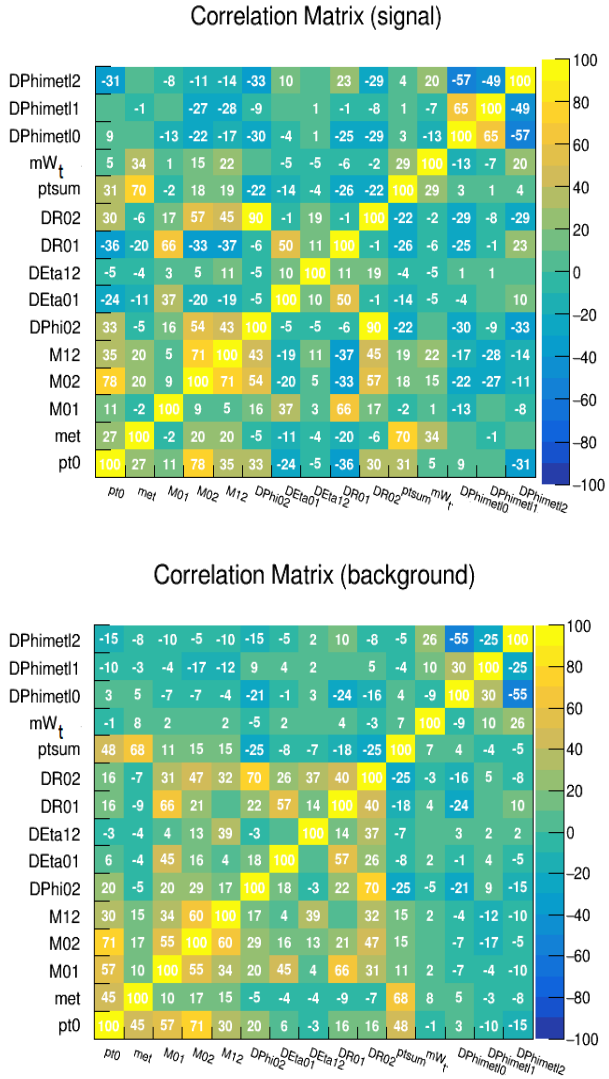
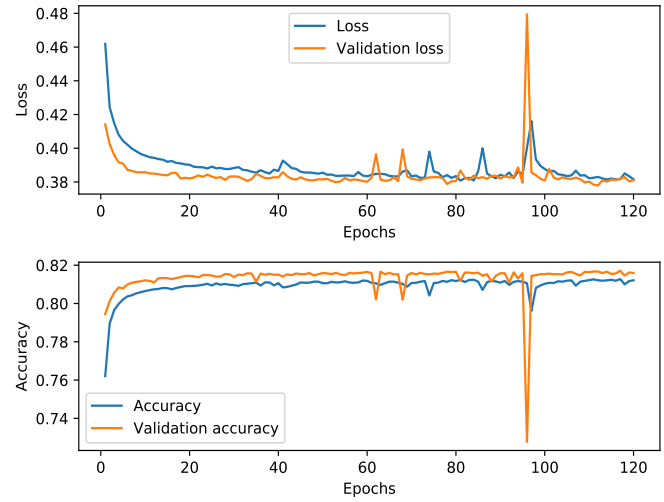
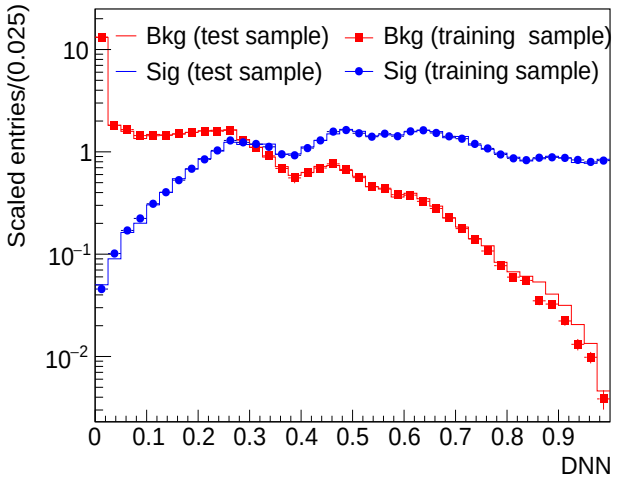


Figure 2.11: Correlation Matrix of the ANN input variables for signal and background.

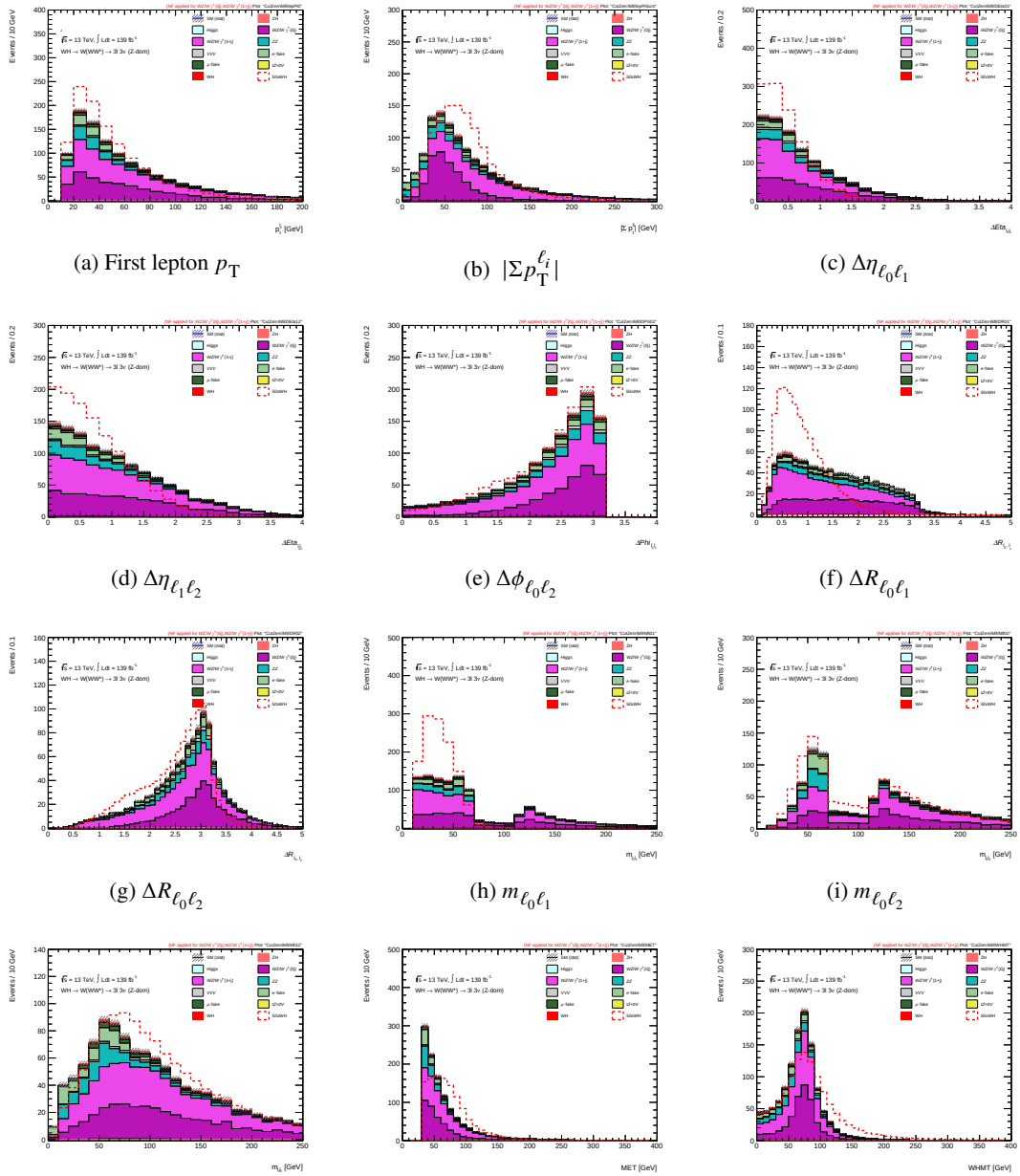


(a)



(b)

Figure 2.12: The ANN epoch evolution (a) and the ANN output distributions for the training and test samples (b).



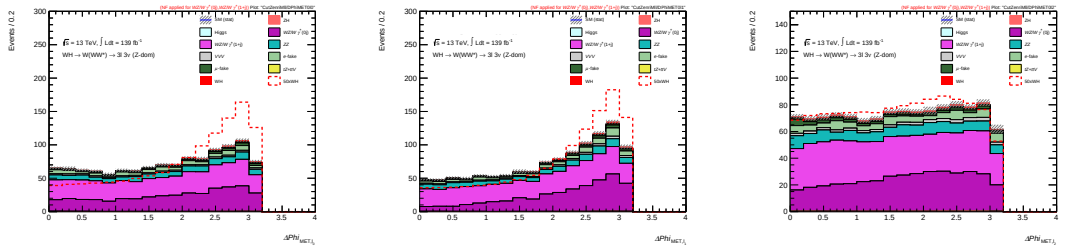
(a) First lepton p_{T} (b) $|\Sigma p_{\mathrm{T}}^{\ell_i}|$ (c) $\Delta\eta_{\ell_0\ell_1}$

(d) $\Delta\eta_{\ell_1\ell_2}$ (e) $\Delta\phi_{\ell_0\ell_2}$ (f) $\Delta R_{\ell_0\ell_1}$

(g) $\Delta R_{\ell_0\ell_2}$ (h) $m_{\ell_0\ell_1}$ (i) $m_{\ell_0\ell_2}$

(j) $m_{\ell_1\ell_2}$ (k) $E_{\mathrm{T}}^{\mathrm{miss}}$ (l) m_{T}^{W}

Figure 2.13: The nominal MC distributions of input variables of the ANN in the $3l$-"Z-dominated" SR. Only MC statistical uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top quark, Z+jets and $Z\gamma$ background. The WH signal is both stacked and superimposed (with a factor of 50). Normalisation factors from a fit to data in the respective WZ CRs (Section 2.9) are applied to the $WZ/W\gamma^*$ processes.



(a) $\Delta\phi_{\ell_0 E_{\mathrm{T}}^{\mathrm{miss}}}$ (b) $\Delta\phi_{\ell_1 E_{\mathrm{T}}^{\mathrm{miss}}}$ (c) $\Delta\phi_{\ell_2 E_{\mathrm{T}}^{\mathrm{miss}}}$

Figure 2.14: The nominal MC distributions of input variables of the ANN in $3l$-"Z-dominated" SR. Only MC statistical uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top quark, Z+jets and $Z\gamma$ background. The WH signal is both stacked and superimposed (with a factor of 50). Normalisation factors from a fit to data in the respective WZ CRs (Section 2.9) are applied to the $WZ/W\gamma^*$ processes.

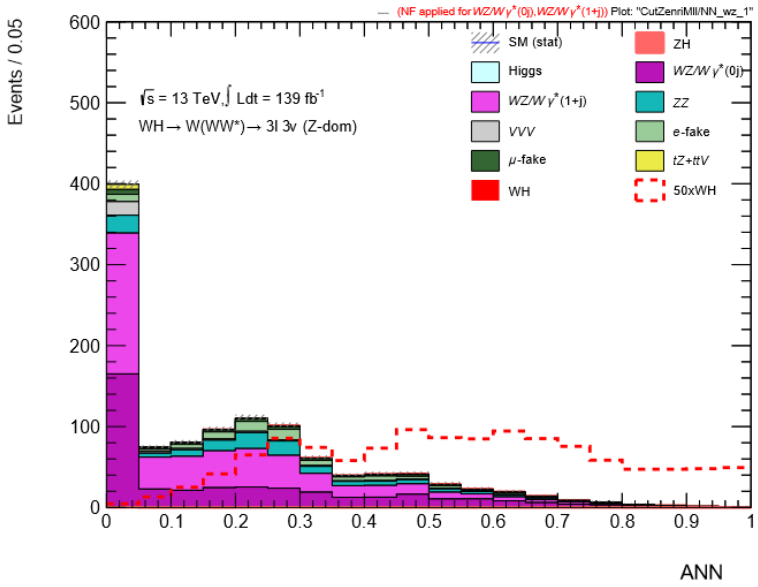


Figure 2.15: The nominal MC distribution of the ANN output in the $3l$-"Z-dominated" SR. Only MC statistical uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top quark, Z+jets and $Z\gamma$ background. The WH signal is both stacked and superimposed (with a factor of 50). Normalisation factors from a fit to data in the respective WZ CRs (Section 2.9) are applied to the $WZ/W\gamma^*$ processes.

2.8.3 Data/MC agreement of the ANN input variables

The data/MC agreement is checked in a region with a negligible expected signal content by defining a vetoed dataset in signal region side-bands.
Other than satisfying requirements in Table 2.3 with at least a $SFOS$, the events of this region fulfil at least one of the cuts in Table 2.14, e.g. the cuts obtained by inverting the selection of $3l$-"Z-dominated" SR, shown in Table 2.9. The MC yields of the WH process, total background and observed data

Table 2.14: The definition of the $3l$-"Z-dominated" SR side band region.

Quantities	Cuts and vetos
b-jets veto	N_{btag}>0
$E_{\mathrm{T}}^{\mathrm{miss}}$ cut	$E_{\mathrm{T}}^{\mathrm{miss}}$ <30 GeV
Dilepton mass cuts	$\lvert m_{\ell\ell} - M_Z \rvert$ < 25 GeV
$m_{\ell\ell}^{min}$	<12GeV

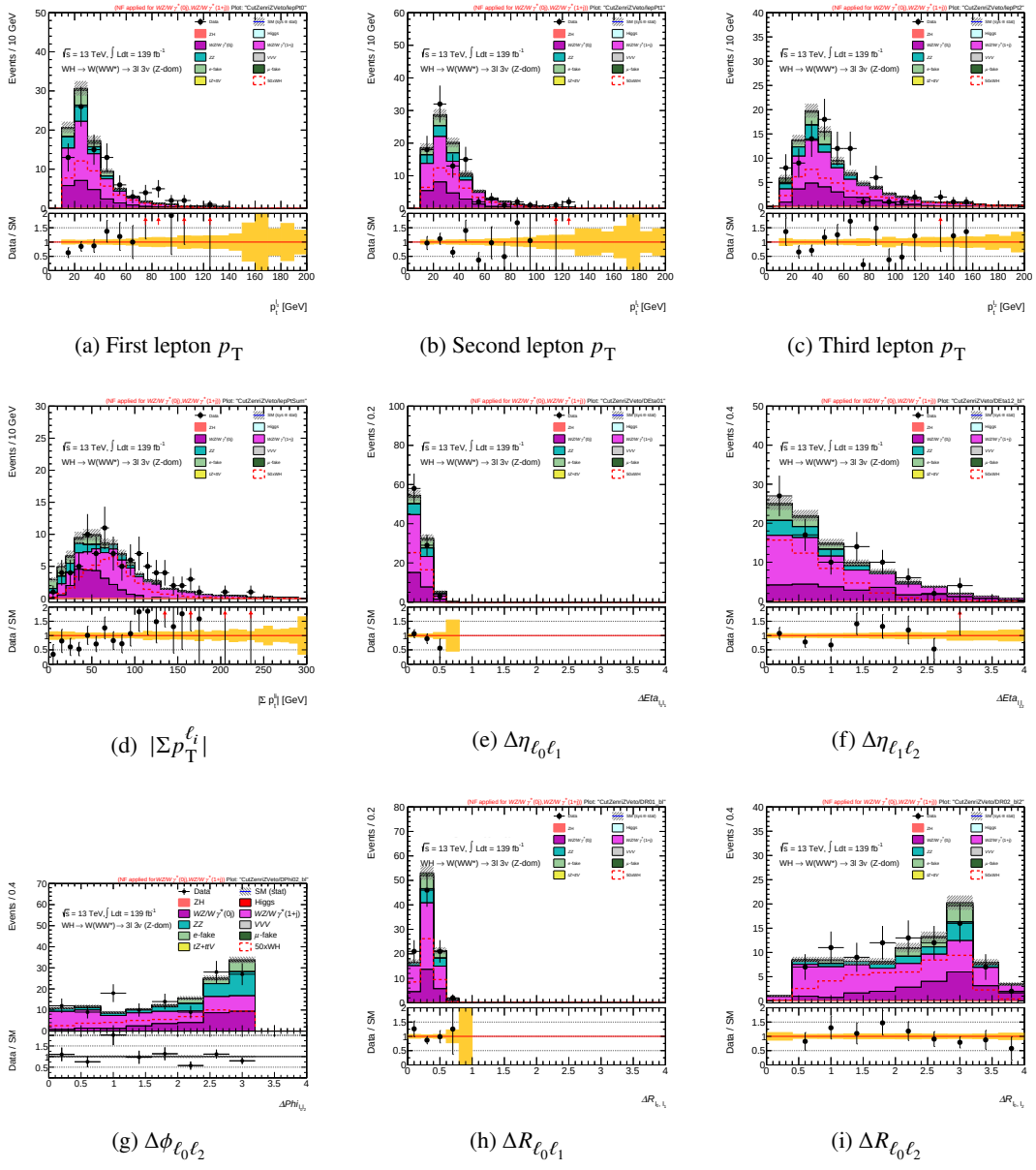
are reported in Table 2.15, after the different cuts and vetos of $3l$-"Z-dominated" SR (in Table 2.9). The data/MC ratio is shown in the last column of Table 2.15, where a good data/MC agreement is observed.
The MC yields of the single background processes that form the total background are reported in Table 2.16, where $WZ/W\gamma^*$ remains the dominant contribution.
A satisfying data/MC agreement is also observed in the distribution shapes of the most relevant variables in Figures 2.16 and 2.17 and the ANN output in Figure 2.18.

Table 2.15: Cutflow for the event selection in the $3l$-"Z-dominated" SR side band region. Yields for the WH signal process, total background and observed data. MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective WZ CRs (Section 2.9) are applied to the $WZ/W\gamma^*$ processes.

$\sqrt{s} = 13TeV$, $\mathcal{L} = 139fb^{-1}$ (Full Run 2)	WH	Total Bkg (dd)	data	Data/TotMC
Scale factors		NFs Applied		
Preselection	58.00 ± 0.10	29002.43 ± 37.87	30401	1.04 ± 0.01
1 or 2 SFOS	19.74 ± 0.06	27625.56 ± 36.91	28883	1.04 ± 0.01
E_T^{miss} > 30 GeV	12.28 ± 0.05	16986.33 ± 26.06	17642	1.03 ± 0.01
bjet-veto	10.50 ± 0.04	14527.47 ± 25.24	14580	1.00 ± 0.01
Z-veto	0.89 ± 0.01	92.54 ± 4.56	90	0.95 ± 0.10

Table 2.16: Cutflow for the event selection in the $3l$-"Z-dominated" SR side band region. Yields for the ZH process (not included in the total background column) and each single background component. MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective WZ CRs (Section 2.9) are applied to the $WZ/W\gamma^*$ processes.

(Full Run 2)	ZH	$WZ/W\gamma^*_{w0jet}$	$WZ/W\gamma^*_{wjets}$	ZZ^*	VVV	tV+ttV	μ-fakes	e-fakes	other Higgs
Scale factors		NF = 1.04	NF = 0.90						
Preselection	94.17 ± 0.47	9937.73 ± 24.99	10655.06 ± 16.77	3999.93 ± 15.19	120.60 ± 0.23	1299.63 ± 2.45	748.71 ± 3.94	2146.03 ± 16.61	94.74 ± 0.46
1 or 2 SFOS	90.00 ± 0.46	9537.90 ± 24.42	10145.81 ± 16.38	3872.39 ± 14.89	66.77 ± 0.20	1222.71 ± 2.34	665.58 ± 3.73	2041.67 ± 16.02	72.72 ± 0.41
E_T^{miss} > 30 GeV	60.67 ± 0.36	6557.31 ± 20.11	7639.23 ± 13.95	1031.10 ± 5.50	55.27 ± 0.19	1048.20 ± 2.18	290.48 ± 2.44	303.97 ± 6.25	60.77 ± 0.34
bjet-veto	51.11 ± 0.34	6442.37 ± 19.91	6674.29 ± 13.40	870.39 ± 5.07	44.35 ± 0.17	98.57 ± 0.56	163.16 ± 1.88	218.99 ± 5.59	15.35 ± 0.29
Z-veto	0.07 ± 0.01	23.98 ± 1.07	47.66 ± 1.04	10.51 ± 0.72	0.22 ± 0.01	0.35 ± 0.04	1.62 ± 0.20	8.12 ± 1.21	0.07 ± 0.02

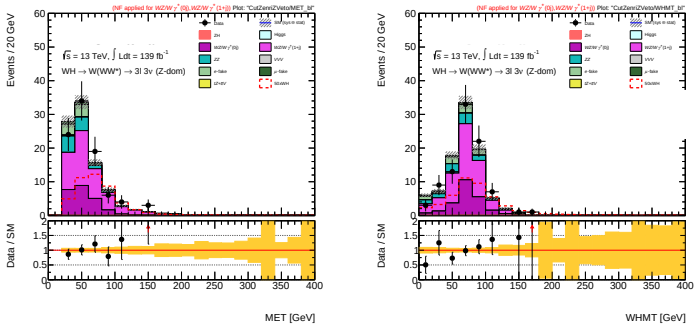


(a) First lepton p_{T} (b) Second lepton p_{T} (c) Third lepton p_{T}

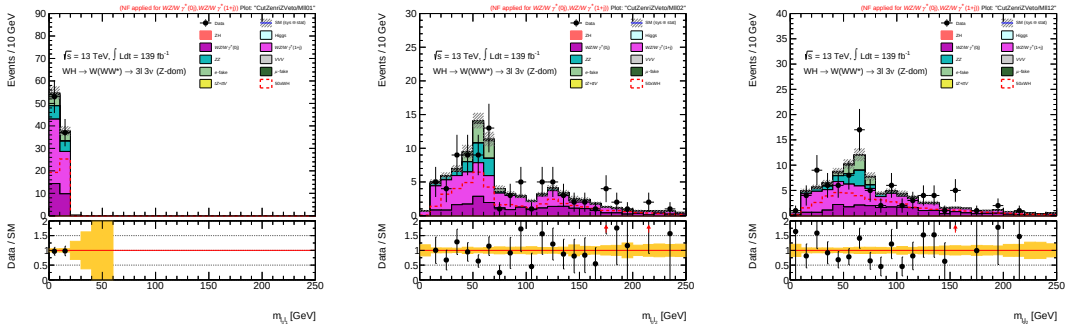
(d) $|\Sigma p_{\mathrm{T}}^{\ell_i}|$ (e) $\Delta\eta_{\ell_0\ell_1}$ (f) $\Delta\eta_{\ell_1\ell_2}$

(g) $\Delta\phi_{\ell_0\ell_2}$ (h) $\Delta R_{\ell_0\ell_1}$ (i) $\Delta R_{\ell_0\ell_2}$

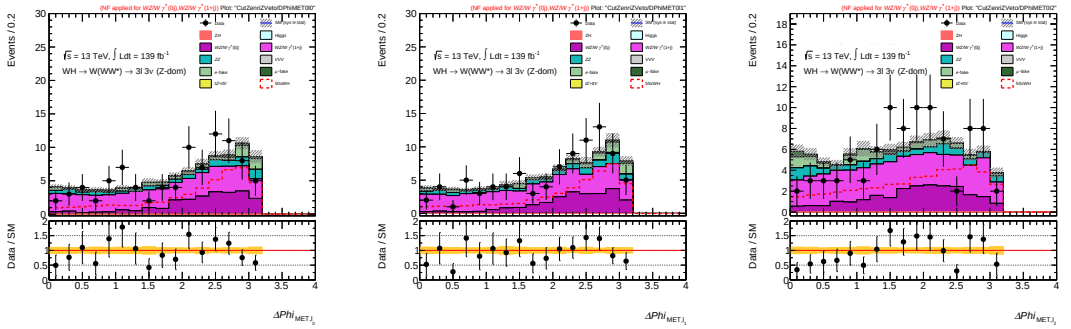
Figure 2.16: The distributions of the relevant variables in the 3l-"Z-dominated" SR side band region. MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top, Z+jets and Zγ background. Normalisation factors from a fit to data in the respective WZ CRs (Section 2.9) are applied to the $WZ/W\gamma^*$ processes.



(a) $E_{\mathrm{T}}^{\mathrm{miss}}$ (b) m_{T}^{W}



(c) $m_{\ell_0\ell_1}$ (d) $m_{\ell_0\ell_2}$ (e) $m_{\ell_1\ell_2}$



(f) $\Delta\phi_{\ell_0 E_{\mathrm{T}}^{\mathrm{miss}}}$ (g) $\Delta\phi_{\ell_1 E_{\mathrm{T}}^{\mathrm{miss}}}$ (h) $\Delta\phi_{\ell_2 E_{\mathrm{T}}^{\mathrm{miss}}}$

Figure 2.17: The distributions of the relevant variables in the 3l-"Z-dominated" SR side band region. MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective WZ CRs (Section 2.9) are applied to the $WZ/W\gamma^*$ processes.

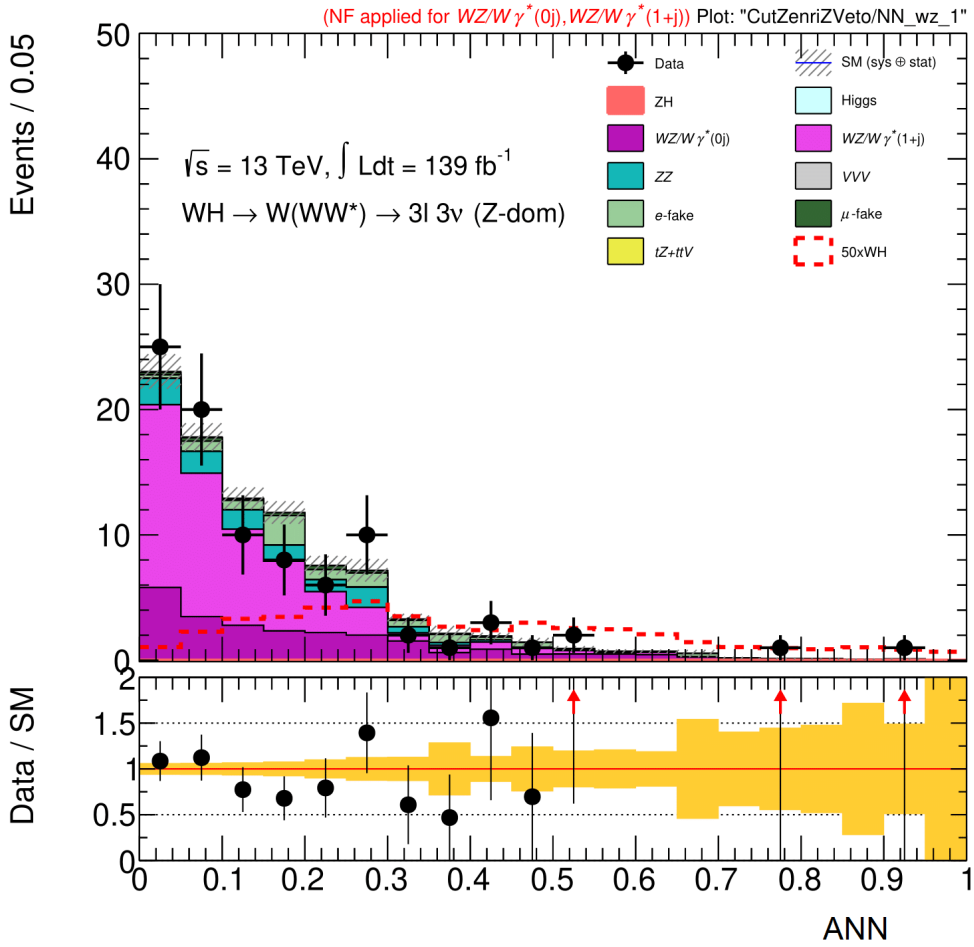


Figure 2.18: The ANN distribution in the 3l-"Z-dominated" SR side band region. Only MC statistical uncertainties are shown. MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons (e-fakes and μ-fakes) from top, Z+jets and Zγ background. Normalisation factors from a fit to data in the respective WZ CRs (Section 2.9) are applied to the $WZ/W\gamma^*$ processes.

2.9 3ℓ WZ Control Regions and Top Validation Region

Further checks on the MC modelling of the ANN output and the variables in Figures 2.16 and 2.17 are made in two $WZ/W\gamma^*$ control regions (WZ_{w0jet} CR and WZ_{wjets} CR) and in a validation region (Zdom-top VR), presented in this section.
The $WZ/W\gamma^*$ process is the main source of background in the $3l$-"Z-dominated" SR, and dedicated control regions are defined to determine the normalisation of the $WZ/W\gamma^*$ background in common to the $2l$ WH and $3l$ WH channels.
Firstly, a $WZ/W\gamma^*$ CR is defined by reversing the Z-veto cut in Table 2.9. It means to require at least one SFOS lepton pair with the invariant mass within 25 GeV from the Z-resonance that assures the orthogonality with respect to $3l$-"Z-dominated" SR. Looking at the distribution of the number of jets (Figure 2.19) in the $WZ/W\gamma^*$ CR, the data/MC ratio decreases as a function of the number of jets.
Then, the $WZ/W\gamma^*$ control region is separated in terms of the number of jets into the WZ_{w0jet} CR

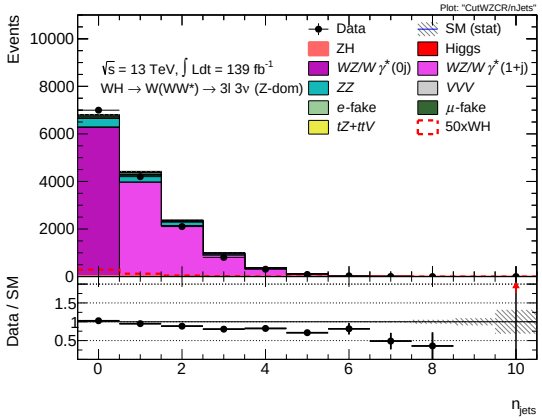


Figure 2.19: The distribution of the number of jets in the $WZ/W\gamma^*$ CR without the application of the normalisation factor on the $WZ/W\gamma^*$ background.

for events without jets and the WZ_{wjets} CR for events with $n_{jets} \geq 1$.
For these two classes of $WZ/W\gamma^*$ events, the normalisation factors are estimated by performing a profile likelihood fit to the data in the respective CRs, and their values are reported in Table 2.17.
In the previous sections, these normalisation factors have been already applied.

Table 2.17: Normalisation factors from the profile likelihood fit to the data of the WZ_{w0jet} CR and WZ_{wjets} CR.

CRs	NF value
WZ_{w0jet} CR	$1.040^{0.074}_{-0.066}$
WZ_{wjets} CR	$0.897^{0.169}_{-0.156}$

Other relevant contributions to the background are the processes that include the top quark. When they involve misidentified lepton candidates ($t\bar{t}$ and tW processes), they are dealt with using data-driven methods (see Section 2.6).

A validation region (Zdom-top VR) is, then, defined to check the modelling of these top backgrounds. It is defined by the presence of at least one SFOS, at least a jet and one b-jet. Table 2.18 also summarises the target backgrounds and the differences in the CRs/VR definitions with respect to the $3l$-"Z-dominated" SR.
Tables 2.19 and 2.20 show the background composition and data yields in the CRs/VR. The data/TotMC

Table 2.18: Summary of the control and validation regions. The WZ CR line is referred to both WZ_{w0jet} CR and WZ_{wjets} CR.

Name	Target background	Changes w.r.t. reference SR
WZ CR	$WZ/W\gamma^*$	$\geq$ 1 SFOS pair with $\lvert m_{\ell\ell} - m_Z \rvert$ <25 GeV
Zdom-top VR	top	at least 1 jet, one b-tagged jet

values for the $WZ/W\gamma^*$ are coherent with the fact that normalisation factors in Table 2.17 have been already applied. In the Zdom-top VR, the agreement between data and MC modelling is satisfying since the VR is statistically limited.

Table 2.19: Cutflow with Monte Carlo and data yields in the WZ_{w0jet} CR, WZ_{wjets} CR and Zdom-Top VR. MC yields for the WH signal process, total background and data. MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective WZ CRs are applied to the $WZ/W\gamma^*$ processes.

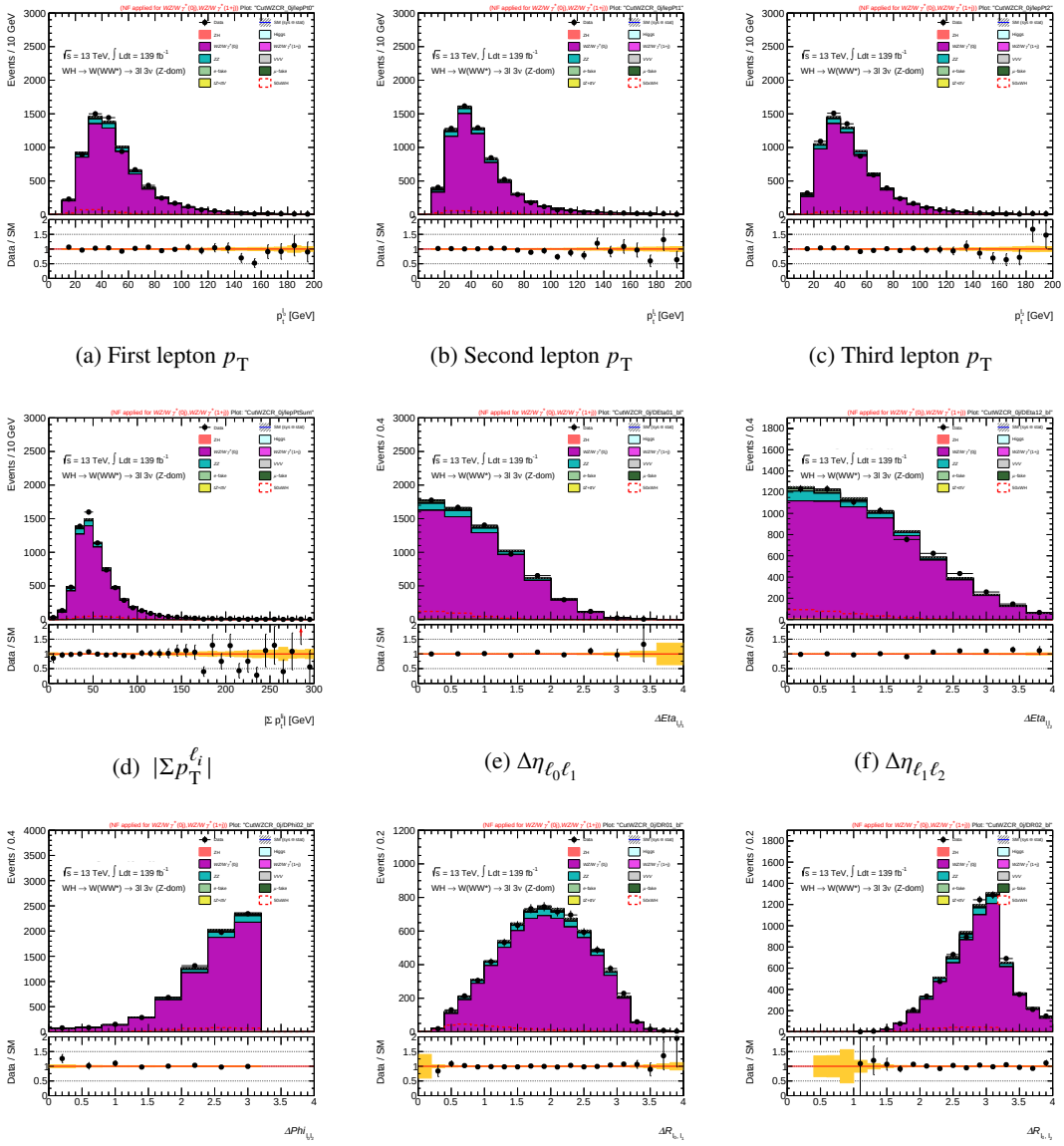
$\sqrt{s} = 13TeV$, $\mathcal{L} = 139fb^{-1}$ (Full Run 2)	WH	Total Bkg	data	Data/TotMC
Scale factors		NFs Applied		
WZ CR 0j	5.72 ± 0.03	6893.61 ± 20.28	6916	1.00 ± 0.01
WZ CR 1j	3.61 ± 0.02	7451.31 ± 14.73	7476	1.00 ± 0.01
Zdom-top VR	0.77 ± 0.01	158.53 ± 13.9	193	1.20 ± 0.14

Table 2.20: Cutflow with Monte Carlo and data yields in the WZ_{w0jet} CR, WZ_{wjets} CR and Zdom-Top VR. MC yields for the ZH process (not included in the total background column) and every single background contribution. MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top, Z+jets and Zγ background. Normalisation factors from a fit to data in the respective WZ CRs are applied to the $WZ/W\gamma^*$ processes.

(Full Run 2)	ZH	$WZ/W\gamma^*_{w0jet}$	$WZ/W\gamma^*_{wjets}$	ZZ^*	VVV	tV+ttV	μ-fakes (dd)	e-fakes (dd)	other Higgs
Scale factors		NF = 1.04	NF = 0.90						
WZ CR 0j	18.36 ± 0.23	6383.63 ± 19.84	0.00	342.21 ± 2.86	15.82 ± 0.09	3.47 ± 0.09	73.32 ± 1.26	69.87 ± 2.81	5.29 ± 0.18
WZ CR 1j	32.43 ± 0.25	0.00	6592.29 ± 13.33	508.28 ± 4.02	28.18 ± 0.14	94.38 ± 0.55	86.71 ± 1.37	131.53 ± 4.57	9.93 ± 0.22
Zdom-top VR	0.25 ± 0.02	0.00	41.80 ± 0.75	5.21 ± 0.45	2.42 ± 0.03	52.71 ± 5.82	32.53 ± 0.78	16.16 ± 0.79	7.70 ± 0.07

The distributions of the relevant kinematic variables in the WZ_{w0jet} CR and WZ_{wjets} CR are shown in Figures 2.22 and 2.23, while the ANN output distribution is shown in Figure 2.24.

The same distributions are shown in Figures 2.25, 2.26 and 2.27 in the Zdom-top VR. In this region, the shape of the invariant masses $m_{\ell 0 \ell 1}$ and $m_{\ell 0 \ell 2}$ is determined by the Z-veto cut applied to $SFOS$ leptons, that has been inverted in the $WZ/W\gamma^*$ CRs. No relevant discrepancies in the data/MC modelling of the variable shape are present, confirming the satisfying data/MC agreement in Table 2.19.

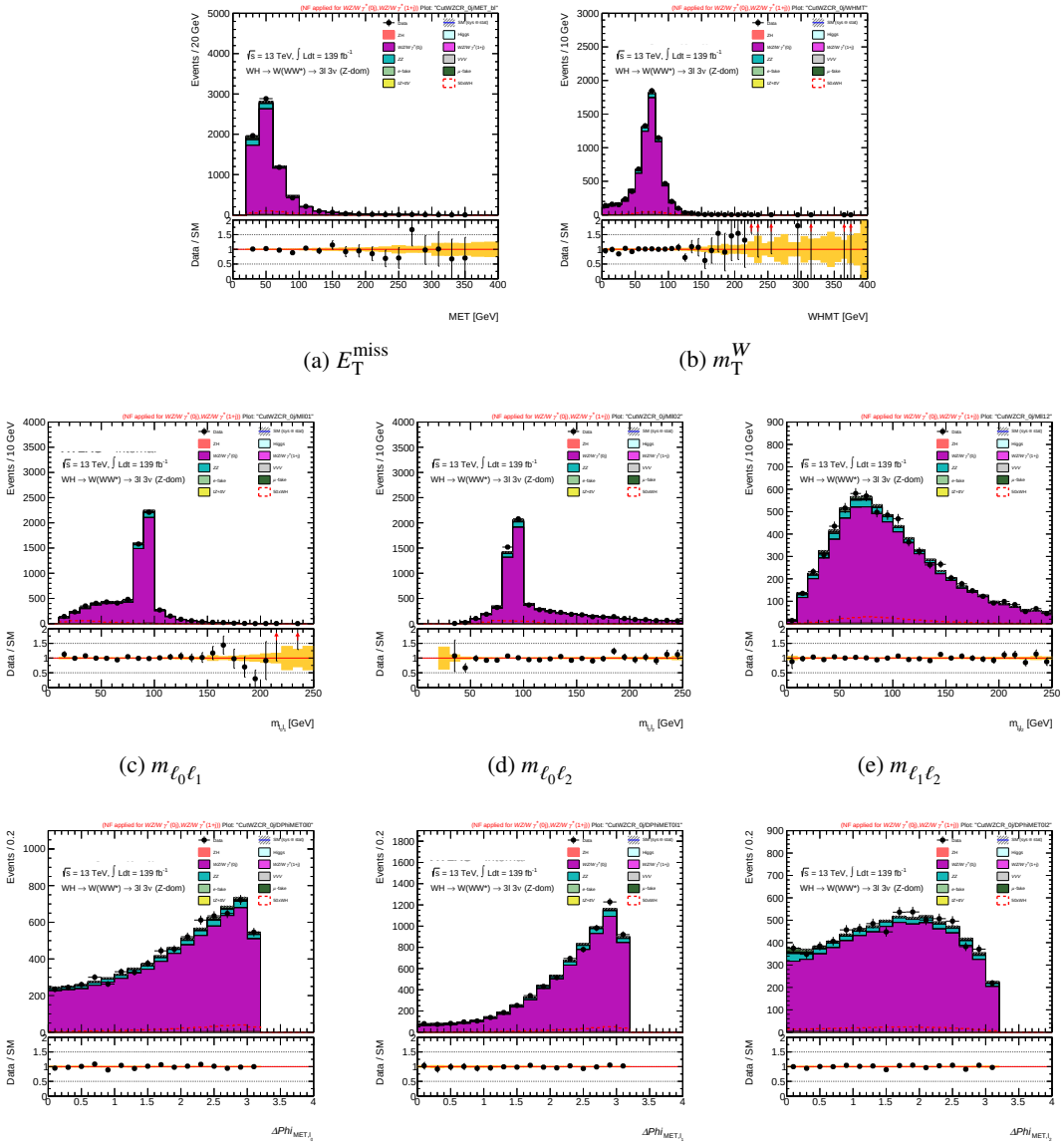


(a) First lepton p_{T} (b) Second lepton p_{T} (c) Third lepton p_{T}

(d) $|\Sigma p_{\mathrm{T}}^{\ell_i}|$ (e) $\Delta\eta_{\ell_0\ell_1}$ (f) $\Delta\eta_{\ell_1\ell_2}$

(g) $\Delta\phi_{\ell_0\ell_2}$ (h) $\Delta R_{\ell_0\ell_1}$ (i) $\Delta R_{\ell_0\ell_2}$

Figure 2.20: The distributions of the relevant variables for $3l$-"Z-dominated" analysis in the WZ_{w0jet} CR without jet. Only MC statistical uncertainties are shown. Data driven methods are used to estimate the misidentified leptons (e-fakes and μ-fakes) from top, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective $WZ/W\gamma^*$ CRs are applied to the $WZ/W\gamma^*$ processes.

