

https://dspace.org/

Fake lepton studies in a search for heavy scalar in dilepton + jets final states with ATLAS detector

Mashishi, Lehumo

2021
https://hdl.handle.net/10539/32433

Downloaded from a DSpace repository. https://dspace.org/

Fake lepton studies in a search for heavy scalar in dilepton + jets final states with ATLAS detector



Institute for Collider Particle Physics
University of Witwatersrand

Lehumo Mashishi

Supervisor: Prof. Bruce Mellado

Co-Supervisor: Dr. Yesenia Hernández Jiménez

A dissertation submitted to the University of
the Witwatersrand in accordance with the requirements of the degree of
Masters of Science in the Faculty of Science

July, 2020

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or processional qualification except as specified.



Lehumo Mashishi, July 2020

Abstract

The estimation of fake lepton background is presented for the physics beyond the Standard Model search in the final state with a pair of different flavour opposite-sign leptons, $e\mu$, accompanied by at least two jets. The search uses the LHC full Run 2 data recorded at a center-of-mass energy of 13 TeV collected with the ATLAS detector, with the corresponding luminosity of 139 fb^{-1}. The search extends the Standard Model by introducing the heavy scalar, H, and the additional Higgs-like scalar, S. One of the challenges in most physics analysis is to estimate background processes that can enter the signal region and be misidentified as the signal. This thesis focuses on the W+jets background which enters the signal region. From the W+jets background process, there is one real lepton from W and one fake which originates from the jet misidentified as a lepton. The Monte Carlo simulation cannot be trusted for the estimation of W+jets background as the rate of misidentification might not be well modelled. Hence one has to develop the data-driven techniques for the W+jets estimation. In the $H \to Sh \to e^{\pm}\mu^{\mp}$+jets analysis, the fake factor method is used for fake estimation.

Dedication

“ To God and my beloved parents for always supporting me, because they are the driving force of my life and career. Without their love none of this would matter. Throughout my life, they have actively supported me in my determination to find and realize my potential, and to make this contribution to the world.”

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor Prof. Bruce Mellado and for giving me the opportunity to carry out my master's thesis at University of Witwatersrand, School of Physiscs. This encouraged me to become more involved in my field and opened the way to new experiences.

I would also like to thank Dr.Yesenia Hernandez Jimenez warmly for her unlimited patience and close guidance in the course of this research, her useful comments and, most importantly, much appreciated help.

Nomenclature

Acronyms	Meaning
ATLAS	A Toroidal LHC AppparatuS
BEH	Brout–Englert–Higgs
BSM	Beyond Standard Model
CERN	Conseil Européen pour la Recherche Nucléaire
CMS	Compact Muon Solenoid
CR	Control Region
CSC	Cathode Strip Chambers
DM	Dark Matter
EWSB	Electroweak Symmetry Breaking
EW	Electroweak
FF	Fake Factor
HTL	High Level Trigger
ICPP	Institute For Collider Particle Physics
JVT	Jet Vertex Tagger
LHC	Large Hadron Collider
LO	Leading Order
MC	Monte-Carlo
MDT	Muon Drift Tubes
MS	Muon Spectrometer
NNLO	Next to Next Leading Order
PDF	Parton Distribution Function
QCD	Quantum Chromodynamics
QFT	Quantum Field Theory
RPC	Resistance Plate Chambers
SMT	Silicon Microstrip Tracker
SR	Signal Region
SM	Standard Model
SSB	Spontaneous Symmetry Breaking
TGC	Thin Gap Chambers
TRT	Transition Radiation Tracker
2HDM	Two Higgs Doublet Model
VBF	Vector Boson Fusion

Symbol	Meaning
h	Higgs Boson
S	Higgs-like Scalar
H	Heavy Scalar
e	Electron
μ	Muon
τ	Tau
ν	Neutrino
ν_e	Electron Neutrino
ν_μ	Muon Neutrino
ν_τ	Tau Neutrino
$W^\pm$	W Boson
Z^0	Z Boson
$\mathcal{L}$	Lagrangian Density
γ	Photon
m	Mass
t	Top quark
fb^{-1}	Inverse femtobarn
g	Gluon
$\sqrt{s}$	Center of Mass Energy
η	Pseudorapidity
p_T	Transverse Momentum

Contents

Declaration i

Abstract ii

Dedication iii

Acknowledgements iv

Nomenclature v

1 Introduction 1

1.1 Outline of the Thesis 2

1.2 Author's Contribution 3

2 Theoretical Background 4

2.1 The Standard Model of Particle Physics 4

2.1.1 Elementary Particles 5

2.1.2 Bosons 5

2.1.3 Fermions 6

2.1.4 Standard Model Lagrangian 6

2.1.5 The Electroweak Theory 7

2.1.6 Spontaneous Symmetry Breaking 8

2.2 The Higgs Boson 9

2.2.1 Production Processes . . . 9
2.2.2 Decay Mechanisms . . . 11
2.3 The Madala Hypothesis . . . 12
2.3.1 Multi-lepton Final States . . . 16

3 The Large Hadron Collider and ATLAS Experiment 18
3.1 The Large Hadron Collider . . . 18
3.1.1 The Accelerator Complex and Performance . . . 19
3.2 The ATLAS Detector . . . 22
3.2.1 The Coordinate System . . . 23
3.2.2 The Inner Detector . . . 24
3.2.3 The Calorimeter System . . . 27
3.2.4 Muon Spectrometer . . . 31
3.2.5 Trigger and Data Acquisition . . . 33

4 Data Set and Simulated Events 35
4.1 Data Set . . . 35
4.2 Trigger . . . 37
4.3 Signal Samples . . . 38
4.4 Background Samples . . . 38

5 Object Reconstruction and Selection 41
5.1 Object Definition . . . 41
5.1.1 Leptons . . . 41
5.1.2 Jets . . . 44
5.2 Event Selection . . . 45

6 W+jets Background Estimation 47
6.1 The W+jets Background Process . . . 47
6.2 W+jets Fake Estimation Strategy . . . 48

6.3 W+jets Control Region 50
6.3.1 Control Region Kinematics 51
6.4 Measurements of Z+jets Fake Factor 52
6.4.1 Event Selection 55
6.4.2 Normalisation of the WZ Background 58
6.4.3 Computation of the Fake Factor 60
6.5 Closure Test 64
6.6 Flavour Composition studies 66
6.6.1 Flavour Matching Scheme 66
6.6.2 Electron Flavour Composition 68
6.6.3 Muon Fake Flavour Composition 70
6.7 Fake Estimation in the Signal Region 71

7 Summary and Conclusions **75**

A Control Plots **77**
A.1 W+jets Control Region 77
A.1.1 Electron, Anti-id 77
A.1.2 Muon, Anti-id 77
A.2 Z+jets Fake Factor Control Region 77

B Cutflows **82**
B.1 Z+jets Fake Factor 82
B.2 W+jets Control Region 84
B.2.1 Electron, Anti-id 84
B.2.2 Muon, Anti-id 84
B.3 Closure Test 89
B.4 Flavor Composition 92
B.4.1 Z+jets Sample 92

B.4.2 W+jets Sample 94

C Additional Material **98**

C.1 Fake Factor Dependency on b-jets 98

References **101**

List of Figures

2.1 The fundamental particles in SM of particle physics. 5

2.2 Mexican-hat-like Higgs potential. 9

2.3 Higgs production Feynman diagrams. 10

2.4 The Higgs boson cross sections and branching ratios as a function of its mass. 11

2.5 Feynman diagrams showing the dominant production mechanism of the heavy scalar, $gg \to H$, and its decay mode, $H \to Sh$. 14

2.6 The negative 2 times the logarithm of the profile likelihood ratios for each of the individual fit results considered in this article, overlaid with that of their combination. 17

3.1 The full CERN accelerator complex. 20

3.2 An overview of a full ATLAS detector and coordinate system. 23

3.3 A schematic cross-section of the ID. 25

3.4 An overview of the calorimeter system. 28

3.5 Accordion-shaped geometry of the LAr calorimeter. 29

3.6 Segmentation in radial depth and pseudorapidity of the tile-calorimeter modules in the central (left) and extended (right) barrels. 30

3.7 An overview of the muon system. 32

3.8 ATLAS Trigger and DAQ systems. 34

4.1 Delivered luminosity in Run 1 and Run 2 (a), total integrated luminosity during Run 2 (b), and mean number of interactions per bunch crossing for the pp collision data from 2015 to 2018 at $\sqrt{s} = 13$ TeV (c) from Ref. [69]. 36

5.1 Definitions of the d_0, $z_0 \sin\theta$ track reconstruction parameters. 42

5.2 Number of jets after the b-jet veto requirement (left) and number of b-jets after requiring at least two jets (right). 46

6.1 Example of CR definition based on lepton isolation requirements. . . . 51

6.2 W+jets control region plots for electron fakes. 53

6.3 W+jets control region plots for muon fakes. 54

6.4 Transverse mass distribution of the additional lepton. 55

6.5 The electron and muon fake quality and the d_0 significance. 56

6.6 Transverse momentum distributions of the additional lepton region with $m_{\mathrm{T}} > 50$ GeV. 59

6.7 Transverse momentum for additional electron and muon in the "id" and "anti-id" category. 61

6.8 pseudorapididty for additional electron and muon in the "id" and "anti-id" category. 62

6.9 Electron and muon FFs as a function of fake candidate transverse momentum. 64

6.10 Comparison of the flavour composition in electron fakes between the Z+jets and W+jets samples. 67

6.11 Comparison of the flavour composition in muon fakes between the Z+jets and W+jets samples. 69

6.12 The n-jets, E_{T}^{miss}, $m_{\ell\ell}$ and m_{T} distributions at the end of the selection. 72

6.13 The $\Delta\phi^{\ell\ell}$, $p_{\mathrm{T}}^{\ell\ell}$, leading lepton η and leading lepton p_{T} distributions at the end of the selection. 73

6.14 The leading and sub-leading lepton p_{T} distributions. 74

A.1 W+jets control region plots for electron fakes at the end of the selection in the $e\mu$ and μe channels. 78

A.2 W+jets control region plots for muon fakes at the end of the selection in the $e\mu$ and μe channels. 79

A.3 Transverse momentum for additional electron and muon in the "id" and "anti-id" categories. 80

A.4 Pseudorapidity for additional electron and muon in the "id" and "anti-id" categories. 81

List of Tables

2.1 Higgs production processes and corresponding cross sections. 11

2.2 Higgs decay channels for a mass of 125 GeV. Values from Ref. [26] . . . 12

2.3 ATLAS and CMS results for binding of important parameters of the Madala hypothesis. 13

2.4 The list of decay modes of the 2HDM scalars and S. 16

2.5 A summary of SM+BSM fit results and their combination. 16

3.1 Beam parameters during Run 1 (2012) and Run 2 (2018) in comparison to the nominal design values 22

3.2 An overview of the intrinsic resolutions of the ID components. 26

3.3 Energy resolutions of the constituent calorimeter sections. 28

4.1 Single lepton triggers for the 2015 to 2018 data set used in the analysis. 37

4.2 Signal samples for $H \to Sh$ analysis at different mass points. 38

4.3 Background samples and corresponding MC generators used, PDF set and underlying-event tune used for each process. 40

5.1 Electron selection. 43

5.2 Muon selection. 43

5.3 Summary event selection in the SR. 45

5.4 Cutflow showing only the main dominant background after the final pre-selection and SR requirements. 46

6.1 The muon and electron "id" and "anti-id" requirements. 49
6.2 The Z+jets FF CR cutflow. 60
6.3 Numerical results for electron FF in the full $|\eta|$ range. 63
6.4 Numerical results for muon FF in the full $|\eta|$ range. 63
6.5 The number of W+jets events in the electron fake CR with FF in the binning of p_{T}. 65
6.6 The number of W+jets events in the muon fake CR with FF in the binning of p_{T}. 65
6.7 Total number of W+jets events in CR of fake electrons and muons together with the number of W+jets events in SR. 66
6.8 Fractions of the "id" electron flavour composition for W+jets and Z+jets MC simulations. 68
6.9 Fractions of the "anti-id" electron flavour for W+jets and Z+jets sample. 68
6.10 The electron FFs in each flavour category for W+jets and Z+jets MC simulation. 68
6.11 Fractions of the muon "anti-id" flavour composition for W+jets and Z+jets. 70
6.12 Fractions of the muon "id" flavour composition for W+jets and Z+jets. 70
6.13 The muon FFs for each flavour category for W+jets and Z+jets from MC. 71
6.14 The electron and muon fake estimation in the SR $e\mu + \mu e$ channel. . . . 72

B.1 Cutflow for the Z+jets fake estimation selection in the $e\mu + \mu e$ channel. 83
B.2 The CR cutflow in the $e\mu$ channel for one id + one anti-id electron. . . 85
B.3 The CR cutflow in the μe channel for one id + one anti-id electron. . . 86
B.4 The CR cutflow in the $e\mu$ channel for one id + one anti-id muon. . . . 87
B.5 The CR cutflow in the μe channel for one id + one anti-id muon. . . . 88

B.6 The CR of one id + one anti-id electron cutflow for closure test in the $e\mu + \mu e$ channel. 90
B.7 The CR of one id + one anti-id muon cutflow for closure test in the $e\mu + \mu e$ channel. 91
B.8 The Z+jets flavour composition cutflow in the $e\mu + \mu e$ channel. 92
B.9 The Z+jets flavour composition percentage cutflow in the $e\mu + \mu e$ channel. 93
B.10 The W+jets flavour composition cutflow in the $e\mu + \mu e$ channel. 94
B.11 The W+jets fake electron flavour composition percentage cutflow in the $e\mu + \mu e$ channel. 95
B.12 The W+jets fake muon flavour composition cutflow in the $e\mu + \mu e$ channel. 96
B.13 The W+jets fake muon flavour composition percentage cutflow in the $e\mu + \mu e$. 96
B.14 The SR cutflow in the $e\mu + \mu e$ channel. 97

C.1 Electron fake yields for the Z+jets selection for the two subcategories with b-jets and without b-jets. 99
C.2 Muon fake yields for the Z+jets selection for the two subcategories with b-jets and without b-jets. 99
C.3 Electron and muon FF comparison with the b-section and b-veto. . . . 100

Chapter 1

Introduction

What are the ultimate building blocks of matter and the rules that govern them? In the last 50 years, scientists have made some very rapid progress and indeed our modern understanding underpinning the Universe can explain the phenomena from the behavior of atoms to how stars burn and this understanding is called the Standard Model, SM, of particle physics. The theory has managed to successfully describe the electromagnetic, weak and strong forces in the framework of quantum field theory. To keep the structure of the gauge interactions at high energy untouched, the SM requires for the symmetry of the system to be broken. In the 1950s, Higgs and others brought the idea of a neutral field called the Brout–Englert–Higgs, BEH, field carried by the particle called Higgs boson [1, 2], which would break the symmetry and give masses to some of the particles, like the carriers of the weak force, and this whole idea is called the BEH mechanism [3].

In 2012, the discovery of the Higgs boson, h, with a mass of approximately 125 GeV by both the ATLAS and the CMS [4, 5] collaborations has unfolded a path to perform studies for the new physics Beyond the Standard Model, BSM. One of the major studies at the Large Hadron Collider, LHC, focuses on the physics analysis BSM. The BSM model extends the SM by introducing the hypothetical scalar,[1] H, and the additional

[1]The hypothetical scalar will also be referred to as the Madala boson.

Higgs-like scalar, S [6–9]. The thesis focuses on the estimation of the fake lepton background for the BSM search in the final state with a pair of different flavor opposite-sign leptons, $e\mu$, accompanied by at least two jets, $H \to Sh \to e^{\pm}\mu^{\mp} + \geq 2$jets. The search uses the LHC full Run 2 data recorded at a center-of-mass-energy of 13 TeV recorded with the ATLAS detector, with the corresponding luminosity of 139 fb^{-1}.

One of the challenges in most physics analysis is to estimate background processes that can enter the Signal Region, SR, and be misidentified as signal events. The W+jets background enters the signal region. From the W+jets background process, there is one real lepton from W and one fake, which originates from the jet misidentified as a lepton. In the $H \to Sh \to e^{\pm}\mu^{\mp} + \geq 2$jets analysis, the fake factor method is used for fake lepton estimation.

1.1 Outline of the Thesis

The main focus of this thesis is to estimate the contribution of the W+jets process as a background process that enters the SR in $H \to Sh$ analysis when a jet is misidentified as a lepton. The W+jets process is produced at a very large rate around 10^4 larger than in the $H \to Sh$ signal. MC simulations for the rate of W+jets background have large uncertainties. Hence, we employ a data-driven method to perform the W+jets background estimation [10]. This technique is called the fake factor method and it is further explained in Chapter 6. Chapter 2 focuses on the theoretical background of the SM: describing the building blocks of nature, fundamental forces, symmetries, the discovery of h, and lastly the Madala hypothesis. Afterward, the description of the LHC and the ATLAS detector, its components and particle identification are explained in Chapter 3. The data sets and simulated events for the $H \to Sh \to e^{\pm}\mu^{\mp} + \geq 2$jets analysis are described in Chapter 4, followed by the object reconstruction and event selection in Chapter 5. In Chapter 6 the estimation method of the fake lepton

background is fully described together with the obtained results. Finally, Chapter 7 gives a summary of the future work, and conclusions.

1.2 Author's Contribution

The $H \rightarrow Sh$ analysis work is performed by the University of Witwatersrand, School of Physics and Institute for Collider Particle Physics, ICPP. This is studied in collaboration with the ATLAS experiment at CERN. The author has been involved in the estimation of the fake lepton background in the $H \rightarrow Sh \rightarrow e^{\pm}\mu^{\mp} + \geq$ 2jets analysis. The estimation is performed using the fake factor method which is fully described in Chapter 6 together with the produced results with the ATLAS full Run 2 data-set.

Chapter 2

Theoretical Background

An outline of the elementary particles of the SM physics will be described in context, along with approaches that build up the theoretical background of the thesis. The reader is referred to the references [11, 12].

2.1 The Standard Model of Particle Physics

The SM [13–16] is a theory that describes the way that particles interact with one another. It describes the electromagnetic, strong nuclear and weak nuclear forces as being mediated by certain fundamental particles. That is, particles interact with one another by exchanging the fundamental particles and this results in the particles repelling or attracting one another. The electromagnetic force describes how electrons interact with the nuclei and it holds atoms and molecules together. The strong force holds together the quarks inside the protons and neutrons. The weak force is what causes β-decay of atoms and allows the build-up of the radioactive isotope, which is a creation of deuterium that helps the sun to burn. From the SM, nature is described using the context of Quantum Field Theory, QFT [17]. The states at which fields are quantised represents the elementary particles. The SM theory also anticipates the existence of h, which is discussed in Section 2.2.

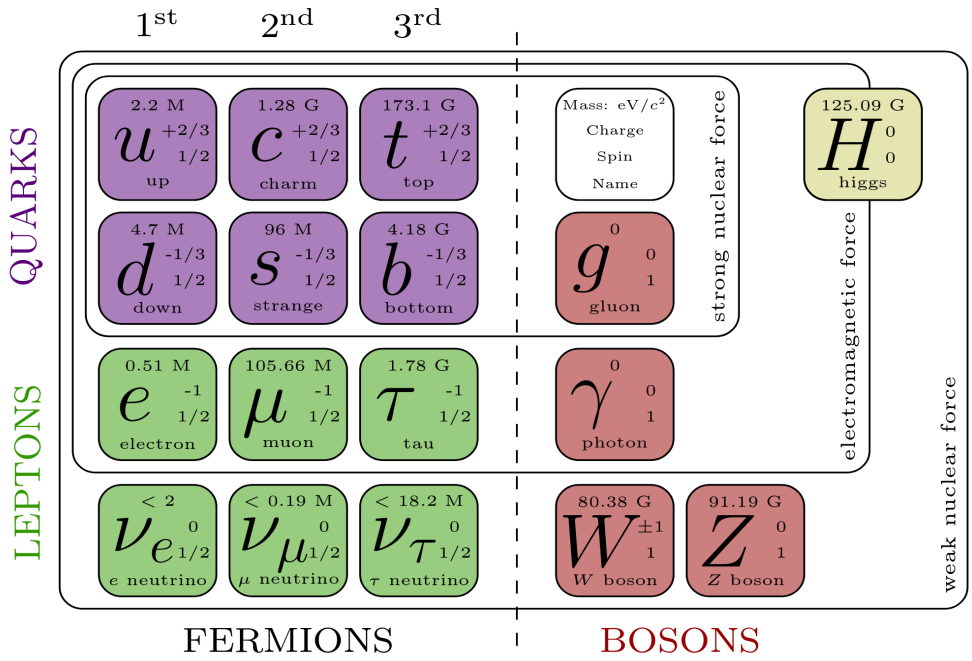


Figure 2.1: The SM of elementary particles, with three generations of matter in the first three columns in purple and green. The gauge bosons in the fourth column in red and the Higgs boson at the far top right corner in yellow.

2.1.1 Elementary Particles

In the SM only two types of elementary particles have a special meaning and those are fermions and bosons, unlike all the particle names, which are related to charge, mass or the force they experience. The fact that they tie together and distinguish fermions and bosons is the idea of subatomic spin. All bosons have a spin that is an integer multiple of 1, spin: 0, 1, 2, etc and, in contrast, all fermions have half-integer of spin: 1/2, 3/2, etc. Figure 2.1 shows all the fundamental particles in the SM of particle physics and the corresponding properties.

2.1.2 Bosons

The bosons include the four forces carrying particles of the SM and these being photons, gluons, W and the Z-bosons. These particles are the gauge bosons of the SM. The *graviton* [18] is proposed in the theory but not yet discovered and it is responsible

for gravitational attraction. The gluon is an exchange particle responsible for the strong interaction, or strong force. The exchange of gluons between quarks binds them together to form larger particles, such as protons and neutrons. Photons are force carriers of the electromagnetic field and W and Z-bosons are force carriers involved in the weak force. The interaction of these bosons with the Higgs bosos is responsible for their masses.

2.1.3 Fermions

Quarks and leptons are fermions. Quarks make up, amongst other things, the protons, and neutrons in the nucleus. The combination of the up and down quarks produces neutrons and protons, two downs and one up for a neutron and two ups and one down for a proton. Leptons include electrons and neutrinos. Electrons are fundamental parts of all atoms occupying energy levels around the nucleus. Neutrinos prevail the universe streaming out from the sun as the consequence of all nucleus actions and reactions. Each of these particles has an anti-particle. The formations of quarks are grouped as: "up", and "down", "charm" and "strange", lastly "top" and "bottom". In the lepton family, we have an electron, e, muon, μ, tau, τ and corresponding neutrinos, electron-neutrino, ν_e, muon-neutrino, ν_μ, tau-neutrino, ν_τ.

2.1.4 Standard Model Lagrangian

The properties of all fundamental particles and their interactions are carried out by the mathematical quantity referred to as the SM Lagrangian. The Lagrangian of the SM is invariant under space-time dependent continuous internal transformations of the group $SU(3) \times SU(2) \times U(1)$, where $U(1)$ is the gauge group that refers to the electromagnetic force, $SU(2)$ to the weak force and $SU(3)$ to the strong force. Equation 2.1 shows the SM Lagrangian and its components:

$$\mathcal{L}_{SM} = \tfrac{1}{4} W_{\mu\nu} \cdot W^{\mu\nu} - \tfrac{1}{4} B_{\mu\nu} B^{\mu\nu} - \tfrac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a$$

$$+ \bar{L}\gamma^{\mu}(i\partial_{\mu} - \tfrac{1}{2}g\tau \cdot W_{\mu} - \tfrac{1}{2}g\prime Y B_{\mu})L + \bar{R}\gamma^{\mu}(i\partial_{\mu} - \tfrac{1}{2}g\prime Y B_{\mu})R$$

$$+ \tfrac{1}{2}\,|(i\partial_{\mu} - \tfrac{1}{2}g\tau \cdot W_{\mu} - \tfrac{1}{2}g\prime Y B_{\mu})\phi|^2 - V(\phi)$$

$$+ g\prime\prime(\bar{q}\gamma^{\mu}T_a q)G^a_{\mu} + (G_1\bar{L}\phi R + G_2\bar{L}\phi_c R + h.c.), \ [3] \tag{2.1}$$

where the 1^{st} term of the equation represents the kinetic energies and self-interactions of the gauge bosons, 2^{nd} term represents kinetic energies and electroweak interactions of fermions, 3^{rd} term represents $W^{\pm}, Z, \gamma$ and masses of h and couplings and 4^{th} term represents the interactions between quarks and gluons, then fermion masses and couplings to h.

2.1.5 The Electroweak Theory

With the experimental observation of W and Z-bosons [19], these particles provided the necessary evidence to unify the electromagnetic force and weak nuclear force into a single force called the "electroweak force". The unification between the electromagnetic and weak interaction is considered and the gauge group that should be considered is not just $SU(2)$, but a direct product of $SU(2)$ and $U(1)$, $SU(2)_L \times U(1)_Y$, where L represents the fields on the left and Y is the interaction or the charge corresponding to that as hyper charge. Photons are represented by a mixture of $SU(2)_L$ and $U(1)_Y$ by gauge fields. This also tells us that the corresponding charges are related and this relation is given by:

$$Q = I_3 + \frac{Y}{2}, \tag{2.2}$$

where Q is the charge of electromagnetic interaction, I_3 is the 3^{rd} projection of isospin, which is charge corresponding to $SU(2)$. The $SU(2)_L \times U(1)_Y$ gauge invariance breaks

[3] G_1 and G_2 are Yukawa couplings. The charge conjugate of the Higgs is defined by $(\phi_c)^{\alpha} = \epsilon^{\alpha\beta}(\phi^y)\beta$, where α, β are $SU(2)$ (weak) indices. This is needed to couple the left-handed quark doublet to the right-handed up-type quarks in order to give a hyper charge neutral term in the Lagrangian. Note that using ϕ^y doesn't work for this purpose, since it doesn't have the correct $SU(2)$ transformation properties to form an $SU(2)$ singlet with $\bar{L}$.

as the massive $W^{\pm}$, Z-bosons and massless photon are introduced in the SM Lagrangian. The gauge symmetry excludes the masses for the gauge bosons. From the BEH mechanism, the gauge symmetry needs to be broken to generate masses. This theory predicts the existence of the h field, which permeates all of the space and acts on particles with mass.

2.1.6 Spontaneous Symmetry Breaking

For the Spontaneous Symmetry Breaking, SSB [20], the $U(1)$ symmetry can be considered as a clean example to look at. From the $U(1)$ model the complex scalar field ϕ has a Lagrangian density, which is invariant under $U(1)$ transformation:

$$\phi(x) \rightarrow e^{i\alpha}\phi(x), \tag{2.3}$$

where α is a constant and it is space-time independent parameter. The Lagrangian density is expressed as follows:

$$\mathcal{L} = \partial_{\mu}\phi\partial^{\mu}\phi^{\star} - V \tag{2.4}$$

with

$$V(\phi) = \mu^2\phi\phi^{\star} + \lambda(\phi\phi^{\star})^2, \tag{2.5}$$

where μ is the mass from the mass term $\phi\phi^{\star}$, and λ is a coupling constant from the interaction of the field with itself. The potential carries the parabolic shape in a case of a square mass greater than zero, $\mu^2 > 0$, and maximum energy state will be that of $\phi(x) = 0$. If the mass is chosen to be negative, $\mu^2 < 0$, the potential takes the Mexican hat shown in Figure 2.2 . Hence, in the ground state $\phi(x) = v$, where $v = \sqrt{\frac{-\mu^2}{\lambda}}$. In this case the field $\phi(x)$ obtains the non-zero vacuum expectation value and the $U(1)$ symmetry is considered to be spontaneously broken. The SSB theory was validated during the discovery of the h particle, which is discussed in Section 2.2.

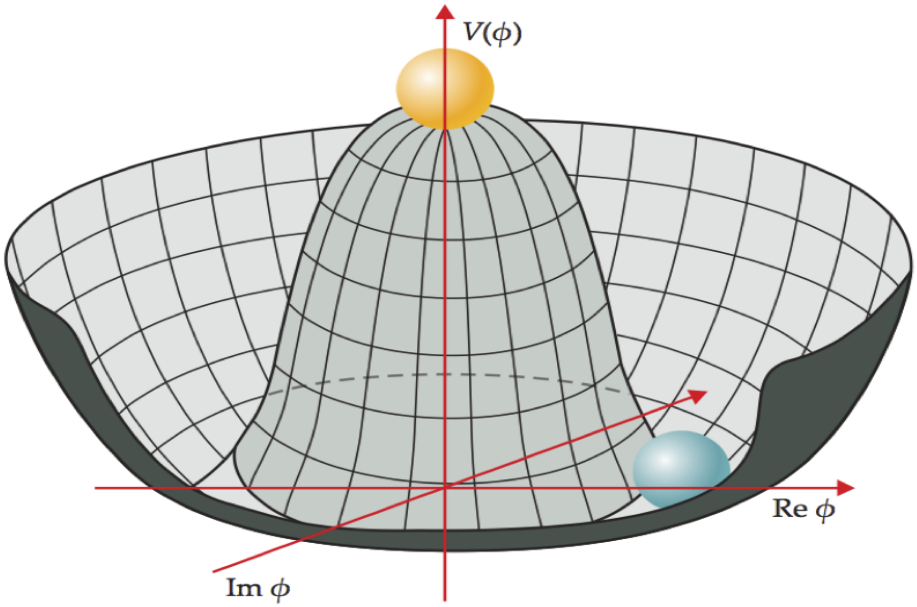


Figure 2.2: Mexican hat shape of Higgs potential that leads to "symmetry breaking".

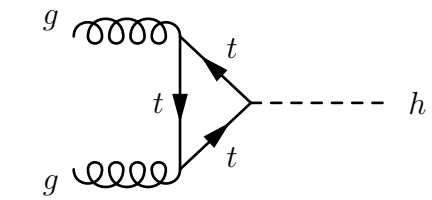
2.2 The Higgs Boson

This section briefly describes the production of the h boson and its decay channels in proton-proton collisions, which are well summarised in Refs [21–23].

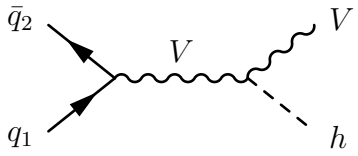
2.2.1 Production Processes

There are several production mechanisms in which the h boson can be observed experimentally, yet the probability of its production in any collision is expected to be very small. A main part of the general Higgs boson study is its production channels. The probability of producing the Higgs is dependent on the cross sections of its various production channels. The calculation of these cross sections is complicated due to the non-fundamental character of the proton, as it is composed of more elementary components, i.e. quarks and gluons. Hence, they are computed at the parton level[3] and then deduced to proton level using parton distribution functions. The Higgs pro-

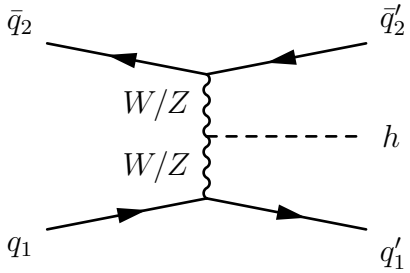
[3]Here, hadrons consist of a number of components that looks like points, identified as quarks and gluons, namely partons.



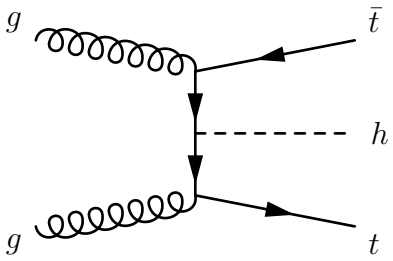
(a) Gluon fusion.



(b) Vector boson associated production.



(c) Vector boson fusion.



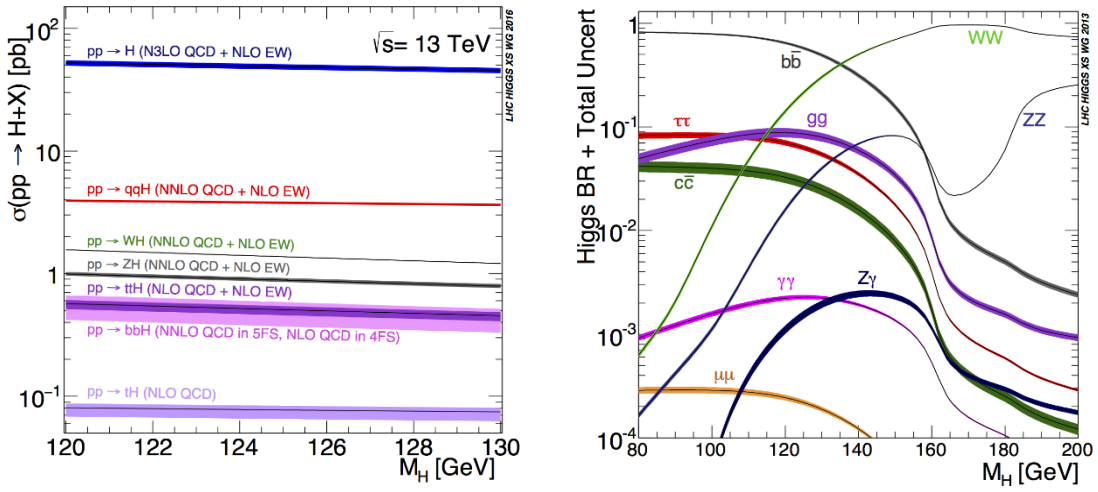
(d) Top associated production.