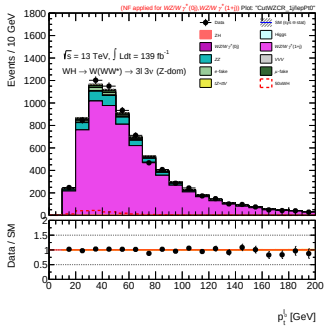


(a) $E_{\mathrm{T}}^{\mathrm{miss}}$ (b) m_{T}^{W}

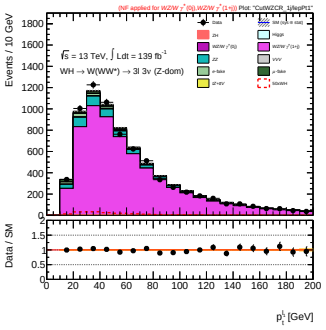
(c) $m_{\ell_0\ell_1}$ (d) $m_{\ell_0\ell_2}$ (e) $m_{\ell_1\ell_2}$

(f) $\Delta\phi_{\ell_0 E_{\mathrm{T}}^{\mathrm{miss}}}$ (g) $\Delta\phi_{\ell_1 E_{\mathrm{T}}^{\mathrm{miss}}}$ (h) $\Delta\phi_{\ell_2 E_{\mathrm{T}}^{\mathrm{miss}}}$

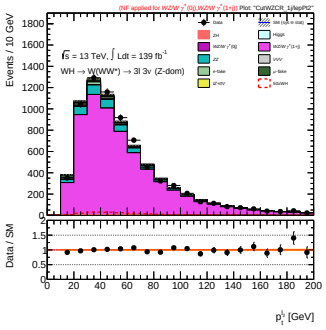
Figure 2.21: The distributions of the relevant variables for 3l-"Z-dominated" analysis in the WZ_{w0jet} CR without jets. MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective $WZ/W\gamma^*$ CRs are applied to the $WZ/W\gamma^*$ processes.



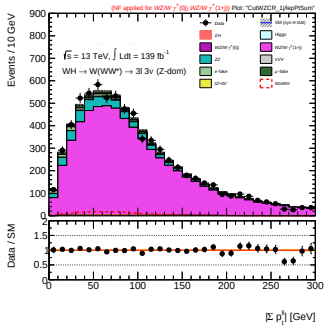
(a) First lepton p_{T}



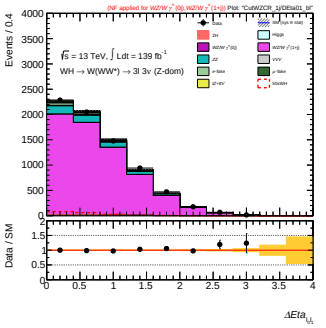
(b) Second lepton p_{T}



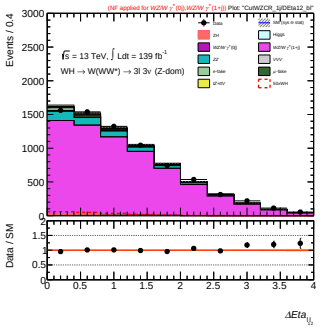
(c) Third lepton p_{T}



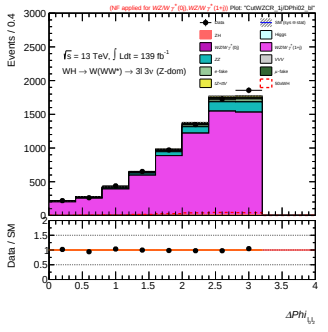
(d) $|\Sigma p_{\mathrm{T}}^{\ell_i}|$



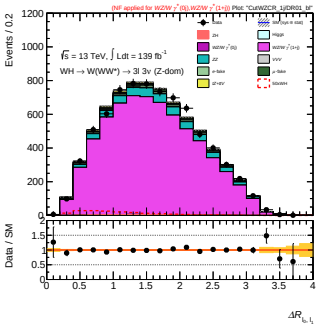
(e) $\Delta\eta_{\ell_0\ell_1}$



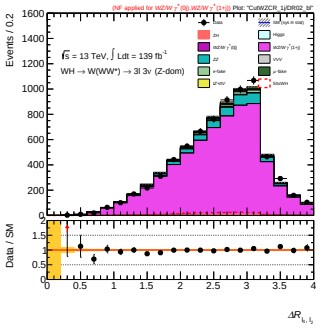
(f) $\Delta\eta_{\ell_1\ell_2}$



(g) $\Delta\phi_{\ell_0\ell_2}$

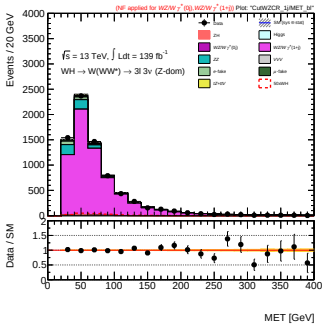


(h) $\Delta R_{\ell_0\ell_1}$

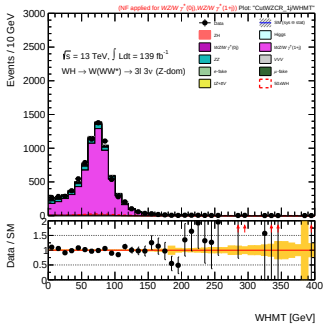


(i) $\Delta R_{\ell_0\ell_2}$

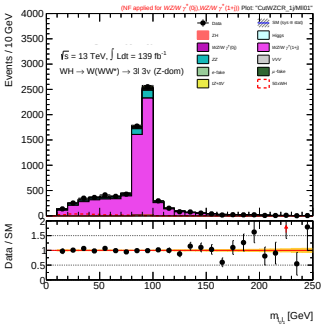
Figure 2.22: The distributions of the relevant variables for 3*l*-"Z-dominated" analysis in the WZ_{wjets} CR with $njets \geq 1$. MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective $WZ/W\gamma^*$ CRs are applied to the $WZ/W\gamma^*$ processes.



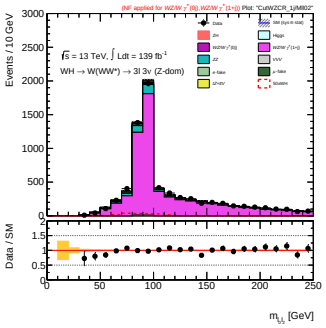
(a) $E_{\mathrm{T}}^{\mathrm{miss}}$



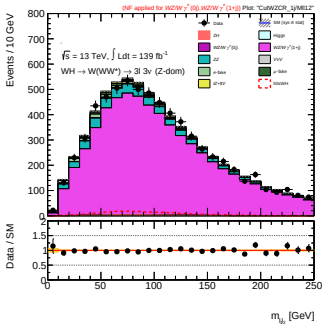
(b) m_{T}^{W}



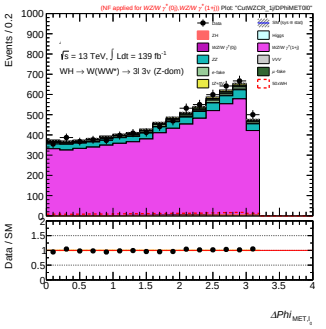
(c) $m_{\ell_0\ell_1}$



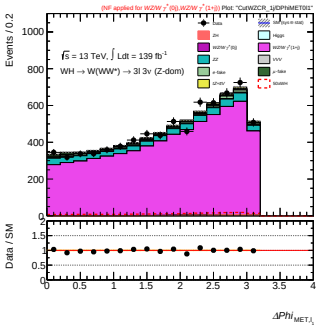
(d) $m_{\ell_0\ell_2}$



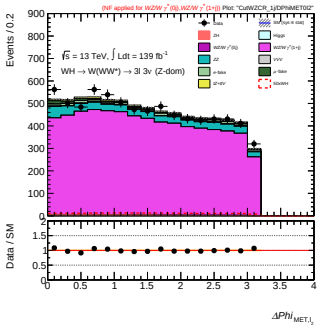
(e) $m_{\ell_1\ell_2}$



(f) $\Delta\phi_{\ell_0 E_{\mathrm{T}}^{\mathrm{miss}}}$

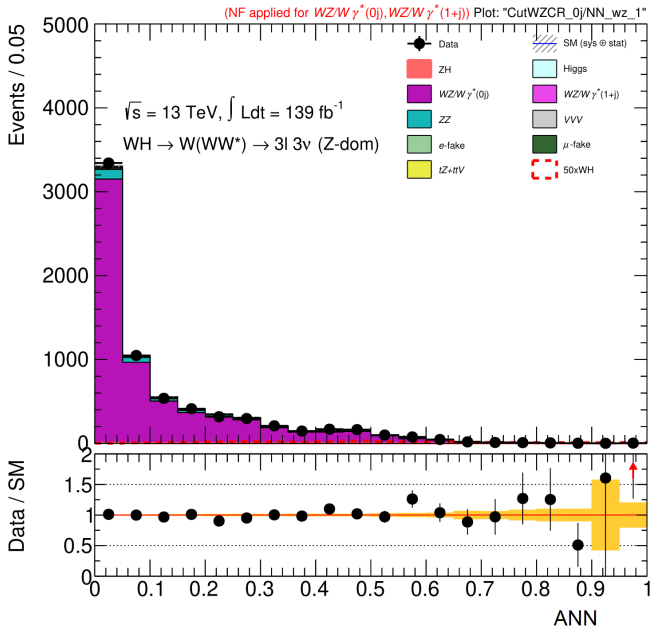


(g) $\Delta\phi_{\ell_1 E_{\mathrm{T}}^{\mathrm{miss}}}$

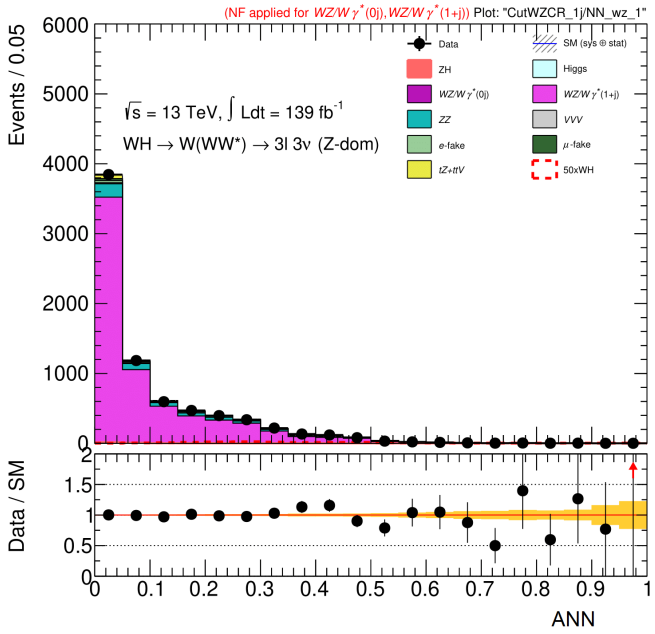


(h) $\Delta\phi_{\ell_2 E_{\mathrm{T}}^{\mathrm{miss}}}$

Figure 2.23: Distributions of the relevant variables for 3l-"Z-dominated" analysis in the WZ_{wjets} CR with $njets \geq 1$. MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective $WZ/W\gamma^*$ CRs are applied to the $WZ/W\gamma^*$ processes.

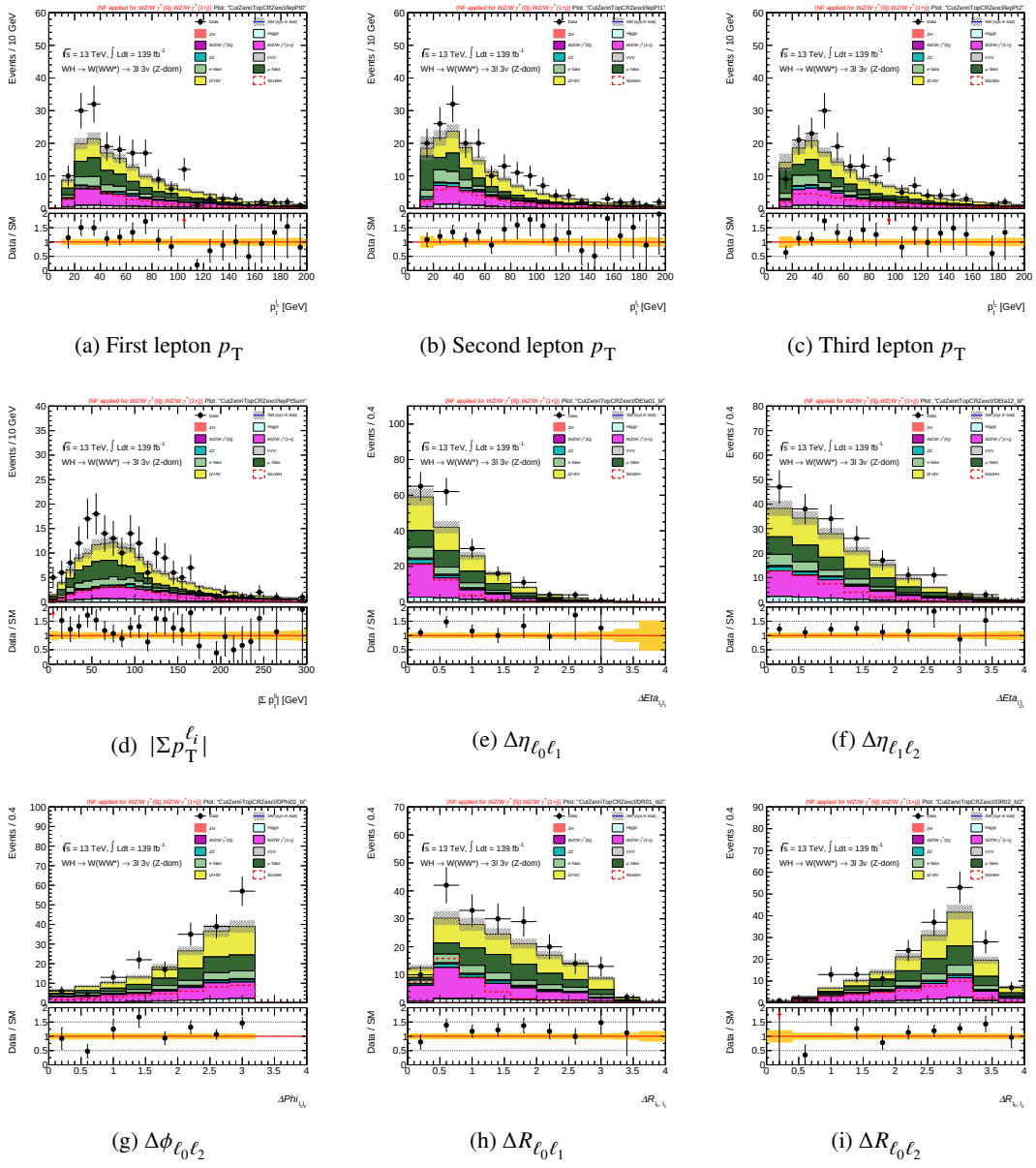


(a) WZ_{w0jet} CR



(b) WZ_{wjets} CR

Figure 2.24: The ANN output distribution in WZ_{w0jet} CR and WZ_{wjets} CR. MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective $WZ/W\gamma^*$ CRs are applied to the $WZ/W\gamma^*$ processes.

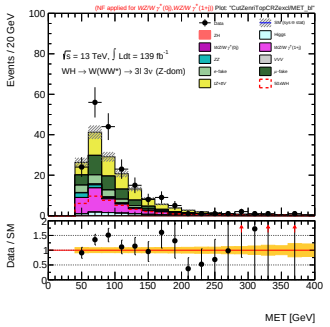


(a) First lepton p_{T} (b) Second lepton p_{T} (c) Third lepton p_{T}

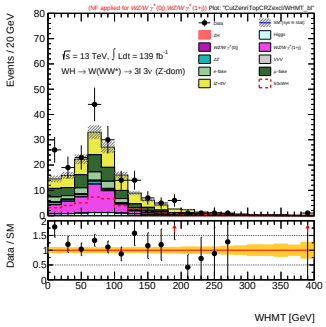
(d) $|\Sigma p_{\mathrm{T}}^{\ell_i}|$ (e) $\Delta\eta_{\ell_0\ell_1}$ (f) $\Delta\eta_{\ell_1\ell_2}$

(g) $\Delta\phi_{\ell_0\ell_2}$ (h) $\Delta R_{\ell_0\ell_1}$ (i) $\Delta R_{\ell_0\ell_2}$

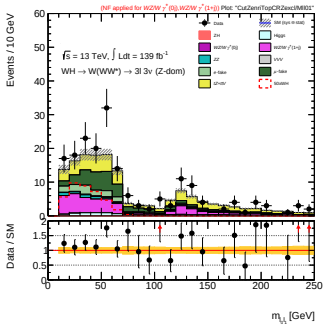
Figure 2.25: Distributions of the relevant variables in the $3l$-"Z-dominated" analysis in the Zdom-Top VR. MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective $WZ/W\gamma^*$ CRs (Section 2.9) are applied to the $WZ/W\gamma^*$ processes.



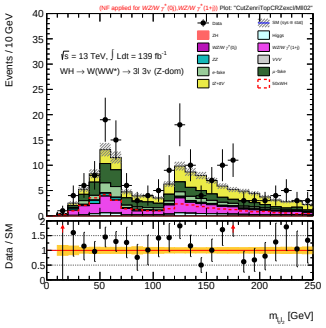
(a) $E_{\mathrm{T}}^{\mathrm{miss}}$



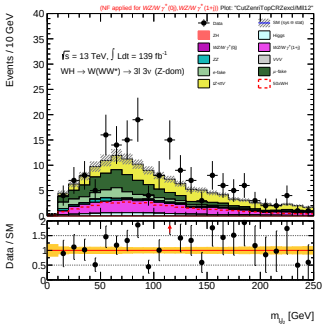
(b) m_{T}^{W}



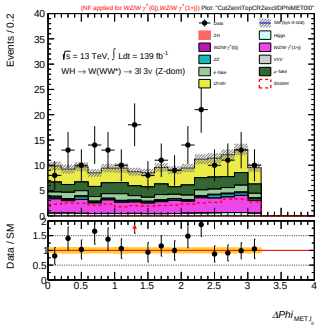
(c) $m_{\ell_0\ell_1}$



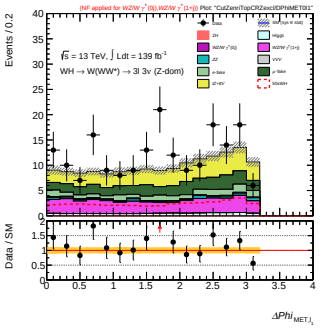
(d) $m_{\ell_0\ell_2}$



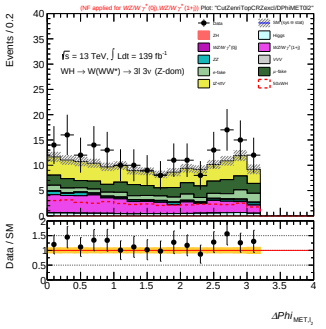
(e) $m_{\ell_1\ell_2}$



(f) $\Delta\phi_{\ell_0 E_{\mathrm{T}}^{\mathrm{miss}}}$



(g) $\Delta\phi_{\ell_1 E_{\mathrm{T}}^{\mathrm{miss}}}$



(h) $\Delta\phi_{\ell_2 E_{\mathrm{T}}^{\mathrm{miss}}}$

Figure 2.26: Distributions of the relevant variables in the 3*l*-"Z-dominated" analysis in the Zdom-Top VR. MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective $WZ/W\gamma^*$ CRs (Section 2.9) are applied to the $WZ/W\gamma^*$ processes.

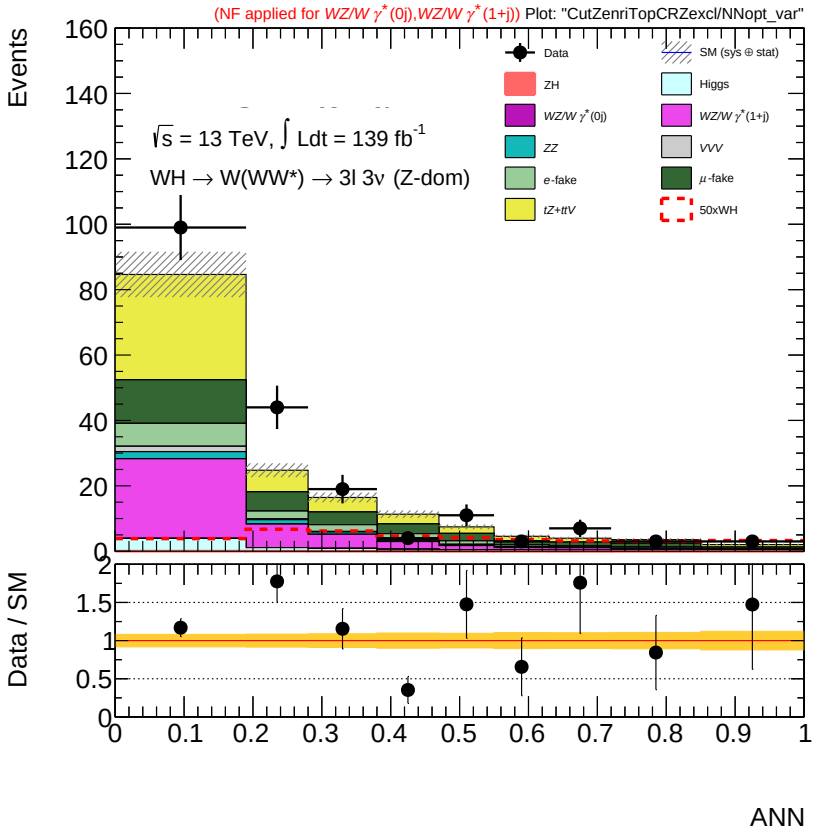


Figure 2.27: ANN distribution of the $3l$-"Z-dominated" analysis in the Zdom-Top VR with optimised binning (see Section 2.11.2). MC statistical and experimental systematic uncertainties are shown. Data driven methods are used to estimate the misidentified leptons, e-fakes and μ-fakes, from top, Z+jets and $Z\gamma$ background. Normalisation factors from a fit to data in the respective $WZ/W\gamma^*$ CRs (Section 2.9) are applied to the $WZ/W\gamma^*$ processes.

2.10 Systematic uncertainties

Systematic uncertainties on the simulated signal and background processes are discussed in this section.
They are divided into theoretical and experimental uncertainties. Considering the MC yield $Y_{MC} = \sigma^{th}_{MC} A^{th} \epsilon^{exp}_{Tot} L^{exp}_{int}$, the systematic uncertainties on cross-section σ^{th}_{MC} and acceptance A^{th} are referred to as theoretical. While the uncertainties on the efficiency ϵ^{exp}_{Tot} and integrated luminosity L^{exp}_{int} are referred to as experimental.

2.10.1 Theoretical uncertainties

For $3l$-"Z-dominated" analysis, the theoretical uncertainties are evaluated on the signal WH process and on the most relevant backgrounds, $WZ/W\gamma^*$ and ZZ^* processes. Their effects are also checked on the shape of the ANN output distribution as the ANN is exploited in the profile likelihood fit in Section 2.11 to evaluate the expected (and observed) signal strength parameter.
The total theoretical systematic uncertainties are reported in Table 2.21. In addition to the cross-section and acceptance uncertainties, other contributions are included depending on whether the processes are signal or backgrounds.

Table 2.21: List of theory systematic uncertainties considered in the $3l$-"Z-dominated" analysis. All uncertainties are given as relative variations on the yields. The yields are assumed to be taken in the region where the MVA analysis is done. Uncertainties from multiple sources are added in quadrature. In the last column, only the relative constraints > 1% are reported.

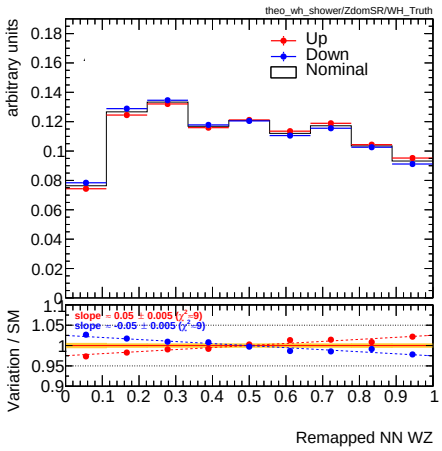
Systematic uncertainty	region	Value [%]	Relative constraint on expected signal strength parameter [%]
WH			
Higgs branching ratio (only in the μ measurement)	SR	±1.5	< 1
Cross section (PDF+α_s, QCD scales: only in the μ measurement)	SR	+2.3/-2.7	< 1
Acceptance (PDF+α_s, QCD scales)	SR	+3.6/-3.8	< 1
Parton shower modelling	SR	±3.8	< 1
Background Processes			
$WZ/W\gamma^*$ without jets			
Total (PDF+α_s, QCD scales, jet merging/resummation scale, parton recoil scheme)	SR	+5.3/-5.6	8.22 (CSSKIN), 1.38 (CKKW)
	WZ_{w0jet} CR	+3.4/-3.9	
$WZ/W\gamma^*$ with at least a jet			
Total (PDF+α_s, QCD scales, jet merging/resummation scale, parton recoil scheme)	SR	+20/-16	< 2.72 (CSSKIN), 0.61 (QSF)
	WZ_{jets} CR	+20/-15	
$qq \to ZZ$			
Total (PDF+α_s, QCD scales, jet merging/resummation scale, parton recoil scheme)	SR	+11/-9	2.17 (QCD scale)
WWW			
Total (MC modelling)	SR	±8.6	< 1

For the signal process, the cross-section normalisation (PDF+α_s and QCD scale) and Higgs branching ratio uncertainties are assigned following the recommendations of the Higgs Working Group [68].
The parton shower modelling uncertainty is also evaluated for the signal process. It is calculated by comparing the WH samples simulated using the nominal Pythia8 and alternative Herwig7 generators at the truth level.
The acceptance uncertainty includes the contributions from the three sources: PDF, strong coupling constant α_s, and QCD scale.
For the signal process, the PDF uncertainty is calculated as the quadrature sum of differences between the nominal and alternative variations of the PDF4LHC [59] PDF set, following the recommendation of the LHC Higgs Working Group.
The strong coupling constant α_s scale uncertainty is evaluated by varying up and down ($\pm$ 0.0015) with respect to the nominal value $\alpha_s(m_Z) = 0.118$.
The QCD scale uncertainty is calculated as the envelope of the up and down variations of the renormalisation (μ_R) and factorisation (μ_F) scales. The following variation combinations (μ_R,μ_F) from the nominal values (μ_R^0, μ_F^0) are considered:

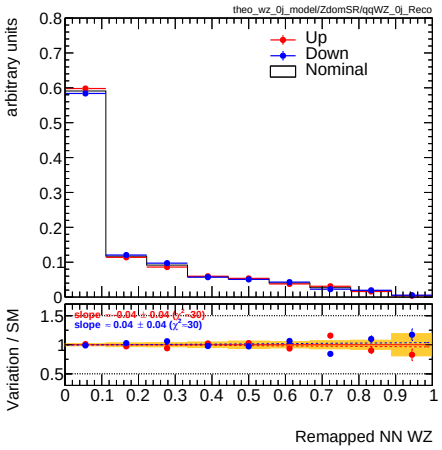
$$(\mu_R, \mu_F) \in \left\{(\mu_R^0, \mu_F^0), (0.5\mu_R^0, \mu_F^0), (\mu_R^0, 0.5\mu_F^0), (2\mu_R^0, \mu_F^0), (\mu_R^0, 2\mu_F^0), (0.5\mu_R^0, 0.5\mu_F^0), (2\mu_R^0, 2\mu_F^0)\right\} \quad (2.21)$$

The uncertainties on the $WZ/W\gamma^*$ yields are evaluated in the WZ_{w0jet} and WZ_{wjets} control regions, as well as in the $3l$-"Z-dominated" SR.
The PDF, strong coupling constant α_s, and QCD scales uncertainties are evaluated using the same procedure. The PDF uncertainty is calculated using 100 alternative variations associated with the NNPDF3.0 PDF set in the case of $WZ/W\gamma^*$ and ZZ^* processes.
Moreover, the jet merging scale (CKKW in Table 2.21), the jet resummation scale (QSF in Table 2.21), and the parton recoil scheme (CSSKIN in Table 2.21) uncertainties are evaluated. They are calculated using truth level comparisons between the nominal and varied samples. In the calculations of the jet merging scale uncertainty, the nominal 30 GeV cut is varied down to 25 GeV and up to 35 GeV. While, in the jet resummation scale uncertainty, the nominal scale is varied up and down by a factor of 2.
In the parton recoil scheme uncertainty, the nominal sample is compared with an alternative sample on which a different parton recoil scheme [69] is applied. The nominal parton recoil scheme is described in Reference [60]. These comparisons are performed at the truth level, and the variations are symmetrised to produce an uncertainty estimate on the nominal samples.
The prescriptions of the uncertainties on the $WZ/W\gamma^*$ yields are also valid for the ZZ^* process. The ZZ^* process is the second most relevant background in the $3l$-"Z-dominated" SR, but it is not so significant as to have a dedicated CR. Therefore the uncertainties on the ZZ^* yields are only evaluated in the $3l$-"Z-dominated" SR.
Moreover, the MC modelling uncertainties are evaluated for $q\bar{q} \rightarrow WZ$ and $q\bar{q} \rightarrow ZZ^*$ backgrounds, by comparing the samples from the nominal Sherpa and alternative Powheg generators. They are not included in the profile likelihood fit on the ANN output distribution, but they are used as checks to individuate shape trends.
Even if its small contribution in the $3l$-"Z-dominated" SR, the WWW process is also considered in the study on the MC modelling uncertainty as it has the same signature of the WH signal process in the SR. For the WWW process sample, the nominal sample from the Sherpa generator is compared with the alternative sample using MadGraph5_aMC@NLO generators.

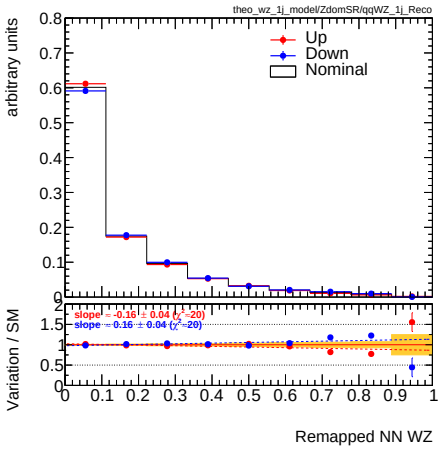
Figure 2.28 shows the ANN distributions on which the parton shower modelling (for the signal WH process) and MC modelling uncertainties (for the background processes) are evaluated. No significant shape variations are observed, except for the trend in the last bins of the WWW process where the statistic is limited in any case. In the last column of Table 2.21, the relative constraints of the theoretical uncertainties on the signal strength of the $3l$-"Z-dominated" channel are reported. The uncertainties on the signal WH process contribute less than 1% to the constraints on the signal strength parameter. Among the background processes, the largest constraints are due to the parton recoil scheme uncertainty on the $WZ/W\gamma^*$ yields and due to the QCD scale uncertainty on the ZZ^* yields.
As reported in the section on the estimation procedures, some uncertainties are evaluated at the truth level. Then, the reconstruction-level and truth-level distributions of the ANN input variables and ANN output are also compared for the WH, $WZ/W\gamma^*$, and ZZ^* processes to additionally check the estimated uncertainties. Most of the cases are compatible, and the observed differences are not relevant.



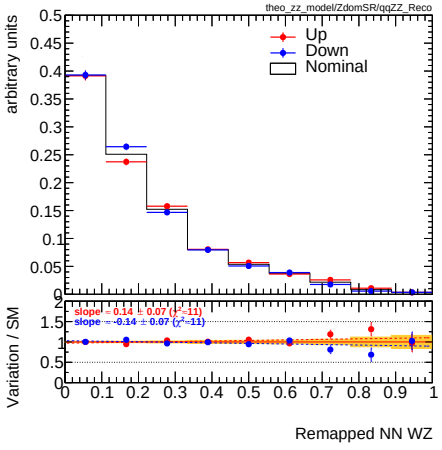
(a) WH - parton shower modelling



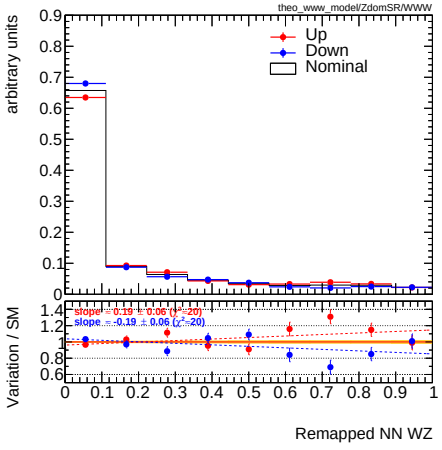
(b) $WZ/W\gamma^*$ with 0 jet - MC modelling



(c) $WZ/W\gamma^*$ with 0 jet - MC modelling



(d) ZZ - MC modelling



(e) WWW - MC modelling

Figure 2.28: Parton shower and MC modelling uncertainties.

2.10.2 Experimental uncertainties

The recommendations of ATLAS Combined Performance groups are followed for evaluating the experimental uncertainties.
The contributions of these uncertainties can be divided into those that affect the integrated luminosity, pile-up, and the reconstructed objects: electrons, muons, jets and E_T^{miss}. The experimental uncertainties are reported in Table 2.22, together with their constraints on the signal strength parameter.

The integral luminosity uncertainty is ±1.7%. It is evaluated for the total integrated luminosity of the full Run 2 pp dataset.
The pile-up uncertainties have two contributions related to the data scale factor μ rescaling and pile-up event rejection.
In the $3l$-"Z-dominated" analysis, the data scale factor μ is varied between $1/0.99$ and $1/1.07$ around the nominal $\mu_0 = 1/1.03$ to evaluate the systematic uncertainty on the data scale factor μ rescaling of Monte Carlo samples.
The pile-up rejection uncertainty is associated with the pile-up jet rejection tool efficiency. It is estimated by varying the JVT criteria [45].
For each reconstructed object, the systematic experimental uncertainties can be additionally split into several contributions according to the uncertainty source.
Following the row order of Table 2.22, the uncertainties associated with the lepton reconstruction are: the trigger efficiency uncertainty, the reconstruction and identification uncertainties, electron energy scale/resolution uncertainties, muon momentum scale/resolution uncertainties, isolation uncertainty and muon track-to-vertex-association (TTVA) uncertainty.
The trigger efficiency is directly estimated from data using the tag-and-probe method described in detail in Reference [70]. For electrons, it is calculated by exploiting the resonances $Z \rightarrow ee$ (above the energy of 15 GeV) or $J/\Psi \rightarrow ee$ (below the energy of 15 GeV). Similarly, the J/Ψ and Z resonances decaying in a dimuon system are used to measure the trigger efficiency of low p_T and moderate p_T muons, respectively. For high p_T muons, two semileptonic $t\bar{t}$ and W+jets events are considered.
Then, the trigger efficiency uncertainty is evaluated by varying one at a time the requirements on the uncertainty sources during the efficiency measurement.
The tag-and-probe method is also used to calculate the reconstruction, identification and isolation efficiencies. The corresponding systematic uncertainties are estimated by varying the requirements on the reconstruction, the identification level[10] or the isolation criteria in the related efficiency definition.
The systematic uncertainty on the energy (momentum) scale/resolution is calculated by: shifting the energy (momentum) by a scale factor before selecting the events; and observing the effect of this shift on the number of events in the final state. The procedure is also repeated on the nominal scaled energy (momentum) value with $\pm 1\sigma$. For muons, another systematic uncertainty is evaluated. It is associated with the correction for the efficiency of the muon TTVA cuts ($\left|d_0\right| / \sigma_{d_0} < 3$ and $\left|\Delta_{z_0} \sin\theta\right| < 0.5$ mm in Section 2.3.2). The energy scale/resolution and flavour tagging contribute to the systematic experimental uncertainties associated with the jets.
The jet energy scale (JES) and resolution (JER) uncertainties are evaluated using the same procedure described for the energy (momentum) scale/resolution uncertainties for the leptons.
Every JES uncertainty source is separately reported in Table 2.22. Beyond the possible systematic effects related to the *in situ* analyses, the other JES uncertainties sources are: the η-intercalibration,

[10] The various identification levels are reported for the leptons in Section 2.3.2.

Table 2.22: Summary of the experimental systematic uncertainties considered and constraints (%) on the expected signal strength parameter. In the last column, only the relative constraints > 0.05% are reported.

Source of systematic uncertainty	Relative constraint on expected signal strength parameter [%]
Event	
Integrated Luminosity	0.18%
Pile up μ value rescaling and pileup rejection	0.25% (rescaling)
Leptons	
Trigger efficiency uncertainty	< 0.05
Reconstruction uncertainty	< 0.05
ID efficiency uncertainty	< 0.05
Isolation efficiency uncertainty	
(electrons)	7.10%
(muons)	4.22%
Jets	
Jet Energy Scale (JES) uncertainty from the in situ analyses	< 0.05
energy scale uncertainty on η-intercalibration (modelling, method non-closure and statistics)	< 0.05
energy scale uncertainty on pile-up (μ dependent, N_{PV} dependent, p_{T} term, density ρ)	< 0.05
energy scale uncertainty on flavour composition	0.13
energy scale uncertainty on samples flavour response	< 0.05
energy scale uncertainty from the behaviour of high-p_{T} jets	< 0.05
energy resolution uncertainty, each for both MC and pseudo-data	0.51, 0.24%, 0.11%
	< 0.05
Jet vertex tagger (JVT) efficiency uncertainty	< 0.05
energy scale uncertainty on b-jets	< 0.05
b-tagging efficiency uncertainties	< 0.05
MET	
track-based soft term related longitudinal resolution uncertainty	0.16
track-based soft term related transverse resolution uncertainty	0.08
track-based soft term related longitudinal scale uncertainty	< 0.05
track MET scale uncertainty due to tracks in jets	< 0.05

the pile-up dependence, the jet flavour composition, and the sample flavour response.
Dedicated uncertainties are also evaluated for the corrections applied to high-p_T jets as well as for flavour tagging efficiency and b-jet energy scale.
The systematic uncertainties associated with the measurement of E_T^{miss} depend on the composition of the hard terms and the magnitude of the corresponding soft term (see Section 2.3.4). The contribution of the hard-term composition is extracted from the scale and resolution uncertainties for the individual contributing objects. The $E_T^{miss,soft}$ systematic uncertainties are based on data-to-MC comparisons of: mean value $\langle \mathcal{P}_{||} \rangle$ (where $\mathcal{P}_{||}$ is geometrically defined in Figure 2.29) for the response; and variance $\langle \mathcal{P}_{||}^2 \rangle - \langle \mathcal{P}_{||} \rangle^2$ ($\langle \mathcal{P}_{\perp}^2 \rangle$) for the longitudinal (perpendicular) resolution, using $Z \rightarrow \mu\mu$ events without jets[11].

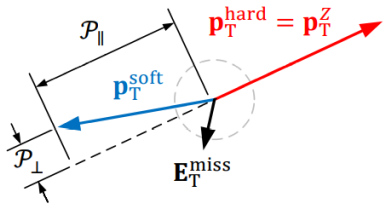


Figure 2.29: Sketch of the $\boldsymbol{p_T^{soft}}$ projections. $\mathcal{P}_{||}$ is the projection into $\boldsymbol{p_T^{hard}}$ and $\mathcal{P}_{\perp}$ is the projection into the perpendicular direction to $\boldsymbol{p_T^{hard}}$.

In general, the systematic experimental uncertainty constraints on the signal strength parameter of the $3l$-"Z-dominated" channel are smaller than contributions of the systematic theoretical uncertainties by comparing the last columns of Tables 2.21 and 2.22.
The integral luminosity and pile-up uncertainties give O(0.1%) contributions to the constraints on the signal strength parameter.
Among the lepton contributions, the isolation uncertainty gives the largest constraint, and it alone is comparable with the magnitude order of the theoretical uncertainties.
When the jet contributions are summed in quadrature, the total contribution is a few per cents giving a comparable contribution as the lepton isolation terms. In contrast, the uncertainties on the E_T^{miss} and b-jets provide negligible contributions.

2.10.3 Systematic uncertainties on misidentified leptons

The scale and fake factor methods (described in Section 2.6) introduce systematic uncertainties on the prediction of the misidentified lepton background. The uncertainties in Tables 2.5 and 2.7 represent the systematic contributions to the misidentified lepton background estimation. The statistical uncertainties σ_{stat} in Tables 2.5 and 2.7 do not provide a significant contribution to the estimation of the systematic uncertainties. The uncertainties $\sigma_{stat+xs}$ in Tables 2.5 and 2.7 give small constraints on the signal strength parameter. They are less than 0.07% in the case of the misidentified muons and

[11] The E_T^{miss} = 0 characterises the $Z \rightarrow \mu\mu$ events. Therefore, $E_T^{miss,soft}$ can be evaluated as opposite hard term.

less than 0.21% in the case of the misidentified electrons. Then, the σ_{stat} and $\sigma_{stat+xs}$ contributions are not reported in Table 2.23.
While the uncertainties on the non-closure test (on the misidentified leptons) in Tables 2.5 and 2.7 give relative constraints on the signal strength parameter of a few per cents in the case of the Scale Factor method applied to the misidentified electrons.

Table 2.23: Most relevant systematic uncertainties due to misidentified leptons and constraints (%) on the expected signal strength parameter.

Source of systematic uncertainty	Relative constraint on expected signal strength parameter [%]
Misidentified leptons	
Scale Factor non closure for misidentified electrons	5.67%
Fake Factor non closure for misidentified muons	0.47%

2.11 Statistical analysis

This section focuses on the method used to evaluate the expected predictions (derived from the Asimov dataset and data in the CRs) and the successive final measurements of the signal strength parameter μ. The definition of the signal strength parameter is given in Section 2.1. This definition can be extended to any process. It can generally be defined as the ratio between the measured events and the Standard Model expected yields.

2.11.1 Profile likelihood procedure

In the $VH \rightarrow V(WW^*)$ analysis, a binned likelihood function is constructed as the product of Poissonian probability terms (P) that predict how the data in the SR and CRs distribute with respect to the expected number of signal and background events. The same likelihood function parametrisation:

$$L(\mu, \mu_b) = P(N|\mu s + \mu_b b_{SR}^{exp}) \times P(M|\mu_b b_{CR}^{exp}) \tag{2.22}$$

is exploited in the Profile likelihood procedure for every single channel of the $VH \rightarrow V(WW^*)$ and for their combination to evaluate the signal strength μ and the normalisation factors for the dominant backgrounds.
In Equation 2.22, the $P(N|\mu s + \mu_b b_{SR}^{exp})$ is the Poissonian probability to find signal and background in the SR, N is the Asimov data (or observed yields) in SR, μ is the signal strength parameter, s and b_{SR}^{exp} are respectively the expected MC signal and background yields in SR.
Then, a zero value for the signal strength parameter ($\mu = 0$) corresponds to an only-background scenario, whereas $\mu = 1$ corresponds to the SM signal expectation.
Similarly, $P(M|\mu_b b_{CR}^{exp})$ is the probability density to find only background in the CR where M are the observed yields in CR and b_{CR}^{exp} are the expected MC background yields in the CR. Finally, the μ_b is the background strength parameter.
The quantities s, b_{SR}^{exp}, and b_{CR}^{exp} are functions of nuisance parameters, θ, that describe the systematic uncertainties.
The parameter μ and the nuisance parameters are estimated by minimising the log-likelihood ratio $\tilde{q_\mu} = -2 \ln \tilde{\lambda}(\mu)$, the adopted test statistics, where $\tilde{\lambda}(\mu)$:

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{L(\mu, \hat{\hat{\vec{\theta}}}(\mu))}{L(\hat{\mu}, \hat{\vec{\theta}})} & \hat{\mu} \geq 0 \,, \\ \frac{L(\mu, \hat{\hat{\vec{\theta}}}(\mu))}{L(0, \hat{\hat{\vec{\theta}}}(0))} & \hat{\mu} < 0 \,. \end{cases} \tag{2.23}$$

In Equation 2.23, the $\hat{\mu}$ and $\hat{\vec{\theta}}$ are the ML estimators of the strength parameter μ and vector of the nuisance parameters $\vec{\theta}$, respectively. The $\hat{\hat{\vec{\theta}}}(\mu)$ and $\hat{\hat{\vec{\theta}}}(0)$ are the conditional ML estimators of the vector $\vec{\theta}$ relative to a given μ and 0. The procedure of choosing specific values of the nuisance parameters for a given value of μ is referred to as "profiling".
The μ measurement corresponds to a hypothesis test with the p-value derived from the sampling distribution $f(\tilde{q}_\mu \,|\mu, \hat{\theta}_\mu)$:

$$p_\mu = \int_{\tilde{q}_{\mu obs}}^{\infty} f(\tilde{q}_\mu \,|\mu, \hat{\theta}_\mu) d\tilde{q}_\mu ; \tag{2.24}$$

and the statistical significance Z is measured from p_μ by replacing $\mu = 0$ in Equation 2.24. It is defined as $Z = \Phi^{-1}(1 - p_0)$, where Φ^{-1} is the quantile of the standard Gaussian. In other words, by Z, the expected or measured p-value can be referred to as the number of Gaussian standard deviations above its mean with an upper-tail probability equal to p.

2.11.2 Results in the 3l-"Z-dominated" and combined analyses

The binning of the ANN output histogram in Figure 2.30 is optimised to have the best significance and the lowest statistical fluctuation using the procedure described in Reference [71] before applying the binned profile likelihood procedure. The fit is performed on the Asimov dataset in SR and data

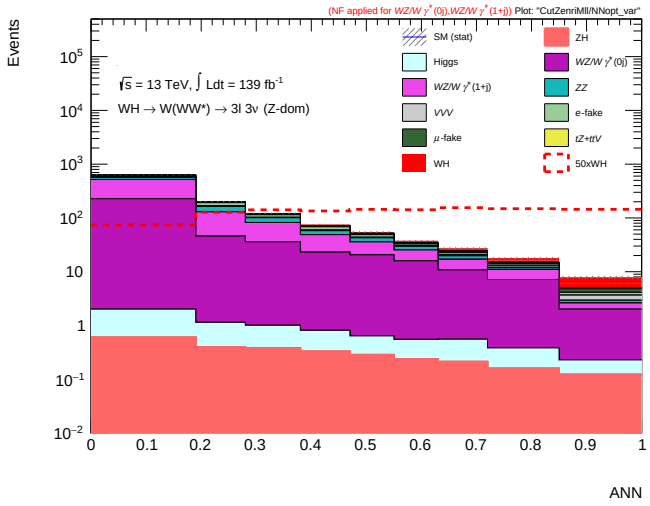


Figure 2.30: The ANN output distribution with optimised binning in the 3l-"Z-dominated".

in CRs including the systematic theoretical and experimental uncertainties. The fit results for the 3l-"Z-dominated" are reported in Table 2.24.

Figure 2.32 shows the post-fit distributions where Asimov data are referred to as data. The

Table 2.24: The expected strength parameter μ_{WH} and significance Z for 3l-"Z-dominated" channel, including experimental and theoretical systematic uncertainties and Asimov data set in the SR.

Parameter	Expected value
μ^{exp}_{Zdom}	$1.00^{+0.78}_{-0.71}$
Significance Z	1.45

uncertainty band in Figure 2.32 represents the post-fit uncertainties in which the statistical contribution is larger than the systematic components, as reported in the breakdown of the uncertainties in Figure 2.31. Among the systematic uncertainties, the largest contribution is the theoretical uncertainties on $WZ/W\gamma^*$ yields without jets (THEOWZ0J in Figure 2.31), followed by the misidentified lepton uncertainties (Fake factor systematic uncertainties in Figure 2.31) and the lepton (and the jet in second

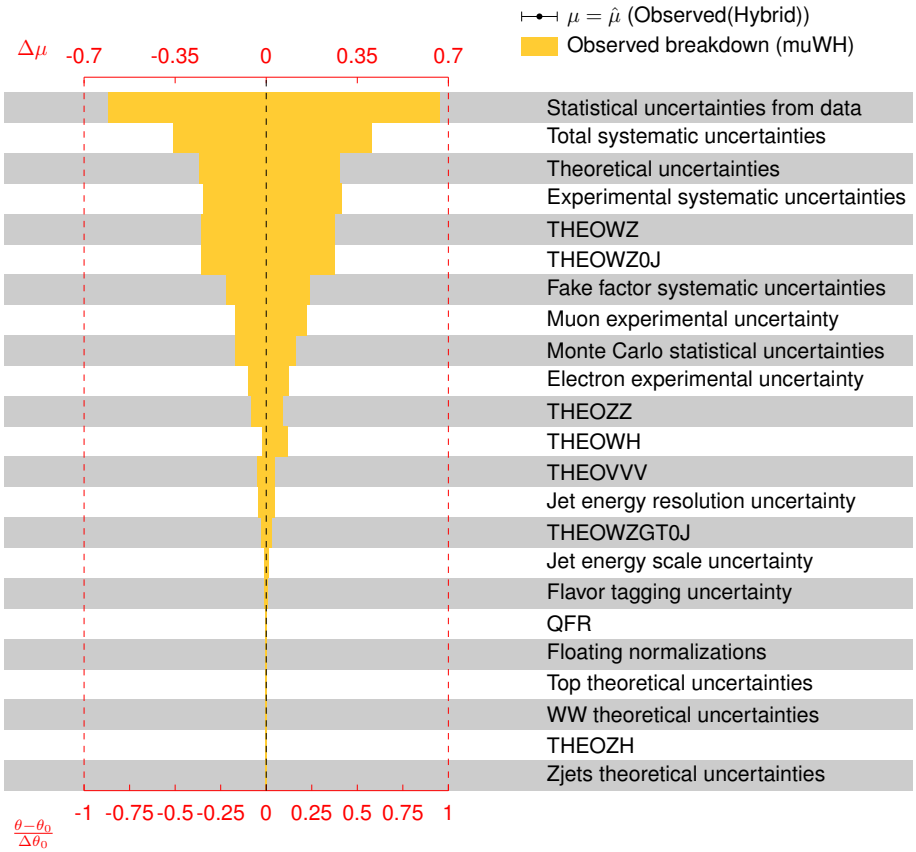


Figure 2.31: breakdown of the uncertainties on the expected signal strength parameter μ^{exp}_{Zdom} for $3l$-"Z-dominated" channel.

place) experimental uncertainties (Muon and Electron experimental uncertainty in Figure 2.31).
Figure 2.33 shows the post-fit ANN distributions in $WZ/W\gamma^*$ CRs that have been used to successfully constrain the NFs (in Section 2.9) of $WZ/W\gamma^*$ background. The results in Table 2.25 are relative to the combinations of the sub-channels of $VH \rightarrow V(WW^*)$ analysis dedicated to one of the two vector bosons (Z or W) in association with the Higgs boson production. The last two rows of Table 2.25 are

Table 2.25: The expected strength parameter μ_{ZH}, μ_{ZH} and μ_{VH} and significance Z for $VH \rightarrow V(WW^*)$ channel, including experimental and theoretical systematic uncertainties and Asimov data set in the SR.

Parameter	Expected value
μ_{ZH}	$1.00^{+0.46}_{-0.39}$
μ_{WH}	$1.00^{+0.31}_{-0.29}$
μ_{VH}	$1.00^{+0.25}_{-0.23}$
Significance Z	5.13

relative to the overall combined fit to all the signal and control regions of the VH analysis, resulting in an expected $\mu_{VH} = 1.00^{+0.25}_{-0.23}$ with the expected significance of 5.13 σ, making $VH \rightarrow V(WW^*)$

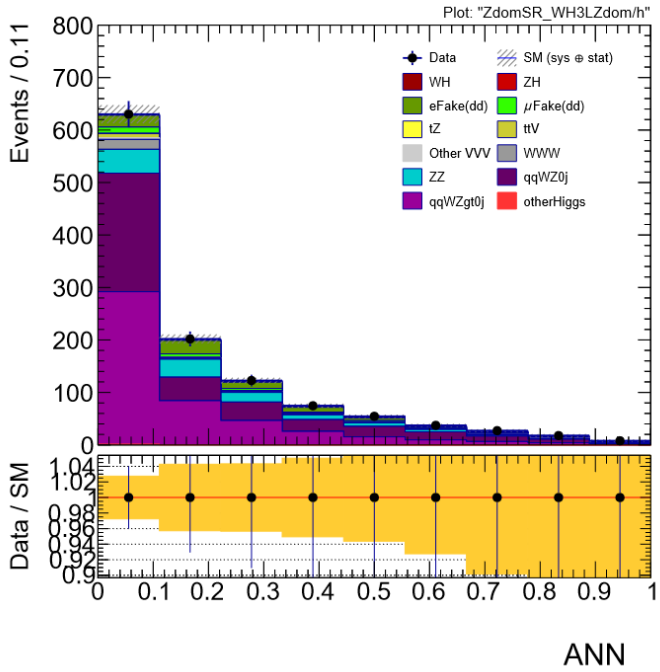
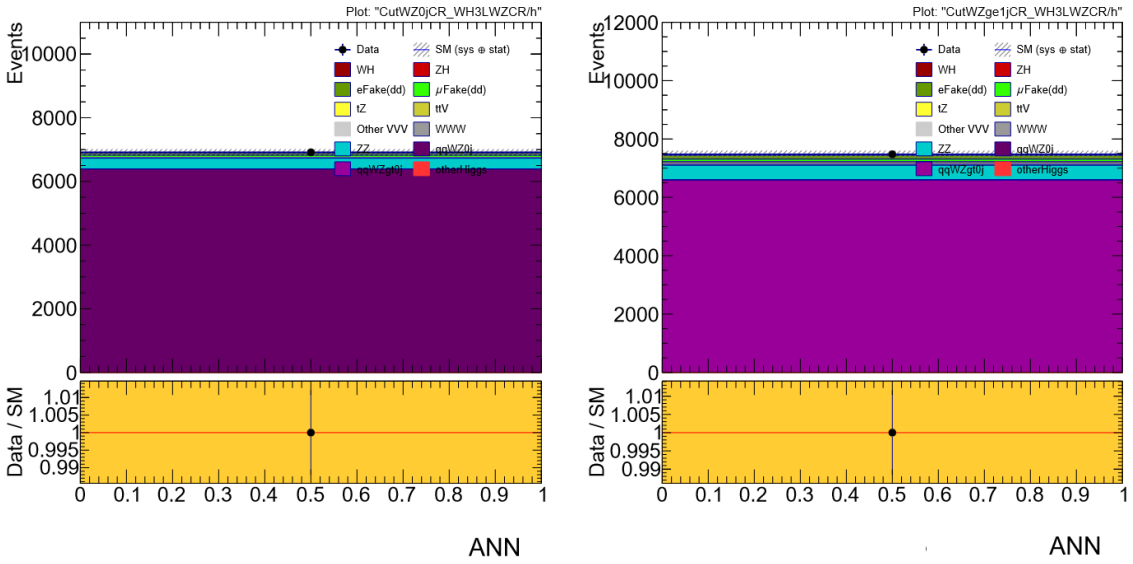


Figure 2.32: Post-fit distribution of ANN output in the 3l-"Z-dominated" SR, with experimental and theoretical systematic uncertainties. In the legend and subplot, data is referred to as Asimov data.

channel potentially sensitive to the observation of VH production events.
The fit considers parameter correlations and non-linearities, leading to asymmetrical error intervals in the tables of this section and Figures 2.34 and 2.35. For the WH process, more narrow intervals are obtained with respect to the ZH process.
The profile likelihood fit is also repeated considering a common signal strength parameter μ_{VH} for the Higgs boson production associated with a generic weak vector boson V, shown in Figure 2.35.



(a) WZ_{w0jet} CR (b) WZ_{wjets} CR

Figure 2.33: Post-fit distribution of ANN output in the WZ_{w0jet} and WZ_{wjets} CRs with experimental and theoretical systematic uncertainties. In the legend and subplot, data is referred to as data in CRs.

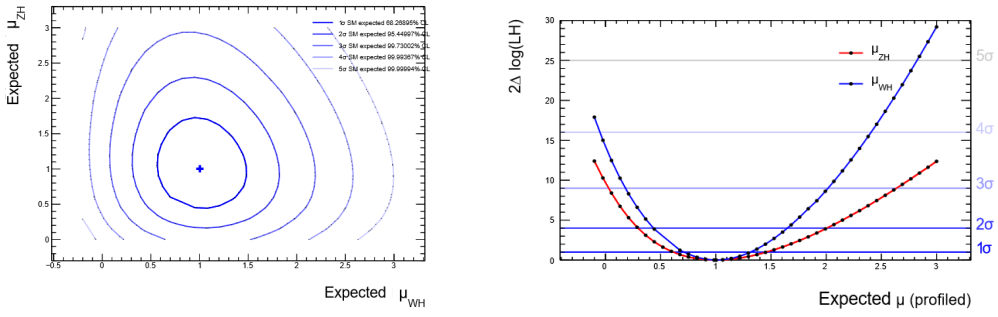


Figure 2.34: The expected significance contours (a) in terms of σ. Profile likelihood, $2\Delta\log(LH)$, curves for μ_{ZH} (combination of $2l$ ZH and $4l$ ZH channels) and μ_{WH} (combination of $2l$ WH and $3l$ WH channels).

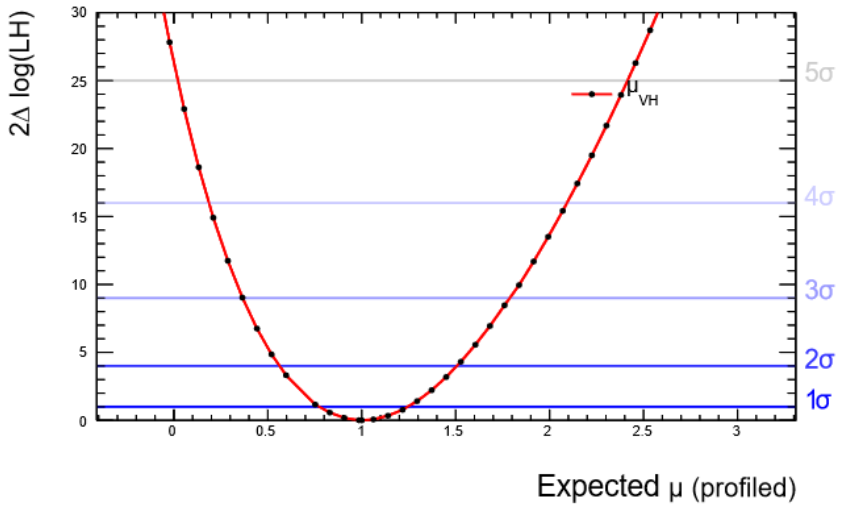


Figure 2.35: The profile likelihood, $2\Delta\log(LH)$, curves for μ_{VH} (combination of $VH \rightarrow V(WW^*)$ subchannels).

2.12 Simplified Template cross-section analysis

The Simplified Template Cross Sections (STXS) is the common framework implemented for the Higgs boson measurements [72] by the LHC experiment collaborations. The STXS measurements are physical cross sections measurements. They aim to complete the results provided by the inclusive LHC analyses using more fine-grained measurements for individual Higgs boson production channels in mutually exclusive kinematic regions (also called *bin*) and a reduction of some theoretical uncertainties. The definition of the *bins* tries to be close as much as possible to the kinematic selection of the individual inclusive analysis to maximise the experimental sensitivity, too.
The *bins* are grouped into configurations called *stages*. The *stages* are organised in progressive order in which the next *stage* has an increased number of *bins* respect the previous. This multi-*stage* organisation easily implements the fine-granularity evolution of the measurements with the increase of the data statistics.
The STXS *stage* 0 is formed by *bins* that correspond most closely to the production mode measurements in LHC Run 1, as sketched in Figure 2.36.
For the analysis of the VH production mechanism, the *stage* 0 selects the interval of Higgs boson

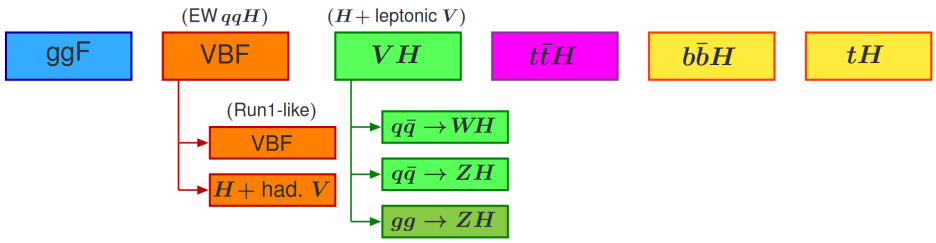


Figure 2.36: The STXS *stage* 0 corresponding to the production mode measurements in LHC Run 1.

rapidity $|Y_H| < 2.5$ that is also split into the $q\bar{q} \to WH$, $q\bar{q} \to ZH$ and $gg \to ZH$ *bins*.
The *stage* 1 refines the *stage* 0. It has been updated by the *stages* 1.1 and 1.2 that are specific for the full Run 2 measurements. An updated uncertainty scheme differs the version 1.2 from version 1.1. The *stage* 1.2 of the VH production mechanism is split into *bins* according to the transverse momentum of the vector boson, ${p_T}^V$, and the number of jets (inside the ${p_T}^V$ *bins*), as shown in Figure 2.37.
Due to the low statistic power of the sub-channels of $VH \to V(WW^*)$ analysis, the high ${p_T}^V$ *bins* are merged in one bin ${p_T}^V \geq 150$ GeV, and the results will be inclusive in the number of jets.

For the $3l$-"Z-dominated" SR, the STXS analysis is performed in two cases where the ${p_T}^W$ range is divided into two and three *bins*. The scheme with two *bins* foresees the $0 < {p_T}^W < 150$ GeV and ${p_T}^W \geq 150$ GeV *bins*. While the scheme with three *bins* foresees the $0 < {p_T}^W < 75$ GeV, $75 \leq {p_T}^W < 150$ GeV and ${p_T}^W \geq 150$ GeV bins. A scheme with four *bins*, in which ${p_T}^W \geq 150$ GeV bin is additionally split, is excluded since the range $p_T^W \geq 250$ GeV is barely populated, as visible in Figure 2.38.
Studies are carried out to define a proxy variable for p_T^W. For $3l$-"Z-dominated" SR, three different candidates are taken into account.
The first candidate is the transverse momentum of the lepton l_2 (${P_T}^{l_2}$), which is the lepton assigned

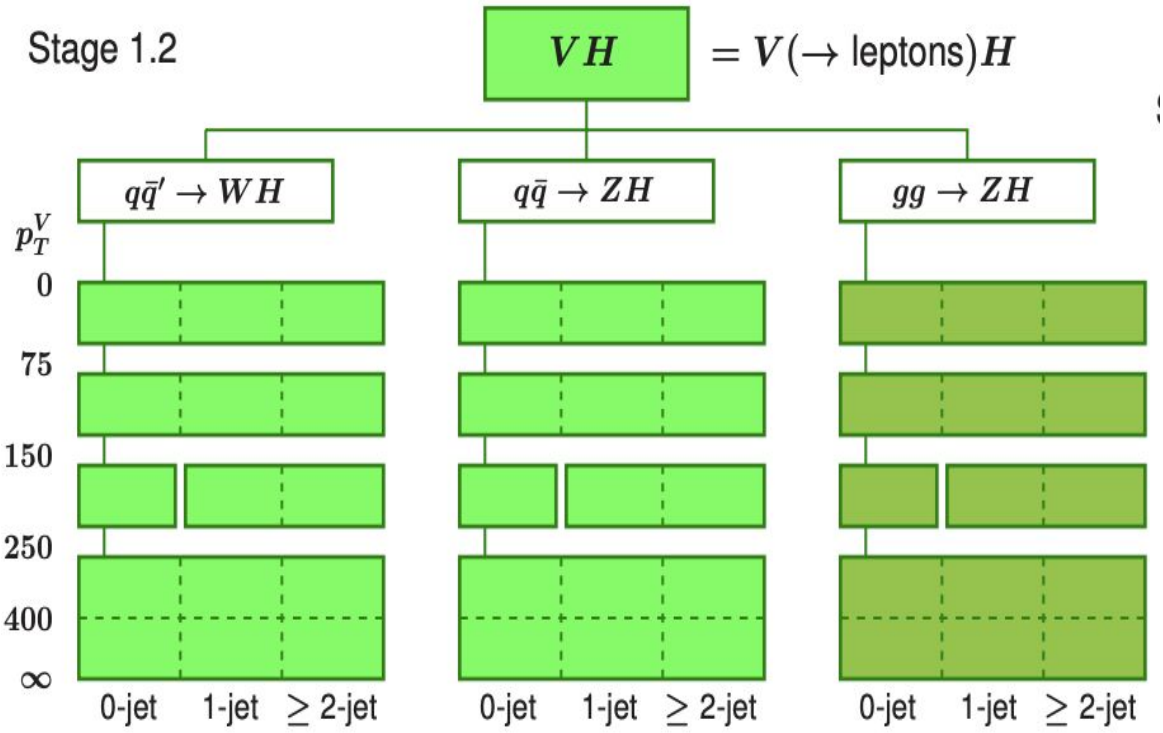


Figure 2.37: The magnified VH related part of the STXS binning diagram at stage 1.1 and 1.2. The white boxes are the recommended *bins*. The dashed lines represent the *bins* in which the MC samples should be split during the data analysis. This splitting is inserted to implement theoretical uncertainties that affect the shape of the observables, like the absolute yield variation in a given *bin* and the migration from a *bin* to the other.

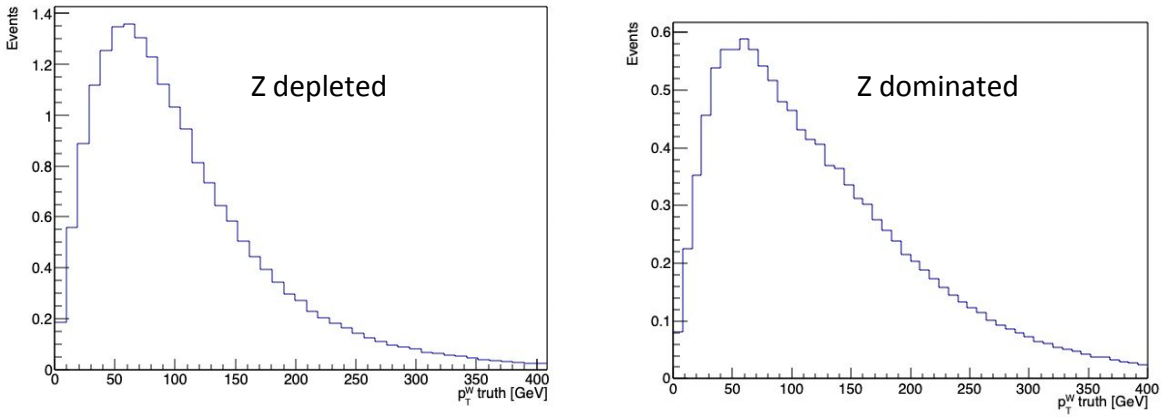


Figure 2.38: p_T^W spectrum of the vector boson in the WH production is shown at truth level for the 3l-"Z-depleted" (left) and 3l-"Z-dominated" (right) SRs.

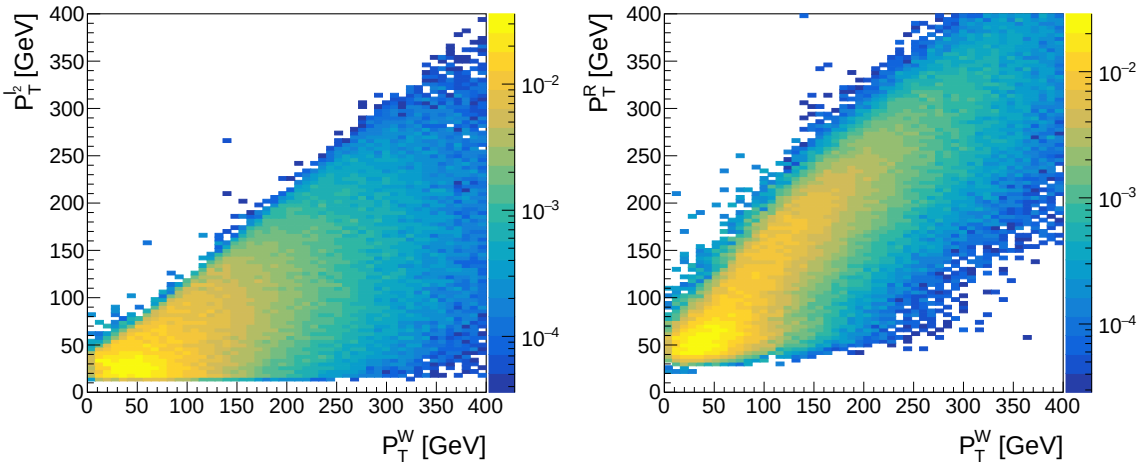


Figure 2.39: Correlation between the $p_T^{l_2}$ (left)/p_T^R (right) and the p_T^W variables in the Z dominated analysis.

to the W boson in the WH process[12]. A second candidate is ${P_T}^R$, the output of a regression neural network (RNN) trained to minimise the difference between ${P_T}^R$ and ${P_T}^W$.
The third candidate is based on a multi-classifier that is described in Section 2.12.1. The multi-classifier categorises the events into four classes: three classes corresponding to the P_T^W *bins* in the *STXS* scheme with 3 *bins* for the WH process; and a class for the dominant background $WZ/W\gamma^*$ process. Looking at Figure 2.39, the correlation at high ${P_T}^W$ is better between ${P_T}^W$ and ${P_T}^R$ than between ${P_T}^W$ and ${P_T}^{l_2}$.
Then, the ${P_T}^{l_2}$ is excluded in the comparison among the results of a statistical analysis in which the proxy candidates are exploited by turns in the profile likelihood fit.

2.12.1 Multi-classifier training

The alternative samples in Section 2.8.2 are used to train the multi-classifier with a different selection with respect to Section 2.8.2. The new selection is reported in Table 2.26.

The strategy of training and selection for the multi-classifier differs with respect to the ANN for $3l$-"Z-dominated" analysis because its aim is to discriminate classes of WH process rather than to classify the WH process from WZ background.
The training sample is further divided into the four classes at the truth level, and they are listed below.

- class/output node 1: WH process with $p_T^W \geq 150$ GeV;
- class/output node 2: WH process with $p_T^W < 75$ GeV;
- class/output node 3: WH process with $75 \leq p_T^W < 150$ GeV;
- class/output node 4: WZ process;

[12] Lepton l_2 is defined in Section 2.5.

Table 2.26: Event selection criteria used to define the alternative samples for the multi-classifier training.

Selection	Criterion
Number of SFOS	2 or 1
Number of b-jets	0
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	> 30
$m_{\ell\ell}$ [GeV]	> 12 (min. SFOS)
$m_{l_0 l_1}$ [GeV]	< 85
$\lvert m_{\ell\ell}^{SFOS} - m_Z \rvert$ [GeV]	> 6 (SFOS)

The statistics of the classes are reported in Table 2.27. For the training of the multi-classifier, a

Table 2.27: Event statistics for training, validation and test of the multi-classifier. The validation and test sample have the same statistics. Their sum is reported in second column.

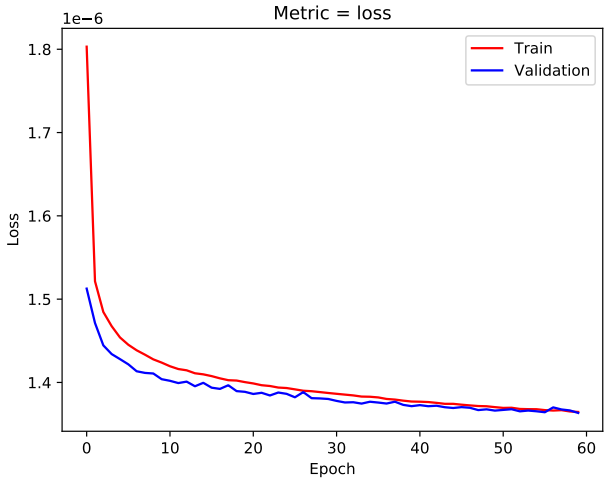
classes	Training	Validation and test
class/output node 1	~111.5k	~27.9k
class/output node 2	~144.4k	~36.1k
class/output node 3	~147.3k	~36.8k
class/output node 4	~1879.4k	~469.84k

customised python environment has been defined with TensorFlow in the backend and Keras.
The input variables include the 15 variables of the ANN for the MVA analysis of the $3l$-"Z-dominated" SR, ${P_T}^{l_2}$ and ${P_T}^{R}$ and 8 discrete variables: the number of jets, the number of $SFOS$ leptons in the events and the lepton flavour combination of each event. Their distributions are reported in Appendix B.
For the multi-classifier, the input data are scaled by means of the sklearn [73] preprocessing method RobustScaler [74]. This scaler subtracts the median and scales the data according to the quantile range, between the 1^{st} quartile (25^{th} quantile) and the 3^{rd} quartile (75^{th} quantile) instead to use the mean and standard deviation. This method reduces the influence of outliers during the scaling.
The architecture is reported in Table 2.28.

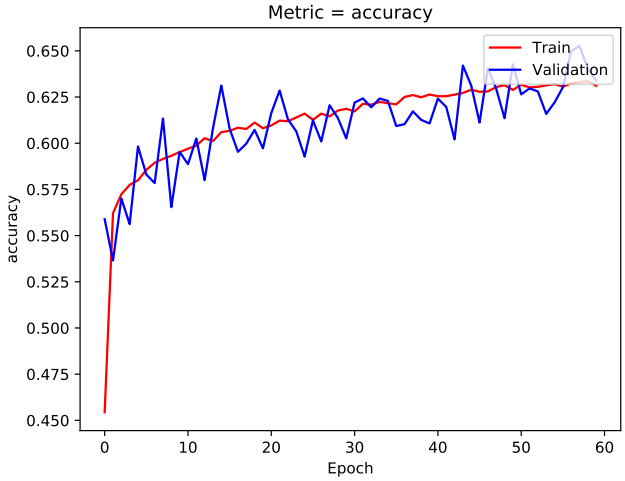
The SDG is replaced by the faster ADAM optimizer to speed up the training, which is repeated several times to optimise the architecture based initially on the ANN model. At the input of each hidden layer, the data pass the Keras API methods [66] of BatchNormalization and GaussianNoise. The BatchNormalization applies a transformation that maintains the mean output close to 0 and the output standard deviation close to 1. The GaussianNoise applies additive zero-centred Gaussian noise to real-valued inputs. It is useful to mitigate overfitting. Moreover, each hidden layer displays an optimised dropout rate found by repeating the training.
The maximum validation accuracy is 0.65, at the 57^{th} and an AUC of 0.89, as visible in Figure 2.40.

Table 2.28: Architecture of the multi-classifier.

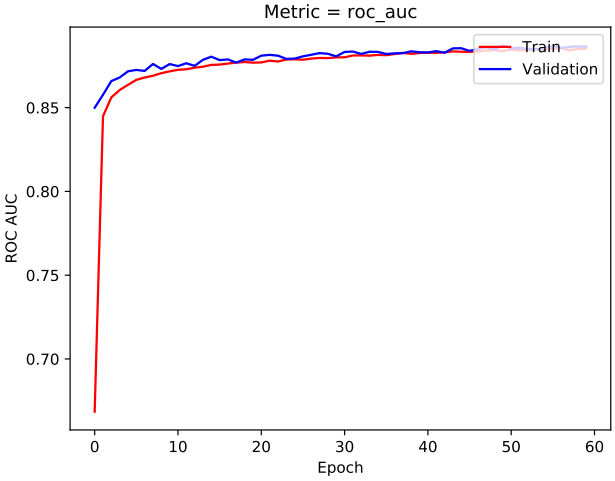
Parameter	Values
number of input variables	25
Kernel initializer	VarianceScaling
number of hidden layers	6
layer activation function	relu
output activation function	softmax
dropout rate (variable)	"drop1": 0.174, "drop2": 0.123, "drop3": 0.164, "drop4": 0.136, "drop5": 0.193, "drop6": 0.177
number of neurons	variable
Layers x Neurons	2x128+2x64+2x32
loss function	categorical crossentropy
optimizer	Adam(lr=0.000869)
batch size	500



(a) Loss function



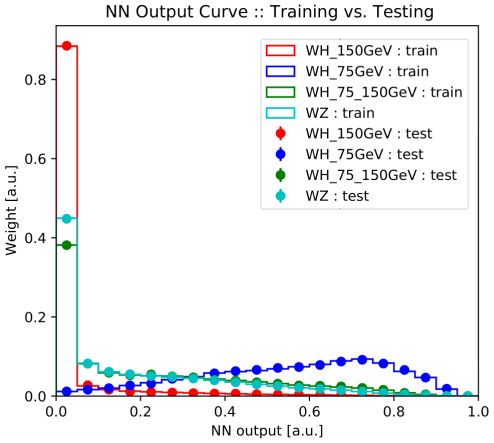
(b) accuracy



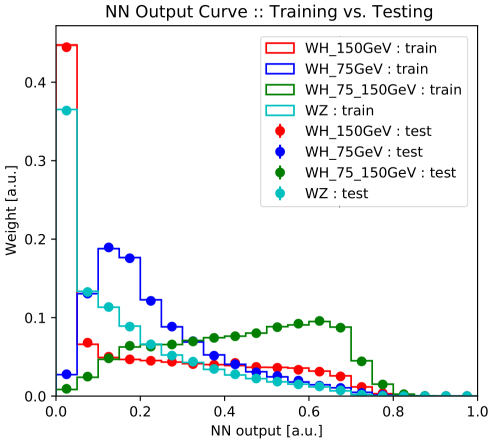
(c) AUC

Figure 2.40: Loss function (a), accuracy (b) AUC of the ROC as a function of the epochs.

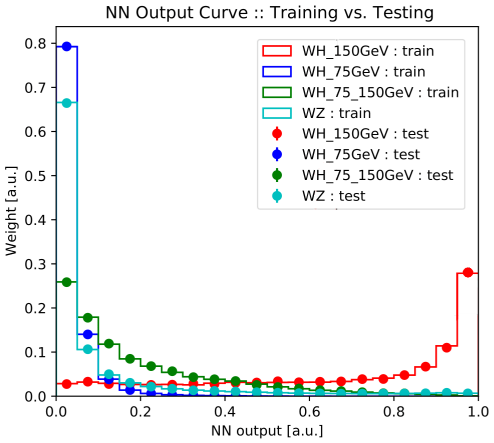
In Figure 2.41, the output distributions are compared between the training and test samples, for each output node. The training and test distributions agree for all the four classes that are singularly well-discriminated versus the rest of the classes.



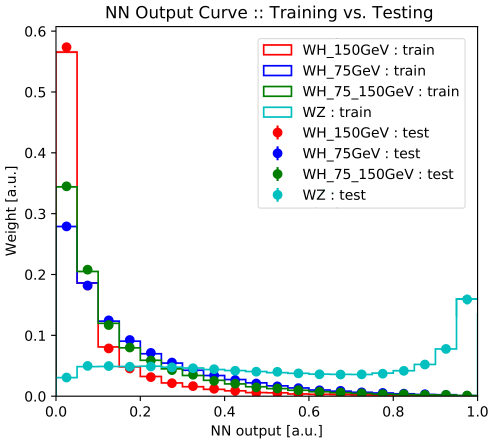
(a) WH process with $p_T^W < 75$ GeV class



(b) WH process with $75 \leq p_T^W < 150$ GeV class vs rest



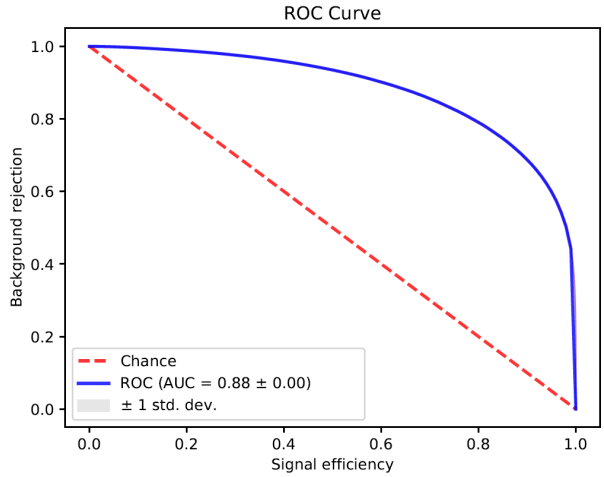
(c) WH process with $p_T^W \geq 150$ GeV vs rest



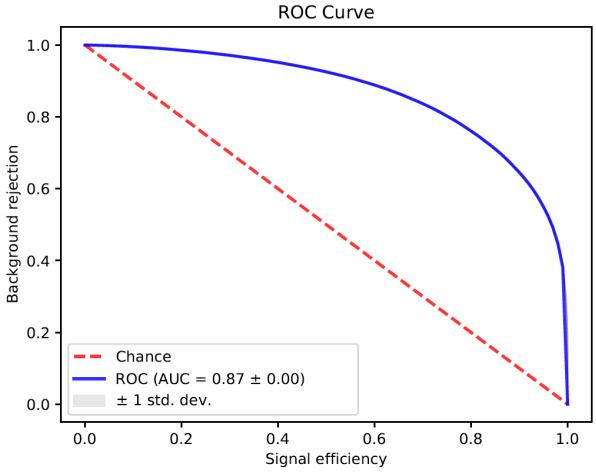
(d) WZ process

Figure 2.41: Training and test output distributions of one output node/class vs rest.

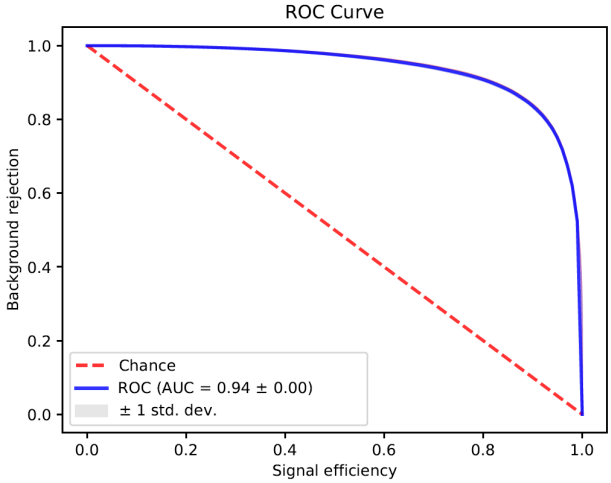
The respective ROC curves of a class vs the rest are shown in Figure 2.42.



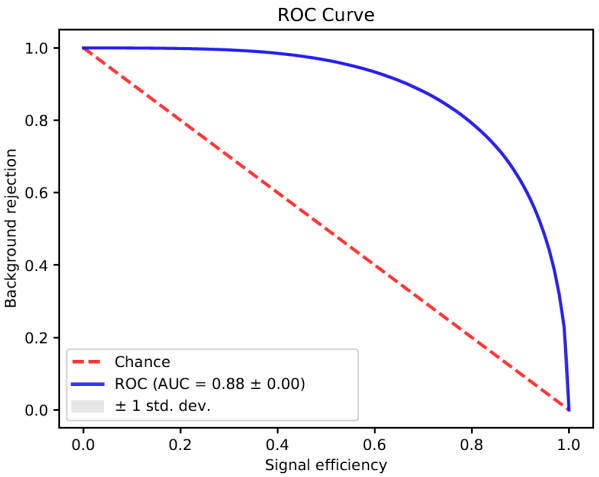
(a) WH process with $p_T^W < 75$ GeV class vs rest



(b) WH process with $75 \leq p_T^W < 150$ GeV class vs rest



(c) WH process with $p_T^W \geq 150$ GeV class vs rest



(d) WZ process vs rest

Figure 2.42: ROC of output nodes/classes in the case of one output node/class vs rest.

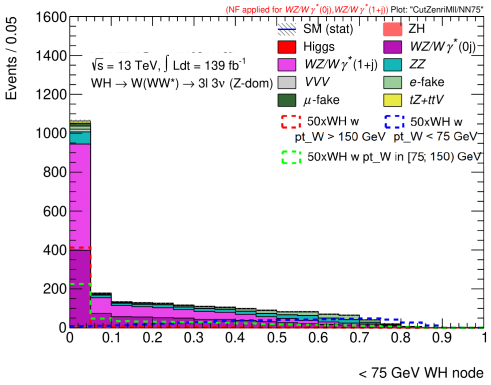
The class-vs-others AUC are reported per each class in Table 2.29. The larger AUC is for the WH process with $p_T^W \geq 150$ GeV class. It means that the efficiency of the multi-classifier to discriminate $p_T^W \geq 150$ GeV class against the other classes is better than in the other cases, reflecting the relative trends of the input variable distribution for the four classes in Appendix B.
It is also interesting to observe that the case of the WZ class versus the rest (i.e. the signal WH process) is comparable with the binary ANN performance in terms of AUC. This observation may suggest that the 15 common input variables have a major role in discriminating the WZ and WH processes with respect to the ${P_T}^{l_2}$, ${P_T}^R$, and the 8 discrete variables.
The multi-classifier is applied to the nominal MC samples that have passed the same selection in Table

Table 2.29: AUC per each class.

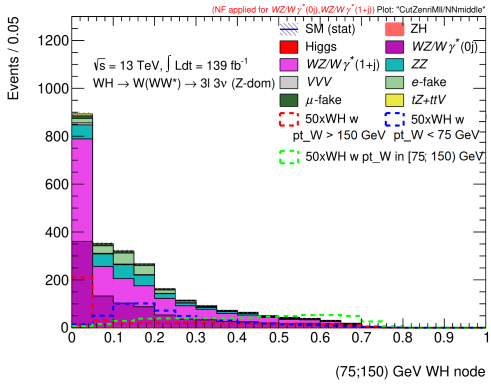
Class	AUC
WH with $p_T^W < 75$ GeV class	0.88
WH with $75 \leq p_T^W < 150$ GeV class	0.87
WH with $p_T^W \geq 150$ GeV class	0.94
WZ class	0.88

2.26.

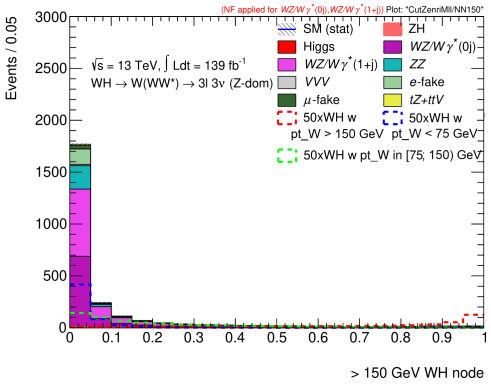
In Figures 2.41 and 2.43, a good agreement is observed between the distributions of the output nodes in the case of alternative (training and test) and nominal (application) MC samples, proving the goodness of the multi-classifier learning.