Figure 2.3: Feynman diagrams for the dominant Higgs boson production modes in high energy pp collisions.

duction mechanisms of interest, namely, the gluon fusion (ggF), vector boson fusion (VBF), associated Vh production (Wh and Zh), Top fusion ($t\bar{t}h$ and th) are shown in Figure 2.3. The most dominant production process is the ggF process in which two gluons fuse and produce h via a loop of virtual heavy quarks (top and bottom quarks), shown in Figure 2.3 (a). Figure 2.3 (b) shows the case whereby h is emitted by a W boson through a process referred to as Vh. In this process, a quark and anti-quark collide and results in a W or Z boson which can emit h if it carries a sufficient amount of energy. Another common process is the VBF process, which is initiated by two quarks or two anti-quarks colliding and exchanging a virtual W or Z boson, which emits h as demonstrated in Figure 2.3 (c). The least common process is the top fusion process in which h is produced in association with a heavy quark and anti-quark pair and $t\bar{t}h$ production is shown in Figure 2.3 (d). The cross sections for various production channels of the Higgs boson are shown in Table 2.1. Figure 2.4 on the left shows the cross section values for different production channels as a function of the Higgs mass, m_h, with the following dominant productions: ggF, VBF, Wh and Zh also the tth.

Table 2.1: The main production channels of the Higgs boson at the LHC, with the corresponding cross sections, for a Higgs mass of 125 GeV [24].

Production	Cross sections	Rel. Uncertainty
ggF	48.6	+4.6%, -6.7%
VBF	3.78	+2.2%, -2.2%
Wh	1.37	+2.6%, -2.6%
Zh	0.88	+4.1%, -3.5%
tth	0.50	+6.8%, -.9.9%



(a) (b)

Figure 2.4: Cross sections for the h boson production as a function of its mass for a low mass range, with bands representing theoretical uncertainties with the observed mass, m_h, of 125 GeV, left, and the branching ratios as a function of its mass for a wide mass range [25], right.

Figure 2.4 on the right shows the branching ratios of h as a function of its mass for various decay mechanisms.

2.2.2 Decay Mechanisms

Due to the proportionality of the coupling strength of the Higgs to the respective mass of the particle it interacts with, the most probable decay paths are the ones consisting of the heaviest possible particles. The main requirements for the decay products are to satisfy the zero-spin and neutral-charge character of the Higgs and conserve the total energy during the decay. The decay can occur either fermionically or bosonically and

Table 2.2: Higgs decay channels for a mass of 125 GeV. Values from Ref. [26]

Decay Channel	Branching Ratio	Rel. Uncertainty
$h \to b\bar{b}$	5.77×10^{-1}	+3.2%, +3.3%
$h \to WW^*$	2.15×10^{-1}	+4.3%, +4.2%
$h \to \tau^+\tau^-$	6.32×10^{-2}	+5.7%, +5.7%
$h \to ZZ^*$	2.64×10^{-2}	+4.3%, +4.1%
$h \to \gamma\gamma$	2.28×10^{-3}	+5.0%, +4.9%

the probability of each decay channel depends, as mentioned, on the mass of the decay products, but also on the type of coupling between them and the Higgs. The observed cross sections of each channel can be used to study how the Higgs couples with the SM particles. Around a mass of 125 GeV, the $b\bar{b}$ is found to have the highest branching ratio. From theory, the $b\bar{b}$ channel has an expected branching ratio of 58% and has been observed through the associated production of h with a vector boson, Vh [27]. The Higgs boson is likely to decay to a pair of W-bosons with a branching ratio of 22% followed by the $\tau\bar{\tau}$ decay with the expected branching ratio of 6%. The decay of h into a Z-boson pair has a branching ratio of 2.6%. The h decay to photons is very rare with approximately 0.3% branching ratio. Table 2.2 gives a summary of the most outstanding decay channels for the study of h.

2.3 The Madala Hypothesis

Since the discovery of the Higgs boson, attention has been taken to experimental measurements to analyse the compatibility of data with SM predictions and theoretical work to understand how any new physics would be revealed. Both ATLAS and CMS took measurements in the Run 1 dataset and some inconsistencies in the results were observed. The University of the Witwatersrand ICPP in collaboration with the Harish-Chandra Research Institute, HRI, has developed a model that provides an explanation for multiple features observed in Run 1 from both ATLAS and CMS. This model is called the Madala hypothesis [6–9]. The word "Madala" comes from a South African

Table 2.3: ATLAS and CMS results for binding of important parameters of the Madala hypothesis [6].

Experiment	Publication	Results
ATLAS	Fiducial cross section measurements on $h \to \gamma\gamma$ [28] and $h \to ZZ^* \to 4\ell$ [29]	Differential Higgs boson p_T spectra
CMS	Fiducial cross section measurements on $h \to \gamma\gamma$ [30] and $h \to ZZ^* \to 4\ell$ [31]	
ATLAS	Limits on $H \to hh \to bb\tau\tau, \gamma\gamma WW^*$, $\gamma\gamma b\bar{b}$ and $b\bar{b}b\bar{b}$ [32]	Di-Higgs boson resonance searches
CMS	Limits on $H \to hh \to \gamma\gamma b\bar{b}$ [33], $b\bar{b}\tau\tau$ [34] and multi-lepton [35]	
ATLAS	Limits on $h \to \gamma\gamma$ [36], Measurements on multi-lepton decay channels [37] and $h \to b\bar{b}$ [38]	Top associated Higgs boson production
CMS	Limits on $h \to \gamma\gamma$ and $h \to b\bar{b}$ [39]	
ATLAS	Limits on $H \to WW$ [40] and $H \to ZZ$ [41]	$H \to VV$
CMS	Limits on $H \to WW$ and $H \to ZZ$ [42]	

Zulu language and it means "old one". The Madala hypothesis introduces the existence of the hypothetical heavy scalar particle, H, named the Madala boson and the additional Higgs-like scalar, S. The phenomenology of these scalars is discussed in Ref [7]. The mass ranges for H and h are as follows: mass of h, $m_h = 125$ GeV, mass of H, $2m_h < m_H < 2m_t$, where m_t represents the mass of a top quark. S is assumed to have a mass $m_h \leq m_S \leq m_H - m_h$, allowing for new processes associated to the decays of H. Table 2.3 shows the ATLAS and CMS measurements, which were considered in the designing of the Madala hypothesis. Results from the Higgs boson $p_{T,h}$ spectra, di-Higgs boson searches, top associated Higgs production and VV decays were used to assist in constraining the relevant parameters of the hypothesis. The excess was observed in the Run 1 data from the transverse momentum of Higgs, $p_{T,h}$, spectrum and some inconsistencies were observed with what the SM predicts. Hence, the $p_{T,h}$ spectrum possibly indicates that the production of h may be accompanied by S and this opens a window for the new physics beyond SM. A heavy scalar H, which can

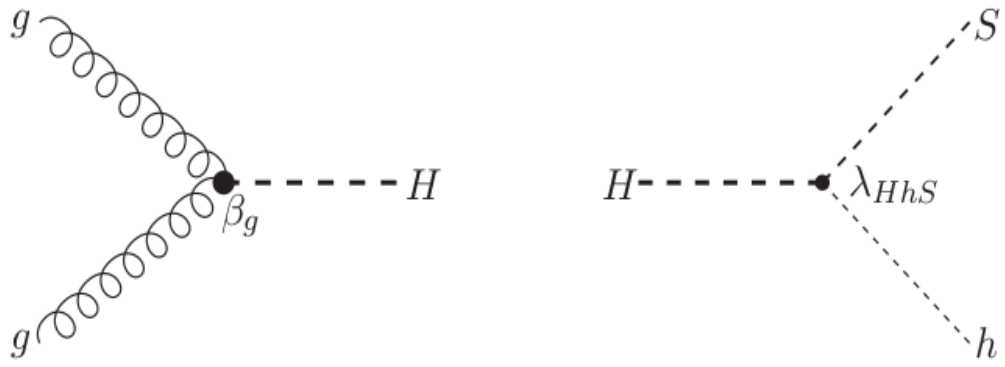


Figure 2.5: Feynman diagrams showing the dominant production mechanism of the heavy scalar, $gg \to H$, and its decay mode, $H \to Sh$.

decay to at least one h together with S is introduced. That is $H \to Sh$ channel, which is the main focus in thesis. Figure 2.5 shows the production mechanism of the heavy scalar, $gg \to H$, and its decay mode, $H \to Sh$.

The SM is simply extended by adding to the Lagrangian, another term, $\mathcal{L}_{BSM}$, with all the BSM interactions and states:

$$\mathcal{L} = \mathcal{L}_{\mathcal{SM}} + \mathcal{L}_{BSM}. \tag{2.6}$$

The Lagrangian sectors in the BSM term are defined as:

$$\mathcal{L}_H = -\frac{1}{4}\beta_g \kappa_{hgg}^{SM} G_{\mu\nu} G^{\mu\nu} H + \beta_V \kappa_{hVV}^{SM} V_\mu V^\mu H, \tag{2.7}$$

$$\mathcal{L}_Y = -\frac{1}{\sqrt{2}}[y_{ttH}\bar{t}tH + y_{bbH}\bar{b}bH], \tag{2.8}$$

$$\mathcal{L}_T = -\frac{1}{2} v \lambda_{Hhh} H h h, \tag{2.9}$$

$$\mathcal{L}_Q = -\frac{1}{4}\lambda_{HHhh} H H h h. \tag{2.10}$$

The three terms $\mathcal{L}_Y$, $\mathcal{L}_T$ and $\mathcal{L}_Q$ represents Yukawa, trilinear, and quartic interactions. The simplified Lagrangian used to describe the production of H is represented by

Equation 2.7. The 1^{st} term of the Equation 2.7 allows for the gluon fusion production mode of H, where β_g denotes the scale factor responsible for the production rate of H through the gluon-fusion and its best fit was at 1.5 ± 0.6 from the Run 1 results [6]. κ_{hgg}^{SM} represents the SM coupling to Higgs gluon fusion and v is the vacuum expectation value, vev = 246 GeV. From the 2^{nd} term in Equation 2.7, the factor β_V represent the couplings of the VVH. The Yukawa couplings amongst H and heavy quarks are indicated in Equation 2.8 and also in Equation 2.9. Equation 2.10 shows the feasible trilinear and quartic couplings amongst h and H. The S boson is assumed to be produced through the decay of H. In principle, S can be included as a singlet scalar that interacts with H and the SM Higgs boson. This would allow the H to produce S bosons in the $H \to SS$ and Sh decay modes. The $H \to Sh$ decay mode is assumed to have a branching ratio of 100%. The Lagrangians for this effective interaction are given by:

$$\mathcal{L}_{HhS} = -\frac{1}{2}v[\lambda_{hhs}hhS + \lambda_{hSS}hSS + \lambda_{HHS}HHS + \lambda_{HSS}HSS + \lambda_{HhS}HhS] \quad (2.11)$$

and

$$\mathcal{L}_S = \frac{1}{4}\kappa_{Sgg}\frac{\alpha_s}{12\pi v}SG^{a\mu\nu}G^a_{\mu\nu} + \frac{1}{4}\kappa_{S\gamma\gamma}\frac{\alpha}{\pi v}SF^{\mu\nu}F_{\mu\nu} + \frac{1}{4}\kappa_{SZZ}\frac{\alpha}{\pi v}SZ^{\mu\nu}Z_{\mu\nu}$$

$$+ \frac{1}{4}\kappa_{SZ_\gamma}\frac{\alpha}{12\pi v}SZ^{\mu\nu}F_{\mu\nu} + \frac{1}{4}\kappa_{SWW}\frac{2\alpha}{\pi S_w^2 v}SW^{+\mu\nu}W^-_{\mu\nu} - \sum_f \kappa_{Sf}\frac{m_f}{v}Sf\bar{f}. \quad (2.12)$$

In Equation 2.11, S is given a vacuum expectation value and couples to the scalar sectors while in Equation 2.12 it is given Higgs-like branching ratios by fixing the parameters in the Lagrangian. Furthermore, the couplings are re-scaled globally to suppress direct production of S. Fixing the branching ratios of S reduces the number of free parameters in the model. A deeper understanding into effective decay modes can be achieved by introducing scalars, H and S, into a two Higgs doublet model, 2HDM, as described in Ref [7]. The results are given by five new scalars: heavy scalar H,

Table 2.4: The list of decay modes of the 2HDM scalars and S [7].

S.No.	Scalars	Decay modes
D.1	h	$b\bar{b}$, $\tau^+\tau^-$, $\mu^+\mu^-$, $s\bar{s}$, $c\bar{c}$, gg, $\gamma\gamma$, Z_γ, W^+W^-, ZZ
D.2	H	D.1, hh, SS, Sh
D.3	A	D.1, $t\bar{t}$, Zh, ZH, ZS, $W^\pm H^\pm$
D.4	$H^\pm$	$W^{\pm h}$, $W^{\pm H}$, $W^\pm S$
D.5	S	D.1, $\chi\chi$

Table 2.5: A summary of SM+BSM fit results and their combination [9].

Selection	Best-fit β_g^2	Significance
ATLAS Run 1 SS $\ell\ell$ and $\ell\ell\ell + b$-jets	6.51 ± 2.99	2.37σ
ATLAS Run 1 OS $e\mu + b$-jets	4.09 ± 1.37	2.99σ
ATLAS Run 2 SS leptons+b-jets	2.22 ± 1.19	2.01σ
CMS Run 2 SS leptons+b-jets	1.41 ± 0.80	1.75σ
CMS Run 2 OS $e\mu$	2.79 ± 0.52	5.45σ
ATLAS Run 2 SS $\ell\ell$ and $\ell\ell\ell + b$-jets	5.42 ± 1.28	4.06σ
ATLAS Run 2 tri-lepton $+E_{\mathrm{T}}^{miss}$	9.05 ± 3.35	2.25σ
CMS Run 2 tri-lepton $+E_{\mathrm{T}}^{miss}$	9.70 ± 3.88	2.36σ
Combination	2.92 ± 0.35	8.04σ

Higgs-like scalar S, pseudoscalar A, and two charged Higgs boson, $H^\pm$. The number of possible final decay states of the five scalars and the Higgs boson are shown in Table 2.4.

2.3.1 Multi-lepton Final States

The Madala hypothesis can accommodate other discrepancies in data on other final states, particularly multi-lepton final states. Within the context of new possibilities, the Madala hypothesis can be used to scrutinize the various ATLAS and CMS di-lepton invariant mass spectra. An excess was expected in events involving two leptons plus jets and E_{T}^{miss} [6, 7]. To understand this excess, a research was conducted based on the decay of S to di-lepton final states [8]. The BSM di-lepton signal was compared with the ATLAS and CMS data and the optimal performance was determined by changing the BSM standardization and adapting the SM+BSM approximation to the results using several measurements [8]. In a continued effort to try and comprehend the anomalies in multi-lepton final states while testing the Madala hypothesis, we consider the same

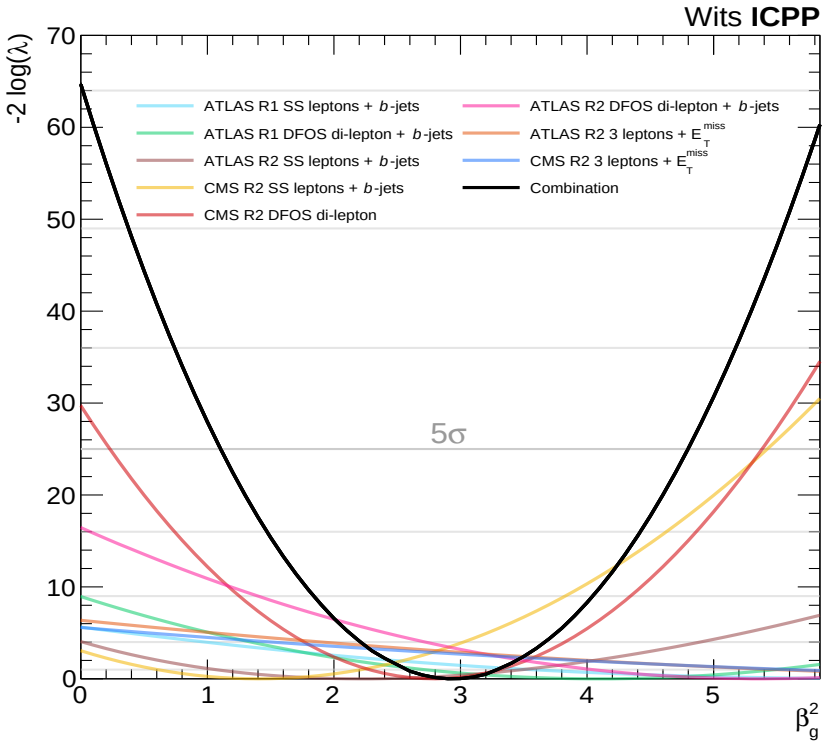


Figure 2.6: The negative 2 times the logarithm of the profile likelihood ratios for each of the individual fit results considered in this article, overlaid with that of their combination [9].

simplified model where H and S are Higgs-like and only consider one parameter [9]. The masses are set respectively to $m_H = 270$ GeV and $m_S = 160$ GeV. The aim is to seek to decide how consistent such differences are. Simultaneous fits are made for all measurements and the profile likelihood ratio is shown for each result and their combination in Table 2.5 and Figure 2.6. The best fit is $\beta_g^2 = 2.92 \pm 0.35$, which represents a deviation from the SM of 8.04 σ. This is an indication of the failure of the current tools to explain multi-lepton data and how our simpler model can catch the effect with one parameter only. Also, the idea of integrating the Madala hypothesis into a 2HDM+S model is now in good shape to be further investigated by experiments in multi-lepton final states.

Chapter 3

The Large Hadron Collider and ATLAS Experiment

The Large Hadron Collider, LHC [43], is a hadron-hadron synchrotron collider at the Conseil Européen pour la Recherche Nucléaire, CERN, near Geneva, that produces collisions at the highest center-of-mass energy, $\sqrt{s}$, ever achieved. It collides hadrons with a different momentum transfer between the partons during each collision and these provide an opportunity to search for new physics beyond the SM across a wide range of energies. This chapter provides a summary of the LHC and gives a detailed description of one of the main LHC experiments: the ATLAS detector [44–47], which provides details of the major systems from which it is constructed.

3.1 The Large Hadron Collider

The LHC is a 27 km superconducting particle accelerator situated about 100 m underground. It is built to produce pp collisions at a center-of-mass energy of up to $\sqrt{s} =$ 14 TeV, however, it was able to produce collisions at 7, 8 TeV in the year 2011-2012 and 13 TeV (2015-2018). In addition, the LHC provides the lead ion, Pb-Pb, and p-Pb collisions. Nonetheless only the pp collisions are considered throughout this the-

sis. Proton-proton beams at the LHC are circulated in opposites directions through separate beam pipes and collided at four interaction points. The beam pipes are maintained at ultra-high vacuum, $\sim 10^{-9}$ mbar, with 1232 dipole magnets that bend the beams into their circular orbit, while the 392 quadrupole magnets focus the beam at the interaction points to ensure collision between the two beams. The dipole magnets produce 8.4 T of the magnetic field by using Niobium Titanium, NbTi, superconducting magnets and such magnets are cooled to $\sim$ 1.9 K using liquid Helium.

The LHC uses sixteen superconducting radio frequency, RF, cavities to accelerate protons and these RF cavities run at a frequency of 40 MHz with an acceleration gradient of 16 MV/m, its beams can handle 2808 of bunches per proton beam where on average, each bunch contains 1.15×10^{11} protons. The LHC requires a very high intensity to observe signatures of new physics. In order to observe these events, a rapid reaction, fine segmented and radiation hard detectors are required. The LHC ring has four main experiments situated at the four interaction points. ATLAS and CMS [48] are both general-purpose detectors with extensive semiconductor-based tracking systems, large calorimeter coverage, and efficient muon detectors. They were designed to discover and measure the Higgs boson and for searches of physics beyond the SM, however, they also pursue a major physics program in SM precision measurements. ALICE, A Large Ion Collider Experiment [49], is designed to study heavy-ion collisions, implementing a large Time Projection Chamber, TPC, along with complementary sub-detectors. LHCb [50] is a one-sided forward arm spectrometer designed to accurately measure the decay products of b-hadron, which are often expected to be formed in the forward region.

3.1.1 The Accelerator Complex and Performance

The accelerator complex at LHC consists of a series of connected accelerators, which gradually increase in size and are built to accelerate particles until they are deposited

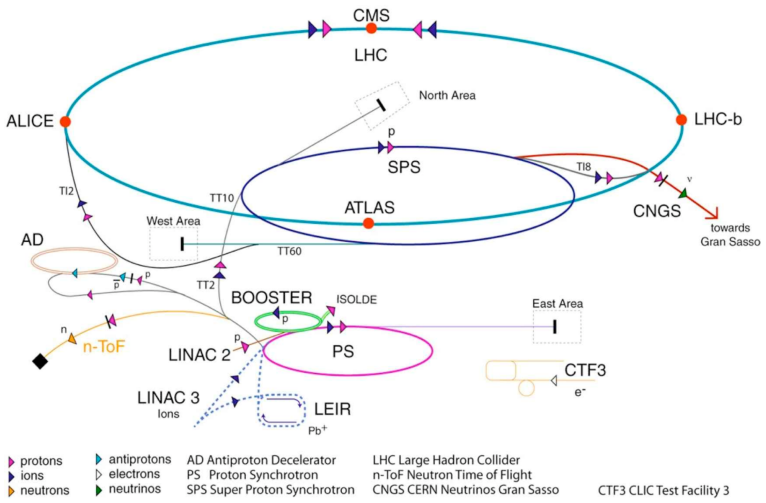


Figure 3.1: The CERN accelerator complex comprising of the multi-stage system of injectors that are used to increase the proton energy. A subset of the many experiments supported by these accelerators is also presented [51].

into the final stage of the chain. Figure 3.1 shows the entire CERN accelerator complex. Protons from the ionization of hydrogen are pumped into the first linear accelerator, LINAC-2, and accelerated to 50 MeV. They are injected into the Proton Synchrotron Booster, PSB, and leave to the Proton Synchrotron, PS, with an energy of 1.4 GeV. The proton beam is then accelerated up to 26 GeV in the PS and transferred to the Super Proton Synchrotron, SPS, and increased in energy again up to 450 GeV. Eventually, the proton bunches are pumped into the LHC in both directions and accelerated to their maximum energy. The bunches are then compressed to 1.71 ns with a bunch space of 25 ns. The Run 1 of the LHC data collection began in March 2010 and ended at the beginning of 2013 with the initial beam collision occurring at a center-of-mass of energy of $\sqrt{s} = 7$ TeV, 2010-2011, then up to 8 TeV, 2012-2013. A major upgrade has been launched for two years and the LHC was restarted in early 2015. LHC Run 2 extended from 2015 until late 2018, with the center-of-mass energy of $\sqrt{s} = 13$ TeV. Besides the beam energy, an important parameter in colliders is called the luminosity. $\mathcal{L}$, which is a measure of the intensity of the collision rate and directly relates the event rate to the cross-section, σ, of a particular physics process:

$$\frac{dN}{dt} = \sigma \cdot \mathcal{L}. \tag{3.1}$$

By integrating Equation 3.1 with respect to time, the total number of events N produced is then obtained.

$$N = \sigma \cdot \int \mathcal{L} dt = \sigma \cdot L, \tag{3.2}$$

where L is the integrated luminosity. The instantaneous luminosity is calculated from measured parameters of the accelerator, given by:

$$\mathcal{L} = \frac{N^2 \gamma_r f_{rev}}{4\pi\beta^* \varepsilon_n} n_b \cdot F, \tag{3.3}$$

where N refers to the number of particles per bunch, γ_r to the relativistic Lorentz parameter, f_{rev} to the revolution frequency, β^* to the beta function at the point of collision, ε_n to the uniform transverse beam emittance, n_b to the number of bunches per beam and F to the luminosity reduction factor due to the beam crossing angle at the point of interaction. Usually, the number of protons per bunch is 1.15×10^{11}, and the LHC is capable of circulating up to 2808 bunches. The revolution frequency is about 11.2 kHz and is defined by the fixed circumference of the LHC ring and by the observed velocity of the circulating protons. The relativistic gamma factor is about 7000, which accounts for the relativistic velocity at which protons are traveling. The normalised transverse beam emittance and the beta function are related to the size and shape of the beam at the interaction point. The geometric reduction factor is the angle at which the beams collide. Table 3.1 shows the different beam parameters during LHC Run 1 and Run 2 compared to the nominal design values. To increase the production rate of events, the luminosity is set high by increasing the bunch population and decreasing the bunch spacing. This leads to multiple inelastic *pp* collisions at the

Table 3.1: Beam parameters during Run 1 (2012) and Run 2 (2018) in comparison to the nominal design values [43].

Parameter	Symbol [Unit]	Design	Run 1	Run 2
Beam energy	[TeV]	7.0	4.0	6.5
Bunch spacing	[ns]	25	50	25
No. bunches	nb	2808	1380	2556
Bunch population	N_b $[10^{11}p]$	1.15	1.6	1.4
Revolution frequency	f_r [kHz]	11.2	11.2	11.2
Normalized emittance	ε_n [μm]	3.73	2.0	2.2
Beta function	β^* [cm]	55	60	25
Crossing Angle	θ_c [μrad]	285	290	260
Peak luminosity	$\mathcal{L}$ $[10^{34}$ cm^{-2}s$^{-1}]$	1.0	0.7	2.6

same time during each bunch crossing in a phenomenon referred to as "pile-up". Pile-up represents the average number of interactions per bunch crossing $\langle\mu\rangle$ [52] .

3.2 The ATLAS Detector

ATLAS, A Toroidal LHC ApparatuS, is a detector designed to tackle a wide range of studies in physics and have the ability to reconstruct a wide range of particles, leptons, jets, missing transverse energy, b quarks, etc. The ATLAS detector is 25 m in diameter, 46 m long, and weighs 7000 tonnes. It is a 4π detector that covers almost the entire solid angle with active instrumentation, aside from a small angular range near the beam pipe's entry and exit point. ATLAS follows a forward-backward symmetric cylindrical geometry and it is constructed from the main barrel region with two enclosed end-caps on each end. The apparatus is typical of collider experiments and provides integrated, complementary subsystems for specific tasks. Each subsystem is designed to be sensitive to the signatures of different traversing particles, in order to achieve the most wholesome picture of particle detection possible. The detector is designed so that particles of a different type, energy, and travel direction would produce different types of signatures on the basis of which they could be reconstructed. Figure 3.2 shows a schematic overview of the full ATLAS detector. The overall structure in Figure 3.2

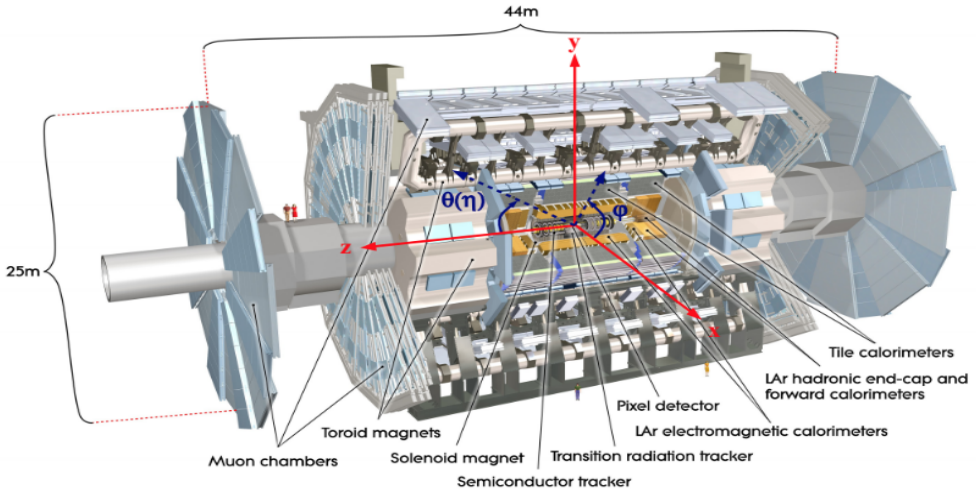


Figure 3.2: An overview of a full ATLAS detector and coordinate system [44–47].

consists of concentric cylinders around the interaction point. The apparatus can be split in three detector systems:

- The Inner Detector, ID [53, 54], composed of the Pixel Detector, the Semi-Conductor Tracker, SCT, and the Transition Radiation Tracker, TRT.
- The calorimeters [55], electromagnetic, hadronic and the forward calorimeter.
- The Muon spectrometer

A short overview of the purpose and technical characteristics of each part will follow.

3.2.1 The Coordinate System

The origin of the coordinate system is defined at the interaction point of the beams. In the cartesian coordinates, the x-axis is defined to point in the direction towards the center of the LHC, the y-axis upwards, and the z-axis along the direction of the beam. The transverse plane, $\theta = 90°$, serves as an important plane of reference where θ is the angle of a track relative to the beam axis. It is used in all particle physics analyses due to momentum conservation in this plane. The pseudorapidity, η, is a very useful

parameter defined as:

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right). \tag{3.4}$$

A track transverse to the beam axis, $\theta = 90°$, is denoted by $\eta = 0$, while a parallel track $\theta = 0°$ by $\eta = \infty$. It is recommended over the polar angle because of its Lorentz invariance under boosts, and it is used to convey the detector area-coverage capabilities. In cases where the particles travel close to the speed of light, the pseudorapidity converges to the definition of rapidity:

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right), \tag{3.5}$$

where p_z is the particle's longitudinal momentum. In such cases, pseudorapidity is chosen over rapidity, since it can be easily measured by polar angle θ. η is typically used along with the cylindrical coordinate r and ϕ. η and ϕ describe the trajectory of a particle from the interaction vertex after its emission, based on the angular separation between the particles:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}. \tag{3.6}$$

3.2.2 The Inner Detector

The ID is the subsystem closest to the interaction point and consists of three independent but complementary sub-detectors: the Pixel Detector, the SCT and the TRT. The ID is placed within a 2 Tesla axial magnetic field across the beam axis produced by a thin superconducting solenoid magnet. Figure 3.3 shows a cross-sectional view of the ID and illustrates the three sub-detectors and position of layers with respect to the beam pipe.

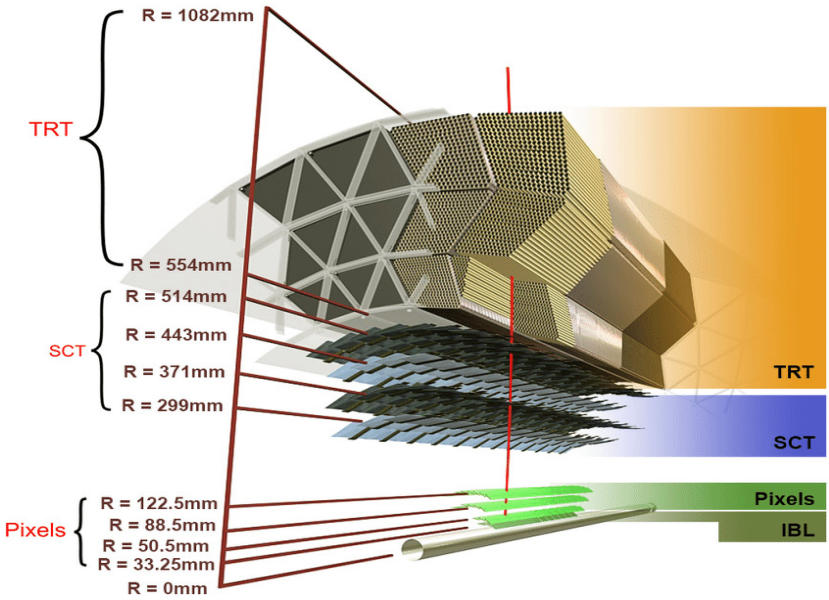


Figure 3.3: A schematic cross-section of the ID in the barrel section of the ATLAS detector, depicting the layers of the subsystems and the distance at which they are located from the interaction point [56].

The ID is built to provide pattern recognition to reconstruct high resolution tracking parameters. The particle trajectories are reconstructed from the hits in the silicon-pixel sensors and these are known as "tracks". The ID also measures the momentum of charged particles across a wide range of energies, where tracks are reconstructed from particles with $p_{\mathrm{T}} > 300$ MeV. The "impact parameter" is an important tracking parameter, this quantity relates to the measurement of secondary vertices and it is an important measure for the identification of jets from heavy-flavoured quarks, such as those from b-quarks. The estimation of W+jets process from the presented BSM physics search relies upon this experimental technique and is described in detail in Chapter 5. The overall measurement and resolution of the momentum given by the ID system is:

$$\frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} = 0.05\% \cdot p_{\mathrm{T}} \oplus 1\%, \tag{3.7}$$

where $\oplus$ denotes the sum of quadrature. A summary of the intrinsic resolutions of the subsystems of the ID is shown in Table 3.2 and the Insertable B-Layer, IBL [57], is the

Table 3.2: An overview of the intrinsic resolutions of the ID components. The sensor element size corresponds to the active detector component of the sensor. The sensor element size and intrinsic resolution are given in terms of (r-ϕ, z) for the Insertable B-Layer, IBL, and pixel detectors, and (r-ϕ) for the SCT and TRT [58].

Detector Component	Sensor Element Size	Intrinsic Resolution
IBL	50×250 μm×μm	8×40 μm×μm
Pixel	50×400 μm×μm	10×115 μm×μm
SCT	80 μm	17 μm
TRT	4000 μm	130 μm

additional pixel layer installed close to the beam pipe during Run 2.

3.2.2.1 Pixel Detector

The pixel detector [59] is situated closest to the beam pipe and consists of three silicon-pixel layers concentrically positioned at different distances from the beam pipe. The silicon-pixel layers and their distances from the beam axis are shown in green color from Figure 3.3. The layers are segmented in r-ϕ and z, where the r-ϕ plane is defined with r as the radial distance from the center of the detector. The three outermost layers support 1744 modules of silicon pixel with a size of 19×63 mm^2 and a thickness of 250 μm. Each module comprises of 46080 readout channels, providing over 80.4 million readout channels in total. These layers produce an intrinsic spatial resolution of 10 μm in the r-ϕ plane orthogonal to the beam axis and 115 μm in the z-axis.