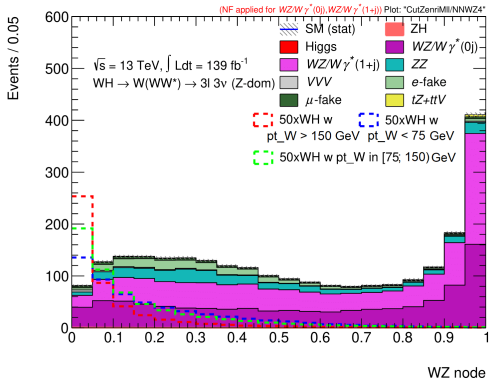
(a) WH process with $p_T^W < 75$ GeV class



(b) WH process with $75 \leq p_T^W < 150$ GeV class



(c) WH process with $p_T^W \geq 150$ GeV class



(d) WZ process

Figure 2.43: Output distributions of the single output node/class.

Finally, the third proxy candidate ξ is defined as a function of the output nodes o_i with $i = \{p_T^W < 75GeV, 75 \leq p_T^W < 150GeV, p_T^W \geq GeV, WZ\}$,:

$$\xi = \ln \frac{o_{p_T^W < 75GeV} + o_{75 \leq p_T^W < 150GeV} + o_{p_T^W \geq 150GeV}}{o_{WZ}}. \tag{2.25}$$

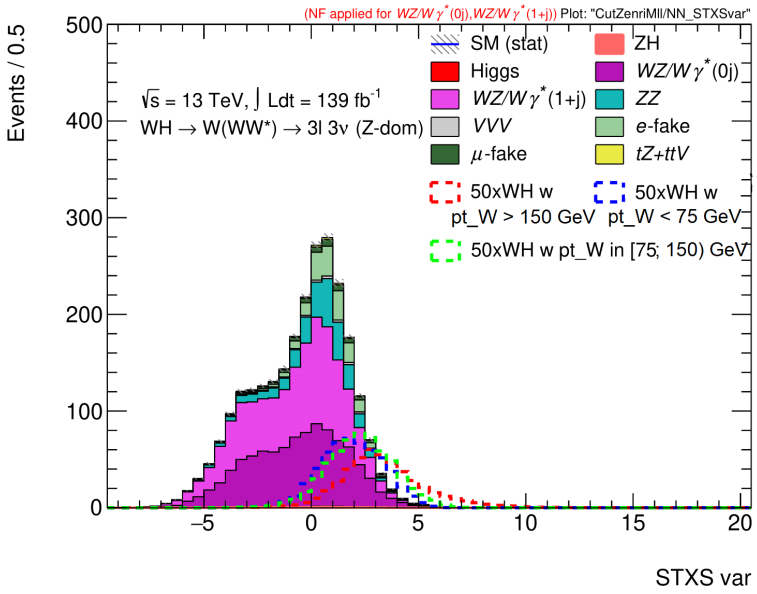


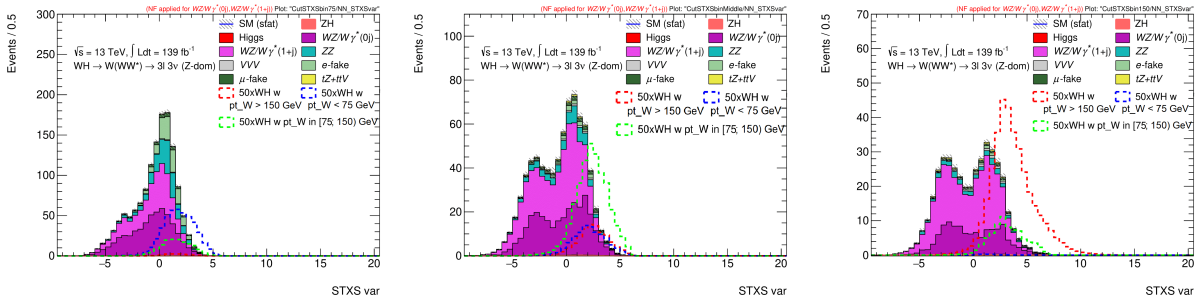
Figure 2.44: ξ distribution after the selection of the $3l$-"Z-dominated"-WH signal region.

The ξ distributions are shown in Figure 2.44 for the p_T^W bins of signal and the different background processes.

Without relying on the truth level P_T^W quantity, the events have been divided into four subsets by comparing the outputs of the multi-classifier output nodes:

- WH process with $p_T^W < 75$ GeV node subset: it includes events that the multi-classifier selects with the highest probability (larger output) to belong in WH process with $p_T^W < 75$ GeV class;
- WH process with $75 < p_T^W < 150$ GeV node subset: it includes events that the multi-classifier selects with the highest probability to belong in the WH process with $75 \leq p_T^W < 150$ GeV class;
- WH process with $p_T^W \geq 150$ GeV node subset: it includes events that the multi-classifier selects with the highest probability to belong in the WH process with $p_T^W \geq 150$ GeV class;
- WZ process node subset: it includes the rest of the events.

The ξ variable is studied in the three p_T^W subset of the WH process. Its distributions in these subsets are shown in Figure 2.45. The backgrounds are distributed towards lower values with respect to the signal classes, although they show a large tail that contaminates the signal. The profile likelihood fit



(a) WH process with $p_T^W < 75$ GeV node subset (b) WH process with $75 < p_T^W < 150$ GeV node subset (c) WH process with $75 < p_T^W < 150$ GeV node subset

Figure 2.45: ξ distribution in the three proxy p_T^W subset of WH process.

with the scheme with two *bins* is performed on ξ. The results are compared with the case in which ξ is replaced with p_T^R in Table 2.30. The two candidates ξ and p_T^R lead similar values for the strength parameters μ in the p_T^W bins, without a significant improvement when ξ replaces p_T^R. Then, it has been considered meaningless to compare the results relative to the scheme with three *bins*, and p_T^R has been confirmed as the p_T^W proxy variable.

Table 2.30: The signal strength parameter from the likelihood fit that exploits ξ (second column) and p_T^R (third column) in the case of the scheme with two *bins*. The systematic uncertainties are not included.

Parameter	ξ	p_T^R
$\mu_{p_T^W<150}$	1.00 ± 0.91	1.00 ± 0.95
$\mu_{p_T^W\geq150}$	1.00 ± 1.05	1.00 ± 1.10

3 Resistive Micromegas detectors for ATLAS Muon Spectrometer

This chapter concerns the resistive Micromegas detectors in the upgraded Muon Spectrometer subsystem (New Small Wheels) of the innermost ATLAS End-Cap regions. They will operate from LHC Run 3. After an introductory description of the detectors and measurement procedures, the results of a study on the principal parameters of the readout Micromegas boards are presented in Section 3.5.1. An overview of the validation tests on the high-voltage detector stability and tracking efficiency is reported in Section 3.6.2 together with their related studies. Finally, Section 3.11 is dedicated to the study on the impact of the electronic channel threshold on the tracking efficiency.

3.1 Muon spectrometer upgrades for Run 3 and HL-LHC

Significant upgrades have been planned for the ATLAS Muon Spectrometer, split in Phase-I and Phase-II during both the LS2 and the LS3 (in Figure 1.1) in order to cope with the operational conditions at the HL-LHC in Run 4 and beyond.
During the LS2, the major Phase-I upgrades of MS have involved the replacement of the Small Wheels in the innermost End-Cap regions of the MS, pointed out by the blue square box in Figure 3.1, new muon detectors in the Barrel Inner Small (BIS) layer and the upgrade of the TDAQ system to include the new detector technologies and the new requirements related with the increase of the LHC luminosity. The newly installed version of Small Wheels, New Small Wheel (NSW), is equipped with Micromegas and small-strip TGC detectors instead of the CSC, MDT, and TGC chambers (described in Section 1.2.5). To further increase the trigger acceptance and robustness, the Phase-I upgrade also includes installing a new detecting system (BIS78 in Figure 3.2) in the Barrel Inner Small (BIS) regions formed by two-multilayer small Muon Drift-Tube chambers (sMDTs) and triplets of thin gap Resistive Plate Chambers (tgRPCs) [75]. It has replaced the old MDT chambers (with a 30 mm diameter tube) of two BIS stations (see the top of Figure 3.2), covering the $1.0 < |\eta| < 1.3$ interval. The tubes of the new sMDTs have 15 mm diameters, while the tgRPCs have 1 mm gas gaps instead of the 2 mm gaps of the previously installed ATLAS RPCs. These RPC chambers are equipped with new front-end electronics, too. Then, this new detecting system represents a preview of a part of the Long shutdown 3 (Phase-II) upgrades when all the current MDT chambers will be replaced in the BIS regions.
Many parts of the TDAQ muon system have been upgraded during the Phase-I upgrade. A new Sector Logic board have been made for the End-Cap to receive inputs from the New Small Wheel detectors, the new BIS RPCs, and the outer layer of the Tile calorimeter. All the inputs will be used to reduce the spurious event rate in the End-Cap. Also, the L1Topo has been upgraded with a new board that

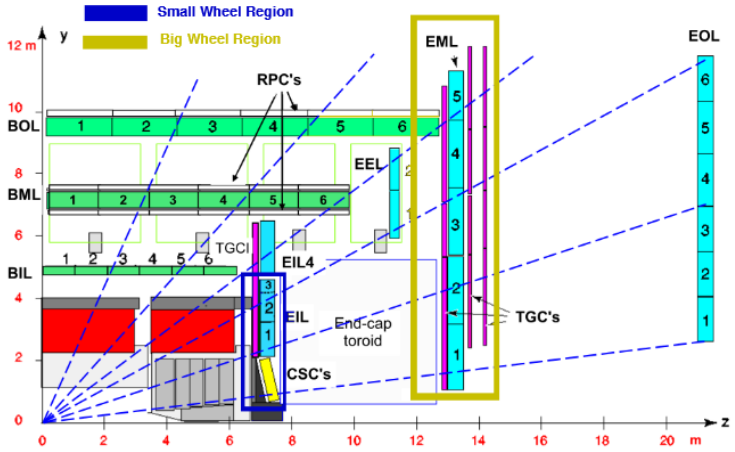
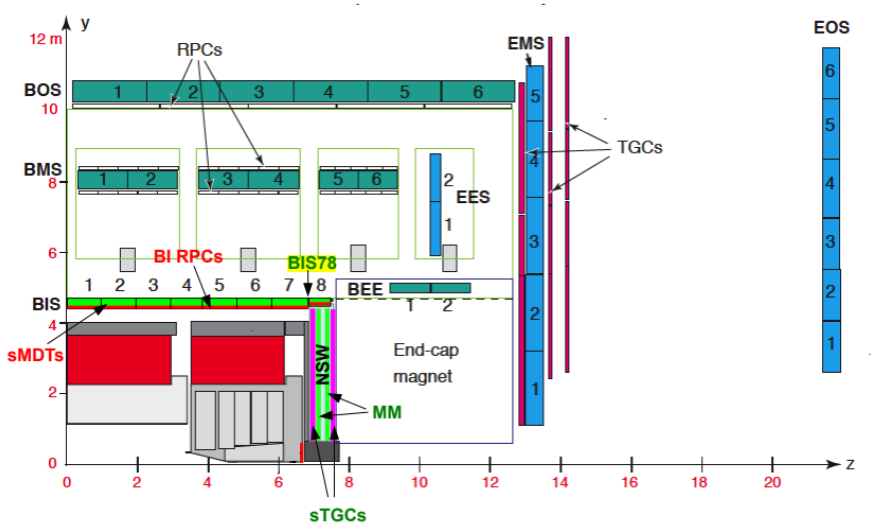
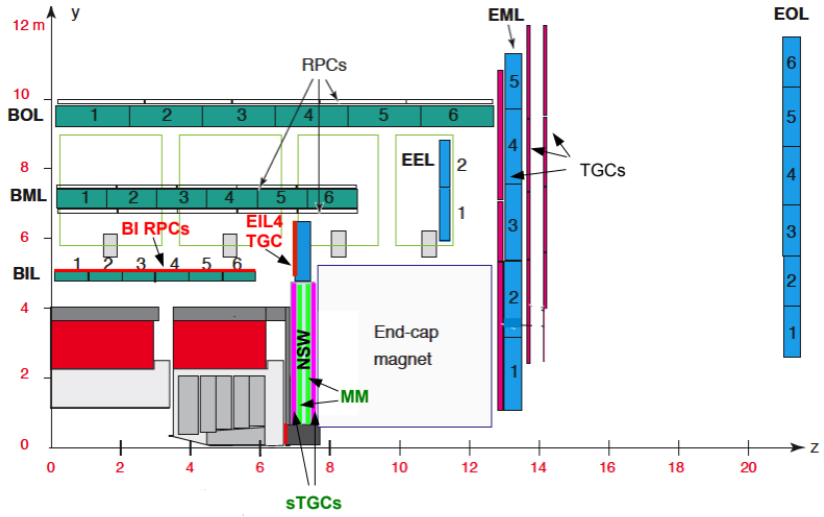


Figure 3.1: A quarter *Z-Y* of ATLAS apparatus with the old MS detector composition. In evidence, the three MS stations in the forward region: End-Cap Inner Large (EIL), End-Cap Middle Large (EML) and End-Cap Outer Large (EOL). The blue (yellow) square individuates the Small (Big) Wheel region. In particular, the Small Wheel subsystems have been updated with new detectors.

performs specific topological algorithms on calorimeter and muon data. A new Muon to Central Trigger Processor Interface (MuCTPI) board will manage the data transfers using the new boards (End-Cap Sector Logic and L1Topo). It is based on the last generation FPGAs. The new readout system is based on the FELIX (Front End LInk eXchange) system [76].
The Phase-II upgrade [75] will conclude the process of adapting the muon spectrometer to the ever increasing performance of the LHC by replacing a large fraction of the front-end, on- and off-detector readout and trigger electronics to manage the higher trigger rates and longer latencies necessary for the new Level 0 trigger. The TDAQ Phase-II baseline architecture will be based on a single-level hardware trigger with a maximum rate of 1 MHz instead of the previous 100 kHz rate during the LHC Run 1 and 2, and 10 μs latency instead of 2.5 μs during the LHC Run 1 and Run 2. If the hadronic trigger rates and the inner pixel detector layer occupancy are higher than expected, the baseline architecture will be evolved into a dual-level hardware trigger architecture.
The rejection of spurious triggers in the Barrel and End-Cap regions and the trigger geometrical coverage of the Barrel region will be improved by installing new detectors. As anticipated before, in the Small Sectors of the Barrel Inner region, all the old MDT detectors will be replaced by the integrated systems of sMDTs and tgRPCs, such as those installed during the Phase-I upgrade. Also, new tgRPC detectors will be mounted on top of existing BIL MDT in the Large sectors of the Barrel Inner region (see the bottom of Figure 3.2). In the Barrel-End-Cap transition region, the Thin Gap Chambers (TGC) EIL4 will be replaced by new TGC EIL4 detectors based on a triplet instead of a doublet.



(a) Small Sector



(b) Large Sector

Figure 3.2: Two r-Z views of the Phase-II ATLAS Muon Spectrometer layout of a Small Sector (top) and a Large Sector (bottom). The drawings show the new detectors to be added in the Phase-II upgrade in red text (BI RPC, sMDT, EIL4 TGC), those to be installed during LS2/Phase-I upgrade in green text (Micromegas and sTGC in the New Small Wheel and BIS78 RPC and sMDT), and those that remain unchanged from the Run 1 layout in black text [75].

3.2 New Small Wheel upgrade for Run 3 and HL-LHC

During Run 1 and Run 2, the previous Small Wheels have operated up to the maximum reached $2.1\text{x}10^{34}$ cm^{-2} s^{-1} instantaneous luminosity. However, their high Level-1 muon trigger rate and transverse momentum resolution in Level-1 muon trigger are limited and inadequate at the higher luminosities expected for Run 3, leading to degraded muon tracking performance at high rates. Moreover, the Small Wheel detectors do not have participated in the Level-1 muon trigger.
Then, the New Small Wheel design provides efficient performance in muon tracking and (also) triggering, compatible with the expected instantaneous and integrated luminosities of Run 3 up to those after the Phase-II upgrade: $5\text{x}10^{34}$ cm^{-2} s^{-1} and 3000 fb^{-1}, respectively. The upgraded NSW have a tracking acceptance of $1.3 < |\eta| < 2.7$ and a trigger acceptance of $1.3 < |\eta| < 2.5$. The New Small Wheels will provide either the muon track reconstruction with high precision and information for the Level-1 trigger under an expected higher background radiation up to 20 kHz/cm^2. The required spatial resolution is about 100 μm for the precision track reconstruction in the offline analysis, while the Level-1 trigger track segments have to be reconstructed online with an angular resolution of approximately 1 mrad.
The Micromegas [77] (MM) and small-strip Thin Gap Chambers [78] (sTGCs) detectors satisfy these requirements. MMs are primarily dedicated to precision tracking, while sTGCs are devoted to the Level-1 trigger function. At the same time, the MM detectors and sTGCs can exchange the tasks, providing a fully redundant detector system for triggering and tracking and in online and offline operations. Considering all the MM tracking layers, their total active area is about 1200 m^2 in the two New Small Wheels[1], segmented into 16 Sectors (8 Large and 8 Small) per wheel to match the MS layout that is divided into Layers and Sectors (see Section 1.2.5). The Large and Small sectors are positioned with an overlap on two different planes, as visible on the left side of Figure 3.3. A sector has a wedge shape, formed by 2 inner MM wedges that are closed by 2 sTGC wedges, as shown on the right side of Figure 3.3. Each wedge has 4 detecting layers.
A MM layer is physically divided into two pieces due to construction reasons. Thus, a MM wedge is formed by two quadruplets with a trapezoidal shape, also called Modules. The nomenclature foresees that the Module that is positioned to larger (smaller) pseudorapidity is called type 1 (type 2). Figure 3.4 (a) and Figure 3.4 (b) summarise the possible Module combinations that can form a MM sector according to the sector size. Moreover, an integrated pair of the inner MM wedges is called Double Wedge in the following sections.

[1] The detecting area of the two NSWs is about twice as large when the STG layers are also added to the MM ones.

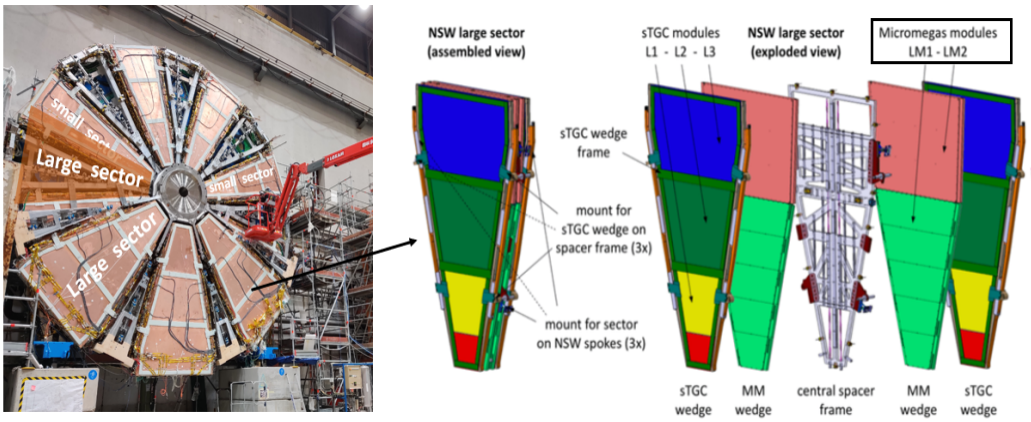
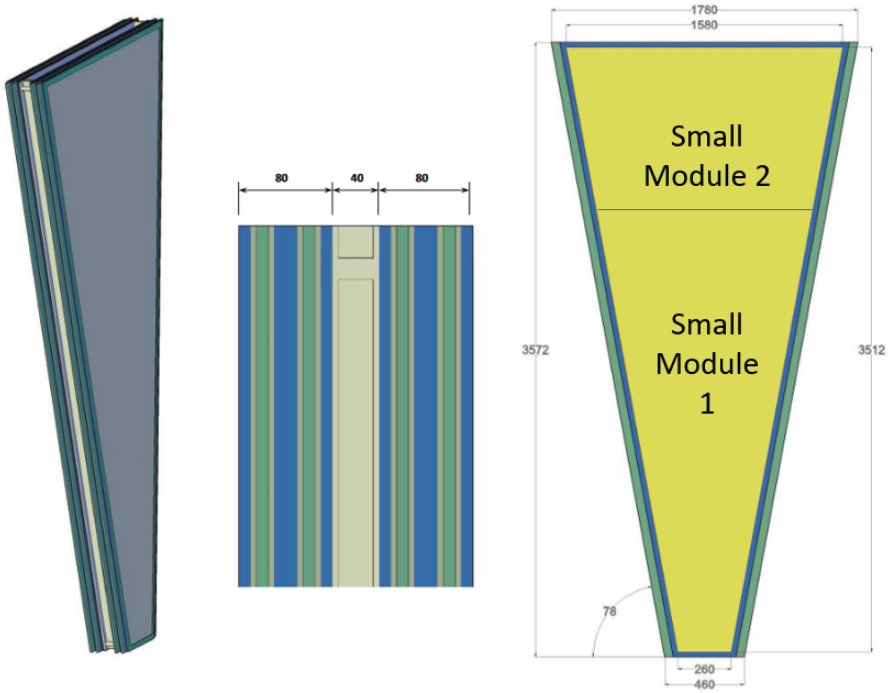
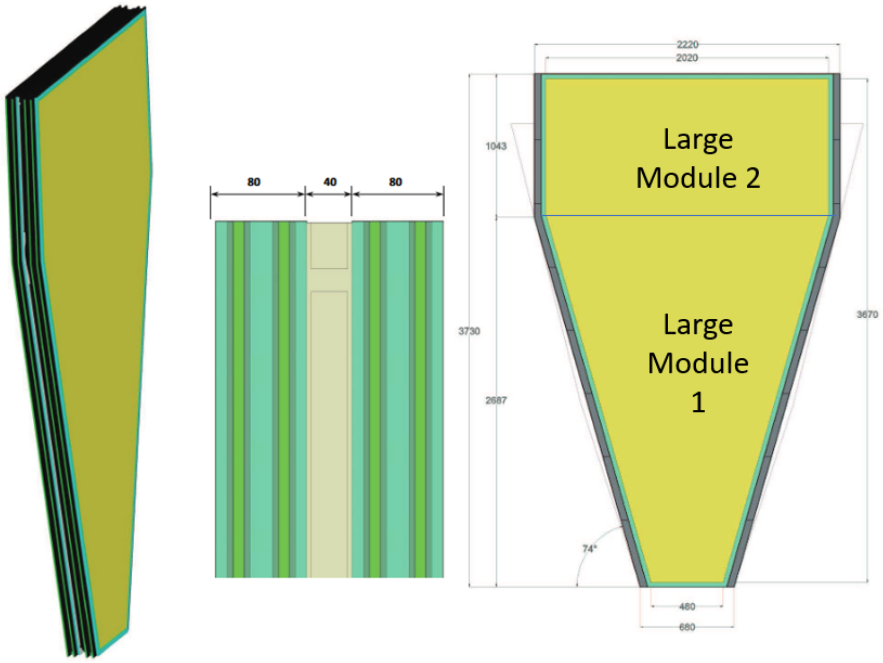


Figure 3.3: Picture of New Small Wheel A when a Large Sector have been mounted (left). A zoom of the structure of a NSW sector with details on the nomenclature of the detectors (right).



(a) Small Wedge



(b) Large Wedge

Figure 3.4: Layout of the MM detectors with reported sizes in millimetres and degrees, including space allocated for services. The active detector area (yellow color), the frame (blue color) and the allowed envelopes (full lines) are indicated. The dimension units are mm (and degree) for lengths (and angles). In evidence, the wedge separation into Module type 1 and type 2.

3.3 New Small Wheel detector technologies

Most PhD activities have involved the construction, characterisation and validation of MM detectors in terms of high voltage (HV) performances[2]. The entire validation chain of HV performances starts with the tests on the Modules in their production sites. Next, all Modules are tested again at CERN after being assembled as Double Wedge with the final service (at CERN BB5 building). This step is described and discussed in detail in Section 3.6. The HV is monitored a third time after being integrated with sTGCs on the mechanical structure (at CERN 191 building).
Detailed descriptions of ATLAS resistive strip Micromegas and sTGC designs are reported in the Technical Design Report [79], together with the entire New Small Wheel project. Therefore, a brief description of the resistive Micromegas technologies of the New Small Wheels will be given in this section before focusing on the studies carried out on the ATLAS resistive strip Micromegas detectors.

3.3.1 Micromegas detectors

The Micromegas (MICRO-MEsh-GAseous-Structure) detectors belong to the Micro Pattern Gaseous Detectors (MPGD) family. The overall MPGDs outperform traditional wire chambers in terms of granularity and rate capability at a reasonable construction effort due to the microelectronics technology application in their manufacturing. They are exploited in several current high energy physics experiments (and technical applications). Many more applications are foreseen because they have managed to fill the gap between the high-performance but expensive solid-state detectors and cheap but rate-limited traditional wire chambers. The Micromegas is a gaseous detector that works in proportional mode. The incident radiation interacts with the gas, and electron and ion pairs are produced. When the electric field is applied, the charged particles move to the corresponding electrodes, and the electrons undergo the phenomenon of Townsend avalanche multiplication. The main Micromegas detector feature is the presence of a thin micro-mesh (Figure 3.5) inside the active volume, which separates the region where the ion-electron pairs are produced around the particle trajectory from the zone where the electron multiplication takes place. Therefore, the region between cathode and micro-mesh behaves as a conversion and charge drift region, called drift gap, and it is a few millimetres wide. In the region between the mesh and the anode, the charge is amplified. Then, this zone is called amplification gap, and it is O(100 μm) wide [77]. To better ensure the uniformity of the electric field intensities, several dielectric pillars are placed in the active area by means of photolithographic techniques that keep the mesh at the same distance from the anode plane in the entire active surface.
The actual (most applied) voltage supply scheme uses a negative voltage on the cathode, mesh at zero potential and a positive voltage on the anode. Then, the voltage applied to the cathode coincides with the drift voltage, as well as the anode voltage with the amplification voltage. Their values are some hundreds of volts. Due to the detector geometry, the electric field in each gap can be approximated to the one present in a planar capacitor. Therefore, the electric fields inside the detector can be very intense even using moderate voltages, considerably reducing the problems of HV management.

[2] For general knowledge, the complete validation of a MM detector also includes passing gas leakage, electronic and alignment tests, but they are not part of this thesis.

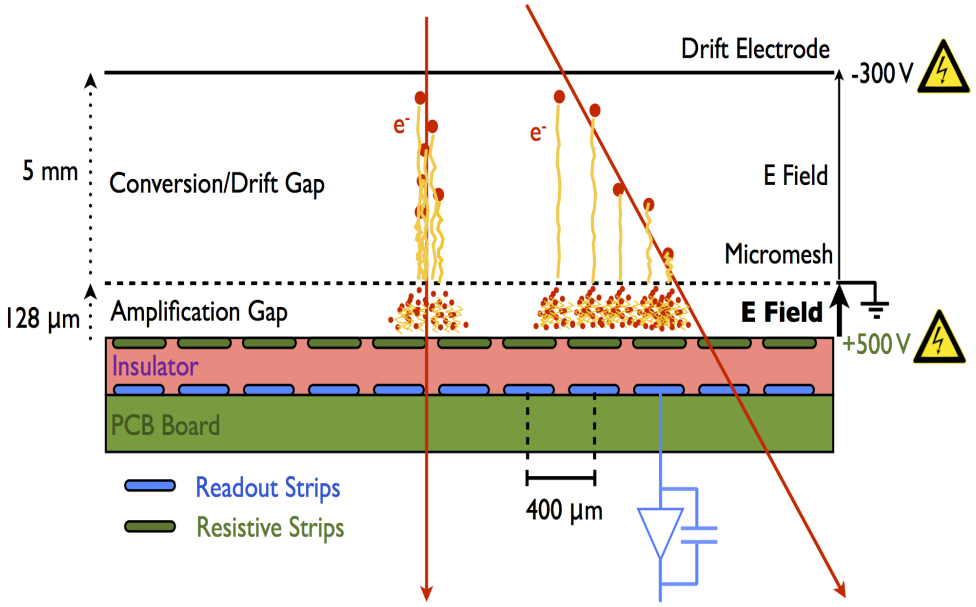


Figure 3.5: Structure of resistive strips Micromegas detector.

The electric field configuration near the micro-mesh is responsible for the efficiency of the electron passage between the two gaps (transparency). In addition to the micro-mesh characteristics (pitch and wire diameter), the transparency is deeply dependent on the ratio between the electric field in the amplification gap and the one in the drift gap. The maximum transparency value is usually observed when this ratio is around 70-100. The micro-mesh presence allows minimising the collection time of positive ions (produced in the amplification gap) because they have to cover only distances about hundreds of micrometres instead of several millimetres, reducing the signal duration (or detector occupancy time) and allowing a high detection rate. The readout electrodes are metallic strips with a few tens micrometers in thickness and hundreds of micrometres in width. They are placed in parallel with a few hundreds of micrometers pitch on a Printed Circuit Board (PCB).

3.4 Resistive strip Micromegas technology

The peculiarities of ATLAS Micromegas detectors are the spark protection resistive layer, large dimensions and industrial production. The readout strips are covered by a thin (50 μm) insulating Kapton layer, on which resistive (about 0.3 MΩ/□) strips are placed overlapping the readout strips by lithographic procedures [80]. In the voltage supply scheme of a resistive strip MM detector, the anode voltage is applied to the resistive strips while the readout strips are floating. Then, the resistive strips quench the possible discharges that may occur in the amplification gap when the avalanche electron number overcomes the Raether limit value (10^7, which defines the spark mode) [80]. They protect the readout strips and electronics. Using the strips rather than a single resistive layer limits the spatial spread of the discharge on the resistive net and the voltage drop on only a few consecutive resistive

strips, improving the spatial resolution and avoiding long dead time. To assure a uniform resistance and a homogeneous HV distribution over the entire plane of ATLAS resistive MM, interconnections between the neighbouring strips are present every 10 mm, forming a ladder scheme as shown in Figure 3.6. The ladder scheme also minimises the effects of possible printing defects on a single resistive strip, like interruption on a resistive strip, with respect to a scheme with fully independent resistive strips because the current can flow through the net. Both the technique of the screen-printing and

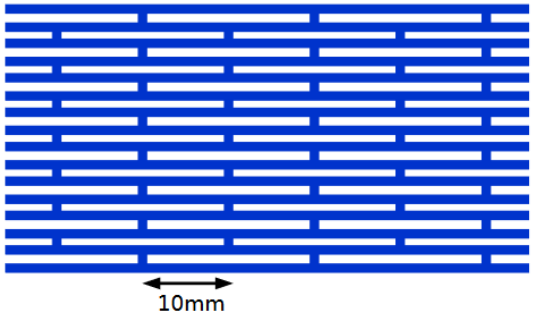


Figure 3.6: Scheme of interconnections in resistive spark protection layer.

sputtering (already used in the MPGD community [81]) have been tested to realise the resistive strip network over the Kapton foil (see Figure 3.7). They have been both valid, although the screen-printing has been considered more suitable for large mass production in terms of cost [82].

Additional studies have proved that the screen-printing resistive layer samples are more robust to the

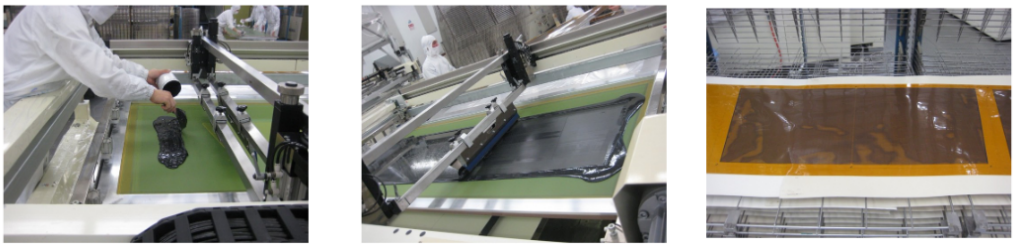


Figure 3.7: Pictures of the screen-printing of a resistive strip Kapton foil. A positive-patterned screen is placed on the Kapton foil and the resistive paste is printed with a roller.

mechanical stress than the sputtering samples (see Reference [83]).

The resistive paste is a composite formed by graphite particulates in an epoxy matrix. Slight different compositions have been used in the resistive foil production, and different tunings have been tried to achieve the most suitable final resistive value. Together with the permeability of the screen, the printing pressure and velocity, the composition and viscosity of the paste contribute to the strip thickness, which defines its bulk resistance. The liquid state of the paste causes the transverse section of the strips to be not a rectangle with sharp edges, as shown in Figure 3.8. During these tests, the limits on the specifications have taken into account that the resistivity changes after the glueing of resistive foil on the bare PCB with the copper readout strips because the glueing is a procedure done at controlled pressure and temperature, and it has given a margin to optimise the final resistivity in a second step.

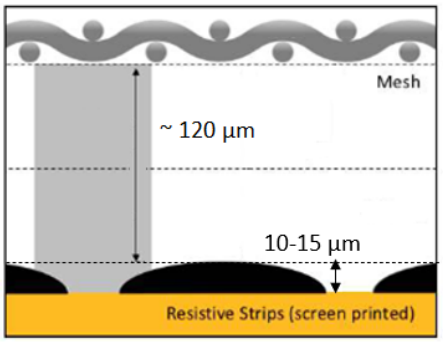


Figure 3.8: Transverse section of screen-printed strip (not in scale).

All the produced resistive foils have undergone a series of quality control tests that evaluate the resistance in several points of the net, providing a 2D map of surface resistivity and the number and severity of the possible defects. The foils have been classified into four quality classes according to the resistivity measurements. The average surface resistivity has to be within the range of [0.28, 2.0] MΩ/□ for all the resistive foils. Foils with 99% of the measurements within the range [0.28, 2.6] MΩ/□ were marked as Grade A foils; foils with 95% of the measurements within the range [0.28, 2.6] MΩ/□ Grade B; foils with 95% of the measurements within the range [0.21, 3.4] MΩ/□ Grade B-; the rest are Grade C.
For the ATLAS Micromegas, only resistive foils of grade A have been selected to be glued on the Printed Circuit Boards (PCB).
Beyond the NSW project, an R&D project (in Appendix E) has been started to push further the resistive Micromegas technology for operations under very high particle flow up to rates of tens MHz/cm^2, three orders of magnitude higher than current applications. The primary challenges of this R&D project have been the miniaturisation of the readout elements, the optimisation of the spark protection system and the stability and robustness under operation.

3.4.1 ATLAS resistive strip Micromegas design

The ATLAS MM design will be described starting from the readout MM board, core of a single MM tracking layer, passing through the Module configuration up to the final MM Double Wedge arrangement, introduced in Section 3.2. The bare PCBs of ATLAS Micromegas detectors consist of an FR4 base, with a trapezoidal shape and different sizes, on which a readout pattern of 1024 copper strips with a 300 μm width and a 0.425 (0.450) mm pitch[3] is realised by a photolithographic procedure.
After that, they undergo a specialised multi-step procedure for the ATLAS Micromegas detectors, done by the production company itself. Firstly, the resistive foil, described in Section 3.4, is aligned and glued on the bare PCB by means of a high pressure technique using a 25 μm glue layer (blue line in Figure 3.9).

[3] The strip pitch varies according to the size of the sector in which the board is destined. The boards for Small (Large) Sectors have 0.425 (0.450) mm pitch.

Next, the connections between the HV input line and the resistive pattern (made by $< 10\,\Omega$ silver

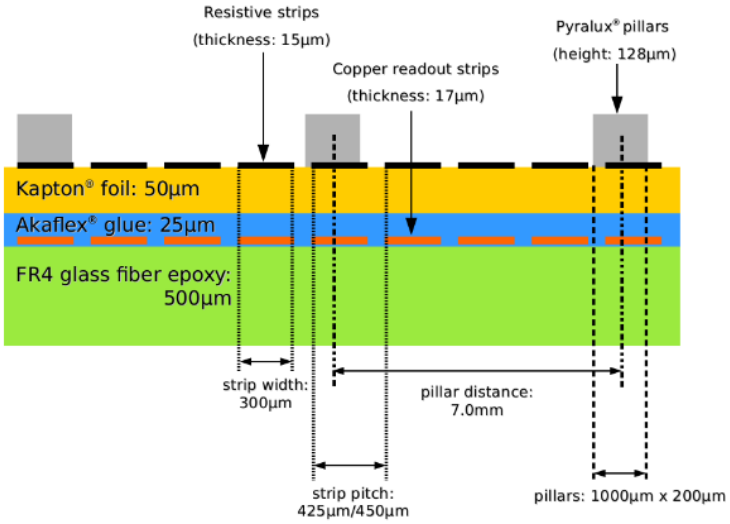


Figure 3.9: Sketch of readout MM board (not in scale) with the described components [84].

conductive paste) are screen-printed. As shown in Figure 3.10 (a), the HV is distributed to the resistive pattern through two HV antennas and silver lines at the two sides of a PCB since the resistive network is separated at the centre.
In the HV connection section (Figure 3.10 (b)) of every board, some dedicated pads are reserved for a low-pass RC filter in the HV supply line to avoid the DC noise from the power supply, with an R of 100 kΩ and a C of 2 nF.
It is followed by the deposition of the pillars and the coverlay over the HV distribution silver lines. The pillars are obtained by a photolithographic procedure from two 64 μm thick dielectric material films (Pyralux) to achieve the amplification gap height. The pillar basis size has been increased with respect to the original design to assure a better adhesion. Studies on the impact of pillar size on the detector performance have been carried out and reported in Reference [84]. The final design has pillars with a rectangular basis of 1000x200 μm^2.
After a curing and rest period, the board is cut and drilled to create the mechanical holes for the assembly and the internal alignment of the quadruplet.
At the end of this procedure, a cycle of quality controls has been performed on the readout Micromegas PCBs (see Section 3.5.1) before realising the readout panels.
An additional step, called passivation, has been introduced in readout panel assembly to reduce the spark rate observed in the first prototypes and modules. The passivation consists of covering a small portion of the active area using a dielectric material. It has been done on the regions with minimum resistance[4] < 1 MΩ localised at the Micromegas readout PCB edges.
The basic module of ATLAS NSW Micromegas detectors is a quadruplet constituted of four layers of resistive strips Micromegas detectors (as shown in Figure 3.11). Its size varies in Small (S) or Large (L) and Module 1-type (M1) or Module 2-type (M2), as shown in Figure 3.4. Independently on the

[4] The minimum resistance is measured by the probes with 1x1 cm^2 surface, as reported in Section 3.5.5.

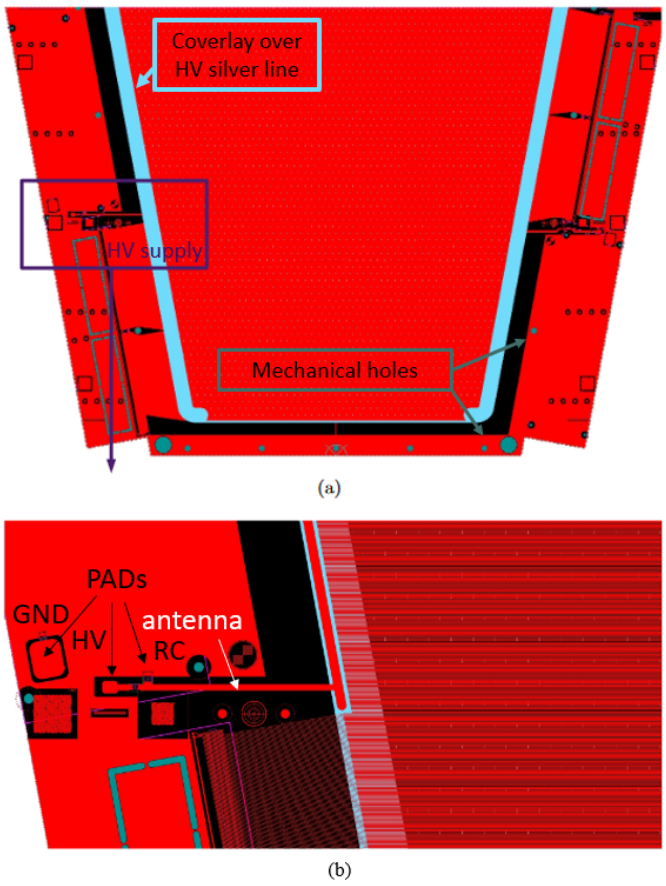


Figure 3.10: Horizontal plane view of a PCB (a). The coverlay over the silver lines (in light blue), the mechanical holes (in grey) for assembly and the connection between HV input line and the resistive pattern (purple) are pointed out. The zoom of HV connection section (b) consisting of an antenna and soldering pads for ground, HV connector and RC filter on PCB board and silver line.

size, a quadruplet or module has two readout (anode) panels, as shown in Figure 3.11. The central structure of a readout panel consists of a 10 mm Al honeycomb with an Al frame that provides a flat and stiff structure. Each side is covered by three (Module 2-type) or five (Module 1-type) Micromegas readout PCBs. The smallest PCB of Small Module 1-type has an active area of 0.26 m^2, and the largest one of Large Module 2-type, of 1.03 m^2. The readout panels in a quadruplet have different strips orientations. One of them is η-type, and the other is $stereo$-type. The strips of each η-panel side are parallel to the direction of the ATLAS η pseudorapidity (eta-panel), giving measurements of the ATLAS coordinate η. While the strips of a $stereo$-panel measure the ATLAS coordinate ϕ, combining the hits on the two sides of the panel that have respectively strips rotated by a 1.5° angle and by a -1.5° angle with respect to the eta-panel strips. The stereo configuration permits the rejection of the ghost tracks.
The micro-meshes of ATLAS Micromegas are stainless steel woven on plain weave with a 30 μm wire diameter and a 70 μm hole pitch, as shown in Figure 3.12.
Different from the most common Micromegas chambers in which the micro-mesh is encapsulated into

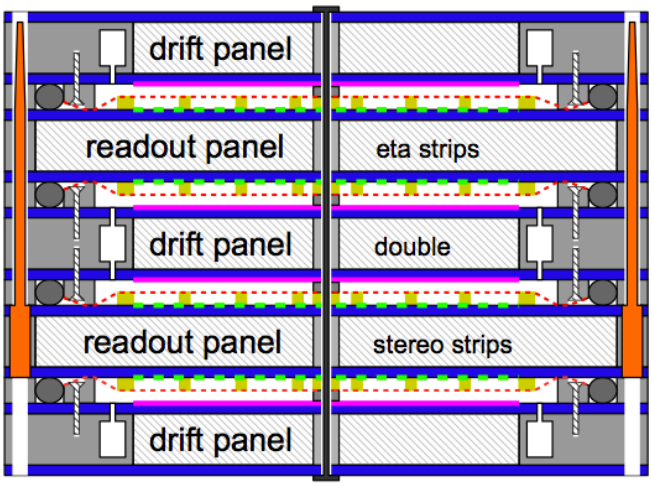


Figure 3.11: Transverse Z-plane of the a Micromegas quadruplet [85].