3.2.2.2 Semi-Conductor Tracker

The Pixel Detector is surrounded by the SCT [60, 61], a silicon strip detector designed to provide eight precision measurements per track in the intermediate radial range to help to measure the momentum, impact parameters and vertex position. The barrel segment consists of four layers of silicon micro-strip modules covering the pseudorapidity range $|\eta| < 1.4$ and nine end cap disks covering $1.4 < |\eta| < 2.5$ on either side. Each module is composed of two strip layers with 80 μm mean strip pitch. The SCT consists of 4088 modules with about 6.3 million readout channels. The intrinsic resolution of

the SCT is 17 μm in the r-ϕ plane and 580 μm in the z-axis.

3.2.2.3 Transition Radiation Tracker

The TRT [62, 63] is the outermost subsystem of the ID. The principle of detection is based on the set of photons that are released when a charged particle traverses regions with different dielectric constants. The TRT consists of 4 mm diameter straw drift tubes and provides coverage of up to $|\eta| = 2$. The straw tubes are parallel to the beam axis in the central region and are 144 cm long. The 37 cm long straws are arranged radially in wheels in the end-cap region. The tubes are filled with the Xe-based gas mixture that facilitates the production of transition radiation as the charged particles pass through the TRT. The straws are positioned in 73 layers within the barrel region and aligned parallel to the direction of the beam, which is then devided into two halves at $\eta = 0$. Also, the end caps of the TRT contain 160 layers of long tubes placed radially around the beam axis. The TRT provides r-ϕ information for the determination of the transverse momentum with an accuracy of 130 μm. A large number of hits, usually 36 per track, are expected to be received by the TRT. This high number of hits and the larger track length compensate for the lower spatial resolution of each hit compared to the silicon detectors.

3.2.3 The Calorimeter System

The next major system of the ATLAS detector is the calorimeter system. Calorimeters measure the energies of incident particles, except muons and neutrinos, and are composed of alternating layers of passive and active material [55]. The calorimeter systems provide coverage of $|\eta| < 4.9$ and separated into towers in both ϕ and η, in the direction towards the center of the detector. A schematic view of the ATLAS calorimeter system is shown in Figure 3.4 The Electromagnetic Calorimeter, ECAL, and the Hadronic Calorimeter, HCAL, are the main systems in a calorimeter and discussed in the next subsections. The relative energy resolution of each calorimeter system is parametrized

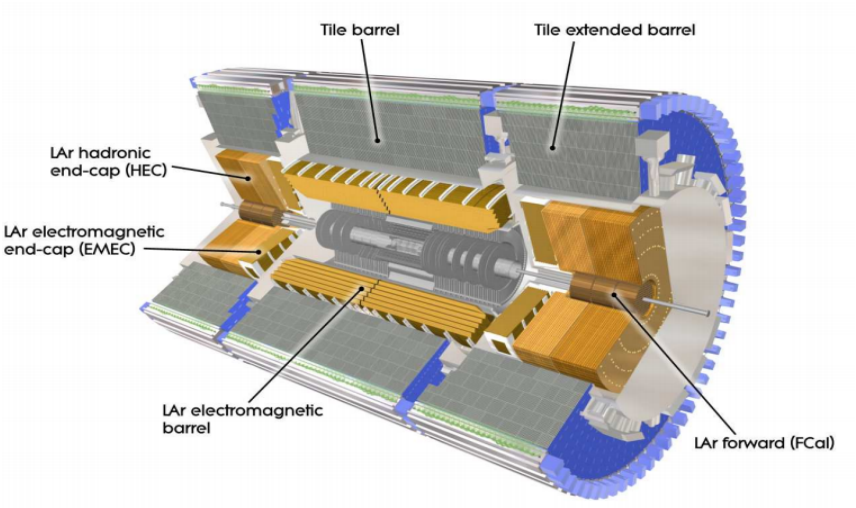


Figure 3.4: An overview of the calorimeter system [64].

Table 3.3: Energy resolutions of the constituent calorimeter sections [55].

Detector Component	Energy Resolution (σ_E/E)	η Coverage
ECAL	$10\%/\sqrt{E}\oplus$ 0.17%	$\lvert\eta\rvert < 3.2$
HCAL Barrel	$50\%/\sqrt{E}\oplus$ 3%	$\lvert\eta\rvert < 3.2$
HCAL end-cap	$50\%/\sqrt{E}\oplus$ 3%	$\lvert\eta\rvert < 3.2$
FCAL	$100\%/\sqrt{E}\oplus$ 10%	$3.1 < \lvert\eta\rvert < 4.9$

as follows:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{3.8}$$

where a is the stochastic contribution of the intrinsic fluctuations in the number and energy of the particles produced in the shower; b stands for the effect of noise from pile-up and read-out electronics; and c is due to systematic effects, such as miscalibration of the detector and also accounting for the inactive material present in the detector. The energy resolution of the different calorimeters, the sampling term, a, the noise term, b, and the constant term, c, are shown in Table 3.3.

3.2.3.1 Electromagnetic Calorimeter

The Electromagnetic, EM, Calorimeter is a sampling calorimeter [65, 66] surrounding the ID and solenoid system. The EM absorbs energy from particles that interact elec-

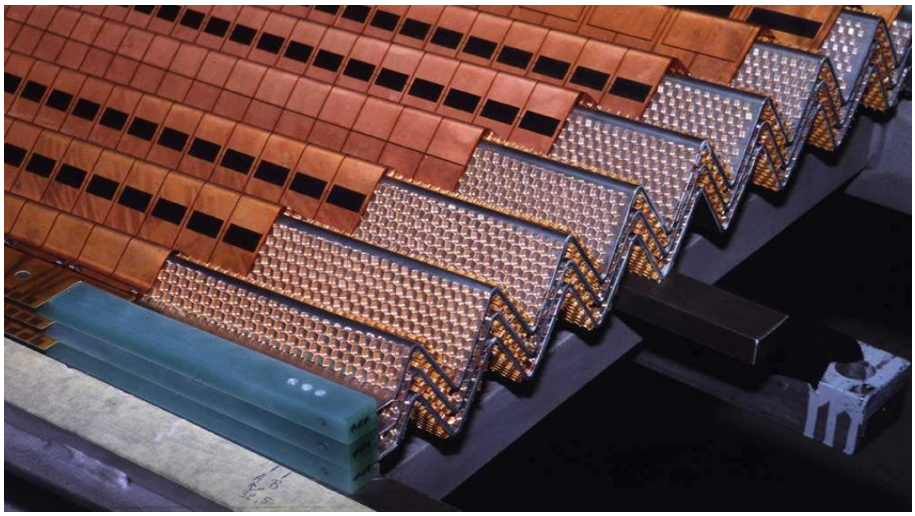


Figure 3.5: Accordion-shaped geometry of the LAr calorimeter [44].

tromagnetically, i.e. charged particles and photons. In this part, photons and electrons are required to lose all their energy, charged hadrons and muons will pass and inject a limited amount of energy. Neutrinos do not interact since they are electrically neutral. The EM calorimeter has a very high precision for the measurements of the absorbed energy as well as the place of the deposition. The EM consists of lead absorbers with an accordion shape, shown in Figure 3.5; Liquid-argon, LAr, as an active medium; and collection electrodes, which are responsible for measuring the deposited energy. This accordion geometry provides complete and uniform coverage in the azimuthal angle ϕ. In this calorimeter, an electromagnetic shower occurs as charged particles interact, resulting in several charged particles ionizing the argon. Within the LAr some electrodes are held at high voltage. The dense absorber is attached to the ground and generates an electric field inside the detector. Under the influence of this electrical field, the ionized argon and electrons head towards the electrodes, inducing an electrical signal that is processed by the read-out electronics. The EM calorimeters cover a pseudorapidity range of $|\eta| < 3.2$ with a gap between $1.37 < |\eta| < 1.52$. The EM calorimeter is divided into three sampling layers in the area $|\eta| < 2.4$, which are longitudinal in the depth of the shower and differ in granularity in the η direction. The first layer is called the strip layer and is divided with strips along the η direction providing high

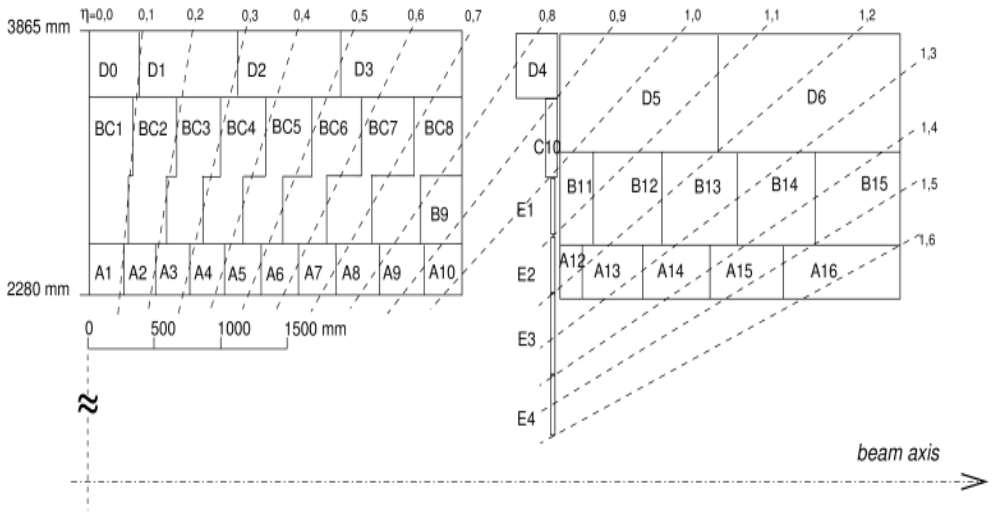


Figure 3.6: Segmentation in radial depth and pseudorapidity of the tile-calorimeter modules in the central (left) and extended (right) barrels. The bottom of the picture corresponds to the inner radius of the tile calorimeter and is parallel to the beam pipe. The tile calorimeter is symmetric about the interaction point at the origin [44].

granularity. This discriminates between showers began by photons or electrons and those caused by neutral pions. The middle layer includes projective cells that absorb most of the energy released by photon and electron showers. The rear layer is denser than the previous one and rectifies the leakage of high-energy showers outside the EM calorimeter.

3.2.3.2 Hadronic Calorimeter

Two different technologies are used for the HCAL: LAr calorimeters in the end-caps, as in the ECAL, and scintillator-tile detectors in the barrel section. Plastic tile scintillator detectors are used in the central section of the barrel called the Tile Barrel, $|\eta| < 1.0$, and the Tile Extended Barrel, $0.8 < |\eta| < 1.7$. The scintillators use steel as an absorbent material due to its high density. The barrel and the extended barrels are azimuthally divided into 64 modules where the layout is shown in Figure 3.6. The tile calorimeter, TileCal [66], is divided into three layers, which are approximately 1.5, 4.1 and 1.8 in the barrel section, and 1.5, 2.6 and 3.3 in the extended barrel. The tile calorimeter has a huge granularity, with cells of $0.1 \times \pi/32$ for the first layer in $\eta \times \phi$.

The sides of the scintillating tiles are readout in two separate photomultiplier tubes using wavelength shifting fibers. The end-cap regions use a LAr calorimeter, where the active medium is LAr, but the absorber material is copper instead of lead. It is placed behind the end cap of the electromagnet and uses the same cryostats to cool the LAr. A coverage of $1.5 < |\eta| < 3.2$ is provided by the end-caps.

3.2.3.3 Foward Calorimeter

The forward calorimeters, FCAL, are high-density liquid argon calorimeters in the far forward regions of the detector. They provide coverage from $3.1 < |\eta| < 4.9$. Each forward calorimeter is split into three layers. The first layer consists of copper absorbers to provide electromagnetic measurements, while the remaining layers utilise tungsten absorbers. The total depth of the forward calorimeter is equivalent to approximately 10 λ.

3.2.4 Muon Spectrometer

The Muon Spectrometer, MS [67], is the outermost part of the ATLAS detector, which covers the $|\eta| < 2.7$ region. The MS is designed to detect and measure the momentum of charged particles that are not already stopped by calorimeters, namely muons. The MS is capable of performing two functions: high precision tracking, and triggering. The muon subsystem is composed of four detector technologies. The Monitored Drift Tubes, MDTs, and Cathode Strip Chambers, CSCs, are responsible for precise monitoring in the area of the barrels and end caps respectively, covering the region $|\eta| < 2.7$ collectively. The Resistive Plate Chambers, RPCs, and Thin Gap Chambers, TGCs, are used as triggering chambers in the barrel region and end caps respectively, providing the triggered ability in the region $|\eta| < 2.4$. Figure 3.7 illustrates the MS with different parts.

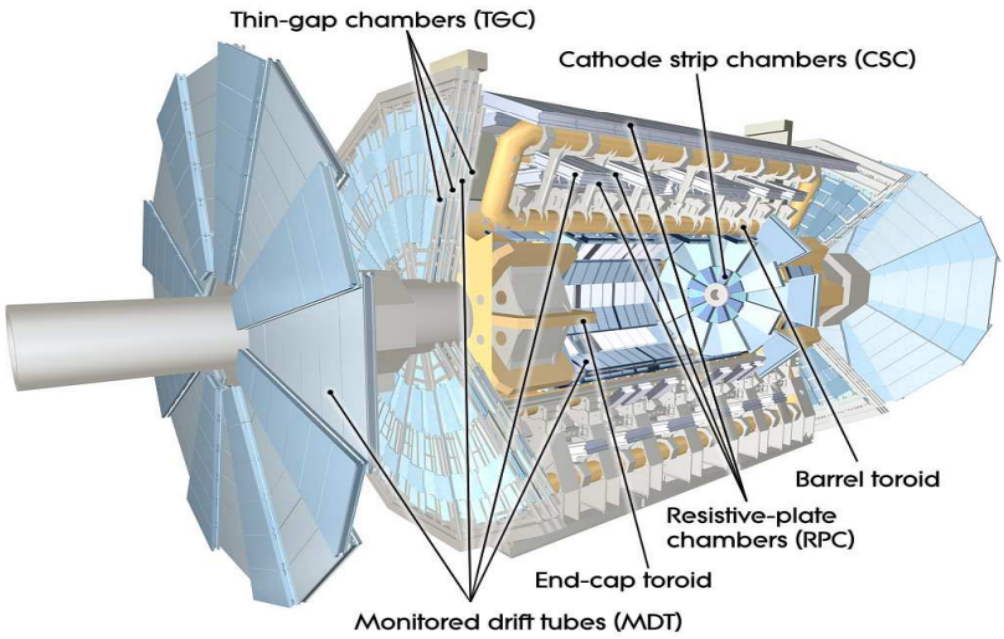


Figure 3.7: An overview of the Muon system [44].

The MDT comprises three cylindrical layers that run perpendicular to the beam-pipe and are responsible for the precise identification of the muon track in the bending plane, η. The second layer provides coverage of $|\eta| < 2.7$, while the third layer provides coverage of $|\eta| < 2.0$. There are 1088 chambers, which provide 5500 m^2 of the total active area. The MDT comprises of three to eight layers of drift tubes capable of averaging 80 μm resolution per tube and 35 μm resolution per chamber.

The CSC is the innermost layer and constructed from multi-wire proportional chambers with cathode planes divided into strips in orthogonal directions. The CSC is configured in two discs per end-cap, each segmented into eight chambers. Each chamber consists of four planes, which allow four measurements of η and ϕ for each of the muon tracks. The time the induced charges take to drift to the collection cathode measures the ϕ coordinate. The spatial resolution of a CSC is 40 μm in the bending plane and about 5 μm in the transverse plane. The coverage of $2.0 < |\eta| < 2.7$ is provided by the CSC layers.

The RPC is used in the barrel region and provides approximately 10 mm of spatial resolution in both bending, $|\eta|$, and non-bending, ϕ, planes. The end-caps use TGCs in the region $1.05 < |\eta| < 2.4$ and provide spatial resolutions of 2-7 mm in the bending plane, $|\eta|$, and 3-7 mm in the non-bending plane, ϕ. The design resolution of the MS is given by 10% for a track p_{T} of 1 TeV.

3.2.5 Trigger and Data Acquisition

The ATLAS trigger and data acquisition system, TDAQ, illustrated in Figure 3.8, is designed to select the physics events in the ATLAS detector. The dominant process in proton collisions is the inelastic scattering. The system is put in place to reject the majority of uninteresting events while selecting the rarer, more interesting physics processes. A multi-level trigger system is implemented, consisting of a Level 1 hardware trigger, L1, and a High-Level software trigger, HLT, where the requirements of both triggers must be met to accept an event and write to disc. The L1 trigger provides direct access to the data from the calorimeters and the MS. The MS provides fast trigger signals that are used for the L1 trigger decision. The maximum L1 trigger rate is 100 kHz, where each event is on average 1.3 MB in size. The event chosen by the L1 trigger is then processed by the HLT, a software-based trigger system. The HLT is designed to perform a more complete examination of the event, where it performs a global reconstruction of the track and receives tracking information from the ID, the calorimeter and the MS.

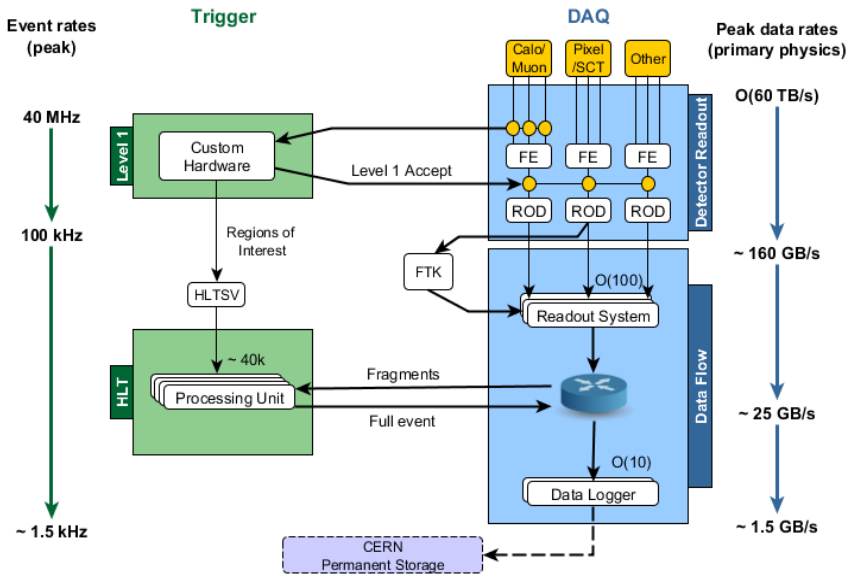


Figure 3.8: ATLAS Trigger and DAQ systems, with the expected peak rates and bandwidths for each component [68].

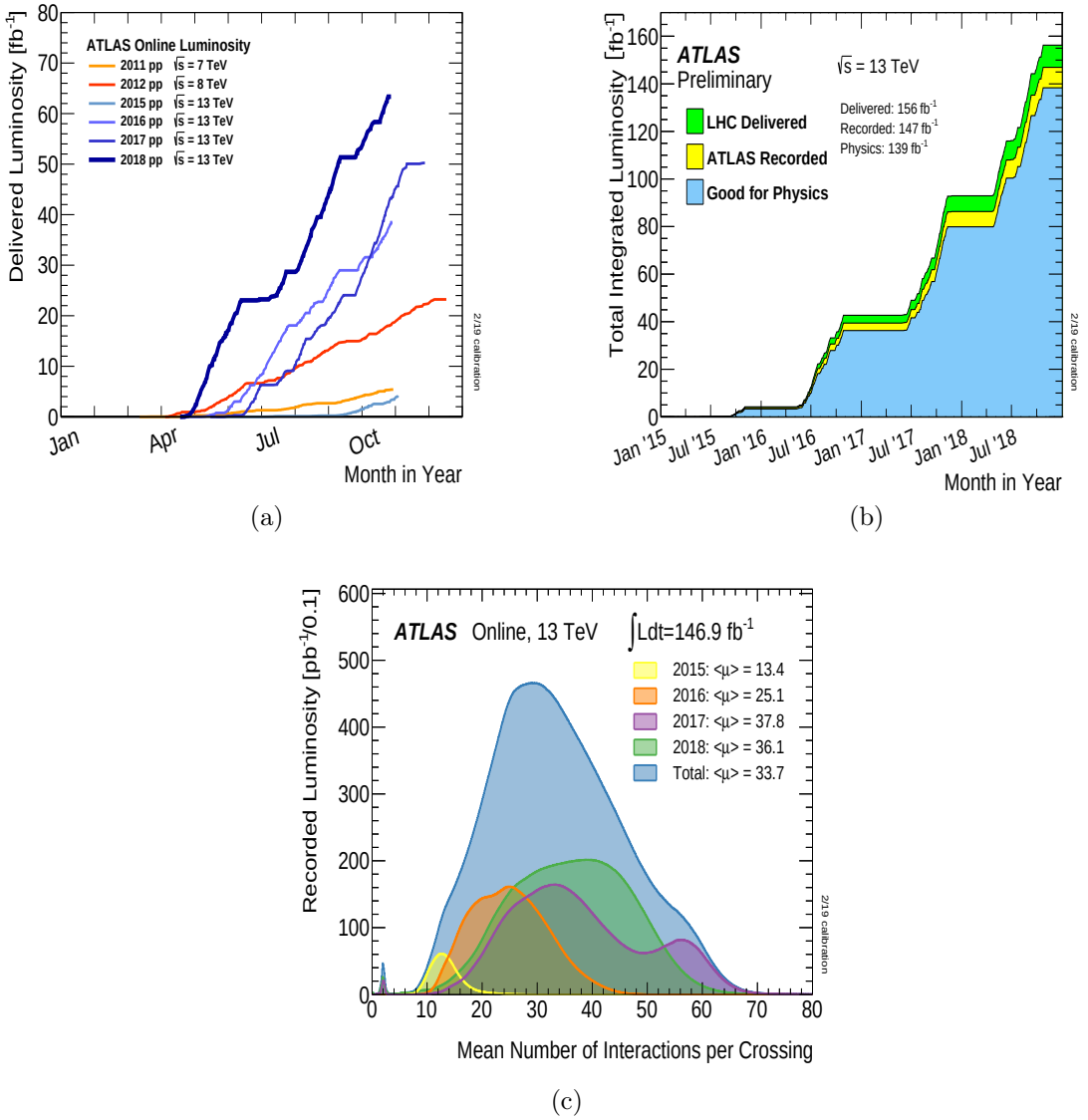
Chapter 4

Data Set and Simulated Events

4.1 Data Set

The data-taking of the ATLAS detector can be divided into two main periods of operation: Run 1 between 2010 and early 2013, and Run 2 between 2015 and 2018. The long shutdown separated the two data-taking periods and was used as a time to upgrade the detector. The total luminosity that the LHC delivered during Run 1 was $\sim$ 5.5 (22.8) fb^{-1} during 2011 (2012) at $\sqrt{s}$ = 7 (8) TeV. During Run 2, the LHC provided 156 fb^{-1} of proton-proton collisions at $\sqrt{s}$ = 13 TeV, of which 147 fb^{-1} was recorded by the ATLAS detector. Of the total recorded, 95% is deemed to be of high enough quality for physics analyses. The BSM search presented in this thesis utilises the full Run 2 data-set of 139 fb^{-1}.

Figure 4.1 (a) depicts the annual delivered luminosity, where each year of data-taking is presented. The total amount of delivered luminosity increases with each year of data-taking. Figure 4.1 (b) depicts the cumulative luminosity delivered by the LHC during Run 2, the total recorded by the ATLAS detector and the total deemed to be of high enough quality for physics analyses. Figure 4.1 (c) shows the average number of interactions per crossing, $\langle\mu\rangle$. The differences between the data-taking years are caused by the differences in LHC machine parameters during operation. The average



(a) (b)

(c)

Figure 4.1: Delivered luminosity in Run 1 and Run 2 (a), total integrated luminosity during Run 2 (b), and mean number of interactions per bunch crossing for the pp collision data from 2015 to 2018 at $\sqrt{s} = 13$ TeV (c) from Ref. [69].

Table 4.1: Single lepton triggers for the 2015 to 2018 data set used in the analysis.

Data-set	Electron Trigger	Muon Triger
2015	HLT_e24_lhmedium_L1EM20VH	HLT_mu20_iloose_L1MU15
	HLT_e60_lhmedium	HLT_mu50
	HLT_e120_lhloose	
2016-2018	HLT_e26_lhtight_nod0_ivarloose	HLT_mu26_ivarmedium
	HLT_e60_lhmedium_nod0	HLT_mu50
	HLT_e140_lhloose_nod0	

number of interactions per bunch crossing for Run 2 was found to be $\langle\mu\rangle = 33.7$.

4.2 Trigger

The trigger mechanism used has shifted between the data-taking periods from 2015 and 2016 to 2018. The single lepton triggers are introduced for an electron or a muon for these years. The single lepton trigger both year periods are shown in Table 4.1. The trigger names "HLT" refer to the high-level trigger based on the software and "L1" to the first-level trigger based on the hardware. The second part of the name, for example "e26" or "mu26", corresponds to the transverse momentum threshold of the lepton, $p_{\mathrm{T}} > 26$ GeV, set at the "HLT", whereas "L1EM20VH" and "L1MU15" correspond to the transverse momentum threshold of the lepton at the L1 level. The conditions for the classification of "lhloose", "lhmedium" and "lhtight" are likely to be established on the lepton. The abbreviation "nod0" defines that the distance from the measured tracks to the primary vertex is not subject to any criterion. Due to the higher luminosity, the limit on the transverse momentum on the lepton is raised from the data set for 2015 to 2018.

Table 4.2: Signal samples for $H \to Sh$ analysis at different mass points.

Process	(m_H, m_S) GeV	Simulation	Generator	Parton Distribution	Cross section (pb)	Events
$H \to Sh$	(240,170)	FastSim	Pythia8	A14NNPDF23LO	1.16E-02	120k
	(350,240)	FastSim	Pythia8	A14NNPDF23LO	2.05E-05	120k
	(400,240)	FastSim	Pythia8	A14NNPDF23LO	1.96E-06	120k
	(240,170)	FullSim	Pythia8	A14NNPDF23LO	1.16E-02	120k

4.3 Signal Samples

The $H \to Sh$ production mode is produced for the parton shower with the MC event generator,[4] Pythia 8.210 [70] and the parton distribution function, PDF, is used with A14NNPDF23LO [71] tune. Samples are produced at different mass points of H and S with m_h at 125 GeV. Table 4.2 shows the signal samples with the corresponding generator and different cross sections at each mass point. The (240,170) GeV mass point is repeated because one is produced with the full simulation and the other with fast simulation. Samples are produced to be inclusive of all possible decays of S and h. A lepton filter is used to generate events with at least two leptons. The lepton filter efficiency is between 4% and 12%. .

4.4 Background Samples

The main sources of SM backgrounds include events from top-quark pair ($t\bar{t}$), associated production of a top quark with a W boson (Wt), single vector boson production (Z and W without or in association with jets), WW production, other diboson (WZ, ZZ, $V\gamma$ with $V = W, Z$) and the Higgs boson $h \to WW$ through ggF production.

The W+jets process modeling is based on a purely data-driven method described in Chapter 6. However, MC samples are used to validate the fake estimation as well as to estimate the sample composition uncertainties, see Section 6.6. For those studies, same MC generators are used to predict the W+jets and Z+jets. Simulated

[4]Event generators such as MC are computer libraries that randomly generate simulated events.

$W(\rightarrow e\nu/\mu\nu/\tau\nu)$+jets events are produced with the POWHEG-BOX v2 [72] generator using the CT10 [73] next-to-leading order (NLO) PDF, interfaced to PYTHIA v8.186 for parton showering, hadronisation and underlying event. As for the WW samples, the AZNLO [74] tune is used for the underlying event together with the CTEQ6L1 [75] PDF set. $Z(\rightarrow ee/\mu\mu/\tau\tau)$+jets production is modelled using the SHERPA v2.2.1 generator [76] with the NNPDF3.0 PDF set [77] at next-to-next-leading order (NNLO). SHERPA includes its own parton showering as well as hadronisation and underlying event modelling. The minimum invariant mass of the dilepton system in these samples is 10 GeV. The samples include up to two additional parton emissions at NLO and up to four at leading-order (LO).

For the generation of $t\bar{t}$ and single top production in the Wt channel, the POWHEG-BOX v2 is used with the CT10 next-to-leading order (NLO) PDF set. For parton shower, hadronisation and underlying event, simulated top events are interfaced to PYTHIA 8.210 for $t\bar{t}$ and PYTHIA 6.428 for single top production, using the A14 [78] tune and the PERUGIA 2012 [79] tune, respectively. The top-quark mass is set to 175.5 GeV.

Background diboson production ($WZ, ZZ, WW, W\gamma,$ and$Z\gamma$) is simulated using the SHERPA v2.2.2 generator with the NNPDF 3.0 PDF set at NNLO. The minimum γ transverse momentum in the $W\gamma$ and $Z\gamma$ samples is set to 7 GeV. The Higgs boson production through ggF events are generated with POWHEG-BOX v2 interfaced to PYTHIA 8 with the AZNLO tune at NNLO accuracy in QCD [80]. Table 4.3 gives a summary of the background processes.

Table 4.3: Background samples and corresponding MC generators used, PDF set and underlying-event tune used for each process.

Process	Event Generator	PDF set	Tune
$gg\mathrm{F}(h \to WW)$	Powheg+Pythia8	NNLO	AZNLO
$t\bar{t}$	Powheg-Box v2	CT10 NLO	A14
Z+jets	Sherpa 2.2.1	NNPDF3.0	Sherpa default
Single-top	Powheg+Pythia	CT10 NLO	Perugia 2012
W+jets	Powheg+Pythia	CT10 NLO	AZNLO
Diboson	Sherpa 2.2.2	NNPDF	AZNLO

Chapter 5

Object Reconstruction and Selection

5.1 Object Definition

The analysis currently relies on the $h \to WW \to \ell\nu\ell\nu$ object definitions, which are briefly described in this chapter: leptons in Section 5.1.1 and jets in Section 5.1.2.

5.1.1 Leptons

The lepton selection has been optimised to maximize the sensitivity of the analysis. Lepton identification choices impact the analysis primarily through the rejection of the W+jets and QCD backgrounds. Leptons are required to originate from the primary vertex that has no less than two associated tracks, each with transverse momentum $p_{\mathrm{T}} > 400$ MeV, where the p_{T} is defined as the magnitude of the component of the momentum orthogonal to the beam axis. If there is more than one primary vertex reconstructed in the event, the one with the largest track $\Sigma\, p_{\mathrm{T}}^2$ is chosen as the hard-scatter primary vertex and is subsequently used for calculation of the main physics objects in this analysis: electrons, muons, jets, and missing transverse momentum. To this end, the absolute value of the longitudinal impact parameter of each lepton track, calculated with respect to the vertex and multiplied by $\sin\theta$ of the track, $|z_0 \sin\theta|$, is required to be less than 0.5 mm. Furthermore, the significance of the transverse impact

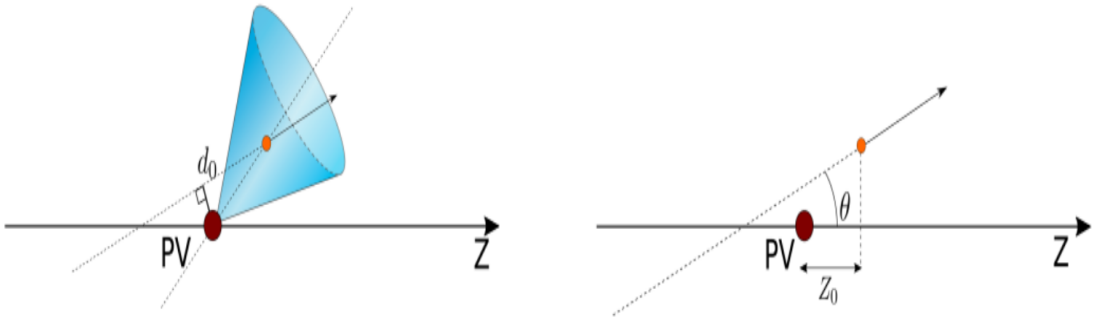


Figure 5.1: Definitions of the d_0, $z_0 \sin\theta$ track reconstruction parameters. PV denotes the primary vertex of the pp collisions.

parameter calculated with respect to the beam line, $|d_0|/\sigma d_0$, is required to be less than three for muons and five for electrons. In the choice of the cuts on the impact parameter significance, the recommendation of the Muon and Egamma combined performance groups are followed [81]. Figure 5.1 illustrates the d_0 and $z_0 \sin\theta$ track reconstruction parameters in pp collisions.

5.1.1.1 Electrons

Electrons are selected from clusters of energy deposits in the calorimeter that match a track reconstructed in the inner detector. They are identified using likelihood identification criteria described in Ref [82]. The levels of identification are categorised as "LooseLH", "MediumLH", and "TightLH", which correspond to approximately 96%, 94% and 88% identification efficiency for an electron with a transverse energy of 100 GeV, where E_{T} is defined as the energy projected into the transverse plane. The electrons used in this analysis are required to have E_{T} greater than 25 GeV and pass "MediumLH". In addition, electrons with $15 < E_{\mathrm{T}} < 25$ GeV are selected if they satisfy the "TightLH" criteria. The pseudorapidity of electrons is required to be within the range of $|\eta| < 2.47$, excluding the transition region between the barrel and endcaps in the liquid argon, LAr, calorimeter, $1.37 < |\eta| < 1.52$. The isolation requirement has been optimised as described in [84] by taking into account the signal and the background efficiency, including the change in the fake factor for the W+jets. Moreover,

Table 5.1: Electron selection.

p_{T} range	Electron ID	Author	Isolation	Impact Parameter
< 25 GeV	TightLH	1	FixedCutTrackCone40	$\lvert z_0 \sin\theta\rvert < 0.5$, $\lvert d_0\rvert/\sigma d_0 < 5$
> 25 GeV	MediumLH	1	Gradient	$\lvert z_0 \sin\theta\rvert < 0.5$, $\lvert d_0\rvert/\sigma d_0 < 5$

Table 5.2: Muon selection.

p_{T} range	Muon ID	Calo Isolation	Track Isolation	Impact Parameter
> 15 GeV	Tight	$E_{\mathrm{T}}^{cone20}/p_{\mathrm{T}} < 0.09$	$E_{\mathrm{T}}^{varcone20}/p_{\mathrm{T}} < 0.06$	$\lvert z_0 \sin\theta\rvert < 0.5$, $\lvert d_0\rvert/\sigma d_0 < 3$

a requirement, Author = 1, has been added to select electrons that are reconstructed exclusively as electrons.