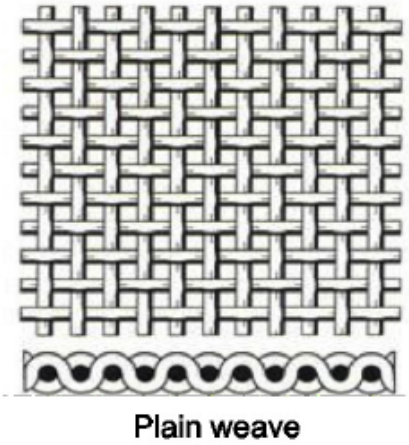


Figure 3.12: ATLAS Micromegas mesh type [83].

the pillars (bulk technique), each quadruplet layer has a mechanically floating [86] micro-mesh directly glued on the drift panel edges. This setting minimises the inactive area on the detector plane with respect to the bulk technique and it simplifies the amplification gap cleaning during the quadruplet assembly.
Dedicated panel cleaning, mesh smoothing and panel drying have been carried out to minimise the presence of possible impurities trapped in the mesh or on the active area of the readout panel, mesh mechanical imperfections and trapped humidity that can cause sparks.
The drift panel constitution and assembly are similar to those of the readout panel but not identical (see Figure 3.13). The differences are related to the different functions of this panel. They principally consist of a flat copper-clad FR4 plane as cathode and an outer frame structure that provides a solid

frame supporting the stretched micro-mesh by an average 10 N/cm tension. It also hosts grooves where the O-rings could be placed to improve the detector tightness and holes for the input and output lines of the gas inflation. Then, a quadruplet is formed by two readout panels (an *eta*-panel and a

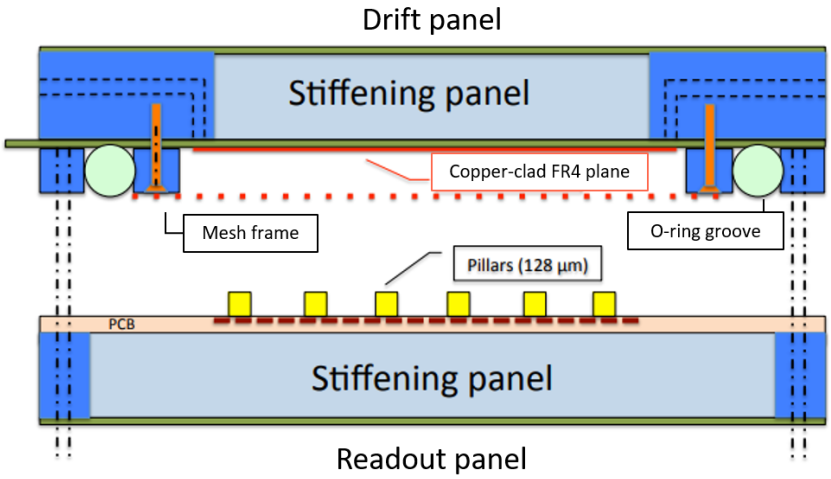


Figure 3.13: Sketch of the transverse of drift and readout panels [79].

stereo-panel) alternated by three drift panels, as shown in Figure 3.11. The central drift panel includes two drift gaps. An assembled quadruplet has a total thickness of 78 mm. During all the assembly steps, the predicted alignment and flatness accuracy have the most strict requirements because the detector will be exploited as a precision tracker.
The overall requirements on the quadruplet alignment tolerances (maximum standard deviation) are [87]:

- 40 μm for the position of the η-strips on the single plane;
- 60 μm for the alignment of the two faces of the readout panel;
- 60 μm for the relative position of the two readout panels;
- $< 34\,\mu$m for the planarity of all panels (on the z coordinate).

They agree with the requirements of the previous MDT and CSC positions reported in Section 1.2.5. For the gas flowing inside the detector, the holes are four in each drift panel, close to the trapezoid vertices. According to Figure 3.14, the gas flows through the input holes near the large basis (inlet line) of Module 2, it exits through the output holes near the small basis, and the gas circulates in Module 1. Along the gas pipes inside the drift panel frame, other inner holes are present with decreasing impedance to permit more uniform inflation. Moreover, readout and drift panels have interconnections in the centre to limit the deformation due to gas overpressure allowing the gas flow among the drift gaps. At the entrance of the holes of the drifts panels, the flux is divided in the manner that the inlet flux of the central drift panel is double with respect to the one in the external drift panels to provide the most uniform inflation as the central drift panel contains two drift gaps. The nominal flow rate

will be 28 (42) NL/h [5] for Small (Large) Modules [88] at the entrance of the flow rate divider. The tolerated gas leak rate must be lower than $10^{-5} \times V$ per minute, where V is the active detector volume.
The final Micromegas sector in Figure 3.3 foresees the ensemble of a pair of Module 1 and Module 2,

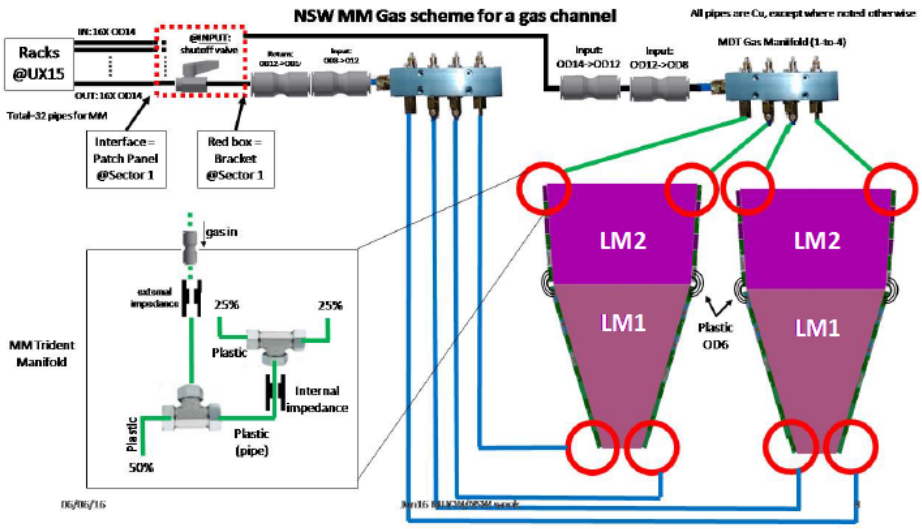
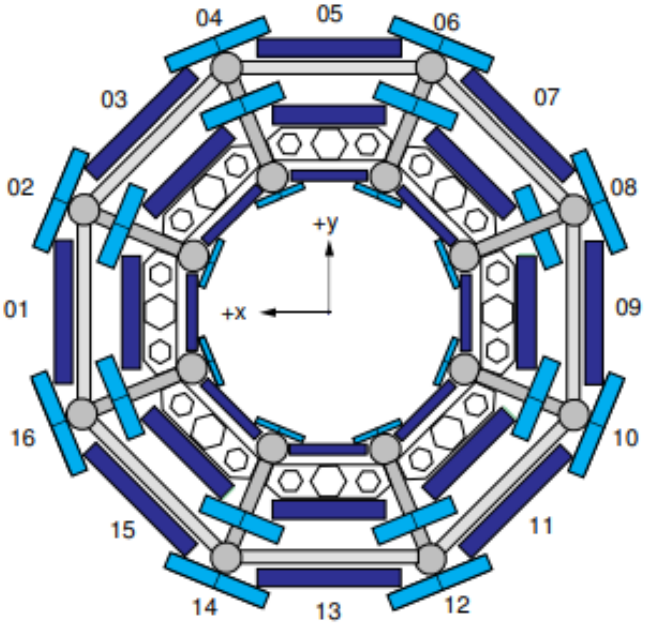


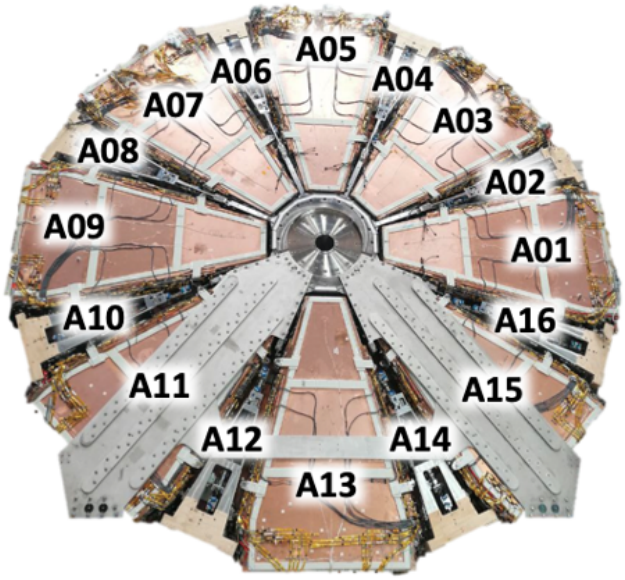
Figure 3.14: Configuration of a single gas channel providing gas mixture to two Micromegas wedges [88].

forming a Double Wedge (DW). A DW is already equipped with the final services (cables, sensors and electronics), including the VMM Front End electronics [89].
The wheel located in the End-Cap region identified by $z > 0$ ($z < 0$) is named Wheel A (C). Small (Large) DWs are placed in even (odd) Sectors of the Muon Spectrometer. The Sector positions in the ATLAS X-Y plane are shown in Figure 3.15 (a), and the respective DWs assume the nomenclature A (C) associated with the Sector number in Figure 3.15 (b).

[5] NL/h = liter at atmospheric pressure per hour



(a)

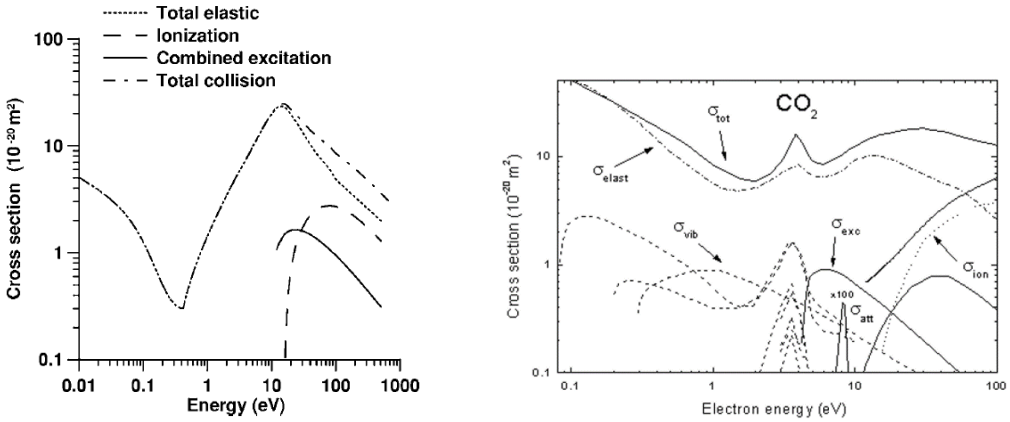


(b)

Figure 3.15: Definition of the 16 sectors in the ATLAS Muon Spectrometer in the ATLAS *X-Y* plane (a). Picture of NSW A with superimposed the sector nomenclature (b).

3.4.2 Gas mixture

Two different gas mixtures are taken into account for the operations of ATLAS Micromegas detectors. They are Ar:CO_2(93%:7%) and Ar:CO_2:iC_4H_{10}(93%:5%:2%). Both mainly consist of Ar gas since, like all the noble gases, this gas has a low electron affinity that minimises the signal losses. However, a small polyatomic gas (or quencher) addition to a noble gas is required to prevent secondary effects that involve the signal electrons and compromise the proportionality of the internal amplification of the signal, specific to a proportional mode gas detector (see Section 3.3.1).
Looking at Figure 3.16, the ionisation cross-section of Ar gas increases with the electron energy up to an electron energy range where the Ar excitation becomes more favourable than the ionisation. Then,



(a) Ar gas (b) Carbon-dioxide

Figure 3.16: Electron cross-section as a function of electron energy in Ar gas [90] (a) and in Carbon-dioxide (CO_2) [91] (b), where σ_{elas} refers to momentum transfer cross-section, σ_{vib} to vibrational, σ_{att} to attachment, σ_{ion} to ionisation and σ_{exc} to excitation cross-sections as well as σ_{tot} refers to the total cross-section.

a noble gas, like Ar gas, can only de-excite by photo-emission. The emitted photon has enough energy to extract an electron from the metallic detector components that can start a different avalanche in the detector. The solution comes with the small addition of a polyatomic gas, like Carbon dioxide, that absorbs the possible Ar de-excitation photons and reduces the possible electron energy growth through its accessible rotational and vibrational energy levels [92] [93]. The CO_2 gas and iC_4H_{10} vapour act as quenchers in the mixtures for ATLAS Micromegas detectors.
The more complex iC_4H_{10} molecule provides a shorter nuclear interaction length and larger minimum loss energy in the path than CO_2 gas (in Table 3.1), resulting in a more efficient quenching. In Figure 3.17, the efficiency curves of ATLAS Micromegas detectors as a function of the amplification voltage show that the Ar:CO_2:iC_4H_{10}(93%:5%:2%) mixture reaches the efficiency plateau at lower voltages than Ar:CO_2(93%:7%) mixture, reducing the electrical stress due to higher voltage operations. The 2% of isobutane vapour already satisfactorily minimises the instabilities of the current observed with the Ar:CO_2(93%:7%) mixture in ATLAS Micromegas detectors, consisting of frequent discharges (see Section 3.6.2). Some first results are in Reference [95].
The iC_4H_{10} vapour can not entirely substitute the fraction of CO_2 gas in the binary mixture with Ar gas because a fraction larger than 2% makes the gas mixture inflammable.

Table 3.1: Atomic/molecular and nuclear properties of Argon, carbon dioxide and isobutane. The gases are evaluated at 20 C and 1 atm (in parentheses) or at STP [square brackets] [94].

Material	$\langle Z/A \rangle$	Nuclear collision length λ_T (g/cm^2)	Nuclear interaction length λ_I (g/cm^2)	Radiation length X_0 (g/cm^2)	$\mathrm{dE/dx}\big\|_{min}$ $\left(\frac{MeV}{gcm^2}\right)$	Density (g/l)	Liquid boiling point at 1 atm (K)
Ar	0.45059	76.4	117.6	(1.519)	19.55	1.396(1.782)	87.28
CO_2	0.49989	62.4	89.7	(1.819)	36.2	[1.977]	
iC_4H_{10}	0.58496	56.4	77.0	(2.239)	45.07	[2.67]	261.42

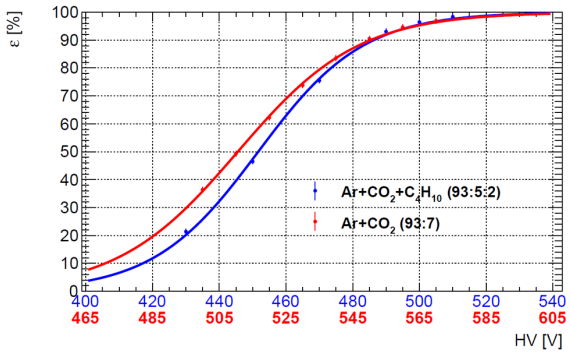


Figure 3.17: Curves of nominal tracking efficiency ϵ[%] as function of the amplification voltage. The red curve is relative to ϵ[%] with Ar:CO_2 (93:7)%, it was measured on Small DWs. The blue curve is relative to ϵ[%] with Ar:CO_2:iC_4H_{10} (93:5:2)%, and it was measured on A13 DW.

Moreover, polyatomic gases, especially those with large molecular structures, can dissipate the gained energy by breaking themselves into smaller molecules that can be reactive with the detector material. Preliminary results of ongoing detector ageing studies prove that the Ar:CO_2:iC_4H_{10} (93:5:2)% mixture does not accelerate the detector ageing with respect to the Ar:CO_2 (93:7)%.

3.5 Statistical studies

In parallel to the production and validation of the ATLAS MM detectors, several statistical studies have been performed on the parameters that characterise the detector to quantify the agreement between the nominal design and sample specifications and to provide the input parameters for the simulations of the detector performance.

3.5.1 Quality analysis of NSW PCBs

Before the final assembly of the readout panel in the different construction sites of the SM1, SM2, LM1 and LM2 Modules, the mechanical and electrical specifications of each produced PCB have

been checked and stored in a database at CERN.

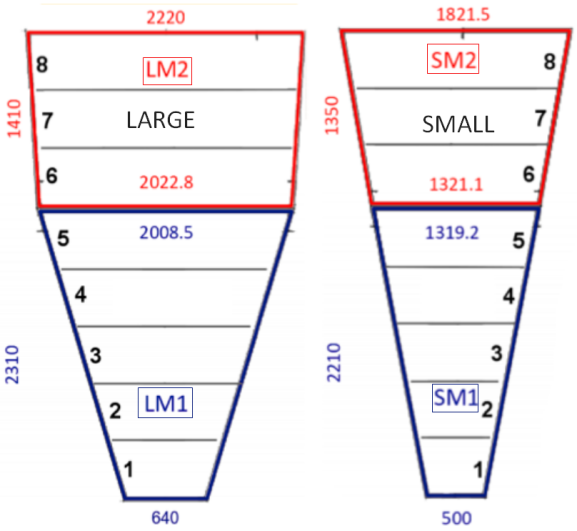


Figure 3.18: Sketch of readout planes in which the PCB subdivisions are visible together with the nominal values of the bases and heights reported for the Large (Left) and Small (Right) Modules.

In the figures of successive sections, the adopted PCB nomenclature is the following: the first letter represents the module size (S-small or L-large), the second letter is about the strip orientation (E-η or S-*stereo*), the first number indicates the PCB size (from 1 to 8), and the remaining numbers identify the single PCB. Two companies have been involved in their production, for a total of ~ 3000 units. ELTOS (IT) company supplied the PCB types: Small Eta from 1 to 8 and the Small Stereo from 1 to 5. ELVIA (FR) produced the remaining types. The boards have been realised using standard PCB manufactures with new specific steps: the resistive layer glueing, the connection between the HV distribution line and resistive layer, and the pillar etching (in Figure 3.19). Therefore, several adaptations on these steps have been made to reach the demanded requirements. The quality controls (QCs) have been performed at CERN. The QCs are grouped into 7 control groups:
visual checks, electric tests, pillar adhesion and height, measurements of resistance and superficial resistivity, measurements of capacitance, alignment of resistive and readout strips, measurements of dimensions and shape.

Without entering into the details, the entire QC set aims to provide the best sample of readout Micromegas boards in terms of electric field uniformity and final tracking performance (dead area and mechanical alignment). The following subsections go shortly through each control group.

3.5.2 Visual checks

During this control, the active area is carefully inspected by eyes and microscope, searching for possible defects that compromise the uniformity of the electric field, like *enclosures, bumps, scratches, bubbles under the dielectric coverlay, screen-printed defects, and no-suitable pillars*. Also, the quality of the HV distribution silver line (SL) to the resistive strips and readout strip terminations are controlled, together with the ratio between the readout strip width and its pitch.

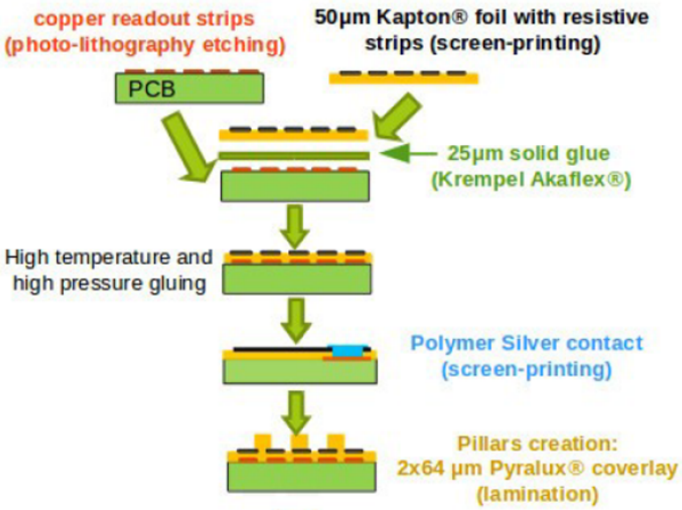


Figure 3.19: Simplified production scheme of ATLAS-MM readout PCB boards.

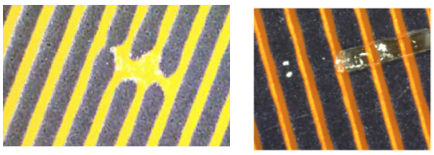


Figure 3.20: Picture of some defects. A not severe screen-printed defect (left) and badly-etched pillar (right).

3.5.3 Electric tests

Electric tests are performed to check the insulation between the coverlay (CL) on top of the HV distribution silver line (SL) and the resistive spark suppression layer, and the insulation between the readout and resistive strips. These tests fail if holes are present in the CL or in the Kapton insulating layer. Consequently, the SL or the readout strips are directly exposed to the charges, starting direct current through them.
After applying a DC ΔV = 1 kV, the tolerances are: a measured > 30 GΩ resistance between the coverlay on top of the HV distribution silver line and the resistive spark suppression layer, and a measured > 3 GΩ resistance in 60 s between the readout and resistive strips.

3.5.4 Measurements of the pillar height

The pillars are dielectric components that define the height of the amplification gap (see Section 3.3.1) and, so, the uniformity of the amplification electric field.
The tolerance on the pillar height non-uniformity has been defined to be less than 5 μm on a 10x10 cm^2 region. This requirement corresponds to a uniform electric field within 4-5% (assuming 120

μm as nominal pillar height). For each PCB board, the pillar height is measured by 4 length gauges, fixed in handy mechanical support, in several positions obtaining a 2D map. There is a systematic difference in mean height between the ELVIA and ELTOS boards, related to slight differences in the manufacturing procedures of the two companies, as shown in Figure 3.21.
In addition, few ELTOS boards have satisfied the requirement with a mean pillar height higher than

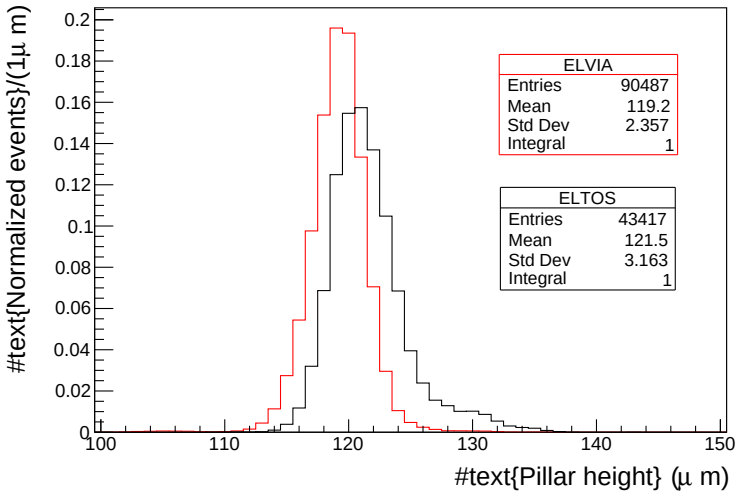
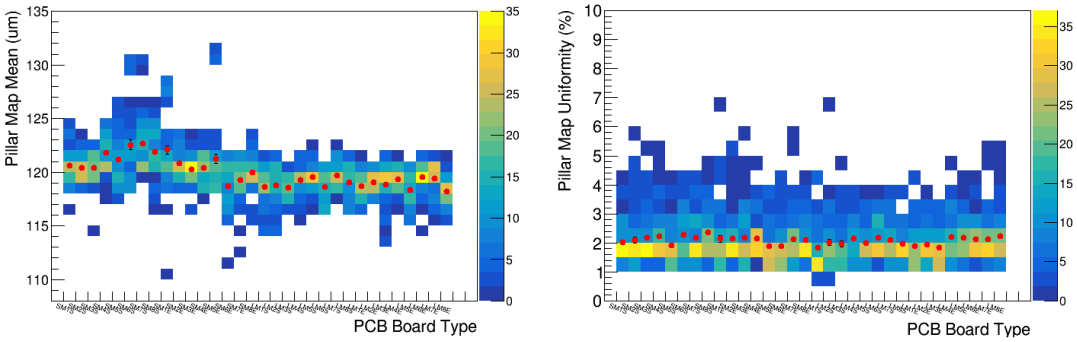


Figure 3.21: Histograms of the map measurements of pillar height per PCB production company: Elvia is represented in red and Eltos in black.

the nominal mean height (130 μm vs the nominal 120 μm) because the starting Pyralux foils have had larger thickness than the design value. It is not critical because a dedicated HV channel supplies each readout board.
The mean value and the uniformity of every map are reported in Figures 3.22 (a) and 3.22 (b) per each type of board, respectively. The uniformity is defined as the ratio between the RMS and the Mean value of the map in per cent. As visible in Figure 3.22 (b), this ratio is generally less than 5%, and its mean values are around 2%. Therefore, the pillar uniformity requirement has been fully satisfied.

3.5.5 Measurements of the resistance and surface resistivity

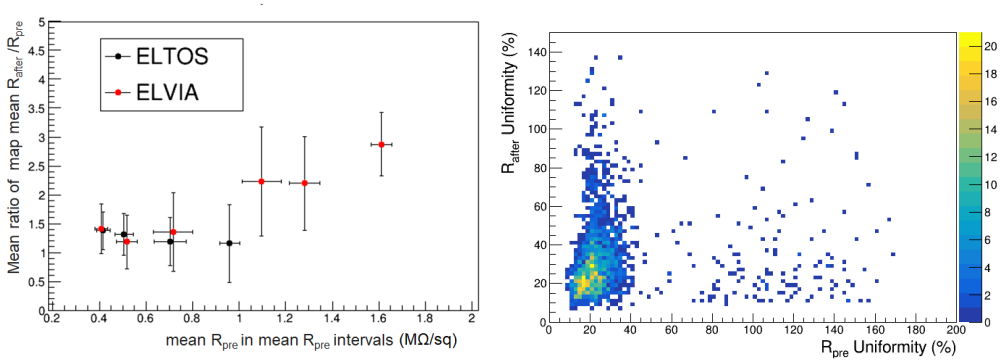
The minimum and maximum resistance are measured using 5 independent probes with a 1x1 cm^2 surface at four different distances from the centre of the board: close to the edges (Left and Right) and in the central zone (Central-Left and Central-Right). In addition, two 2D maps of surface resistivity of every resistive layer are acquired to monitor the local values and their uniformity before and after its glueing on the PCB.
The two maps are acquired by two identical tools and in different places. The first measurement happens soon after the resistive layer production in Japan. The second map is acquired at CERN few months after glueing the resistive layer on the PCB.
By design, a requirement is that the voltage drop between the centre and the edge has to be lower



(a) (b)

Figure 3.22: Pillar height (a) and uniformity (b) histograms per each PCB type; uniformity is RMS/mean. Red markers identify the mean value of histograms.

than 5 V, assuming an O(μA) current. An additional QC requirement has been introduced on the mean surface resistivity after the glueing (R_{after}). The ratio between R_{after} and the mean surface resistivity before glueing (R_{pre}) has to be lower than 3(3.5) because boards with a larger ratio have shown damaged resistive strips after the glueing.
The foil resistivity increases when it is pressed at a high temperature, but the increase is not easily tunable. It depends on the initial resistivity as visible in the trend of the $\frac{R_{after}}{R_{pre}}$ ratio as a function of mean values of the R_{pre} sub-samples with equal statistics. As visible in Figure 3.23 (a), this increase is greater when foil resistivity R_{pre} is larger. In some cases, the resistivity increase also affects the uniformity, as shown in Figure 3.23 (b).
As reported in Section 3.4, resistive paste composition differs slightly among the production batches.



(a) (b)

Figure 3.23: Mean ratio between the map mean R_{after} and R_{pre} as a function of mean R_{pre}. The error bars are filled with RMS to give info about width of distributions (a). R_{pre} uniformity vs R_{after} uniformity (b).

Then, a dedicated study has been carried out to investigate the impact of these differences on the board resistivity. In the following figures, the convention of batch0$_a$ and batch0$_b$ is kept together with chronological enumeration after 0: batch1, batch2 and batch4. The batch0$_a$ is relative to the resistive

paste of pre-series production and beginning of series production, and batch0$_b$ is relative to the first paste only used in series production. The differences per batch are visible before (in Figure 3.24 (a)) and after (in Figure 3.24 (b)) the foil glueing. The mean R_{pre} relative to different batches is generally lower or compatible with the values of batch0$_b$ (full black in Figure 3.24 (a)). The mean R_{after} is, in general, higher or compatible with the values of batch0$_b$ (full black Figure 3.24 (b)). However, these differences are not relevant to the QC requirements.

Finally, the histograms of the mean surface resistivity and uniformity are respectively shown in Figure

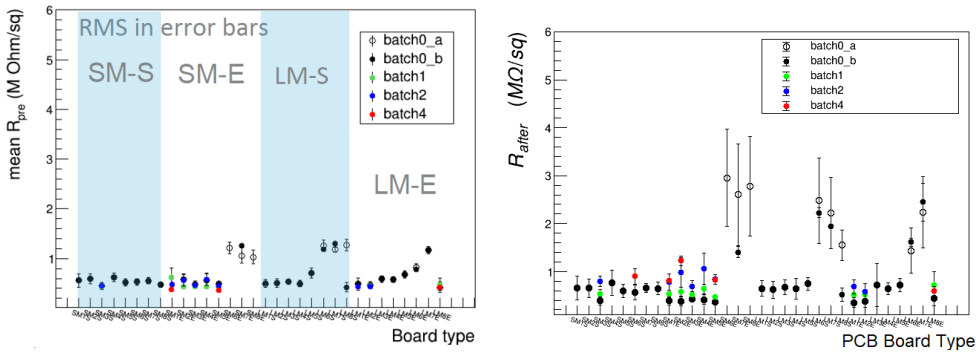
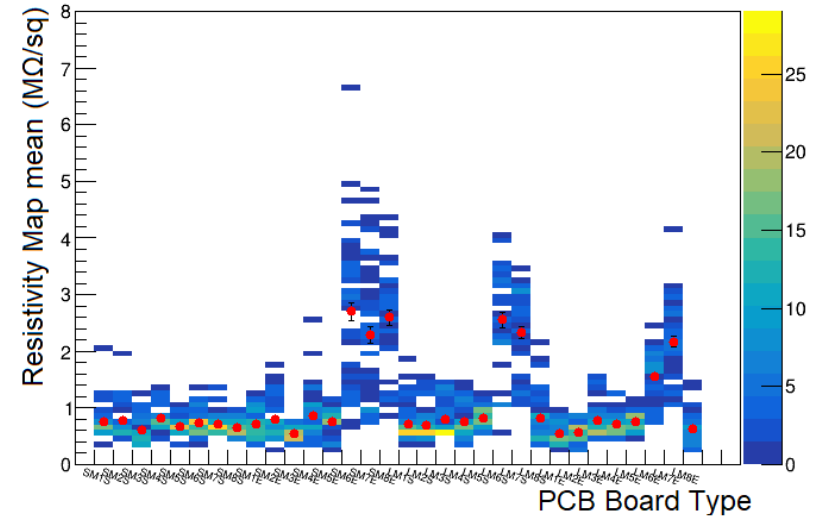
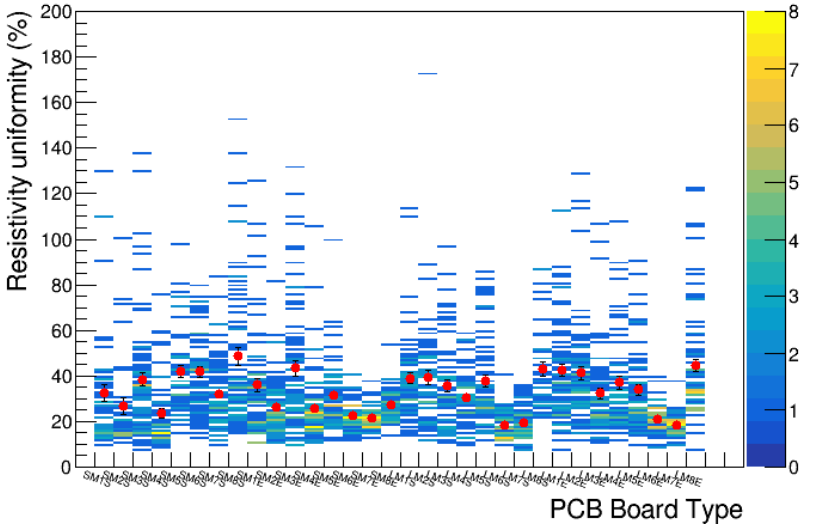


Figure 3.24: Map mean R_{pre} (a) and R_{after} (b) per each PCB type and resistive foil batch. The batch id number is in chronological ordered, batch0$_a$ is the first batch up to batch4.

3.25 (a) and Figure 3.25 (b) per each board type.



(a)



(b)

Figure 3.25: Surface resistivity (a) and uniformity (b) histograms per each PCB type. Red markers identify the mean value of histograms.

Categorisation of the ATLAS Micromegas resistive boards

Similar to the grade classification of the resistive foils in Section 3.4, the ATLAS Micromegas resistive boards have been divided into categories that could represent the uniformity of the resistivity map in a compact way. The definitions of the categories and uniformity differ from the grades and resistivity uniformity (%) reported in Figure 3.25 (b). The being-introduced definitions are based on the mode ($Mode$) and FWHM ($FWHM$) of the histogram in Figure 3.26 of all resistive map values of the ratio u, defined as the ratio between a single measurement and the mean value of its resistivity map.
The most uniform category (CAT) 1 includes the resistive maps with 90% of the measurements with

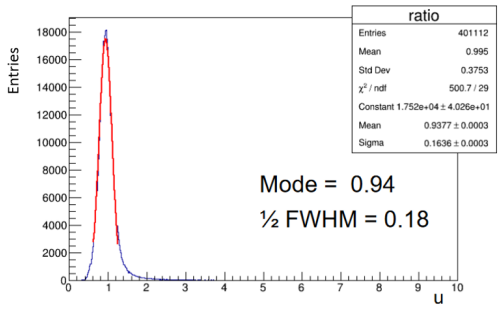


Figure 3.26: Histogram of the ratio u for all the resistivity measurements.

the ratio u in the interval $[Mode - 1/2FWHM, Mode + 1/2FWHM]$, the resistive maps of CAT 2 have 90% of the measurements with the ratio u in the interval $[Mode - 1FWHM, Mode + 1FWHM]$. While the resistive maps of CAT 3 have 90% of the measurements with ratio u in the $[Mode - 3/2FWHM, Mode + 3/2FWHM]$ interval. The rest of the resistive maps are attributed to CAT 4. The uniformity slightly improves in the batch 2, as visible from the histogram parameters in the legends of the right side of Figure 3.27. The same categorisation has been repeated on the measurements of the initial resistive maps of the resistive foils, based on the mode and FWHM of the histogram of the corresponding ratio u_{pre}. To evidence the different case with respect to the maps of resistivity after the glueing on the PCB, the uniformity categories relative to resistivity maps of the resistive foils before glueing are indicated as PRE CAT on the left side of Figure 3.27. The uniformity of resistive foils before the glueing has slightly worsened from batches 0 to batch 1 and 2, as observable on the left side of Figure 3.27. However, these observed variations in the uniformity of the resistivity maps before and after glueing are negligible. They do not introduce systematic contributions to the uniformity of the overall sample, in agreement with the conclusions in the mean resistivity (in Figure 3.24) discussed before.

3.5.6 Measurements of the readout strip capacitance

The readout strip capacitance is measured by an LCR-meter with respect to the HV line input. The test identifies unconnected readout strips (≤ 20 pF). The requirements on a readout MM board are: no more than 2 consecutive unconnected strips; and more than 98% connected strips in each half of the board. This measurement can not give direct access to the capacitance value because the resistance of

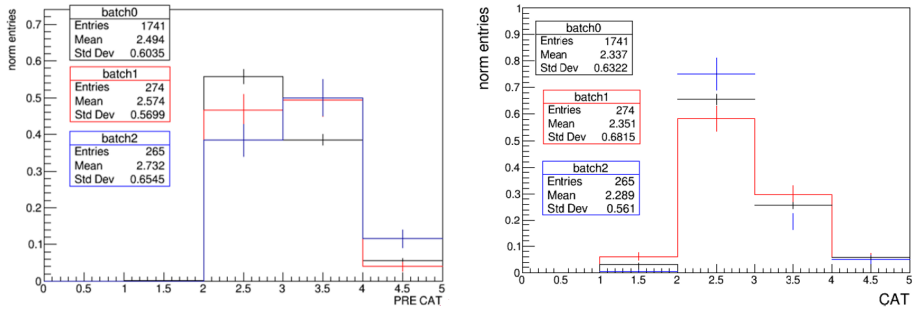


Figure 3.27: Histograms of the uniformity categories PRE CAT (left) and CAT (right) per batch.

the resistive layer is included in it.
To qualitatively describe the relation between the measured capacitance and the above resistance, the LCR-meter is approximated as a Wheatstone bridge circuit (in Figure 3.28) and the anode board as an RC series (with $\tau_{det} = R_{det}C_{det}$) in the branch to be balanced.
Then, the effective capacitance C_{det} of the anode board is equal to Equation 3.1, where i is the current

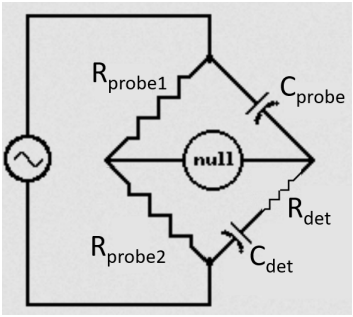


Figure 3.28: Equivalent Wheatstone bridge circuit of the LCR-meter (probe) and detector approximated as an RC series.

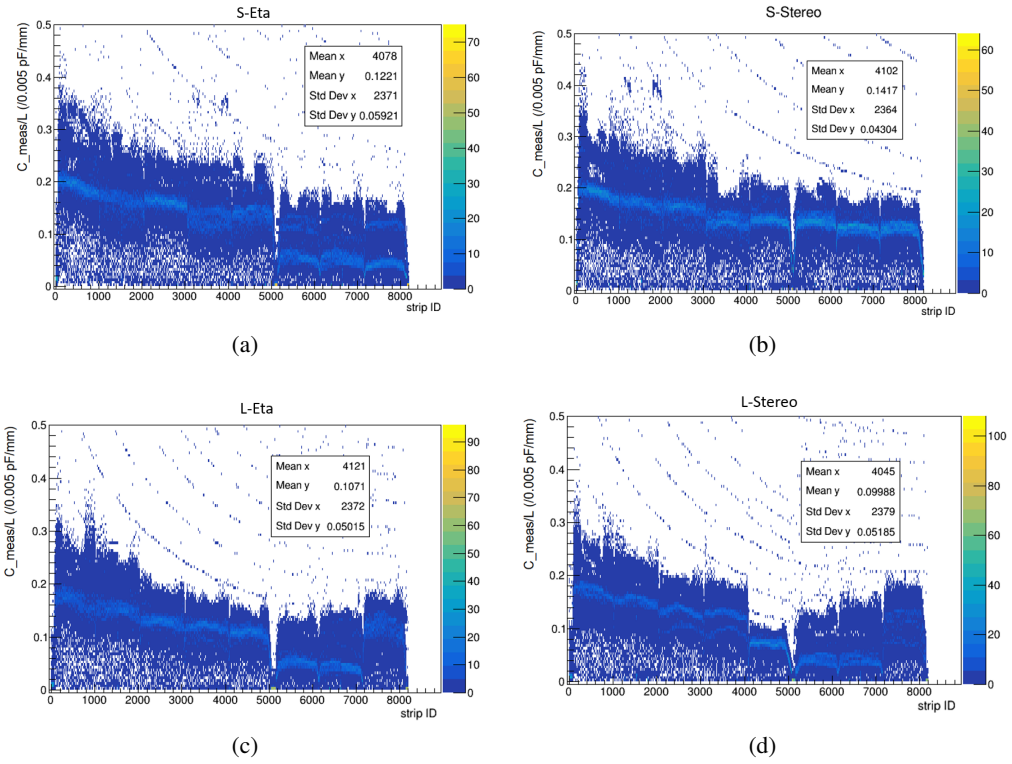
through the branch. The used LCR-meter has an internal oscillating generator with angular frequency ω. Assuming that the measurement time is larger than the circuit characteristic time τ, the functional dependence between the C_{meas} and R_{det} become like in Equation 3.2, at a fixed time of measurement t, length of the strip and given ω.

$$C_{det} = \left[\frac{1}{C_{meas}} - \frac{R_{det}}{i}\frac{\mathrm{d}i}{\mathrm{d}t}\right]^{-1} \tag{3.1}$$

$$\frac{1}{C_{meas}} = \mathrm{coeff}(\omega)|_t R_{det} + \frac{1}{C_{det}} \tag{3.2}$$

Beyond the dependence on the R_{det}, the C_{meas} varies as the strip length. The 2D histograms in Figure 3.29 between the strip ID and C_{meas} per unit strip length are quite uniform for the strips at the centre of the boards.
However, the approximation does not consider the separation at the centre of the resistive layer, i.e. only the capacitance of half strip is in series with the detector resistance. This consideration may

partly explain the decreasing trend observed from the shortest to the longest strips in Figure 3.29. The contribution of the R_{det} is visible in Figure 3.30 and in the cases of SM6E[6], SM7E, SM8E, LM6S, LM7S and LM6E, LM7E board types that have a larger mean resistivity in Figure 3.25 (a).



(a) (b)

(c) (d)

Figure 3.29: 2D Histograms between the strip ID and C_{meas} per unit strip length for the different sizes of modules Small (S) and Large (L) and strip orientations Eta and Stereo. Figure (a) is relative to S-Eta, (b) is S-Stereo, (c) is L-Eta and (d) is L-Stereo. The SM6E, SM7E and SM8E sizes correspond to the strip ID belonging to (5120, 8192) in Figure (a). The LM6E and LM7E sizes and the LM6S and LM7S sizes correspond to the strip IDs from 5129 to 7168 in Figure (c) and Figure (d), respectively.

[6] SM6E indicates Small Module PCB 6 strip orientation *eta* and analogous for the other acronyms.

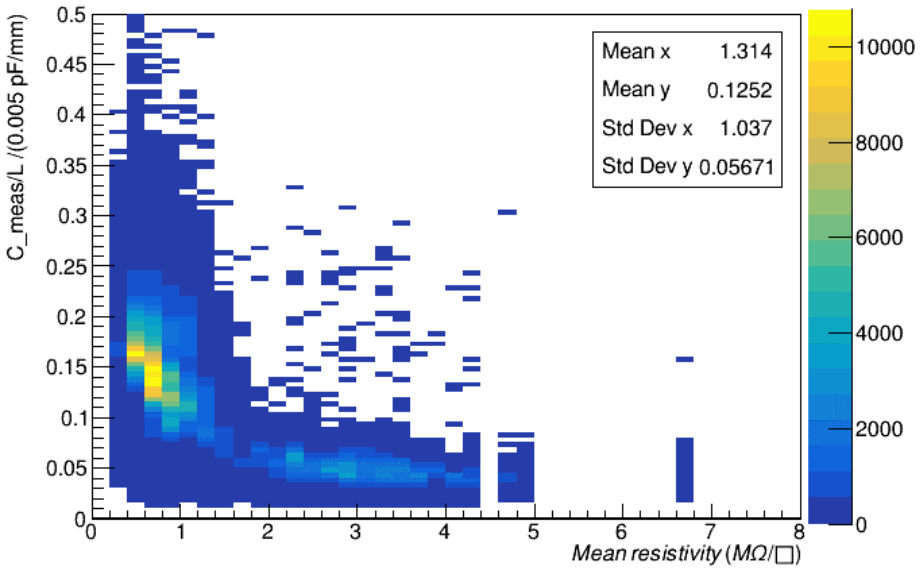


Figure 3.30: Capacity C_{meas} per unit strip length as a function of mean PCB resistivity. The relation between the R_{det} and the mean PCB resistivity is exploited to prove the correlation between C_{meas} and R_{det}.

3.5.7 Alignment between resistive and readout strips

The alignment between the resistive and readout strips affects the accuracy of the track projection on a readout plane because the signal is capacitively inducted on readout strips [96]. The requested maximum misalignment in strip direction is 0.06°, corresponding to a 1 mm shift in the perpendicular strip direction over ~2 m of strip length and a 3 mm maximum shift in the longitudinal strip direction (see Figure 3.31).

During the QC validation, no boards have been rejected due to large misalignment between strips.

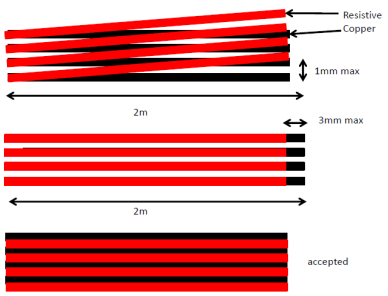
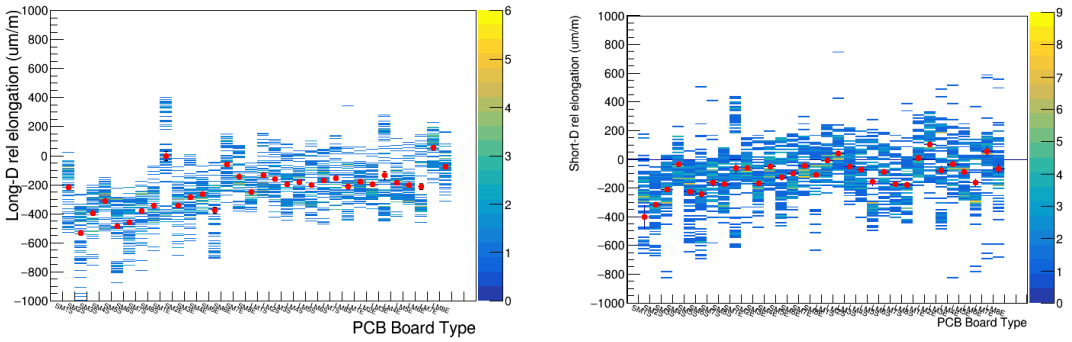


Figure 3.31: Maximum accepted misalignment between resistive and readout strips.

3.5.8 Measurements of the readout board elongations and shape

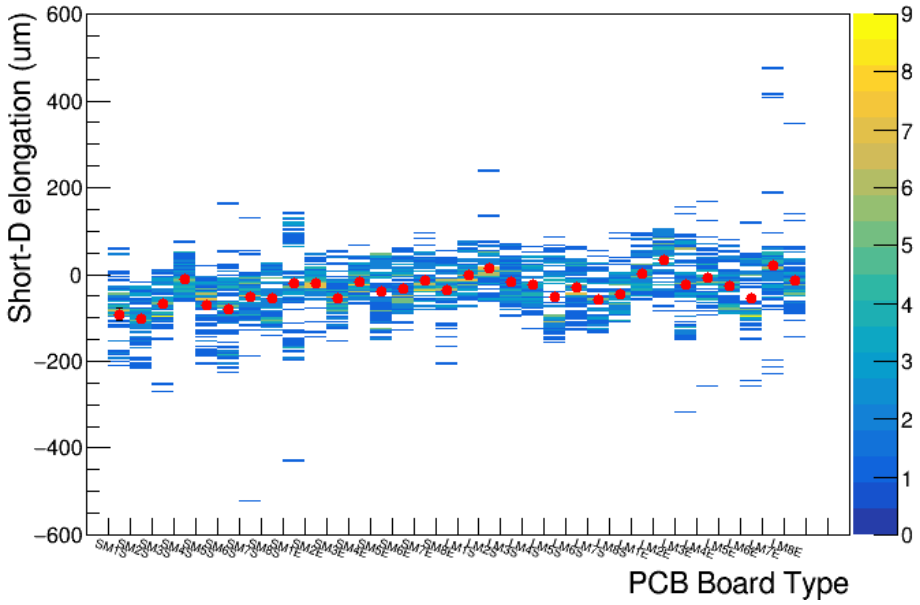
Deformations of the readout MM board in time (from the resistive foil glueing up to readout panel assembly) are expected because the FR4 of PCB is a hydrophilic polymeric material. Specific corrections have been applied to the etching masks of the readout strips based on the relative humidity of the production company sites.
The readout MM board dimensions have been measured at CERN together with the edge precision of the readout MM board bases several weeks after the resistive foil glueing. Their values are stored in a database to keep information on the deformation of the sector, from the PCB level up to the Module level, and to assure the required mechanical alignment precision on the wheel reported in Section 1.2.5. For the strip (or long) direction, the relative 1D elongation (μm/m) is shown as a function of the board type in Figure 3.32 (a). While Figure 3.32 (b) shows the relative 1D elongation (μm/m) in the short direction, perpendicular to the strip.
The elongation measurements have been acquired in different temperature and relative humidity (RH) conditions. Even though no corrections have been applied to scale the measurements at fixed temperature and RH, most of the measurements satisfy the imposed requirements.
A measured relative elongation in the long direction has to be within the interval ±500 μm/m around the mean value (red markers in Figure 3.32-Left) of the distribution per PCB type. The elongation in



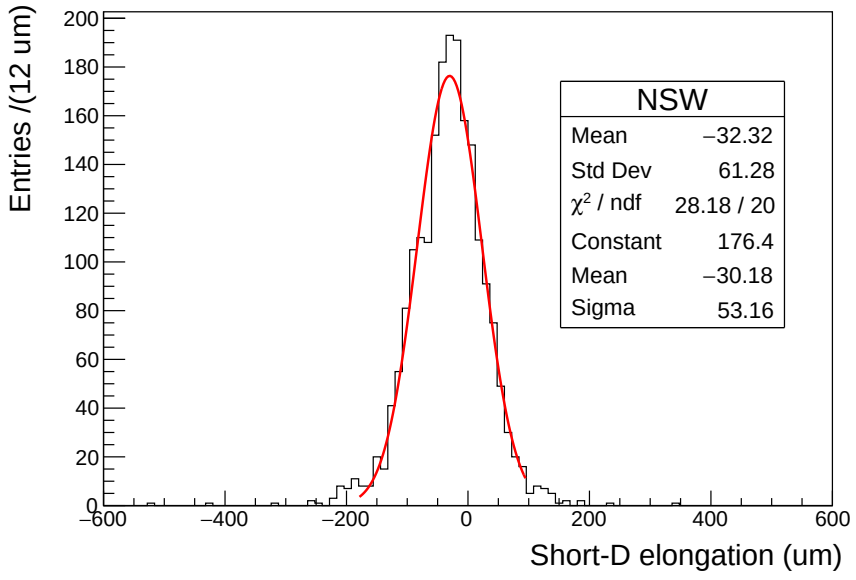
(a) (b)

Figure 3.32: Relative elongation in the long direction (a) and in the short direction (b) per each anode board type (with 20 μm as the precision of the measurement). On the y-axis, the bin width is 10 μ/m. Red markers identify the mean value of histograms.

the short direction (in Figure 3.33 (a)) is the most critical among the two as it influences the final alignment precision on the η of the assembled sector.
The measured elongations in the short direction satisfy the requirement to be within the interval ±100 μm around the mean value (red markers in Figure 3.33 (a)) of the distribution per PCB type. In the 7.6% of the entries in Figure 3.33 (b) (relative to readout MM boards in the final NSW detectors), an exception has been made in the acceptance of the boards on the basis that the magnitude of the relative elongation is still below 0.1% (see Figure 3.32 (b)).



(a)



(b)

Figure 3.33: Elongation in the short direction (a) per each anode board type (with 20 μm as the precision of the measurement). On the y-axis, the bin width is 10 μm, and red markers identify the mean value of histograms (b). Histogram of elongation in the short direction for the readout MM boards in the final NSW detectors.

3.6 Validation tests of the ATLAS Micromegas detectors

Each Module, formed by four independent resistive Micromegas layers, has been tested in terms of high voltage behaviour and tracking efficiency before (with temporary FE electronics) and after integrating with other modules to make an entire Micromegas wedge, fully equipped with cooling channels, sensor and front-end electronics.

3.6.1 Validation tests of Micromegas quadruplets

For the last Modules, their validation tests of high voltage behaviour and tracking efficiency have been performed at the CERN site for the demands of the assembly schedule. The experimental set-up of the CERN validation tests is described in Appendix C.
The protocol of the validation in terms of high voltage performance consists of a supervised voltage ramp up and the monitoring of the time trend of the current in a few days, at the maximum reachable voltages. The nominal working point of the amplification voltage for the ATLAS Micromegas is 570 V with the gas mixture Ar:CO_2(93%:7%).
The ramp up is performed through voltage steps (50 V,100 V,300 V,400 V,450 V,470 V,490 V, V_{step-1}+10 V, ..., 570 V). During the test, a spark is defined as a current variation greater than 100 nA, and the current-time trend (see Figure 3.34) has to have an average value and a spark rate smaller than 50 nA and 6 sparks/min, respectively. The ramp up is slowly performed because some HV sectors do

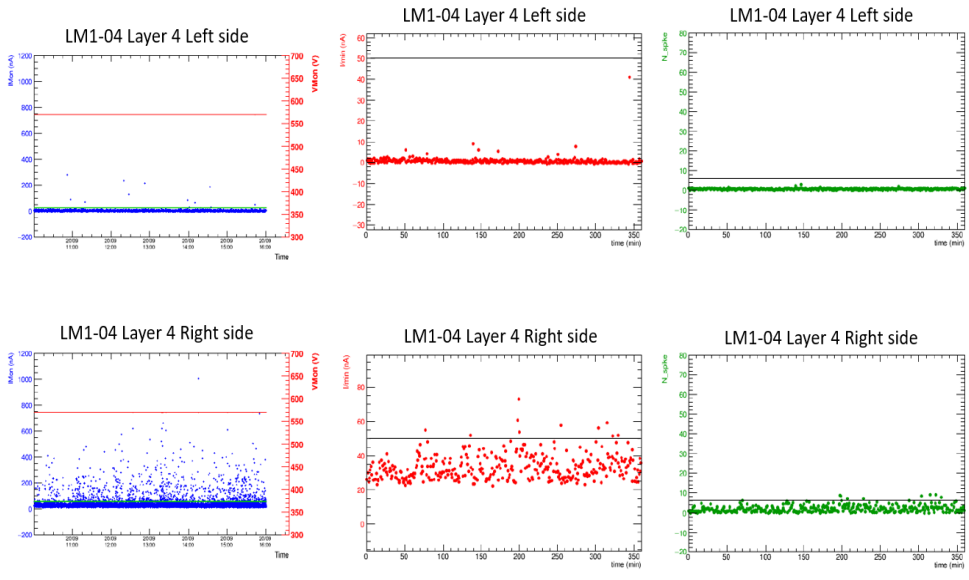


Figure 3.34: Example of HV validation plots with Ar:CO_2(93%:7%) for two HV sectors of the LM1-04 Module, which has not been included in the NSW. On the left, the time trends of the current I_{mon} (in blue) and voltage V_{mon} (in red) of the resistive strips. In the centre, the time trend of the average (in 60 min) current (in red). On the right, the number of sparks (in green) as a function of the time. The black lines represent the tolerated upper limits, i.e. 50 nA for the average current and 6 sparks/min for spark rate.

not firmly operate at 570 V. The lowest values are set to find possible short circuits or $O(< 10M\Omega)$ parasitic resistance in the external RC filter, in general relating to its bad soldering. While some constructive defects or trapped impurities set the stable HV working point between ~ 430 and 550 V. They are empirically identified through the resistance value of $O(> 10M\Omega)$ between the RC input and ground, using a tester with ΔV = 500 V.
In Section 3.4.2, it has been reported that the current instabilities are satisfactorily minimised by replacing 2% of the carbon dioxide with isobutane vapour in the gas mixture, which is also visible comparing the two sides of Figure 3.35.
Part of the NSW Modules has also been tested and validated with the Ar:CO_2:iC_4H_{10}(93%:5%:2%)

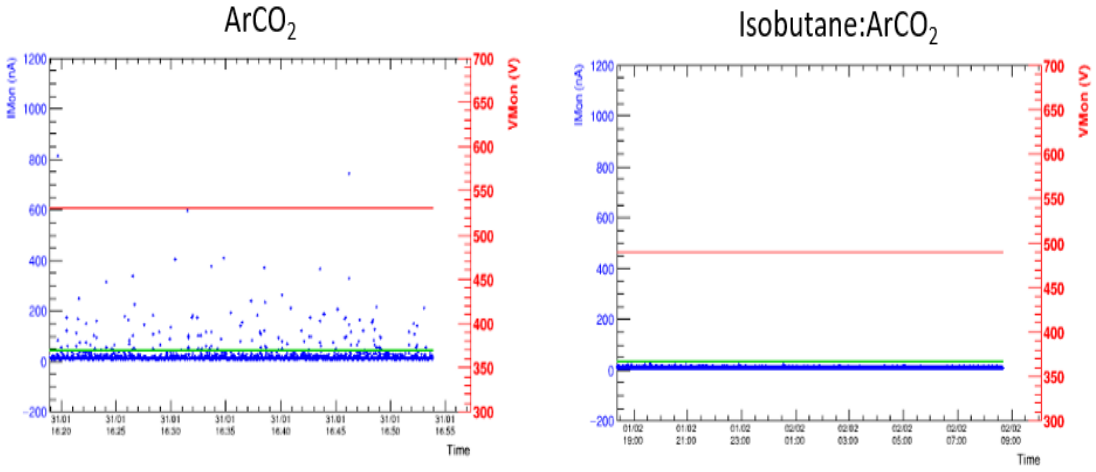


Figure 3.35: Current time trends of the Layer 3-PCB 4 right side sector of LM1-31 Module, with Ar:CO_2(93:7)% at 530 V (left) and with Ar:CO_2:iC_4H_{10}(93%:5%:2%) at 490 V (right), as amplification voltage that corresponds to a similar or greater gain factor respect the case Ar:CO_2(93:7)% at 530 V.

mixture, like the case shown in Figure 3.35.
As the isobutane has higher attenuation with respect to the carbon dioxide, the spark definition and upper tolerance limits have been changed to perform meaningful comparisons with the Ar:CO_2 mixture. When the mixture is Ar:CO_2:iC_4H_{10}(93%:5%:2%), a current variation greater than 25 nA is considered a spark. The maximum average current and the spark rate are 10 nA and 0.6 spark/min, respectively.
After finding the stable HV configuration, the validation tests foresee the measurements of tracking efficiency of all the Module layers.
Firstly, a track from the cosmic rays, detected by the trigger system in Appendix C, is reconstructed in three layers of a Module, excluding a layer each time. The first coordinate of the track is directly measured from the hits on the η layers. The second coordinate is defined by means of Equation 3.3 for the η layers and Equation3.4 for the stereo layers, assuming zero the distance between the most inner η readout layer and the stereo readout layers.

$$y = \frac{x_{stereo1} - x_{stereo2}}{2\sin(\theta)} \tag{3.3}$$

$$y_{1(2)} = \frac{x_{inner-\eta}\cos(\theta) - x_{stereo2(1)}}{\sin(\theta)} \tag{3.4}$$

Then, the reconstruction algorithm searches for possible clusters on the excluded layer in the fiducial section (or bin) around the projection of the previously reconstructed track in the other three layers. Finally, the efficiency of the excluded layer is evaluated as the ratio between the number of the measured clusters in the fiducial sections and the number of the projections.
For the 1D efficiency map, the fiducial bin width is ±20 mm for both the η and $stereo$ layers (i.e. a group of ∼ 100 strips).
In the 2D efficiency map, the bin width on the stereo layer is increased to ±25 mm (i.e. a group of ∼ 125 $stereo$-strips), to take into account the relative rotation (±1.5°) between the η and stereo strips. The bin width is the same for the η and stereo for the y-direction (± 5 mm), as visible in Figure 3.36.
None of the validated modules at CERN has been rejected due to low efficiency.

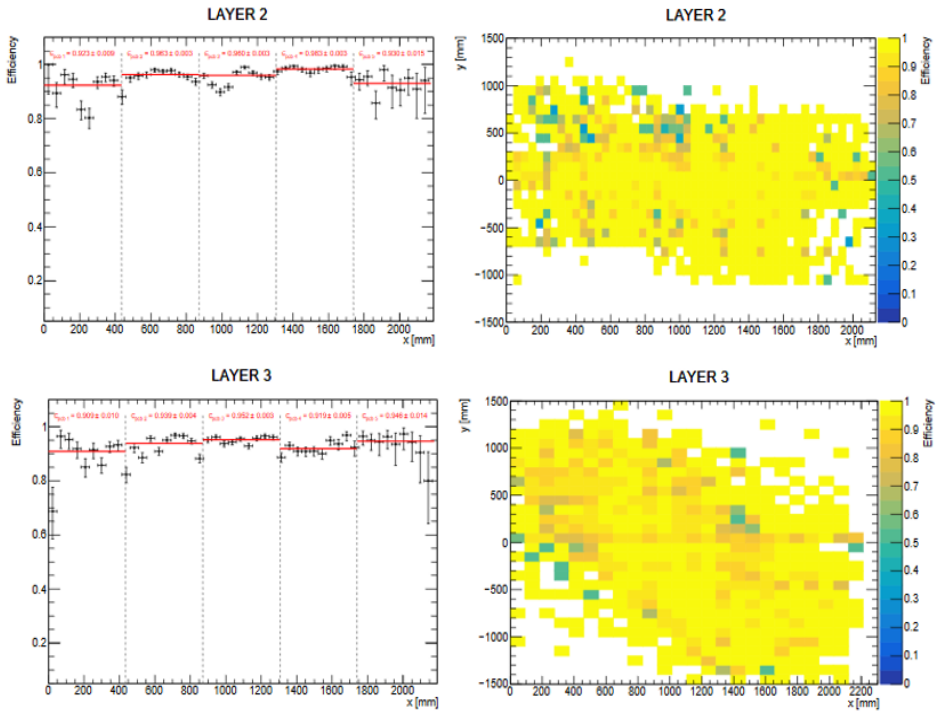


Figure 3.36: Tracking efficiency maps, realised by the software implemented at INFN Laboratory of Frascati.

3.6.2 Validation tests of Micromegas Double Wedges

The measurements described in Section 3.6.1 are repeated on the Micromegas Double Wedges (DW), and the relative set up is reported in Appendix D. The validation criteria on the HV stability are not subjected to modifications. The tolerances and set limits are the same in Section 3.6.1 because the validation of a Double Wedge means to validate two pairs of Large (or Small size) Modules of type 1 and 2, operating at the same time.
In terms of tracking efficiency, the number of layers is eight and each strip is independently read by a VMM [89] FE electronic channel during the validation tests of Double Wedges. Therefore, the cluster and track definitions change with respect to the previous section. Moreover, two different groups of the ATLAS collaboration worked on these two classes of tests (HV and data acquisition).

HV stability and current-time trends

The experience has shown that the results of the first validation of a single Module at the construction sites or CERN (in Section 3.6.1) are preliminary. After that, the HV map may change in time in positive and negative ways due to trapped impurities, residual humidity, or because the operation period has not lasted enough to monitor all the variations in time up to the stable behaviour of an HV sector.
To minimise the second case, the average time duration of a double wedge lasted about 10 days with 4-5 days of continued HV operation. The HV map also changes because an HV channel supplies the two sides of a PCB, and the minimum value between the sides defines the amplification voltage of the entire PCB. This HV supply configuration is also called Half Granularity (HG) scheme.
Therefore the validation procedure foresees the recording of a new HV map of each HV sector, plugging one side (between left and right) at the time and an HV map of the HG scheme. During the validation of the double wedges, different current-time trends are observed. The occurrence frequency is roughly estimated based on the average number of events observed during the validation of a DW. They are shortly described and enlisted from the most to less frequent trends below:

1. **unstable baselines**: the current rapidly varies in such a way that the monitored trend appears like a sequence of capacitor discharges (filter C of power supplier).
   This trend is observed in every Module type (Large/Small and Modules 1/2). Its occurrence frequency (%) is about 23%, considering ten interested HG sectors per Small DW and that the average frequency per DW is quite double in the Large DWs.
   This trend is not worrying if the average current is smaller than 40 nA.
   In the worst cases, possible parasitic components make an RC series charging and discharging in the current-time trend with a maximum current of O(100-1000 nA), as shown on the left side of Figure 3.37. Spikes or fast current variations often superimpose making the cumulative trend more irregular. Then, they make difficult to recognise if this trend corresponds to impurities that can disappear by burning, like the trend on the right side of Figure 3.37, or floating metallic defects. Empirically, the overall trend of a not metallic defect burning decreases in the first tens minutes, and the current has to be stable after. If another similar event with equal or greater maximum current occurs, it means that the defect is floating or starting to behave as a dielectric. Therefore, it charges and discharges in a limited time.

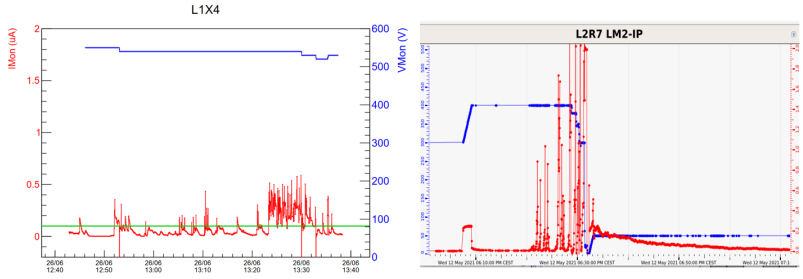


Figure 3.37: Examples of unstable baselines. The current-time trend (red) and voltage-time (blue) of an HV sector of DW C15 (left) and DW C13 (right). The threshold of 100 nA is indicated in green. Possible parasitic components make an RC series charging and discharging in the current-time trend (left). The final worsening may correspond to the burning of the impurity behaved as parasitic component. The current-time trend (right) reasonably corresponds to the burning of some impurities.

Some sectors have been treated with pure Ar rising their voltage with a low ramp up at a fixed current to perform the impurity burning more safely.

2. **fast spikes**: their effects are the most unpredictable according to their causes. In some cases, a positive result is observed after an intense spark has occurred because it has released some accumulated charge in the capacitor components. The HV stability that can improve by using this treatment is usually when a bunch of sparks with amplitude at the limit of $I0$ appears a few minutes after a set HV change.
   Other times, a worsened increase in the spark rate is observed after a sudden huge spike, which can also be individuated as an inflexion point in the voltage-time trend.
   Their occurrence frequency (%) is about 13%, with an average value of eight interested HG sectors per DW.
   Spikes with an amplitude (rate) smaller than 1.5 μA (6 sparks/min) are tolerated for a sector at 570 V. While the voltage is decreased when the spike amplitude is close to the limit of 2 μA with a constant rate of 1-2 sparks/min in 24-48 hours.

3. **not-ohmic offsets**: floating defects (like flying mesh wire) make a shortcut in the amplification gap when the voltage overcomes a certain threshold. Even if they may be considered part of the unstable baseline category, their signature is an average current that does not linearly increase with the voltage. They are always dealt with by decreasing the voltage. Their occurrence frequency (%) is around 6%. They have been about three times more frequent in the Large DWs than in the Small DWs.

4. **slow diverging currents**: some Large Module 1 sectors have shown a very slow current increase that became clearly visible after a few days, and they have been successively operated by decreasing the HV. Its cause is not clear yet. It is often anticipated by an isolated spike, like in Figure 3.38. The current slowly increases, or it becomes more unstable up to be a clear visible variation. They are dealt with by decreasing the voltage. These trends mainly involve the sectors with PCB size smaller than 4. Their occurrence frequency is about 4%.

Figure 3.38: Current-time trend of a HV sector of LM1-04 in Grafana dashboard, acquired during the validation of the quadruplet. Example of slow diverging current, the current decreased after a day within the tolerances.

5. **Faulty soldering of external filter capacitor**: unstable current-time trends have been caused sometimes by faulty soldering of the external filter capacitor, which created parasitic resistance toward the detector mass. They are detected by resistance lower than a few MΩ. They have been solved by replacing the external RC filter.
   The occurrence frequency (%) is < 2%, considering an average value of two RC filters per DW.

6. **time-limited step current increases**: some time-limited O(1 uA) current steps, like sporadic shortcuts, have been rarely observed. The causes may be different, like trapped water drops or damaged pillars. They are dealt with decreasing the voltage. Isolated variations, like those in Figure 3.39 (left), are not worrying, and a voltage decrease is not required. While Figure 3.39 (right) shows one of the worst cases, and they occurred very few times in the Large Module 2 sectors. Their frequency is < 1%.

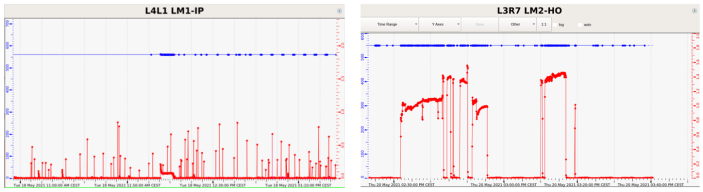


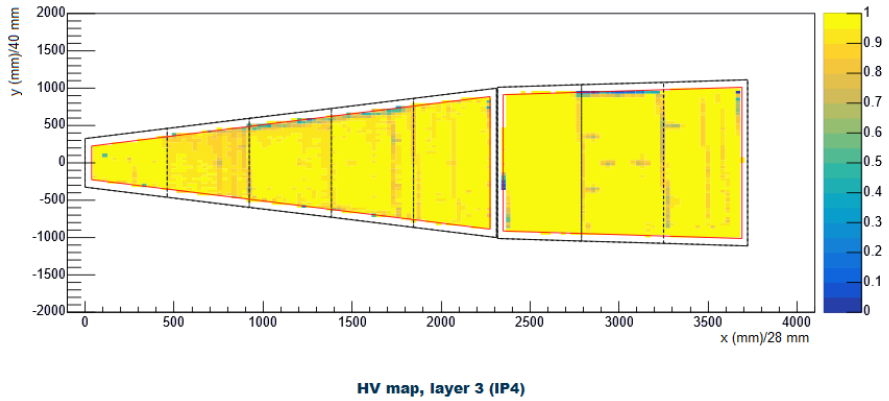
Figure 3.39: Examples of time-limited step current increase. The current-time trend (red) and voltage-time (blue) of a Large Module 1 sector (left) and a Large Module 2 sector (right). The current variation (left) is not worrying.

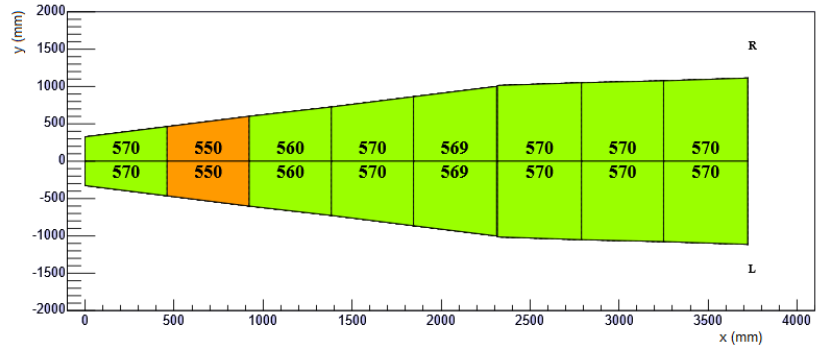
Ohmic offsets (proportional to the voltage) have been observed, too. In these cases, the region of the resistive plane involved in the current spreading is clearly visible from the efficiency maps as an inefficient spot, as shown later in the bottom part of Figure 3.40.

Tracking efficiency

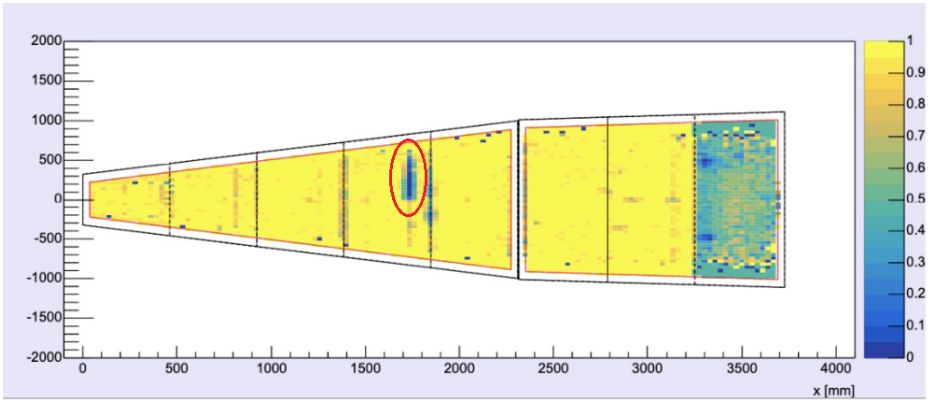
The measurement procedure for the tracking efficiency is the same as the one reported in Section 3.6.1. It only differs by the number of tracking layers and readout granularity, as already discussed at the beginnings of Section 3.6.2. Then, excluding the interested layer, tracks of cosmic rays are

reconstructed in at least six layers with the additional requirements: a maximum number of clusters in an event less than 30, a maximum number of clusters in a layer less than 10, and a cut on the χ^2 equal to 10. Their projections on the excluded layer individuate the fiducial sections where the algorithm searches the possible clusters. For the DW tests, the widths of these fiducial sections have to be less than 10 mm in both η and $stereo$-views.
Again, the efficiency is measured as the ratio between the number of clusters in the fiducial sections and the number of track projections. The requirement for the tracking efficiency to pass the validation is that each tracking layer has an average efficiency larger than 80%. An example of the 2D efficiency map is shown in Figure 3.40 (top), relative to a $stereo$ layer of Double Wedge C01. The efficiency is higher than 95% in most parts of the layer except for the passivated edges, the region close to the gas interconnections and the PCB2 area because its stable HV is 550 V with Ar:CO_2 mixture, as visible in Figure 3.40 (middle).
The efficiency is a function of the amplification voltage (see Figure 3.17), and the nominal efficiency corresponding to 550 V is smaller than 570 V. This is confirmed by the average measured efficiency of $\sim$ 92% in the PCB2 area, and it is also visible from the efficiency as a function of the strip position in Figure 3.41.





(a) top and middle



(b) bottom

Figure 3.40: 2D map of tracking efficiency of a *stereo*-view layer (top Figure), where the inefficient regions close to the edges correspond to the passivated area, and its HV map (middle Figure). Example of inefficient spot (in red oval in bottom Figure) corresponding to an ohmic offset in the current trend (Section 3.6.2).

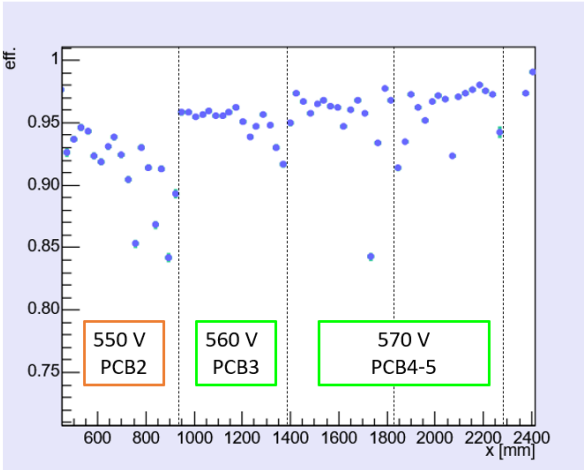


Figure 3.41: Tracking efficiency as a function of the strip position for PCB2, PCB3 and PCB4 of the stereo layer in Figure 3.40.

The average efficiency over all the DW layers can quantify some observations on the efficiency maps of DWs in the New Small Wheel subsystems. Figure 3.42 shows that the mean efficiency over all the layers is, on average, ~ 90%. A Small DW generally has a higher mean efficiency than a Large DW, with the Double Wedge A11 at the limit of the acceptance. This trend is discussed in correlation with the HV working points of the readout MM boards in the next Section 3.7.

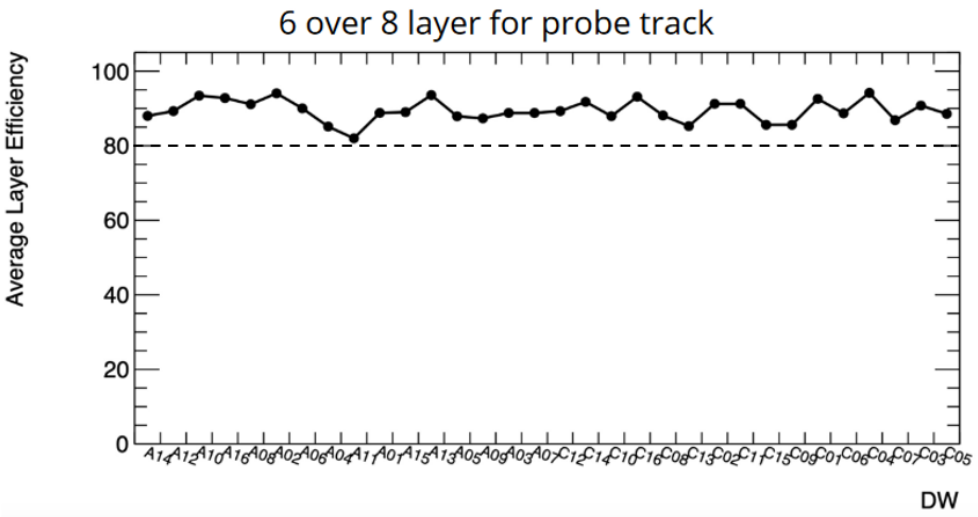
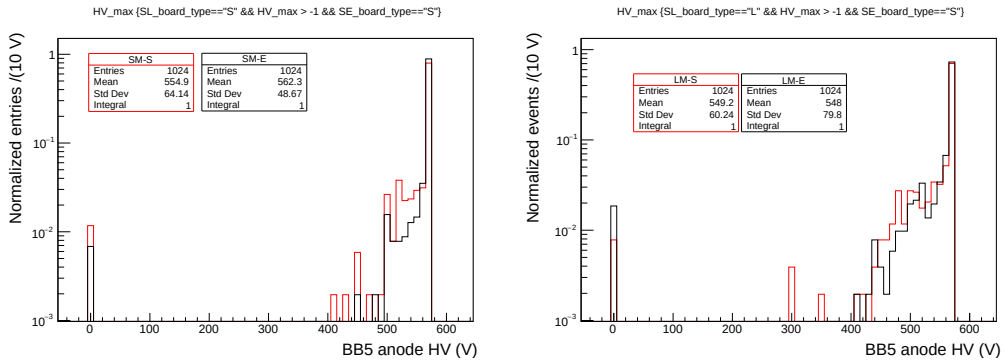


Figure 3.42: Average efficiency over all eight layers as a function of the DW identification name. The order is chronological on the axis of DW names. Even (Odd) number refers to a Small (Large) DW.

3.7 Validated HV setting of the Micromegas Double Wedges

The procedures described in Section 3.6.2 are repeated on the 32 Micromegas Double Wedges for wheels A and C. In Figures 3.43 (a) and (b), the anode voltage[7] HV histograms show that the Large Modules have more HV sector at voltages lower than 570 V than Small Modules, as anticipated in previous Section 3.6.2. The percentage of sectors with an $HV \leq 550$ V is 22%(13%) for the case of the Large (Small) DWs. While the histograms do not systematically differ between the two strip orientations: η and $stereo$ views.

In Figure 3.44, the voltages are reported in two 2D histograms to visualise the HV distribution per



(a) Small DWs (b) Large DWs

Figure 3.43: Micromegas anode HV histogram for the η ($stereo$) view in black (red) of the Small DWs (a) and Large DWs (b). The histograms are normalised by the number of all HV sectors equal to 1024 to include the temporarily-unplugged sectors in the HG configuration.

each side of every DW. A histogram for each side takes into account the cases in which only one HV sector of the PCB is supplied.

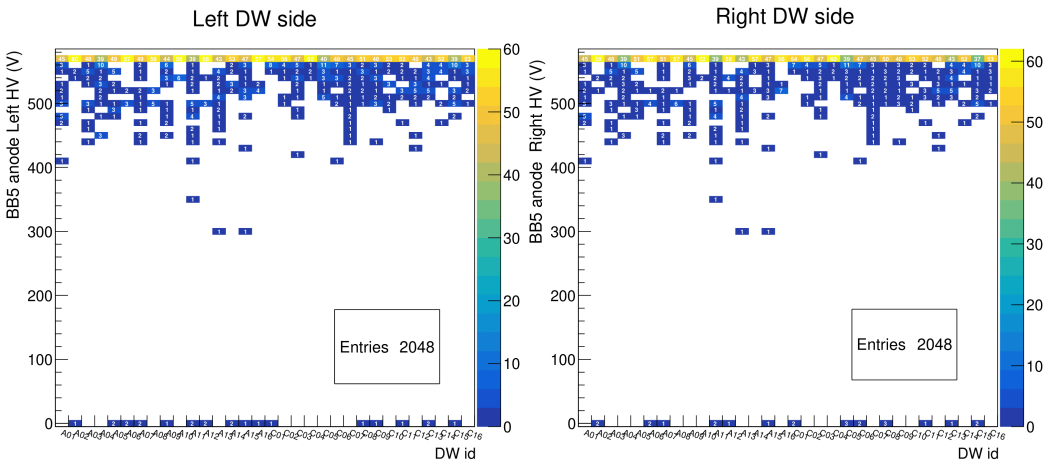
During the tests with Ar:CO_2 (93:7)%, the geometrical acceptance has been sometimes overshadowed by the requirement to reach the maximum overall tracking efficiency on a layer by unplugging one of the two HV sectors of a PCB.

The total efficiency is higher with only one sector at voltages greater than 540 V rather than two sectors at a working point lower than 500 V. In these cases, the efficiency times 0.5 of only a sector is greater than the efficiency of the two PCB sectors at the minimum stable voltage, referring to the red curve in Figure 3.17. It has been required that only an HV sector can be removed on among the PCBs of same size over the different layers to assure at least seven tracking layers.

Every time that a DW with an unplugged HV sector has also been validated with the Ar:CO_2:iC_4H_{10} (93:5:2)% mixture, the disconnected sector has been plugged again in parallel to the HV sector of the other side of PCB at amplification greater than 480 V, corresponding to a nominal efficiency greater than 84% (blue curve in Figure 3.17). Based on the results with Ar:CO_2:iC_4H_{10} (93:5:2)% mixture and on the fact that the HV picture validated in BB5 is not the final, it has been decided to supply all the HV sectors during the successive tests on the HV stability when the Micromegas DWs have been

[7] The positive anode voltage is applied to the resistive strips, as reported in Section 3.4.

integrated with the sTGC wedges on the mechanical wheel structure.



(a) Left (b) Right

Figure 3.44: 2D histograms of the amplification voltages (on the y-axis) for each DW (its ID on the x-axis).

3.8 Amplification voltage dependence on readout Micromegas board parameters

The anode HV is also referred to as amplification voltage (see Section 3.3.1) in the following studies about the possible correlations between the anode HV dependence and the PCB parameters since they are characteristic of the resistive strip Micromegas technology.
In these studies, only the sectors with an amplification voltage greater than 430 V have been considered since smaller values are reasonably due to defects in detector components rather than due to the value of the readout Micromegas board parameters.
When the amplification voltages are rearranged as a function of the readout Micromegas board type (in Figure 3.45), the readout Micromegas board types with the lowest mean amplification voltages are L-Eta/Stereo 1 and 2.
The pillar height defines the intensity of the amplification electric field. Therefore, it is interesting to look at the 2D histogram between the amplification voltage values of Section 3.7 and the mean pillar height of the readout Micromegas boards in Figure 3.46. As visible in Figure 3.46, there is not a significant correlation between the amplification voltage and mean pillar height of the readout Micromegas boards. While a stronger correlation is observed between the amplification voltage and the mean resistivity of the readout Micromegas boards in Figure 3.47. From the intersection between the profile of the 2D histogram (of the mean resistivity and the amplification voltage) in Figure 3.47 and the 560 V line, the minimum of the mean resistivity is about 0.7 MΩ/□ corresponding on average to a stable amplification voltage ≥ 560 V.

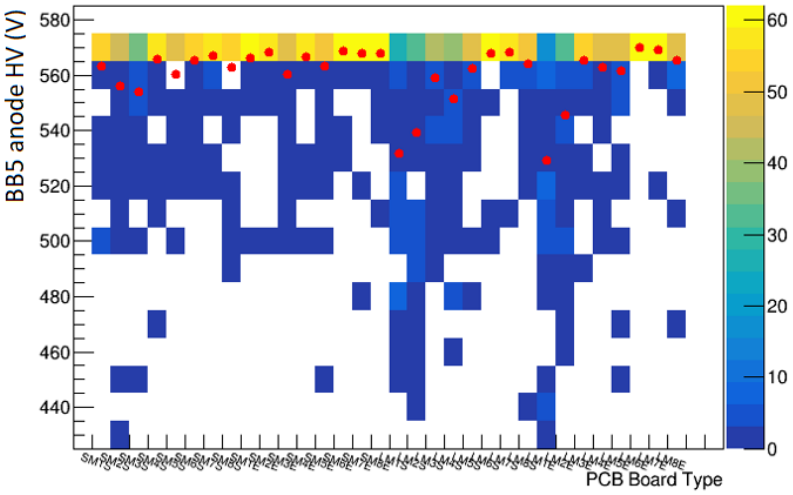


Figure 3.45: 2D histograms of the amplification voltages (on the y-axis) for each anode board type (on the x-axis). The red markers represent the mean values of the amplification distributions with minimum value 430 V and fixed anode board type.

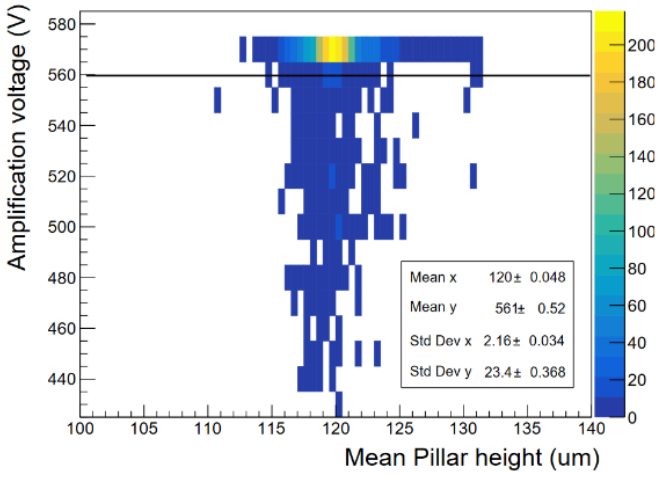


Figure 3.46: 2D histogram between mean pillar height of the readout Micromegas board and amplification voltage. Bin width are 1 μm and 10 V, respectively. The black line individuates the voltage 560 V where the detector is fully efficient.