5.1.1.2 Muons

Muons are reconstructed by combining ID and muon spectrometer (MS) tracks that have consistent trajectories and curvatures. The combination is performed through an overall fit using the hits of the inner detector track, the energy loss in the calorimeter, and the hits of the track in the muon system. In this way the combined muons are reconstructed [83]. Based on the quality of the reconstruction and identification, muon candidates are defined as "Loose", "Medium" and "Tight", with increasing purity. The muon candidates used in this analysis are required to pass "Tight" selection and also have $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$. The isolation requirement for muon has also been optimised as for electrons, details in Ref.[84].

The results of the electron and muon selection are summarised in Table 5.1 and Table 5.2. From these tables isolation refers to the quantity of tracks and deposition of calorimeter energy near the lepton. Specifically, the selection is based on the so-called "FixedCutTrackCone40", "gradient isolation" also the track and calorimeter isolation "$E_{\mathrm{T}}^{varcone20}/p_{\mathrm{T}}$" and "$E_{\mathrm{T}}^{cone20}/p_{\mathrm{T}}$" that incorporates the parameter's η dependency.

5.1.2 Jets

The $H \to Sh \to \ell\nu\ell\nu$ analysis selects at least two jets, hence jets are an important part of the object definition. Jets used for the analysis are referred to as "tagging" jets. Jets are reconstructed from three-dimensional clusters of energy deposits in the calorimeter using the anti-k_t algorithm [85] with distance parameters of $R = 0.4$. The four momenta of the jets are calculated as the sum of the four momenta of their constituents, which are assumed to be massless. The momenta of jets are corrected for losses in passive material, the non-compensating response of the calorimeter, and contributions from pile-up [86].

The jets in this analysis are required to have $p_{\mathrm{T}} > 30$ GeV in the full η range. Jets with $|\eta| > 4.5$ are discarded. For jets with $p_{\mathrm{T}} < 30$ GeV and $|\eta| > 2.4$, the calibrated "Jet vertex tagger", JVT, variable is required to be larger than 0.59, where the JVT is a multivariable tagger used to suppress jets from pile-up events [87]. Furthermore, to reduce the pileup in the forward region the ForwardJVT is also applied [88]. Forward JVT method is designed to identify pileup jets in the forward region, $|\eta| > 2.5$, where there is no tracking information by using topology information. The working point has an efficiency of 90%.

5.1.2.1 b-tagged jets

Jets containing b-hadrons are identified using the MV2C10 b-tagging algorithm [89] with an efficiency of 85%, determined from $t\bar{t}$ simulated events. The jets recognised as b-quark-induced are called b-jets in this thesis. The corresponding rejection for selecting jets originating from a light quark or gluon is 33.53. The rejection for selecting jets containing c-hadrons is approximately 3.1. It was also decided to apply the b-jet veto to jets below the jet p_{T} threshold of the analysis, 20 GeV$< p_{\mathrm{T}} <$ 30 GeV, to improve the suppression of top-quark background.

Table 5.3: Summary of the baseline event selection in the SR.

Selection requirement	SR
Leading lep p_{T}	> 27 GeV
Sub-leading lep p_{T}	> 15 GeV
Number of jets, $p_{\mathrm{T}} > 30$ GeV	$\geqslant 2$
b-tagged jet multiplicity, $p_{\mathrm{T}} > 20$ GeV	$= 0$
$m_{\tau\tau}$	< 70 GeV
m_{T}	< 300 GeV

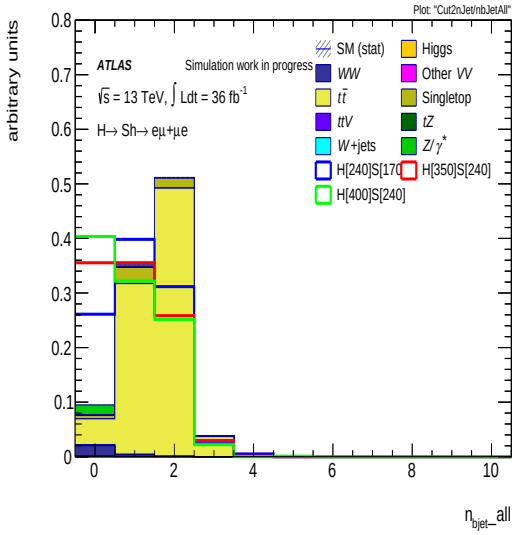
5.2 Event Selection

The pre-selection requires the following cuts: the good run list, GRL, jet cleaning, overlap: $V\gamma$ or Vjets, trigger, and trigger matching. After these requirements, a selection of exactly two leptons with minimal selection on the invariant mass, $m_{\ell\ell}$, is applied. Leptons should have different flavor and opposite electric charge, DFOS, $e^{\pm}\mu^{\mp}$. The transverse momentum of the leading lepton, $p_{\mathrm{T}}^{\ell 0}$, must be greater than 27 GeV and the sub-leading lepton transverse momentum, $p_{\mathrm{T}}^{\ell 1}$, greater than 15 GeV. After the pre-selection, the following cuts are applied to define the full SR: at least two jets, b-veto, the invariant mass of tau pair, $m_{\tau\tau}$ defined from collinear mass approximation [90] followed by the transverse mass, m_{T}, defined as $m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}E_{\mathrm{T}}^{miss}(1 - cos\Delta\phi\ell\nu)}$. The summary of the event selection in the SR is shown in Table 5.3.

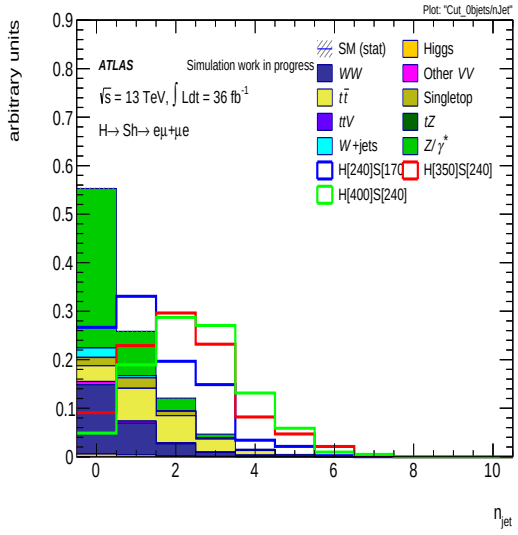
Table 5.4 shows only the dominant backgrounds at the pre-selection and the SR. These dominant backgrounds are top ($t\bar{t}$ and single-top), dibosons (WW and other VV), Z/γ and W+jets. The total backgrounds include other background processes such as triboson, $V\gamma$ and Higgs. The b-veto is applied to reject the top background. Top is the main dominant process with about 75% of the total background at the final selection, $m_{\mathrm{T}} > 50$ GeV. This is followed by dibosons, WW and other VV with 19%, the Z/γ with 6% and W+jets with 1.2% of the total background shown in Table 5.4.

Table 5.4: Cutflow showing only the main dominant background after the final pre-selection and SR requirements. The top part shows the pre-selection followed by SR selections.

Cuts	dibosons	top	Z/γ^*	W+jets MC	AllBkg
pre-selection	132302.33 ± 165.20	1135402.86 ± 402.65	216553.79 ± 432.84	13286.23 ± 364.38	1507582.75 ± 658.81
>=2jets	19519.39 ± 39.80	814648.53 ± 318.70	15928.47 ± 92.88	1011.79 ± 100.16	852445.14 ± 280.25
b-jets=0	16135.63 ± 37.61	56130.59 ± 90.81	12681.77 ± 86.32	826.82 ± 90.13	86751.19 ± 135.69
$M_{\mathrm{tt}} < 70$ GeV	8952.87 ± 24.11	33108.17 ± 69.44	4288.00 ± 58.18	513.39 ± 70.49	47574.58 ± 97.81
$M_{\mathrm{T}} < 300$ GeV	8544.59 ± 23.50	31875.67 ± 68.16	2963.41 ± 47.10	479.74 ± 68.40	44503.23 ± 89.86



(a) Number of b-jets after at least two jets



(b) Number of jets after the b-jets veto

Figure 5.2: Kinematic distributions of number of jets after the b-jets veto requirement on (left) and b-jets after requiring at least two jets (right).

Figure 5.2 illustrates the number of b-jets after requiring at least two jets and the number of jets after the b-veto requirement. Signal dominates in events without b-jets. The distributions show the MC only comparing the SM background and signal normalised to the total background with the luminosity of 36 fb^{-1}. All the background is estimated from the MC except the W+jets, which is estimated using the data-driven method described in the next chapter.

Chapter 6

W+jets Background Estimation

A challenging source of background is formed by processes that enter the SR as fakes. Collisions producing W bosons in association with one or more jets referred here as W+jets are accepted into the SR that requires two identified leptons when a jet is misidentified as a prompt lepton. The probability of such misidentification is small, but the relatively large cross section for W+jets events makes such process a significant background. The so-called FF method will be presented in this chapter for estimating the W+jets contribution in the SR. The estimations have been carried out with ATLAS full Run 2 dataset.

6.1 The W+jets Background Process

The present research focuses on the W+jets process, where a W boson is generated in conjunction with a jet. If the W decays leptonically and the jet is misidentified by its detector signature as a lepton, the process enters the SR as it resembles the two-lepton signature. Before addressing the reduction of this background from the SR, it is necessary to examine how large a contribution is to be expected and whether it poses a significant problem. In ATLAS, the jet suppression is at a rate of 10^{-5} [91]; only jets are misidentified as leptons in the detector response tails. Nevertheless, despite the small

rate of fake lepton, due to the large production cross section of jets at the LHC, there may be a significant level of background from misidentification. MC simulations are not supposed to accurately model the W+jets, as a detailed theory forecast will entail accurate simulations of the misidentified particles along with an accurate model of the frequency of misidentification. As mentioned, jet misidentification occurs at small rates making it difficult to model it theoretically and one has to develop data-driven techniques to estimate it.

6.2 W+jets Fake Estimation Strategy

The basic idea of the FF approach is to select a control sample of events enriched in the W+jets background, also known as the control region, CR, and then use an extrapolation factor to relate these events to the background in the SR. The method is data-driven given that the control sample is chosen in data, and the extrapolation factor is measured with data. To collect background events more efficiently, the SR particle selection is replaced in the CR with a selection of particles, for which misidentification occurs more frequently.

The CR is described as the same as the SR, apart from some parameters that need to be loosened compared to the full range of particles in the SR. Specifically, the CR consists of two leptons where one of the leptons satisfies a tight criteria, these lepton candidates are referred to as "identified" or "id", while the other does not meet these criteria, it fulfills loose set of requirements, denoted as "anti-identified" or "anti-id". Table 6.1 presents the summary of the requirements. The Fake Factor, FF, is defined as the ratio of the number of "id" lepton to the number of "anti-id":

$$FF = \frac{N_{id}}{N_{anti-id}}. \tag{6.1}$$

It is measured separately for electron and muon in a region enriched in Z+jets process.

Table 6.1: Requirements for "id" and "anti-id" electrons (left) and muons (right).

id electron	anti-id electron	id muon	anti-id muon
$p_{\mathrm{T}} > 15$ GeV		$p_{\mathrm{T}} > 15$ GeV	
$\lvert\eta\rvert < 2.47$, excluding $1.37 < \lvert\eta\rvert < 1.52$		$\lvert\eta\rvert < 2.45$	
$\lvert z_0 \sin\theta\rvert < 0.5$ mm		$\lvert z_0 \sin\theta\rvert < 0.5$ mm	
PassLHTight if $p_{\mathrm{T}} < 25$ GeV PassLHMedium if $p_{\mathrm{T}} > 25$ GeV	PassLHLoose	Pass Quality Tight	Pass Quality Medium
$\lvert d_0\rvert/\sigma(d_0) < 5$		$\lvert d_0\rvert/\sigma(d_0) < 3$	$\lvert d_0\rvert/\sigma(d_0) < 15$
Pass FixedCutTrackCone40 isolation if $p_{\mathrm{T}} < 25$ GeV Gradient isolation if $p_{\mathrm{T}} > 25$ GeV		$E_{\mathrm{T}}^{cone20}/p_{\mathrm{T}} < 0.09$ $E_{\mathrm{T}}^{varcone20}/p_{\mathrm{T}} < 0.06$	
Author==1			
	Veto against identified Electron		Veto against identified Muon

It measures the ratio of the number of objects passing the lepton identification selection to the number of objects passing the "anti-id" requirements but failing the "id" criteria. The FF is not a probability because the numerator is not a subset of the denominator. The veto against identified electron or muon implies that the "anti-id" selection should fail one of the "id" requirements on isolation or $|d_0|/\sigma(d_0)$. The FF relates background misidentified with this criteria, to background misidentified as passing the full particle selection of the SR. The main concept of the FF method can be summarised in the following relation:

$$N_{id+id}^{W+jets} = FF \times N_{id+anti-id}^{W+jets}, \tag{6.2}$$

where N_{id+id}^{W+jets} corresponds to the W+jets events that enter the SR. $N_{id+anti-id}^{W+jets}$, is the number of W+jets events where one of the lepton signatures satisfies all signal lepton conditions and the other fails, but instead meets a looser set of requirements. The $N_{id+anti-id}^{W+jets}$ is estimated in the CR. The extrapolation linearity is an assumption based on the CR definition; provided that the CR is only slightly different from the SR, it can be concluded that a polynomial, i.e. linear, first-order extrapolation is a

reasonable approximation. In the $H \to Sh$ analysis, electrons and muons are included in the final state of interest. The background of W+jets may arise from the jet being misidentified as either an electron or a muon. Since the criteria for identifying a particle as an electron or muon are different, it is expected that the FF will be different for jet misidentification as an electron or muon and separate calculations have to be performed.

6.3 W+jets Control Region

As already stated, the CR is characterised as rich in W+jets events, i.e. approximately 85% to 90% of the events must come from this process [92] and do not interfere with the $H \to Sh$ signal. It is also necessary to subtract EW processes: WW, WZ, and ZZ, as well as $t\bar{t}$ and single top from the CR, as only fake signatures should be studied. In the W+jets CR there is one-"id" plus one-"anti-id" lepton and same requirements as in the SR defined in Table 5.3 are applied.

The definition of "id" or "anti-id" for the choice of CR particles is based on the identification and isolation criteria. The selection requirements for electrons and muons with respect to these parameters vary with respect to p_{T} and η, hence the FF calculations, as well as lepton flavour, are binned with respect to these variables. The difference is focused on how muons are reconstructed from specific detector subsystems, mainly muon spectrometer and internal detector. For the distinction between "id" and "anti-id" leptons, the specific cut definition is not fixed, but is selected to optimise the analysis. A tight definition can limit the statistics, while with a broad one the objects that enter the CR may not indicate similar characteristics of the SR objects.

Figure 6.1 shows an illustration of the CR definition based on the parameter of isolation. The real lepton of the W+jets process will pass the "id" SR requirements. In some cases, the jet will pass all "id" requirements, entering the SR, whereas in some

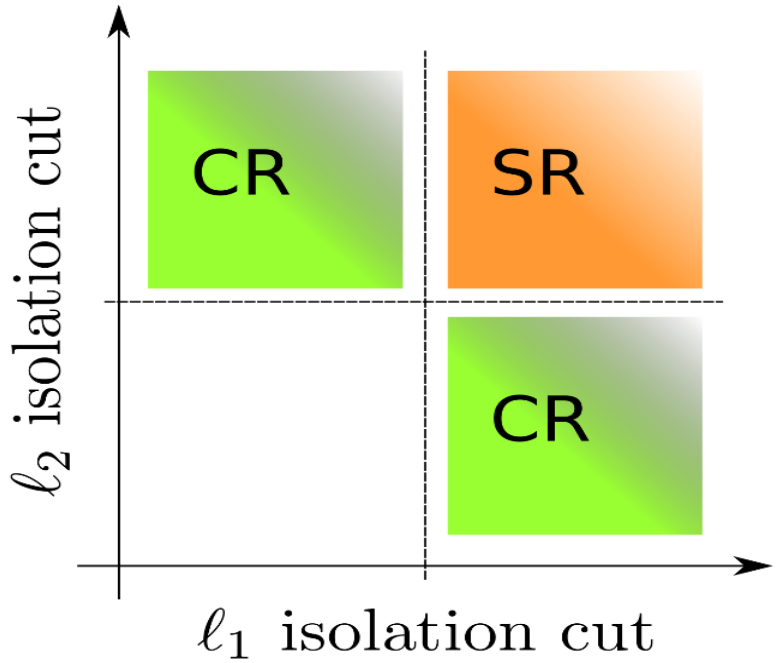


Figure 6.1: Example of CR definition based on lepton isolation requirements. For the W+jets process, the real lepton originating from the W will pass the "id" requirements. The jet will in some cases pass all "id" requirements, entering the SR, while in some cases it will only pass the looser requirements.

cases the looser requirements will only pass thus entering the W+jets CR.

6.3.1 Control Region Kinematics

In this section, distributions in the W+jets control region, 1-id + 1-anti-id, are presented, showing the estimation of the W+jets background process. This region is enriched with fakes. The estimation of W+jets background is given by Equation 6.3:

$$W + jets^{Estimate} = Data - EW^{MC}. \tag{6.3}$$

EW MC processes are subtracted from data and the difference is taken as the W+jets estimation. This estimation is split into fake lepton: electron and muon.

6.3.1.1 Electron, Anti-id

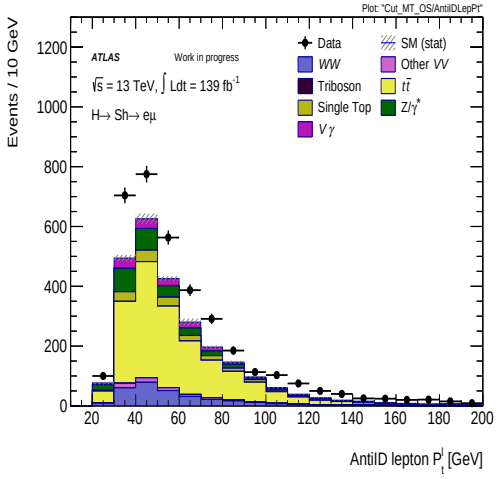
The contribution of the fake electron in the W+jets CR is shown in Figure 6.2. The plots show the composition of data and EW subtraction; the W+jets is taken as the difference between the data and the other processes. The electron fake candidate p_{T} and η are shown for $e\mu$ and μe channels respectively. The $e\mu$ and μe channels are ordered by p_{T} of leptons ℓ_o, ℓ_1 where ℓ_o represents the leading lepton and ℓ_1 the sub-leading lepton. For the $e\mu$ channel, the leading lepton is an electron and muon is the sub-leading. The μe case represents the other way round, the muon is the leading and electron is the sub-leading. The $e\mu$ channel has a lower contribution of the fake electron than in the μe channel as expected. More fake electrons come from the sub-leading electron than the leading. The other kinematic distributions for the W+jets CR of the electron fakes are shown in Appendix A Section A.1.

6.3.1.2 Muon, Anti-id

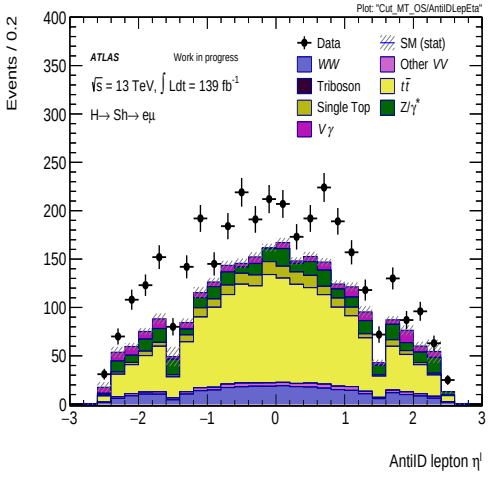
Figure 6.3 shows the muon fake contribution from the muon fake candidate p_{T} and η are shown for $e\mu$ and μe channels respectively. A larger fake muon contribution is observed in the $e\mu$ channel. The μe channel gives a small contribution. Other kinematic distributions for the W+jets CR of the muon fake are shown in Appendix A Section A.1.

6.4 Measurements of Z+jets Fake Factor

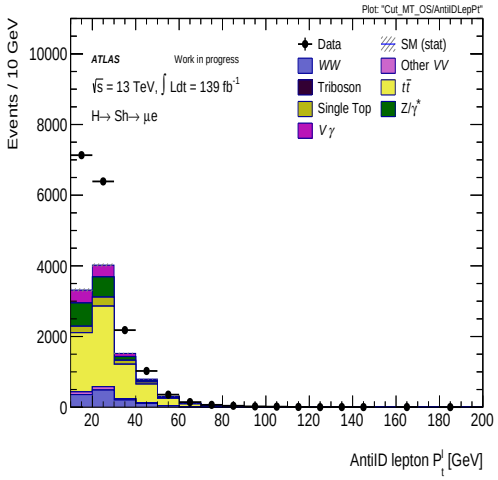
The nominal FF estimation is done in a sample of Z+jets events, where an additional lepton candidate, apart from the two leptons decaying from the Z-boson, is present. The sample of additional lepton candidate is subdivided into two sub-samples, depending on whether the additional lepton passes the full "id" or "anti-id" requirements detailed in Table 6.1. Both samples are binned double-differentially in the transverse momentum and pseudorapididty of the lepton candidate, being the FF the ratio be-



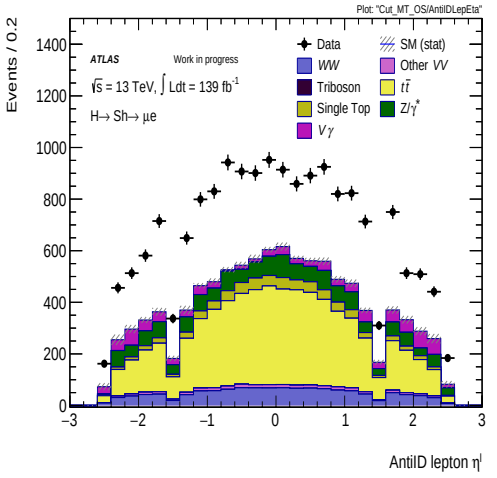
(a) $e\mu$: electron p_{T}



(b) $e\mu$: electron η

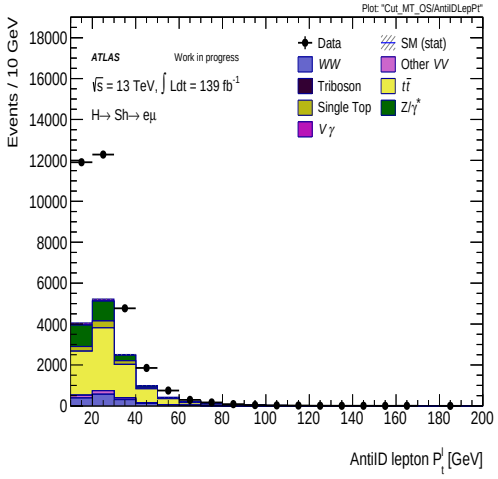


(c) μe: electron p_{T}

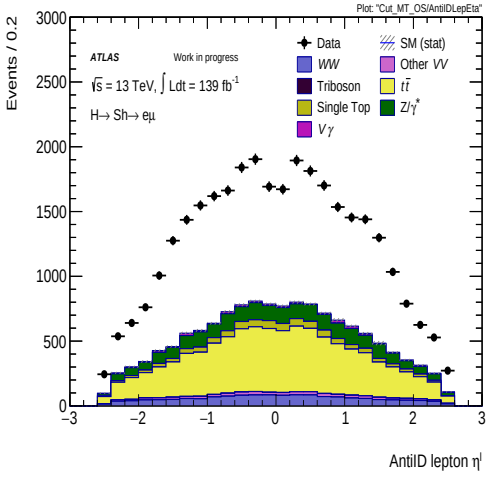


(d) μe: electron η

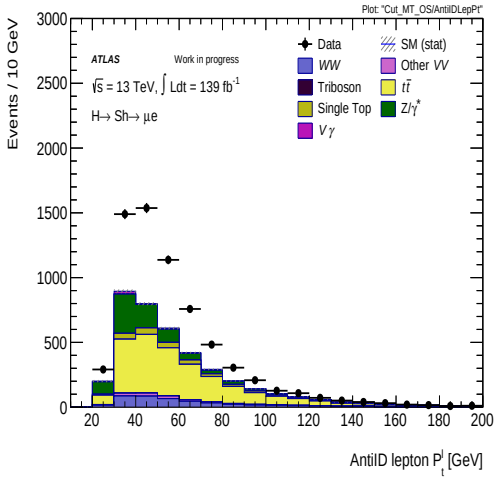
Figure 6.2: W+jets control region plots for electron fakes in the $e\mu$ and μe channel in the top and bottom, respectively. The plots show the composition of data and EW subtraction; the W+jets is taken as the difference between the data and the other processes. The electron fake p_{T} distribution is shown in the left and electron fake η in the right.



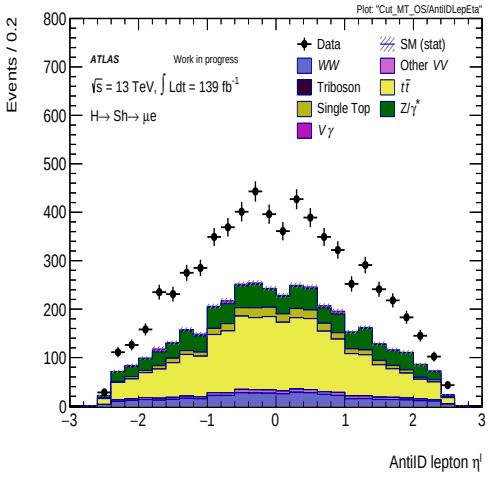
(a) $e\mu$: muon p_{T}



(b) $e\mu$: muon η

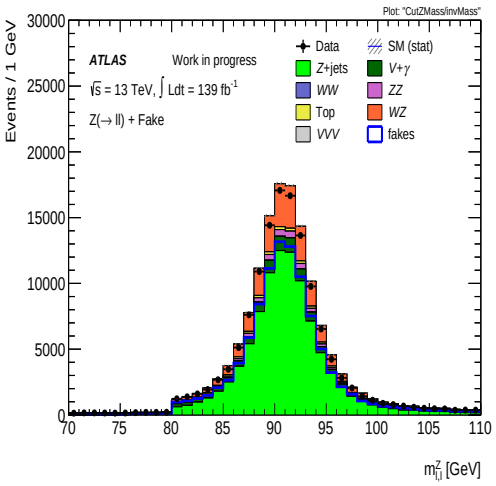


(c) μe: muon p_{T}

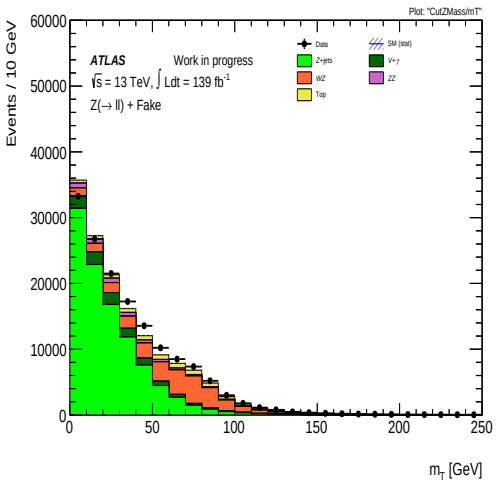


(d) μe: muon η

Figure 6.3: W+jets control region plots for muon fakes in the $e\mu$ and μe channel in the top and bottom, respectively. The plots show the composition of data and EW subtraction; the W+jets is taken as the difference between the data and the other processes. The muon fake p_{T} distribution is shown in the left and muon fake η in the right.



(a) $ee + \mu\mu$, $m_{\ell\ell}$: linear scale



(b) $ee + \mu\mu$, 3^{rd}-lepton m_{T}: linear scale

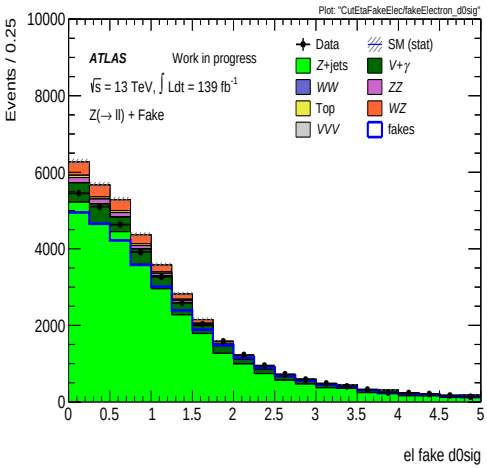
Figure 6.4: The invariant mass distribution of the Z decay lepton candidates (a) and transverse mass distributions of the additional lepton (b).

tween these two-dimensional distributions.

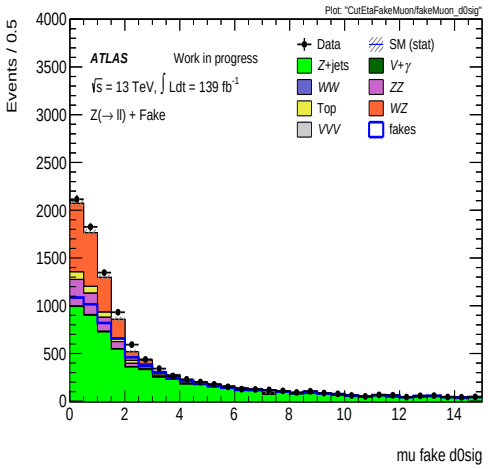
6.4.1 Event Selection

Z+jets events are selected in a sample of events containing exactly three lepton candidates. Events are triggered using the single lepton trigger selection described in Chapter 4 Table 4.1. The two leading lepton candidates are pre-selected as candidates for the $Z \to \ell\ell$ decay leptons, and thus they are required to have same flavour and opposite charge. Both leptons are required to pass the identification cuts described in Table 6.1 and to have a transverse momentum $p_{\mathrm{T}} > 27$ GeV. The di-lepton pair is required to be consistent with a Z boson decay. The $m_{\ell\ell}$ is thus required to be within the interval 70 GeV $< m_{\ell\ell} <$ 110 GeV. In order to reduce the amount of $V\gamma$ background in the case of the additional lepton being an electron, the window cut is further tightened to 80 GeV $< m_{\ell\ell} <$ 110 GeV. Figure 6.4 (a) shows $m_{\ell\ell}$ distribution after the requirement.

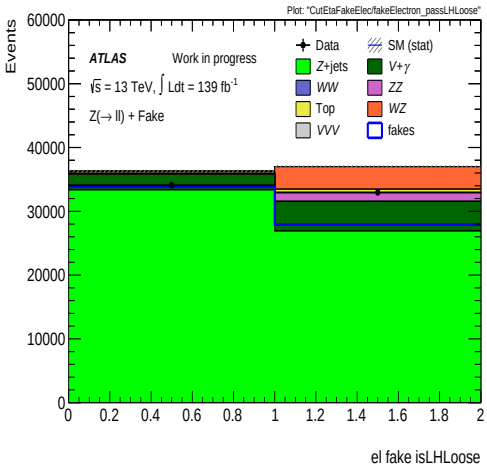
Additionally, in order to reduce the amount of three prompt lepton processes, the event is vetoed if it contains an additional lepton with $m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}E_{\mathrm{T}}^{miss}(1 - cos\Delta\phi\ell\nu)} > 50$ GeV.



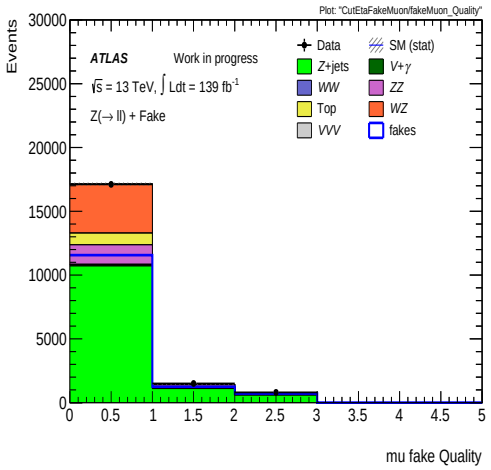
(a) electron d_0-significance



(b) muon d_0-significance



(c) electron LHLoose



(d) muon quality

Figure 6.5: The distributions for the fake electron and muon d_0 significance are shown at the top. Electron and muon quality are shown at the bottom for the full η range.

This requirement highly rejects the WZ background which is reduced by 68% after the cut is applied. Figure 6.4 (b) shows m_{T} distribution of the additional lepton. For events passing the set of requirements above, only those with exactly one additional lepton, e or μ, are selected. The third lepton is then classified as "id" or "anti-id" attending to the definitions in Table 6.1.

The d_0 significance distributions for the fake electron and muon are shown in Figure 6.5 (a) and (b) for the full η range. The electron and muon quality is shown in Figure 6.5 (c) and (d). Following the requirements in Table 6.1, the two variables d_0 significance and lepton quality help to distinguish between "id" and "anti-id". From Figure 6.5 (b) a region where fake muon d_0 significance is less than three represents the "id" region with most of the contribution from the WZ process, as expected. A region where the d_0 significance is greater than three represents the "anti-id" composed by the Z+jets. Figure 6.5 (d) shows the fake muon quality. The three bins represents the "Tight", "Medium" and "Loose" selections. The first bin represents the muon "id" region which has most of the WZ contribution as expected. The second bin represents the "anti-id" region expected to have a higher contribution of the Z+jets. Figure 6.5 (c) shows the fake electron "LHLoose" quality and this region is composed of the Z+jets when the requirement is not applied.