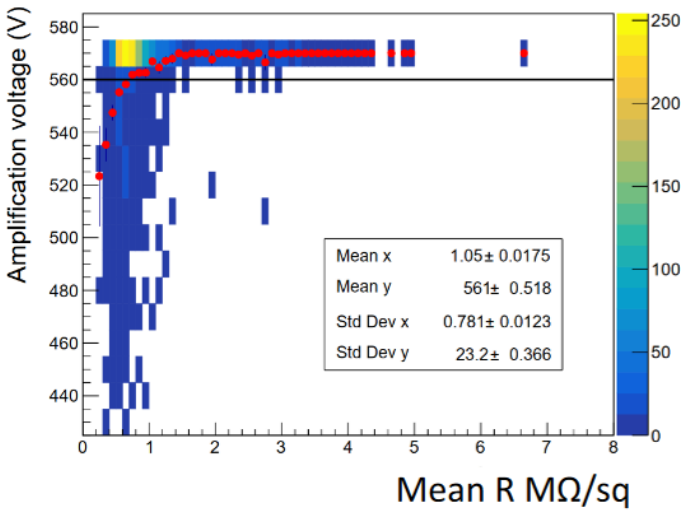


Figure 3.47: 2D histogram between mean resistivity R of the readout Micromegas board and the amplification voltage. The bin widths are 0.1 MΩ/□ and 10 V, respectively. The black line individuates the voltage 560 V where the detector is fully efficient. The red dot marks represent the mean values of the amplification voltages in the mean resistivity bins.

3.9 Front-End electronic of ATLAS Micromegas detector

The FE electronic of ATLAS Micromegas is based on the VMM3 ASIC. It is composed of 64 linear FE channels. The analogic line of each channel consists of a low-noise charge amplifier (CA) and a shaper at the output of the amplifier, with a baseline stabilizer. The gain of the CA and peaking time of shaper are settable. They are set respectively to 9 mV/fC and 25 ns for the ATLAS Micromegas detectors. The amplified and shaped signal is sent to an ADC (and in a TAC) to measure its peak amplitude (and its arrival time) and in a discriminator to build a 0-level logic and to provide the Time-over-Threshold (ToT) and Threshold-to-Peak (TtP) digital outputs.
The VMM channels share the bias circuits, a temperature sensor, a test-pulse generator, a mixed-signal multiplexer, the control logic and two 10-bit DAC for adjusting the test pulse and the threshold voltage of the discriminators. The discriminator of the VMM3 is common to the VMM1, and it is detailed in Reference [89]. In particular, it has two functions. The sub-hysteresis function allows the discrimination of smaller signals than the hysteresis of the comparator circuit to keep a suitable dynamic range at high comparator speed. The second function is the neighbour processing logic allows the signal processing of the closest strips to the one on which the signal exceeds the discriminator threshold, even if the neighbour signals do not overcome the thresholds.
In the VMM3, the discriminator threshold can also be individually adjusted with an individual channel 5-bit trimming DAC, decreasing the threshold in the [0,30] mV range to reduce the threshold spread among the VMM channels, as visible in Figure 3.48. The threshold of a readout channel is evaluated

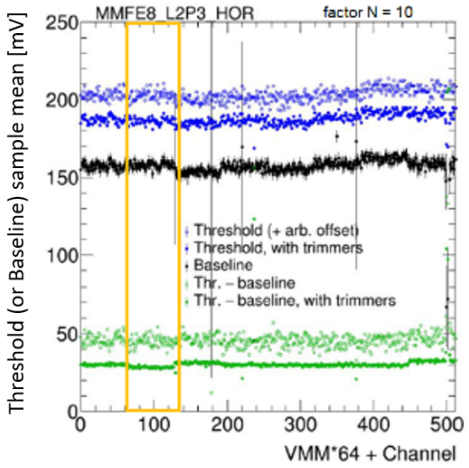


Figure 3.48: Example of the baseline (black), the threshold (blue) and the effective threshold (green) as a function of readout channel indices of a group of 8 VMM, mounted on a same front-end card (MMFE8). The empty (full) markers of the thresholds are not corrected (are corrected) by the trimmer.

at the VMM-level (group of 64 consecutive strips). According to Equation 3.5, its definition is built on the median ($\text{median}_{baseline}$) and the rms ($\text{RMS}_{baseline}$) of the VMM-level baseline distribution. Depending on the strip length, the $\text{RMS}_{baseline}$ varies from $\sim$ 1 000 ENC[8] to $\sim$ 3 000 ENC for DW

[8] The ENC is the acronym of Equivalent Noise Charge. It expresses electronic noise in terms of number of electrons. Considering the used CA gain, 1000 ENC is about 1.44 mV, useful to compare with thresholds that is expressed in mV.

A07, and it is multiplied by N, the threshold factor (settable parameter).

$$\text{threshold}_{\text{VMM channel}} = \text{N} \times \text{RMS}_{baseline} + \text{median}_{baseline} \tag{3.5}$$

The effective threshold is defined as the difference between the threshold with an applied trimmer (Trim) and the baseline, according to Equation 3.6.

$$\text{threshold}^{eff}_{\text{VMM channel}} = \text{threshold}_{\text{VMM channel}} + \text{Trim} - \text{baseline}_{\text{VMM channel}} \tag{3.6}$$

The channels also have in common the Address in Real Time (ART), which reports the flag and the address of the first channel above-threshold, and it self-resets in 40 ns.

3.10 Noise dedicated studies on Micromegas Double Wedge A07

Some studies on the VMM ASIC FE electronics have been dedicated to investigating and improving the noise level of the ATLAS Micromegas detectors when they are mounted on the wheel mechanical structures, together with the sTGCs. A part of them has been performed on the alone ATLAS Micromegas DWs (before the integration with sTGCs) to investigate the noise and the detector performance when larger thresholds than the actual nominal values are applied.
Each ATLAS readout plane has 8192 readout strips with different lengths from 426.7 mm to 1990.0 mm for the Large modules (from 284.0 to 1579.6 for Small modules).
During the study on Micromegas A07 sector before its integration on the wheel, the N factor in Equation 3.5 has been varied among $\{8, 9, 10, 11, 12, 15\}$ with a fixed VMM gain (equal to 9 mV/fC). The actual nominal threshold factor is 9. If the performances are still acceptable when tighter thresholds are applied, a higher N factor can be set to more efficiently suppress possible noise coming from interferences with other systems of the NSW. Between factors 9 and 12, the average effective threshold increases by $\sim$ 10 mV (see the left side of Figure 3.49). The average strip occupancy has confirmed that factor 9 is the minimum applicable factor to have an acceptable uniformity of the occupancy on the plane. Indeed, factor 8 can not suppress the noise on the longest strips (in the right side of Figure 3.49). In correspondence with factor 15, a loss of events on the shortest strips is visible on the right side of Figure 3.49.

3.11 Tracking efficiency as a function of the effective threshold

During this study, the cosmic rays have been used as the radiation source and a system of scintillator counters as the external trigger system (in Appendix D). The study has been performed with the HV sectors at the nominal (570 V) voltage and with the Ar:CO_2 (93%:7%) gas mixture. The tracks have been reconstructed according to the selection as reported in the caption of Figure 3.50. The tracking efficiency has been evaluated at different average levels (VMM = group of 64 strips, MMFE8 = group of 8 VMM, and PCB = group of 2 MMFE8). The efficiency at the PCB level decreases by $\sim$3% from 20 to 60 mV effective threshold, and it is always above 80%. No significant occupancy and efficiency losses are observed by increasing the thresholds of $\sim$ 10 mV with respect to the nominal values. Specifically, the average PCB efficiency drop at the nominal voltage is around 2% from the

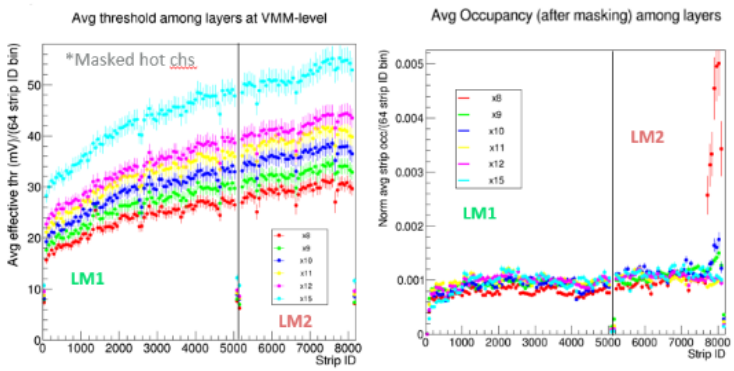


Figure 3.49: Average threshold at VMM-level on the A07 layers for N = {8, 9, 10, 11, 12, 15} (left). Average occupancy on the A07 layers (right).

nominal factor 9 to factor 12. Meanwhile, some hardware interventions have been made on the noise causes improving the ground of the NSW sectors to the point that it has not been necessary to apply higher thresholds on the electronic readout thresholds.

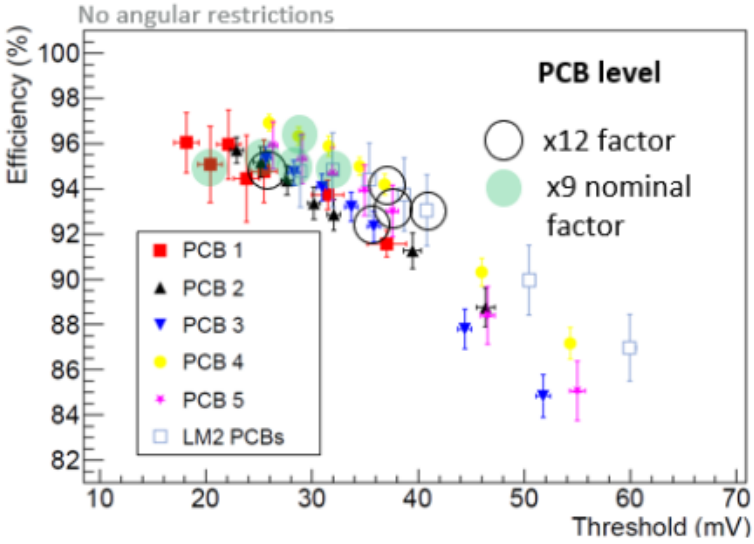


Figure 3.50: Efficiency at PCB-level as a function of the average effective threshold at PCB-level on the A07 layers. Full green dots are relative to the dataset with factor N = 9 and empty black dots are relative to N = 12. The tracking requirements are: no restriction on the cluster multiplicity; cut (> 1.2 fC) on the minimum charge of the strip with the maximum charge in the cluster (track projection on the readout plane) and a cut (> 0.4 fC) on the minimum charge of the other strips of the cluster. Last requests are a maximum number of clusters in an event less than 30 and a maximum number of clusters in a layer less than 10.

Conclusion

The thesis has concerned two different topics of the ATLAS experiment. The first topic is related to the study of Higgs boson production in association with W boson with subsequent $H \rightarrow WW^*$ decay and three leptons in the final state with at least one flavour and opposite charge sign pair ($3l$-"Z-dominated" channel). This study aims to determine the inclusive cross-section and its corresponding measurement in the scope of the common LHC framework *Simplified Template Cross-Section* ($STXS$). The second topic is devoted to the validation of the resistive Micromegas detectors that have been installed in the innermost End-Cap regions of the ATLAS Muon Spectrometer.
In the first part of the thesis, the entire procedure of the signal strength parameter μ measurement is described for the $3l$-"Z-dominated" channel, where μ is defined as the ratio between the observed signal yields and its SM prediction. The procedure can be divided into: a unique event pre-selection for all the sub-channels of the combined $VH \rightarrow VWW^*$ analysis (where the associated V boson can be a W or Z boson), a customised multivariate analysis for every sub-channel, and a common statistical analysis performed on the single sub-channel and their combination.
In particular, my contribution has concerned the multivariate analysis of $3l$-"Z-dominated" channel. A binary neural network classifier (ANN) has been optimised and validated to discriminate the WH signal process in the $3l$-"Z-dominated" signal region against the dominant irreducible backgrounds. It has also shown a satisfying background rejection on the reducible component.
The ANN output has been exploited as discriminant on which the profile likelihood fit is applied to estimate the expected signal strength parameter μ^{exp}_{Zdom} and the significance Z of $3l$-"Z-dominated" channel together with the normalisation factors for the main background, $WH/W\gamma^*$ process. The results have been obtained using the Asimov dataset in the signal region and data in the control regions, and they are:

$$\mu^{exp}_{Zdom} = 1.00^{0.78}_{-0.71}$$
$$Z = 1.45\sigma$$

The μ^{exp}_{Zdom} evaluation is principally affected by the statistical uncertainties rather than the systematic ones in which the experimental and theoretical contributions are comparable. The statistical term contributes to $\sim$ 87% of total symmetric uncertainty. The fit has been repeated including all the $VH \rightarrow VWW^*$ sub-channels to estimate the combined signal strength parameters μ^{exp}_{WH} and μ^{exp}_{ZH}, one for each weak vector V boson:

$$\mu^{exp}_{WH} = 1.00^{0.31}_{-0.29}$$
$$\mu^{exp}_{ZH} = 1.00^{0.46}_{-0.39}$$

They are associated with an expected significance of 5.13σ. The signal strength parameters are principally affected by the statistical uncertainties in the combined fit, especially for the evaluation μ_{ZH}^{exp}. The reported uncertainties include both the statistical and systematic contributions.
The $STXS$ measurements are in progress, and they exploit multivariate discriminators, too. Differently from the inclusive analysis, $STXS$ provides cross-section measurements in optimised bins of the transverse momentum p_T^W of the associated W boson in the Higgs production WH process. An interesting study has been carried out to find the best p_T^W proxy between the output of a regression Neural Network p_T^R (optimised to tag p_T^W by another analysis group) and the variable ξ, based on a neural network multi-classifier with p_T^R in input. Their corresponding expected signal strength parameters have been compared, and no significant improvements have been observed when ξ replaces p_T^R. Then, p_T^R is chosen as the final proxy variable.
The presented work has been preparatory to the inclusive cross-section and $STXS$ measurements of the single $WH \rightarrow W(WW^*)$ and combined $VH \rightarrow V(WW^*)$ in which the Asimov data will be replaced with real data. The $VH \rightarrow V(WW^*)$ analysis has started to replace the Asimov data with real data in one of its sub-channel. Then, the observed evaluations of the signal strength parameters will be public soon.

In the second part of this thesis, my contributions to the results of the quality controls on the ATLAS readout Micromegas (MM) boards and validation tests of the MM detector high voltage (HV) stability have been presented. The upgraded detector-equipped structures, New Small Wheels, are now present in the innermost End-Cap region of the ATLAS Muon Spectrometer.
The quality of about 3000 ATLAS readout MM boards has been controlled through several parameters. The design specifics have been respected satisfyingly by all the 2048 boards that have been integrated into the final detectors.
Before the integration in New Small Wheels, the MM detectors have been validated in terms of the high voltage (or amplification) stability and the tracking efficiency that has been measured by applying the stable high voltage configuration with the Ar:CO_2(93%:7%) gas mixture. With this mixture, the HV working point (or amplification voltage) of the ATLAS MM detectors is 570 V with a nominal $\sim$ 98% efficiency. Most HV sectors have been validated with 570 V. The amplification voltage of 560 V is still an acceptable working point that provides detector performances close to the nominal ones. The percentage of the HV sectors that have been stable at an amplification voltage lower than 560 V is 22% (13%) for the case of Large (Small) Micromegas sectors. Then, several studies have been carried out to investigate the causes and solutions to the observed HV instabilities of these sectors. In particular, the presented study has focused on the readout MM board parameters that principally affect the amplification voltages. They are the pillar height and surface resistivity of the ATLAS readout MM board. The correlation studies have shown no significant correlation between the mean pillar heights of the readout MM boards and the validated amplification voltages lower than 560 V.
Moreover, as the pillar height and electric field intensity are inversely proportional in the first approximation, the requirements for their uniformities are the same. The pillar height uniformity respects the imposed tolerances. Defined as the ratio between the RMS and mean value of the pillar height distribution of a readout MM board, each uniformity measurement is less than 5% with a mean value of $\sim$ 2% per every readout MM board type.
It is observed that the surface resistivity of the ATLAS readout MM board is the main parameter that influences the validated amplification voltages and the high voltage detector stability. The results of

the correlation studies have shown that a mean resistivity greater than 0.7 MΩ/□ statistically assures a stable high voltage behaviour at voltages equal to or greater than 560 V.
Studies on the detector performances have also been carried out with a more quenching than Ar:CO_2(93%:7%) gas mixture to improve the high voltage stability of all the MM detectors. It is observed that a 2% addition of isobutane vapour (replacing 2% CO_2) in the gaseous mixture already successfully stabilises the high voltage behaviour of the less robust sectors tested with both the gas mixture. The results of these studies have supported the decision to use Ar:CO_2:iC_4H_{10}(93%:5%:2%) mixture during Run 3.

Acknowledgments

I would like to express my special thanks to all people that helped and gave precious advises to successfully carry out the PhD activities. Particular acknowledgements go to the VH analysis group, Roma Tre, Frascati and Naples INFN sections, GDD group, 188 and BB5 teams, and CERN EP-ADE group.

Bibliography

[1] ATLAS Coll., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on pp. 7, 9, 11, 13, 18–23).

[2] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001 (cit. on pp. 7, 9).

[3] ATLAS Coll., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: 1207.7214 [hep-ex] (cit. on pp. 7, 9, 23).

[4] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: 1207.7235 [hep-ex] (cit. on pp. 7, 9, 23).

[5] ATLAS Coll., *Measurement of the production cross section for a Higgs boson in association with a vector boson in the $H \to WW^* \to \ell\nu\ell\nu$ channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **798** (2019) 134949, arXiv: 1903.10052 [hep-ex] (cit. on p. 7).

[6] ATLAS Coll., *ATLAS Letter of Intent for a General-Purpose pp Experiment at the Large Hadronic Collider at CERN*, (1992), URL: http://cds.cern.ch/record/291061 (cit. on p. 9).

[7] CMS Coll., *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004 (cit. on p. 9).

[8] ALICE Coll., *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002 (cit. on p. 9).

[9] LHCb Coll., *The LHCb Detector at the LHC*, JINST **3** (2008) S08005 (cit. on p. 9).

[10] CERN website, *HL-LHC Industry*, URL: https://project-hl-lhc-industry.web.cern.ch/content/project-schedule (cit. on p. 10).

[11] ATLAS Coll., *ATLAS Insertable B-Layer Technical Design Report*, (2010), URL: https://cds.cern.ch/record/1291633 (cit. on p. 14).

[12] ATLAS Coll., *Production and integration of the ATLAS Insertable B-Layer*, JINST **13** (2018) T05008 (cit. on p. 14).

[13] ATLAS Coll., *ATLAS Inner Detector Technical Design Report*, (1997), URL: https://cds.cern.ch/record/331063/ (cit. on pp. 14, 16).

[14] M. Abbrescia et al., *Study of long-term performance of CMS RPC under irradiation at the CERN GIF*, NIM Research Sec. A **533** (2004) 102 (cit. on p. 21).

[15] ATLAS Coll.,
Performance of the ATLAS RPC detector and Level-1 muon barrel trigger at $\sqrt{s} = 13$ TeV,
JINST **16** (2021) P07029, arXiv: 2103.01029 (cit. on p. 21).

[16] R. Blair et al., *The ATLAS High Level Trigger Region of Interest Builder*,
JINST **3** (2008) P04001 (cit. on p. 23).

[17] T. W. B. Kibble, *Spontaneous symmetry breaking in gauge theories*,
Philosophical transactions of the Royal Society A (2015) (cit. on p. 23).

[18] F. Gianotti et T.S. Virdee, *The discovery and measurements of a Higgs boson*,
Philosophical transactions of the Royal Society A (2015) (cit. on p. 23).

[19] T. Hambye and K. Riesselmann,
Matching conditions and Higgs boson mass upper bounds reexamined,
Phys. Rev. D **55** (1997) 7255 (cit. on p. 24).

[20] L. H. W. Group, *CERN Yellow Reports 4*, URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHWG?redirectedfrom=LHCPhysics.LHCHXSWG#SM_Higgs
(cit. on pp. 25, 26).

[21] ATLAS and CMS Coll., *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*,
Phys. Rev. Lett. **114** (2015) 191803, arXiv: 1503.07589 [hep-ex] (cit. on pp. 24, 26).

[22] LHC Higgs Cross Section Working Group, *Higgs cross sections for HL-LHC and HE-LHC*,
URL:
https://twiki.cern.ch/twiki/bin/view/LHCPhysics/HiggsEuropeanStrategy
(cit. on p. 26).

[23] ATLAS Coll., *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector*,
Phys. Lett. B **786** (2018) 59, arXiv: 1808.08238 (cit. on pp. 25–27).

[24] ATLAS Coll.,
ATLAS Sensitivity Prospects for Higgs Boson Production at the LHC Running at 7 TeV,
(2010), URL: https://cdsweb.cern.ch/record/1278455 (cit. on p. 26).

[25] ATLAS Coll., *Measurements of gluon–gluon fusion and vector-boson fusion Higgs boson production cross-sections in the $H \to WW^* \to e\nu\mu\nu$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **789** (2019) 508, arXiv: 1808.09054
(cit. on p. 27).

[26] ATLAS Coll., *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018) 173, arXiv: 1806.00425
(cit. on p. 27).

[27] ATLAS Coll., *Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
Phys. Rev. D **97** (2018) 072016, arXiv: 1712.08895 (cit. on p. 27).

[28] ATLAS Coll., *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, Phys. Rev. D **97** (2018) 072003, arXiv: 1712.08891
(cit. on pp. 27, 28).

[29] ATLAS Coll., *Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Rev. D **98** (2018) 052005, arXiv: 1802.04146 (cit. on p. 27).

[30] ATLAS Coll., *A combination of measurements of Higgs boson production and decay using up to 139 fb^{-1} of proton-proton collision data at $\sqrt{s}$= 13 TeV collected with the ATLAS experiment*, submitted to ICHEP (2020), URL: http://cds.cern.ch/record/2725733 (cit. on p. 27).

[31] ATLAS Coll., *Measurements of Higgs bosons decaying to bottom quarks from vector boson fusion production with the ATLAS experiment at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **81** (2020) 537, arXiv: 2011.08280 (cit. on p. 28).

[32] ATLAS Coll., *Measurements of WH and ZH production in the $H \to b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 178, arXiv: 2007.02873 (cit. on p. 28).

[33] ATLAS Coll., *Measurement of the Higgs boson decaying into bquarks in associated production with a top-quark pair in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, submitted to JHEP (2021), arXiv: 2111.06712 (cit. on p. 28).

[34] ATLAS Coll., *Measurements of gluon fusion and vector-boson-fusion production of the Higgs boson in $H \to WW^* \to e\nu\mu\nu$ decays using pp collisions at $\sqrt{s}$=13 TeV with the ATLAS detector*, (2021), URL: https://cds.cern.ch/record/2759651 (cit. on p. 28).

[35] ATLAS Coll., *Measurement of the production cross section for a Higgs boson in association with a vector boson in the $H \to WW^* \to l\nu l\nu$ channel in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Lett. B **798** (2019) 134949 (cit. on pp. 28, 29).

[36] ATLAS Coll., *Measurements of Higgs boson production cross-section in the $H \to \tau^+\tau^-$ decays channel in pp collisions at $\sqrt{s}$=13 TeV with the ATLAS detector*, submitted to JHEP (2021), arXiv: 2201.08269 (cit. on p. 28).

[37] ATLAS Coll., *Measurement of the properties of Higgs boson production at $\sqrt{s}$ = 13 TeV in the $H \to \gamma\gamma$ channel using 139 fb^{-1} of pp collision data with the ATLAS experiment*, submitted to ICHEP (2020), URL: https://cds.cern.ch/record/2725727 (cit. on p. 28).

[38] ATLAS Coll., *Higgs boson production cross-section measurements and their EFT interpretation in the 4l decay channel at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Eur. Phys. J. C **80** (2020) 957 (cit. on p. 28).

[39] ATLAS Coll., *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15** (2020) P04003 (cit. on pp. 30, 31).

[40] B. Salvachua, *Overview of proton-proton physics during Run 2*, (2019), URL: https://cds.cern.ch/record/2750272 (cit. on p. 31).

[41] J. Wenninger, *Operation and Configuration of the LHC in Run 2*, (2019), URL: https://cds.cern.ch/record/2668326 (cit. on p. 32).

[42] ATLAS Coll., *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s}$ = 13 *TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv: 1603.05598 [hep-ex] (cit. on p. 33).

[43] ATLAS Coll., *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 (cit. on pp. 34, 35).

[44] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph] (cit. on p. 35).

[45] ATLAS Coll., *Tagging and suppression of pileup jets*, ATL-PHYS-PUB-2014-001, 2014, URL: https://cds.cern.ch/record/1643929 (cit. on pp. 35, 83).

[46] ATLAS Coll., *Forward Jet Vertex Tagging: A new technique for the identification and rejection of forward pileup jets*, ATL-PHYS-PUB-2015-034, 2015, URL: https://cds.cern.ch/record/2042098 (cit. on p. 35).

[47] P. Windischhofer, *Heavy-flavour jet tagging in ATLAS*, (2020), URL: https://cds.cern.ch/record/2706702 (cit. on pp. 35, 36).

[48] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at* $\sqrt{s}$ = 13 *TeV*, Eur. Phys. J. C **78** (2018) 903, arXiv: 1802.08168 [hep-ex] (cit. on p. 36).

[49] T. Sjöstrand et al., *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ex] (cit. on p. 38).

[50] S. Frixione et al., *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ex] (cit. on p. 38).

[51] E. Bothmann et al., *Event Generation with Sherpa 2.2*, SciPost Phys **7** (2019) 034, arXiv: 1905.09127 [hep-ex] (cit. on p. 38).

[52] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, andtheir matching to parton shower simulations*, JHEP **07** (2014) 79, arXiv: 1405.0301 [hep-ex] (cit. on p. 38).

[53] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2010) 152 (cit. on p. 38).

[54] ATLAS Coll., *Higgs Working Group*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/HiggsPublicResults (cit. on p. 39).

[55] k. Hamilton et al., *MINLO: Multi-scale improved NLO*, JHEP **10** (2012) 155, arXiv: 1206.3572 [hep-ex] (cit. on p. 39).

[56] K. Hamilton et al., *NNLOPS simulation of Higgs boson production*, JHEP **10** (2013) 222, arXiv: 1309.0017 [hep-ex] (cit. on p. 39).

[57] C. Anastasiou et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, JHEP **05** (2016) 58, arXiv: 1602.00695 [hep-ex] (cit. on p. 39).

[58] H. L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv: 1007.2241 (cit. on p. 40).

[59] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G: Nucl. Part. Phys. **43** (2016) 023001, arXiv: 1510.03865 (cit. on pp. 40, 80).

[60] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: 0709.1027 (cit. on pp. 40, 80).

[61] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 40, arXiv: 1410.8849 (cit. on p. 40).

[62] ATLAS Coll., *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **09** (2014) 145, arXiv: 1406.3660 (cit. on p. 40).

[63] ATLAS Coll., *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419 (cit. on p. 40).

[64] S. Agostinelli et al., *Geant4—a simulation toolkit*, NIM A **506** (2003) 250 (cit. on p. 40).

[65] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, (2018), arXiv: arXiv:physics/0703039 (cit. on pp. 51, 55).

[66] F. Chollet et al., *Keras*, URL: https://keras.io/ (cit. on pp. 51, 96).

[67] M. Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, (2015), URL: https://www.tensorflow.org/ (cit. on p. 51).

[68] LHC Higgs Working Group, *CERN Yellow Reports*, URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHWGLibrary (cit. on p. 80).

[69] S. Höche, F. Krauss, Schumann and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009), arXiv: 0903.1219 (cit. on p. 80).

[70] ATLAS Coll., *Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data*, Eur. Phys. J. C **77** (2017) 195, arXiv: 1612.01456 [hep-ex] (cit. on p. 83).

[71] G. Aad et al., *Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated (W/Z) H production with the ATLAS detector*, JHEP **01** (2015) 69, arXiv: 1409.6212 (cit. on p. 88).

[72] N. Berger et al., *Simplified Template Cross Sections-Stage 1.1*, (2019), arXiv: 1906.02754 (cit. on p. 93).

[73] scikit-learn, *scikit-learn, Machine Learning in Python*, URL: https://scikit-learn.org/stable (cit. on p. 96).

[74] scikit-learn, *Robust Scaler method*, URL: ttps://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.RobustScaler.html#sklearn.preprocessing.RobustScaler (cit. on p. 96).

[75] ATLAS Coll.,
Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometert,
CERN-LHCC-2017-017 ; ATLAS-TDR-026, 2017,
URL: https://cds.cern.ch/record/2285580 (cit. on pp. 105–107).

[76] W. Wu, *FELIX: the New Detector Interface for the ATLAS Experiment*,
IEEE Transactions on Nuclear Science **66** (2019) 986 (cit. on p. 106).

[77] Y. Giomataris, *MICROMEGAS: a high-granularity position sensitive gaseous detector for high particle-flux environments*, NIM Research Sec. A **376** (1996) 29 (cit. on pp. 108, 111).

[78] S. Majewski et al, *A thin multiwire chamber operating in the high multiplication mode*,
NIM Research **217** (1983) 265 (cit. on p. 108).

[79] ATLAS Coll., *New Small Wheel Technical Design Report*,
CERN-LHCC-2013-006 ; ATLAS-TDR-020, 2013,
URL: https://cds.cern.ch/record/1552862 (cit. on pp. 111, 118).

[80] T. Alexopoulos et al., *A spark-resistant bulk-micromegas chamber for high-rate applications*,
NIM Research Sec. A **640** (2011) 110 (cit. on p. 112).

[81] A. Ochi, *Carbon Sputtering Technology for MPGD detectors*, PoS(TIPP2014) (2014) 351,
URL: https://pos.sissa.it/213/351/pdf (cit. on p. 113).

[82] A. Ochi, *Development of large area resistive electrodes for ATLAS NSW Micromegas*,
Eur. Phys. J. C. **174** (2018) 03001 (cit. on p. 113).

[83] F. Kuger, *Signal Formation Processes in Micromegas Detectors and Quality Control for large size Detector Construction for the ATLAS New Small Wheel*, PhD Thesis, 2017,
URL: https://cds.cern.ch/record/2277011 (cit. on pp. 113, 117).

[84] O. Sidiropoulou, *Characterization of the ATLAS-type Micromegas Detectors*, PhD Thesis,
2018, URL: https://cds.cern.ch/record/2636918 (cit. on p. 115).

[85] P. Iengo, *"Micromegas for ATLAS New Small Wheel upgrade - status report*, (2015)
(cit. on p. 117).

[86] M. Bianco et al., *Construction of two large-size four-plane Micromegas detectors*,
NIM Research Sec. A **814** (2016) (cit. on p. 117).

[87] M. Iodice,
Resistive Micromegas for the Muon Spectrometer Upgrade of the ATLAS Experiment,
PoS(ICHEP2016), 2016 275, URL: https://pos.sissa.it/282/275/pdf (cit. on p. 118).

[88] E. N. Gazis et al., *Design of the ATLAS New Small Wheel Gas Distribution System for the Micromegas Detector Modules*, PoS(LHCP2016), 2016,
URL: https://pos.sissa.it/276/198/pdf (cit. on p. 119).

[89] G. D. Geronimo, *VMM1—An ASIC for Micropattern Detectors*,
IEEE Transactions on Nuclear Science **60** (2013) 2314 (cit. on pp. 119, 138, 148).

[90] I. K. et al., *Influence of electron recapture by the cathode upon the discharge characteristics in dc planar magnetrons*, Phys. Rev. E **72** (2005) 056402 (cit. on p. 121).

[91] T. W. et al., *Semiempirical analysis of electron scattering cross sections in N_2O and CO_2*,
Czechoslovak Journal of Physics **54** (2006) C742 (cit. on p. 121).

[92] A. Sharma, *Properties of some gas mixtures used in tracking detectors*, SLAC-J-ICFA-16-3, 1998,
URL: https://www.slac.stanford.edu/pubs/icfa/summer98/paper3/paper3.pdf (cit. on p. 121).

[93] W. Blum et L. Rolandi, *Particle Detection with Drift Chambers*, Book, 1993 (cit. on p. 121).

[94] PDG, *Atomic and Nuclear properties of materials*,
URL: https://pdg.lbl.gov/2004/reviews/atomicrpp.pdf (cit. on p. 122).

[95] F. Vogel,
Test of ATLAS Micromegas detectors with a ternary gas mixture at the CERN GIF++ facility, ATL-MUON-PROC-2021-009, 2021, URL: https://cds.cern.ch/record/2783915/ (cit. on p. 121).

[96] T. Alexopoulos et al., *Performance studies of resistive-strip bulk micromegas detectors in view of the ATLAS New Small Wheel upgrade*, NIM Research Sec. A **937** (2019) 125 (cit. on p. 132).

[97] T. Beermann et al., *Implementation of ATLAS Distributed Computing monitoring dashboards using InfluxDB and Grafana*, EPJ Web of Conferences **245** (2020) 03031 (cit. on p. 175).

[98] S. Martoiu et al., *Development of the scalable readout system for micro-pattern gas detectors and other applications*, JINST **8** (2013) C03015 (cit. on p. 175).

[99] A. Barriuso Poy et al., *The detector control system of the ATLAS experiment*, JINST **3** (2008) (cit. on p. 177).

[100] M. Alviggi et al.,
Pixelated resistive bulk micromegas for tracking systems in high rate environment, JINST **15** (2020) C06035 (cit. on pp. 179, 180).

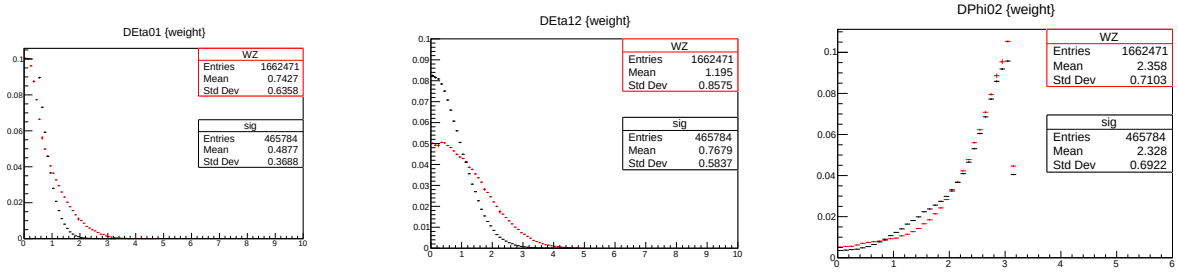
[101] M. Alviggi et al., *Small-pad resistive Micromegas for high rate environment: Performance of different resistive protection concepts*, NIM Research Sec. A **936** (2019) (cit. on pp. 179, 180).

[102] M. T. Camerlingo et al.,
Rate capability and stability studies on small-Pad resistive Micromegas, PoS(ICHEP2020), 2020 825 (cit. on p. 180).

Appendices

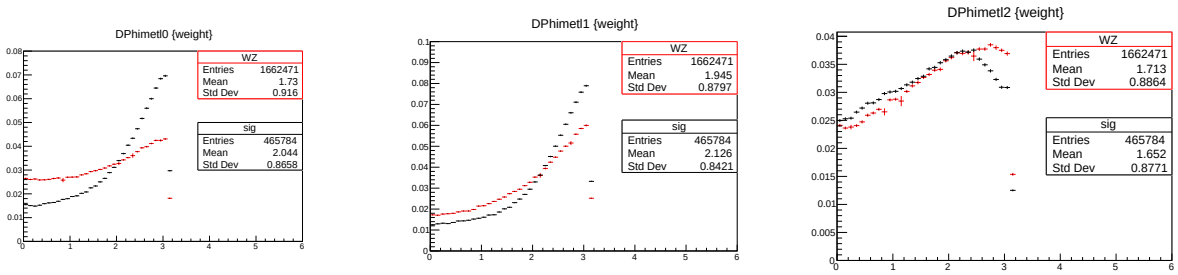
A Distribution of the alternative samples for the ANN training

The original samples from which the training and test/validation samples for the ANN implementation have been scaled. Every variable has had its mean subtracted, and it has been scaled to its standard deviation, as described in Section 2.8.2.
The mean and standard deviation are relative to the variable distribution that includes all the signal and background events. The distributions of the scaled 15 ANN input variables are then reported in Figures A.1 to A.5 for each process. The WH process is illustrated in black, and the WZ process in red. The training and test/validation samples are successively made through a random division of these distributions. The variables that show a good separation between signal and background processes are: the azimuthal separations $\Delta\Phi_{l_i E_T^{miss}}$ between the lepton l_i and E_T^{miss}, the angular distance $\Delta R_{l_0 l_1}$, $m_{l_0 l_1}$, $\left|\sum_i p_T^{l_i}\right|$, E_T^{miss} and m_T^W.



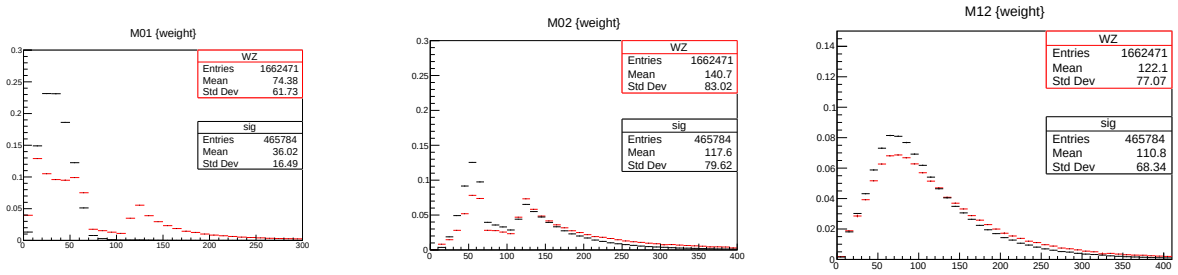
(a) $\Delta\eta_{l_0 l_1}$ (b) $\Delta\eta_{l_1 l_2}$ (c) $\Delta\phi_{l_0 l_2}$

Figure A.1: Angular variables of couple of charged leptons.



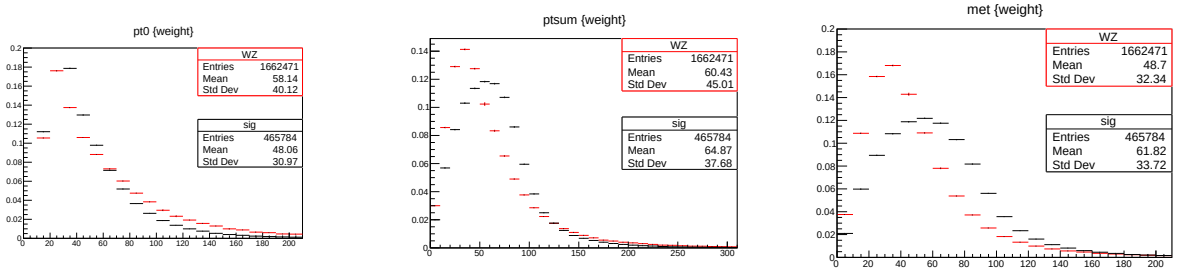
(a) $\Delta\phi_{l_0 E_T^{miss}}$ (b) $\Delta\phi_{l_1 E_T^{miss}}$ (c) $\Delta\phi_{l_2 E_T^{miss}}$

Figure A.2: Angular variables of couple of a charged lepton and E_T^{miss}.



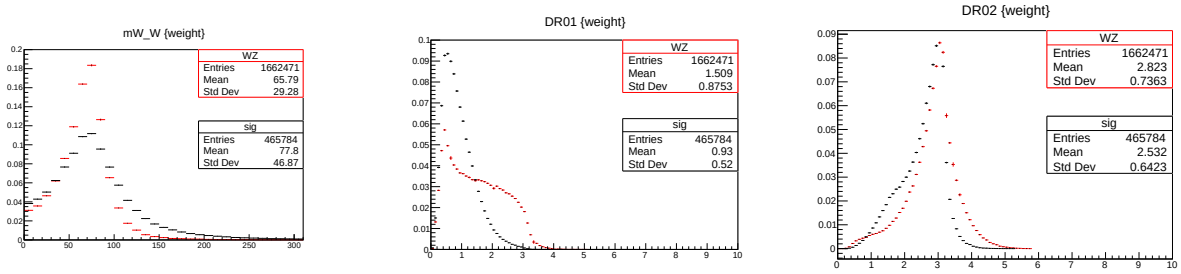
(a) $m_{l_0 l_1}$ (b) $m_{l_0 l2}$ (c) $m_{l_1 l_2}$

Figure A.3: Invariant masses of couple of charged leptons.



(a) $p_T^{l_0}$ (b) $\left|\sum_i p_T^{l_i}\right|$ (c) E_T^{miss}

Figure A.4: Transverse momentum and energy variables.



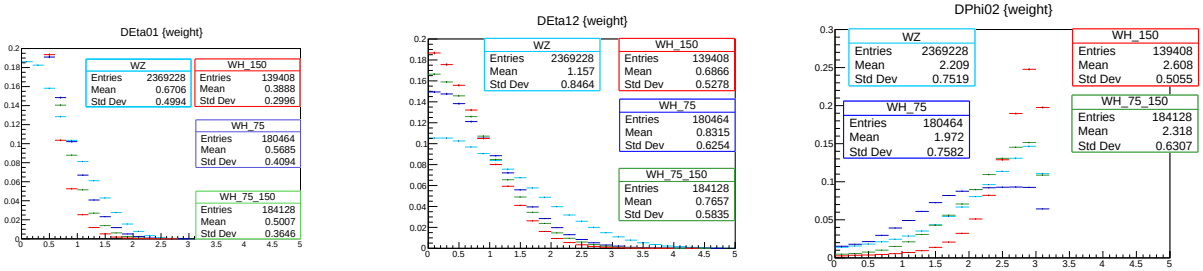
(a) m_T^W (b) $\Delta R_{l_0 l_1}$ (c) $\Delta R_{l_0 l_2}$

Figure A.5: Invariant mass of associated W boson and angular distances ΔR.