The set of prompt lepton backgrounds considered for the Z+jets FF includes $V\gamma$, diboson production processes such as WW, WZ, and ZZ, as well as $t\bar{t}$ and single top. The $V\gamma$ background is estimated using a dedicated MC sample generated with Sherpa. To avoid overlaps of the $V+$ jets samples and the $V+\gamma$ samples, events in the nominal $V+$ jets samples, Powheg and Sherpa, are removed if an additional photon is found at a distance $\Delta R > 0.1$ from the leptons originating from the V boson decay. This removes the overlap with the $V+\gamma$ sample, which was generated using an explicit cut of $\Delta R > 0.1$ at the event generation level. Triboson processes and $t\bar{t}$ production

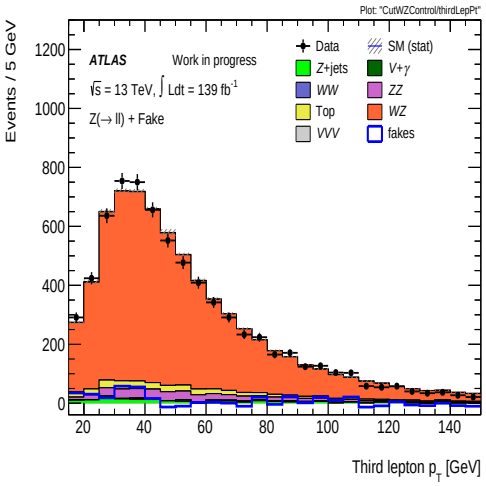
in association with a vector boson are also considered, although they represent a small percentage of the events in the total yield. These backgrounds are estimated from MC simulation and they are subtracted from the data. Once leptons in the "id" and "anti-id" category are selected, the FFs are obtained from Equation 6.1.

6.4.2 Normalisation of the WZ Background

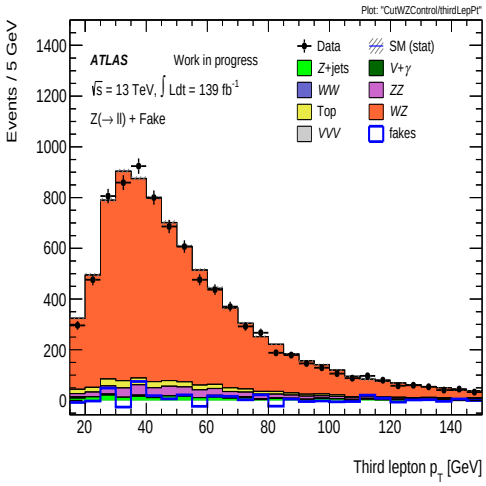
Previous ATLAS measurements of the WZ production cross section [93], as well as studies made within the HZZ group [94] suggest that the MC predictions underestimate the WZ cross section. To check how this effect affects our fake estimation and to derive a convenient normalisation factor for the SHERPA sample, a WZ enriched sample is selected by requiring the event to have an additional, fully identified lepton with $m_{\mathrm{T}} > 50$ GeV besides the two leptons from the Z decay. As explained above, the two leptons from the Z decay are selected as the pair with an invariant mass closest to the Z mass pole. To calculate the WZ normalisation factor, N_f^{WZ}, the data events are subtracted from all the MC background samples except WZ and divided by the WZ background from MC simulation. All the calculation is performed at the WZ CR, $m_{\mathrm{T}} > 50$ GeV, using Equation 6.4:

$$N_f^{WZ} = \frac{Data - MC}{WZ}. \tag{6.4}$$

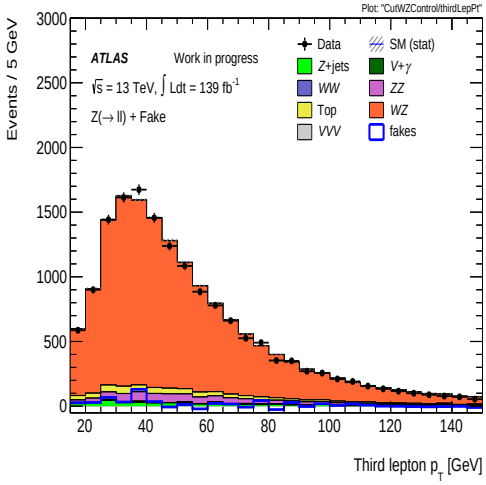
This calculation often relies on different Z+jets generators, POWHEG and SHERPA. In this case, the SHERPA sample performs better accuracy with the calculated N_f^{WZ} of 1.00 ± 0.02 with statistical uncertainties only. Kinematic distributions in the WZ enriched region have been checked to validate the modeling. Figure 6.6 shows the kinematic distributions for the third lepton in the ee, $\mu\mu$ and ee+$\mu\mu$ channel after the normalisation has been applied. From the plots, it is visible that a reasonable agreement between data and MC is obtained after applying the normalisation factor. Note that no systematic uncertainties are performed at this analysis level.



(a) ee, 3^{rd}-lepton p_{T}: linear scale



(b) $\mu\mu$, 3^{rd}-lepton p_{T}: linear scale



(c) $ee + \mu\mu$, 3^{rd}-lepton p_{T}: linear scale

Figure 6.6: Transverse momentum distributions of the additional lepton in the region with $m_{\mathrm{T}} > 50$ GeV after the normalisation factor is applied to the WZ background, (a) shows the $Z \to ee$ channel, (b) the $Z \to \mu\mu$ channel and (c) the combined $ee + \mu\mu$ channel.

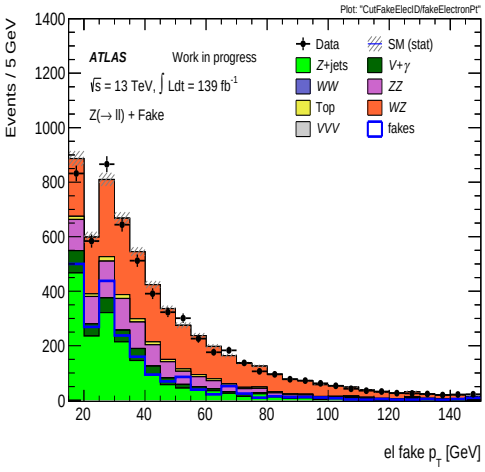
Table 6.2: The Z+jets FF CR showing the electron and the muon fake type in the "id" and "anti-id" regions. The normalisation of the WZ background is applied. The fake yield is shown in the last column.

Cuts	WZ	ZZ	Top	Data	Data − prompt
Fake type: e	3893.80 ± 15.19	1558.82 ± 12.66	181.34 ± 1.21	79471	73828.67 ± 282.60
ID	2851.74 ± 12.63	956.81 ± 6.79	120.55 ± 0.99	6045	2109.63 ± 79.06
Anti-ID	586.99 ± 7.10	369.06 ± 8.02	31.07 ± 0.51	18766	17777.70 ± 137.41
Fake type: μ	4166.22 ± 14.56	1672.63 ± 17.56	175.53 ± 1.21	5157	728.09 ± 73.92
ID	3270.20 ± 13.08	1020.12 ± 11.62	131.28 ± 1.04	996	410.67 ± 33.04
Anti-ID	745.83 ± 5.87	599.62 ± 12.87	38.68 ± 0.58	13514	12128.29 ± 117.11

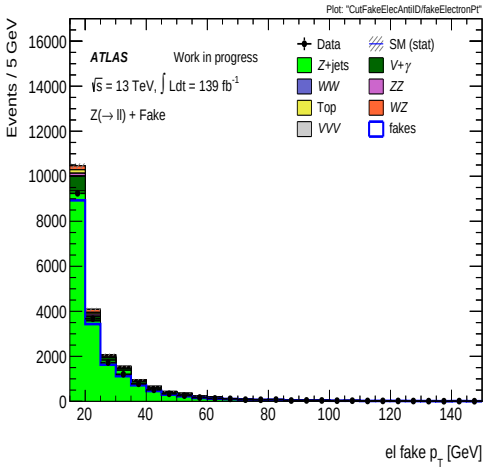
6.4.3 Computation of the Fake Factor

As mentioned previously in Section 6.2 the FF is obtained from Equation 6.1. Table 6.2 shows the Z+jets FF CR with the electron and muon fake type in the "id" and "anti-id" regions. The fake contribution in the last column is obtained by subtracting MC processes with only three-prompt leptons from the data. These MC backgrounds processes are WZ, ZZ, and some top contribution: the $t\bar{t}+W$ and $t\bar{t}+Z$. The WZ background is dominant in the "id" region as expected. The "anti-id" region is dominated by the Z+jets, fakes. The full cutflow of the Z+jets selection used for the fake estimation is shown in Appendix B Table B.1.

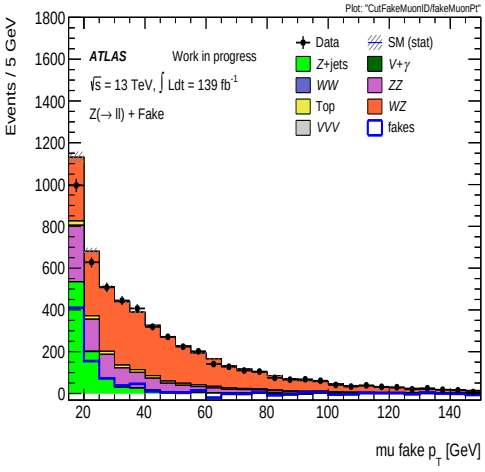
Figure 6.7 and Figure 6.8 show the transverse momentum and pseudorapididty of the "id" and "anti-id" fake candidate leptons for electrons and muons, separately. All the plots show the data in black dots and three-prompt lepton processes in the coloured histograms, except the green histogram, which is showing the Z+jets prediction by Sherpa. This is included only for comparison. The blue line shows the fakes contribution computed as the data minus the processes with three-prompt leptons. The electron p_{T} and η spectra in the "id" show that the fake contribution differs in each p_{T} bin and also depends on the η. Hence this motivates the computation of FF in p_{T} bins and split in η. The logarithmic scale plots are shown in Appendix A Section A.2.



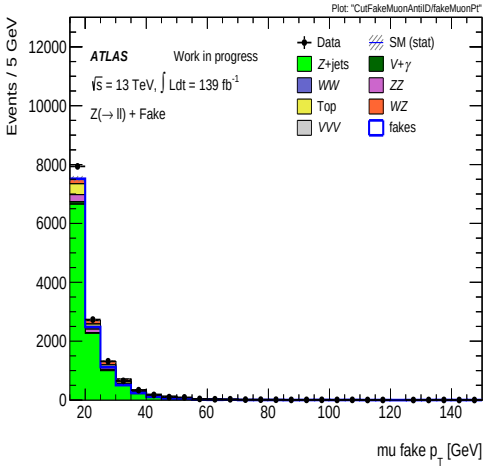
(a) "id" electron p_{T}: linear scale



(b) "anti-id" electron p_{T}: linear scale

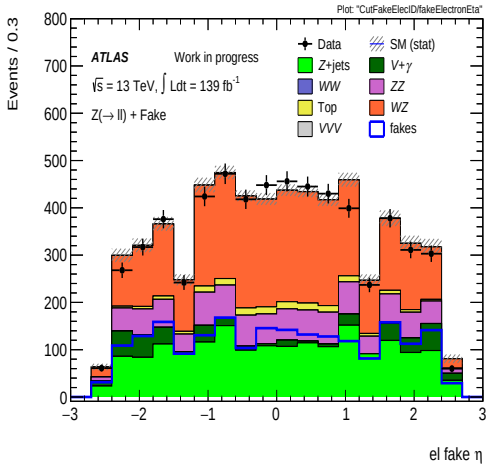


(c) "id" muon p_{T}: linear scale

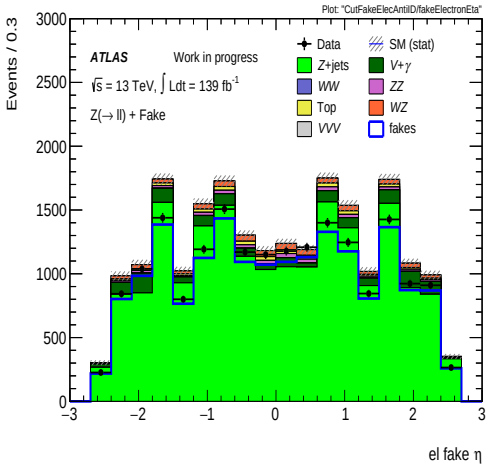


(d) "anti-id" muon p_{T}: linear scale

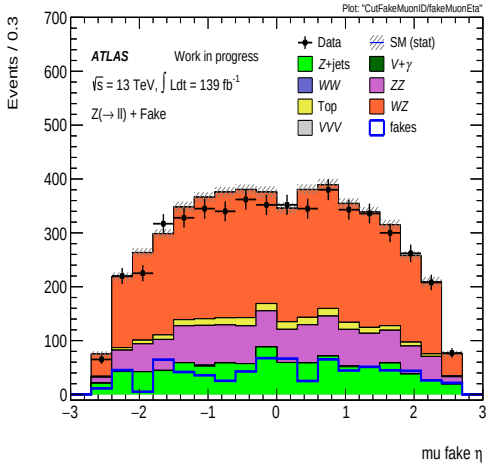
Figure 6.7: Transverse momentum for additional electron and muon in the "id" and "anti-id" categories. The fake contribution shown in blue line is computed by subtracting from the data the MC processes with three- prompt leptons only. The normalisation factor is applied to the WZ background.



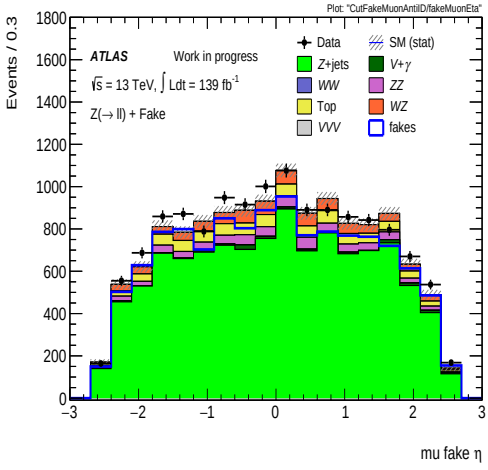
(a) "id" electron η: linear scale



(b) "anti-id" electron η: linear scale



(c) "id" muon η: linear scale



(d) "anti-id" muon η: linear scale

Figure 6.8: pseudorapididty for additional electron and muon in the "id" and "anti-id" categories. The fake contribution shown in blue line is computed by subtracting from the data the MC processes with only three-prompt leptons. The normalisation factor is applied to the WZ background.

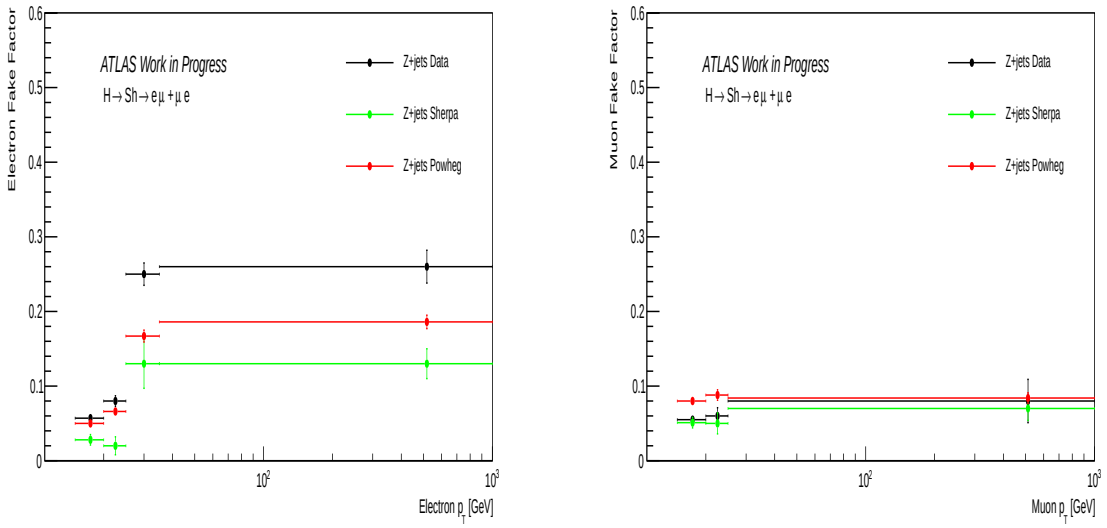
Table 6.3: Numerical results for electron FF in the full $|\eta|$ range, the central and forward regions, $|\eta| < 1.5$ and $1.5 < |\eta| < 2.5$, together with Monte Carlo expectations from SHERPA and POWHEG. The uncertainties are statistical only.

p_{T} range, GeV	Data − prompt	SHERPA	POWHEG
15 - 20	0.056 ± 0.003	0.028 ± 0.007	0.050 ± 0.002
20 - 25	0.078 ± 0.007	0.020 ± 0.012	0.066 ± 0.005
25 - 35	0.250 ± 0.015	0.130 ± 0.033	0.167 ± 0.008
35 - 1000	0.260 ± 0.022	0.130 ± 0.020	0.186 ± 0.009

Table 6.4: Numerical results for muon FF in the full $|\eta|$ range, $|\eta| < 2.5$ together with MC expectations from SHERPA and POWHEG. The uncertainties are statistical only.

p_{T} range, GeV	Data − prompt	SHERPA	POWHEG
15 - 20	0.055 ± 0.004	0.051 ± 0.007	0.080 ± 0.004
20 - 25	0.060 ± 0.011	0.050 ± 0.014	0.088 ± 0.007
25 - 1000	0.080 ± 0.029	0.070 ± 0.016	0.084 ± 0.007

The FF is finally computed as a binned ratio in p_{T} using four bins in p_{T}, [15, 20, 25, 35, 1000], [15, 20, 25, 1000] GeV for electrons and muons, respectively. For electrons, the FF is computed separately in the central and forward regions, $|\eta| < 1.5$ and $1.5 < |\eta| < 2.5$, excluding the crack region $1.37 < |\eta| < 1.52$. For muons, there is no statistically significant difference observed between low and high pseudorapididty regions. Thus the FF is integrated over $|\eta|$ to gain statistical precision. Table 6.3 contains the numerical results for the electron FFs, while Table 6.4 contains the numerical results for the muon FFs. The FFs obtained from data are here compared with Z+jets MC prediction using POWHEG and SHERPA. The p_{T}-differential distributions of the electron FF for the inclusive pseudorapididty bin are shown in Figure 6.9. It is visible that the fake estimation in Z+jets events is dominated by statistical uncertainties, the effect is more emphasised for the muon FFs. It can be noted that the FFs estimated with MC generators reasonably reproduce the FFs in data for muons. For electrons in the full η range, the MC predictions are mostly below the data. Within statistical uncertainties, the FFs are in a reasonable agreement for the muon case. This illustrates that using a MC generator to estimate the fakes instead of a data-driven technique may underestimate the number of fake leptons in the analysis.



(a) electron (b) muon

Figure 6.9: Differential distributions of the electron and muon FF as a function of the fake candidate transverse momentum in the pseudorapididty bin. MC expectations from POWHEG and SHERPA are also shown. The uncertainties are statistical only.

6.5 Closure Test

The closure test is the method used to validate the fake contribution in the CR after applying the obtained FFs. To perform this validation the number of W+jets events in the SR, 2-id leptons, are compared to the number of W+jets events in the CR of id + ant-id after applying corresponding electron and muon FFs. For the closure test, only W+jets MC are used. The closure test is performed at the end of the SR selection. Equation 6.5 relates the number of W+jets in the SR with the W+jets in the CR when the FFs are applied. The relative difference between the values obtained in Equation 6.5 is taken as the uncertainties from the closure test:

$$N_{id+id}^{W+jets(MC)} = FF_e \times N_{id+anti-id,e}^{W+jets(MC)} + FF_\mu \times N_{id+anti-id,\mu}^{W+jets(MC)}, \tag{6.5}$$

where the term FF_e represents the electron FFs and the anti-id in this case is the fake electron, $N_{id+anti-id,e}^{W+jets(MC)}$, and FF_μ represents the muon FFs and the "anti-id" in

Table 6.5: The number of W+jets events in the electron fake CR with FF in the binning of p_{T}. The column shows the W+jets events after applying FF in each p_{T} bins. The last raw represents the sum of the events for all the p_{T} bins.

Electron p_{T} range	FF_e	W+jets MC	W+jets×FF$_e$
15-20	0.056 ± 0.003	1068.51 ± 99.73	59.83 ± 8.79
20-25	0.078 ± 0.007	433.52 ± 66.30	33.81 ± 8.21
25-35	0.250 ± 0.015	361.11 ± 59.69	90.27 ± 20.34
25-1000	0.250 ± 0.022	489.13 ± 71.99	122.28 ± 28.76
Tota Events			306.20 ± 17.11

Table 6.6: The number of W+jets events in the muon fake CR with FF in the binning of p_{T}. The column shows the W+jets events after applying FF in each p_{T} binning. The last raw represents the sum of the events for all the p_{T} binning.

Muon p_T range	FF_μ	W+jets MC	W+jets×FF$_\mu$
15-20	0.055 ± 0.003	639.49 ± 73.51	35.17 ± 6.60
20-25	0.060 ± 0.011	227.61 ± 45.97	13.66 ± 5.26
25-1000	0.080 ± 0.029	181.38 ± 38.99	14.51 ± 8.38
Total Events			63.34 ± 20.24

this case is the fake muon, $N^{W+jets(MC)}_{id+anti-id,\mu}$. Table 6.5 and Table 6.6 show the number of W+jets events for fake electron and muon at the end of the selection in each p_{T} bin. The total number of W+jets events in the SR is obtained by adding together the last rows of Table 6.5 and Table 6.6. To evaluate the corresponding uncertainty from the closure test the events in the CR and SR in Table 6.7 are compared. The closure test shows that the results are compatible with about $\pm 1\ \sigma$ given the associated statistical uncertainty. The uncertainty associated to the closure test is obtained from the relative difference and it is found to be about ∼26%.

The full cutflow for the closure test showing W+jets events in the CR of id + anti-id electron and muon at the end of the selection is found in the Appendix B Section B.3.

Table 6.7: The summation of the number of W+jets events in CR of fake electrons and muons together with the number of W+jets events in SR at the end of the selection.

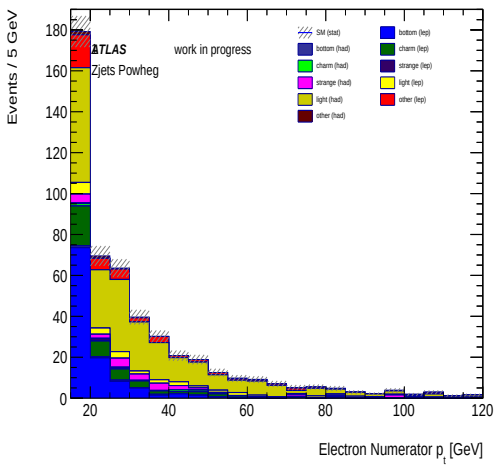
$\sum W$+jets×FF$_{e,\mu}$	W+jets SR from MC
369.54 ± 37.55	479.74 ± 68.40

6.6 Flavour Composition studies

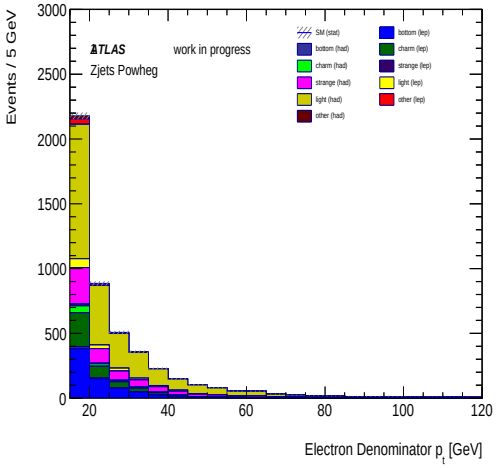
A natural complication in the accurate estimation of the fake background contribution is that fake leptons can come from many different sources whose relative abundance varies between samples. If these fake sources have separate chances of becoming id and anti-id, this, in turn, will manifest as a discrepancy between the FF of the sample. To understand this discrepancy, a flavour composition study of fake leptons based on MC is carried out on the samples of W+jets and Z+jets Powheg. In the W+jets samples, fake leptons are selected by performing an event selection similar to the analysis selection. Furthermore, the reconstructed lepton from the leptonically decaying W is identified by truth matching, leaving the other lepton to be classified as the fake. Truth refers to an object in the MC simulation, which can be matched to the corresponding reconstructed objects. For example, one can ask for a reconstructed object what "true" particle it corresponds to. An event selection similar to the FF measurement of Z+jets presented in Section 6.4 is performed. The two Z candidate leptons are also matched to the Z truth and the additional lepton is classified as the fake lepton.

6.6.1 Flavour Matching Scheme

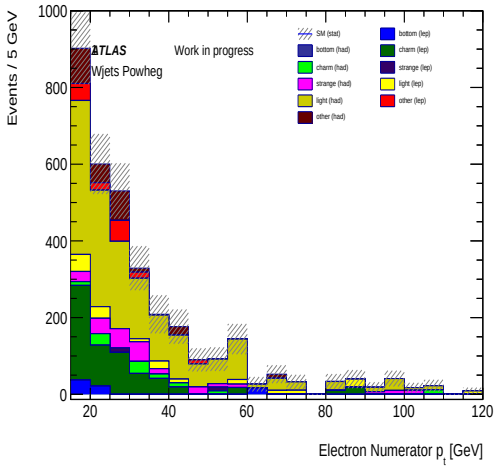
A fake flavour matching scheme is designed based on truth objects found to be in close proximity to the reconstructed fake-lepton candidate to study the flavour composition of the fake lepton samples. The algorithm begins by searching for a bottom truth object, either quark or meson/baryon state, in ΔR(lep, jet) < 0.4. If none is detected, then the process repeats in that order for charm, strange, and light objects. If still nothing is detected, the fake lepton is marked as "other". The fake lepton is marked



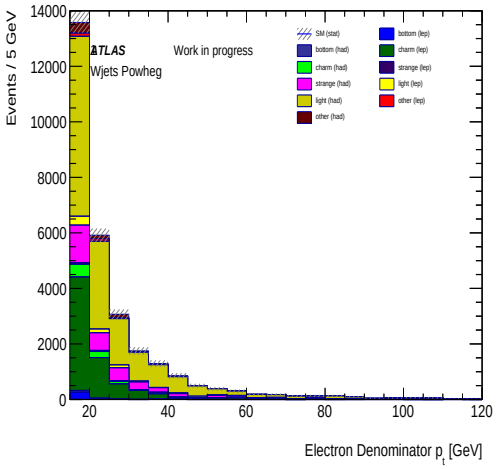
(a) "id" electron p_{T}



(b) "anti-id" electron p_{T}



(c) "id" electron p_{T}



(d) "anti-id" electron p_{T}

Figure 6.10: Comparison of the flavour composition in electron fakes between the Z+jets, top, and W+jets, bottom, samples. The left shows fake "id" leptons, while the right shows fake "anti-id" leptons. The composition of electron fakes is mostly dominated by light flavour jets.

Table 6.8: Table showing the "id" electron flavour percentages for W+jets and Z+jets from MC simulation.

Sample	bottom%	charm%	strange%	light%	other%
Z+jets	22.97 ± 1.44	9.80 ± 0.81	6.33 ± 0.61	52.25 ± 2.17	8.86 ± 0.77
W+jets SR	2.18 ± 0.86	22.13 ± 2.89	7.81 ± 1.59	55.98 ± 5.07	11.90 ± 2.07

Table 6.9: Table showing the "anti-id" electron flavour percentages for W+jets and Z+jets MC simulation.

Sample	bottom%	charm%	strange%	light%	other%
Z+jets	15.33 ± 0.37	11.78 ± 0.29	15.46 ± 0.33	55.34 ± 0.73	2.09 ± 0.11
W+jets CR	1.42 ± 0.23	26.90 ± 1.11	12.52 ± 0.70	55.71 ± 1.74	3.41 ± 0.36

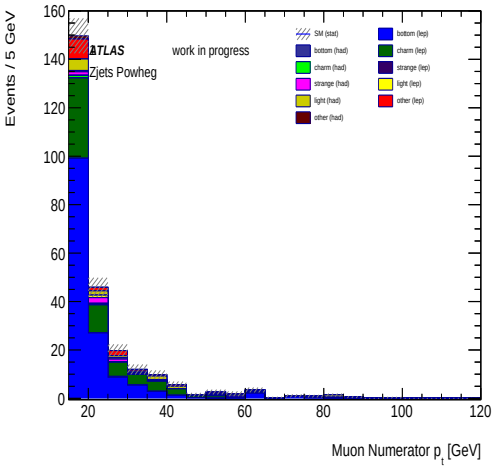
as "leptonic" at each stage of the process if there is a truth lepton within the range of $\Delta R < 0.03$. If not, it is classified as "hadronic".

6.6.2 Electron Flavour Composition

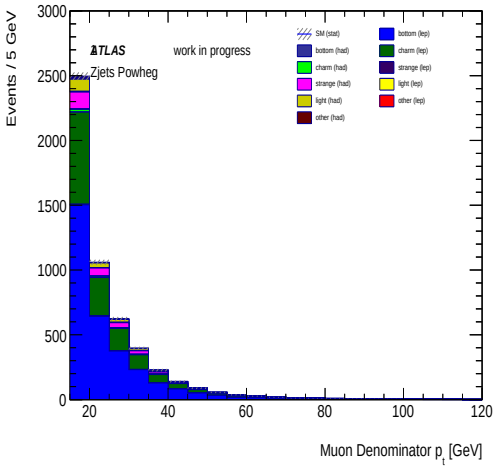
Figure 6.10 shows electron fake flavour composition in the p_T spectrum for SR, CR W+jets and Z+jets samples. The combination of the leptonic and hadronic flavour is represented by the coloured histograms for "bottom", "charm", "strange", "light" and "other". There are some discrepancies in the bottom flavour composition for low p_T electrons between "id" and "anti-id". The flavour fractions of "id" and "anti-id" electrons are listed in Table 6.8 and Table 6.9 respectively. The highest electron fake flavour composition comes from the light flavour in both Z+jets and W+jets samples for "id" and "anti-id". The second highest flavour fraction is from the heavy flavours represented by the bottom, charm and strange. The smallest contribution arises from the "other". Table 6.10 lists the electron FFs separated by flavour in each sample. The FFs are mostly similar except for strange and this is acceptable given the relatively low

Table 6.10: The electron FFs in each flavour category for W+jets and Z+jets MC simulation.

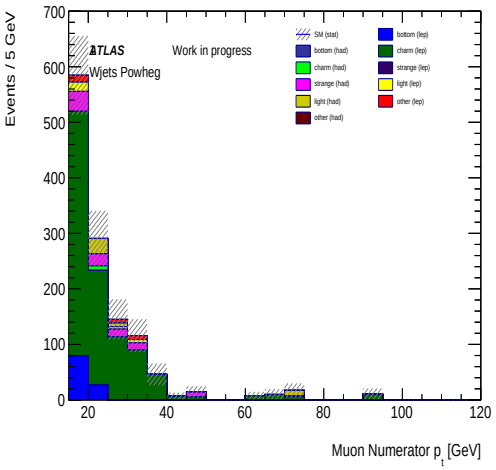
Sample	bottom FF	charm FF	strange FF	light FF	other FF
Z+jets	0.157 ± 0.010	0.085 ± 0.007	0.043 ± 0.004	0.099 ± 0.003	0.443 ± 0.044
W+jets CR	0.180 ± 0.077	0.100 ± 0.012	0.070 ± 0.015	0.119 ± 0.009	0.410 ± 0.780



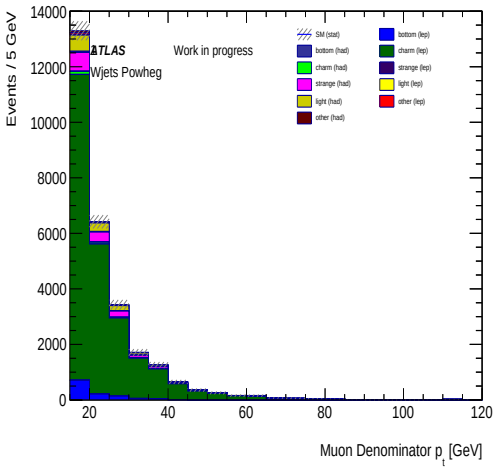
(a) "id" muon p_{T}



(b) "anti-id" muon p_{T}



(c) "id" muon p_{T}



(d) "anti-id" muon p_{T}

Figure 6.11: Comparison of the flavour composition in muon fakes between the Z+jets, top, and W+jets, bottom, samples. The left shows fake "id" leptons, while the right shows fake "anti-id" leptons. The composition of muon fakes is dominated by the leptonic decay of heavy flavour mesons. The W+jets sample contains the largest contribution from c-mesons, which is expected due to an asymmetry in the W+charm process

Table 6.11: Table showing the fractions of the muon "anti-id" flavour for W+jets and Z+jets MC simulation.

Sample	bottom%	charm%	strange%	light%	other%
Z+jets	59.93 ± 0.83	29.27 ± 0.48	6.09 ± 0.19	4.01 ± 0.15	0.71 ± 0.06
W+jets CR	4.15 ± 0.36	84.46 ± 2.24	5.57 ± 0.44	4.90± 0.41	0.93 ± 0.17

Table 6.12: Table showing the fractions of the muon "id" flavour composition for W+jets and Z+jets samples from MC simulation.

Sample	bottom%	charm%	strange%	light%	other%
Z+jets	57.73 ± 3.61	26.00 ± 2.01	3.45 ± 0.63	6.51 ± 0.89	6.31 ± 0.89
W+jets SR	8.47 ± 2.46	76.17 ± 9.64	7.50 ± 2.32	5.79 ± 1.20	2.08 ± 1.06

contribution ∼10% of strange flavour from Table 6.8 and Table 6.9.

6.6.3 Muon Fake Flavour Composition

Figure 6.11 shows the combination of the leptonic and hadronic muon fake flavour composition in the p_{T} spectrum for SR, CR W+jets and Z+jets samples from MC simulation. The combination of the leptonic and hadronic flavour is represented by the coloured histograms for "bottom", "charm", "strange", "light" and "other". The flavour fractions for "id" and "anti-id" muons are summarised in Table 6.11 and Table 6.12 respectively. The composition of muon fakes is largely dominated by heavy flavour decay with large fraction coming from the "bottom" in the Z+jets CR. The fraction of charm is dominant in both W+jets CR and SR. Finally, Table 6.13 lists the fake muon FFs separated by flavour for each sample. The difference in FFs are not negligible hence, these flavour differences are taken into account as the uncertainty of this method from the closure test results in Table 6.7.

The full cutflows on electron and muon flavour composition studies are shown in Appendix B Section B.4. Other studies on the FF dependency on b-jets are presented in Appendix C in order to validate the study of the flavour composition.

Table 6.13: The muon FFs for each flavour category for W+jets and Z+jets from MC.

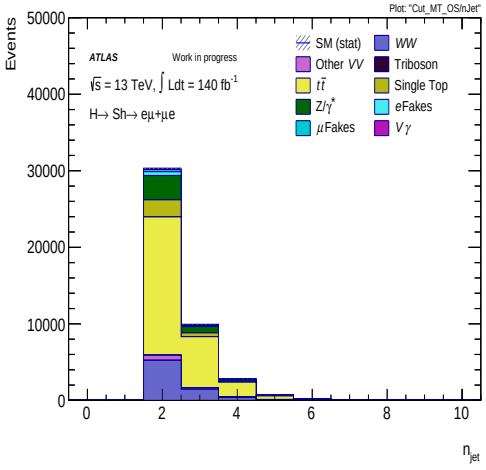
Sample	bottom FF	charm FF	strange FF	light FF	other FF
Z+jets	0.048 ± 0.002	0.044± 0.003	0.028 ± 0.005	0.080 ± 0.011	0.437 ± 0.071
W+jets CR	0.090 ± 0.027	0.040 ± 0.004	0.060 ± 0.018	0.050 ± 0.018	0.100 ± 0.053

6.7 Fake Estimation in the Signal Region

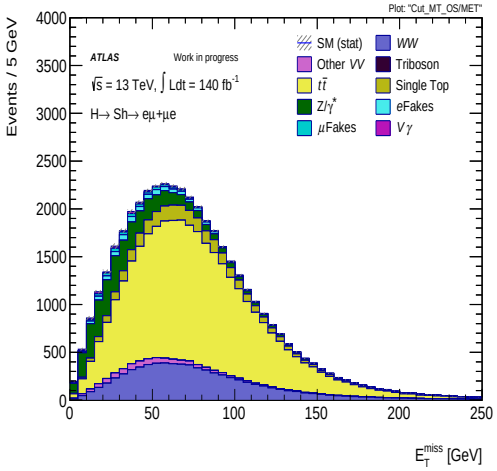
In order to estimate the fake contribution in the SR, the calculated FFs from data in Table 6.3 and Table 6.4 are applied to the W+jets CR. Table 6.14 shows the split in W+jets electron and muon fake contribution in the SR through the end of the selection described in Chapter 5 Table 5.3. The last column shows the total background. The fake estimation yield is found to be $< 3\%$ of the total background and dominated by electrons as expected. Figure 6.12 and Figure 6.13 show the kinematic distributions at the end of selection in the SR for the $e\mu + \mu e$ channel. Figure 6.12 (a) shows the distribution of the number of jets. Most of the fakes are concentrated in the two jet bin and there are almost no fakes at high jet multiplicity. Figure 6.12 (b) illustrates the E_{T}^{miss} distribution with fakes concentrated at $E_{\mathrm{T}}^{miss} < 60$ GeV with no contributions at a region with high E_{T}^{miss}. Figure 6.12 (c) shows an important variable for the analysis, the $m_{\ell\ell}$. Fakes are located at the low $m_{\ell\ell}$ region, $m_{\ell\ell} < 100$ GeV and this is the region where the $H \rightarrow Sh$ signal is concentrated, hence it is crucial to have a good fake estimation in this region. The m_{T} distribution in Figure 6.12 (d) shows fakes in the region 100 GeV $< m_{\mathrm{T}} <$ 160 GeV. Figure 6.13 (a) shows the $\Delta\phi^{\ell\ell}$ spectrum with a small contribution of fakes throughout $\Delta\phi^{\ell\ell} < 3$ rad. Figure 6.13 (b) shows the transverse mass of the di-lepton system, $p_{\mathrm{T}}^{\ell\ell}$, fakes are concentrated in the 30 GeV $< p_{\mathrm{T}}^{\ell\ell} <$ 60 GeV region. The leading lepton η distribution in Figure 6.13 (c) also shows a small fake contribution throughout the η spectrum. The leading lepton p_{T} spectrum in Figure 6.13 (d) shows fakes are concentrated in the low $p_{\mathrm{T}}^{\ell 0}$ region below 80 GeV. Finally Figure 6.14 shows the leading and the sub-leading lepton p_{T} distributions in the $e\mu$ and μe channels. In the $e\mu$ case, the composition of the muon is dominant. For the μe, more contribution arises from electron as expected.

Table 6.14: Cutflow for the electron and muon fake estimation in the $e\mu + \mu e$ channel. The upper part shows the common initial pre-selection, channel selection, followed by the baseline SR selection.

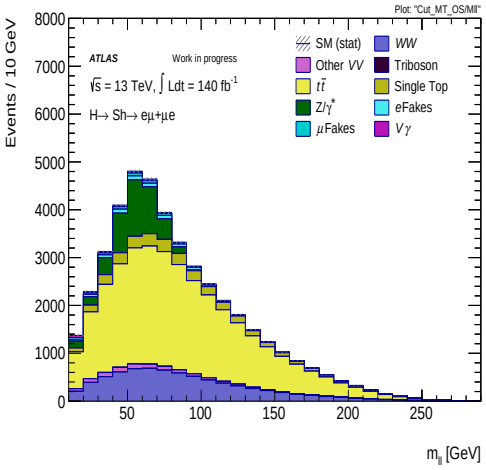
$\sqrt{13}$ TeV, 140 fb^{-1}	eFakes	μFakes	Total bkg
n-jets >= 2	3805.39 ± 80.78	2518.69 ± 36.13	858173.69 ± 283.35
n-$b$jets = 0	1294.88 ± 33.23	694.10 ± 14.64	88238.11 ± 121.01
$m_{\tau\tau} < 70$ GeV	746.70 ± 24.34	419.22 ± 10.64	48452.14 ± 86.06
$m_{\mathrm{T}} < 300$ GeV	701.11 ± 23.68	414.21 ± 10.47	44618.21 ± 84.05



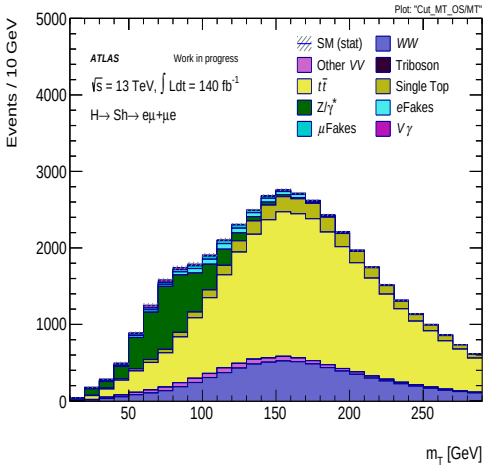
(a) $e\mu + \mu e$: n-jets



(b) $e\mu + \mu e$: E_{T}^{miss}

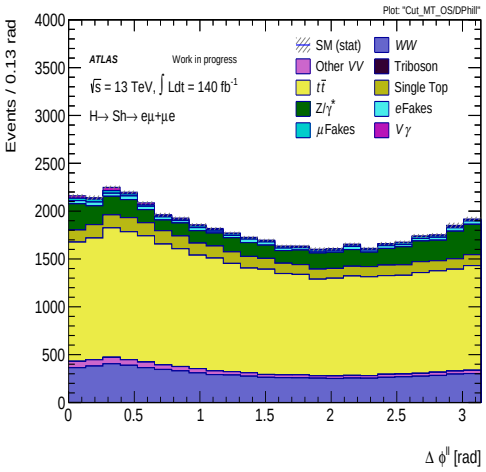


(c) $e\mu + \mu e$: $m_{\ell\ell}$

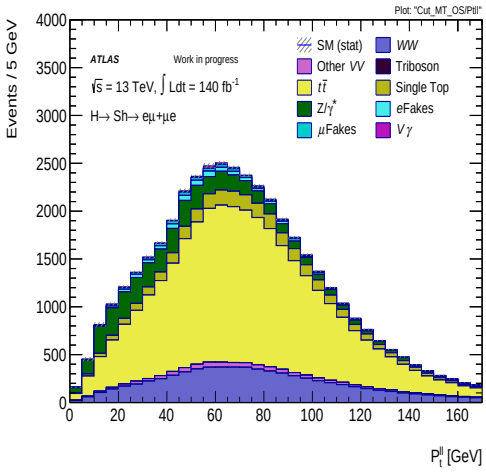


(d) $e\mu + \mu e$: m_T

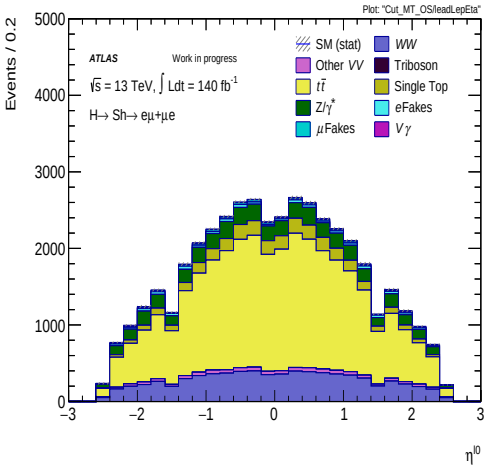
Figure 6.12: The njets, E_{T}^{miss}, $m_{\ell\ell}$ and m_{T} distributions at the end of the selection, the light blue background colour represents the contribution from the electron fakes while the dark blue represents the muon fakes.



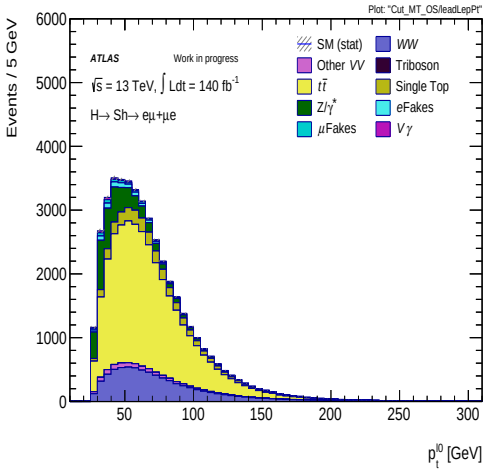
(a) $e\mu + \mu e$: $\Delta\phi^{\ell\ell}$



(b) $e\mu + \mu e$: $p_{\mathrm{T}}^{\ell\ell}$

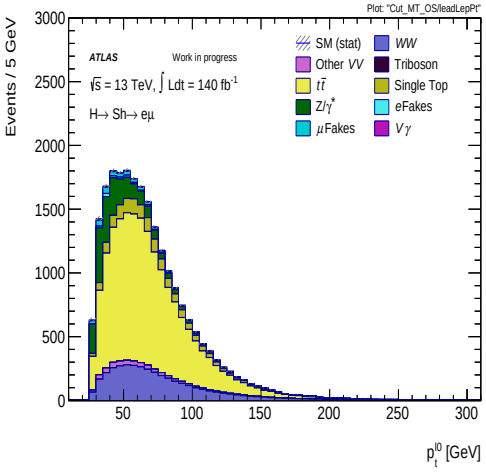


(c) $e\mu + \mu e$: leading lepton η

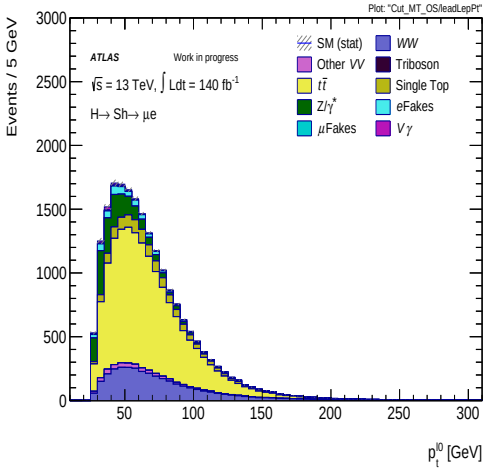


(d) $e\mu + \mu e$: leading lepton p_{T}

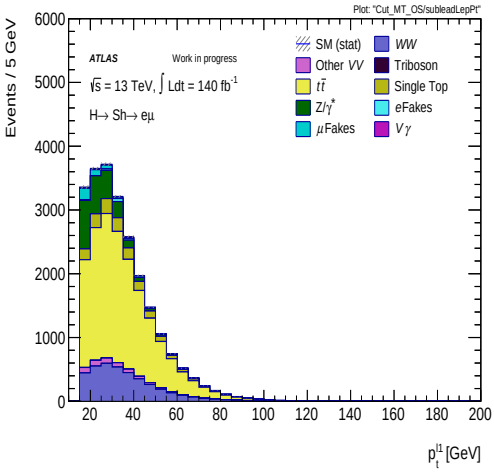
Figure 6.13: The $\Delta\phi^{\ell\ell}$, $p_{\mathrm{T}}^{\ell\ell}$, leading lepton η and leading lepton p_{T} distributions, the light blue background colour represents the contribution from the electron fakes while the dark blue represents the muon fakes.



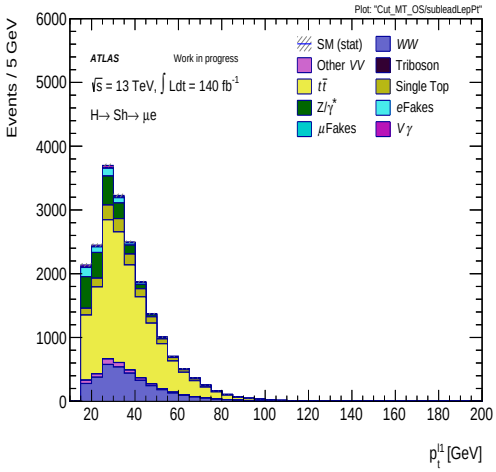
(a) $e\mu$: leading lepton p_{T}



(b) μe: leading lepton p_{T}



(c) $e\mu$: sub-leading lepton p_{T}



(d) μe: sub-leading lepton p_T

Figure 6.14: The leading and sub-leading lepton p_{T} distributions, the light blue background colour represents the contribution from the electron fakes while the dark blue represent the muon fakes.

Chapter 7

Summary and Conclusions

This thesis focuses on the estimation of fake lepton backgrounds in the $H \to Sh \to e^{\pm}\mu^{\mp} + \geq 2$ jets analysis in ATLAS. Background processes can produce signatures in the detector that are similar to that of the Higgs-like scalar decay, hindering the analysis. An important background contribution can arise from the misidentification of an object. The present work focuses on the estimation of the contribution of fakes from the W+jets process that occurs when a jet is misidentified as a lepton: electron or muon. The jet misidentification study is performed using the data-driven Fake Factor method, which selects a control sample of events enriched in the W+jets background and then uses an extrapolation factor to relate these events in the $H \to Sh$ signal region. The fact that misidentification is a rare process that is not expected to be predicted accurately by simulation is motivated by relying on data rather than MC.

The author took part in the first estimation of the fake factor with the Z+jets full Run 2 data samples. The fake factors are measured in bins of transverse momentum and pseudorapidity of the misidentified object. The results on the fake factor obtained from Z+jets in data are compared to the calculation of the fake factors in the MC Z+jets with different generators, Powheg and Sherpa. It is observed that within statistical uncertainties, the fake factors agree in all three samples. The closure test

is performed to check the validity of the method. This is computed by comparing the number of W+jets from MC simulation in the CR and SR at the end of the selection. The closure test shows that the results are compatible with $\sim \pm 1\ \sigma$ given the associated statistical uncertainty. The uncertainty associated with the closure test is obtained from relative difference and it is found to be about 26%.

The flavour composition is performed for both electron and muon in W+jets and Z+jets sample using flavour matching scheme. For the electron fakes, most contribution comes from the light flavour. The flavour composition is highest for W+jets as compared to Z+jets. For the muon fakes, most contribution comes from the heavy flavour jets. The estimation of the fake yield in the SR is found to be very small: $<$ 3% of the total background.

Appendix A

Control Plots

A.1 W+jets Control Region

A.1.1 Electron, Anti-id

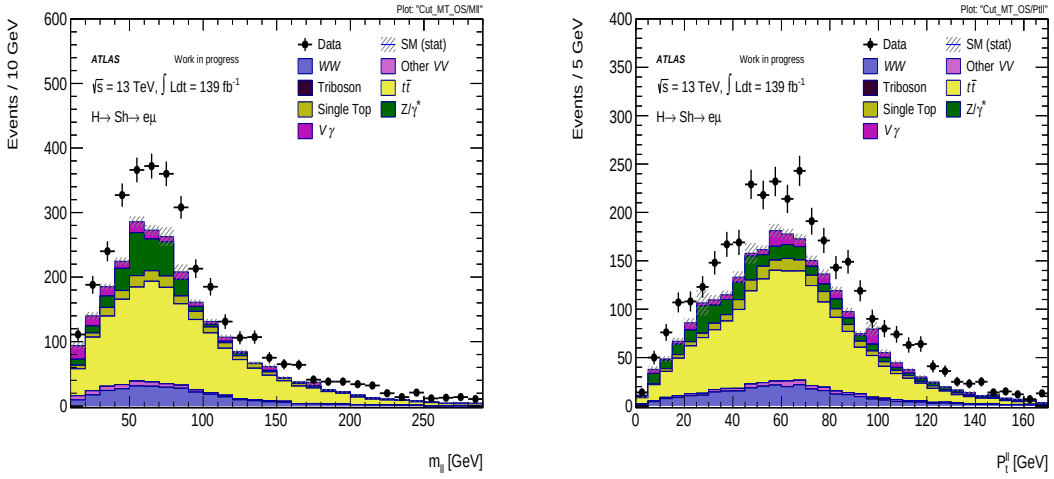
Other kinematic distribution for the W+jets CR are shown in Figure A.1, the $m_{\ell\ell}$ and $p_{\mathrm{T}}^{\ell\ell}$ in $e\mu$ and μe channels respectively.

A.1.2 Muon, Anti-id

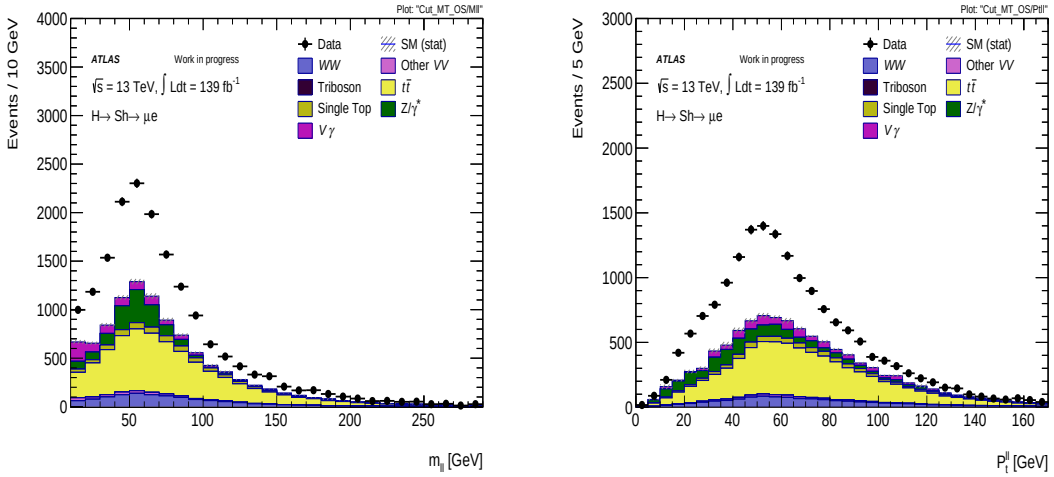
Other kinematic distribution for the W+jets CR are shown in Figure A.2, the $m_{\ell\ell}$ and $p_{\mathrm{T}}^{\ell\ell}$ in $e\mu$ and μe channels respectively.

A.2 Z+jets Fake Factor Control Region

Other kinematic distributions for the Z+jets fake factor CR are shown in Figure A.3 and Figure A.4.

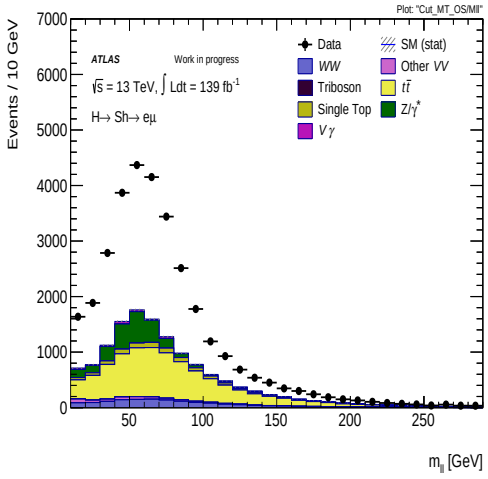


(a) $e\mu$: $m_{\ell\ell}$ (b) $e\mu$: $p_{\mathrm{T}}^{\ell\ell}$

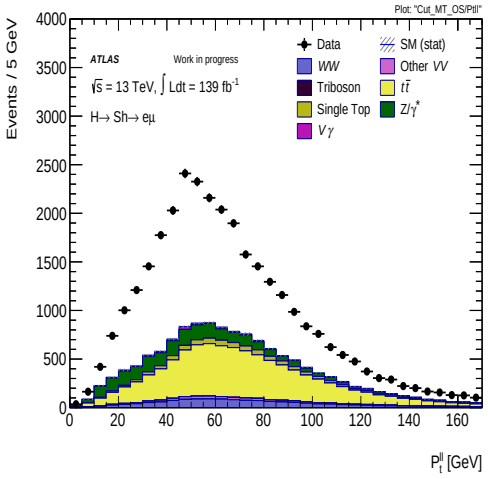


(c) μe: $m_{\ell\ell}$ (d) μe: $p_{\mathrm{T}}^{\ell\ell}$

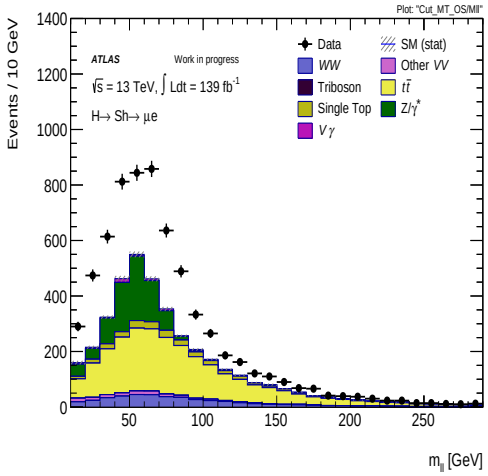
Figure A.1: W+jets control region plots for electron fakes at the end of the selection in the $e\mu$ and μe channel top and bottom respectively. The plots show the composition of data and EW subtraction the W+jets is taken as the difference between the data and the other processes. The $m_{\ell\ell}$ distribution is shown in the left and the $p_{\mathrm{T}}^{\ell\ell}$ in the right.



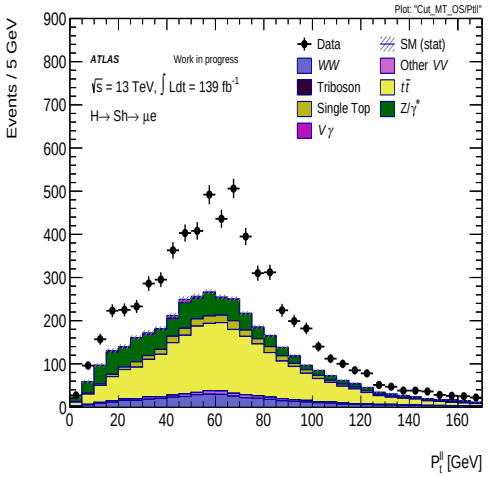
(a) $e\mu$: $m_{\ell\ell}$



(b) $e\mu$: $p_{\mathrm{T}}^{\ell\ell}$

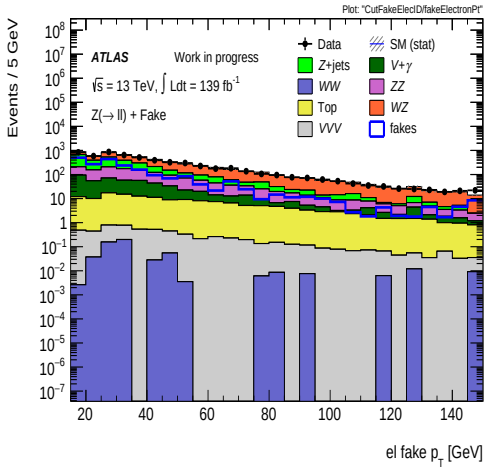


(c) μe: $m_{\ell\ell}$

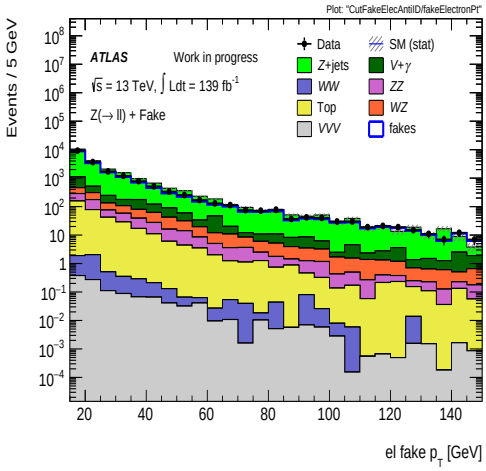


(d) μe: $p_{\mathrm{T}}^{\ell\ell}$

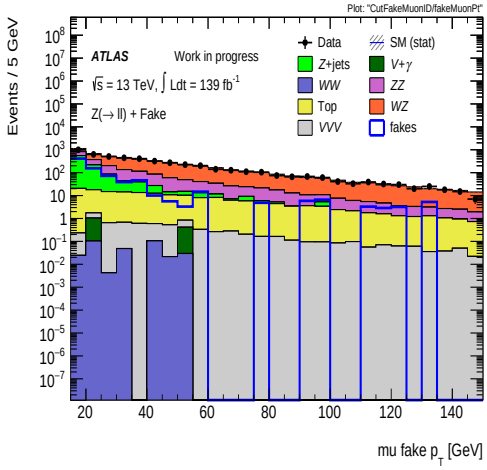
Figure A.2: W+jets control region plots for muon fakes at the end of the selection in the $e\mu$ and μe channel top and bottom respectively. The plots show the composition of data and EW subtraction the W+jets is taken as the difference between the data and the other processes. The $m_{\ell\ell}$ distribution is shown in the left and the $p_{\mathrm{T}}^{\ell\ell}$ in the right.



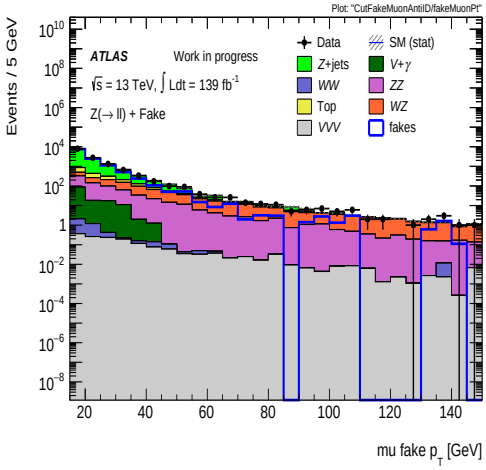
(a) id e fake p_{T}: log scale



(b) anti-id e fake p_{T}: log scale

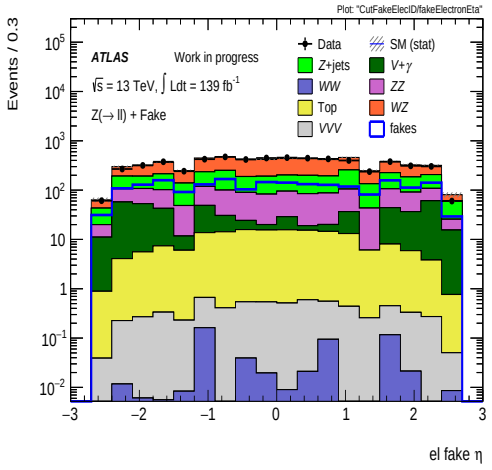


(c) id μ fake p_{T}: log scale

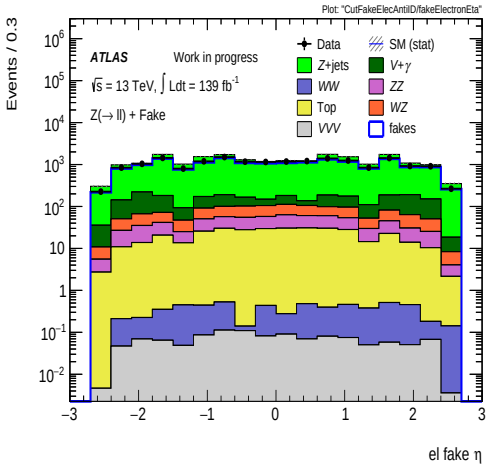


(d) anti-id μ fake p_{T}: log scale

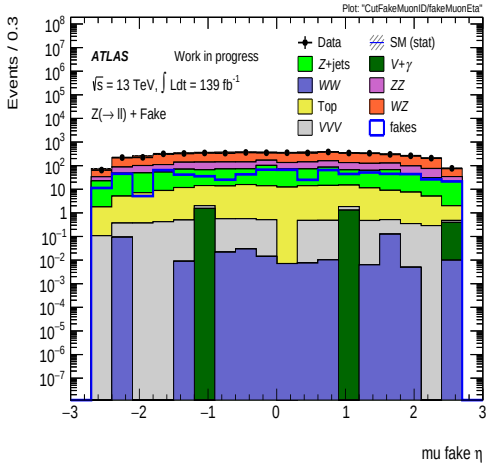
Figure A.3: Transverse momentum for additional electron and muon in the "id" and "anti-id" category. The fake contribution shown in blue line is computed by subtracting from the data the EW processes with three prompt leptons only. The normalization factor is applied to the WZ background. The plots are shown with linear scale.



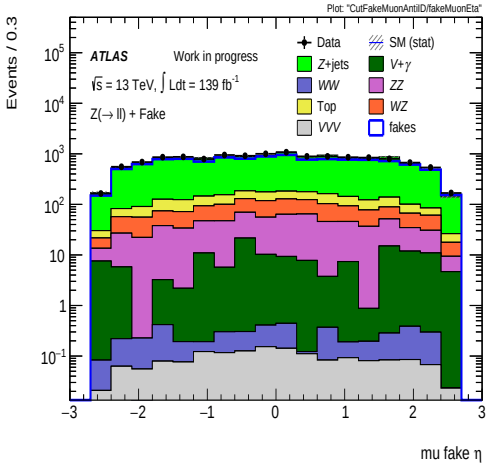
(a) id e fake η: log scale



(b) anti-id e fake η: log scale



(c) id μ fake η: log scale



(d) anti-id μ fake η: log scale

Figure A.4: Pseudorapidity for additional electron and muon in the "id" and "anti-id" categories. The fake contribution shown in blue line is computed by subtracting from the data the EW processes with three prompt leptons only. The normalization factor is applied to the WZ background. The plots are shown with linear scale.

Appendix B

Cutflows

B.1 Z+jets Fake Factor

Table B.1 shows the cutflow for the Z+jets fake estimation selection in the $e\mu + \mu e$ channel. The upper part of the cutflow shows the common selection, followed the fake electron yields in the middle and the fake muon yields in the bottom. Z+jets in the table are estimated with SHERPA and POWHEG. The normalization factor is applied to the WZ background. The Z+jets fake contribution is taken as the data minus the background processes with three prompt leptons in the third column. The rightmost column shows the fake purity, taken as the fake yield divided by the data yield.

Table B.1: Cutflow for the Z+jets fake estimation selection in the $e\mu + \mu e$ channel.