B Distribution of the alternative samples for the multiclassifier training

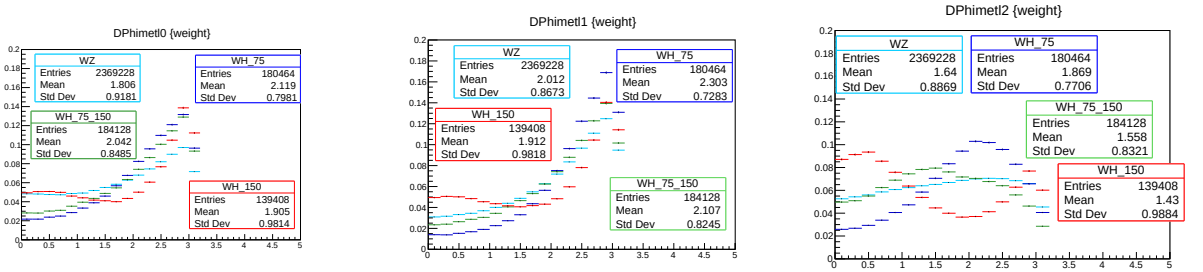
The multi-classifier training and test/validation samples have passed a dedicated selection of cut and vetos, different from the $3l$-"Z-dominated" SR definition.
As described in Section 2.12.1, the multi-classifier has four classes: the WH process with $p_T^W < 75$ GeV (in blue in the next figures), the WH process with $75 \leq p_T^W < 150$ GeV (in green in the next figures), the WH process with $p_T^W \geq 150$ GeV (red in the next figures) and the WZ process (in light blue in the next figures). The optimisation studies for the set of the multi-classifier input variables have started from the 15 input variables of ANN in Figure B.1 to B.5, and other 10 variables have been added.
The relative trends among the distributions for the four truth-level classes show that the starting core of 15 ANN input variables is a valid set even after this different selection.
The variables that show the best separation among the classes are: the azimuthal separations $\Delta\Phi_{l_i E_T^{miss}}$ between the lepton l_i and E_T^{miss}, the angular distance $\Delta R_{l_0 l_1}$ and $\Delta R_{l_0 l_2}$, $m_{l_0 l_2}$ and $m_{l_1 l_2}$, $\left|\sum_i p_T^{l_i}\right|$, E_T^{miss} and m_T^W.
The transverse momentum of lepton l_2 and the RNN output p_T^R in Figure B.6 have been included for their capability to discriminate the truth-level p_T^W classes of the WH process. Indeed, they have competed in the selection of the p_T^W proxy variable. The 8 additional discrete variables in Figures B.7 to B.9 do not help to discriminate the truth level p_T^W classes of the WH process.

B Distribution of the alternative samples for the multiclassifier training



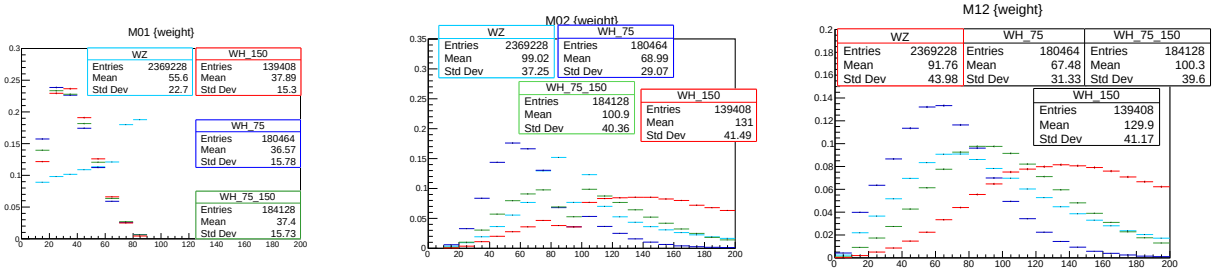
(a) $\Delta\eta_{01}$ (b) $\Delta\eta_{12}$ (c) $\Delta\phi_{02}$

Figure B.1: Angular variables of couple of charged leptons.



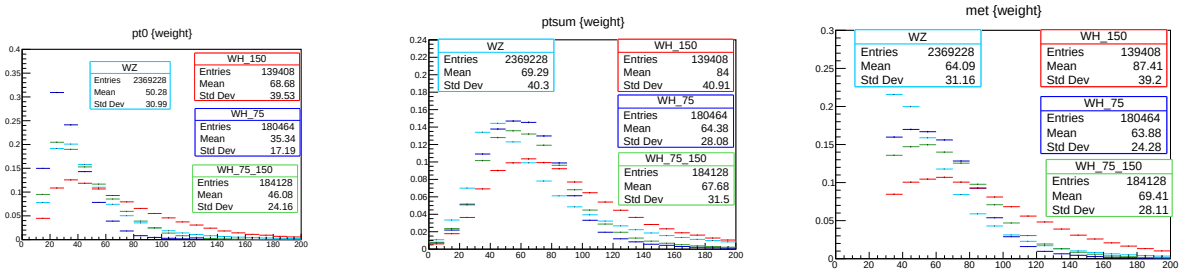
(a) $\Delta\phi_{MET0}$ (b) $\Delta\phi_{MET1}$ (c) $\Delta\phi_{MET2}$

Figure B.2: Angular variables of couple of a charged lepton and MET.



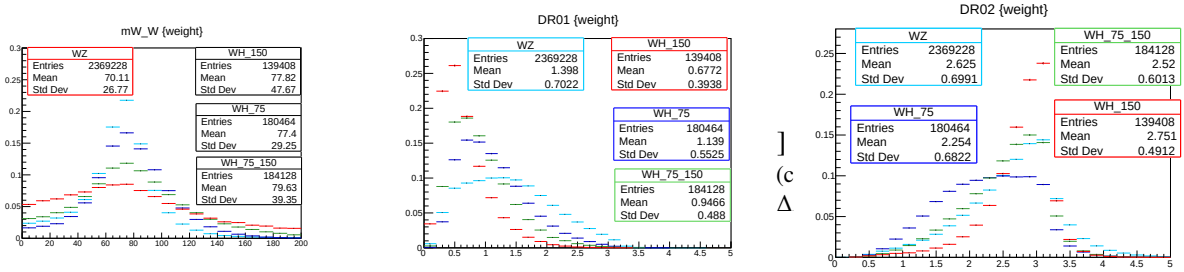
(a) M_{01} (b) M_{02} (c) M_{12}

Figure B.3: Invariant masses of couple of charged leptons.



(a) p_{T0} (b) p_{sumT} (c) MET

Figure B.4: momentum and energy variables



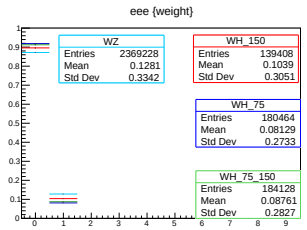
(a) m_{TW} (b) ΔR_{01}

Figure B.5: invariant mass of associated W boson and angular distance ΔR.

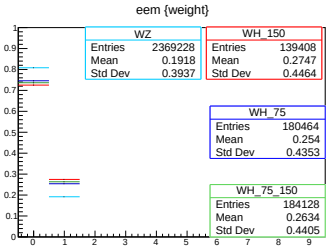
(a) $p_T^{l_2}$ (b) p_T^R

Figure B.6: Transverse momentum of l_2 lepton and output of regression neural network.

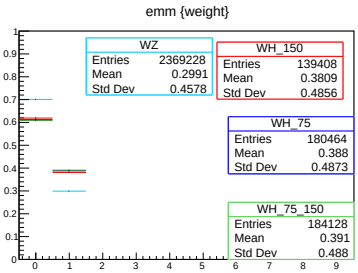
B Distribution of the alternative samples for the multiclassifier training



(a) eee

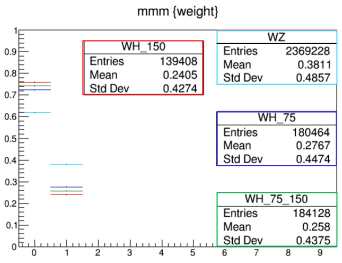


(b) eem

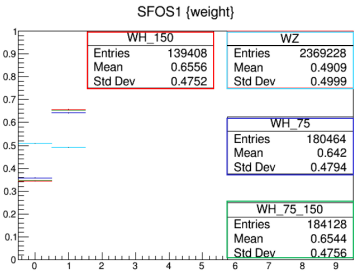


(c) emm

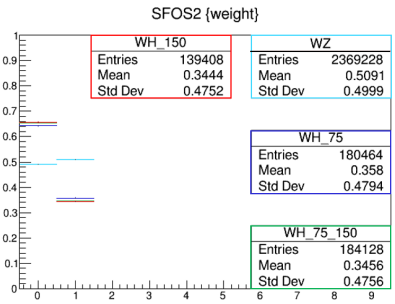
Figure B.7: Discrete variables relative to lepton flavour and $SFOS$ pairs.



(a) mmm

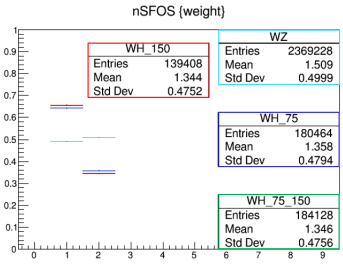


(b) $SFOS1$

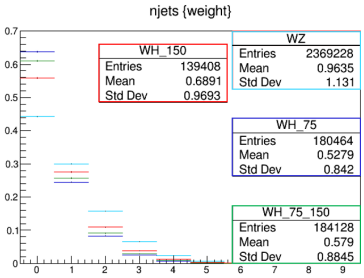


(c) $SFOS2$

Figure B.8: Discrete variables relative to lepton flavour and $SFOS$ pairs.



(a) n_{SFOS}



(b) n_{jet}

Figure B.9: Discrete variables relative to $SFOS$ pairs and number of jets.

C CERN set-up for the validation of the ATLAS Micromegas quadruplets

The stand for the validation test of the ATLAS Micromegas quadruplets is placed in the CERN BB5(889) building. It is equipped with a table customised for all module sizes (Figure C.1) and a mechanical structure to sit the scintillator counters of the trigger system (Figure C.1). The stand has a dedicated Ar:CO_2 gas line in which a VAISALA humidity sensor (read by picolog software) is present. Two independent mainframes, two CAEN boards and a single CAEN supply module form the high voltage power supply system. One board polarity is negative to supply each drift plane independently at -300 V voltage. The resistive strips are supplied with positive voltages using a CAEN board A7038, customised for the final NSW power supply. A single positive HV channel supplies one of the HV sectors of the module (or half PCB).
The CAEN supply module is for the photomultipliers (coupled with the trigger system scintillators). For the Ar:CO_2(93%:7%) mixture, the set limits on the supply boards are 30 s as trip time and 2 μA as a max $I0$. Therefore, if the current overcomes the $I0$ for more than 30 s, the HV is automatically switched off. With the available set-up, a spark is defined as a sampled variation of current greater than 100 nA with respect to the baseline of power supply without load. The tolerated spark rate is less than 6 sparks per minute.
The customised specifications for the Ar:CO_2:iC_4H_{10}(93:5:2)% foresees a 10 s trip time for the same maximum $I0$ (2 μA). The spark threshold moves to 25 nA, and the max acceptable spark rate is set to 0.6 sparks per minute.
The current on the resistive strips is real-time monitored by a GRAFANA dashboard [97], interfaced with the GECO2020 CAEN software, and stored by a C-language code developed at LNF, based on the CAEN HV Wrapper Library functions. To acquire the signal from the readout strips, a DAQ system based on the hybrid APV-25 chip, composed of an SRS crate ([98]), 5 FECs with ADC cards, and a slow controller to manage the communication between the FECs and the serial port of the PC. Adapter cards with zebra connectors have been used to match one APV channel with a pair of readout strips.

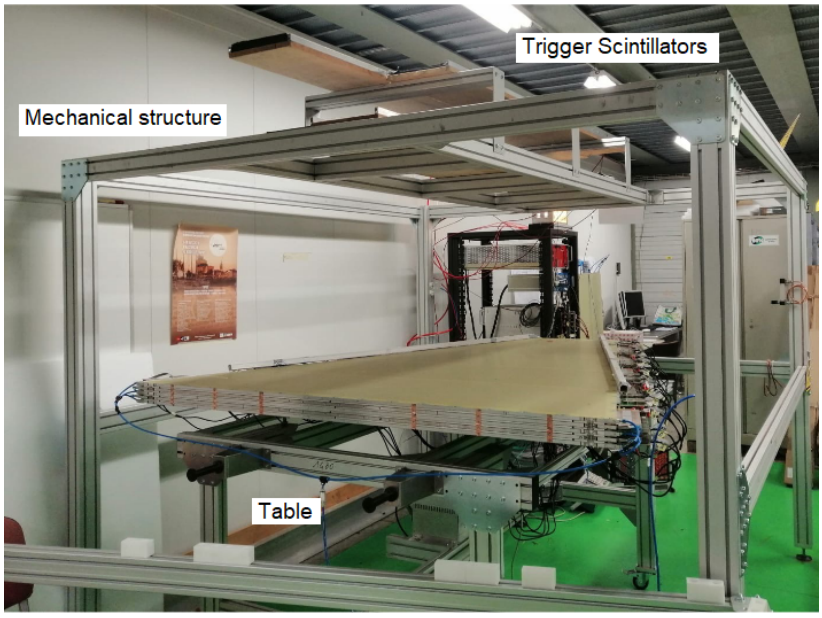


Figure C.1: Picture of the CERN stand for the validation of the Micromegas modules.

D CERN set-up for the validation of the ATLAS Micromegas Double Wedges

In the same BB5(889) building, another stand is present for the validation of the ATLAS Double Wedges. Its set-up is similar to the one in Appendix C because the performed tests are the same, with the big difference that the final HV cables/connectors and the final front-end are soldered or mounted on the modules of the ATLAS Double Wedges. Therefore, the structure permits to sit the scintillator counters of the trigger system, and it has two trolleys to place the detectors. The stand has two gas lines for the binary Ar:CO_2(93:7)% and ternary Ar:CO_2:iC_4H_{10}(93:5:2)% mixtures. The humidity is monitored by VAISALA humidity sensors, read by a Arduino circuit, and the overpressure with respect to the atmospheric value is measured using differential manometers at the output. Up to the patch panel, the HV supply system is similar to the one described in the previous Appendix C. For the readout strips supply, CAEN board A7038 is used in combination with customised splitter boxes. By means of these splitter-boxes, a single positive channel supplies two HV sectors of the Double Wedge, i.e. the resistive strips of a PCB from the two antennas at its sides. A single channel of the supply board with negative board supplies in parallel four sides of drift planes at -300 V voltage.
For the real-time current monitoring and data storing at this stand, there are two redundant software. A dedicated Grafana dashboard, interfaced with the CAEN software, is used principally to monitor the time trends, average values and spark rate. A temporary Detector Control System [99], based on WinCC, is used to store the data and control the HV boards, too. For the final VMM3 Front-End electronics, the stand is also equipped with a cooling system, modules for the Low Voltage and digital data acquisition.

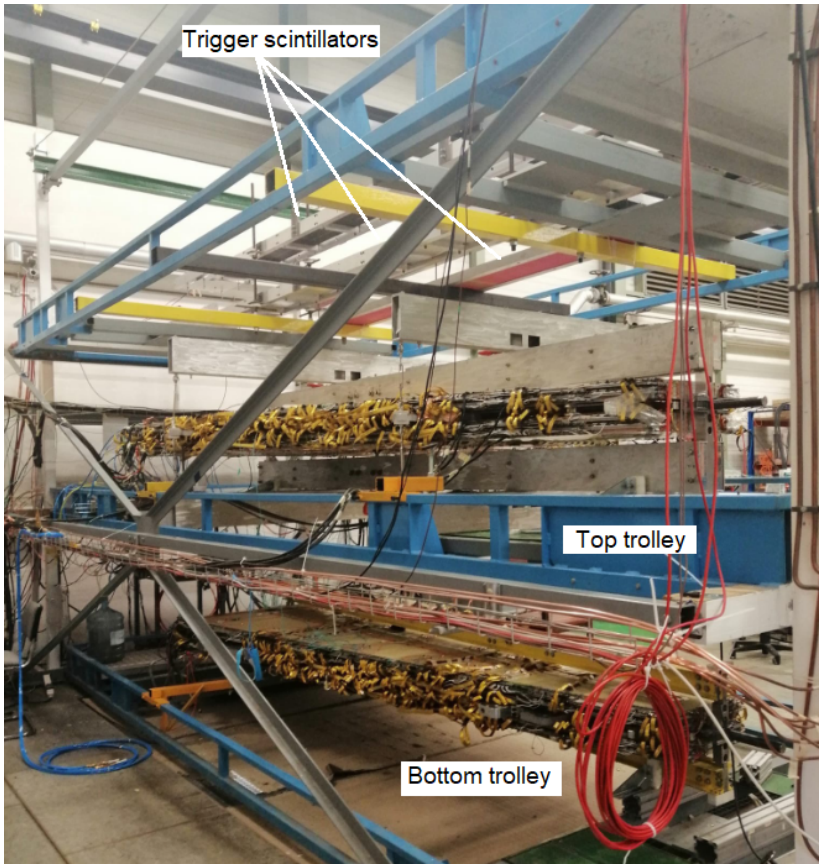


Figure D.1: Picture of the CERN stand for the validation of the Micromegas Double Wedges.

E On-going R&D on resistive Micromegas detectors

The R&D on Resistive High granularity Micromegas (RHUM) aims to develop resistive Micromegas detectors with high granularity readout (pad O(mm^2)) capable of operating efficiently at rates of O(10 MHz/cm^2). To achieve this goal, the readout electrodes have been minimised in comparison with the standard Micromegas resistive strip detectors, to reduce their occupancy (small pads of order 3 mm^2). In addition, different resistive spark suppression layouts have been optimised and implemented in different prototypes.

E.1 RHUM prototypes

Several small-pad Micromegas detectors have been built with similar anode planes: segmented in matrices of 48x16 readout pads with a rectangular shape (0.8x2.8 mm^2) and a pitch of 1 and 3 mm in the two coordinates, respectively. The active surface is 4.8x4.8 cm^2 with 768 electronics channels routed off-detector for readout. Consequently, a new sparks suppression resistive layout is needed to match the pad geometry. Two different schemes have been implemented (see the left and right sides of Figure E.1). The first scheme (shown on the left side of Figure E.1) involves a pad-patterned layer with embedded resistors for each readout pad, and it is referred to as PAD-P. The second scheme (shown on the right side of Figure E.1) exploits a double layer of Diamond-Like-Carbon (DLC) resistive foils, which sizes are the same as the active area. Since 2016, long-term characterisation and performance studies of these detectors have been carried out. A complete comparison between the performances of the prototypes with these different resistive layouts is reported in References [100] and [101].

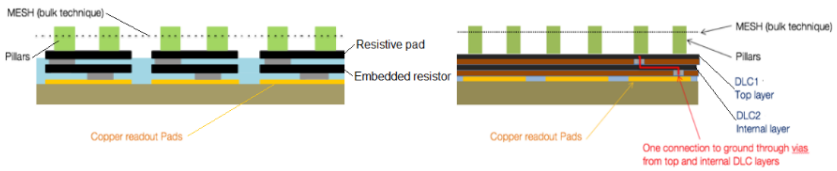


Figure E.1: Sketches of the transverse sections of the pad-patterned scheme with embedded resistors (left) and the scheme with a double layer of Diamond-Like-Carbon (DLC) resistive foils (right) [101].

In all the presented studies, the gas mixture is Ar:CO_2(93%:7%).
The PAD-P detector is stable and robust against discharges. It has a fair spatial resolution of 190 μm in the 1 mm pitch coordinate and an energy resolution of about 48% (FWHM). The PAD-P gain is affected by the charging-up effect (decreasing by about 20% at a particle rate of $\sim$ 1 MHz/cm^2) due

to the exposed dielectrics in the amplification gap. This is also visible in the time evolution of the current [100]. Trying to improve the performance of the PAD-P scheme and following the general interest in the Diamond-Like-Carbon application in MPGD, the second layout has been designed like the sketch on the right side of Figure E.1, where the pad-patterned planes and intermediated embedded resistors are replaced by two continuous DLC foils. The spatial resolution improves (below 100 μm on the same precision coordinate) because of a larger spread of the charge among the pads. The energy resolution is also better thanks to a more uniform electric field, which is not affected by edge effects on the resistive layer pads. The charging-up effect is minimised [101]. In this scheme, a net of low resistive silver vias evacuates the current from the double layer of DLC to the ground, avoiding the reduction of the rate capability due to the current spread on all the planes. Several DLC prototypes have been built, with different techniques and different values of the average resistivity of the DLC foils [100]. Among those, the best performances are obtained with the DLC20-6 mm vias pitch detector, with approximately 20 M$\Omega/\square$ resistivity.

An upgraded version of the second scheme has been realised by means of the Sequential Built-Up technique (SBU), and a prototype PAD-P-3 has been made with a third scheme, mixing some features of the first two schemes.

E.2 Rate capability measurements

In this phase of technology development, several routine measurements on the main parameters have been repeated to track the time evolution of the detector performance and to compare the different prototypes. For precision tracking application, the considered parameters are: gain, energetic resolution, spatial resolution, tracking efficiency, rate capability and stability. In the case of rate capability measurements, the difficult aspect is to build a set-up and procedure that are able to manage very broad ranges of currents and rates. The procedure is described in the following Section E.2.1, together with some results on the rate capability and stability of the first prototypes in Sections E.2.2 and E.3, which have been presented in Reference [102].

E.2.1 Procedure of rate extrapolation vs current

A scan has been performed varying the current of an X-ray gun with a Cu target (~8 keV photons) to investigate the rate capability of small Pad resistive Micromegas detectors. The experimental setup (shown in Figure E.2) also includes:

- A power supply module (CAEN N147H) with a 0.5 nA sensitivity error and a 20 μA maximum of the range (I set).
- A Multi Channel Analyzer (Amptek MCA-8000);
- A charge sensitive preamplifier with an amplifier stage (CERN RD51 APIC);
- A picoammeter (Keithley 6487);

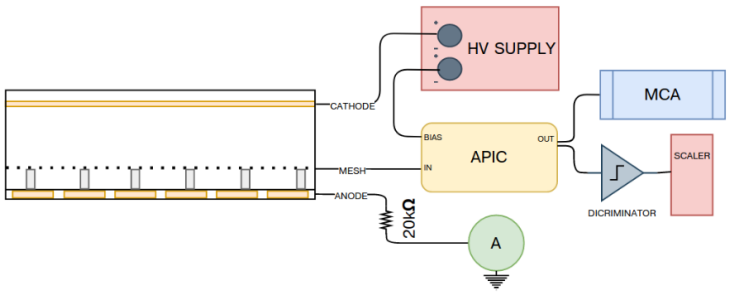


Figure E.2: Sketch of experimental setup for X-ray measurements.

Beyond providing the voltage on the mesh and resistive electrodes, the power supply module is used to directly acquire the current on the mesh. It is connected to a PC, and the current values are stored by a Labview VI with a O(1 Hz) sampling frequency. The picoammeter has been occasionally inserted in the set-up to check if the current is correctly acquired when its values are close to the low limit of the power supply. The APIC amplifies and shapes the signal from the mesh, whose occurrence (rate) and amplitude are acquired by the MCA with fast and slow shaping, respectively. Also, the MCA counter has been tested, performing redundant measurements with MCA and an external scaler in a rate range less than 300 kHz, where responses of the available scaler and MCA are linear. Thanks to the combination of a correction on the dead time of MCA and the absolute peak acquisition mode, the MCA response is still linear up to ~ 500 kHz. For higher rates, the response functions of both the instruments (MCA and scaler) are not linear. Then, a procedure of extrapolation is applied to the rates higher than 300-500 kHz in the gain and rate capability measurements. Linear regression between the measured rate and current of the X-ray gun is performed in a range of low values of current of the X-ray gun, as shown in Figure E.3. Then, the fit parameters are used to extrapolate the high rates.

A set of copper masks with holes of different sizes (0.071 cm^2, 0.79 cm^2 circular, or 1.9x1.9 cm^2

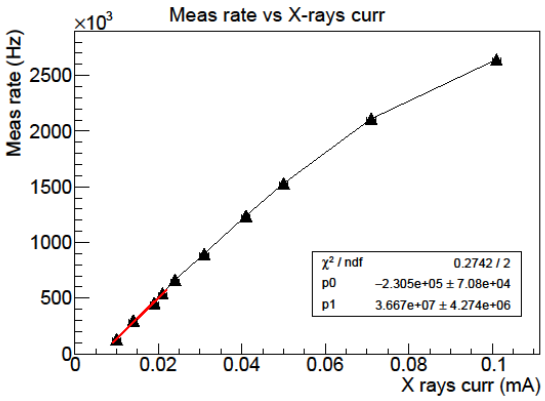


Figure E.3: (MCA) Measured rate as a function of X-ray current, in PAD-P when a Cu mask with 0.79 cm^2 circular hole is placed above the active area (April 2019).

squared holes) has been used to investigate the dependence on the irradiated area for a possible application of the MM-Pad technology to large areas.

It has been observed that the X-ray sensitivity does not allow to acquire a sufficient number of measurements in the range of rates lower than 300 kHz to perform a correct linear regression when the area increases. The fit can be performed using the mask with the 0.79 cm^2 hole area as the maximum limit area among the other masks.
Then, for the larger areas, the rate extrapolation is carried out using a combination of two linear fits. The combined fits are based on the linear proportionality between the current on the mesh and rate (typical of the proportional mode gaseous detector), as reported in Figure E.5.
As shown in Figure E.5, a first linear fit (Equation E.1) is performed between the current on the mesh

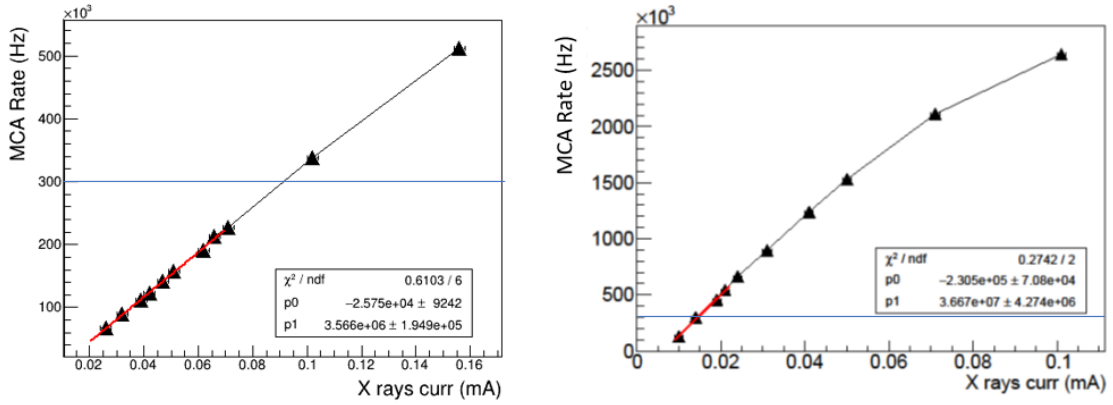


Figure E.4: Measured rate as a function of X-ray current with 0.071 cm^2 circular area (left-August 2019) and measured rate as function of X-ray current with 0.79 cm^2 circular area (right-April 2019) in PAD-P at 527 V.

(I in Equation E.1) and the measured rate (r in Equation E.1) in the range < 500 kHz with the 0.79 cm^2 (or 0.071 cm^2) where the linear response of scaler/MCA counter voltage is linear and the drop due to the impedance (2.7 MΩ) of the APIC input filter is negligible.

$$I = br + a \quad \text{where} \quad b = n_0 eG \tag{E.1}$$

The parameter b is proportional to the gain factor G[1]. No physical meaning is attributed to the offset parameter a. The statistical fluctuations associated with the fit range make less probable the interpretation of a as residual dark current, surviving the subtraction of the baseline current (without X-ray irradiation) in each I value before the fit. Then, a linear fit is performed between the current on the mesh (I in Equation E.2) and X-ray current (X in Equation E.2) for each investigated area. Their linear relation is not affected by the limitations of the acquisition set-up.

$$I = cX + d \tag{E.2}$$

Where c is the proportional factor between the two currents and d is a residual fit offset. By combining the fit parameters of Equations E.1 and E.2, the researched relation that overcomes the set-up limitations for the large areas is built as:

$$r = \frac{c}{b}X + \frac{d-a}{b} \tag{E.3}$$

[1] The other symbols n_0 and e in Equation E.1 respectively indicate the number of the average ionisation pairs and elementary charge.

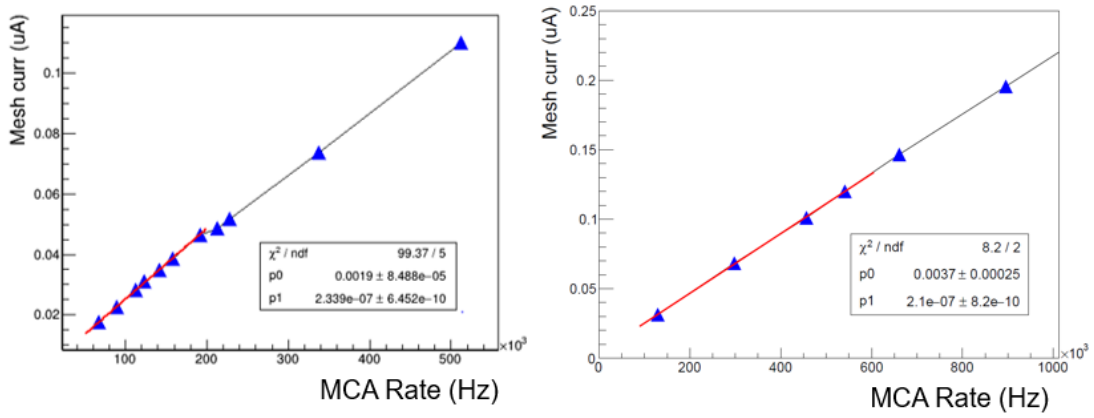


Figure E.5: Mesh current as a function of measured rate with 0.071 cm^2 circular area (left-August 2019) and mesh current as a function of measured rate with 0.79 cm^2 circular area (right-April 2019) in PAD-P at 527 V. In the labels, $p0$ is parameter a and $p1$ is parameter b of Equation E.1.

The combined $\frac{d-a}{b}$ term is compatible with zero in a few cases, and the contribution of the $(d-a)$ constant on the error is dominant for small values of X-ray current. The relative error decreases up to 5-7% at the highest rates. The magnitude of the relative errors on the low extrapolated values is around 20% (relative to a $\sim$ 0.020 mA X-ray current). The final extrapolated rates r are reported on the horizontal axis of Figure E.6, where the current on the mesh acquired supplying the detector directly, without APIC in the acquisition chain in Figure E.2. Then, Figure E.6 shows the current on the mesh as a function of r in a range that exceeds 500 Hz (divided by the mask hole). The

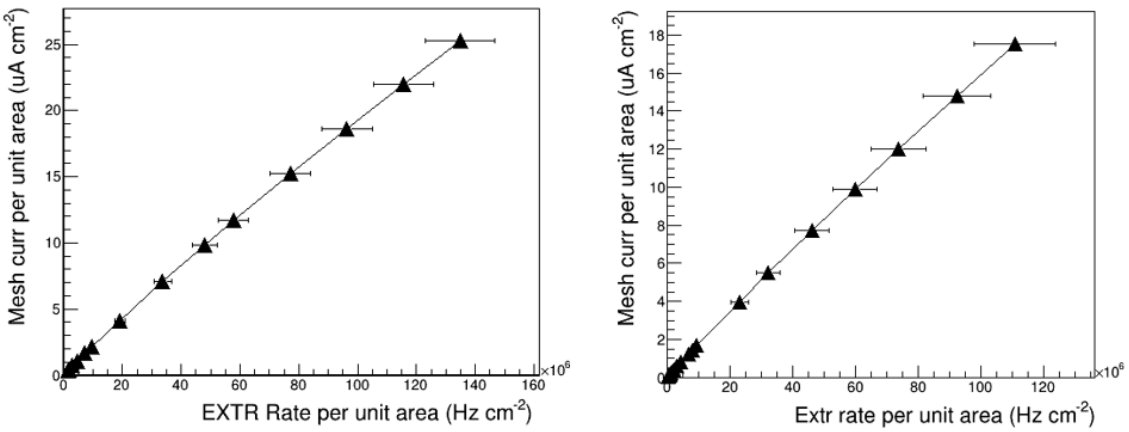


Figure E.6: Mesh current per unit area as function of X-ray current with 0.071 cm^2 circular area (left-August 2019) and mesh current as a function of X-ray current with 0.79 cm^2 circular area (right-April 2019) in PAD-P 527 V.

extrapolated results are subject to possible environment and noise variations between two different

acquisition sessions because the measurements are carried out in a non-controlled environment. The environment variations affect the gain factor that can be slightly different between the two sessions at a fixed amplification voltage. While the noise variations principally affect the rate measurements. The first effect is minimised by performing the same measurement at different amplification voltages in each acquisition session. The impact of the noise variation is reduced by optimising the MCA threshold in every acquisition session.
For example, the measurements for the 0.071 cm^2 and 0.79 cm^2 areas are acquired in different sessions. Their corresponding extrapolated rates differ by about -20% between the 0.071 cm^2 area (left side of Figure E.6) and the 0.79 cm^2 area (right side of Figure E.6) in all the investigated ranges.
In addition, the 0.071 cm^2 area is more sensitive to possible local inhomogeneities in the active area than the 0.79 cm^2 case, where the possible inhomogeneities are mediated by the larger area. This more sensitivity of 0.071 cm^2 area may contribute to the differences observed between the two session days. Then, the 0.79 cm^2 area has been chosen as the best mask hole to use in the first fit of the extrapolation rate procedure by compromising the set-up limits and variations due to possible inhomogeneities (or shifts) of the irradiated area.
Then, the results of direct comparison between measurements obtained during the same session with similar conditions of acquisition are reported in the following sections.

E.2.2 Study on the rate capability

The rate capability is defined as the gain stability with respect to the particle rate. As the rate increases, the gain drops in the small-pad Micromegas detectors since the current on the resistive elements or charging-up effect becomes significant. Therefore, the relative gain with respect to its value at low rates is studied as a function of the rate to find the rate value at which the gain drops below a limit (20%) at different amplification voltages. The relative gain is reported as a function of hit rates using a ^{55}Fe radioactive source (circles markers) to reach the lowest rate values or an X-ray gun (triangle markers) in the rest of the rate range in Figure E.7, for PAD-P and DLC20, as the best representative of the DLC family.
PAD-P requires higher voltage with respect to DLC20 to set the same gain factor (about 20 V more) due to boundary effects in the amplification electric field configuration.
In the rate range below 0.5 MHz/cm^2, the gain trend is mainly affected by charging-up in PAD-P and its drop is already significant at 540 V amplification voltage, where the gain at the low rate is $\sim 10^4$ (black in Figure E.7). Conversely, the gain trend is quite stable in this range in DLC20, where there is no exposed dielectric material in the amplification gap that can charge up (except for the pillars). In the full X-ray rate range, the trend of PAD-P visibly changes above 5 MHz/cm^2 because the charging-up and ohmic contributions behave differently. The latter becomes more relevant as the rate increases while the charging-up saturates. In the high rate regime, a gain drop of 20% is reached for PAD-P at approximately 25 MHz/cm^2 operating the detector at 530 V (corresponding to a gain of 6500 at low rates). The same drop is reached for DLC20 (at equivalent gain i.e. HV=510 V) around 20 MHz/cm^2. At higher rates, PAD-P has a smaller gain drop than DLC20.
The study with X-ray gun is repeated by varying the exposure area (0.79 cm^2 circle, 3.69 cm^2 square, 9 cm^2 circle) and reported in Figure E.8. PAD-P gain drop is independent on the irradiated area. While, in the DLC prototype, the drop dependence on the area becomes negligible when the irradiated area is significantly larger than the cell defined by the vias pitch.

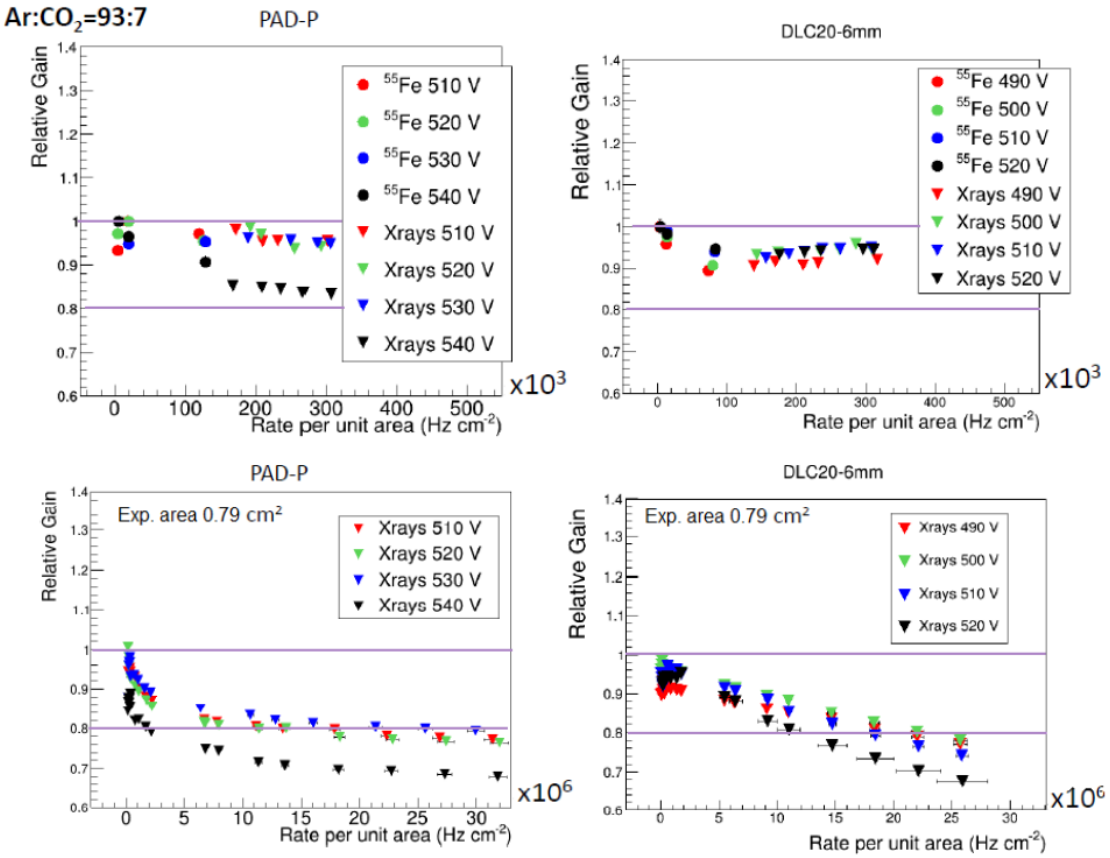


Figure E.7: Relative gain G/G_0 as a function of rate per unit area (^{55}Fe circle markers and X-ray gun triangle markers) where G_0 is the gain at low rates. Zoom of G/G_0 in the low rate per unit area (top) and in the full investigated X-ray range (bottom)

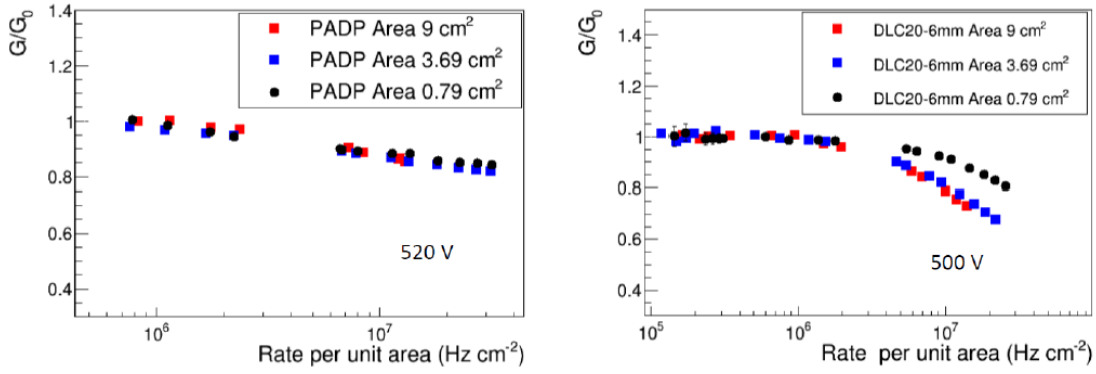


Figure E.8: Relative gain G/G_0 as a function of rate per unit area for different exposure areas (0.79 cm^2 circle, 3.69 cm^2 square, 9 cm^2 circle) where G_0 is the gain at low rates.

E.3 Study on stability

During the test beam in the year 2019 at the PSI πM1 beamline (300 MeV/c positive charged pion), a first study on stability has been carried out with charged particles. The gain curves on the left side of Figure E.9 as a function of the amplification have been produced by the test beam data. They have been used to plot the spark rate probability (pr^{spr} in Equation E.4) as a function of the gain. In this study, a current increase is counted as a spark when larger than 30% of the stable value at a fixed amplification value. Consequently, the spark probability(/pion/area) is defined as the number of sparks per the time acquisition interval, per area and particle rate (see the right side of Figure E.9).

$$pr^{spr} = \frac{\text{number of sparks}}{\text{pion rate} \cdot \text{Time window} \cdot \text{Area}} \tag{E.4}$$

PAD-P is the most robust prototype, as visible on the right side of Figure E.9. From this study and all

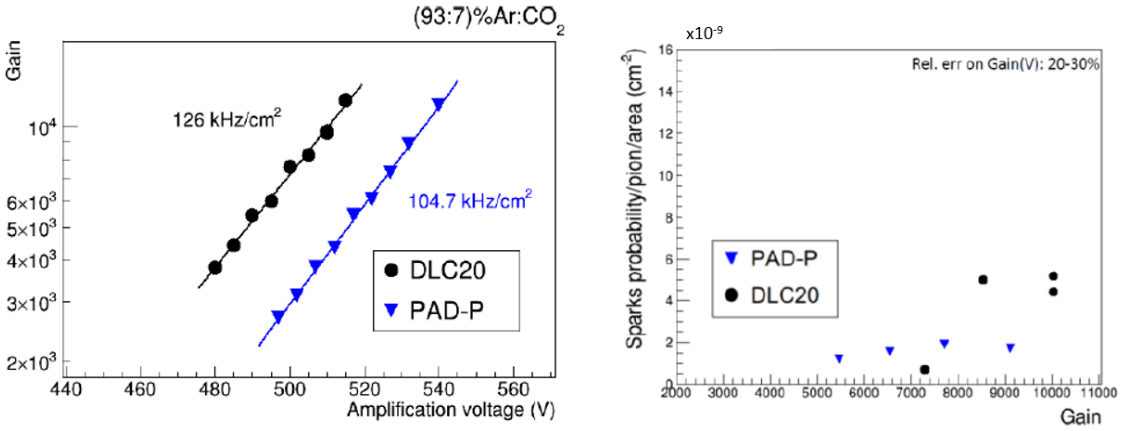


Figure E.9: Gain as a function of amplification voltage (left) and spark rate probability per incident particle (pions) and per unit area as a function of gain measured during test beam a 300 MeV/c charged pion beam (right).

previous tests, the PAD-P detector shows very good stability and the possibility to operate at gains larger then 10^4. Also, the DLC20 has a low spike probability (lower than $5x10^{-9}cm^{-2}$ per incident pion) up to a gain of 10^4. However, if it operates at larger gains than 10^4, its HV behaviour could be subject to instabilities. Moreover, the comparison between PAD-P and the first DLC prototypes is not conclusive because low resistance spots have been found in the DLC prototypes during their characterization tests. They consist of some current evacuation silver vias that are in part directly exposed to the charges in the amplification gap, as shown on the left side of Figure E.10. This constructive defect is solved by the most recent Sequential-Build-Up (SBU) technique, as shown on the right side of Figure E.10.

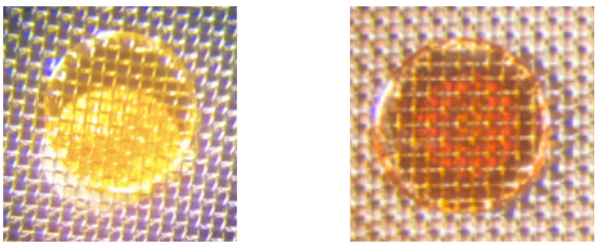


Figure E.10: Exposed vias in DLC prototype (left) and covered vias in a SBU prototype (right).