$\sqrt{s} = 137 TeV$, (Full Run 2)	Z+jets Sherpa	Z+jets Powheg	$V+\gamma$	WW	$t\bar{t}$	single top	WZ	ZZ	$t\bar{t}+W$	$t\bar{t}+Z$	VVV	Data	fakes	ZFake purity(%)
Channel Selection	325393.22 ± 2012.42	321108.84 ± 571.95	68961.29 ± 476.93	867.95 ± 7.93	45431.60 ± 56.86	3315.91 ± 22.40	45539.85 ± 69.87	14750.29 ± 111.80	522.31 ± 2.02	1717.35 ± 3.09	234.34 ± 0.92	500245	437480.86 ± 719.47	87.45 ± 0.19
lll final state	325393.22 ± 2012.42	321108.84 ± 571.95	68961.29 ± 476.93	867.95 ± 7.93	45431.60 ± 56.86	3315.91 ± 22.40	45539.85 ± 69.87	14750.29 ± 111.80	522.31 ± 2.02	1717.35 ± 3.09	234.34 ± 0.92	500245	437480.86 ± 719.47	87.45 ± 0.19
ZCand $P_t > 27$ GeV	204580.11 ± 1611.41	202842.93 ± 457.24	38364.15 ± 360.96	709.58 ± 7.01	35135.17 ± 50.15	2598.49 ± 19.90	41271.10 ± 64.68	9752.23 ± 105.67	471.19 ± 1.92	1606.99 ± 2.99	218.54 ± 0.89	328733	275412.94 ± 586.60	83.78 ± 0.23
Overlap: Vgamma/Vjets	168679.83 ± 1461.20	168321.71 ± 417.62	38364.15 ± 360.96	709.58 ± 7.01	35135.17 ± 50.15	2598.49 ± 19.90	41271.10 ± 64.68	9752.23 ± 105.67	471.19 ± 1.92	1606.99 ± 2.99	218.54 ± 0.89	328733	275412.94 ± 586.60	83.78 ± 0.23
Z-tagging	94395.99 ± 1064.83	94513.69 ± 313.06	9674.36 ± 177.75	76.70 ± 2.27	3047.66 ± 14.80	222.23 ± 5.84	25491.10 ± 39.51	4458.72 ± 23.58	58.77 ± 0.67	935.98 ± 2.25	67.78 ± 0.40	132721	101708.65 ± 367.21	76.63 ± 0.35
FakeLep, pT>15 GeV	94395.99 ± 1064.83	94513.69 ± 313.06	9674.36 ± 177.75	76.70 ± 2.27	3047.66 ± 14.80	222.23 ± 5.84	25491.10 ± 39.51	4458.72 ± 23.58	58.77 ± 0.67	935.98 ± 2.25	67.78 ± 0.40	132721	101708.65 ± 367.21	76.63 ± 0.35
WZ veto	83241.55 ± 1037.24	85910.92 ± 298.82	7778.27 ± 162.53	27.23 ± 1.46	1397.25 ± 10.02	107.69 ± 4.07	8060.02 ± 21.04	3231.45 ± 21.65	14.81 ± 0.34	342.06 ± 1.36	17.57 ± 0.19	99789	88123.09 ± 317.34	88.31 ± 0.42
fake type: electron	71442.55 ± 1008.54	72928.31 ± 271.90	7625.68 ± 160.41	23.55 ± 1.39	616.87 ± 6.67	43.30 ± 2.57	3893.80 ± 15.19	1558.82 ± 12.66	6.63 ± 0.23	174.71 ± 0.98	8.36 ± 0.13	79471	73828.67 ± 282.60	92.90 ± 0.48
fake el η cut	57393.29 ± 803.30	60255.50 ± 247.51	7101.19 ± 155.30	19.36 ± 1.29	559.72 ± 6.35	40.35 ± 2.48	3668.10 ± 14.88	1468.67 ± 12.36	6.21 ± 0.22	163.62 ± 0.94	7.91 ± 0.13	67055	61740.50 ± 259.67	92.07 ± 0.53
ID cuts	989.65 ± 113.03	1787.08 ± 43.08	399.46 ± 34.82	0.47 ± 0.19	36.26 ± 1.62	3.25 ± 0.71	2851.74 ± 12.63	956.81 ± 6.79	4.31 ± 0.18	116.24 ± 0.79	6.27 ± 0.11	6045	2109.63 ± 79.06	34.90 ± 1.38
15 < pT <20 GeV	249.66 ± 58.56	466.92 ± 22.28	81.91 ± 15.05	0.00 ± 0.03	5.66 ± 0.65	0.58 ± 0.29	211.76 ± 4.73	114.36 ± 2.57	0.27 ± 0.04	5.02 ± 0.17	0.48 ± 0.03	832	500.11 ± 29.34	60.11 ± 4.10
20 < pT <25 GeV	71.32 ± 32.89	235.40 ± 15.78	44.69 ± 11.12	0.04 ± 0.02	3.37 ± 0.48	0.15 ± 0.15	207.96 ± 3.34	100.51 ± 2.10	0.28 ± 0.05	5.82 ± 0.18	0.40 ± 0.03	584	269.02 ± 24.49	46.07 ± 4.61
25 <pT <35 GeV	297.15 ± 74.69	535.49 ± 23.59	96.63 ± 21.57	0.36 ± 0.16	10.57 ± 0.89	0.82 ± 0.37	562.97 ± 6.13	251.05 ± 3.64	0.72 ± 0.08	18.64 ± 0.32	1.21 ± 0.05	1510	675.40 ± 39.51	44.73 ± 2.86
pT > 35 GeV	371.52 ± 51.85	549.28 ± 23.54	176.23 ± 19.92	0.07 ± 0.10	16.67 ± 1.09	1.70 ± 0.52	1869.05 ± 9.40	490.88 ± 4.67	3.04 ± 0.15	86.76 ± 0.68	4.17 ± 0.09	3119	665.10 ± 56.83	21.32 ± 1.86
Anti-ID cuts	16911.80 ± 429.46	19081.12 ± 140.75	1378.35 ± 69.34	4.84 ± 0.68	299.89 ± 4.65	21.34 ± 1.78	586.99 ± 7.10	369.06 ± 8.02	1.14 ± 0.10	29.93 ± 0.41	1.17 ± 0.05	18766	17777.70 ± 137.41	94.73 ± 1.01
15 < pT <20 GeV	8932.91 ± 328.06	9348.00 ± 98.73	653.26 ± 48.33	1.52 ± 0.52	136.48 ± 3.13	10.13 ± 1.23	173.73 ± 4.62	130.20 ± 4.51	0.37 ± 0.06	8.19 ± 0.21	0.38 ± 0.03	9243	8930.15 ± 96.36	96.62 ± 1.45
20 < pT <25 GeV	2879.68 ± 187.37	3568.42 ± 61.06	207.96 ± 28.11	1.77 ± 0.36	65.37 ± 2.18	4.64 ± 0.83	133.62 ± 3.05	97.72 ± 4.57	0.29 ± 0.05	6.01 ± 0.18	0.27 ± 0.02	3672	3434.08 ± 60.85	93.52 ± 2.26
25 <pT <35 GeV	2336.88 ± 157.78	3205.83 ± 57.51	183.00 ± 28.01	0.66 ± 0.20	59.49 ± 2.06	5.12 ± 0.88	106.75 ± 2.66	62.98 ± 3.24	0.17 ± 0.05	5.42 ± 0.17	0.20 ± 0.02	2925	2749.49 ± 54.25	94.00 ± 2.54
pT > 35 GeV	2762.34 ± 129.67	2958.87 ± 55.03	334.13 ± 29.96	0.88 ± 0.15	38.56 ± 1.68	1.45 ± 0.46	172.90 ± 3.56	78.16 ± 3.54	0.32 ± 0.05	10.32 ± 0.25	0.32 ± 0.03	2926	2663.99 ± 54.33	91.05 ± 2.51
fake type: muon	11799.00 ± 242.31	12982.62 ± 123.95	152.59 ± 26.19	3.68 ± 0.47	780.38 ± 7.48	64.39 ± 3.16	4166.22 ± 14.56	1672.63 ± 17.56	8.18 ± 0.26	167.35 ± 0.95	9.21 ± 0.14	20318	14294.42 ± 144.36	70.35 ± 0.87
fake mu η cut	11068.39 ± 238.57	12409.41 ± 121.09	137.42 ± 25.15	3.47 ± 0.45	743.50 ± 7.30	60.36 ± 3.04	4071.04 ± 14.41	1634.85 ± 17.37	7.99 ± 0.25	164.96 ± 0.94	9.01 ± 0.14	19417	13529.15 ± 141.16	69.68 ± 0.88
ID cuts	537.19 ± 58.14	899.61 ± 32.38	0.88 ± 3.89	0.34 ± 0.14	41.13 ± 1.73	2.97 ± 0.68	3270.20 ± 13.08	1020.12 ± 11.62	5.86 ± 0.22	125.42 ± 0.82	7.31 ± 0.13	5157	728.09 ± 73.92	14.12 ± 1.45
15 < pT <20 GeV	325.90 ± 44.15	534.27 ± 24.83	−0.43 ± 3.76	0.02 ± 0.01	11.09 ± 0.88	0.53 ± 0.28	304.71 ± 6.48	271.31 ± 7.31	0.51 ± 0.07	8.16 ± 0.21	0.64 ± 0.04	996	410.67 ± 33.04	41.23 ± 3.57
20 < pT <25 GeV	97.64 ± 28.73	200.38 ± 15.56	0.93 ± 0.90	0.10 ± 0.09	6.16 ± 0.67	0.33 ± 0.24	308.67 ± 4.07	153.86 ± 5.29	0.48 ± 0.06	9.21 ± 0.22	0.71 ± 0.04	628	155.08 ± 25.93	24.69 ± 4.25
pT > 25 GeV	113.65 ± 24.61	164.96 ± 13.77	0.38 ± 0.38	0.21 ± 0.11	23.88 ± 1.32	2.11 ± 0.57	2656.83 ± 10.61	594.95 ± 7.32	4.86 ± 0.20	108.05 ± 0.76	5.96 ± 0.11	3533	162.34 ± 60.83	4.59 ± 1.72
Anti-ID cuts	9973.61 ± 218.99	10878.15 ± 113.78	132.81 ± 23.79	3.08 ± 0.43	677.05 ± 6.97	55.16 ± 2.90	745.83 ± 5.87	599.62 ± 12.87	2.00 ± 0.13	36.68 ± 0.45	1.58 ± 0.06	13514	12128.29 ± 117.11	89.75 ± 1.16
15 < pT <20 GeV	6384.16 ± 174.50	6652.58 ± 88.69	84.73 ± 18.93	1.72 ± 0.33	331.49 ± 4.88	27.38 ± 2.05	170.56 ± 3.21	239.20 ± 7.86	0.65 ± 0.07	9.71 ± 0.23	0.36 ± 0.03	7938	7517.52 ± 89.50	94.70 ± 1.55
20 < pT <25 GeV	2032.72 ± 102.20	2266.67 ± 52.31	17.05 ± 10.42	0.94 ± 0.24	153.51 ± 3.31	11.57 ± 1.33	118.65 ± 2.41	124.95 ± 5.48	0.35 ± 0.05	6.31 ± 0.18	0.26 ± 0.02	2733	2482.48 ± 52.62	90.83 ± 2.59
pT > 25 GeV	1556.73 ± 84.03	1958.90 ± 48.41	31.04 ± 9.94	0.42 ± 0.12	192.65 ± 3.71	16.20 ± 1.57	456.61 ± 4.27	235.48 ± 8.59	1.00 ± 0.09	20.66 ± 0.34	0.96 ± 0.05	2843	2128.29 ± 54.18	74.86 ± 2.37
WZ control	357.33 ± 44.02	431.71 ± 21.13	82.28 ± 12.38	1.91 ± 0.35	163.16 ± 3.43	12.08 ± 1.35	14489.60 ± 29.81	950.61 ± 6.34	35.22 ± 0.52	474.32 ± 1.60	42.69 ± 0.32	16417	424.55 ± 131.72	2.59 ± 0.80

B.2 W+jets Control Region

B.2.1 Electron, Anti-id

Table B.2 and Table B.3 shows the CR cutflow in the $e\mu$ and μe channels for one id + one anti-id electron. The upper part of the cutflow shows the common selection, followed by the baseline selection. The second column shows composition of the W+jets data-driven which is obtained by EW subtraction from data. Column 3 shows the fake muon estimation represented by W+jets data-driven, and most composition comes from the sub-leading electron case, μe.

B.2.2 Muon, Anti-id

Table B.4 and Table B.5 shows the CR cutflow in the $e\mu$ channel for one id + one anti-id muon. The upper part of the cutflow shows the common selection, followed by the baseline selection. The second column shows composition of the W+jets data-driven which is obtained by EW subtraction from data. Column 3 shows the fake muon estimation represented by W+jets data-driven, and most composition comes from the sub-leading muon case, $e\mu$.

Table B.2: The CR cutflow in the $e\mu$ channel for one id + one anti-id electron.

$\sqrt{s} = 13TeV$, (Full Run 2)	WW	Other VV	$V\gamma$	Triboson	$t\bar{t}$	Single Top	Z/γ^*	W+jets DD	data
Channel Selection	109135.98 ± 108.45	62393.97 ± 196.40	114061.69 ± 714.45	251.67 ± 1.00	1033047.84 ± 218.56	103191.27 ± 126.04	739668.03 ± 1547.06	3587669.55 ± 2960.95	5749420
$p_t^{lead} > 27$ GeV	104950.45 ± 106.14	56809.01 ± 191.22	83528.87 ± 664.06	246.41 ± 0.99	1001704.19 ± 215.32	100694.70 ± 124.56	579111.69 ± 1385.50	2694195.68 ± 2662.99	4621241
$p_t^{sublead} > 15$ GeV	95825.72 ± 101.71	44182.34 ± 183.49	66927.69 ± 631.02	227.40 ± 0.95	889939.45 ± 203.25	90251.22 ± 118.08	399796.83 ± 1093.12	1561067.35 ± 2200.14	3148218
GRL Selection	95825.72 ± 101.71	44182.34 ± 183.49	66927.69 ± 631.02	227.40 ± 0.95	889939.45 ± 203.25	90251.22 ± 118.08	399796.83 ± 1093.12	1561067.35 ± 2200.14	3148218
Jet Cleaning	95825.72 ± 101.71	44182.34 ± 183.49	66927.69 ± 631.02	227.40 ± 0.95	889939.45 ± 203.25	90251.22 ± 118.08	399796.83 ± 1093.12	1561067.35 ± 2200.14	3148218
Overlap: Vgamma/Vjets	95825.72 ± 101.71	44182.34 ± 183.49	66927.69 ± 631.02	227.40 ± 0.95	889939.45 ± 203.25	90251.22 ± 118.08	372283.68 ± 936.10	1588580.49 ± 2126.49	3148218
Trigger Selection	91056.65 ± 99.22	40768.45 ± 169.70	44835.22 ± 595.68	222.71 ± 0.94	846045.66 ± 198.35	86565.93 ± 115.71	314328.23 ± 854.76	756239.16 ± 1832.18	2180062
Trigger Matching	89775.13 ± 98.40	38851.71 ± 132.21	36259.75 ± 570.49	216.66 ± 0.93	833491.53 ± 196.71	85497.94 ± 114.88	290844.20 ± 768.14	722989.08 ± 1758.55	2097926
Only two Leptons	89620.77 ± 98.37	21689.88 ± 127.88	31048.21 ± 256.70	109.55 ± 0.70	815028.58 ± 194.53	83972.26 ± 113.97	263091.51 ± 655.56	723065.23 ± 1612.49	2027626
OS Leptons	87986.59 ± 97.77	11557.64 ± 90.13	18030.82 ± 194.30	77.11 ± 0.60	793163.76 ± 191.91	82090.05 ± 112.73	247324.32 ± 578.04	471964.71 ± 1466.73	1712195
3rd lepton veto + OS	87986.59 ± 97.77	11557.64 ± 90.13	18030.82 ± 194.30	77.11 ± 0.60	793163.76 ± 191.91	82090.05 ± 112.73	247324.32 ± 578.04	471964.71 ± 1466.73	1712195
W+jets, 1ID and 1AntiID	18717.67 ± 45.30	4675.03 ± 85.55	4955.23 ± 121.81	15.48 ± 0.27	197733.06 ± 95.26	19244.82 ± 54.19	104209.24 ± 342.84	283429.48 ± 886.98	632980
FakeEl	3808.72 ± 19.93	586.54 ± 31.48	1589.28 ± 58.06	2.71 ± 0.12	34334.43 ± 39.78	3195.83 ± 22.13	11119.55 ± 160.39	14731.94 ± 319.26	69369
$M_{\ell\ell} > 10$ GeV	3805.90 ± 19.92	583.41 ± 31.48	1581.75 ± 58.01	2.71 ± 0.12	34310.16 ± 39.77	3193.76 ± 22.13	11111.98 ± 160.36	14718.33 ± 319.14	69308
>=2jets	582.17 ± 4.72	151.71 ± 3.00	366.66 ± 21.51	0.92 ± 0.08	28181.04 ± 35.99	1527.79 ± 15.25	1391.28 ± 47.93	5721.42 ± 205.53	37923
nBjets=0	499.09 ± 4.47	123.05 ± 2.60	308.37 ± 19.96	0.72 ± 0.06	2441.19 ± 10.73	271.16 ± 6.47	1139.50 ± 45.46	2015.93 ± 97.20	6799
$M_{\ell\ell} < 70$ GeV	326.97 ± 3.71	74.93 ± 2.08	171.21 ± 15.08	0.46 ± 0.05	1625.61 ± 8.74	179.78 ± 5.26	420.15 ± 28.43	1129.89 ± 71.32	3929
$M_T < 300$ GeV	308.22 ± 3.51	68.64 ± 1.89	158.92 ± 14.35	0.44 ± 0.05	1557.70 ± 8.56	172.74 ± 5.16	294.72 ± 20.46	1020.63 ± 65.74	3582

Table B.3: The CR cutflow in the μe channel for one id + one anti-id electron.

$\sqrt{s} = 13TeV$, (Full Run 2)	WW	Other VV	$V\gamma$	Triboson	$t\bar{t}$	Single Top	Z/γ^*	W+jets DD	data
Channel Selection	102041.60 ± 106.23	50598.36 ± 150.19	364557.69 ± 1158.53	240.35 ± 0.98	933481.24 ± 207.51	93241.88 ± 119.73	1156960.06 ± 3176.96	6300829.82 ± 4530.84	9001951
$p_t^{lead} > 27$ GeV	97900.16 ± 104.09	46388.83 ± 136.14	305471.30 ± 1108.87	235.47 ± 0.97	902864.79 ± 204.23	90862.63 ± 118.26	972089.16 ± 2990.93	4576364.66 ± 4153.61	6992177
$p_t^{sublead} > 15$ GeV	88984.29 ± 99.43	38038.74 ± 132.49	175748.94 ± 926.18	217.51 ± 0.94	803600.69 ± 192.95	82020.08 ± 112.46	496598.95 ± 1847.81	1365485.81 ± 2720.34	3050695
GRL Selection	88984.29 ± 99.43	38038.74 ± 132.49	175748.94 ± 926.18	217.51 ± 0.94	803600.69 ± 192.95	82020.08 ± 112.46	496598.95 ± 1847.81	1365485.81 ± 2720.34	3050695
Jet Cleaning	88984.29 ± 99.43	38038.74 ± 132.49	175748.94 ± 926.18	217.51 ± 0.94	803600.69 ± 192.95	82020.08 ± 112.46	496598.95 ± 1847.81	1365485.81 ± 2720.34	3050695
Overlap: Vgamma/Vjets	88984.29 ± 99.43	38038.74 ± 132.49	175748.94 ± 926.18	217.51 ± 0.94	803600.69 ± 192.95	82020.08 ± 112.46	416737.75 ± 1521.57	1445347.01 ± 2510.18	3050695
Trigger Selection	85966.75 ± 97.87	36056.73 ± 131.29	161205.98 ± 905.29	213.30 ± 0.93	769760.51 ± 188.98	79047.88 ± 110.45	390759.69 ± 1469.20	963351.15 ± 2353.55	2486362
Trigger Matching	84368.14 ± 96.68	34012.92 ± 130.46	152473.11 ± 882.55	204.23 ± 0.91	752399.18 ± 186.12	77428.97 ± 108.76	363394.72 ± 1387.77	969876.72 ± 2282.94	2434158
Only two Leptons	84176.35 ± 96.61	17460.62 ± 126.49	124801.33 ± 637.09	101.58 ± 0.68	741185.97 ± 184.75	76457.92 ± 108.21	313466.06 ± 1168.14	972998.17 ± 2042.61	2330648
OS Leptons	82587.58 ± 96.03	9199.96 ± 66.63	64799.30 ± 491.67	72.03 ± 0.58	726424.08 ± 182.91	75167.74 ± 107.33	274999.13 ± 966.90	599879.17 ± 1751.69	1833129
3rd lepton veto + OS	82587.58 ± 96.03	9199.96 ± 66.63	64799.30 ± 491.67	72.03 ± 0.58	726424.08 ± 182.91	75167.74 ± 107.33	274999.13 ± 966.90	599879.17 ± 1751.69	1833129
W+jets, 1ID and 1AntiID	17113.44 ± 42.84	2998.73 ± 48.63	12705.44 ± 174.88	13.95 ± 0.26	169834.34 ± 87.60	16537.89 ± 49.76	95156.79 ± 512.88	144439.42 ± 875.66	458800
FakeEl	11943.40 ± 36.22	2086.94 ± 46.56	11773.33 ± 170.37	9.70 ± 0.21	121122.83 ± 73.76	11613.46 ± 41.54	59047.41 ± 489.76	115394.93 ± 782.64	332992
$M_{\ell\ell} > 10$ GeV	11837.37 ± 36.07	2001.08 ± 16.05	11251.16 ± 167.85	9.62 ± 0.21	120175.82 ± 73.47	11523.27 ± 41.39	58877.79 ± 489.45	114535.88 ± 778.86	330212
>=2jets	2205.58 ± 9.04	521.14 ± 6.29	1697.04 ± 52.98	3.55 ± 0.12	98090.38 ± 66.28	5679.67 ± 28.97	6072.12 ± 100.54	31743.52 ± 405.32	146013
nBjets=0	1890.12 ± 8.55	422.22 ± 5.52	1413.34 ± 49.04	2.72 ± 0.11	8704.52 ± 20.01	957.55 ± 12.01	5074.62 ± 96.75	12634.91 ± 208.60	31100
$M_{\mathrm{tt}} < 70$ GeV	1287.22 ± 7.30	261.42 ± 4.01	971.61 ± 41.53	1.89 ± 0.09	6075.53 ± 16.69	673.49 ± 10.04	2101.37 ± 72.08	8021.46 ± 163.59	19394
$M_T < 300$ GeV	1220.86 ± 7.10	236.82 ± 3.82	851.34 ± 37.89	1.83 ± 0.09	5824.02 ± 16.34	646.83 ± 9.84	1412.95 ± 56.05	7201.35 ± 149.67	17396

Table B.4: The CR cutflow in the $e\mu$ channel for one id + one anti-id muon.

$\sqrt{s} = 13TeV$, $\mathcal{L} = 139fb^{-1}$ (Full Run 2)	WW	Other VV	$V\gamma$	Triboson	$t\bar{t}$	Single Top	Z/γ^*	W+jets DD	data
Channel Selection	109135.98 ± 108.45	62393.97 ± 196.40	114061.69 ± 714.45	251.67 ± 1.00	1033047.84 ± 218.56	103191.27 ± 126.04	739668.03 ± 1547.06	3587669.55 ± 2960.95	5749420
$p_t^{lead} > 27$ GeV	104950.45 ± 106.14	56809.01 ± 191.22	83528.87 ± 664.06	246.41 ± 0.99	1001704.19 ± 215.32	100694.70 ± 124.56	579111.69 ± 1385.50	2694195.68 ± 2662.99	4621241
$p_t^{sublead} > 15$ GeV	95825.72 ± 101.71	44182.34 ± 183.49	66927.69 ± 631.02	227.40 ± 0.95	889939.45 ± 203.25	90251.22 ± 118.08	399796.83 ± 1093.12	1561067.35 ± 2200.14	3148218
GRL Selection	95825.72 ± 101.71	44182.34 ± 183.49	66927.69 ± 631.02	227.40 ± 0.95	889939.45 ± 203.25	90251.22 ± 118.08	399796.83 ± 1093.12	1561067.35 ± 2200.14	3148218
Jet Cleaning	95825.72 ± 101.71	44182.34 ± 183.49	66927.69 ± 631.02	227.40 ± 0.95	889939.45 ± 203.25	90251.22 ± 118.08	399796.83 ± 1093.12	1561067.35 ± 2200.14	3148218
Overlap: Vgamma/Vjets	95825.72 ± 101.71	44182.34 ± 183.49	66927.69 ± 631.02	227.40 ± 0.95	889939.45 ± 203.25	90251.22 ± 118.08	372283.68 ± 936.10	1588580.49 ± 2126.49	3148218
Trigger Selection	91056.65 ± 99.22	40768.45 ± 169.70	44835.22 ± 595.68	222.71 ± 0.94	846045.66 ± 198.35	86565.93 ± 115.71	314328.23 ± 854.76	756239.16 ± 1832.18	2180062
Trigger Matching	89775.13 ± 98.40	38851.71 ± 132.21	36259.75 ± 570.49	216.66 ± 0.93	833491.53 ± 196.71	85497.94 ± 114.88	290844.20 ± 768.14	722989.08 ± 1758.55	2097926
Only two Leptons	89620.77 ± 98.37	21689.88 ± 127.88	31048.21 ± 256.70	109.55 ± 0.70	815028.58 ± 194.53	83972.26 ± 113.97	263091.51 ± 655.56	723065.23 ± 1612.49	2027626
OS Leptons	87986.59 ± 97.77	11557.64 ± 90.13	18030.82 ± 194.30	77.11 ± 0.60	793163.76 ± 191.91	82090.05 ± 112.73	247324.32 ± 578.04	471964.71 ± 1466.73	1712195
3rd lepton veto + OS	87986.59 ± 97.77	11557.64 ± 90.13	18030.82 ± 194.30	77.11 ± 0.60	793163.76 ± 191.91	82090.05 ± 112.73	247324.32 ± 578.04	471964.71 ± 1466.73	1712195
W+jets, 1ID and 1AntiID	18717.67 ± 45.30	4675.03 ± 85.55	4955.23 ± 121.81	15.48 ± 0.27	197733.06 ± 95.26	19244.82 ± 54.19	104209.24 ± 342.84	283429.48 ± 886.98	632980
FakeMu	14908.95 ± 40.68	4088.48 ± 79.55	3365.95 ± 107.08	12.76 ± 0.24	163398.62 ± 86.56	16049.00 ± 49.46	93089.69 ± 303.01	268697.54 ± 827.54	563611
$M_{\ell\ell} > 10$ GeV	14653.19 ± 40.39	3711.52 ± 40.24	3225.93 ± 105.80	12.52 ± 0.24	160795.14 ± 85.87	15790.32 ± 49.07	92745.08 ± 299.35	253272.31 ± 811.20	544206
>=2jets	2775.51 ± 9.86	978.79 ± 9.67	753.46 ± 29.67	4.82 ± 0.15	130049.08 ± 77.10	7785.29 ± 34.35	11432.51 ± 85.71	93271.54 ± 512.44	247051
nBjets=0	2370.16 ± 9.30	781.56 ± 8.97	593.26 ± 25.56	3.86 ± 0.13	12441.07 ± 24.13	1357.52 ± 14.48	9253.60 ± 75.60	33721.96 ± 260.48	60523
$M_{t\ell} < 70$ GeV	1549.77 ± 7.75	488.63 ± 5.97	310.38 ± 18.34	2.46 ± 0.11	8479.85 ± 19.91	926.88 ± 11.94	3555.89 ± 47.12	20816.15 ± 198.30	36130
$M_T < 300$ GeV	1474.26 ± 7.52	443.15 ± 5.66	262.54 ± 17.06	2.39 ± 0.10	8121.17 ± 19.48	890.53 ± 11.71	2406.08 ± 39.07	18617.88 ± 186.12	32218

Table B.5: The CR cutflow in the μe channel for one id + one anti-id muon.

$\sqrt{s} = 13TeV$, $\mathcal{L} = 139fb^{-1}$ (Full Run 2)	WW	Other VV	$V\gamma$	Triboson	$t\bar{t}$	Single Top	Z/γ^*	W+jets DD	data
Channel Selection	102041.60 ± 106.23	50598.36 ± 150.19	364557.69 ± 1158.53	240.35 ± 0.98	933481.24 ± 207.51	93241.88 ± 119.73	1156960.06 ± 3176.96	6300829.82 ± 4530.84	9001951
$p_t^{lead} > 27$ GeV	97900.16 ± 104.09	46388.83 ± 136.14	305471.30 ± 1108.87	235.47 ± 0.97	902864.79 ± 204.23	90862.63 ± 118.26	972089.16 ± 2990.93	457636.66 ± 4153.61	6992177
$p_t^{sublead} > 15$ GeV	88984.29 ± 99.43	38038.74 ± 132.49	175748.94 ± 926.18	217.51 ± 0.94	803600.69 ± 192.95	82020.08 ± 112.46	496598.95 ± 1847.81	1365485.81 ± 2720.34	3050695
GRL Selection	88984.29 ± 99.43	38038.74 ± 132.49	175748.94 ± 926.18	217.51 ± 0.94	803600.69 ± 192.95	82020.08 ± 112.46	496598.95 ± 1847.81	1365485.81 ± 2720.34	3050695
Jet Cleaning	88984.29 ± 99.43	38038.74 ± 132.49	175748.94 ± 926.18	217.51 ± 0.94	803600.69 ± 192.95	82020.08 ± 112.46	496598.95 ± 1847.81	1365485.81 ± 2720.34	3050695
Overlap: Vgamma/Vjets	88984.29 ± 99.43	38038.74 ± 132.49	175748.94 ± 926.18	217.51 ± 0.94	803600.69 ± 192.95	82020.08 ± 112.46	416737.75 ± 1521.57	1445347.01 ± 2510.18	3050695
Trigger Selection	85966.75 ± 97.87	36056.73 ± 131.29	161205.98 ± 905.29	213.30 ± 0.93	769760.51 ± 188.98	79047.88 ± 110.45	390759.69 ± 1469.20	963351.15 ± 2353.55	2486362
Trigger Matching	84368.14 ± 96.68	34012.92 ± 130.46	152473.11 ± 882.55	204.23 ± 0.91	752399.18 ± 186.12	77428.97 ± 108.76	363394.72 ± 1387.77	969876.72 ± 2282.94	2434158
Only two Leptons	84176.35 ± 96.61	17460.62 ± 126.49	124801.33 ± 637.09	101.58 ± 0.68	741185.97 ± 184.75	76457.92 ± 108.21	313466.06 ± 1168.14	972998.17 ± 2042.61	2330648
OS Leptons	82587.58 ± 96.03	9199.96 ± 66.63	64799.30 ± 491.67	72.03 ± 0.58	726424.08 ± 182.91	75167.74 ± 107.33	274999.13 ± 966.90	599879.17 ± 1751.69	1833129
3rd lepton veto + OS	82587.58 ± 96.03	9199.96 ± 66.63	64799.30 ± 491.67	72.03 ± 0.58	726424.08 ± 182.91	75167.74 ± 107.33	274999.13 ± 966.90	599879.17 ± 1751.69	1833129
W+jets, 1ID and 1AntiID	17113.44 ± 42.84	2998.73 ± 48.63	12705.44 ± 174.88	13.95 ± 0.26	169834.34 ± 87.60	16537.89 ± 49.76	95156.79 ± 512.88	144439.42 ± 875.66	458800
FakeMu	5170.04 ± 22.88	911.79 ± 14.02	932.11 ± 39.47	4.25 ± 0.14	48711.51 ± 47.27	4924.43 ± 27.40	36109.38 ± 152.25	29044.49 ± 392.75	125808
$M_{\ell\ell} > 10$ GeV	5124.43 ± 22.79	866.43 ± 12.17	898.68 ± 38.63	4.21 ± 0.14	48297.94 ± 47.07	4883.28 ± 27.28	35936.48 ± 149.93	26612.55 ± 387.59	122624
>=2jets	872.09 ± 5.57	282.16 ± 4.79	209.83 ± 16.05	1.53 ± 0.08	39218.39 ± 42.35	2410.04 ± 19.12	4449.02 ± 44.49	13545.95 ± 255.81	60989
nBjets=0	747.45 ± 5.27	226.34 ± 4.41	166.45 ± 14.30	1.23 ± 0.08	3605.29 ± 13.00	430.12 ± 8.12	3671.30 ± 41.45	4823.83 ± 126.00	13672
$M_{t\ell} < 70$ GeV	463.81 ± 4.27	129.76 ± 3.53	75.66 ± 9.01	0.79 ± 0.06	2362.45 ± 10.51	276.96 ± 6.49	1304.99 ± 23.73	3117.58 ± 92.52	7732
$M_T < 300$ GeV	438.07 ± 4.13	116.87 ± 3.38	59.59 ± 8.12	0.76 ± 0.06	2256.26 ± 10.27	264.67 ± 6.34	840.53 ± 17.98	2753.25 ± 85.40	6730

B.3 Closure Test

The closure test is performed at the end of selection for CR of fake electron and muon. Table B.6 and Table B.14 shows the CR of one id + one anti-id electron and muon cutflow for closure test in the $e\mu + \mu e$ channel. The upper part of the cutflow shows the common selection, followed by the CR baseline selection. Only the W+jets MC are considered.

Table B.6: The CR of one id + one anti-id electron cutflow for closure test in the $e\mu + \mu e$ channel. The upper part of the cutflow shows the common selection, followed by the CR baseline selection.

$\sqrt{s} = 13$TeV, (Full Run 2)	Higgs	WW	Other VV	$V\gamma$	Triboson	$t\bar{t}$	Single Top	Z/γ^*	W+jets MC	AllBkg	Fakes DD	data
Channel Selection	16043.40 ± 12.56	2105163.72 ± 149.30	104007.47 ± 244.63	313121.58 ± 855.55	483.65 ± 1.39	1839525.52 ± 363.92	186873.78 ± 169.54	1821464.44 ± 3443.21	2244589.23 ± 4795.10	6722612.78 ± 5985.33	6441785.45 ± 4873.49	10919809
$p_T^{lead} > 27$ GeV	14402.15 ± 11.88	197395.82 ± 146.32	9539.98 ± 233.20	250659.10 ± 789.96	473.71 ± 1.38	1774602.69 ± 358.50	182356.12 ± 167.55	1480713.45 ± 3206.05	1963724.62 ± 4480.27	5959687.63 ± 5586.43	4733846.99 ± 4457.03	8729810
$p_T^{sublead} > 15$ GeV	11082.35 ± 10.39	179885.33 ± 139.99	7622091 ± 223.66	140695.85 ± 590.28	437.01 ± 1.32	1580605.34 ± 338.72	164252.94 ± 159.22	856469.48 ± 2093.48	656920.99 ± 2598.84	3666570.21 ± 3419.78	1769727.78 ± 3117.74	4779377
GRL Selection	11082.35 ± 10.39	179885.33 ± 139.99	7622091 ± 223.66	140695.85 ± 590.28	437.01 ± 1.32	1580605.34 ± 338.72	164252.94 ± 159.22	856469.48 ± 2093.48	656920.99 ± 2598.84	3666570.21 ± 3419.78	1769727.78 ± 3117.74	4779377
Jet Cleaning	11082.35 ± 10.39	179885.33 ± 139.99	7622091 ± 223.66	140695.85 ± 590.28	437.01 ± 1.32	1580605.34 ± 338.72	164252.94 ± 159.22	856469.48 ± 2093.48	656920.99 ± 2598.84	3666570.21 ± 3419.78	1769727.78 ± 3117.74	4779377
Overlap, Vgamma/Vjets	11082.35 ± 10.39	179885.33 ± 139.99	7622091 ± 223.66	140695.85 ± 590.28	437.01 ± 1.32	1580605.34 ± 338.72	164252.94 ± 159.22	746304.29 ± 1700.51	656920.99 ± 2598.84	3556405.02 ± 3194.36	1879892.97 ± 2868.68	4779377
Trigger Selection	10075.09 ± 9.89	172261.25 ± 137.18	7161105 ± 211.55	118804.64 ± 556.05	428.42 ± 1.31	1512209.58 ± 331.38	158154.21 ± 156.30	666256.77 ± 1617.15	590765.78 ± 2464.79	3200666.78 ± 3032.74	1146244.00 ± 2641.67	3856145
Trigger Matching	9763.40 ± 9.67	169542.04 ± 135.77	6790171 ± 194.04	107247.26 ± 517.62	413.56 ± 1.29	1486084.54 ± 327.02	155708.47 ± 154.62	618004.01 ± 1502.85	566740.18 ± 2392.88	3181405.16 ± 2905.06	1128355.02 ± 2540.97	3743020
Only two Leptons	9520.95 ± 9.65	169251.59 ± 135.70	3656699 ± 188.45	90451.09 ± 468.01	214.47 ± 0.99	1472339.31 ± 325.52	154344.63 ± 154.11	556553.31 ± 1280.54	566554.99 ± 2392.49	3055797.33 ± 2786.85	1123786.66 ± 2378.14	3613029
OS Leptons	9240.35 ± 9.59	166114.08 ± 134.88	1914193 ± 73.59	48704.45 ± 351.09	151.00 ± 0.84	1446776.26 ± 322.69	151924.74 ± 152.95	507594.62 ± 1068.03	366881.75 ± 1924.09	2716529.19 ± 2262.14	694480.56 ± 2111.70	3044128
3rd lepton veto + OS	9240.35 ± 9.59	166114.08 ± 134.88	1914193 ± 73.59	48704.45 ± 351.09	151.00 ± 0.84	1446776.26 ± 322.69	151924.74 ± 152.95	507594.62 ± 1068.03	366881.75 ± 1924.09	2716529.19 ± 2262.14	694480.56 ± 2111.70	3044128
W+jets, 1ID and 1AntiID	2799.07 ± 5.27	34295.68 ± 61.04	6325.14 ± 41.77	10297.25 ± 137.04	29.29 ± 0.37	331789.34 ± 152.99	33181.31 ± 70.88	192333.07 ± 596.76	132652.57 ± 1134.27	743702.74 ± 1302.07	258077.84 ± 1130.47	869128
FakeEl	1318.41 ± 3.59	15502.06 ± 40.94	2477.99 ± 22.72	801.782 ± 123.52	12.78 ± 0.25	151106.21 ± 102.66	14540.80 ± 46.70	70245.32 ± 511.08	89547.64 ± 950.46	352829.04 ± 1093.05	125854.60 ± 824.94	389136
$M_{ll} > 10$ GeV	1269.93 ± 3.52	15395.51 ± 40.81	2437.20 ± 22.35	769.181 ± 121.55	12.69 ± 0.25	150234.55 ± 102.36	14450.49 ± 46.56	70072.33 ± 510.77	89086.30 ± 947.98	350650.81 ± 1090.48	124850.49 ± 822.74	386415
>=2jets	263.28 ± 1.54	2762.36 ± 10.25	679.47 ± 13.65	1238.13 ± 43.75	4.67 ± 0.15	123811.82 ± 92.75	7165.32 ± 32.70	7962.71 ± 120.09	5320.02 ± 232.08	149207.78 ± 283.13	34279.24 ± 452.18	178167
nBjets=0	191.57 ± 1.36	2258.25 ± 9.44	525.36 ± 13.38	984.52 ± 40.42	3.47 ± 0.13	7223.01 ± 22.94	902.02 ± 11.81	6287.09 ± 111.98	4423.32 ± 211.09	22798.61 ± 244.28	12264.71 ± 213.89	30640
$M_{tt} < 70$ GeV	165.31 ± 1.27	1523.70 ± 7.99	323.33 ± 12.65	649.07 ± 34.69	2.33 ± 0.10	4985.79 ± 19.02	626.28 ± 9.81	2588.53 ± 80.91	2574.15 ± 159.36	13433.49 ± 183.93	7614.66 ± 164.04	18474
$M_T < 300$ GeV	154.72 ± 1.23	1442.27 ± 7.75	295.04 ± 12.54	566.95 ± 32.36	2.24 ± 0.10	4774.67 ± 18.62	600.57 ± 9.61	1692.68 ± 62.14	2352.28 ± 151.95	11880.78 ± 169.27	6946.50 ± 148.46	16475
FakeEl, $15 < p_T^{e-fake} < 20$ GeV	46.77 ± 0.66	335.78 ± 3.40	87.63 ± 11.80	204.38 ± 19.31	0.50 ± 0.05	1087.82 ± 8.58	131.60 ± 4.33	670.02 ± 42.30	1068.51 ± 99.73	3633.00 ± 111.14	3255.51 ± 90.69	5820
FakeEl, $20 < p_T^{e-fake} < 25$ GeV	39.07 ± 0.62	325.26 ± 3.75	61.64 ± 2.22	112.37 ± 11.73	0.41 ± 0.04	1034.32 ± 8.59	134.17 ± 4.53	408.47 ± 32.76	433.52 ± 66.30	2549.23 ± 75.63	1444.30 ± 69.89	3560
FakeEl, $25 < p_T^{e-fake} < 35$ GeV	27.42 ± 0.52	274.12 ± 3.86	52.10 ± 2.22	110.34 ± 20.31	0.49 ± 0.05	903.83 ± 8.19	115.24 ± 4.26	282.59 ± 22.22	361.11 ± 59.69	2127.24 ± 67.63	1027.87 ± 61.69	2794
FakeEl, $p_T^{e-fake} > 35$ GeV	41.47 ± 0.65	507.11 ± 4.43	93.68 ± 2.85	139.86 ± 11.13	0.84 ± 0.06	1748.70 ± 11.50	219.57 ± 5.90	330.95 ± 22.47	489.13 ± 71.99	3571.31 ± 77.51	1218.82 ± 71.59	4301

Table B.7: The CR of one id + one anti-id muon cutflow for closure test in the $e\mu + \mu e$ channel. The upper part of the cutflow shows the common selection, followed by the CR baseline selection.

$\sqrt{s}$ = 13TeV, (Full Run 2)	Higgs	WW	Other VV	$V\gamma$	Triboson	$t\bar{t}$	Single Top	Z/γ^*	W+jets MC	All Bkg	Fakes DD	data
Channel Selection	16043.40 ± 12.56	205503.72 ± 149.30	104007.47 ± 244.63	313121.58 ± 855.55	483.65 ± 1.39	1830525.52 ± 363.92	186873.78 ± 169.54	1821464.44 ± 3443.21	2244589.23 ± 4795.10	6722612.78 ± 5985.33	6441785.45 ± 4873.49	10919809
$p_T^{lead} > 27$ GeV	14402.15 ± 11.88	197395.82 ± 146.32	95359.98 ± 233.20	250659.10 ± 789.96	473.71 ± 1.38	1774602.69 ± 358.50	182356.12 ± 167.55	1480713.45 ± 3206.05	1963724.62 ± 4480.27	5959687.63 ± 5586.43	4733846.99 ± 4457.03	8729810
$p_T^{sublead} > 15$ GeV	11082.35 ± 10.39	179885.33 ± 139.99	76220.91 ± 223.66	140695.85 ± 590.28	437.01 ± 1.32	1580605.34 ± 338.72	164252.94 ± 159.22	856469.48 ± 2003.48	656920.99 ± 2508.84	3666570.21 ± 3419.78	1769727.78 ± 3117.74	4779377
GRL Selection	11082.35 ± 10.39	179885.33 ± 139.99	76220.91 ± 223.66	140695.85 ± 590.28	437.01 ± 1.32	1580605.34 ± 338.72	164252.94 ± 159.22	856469.48 ± 2003.48	656920.99 ± 2508.84	3666570.21 ± 3419.78	1769727.78 ± 3117.74	4779377
Jet Cleaning	11082.35 ± 10.39	179885.33 ± 139.99	76220.91 ± 223.66	140695.85 ± 590.28	437.01 ± 1.32	1580605.34 ± 338.72	164252.94 ± 159.22	856469.48 ± 2003.48	656920.99 ± 2508.84	3666570.21 ± 3419.78	1769727.78 ± 3117.74	4779377
Overlap: Vgamma Vjets	11082.35 ± 10.39	179885.33 ± 139.99	76220.91 ± 223.66	140695.85 ± 590.28	437.01 ± 1.32	1580605.34 ± 338.72	164252.94 ± 159.22	746304.29 ± 1700.51	656920.99 ± 2508.84	3556405.02 ± 3194.36	1879892.97 ± 2868.68	4779377
Trigger Selection	10075.09 ± 9.89	172361.25 ± 137.18	71611.05 ± 211.55	118804.64 ± 556.05	428.42 ± 1.31	1512209.58 ± 331.38	158154.21 ± 156.30	660256.77 ± 1617.15	590765.78 ± 2464.79	3300666.78 ± 3032.74	1146244.00 ± 2641.67	3856145
Trigger Matching	9763.40 ± 9.67	169542.04 ± 135.77	67901.71 ± 194.04	107247.26 ± 517.62	413.56 ± 1.29	1486084.54 ± 327.02	155708.47 ± 154.62	618004.01 ± 1502.85	566740.18 ± 2392.88	3181405.16 ± 2905.06	1128355.02 ± 2540.97	3743020
Only two Leptons	9520.95 ± 9.65	169251.59 ± 135.70	36566.99 ± 188.45	90451.09 ± 468.01	214.47 ± 0.99	1472339.31 ± 325.52	154344.63 ± 154.11	536553.31 ± 1280.54	566554.99 ± 2392.49	3055797.33 ± 2786.85	1123786.66 ± 2378.14	3613029
OS Leptons	9240.35 ± 9.59	166114.08 ± 134.88	19141.93 ± 73.59	48704.45 ± 351.09	151.00 ± 0.84	1446776.26 ± 322.69	151924.74 ± 152.95	507594.62 ± 1068.03	366881.75 ± 1924.09	2716529.19 ± 2262.14	694480.56 ± 2111.70	3044128
3rd lepton veto + OS	9240.35 ± 9.59	166114.08 ± 134.88	19141.93 ± 73.59	48704.45 ± 351.09	151.00 ± 0.84	1446776.26 ± 322.69	151924.74 ± 152.95	507594.62 ± 1068.03	366881.75 ± 1924.09	2716529.19 ± 2262.14	694480.56 ± 2111.70	3044128
W+jets, 1ID and 1AntiID	2799.07 ± 5.27	34295.68 ± 61.04	6325.14 ± 41.77	10297.25 ± 137.04	29.29 ± 0.37	331789.34 ± 152.99	33181.31 ± 70.88	192333.07 ± 596.76	132652.57 ± 1134.27	743702.74 ± 1302.07	258077.84 ± 1130.47	869128
FakeMu	1480.66 ± 3.86	18793.62 ± 45.27	3847.15 ± 35.05	2279.43 ± 59.36	16.52 ± 0.27	180623.12 ± 113.42	18640.51 ± 53.32	122087.75 ± 308.09	43104.94 ± 619.02	390873.69 ± 707.56	132223.24 ± 772.94	479992
$M_{ll} > 10$ GeV	1354.86 ± 3.69	18494.07 ± 44.98	3580.65 ± 33.09	2181.74 ± 57.98	16.22 ± 0.27	177562.41 ± 112.46	18337.44 ± 52.89	121548.38 ± 303.29	41617.59 ± 607.00	384693.37 ± 694.53	117668.23 ± 758.07	460744
>=2jets	327.34 ± 1.75	3392.89 ± 11.09	1011.15 ± 10.23	434.37 ± 22.24	6.32 ± 0.17	146764.57 ± 102.06	9181.03 ± 37.34	14350.67 ± 89.90	2596.91 ± 153.77	178065.25 ± 210.39	39419.66 ± 485.29	214888
nBjets=0	238.67 ± 1.55	2775.57 ± 10.22	766.89 ± 9.00	329.63 ± 18.99	4.80 ± 0.15	8539.83 ± 25.20	1131.11 ± 13.34	11225.19 ± 80.70	2020.46 ± 136.43	27032.16 ± 162.75	11444.30 ± 210.54	36456
$M_{ll} < 70$ GeV	209.90 ± 1.45	1789.97 ± 8.44	461.22 ± 7.57	161.77 ± 12.77	3.06 ± 0.12	5681.03 ± 20.51	751.07 ± 10.85	4095.24 ± 50.40	1172.08 ± 101.55	14325.34 ± 116.98	6978.74 ± 153.31	20132
$M_T < 300$ GeV	196.50 ± 1.41	1700.56 ± 8.19	418.68 ± 7.31	124.65 ± 11.56	2.96 ± 0.12	5440.72 ± 20.08	720.60 ± 10.62	2638.69 ± 41.51	1048.48 ± 95.06	12291.84 ± 107.39	6135.64 ± 140.97	17379
FakeMu, $15 < p_T^{m-fake} < 20$ GeV	46.59 ± 0.68	349.25 ± 4.03	101.03 ± 3.48	37.09 ± 5.80	0.62 ± 0.05	1217.74 ± 9.41	148.74 ± 4.80	870.34 ± 27.78	639.49 ± 73.51	3410.90 ± 79.69	3214.59 ± 83.26	5986
FakeMu, $20 < p_T^{m-fake} < 25$ GeV	33.50 ± 0.58	282.54 ± 3.21	75.94 ± 3.15	28.40 ± 6.29	0.40 ± 0.04	939.40 ± 8.35	117.91 ± 4.27	473.76 ± 16.01	227.61 ± 45.97	2179.48 ± 50.18	1382.13 ± 61.14	3334
FakeMu, $p_T^{m-fake} > 25$ GeV	116.41 ± 1.09	1068.76 ± 6.37	241.71 ± 5.60	59.16 ± 7.78	1.93 ± 0.09	3283.58 ± 15.65	453.94 ± 8.46	1294.59 ± 26.36	181.38 ± 38.99	6701.47 ± 51.62	1538.92 ± 95.94	8059

B.4 Flavor Composition

B.4.1 Z+jets Sample

Table B.8: The Z+jets flavour composition cutflow in the $e\mu + \mu e$ channel showing all the fraction of the fake flavor in each column. The upper part of the cut-flow shows the common selection together with the electron and muon fake selection, followed by the id and anti-id selections and the split in p_{T} binning.

$\sqrt{s} = 13TeV$, (Full Run 2)	Zjets Powheg	bottom	charm	strange	light	other
Channel Selection	89201.355 ± 166.566	178402.710 ± 235.560	178402.710 ± 235.560	178402.710 ± 235.560	178402.710 ± 235.560	178402.710 ± 235.560
lll final state	89201.355 ± 166.566	178402.710 ± 235.560	178402.710 ± 235.560	178402.710 ± 235.560	178402.710 ± 235.560	178402.710 ± 235.560
Vgamma/Vjets Overlap	66368.871 ± 144.675	8655.348 ± 58.164	5981.036 ± 43.721	8294.694 ± 49.871	21252.043 ± 79.970	22185.750 ± 82.183
ZCand $p_T > 27$ GeV	45410.129 ± 119.631	6472.013 ± 50.085	4455.914 ± 37.339	6741.187 ± 44.874	17491.822 ± 72.185	10249.193 ± 56.433
Z-tagging	25913.208 ± 89.707	4733.934 ± 42.529	3209.792 ± 31.489	4902.585 ± 37.886	12577.661 ± 60.546	489.236 ± 12.083
fake type: electron	19491.656 ± 76.243	1192.836 ± 21.223	1498.109 ± 21.477	4318.882 ± 35.568	12120.948 ± 59.451	360.882 ± 10.090
fake el η cut	15881.334 ± 68.865	1117.309 ± 20.518	1287.877 ± 19.900	3272.165 ± 30.915	9865.015 ± 53.611	338.967 ± 9.770
Truth-matched Z-candidate leptons	15707.370 ± 68.540	1109.255 ± 20.452	1279.728 ± 19.845	3249.680 ± 30.824	9777.367 ± 53.407	291.341 ± 9.124
ID cuts	502.529 ± 12.559	115.424 ± 6.646	48.228 ± 3.861	31.800 ± 2.975	262.554 ± 8.718	44.524 ± 3.714
$15 < p_T^{fake} < 20$ GeV	179.188 ± 7.635	74.418 ± 5.248	20.702 ± 2.566	4.620 ± 1.109	61.785 ± 4.210	17.664 ± 2.284
$20 < p_T^{fake} < 25$ GeV	69.513 ± 4.855	20.261 ± 2.978	7.936 ± 1.602	3.043 ± 0.894	31.518 ± 3.038	6.755 ± 1.450
$p_T^{fake} < 25$ GeV	248.701 ± 9.048	94.678 ± 6.034	28.637 ± 3.025	7.663 ± 1.424	93.303 ± 5.192	24.419 ± 2.706
$p_T^{fake} > 25$ GeV	253.828 ± 8.710	20.746 ± 2.787	19.590 ± 2.400	24.136 ± 2.612	169.251 ± 7.004	20.105 ± 2.543
$25 < p_T^{fake} < 35$ GeV	103.162 ± 5.662	13.718 ± 2.318	9.261 ± 1.677	8.332 ± 1.526	63.951 ± 4.318	7.901 ± 1.701
$p_T^{fake} > 35$ GeV	150.666 ± 6.619	7.028 ± 1.548	10.330 ± 1.717	15.804 ± 2.120	105.300 ± 5.514	12.204 ± 1.891
Anti-ID cuts	4809.408 ± 38.304	737.242 ± 16.571	566.460 ± 13.254	743.393 ± 14.699	2661.702 ± 27.788	100.611 ± 5.360
$15 < p_T^{fake} < 20$ GeV	2178.963 ± 25.823	397.399 ± 12.090	315.143 ± 9.912	294.822 ± 9.207	1106.465 ± 17.866	65.134 ± 4.294
$20 < p_T^{fake} < 25$ GeV	885.331 ± 16.420	155.018 ± 7.570	110.191 ± 5.723	117.239 ± 5.797	487.331 ± 11.891	15.552 ± 2.132
$p_T^{fake} < 25$ GeV	3064.294 ± 30.601	552.416 ± 14.264	425.334 ± 11.446	412.061 ± 10.880	1593.796 ± 21.461	80.686 ± 4.794
$p_T^{fake} > 25$ GeV	1745.114 ± 23.040	184.825 ± 8.434	141.126 ± 6.683	331.331 ± 9.883	1067.905 ± 17.652	19.926 ± 2.397
$25 < p_T^{fake} < 35$ GeV	864.189 ± 16.230	127.428 ± 6.913	89.338 ± 5.303	134.038 ± 6.271	501.525 ± 12.031	11.861 ± 1.848
$p_T^{fake} > 35$ GeV	880.924 ± 16.353	57.397 ± 4.831	51.789 ± 4.068	197.294 ± 7.639	566.380 ± 12.917	8.064 ± 1.526
Electron Fake-Factor	0.104 ± 0.003	0.157 ± 0.010	0.085 ± 0.007	0.043 ± 0.004	0.099 ± 0.003	0.443 ± 0.044
Electron Fake-Factor for $15 < p_T^{fake} < 20$ GeV	0.082 ± 0.004	0.187 ± 0.014	0.066 ± 0.008	0.016 ± 0.004	0.056 ± 0.004	0.271 ± 0.039
Electron Fake-Factor for $20 < p_T^{fake} < 25$ GeV	0.079 ± 0.006	0.131 ± 0.020	0.072 ± 0.015	0.026 ± 0.008	0.065 ± 0.006	0.434 ± 0.111
Electron Fake-Factor for $p_T^{fake} < 25$ GeV	0.081 ± 0.003	0.171 ± 0.012	0.067 ± 0.007	0.019 ± 0.003	0.059 ± 0.003	0.303 ± 0.038
Electron Fake-Factor for $p_T^{fake} > 25$ GeV	0.145 ± 0.005	0.112 ± 0.016	0.139 ± 0.018	0.073 ± 0.008	0.158 ± 0.007	1.009 ± 0.176
Electron Fake-Factor for $25 < p_T^{fake} < 35$ GeV	0.119 ± 0.007	0.108 ± 0.019	0.104 ± 0.020	0.062 ± 0.012	0.128 ± 0.009	0.666 ± 0.177
Electron Fake-Factor for $p_T^{fake} > 35$ GeV	0.171 ± 0.008	0.122 ± 0.029	0.199 ± 0.037	0.080 ± 0.011	0.186 ± 0.011	1.513 ± 0.370
fake type: muon	6421.552 ± 47.270	3541.098 ± 36.856	1711.683 ± 23.028	583.703 ± 13.046	456.714 ± 11.463	128.355 ± 6.648
fake mu η cut	6228.867 ± 46.523	3425.507 ± 36.232	1651.649 ± 22.601	577.461 ± 12.976	453.317 ± 11.422	120.933 ± 6.485
Truth-matched Z-candidate leptons	6184.147 ± 46.288	3405.141 ± 36.099	1641.767 ± 22.546	574.268 ± 12.947	449.000 ± 11.376	113.971 ± 5.836
ID cuts	257.344 ± 9.451	148.571 ± 7.502	66.905 ± 4.557	8.882 ± 1.594	16.757 ± 2.200	16.229 ± 2.210
$15 < p_T^{fake} < 20$ GeV	149.722 ± 7.237	99.104 ± 6.117	34.274 ± 3.188	1.483 ± 0.608	5.288 ± 1.202	9.574 ± 1.726
$20 < p_T^{fake} < 25$ GeV	45.848 ± 3.989	27.030 ± 3.145	11.600 ± 1.972	2.980 ± 0.953	2.792 ± 0.892	1.446 ± 0.652
$p_T^{fake} < 25$ GeV	195.570 ± 8.263	126.134 ± 6.878	45.874 ± 3.749	4.463 ± 1.131	8.080 ± 1.497	11.019 ± 1.845
$p_T^{fake} > 25$ GeV	61.774 ± 4.586	22.437 ± 2.995	21.030 ± 2.592	4.419 ± 1.124	8.677 ± 1.613	5.210 ± 1.216
Anti-ID cuts	5214.209 ± 42.859	3124.919 ± 34.640	1525.924 ± 21.742	317.438 ± 9.653	208.821 ± 7.757	37.106 ± 3.298
$15 < p_T^{fake} < 20$ GeV	2496.045 ± 29.383	1507.082 ± 23.735	733.179 ± 15.020	133.603 ± 6.248	99.480 ± 5.375	22.701 ± 2.553
$20 < p_T^{fake} < 25$ GeV	1059.713 ± 19.387	644.800 ± 15.826	306.788 ± 9.695	63.828 ± 4.322	39.264 ± 3.365	5.032 ± 1.178
$p_T^{fake} < 25$ GeV	3555.758 ± 35.202	2151.882 ± 28.527	1039.968 ± 17.877	197.431 ± 7.597	138.744 ± 6.342	27.733 ± 2.812
$p_T^{fake} > 25$ GeV	1658.451 ± 24.448	973.037 ± 19.651	485.957 ± 12.375	120.007 ± 5.955	70.077 ± 4.467	9.373 ± 1.723
Muon Fake-Factor	0.049 ± 0.002	0.048 ± 0.002	0.044 ± 0.003	0.028 ± 0.005	0.080 ± 0.011	0.437 ± 0.071
Muon Fake-Factor for $15 < p_T^{fake} < 20$ GeV	0.060 ± 0.003	0.066 ± 0.004	0.047 ± 0.004	0.011 ± 0.005	0.053 ± 0.012	0.422 ± 0.090
Muon Fake-Factor for $20 < p_T^{fake} < 25$ GeV	0.043 ± 0.004	0.042 ± 0.005	0.038 ± 0.007	0.047 ± 0.015	0.071 ± 0.024	0.287 ± 0.146
Muon Fake-Factor for $p_T^{fake} < 25$ GeV	0.055 ± 0.002	0.059 ± 0.003	0.044 ± 0.004	0.023 ± 0.006	0.058 ± 0.011	0.397 ± 0.078
Muon Fake-Factor for $p_T^{fake} > 25$ GeV	0.037 ± 0.003	0.023 ± 0.003	0.043 ± 0.005	0.037 ± 0.010	0.124 ± 0.024	0.556 ± 0.165

Table B.9: The Z+jets flavour composition percentage cutflow in the $e\mu + \mu e$ channel showing all the fraction of the fake flavor in each column. The upper part of the cutflow shows the common selection together with the electron and muon fake selection, followed by the id and anti-id selections and the split in p_{T} binning.

$\sqrt{s} = 13TeV$, (Full Run 2)	Zjets Powheg %	bottom %	charm %	strange %	light %	other %
Channel Selection	100.000 ± 0.264	200.000 ± 0.457	200.000 ± 0.457	200.000 ± 0.457	200.000 ± 0.457	200.000 ± 0.457
lll final state	100.000 ± 0.264	200.000 ± 0.457	200.000 ± 0.457	200.000 ± 0.457	200.000 ± 0.457	200.000 ± 0.457
Vgamma/Vjets Overlap	100.000 ± 0.308	13.041 ± 0.092	9.012 ± 0.069	12.498 ± 0.080	32.021 ± 0.139	33.428 ± 0.144
ZCand $p_T > 27$ GeV	100.000 ± 0.373	14.252 ± 0.117	9.813 ± 0.086	14.845 ± 0.106	38.520 ± 0.189	22.570 ± 0.138
Z-tagging	100.000 ± 0.490	18.268 ± 0.176	12.387 ± 0.129	18.919 ± 0.160	48.538 ± 0.288	1.888 ± 0.047
fake type: electron	100.000 ± 0.553	6.120 ± 0.111	7.686 ± 0.114	22.158 ± 0.202	62.185 ± 0.390	1.851 ± 0.052
fake el η cut	100.000 ± 0.613	7.035 ± 0.133	8.109 ± 0.130	20.604 ± 0.214	62.117 ± 0.432	2.134 ± 0.062
Truth-matched Z-candidate leptons	100.000 ± 0.617	7.062 ± 0.134	8.147 ± 0.131	20.689 ± 0.216	62.247 ± 0.435	1.855 ± 0.059
ID cuts	100.000 ± 3.534	22.969 ± 1.442	9.597 ± 0.805	6.328 ± 0.613	52.247 ± 2.171	8.860 ± 0.771
$15 < p_T^{fake} < 20$ GeV	100.000 ± 6.026	41.530 ± 3.422	11.553 ± 1.514	2.578 ± 0.628	34.480 ± 2.771	9.858 ± 1.342
$20 < p_T^{fake} < 25$ GeV	100.000 ± 9.876	29.147 ± 4.743	11.416 ± 2.438	4.378 ± 1.322	45.342 ± 5.397	9.718 ± 2.194
$p_T^{fake} < 25$ GeV	100.000 ± 5.145	38.069 ± 2.794	11.515 ± 1.286	3.081 ± 0.584	37.516 ± 2.494	9.819 ± 1.145
$p_T^{fake} > 25$ GeV	100.000 ± 4.853	8.173 ± 1.133	7.718 ± 0.982	9.509 ± 1.080	66.679 ± 3.585	7.921 ± 1.038
$25 < p_T^{fake} < 35$ GeV	100.000 ± 7.761	13.297 ± 2.362	8.977 ± 1.699	8.077 ± 1.544	61.990 ± 5.394	7.658 ± 1.702
$p_T^{fake} > 35$ GeV	100.000 ± 6.213	4.664 ± 1.048	6.856 ± 1.179	10.490 ± 1.481	69.890 ± 4.777	8.100 ± 1.304
Anti-ID cuts	100.000 ± 1.126	15.329 ± 0.366	11.778 ± 0.291	15.457 ± 0.329	55.344 ± 0.727	2.092 ± 0.113
$15 < p_T^{fake} < 20$ GeV	100.000 ± 1.676	18.238 ± 0.595	14.463 ± 0.486	13.530 ± 0.452	50.779 ± 1.017	2.989 ± 0.200
$20 < p_T^{fake} < 25$ GeV	100.000 ± 2.623	17.510 ± 0.915	12.446 ± 0.686	13.242 ± 0.699	55.045 ± 1.687	1.757 ± 0.243
$p_T^{fake} < 25$ GeV	100.000 ± 1.412	18.028 ± 0.499	13.880 ± 0.398	13.447 ± 0.380	52.012 ± 0.872	2.633 ± 0.159
$p_T^{fake} > 25$ GeV	100.000 ± 1.867	10.591 ± 0.503	8.087 ± 0.398	18.986 ± 0.619	61.194 ± 1.295	1.142 ± 0.138
$25 < p_T^{fake} < 35$ GeV	100.000 ± 2.656	14.745 ± 0.847	10.338 ± 0.644	15.510 ± 0.782	58.034 ± 1.768	1.373 ± 0.215
$p_T^{fake} > 35$ GeV	100.000 ± 2.625	6.516 ± 0.562	5.879 ± 0.474	22.396 ± 0.962	64.294 ± 1.891	0.915 ± 0.174
Electron Fake-Factor	100.000 ± 3.534	22.969 ± 1.442	9.597 ± 0.805	6.328 ± 0.613	52.247 ± 2.171	8.860 ± 0.771
Electron Fake-Factor for $15 < p_T^{fake} < 20$ GeV	100.000 ± 6.026	41.530 ± 3.422	11.553 ± 1.514	2.578 ± 0.628	34.480 ± 2.771	9.858 ± 1.342
Electron Fake-Factor for $20 < p_T^{fake} < 25$ GeV	100.000 ± 9.876	29.147 ± 4.743	11.416 ± 2.438	4.378 ± 1.322	45.342 ± 5.397	9.718 ± 2.194
Electron Fake-Factor for $p_T^{fake} < 25$ GeV	100.000 ± 5.145	38.069 ± 2.794	11.515 ± 1.286	3.081 ± 0.584	37.516 ± 2.494	9.819 ± 1.145
Electron Fake-Factor for $p_T^{fake} > 25$ GeV	100.000 ± 4.853	8.173 ± 1.133	7.718 ± 0.982	9.509 ± 1.080	66.679 ± 3.585	7.921 ± 1.038
Electron Fake-Factor for $25 < p_T^{fake} < 35$ GeV	100.000 ± 7.761	13.297 ± 2.362	8.977 ± 1.699	8.077 ± 1.544	61.990 ± 5.394	7.658 ± 1.702
Electron Fake-Factor for $p_T^{fake} > 35$ GeV	100.000 ± 6.213	4.664 ± 1.048	6.856 ± 1.179	10.490 ± 1.481	69.890 ± 4.777	8.100 ± 1.304
fake type: muon	100.000 ± 1.041	55.144 ± 0.703	26.655 ± 0.409	9.090 ± 0.214	7.112 ± 0.186	1.999 ± 0.105
fake mu η cut	100.000 ± 1.056	54.994 ± 0.712	26.516 ± 0.413	9.271 ± 0.220	7.278 ± 0.191	1.941 ± 0.105
Truth-matched Z-candidate leptons	100.000 ± 1.059	55.062 ± 0.715	26.548 ± 0.415	9.286 ± 0.221	7.261 ± 0.192	1.843 ± 0.095
ID cuts	100.000 ± 5.193	57.732 ± 3.605	25.998 ± 2.012	3.451 ± 0.632	6.512 ± 0.888	6.306 ± 0.889
$15 < p_T^{fake} < 20$ GeV	100.000 ± 6.836	66.192 ± 5.189	22.892 ± 2.400	0.990 ± 0.409	3.532 ± 0.821	6.394 ± 1.194
$20 < p_T^{fake} < 25$ GeV	100.000 ± 12.304	58.955 ± 8.566	25.301 ± 4.831	6.501 ± 2.155	6.090 ± 2.016	3.154 ± 1.449
$p_T^{fake} < 25$ GeV	100.000 ± 5.975	64.495 ± 4.449	23.457 ± 2.158	2.282 ± 0.586	4.131 ± 0.785	5.635 ± 0.973
$p_T^{fake} > 25$ GeV	100.000 ± 10.499	36.322 ± 5.548	34.044 ± 4.898	7.153 ± 1.895	14.047 ± 2.812	8.434 ± 2.065
Anti-ID cuts	100.000 ± 1.162	59.931 ± 0.827	29.265 ± 0.481	6.088 ± 0.192	4.005 ± 0.152	0.712 ± 0.064
$15 < p_T^{fake} < 20$ GeV	100.000 ± 1.665	60.379 ± 1.187	29.374 ± 0.694	5.353 ± 0.258	3.985 ± 0.220	0.909 ± 0.103
$20 < p_T^{fake} < 25$ GeV	100.000 ± 2.587	60.847 ± 1.863	28.950 ± 1.057	6.023 ± 0.423	3.705 ± 0.325	0.475 ± 0.112
$p_T^{fake} < 25$ GeV	100.000 ± 1.400	60.518 ± 1.001	29.247 ± 0.580	5.552 ± 0.221	3.902 ± 0.182	0.780 ± 0.079
$p_T^{fake} > 25$ GeV	100.000 ± 2.085	58.671 ± 1.467	29.302 ± 0.862	7.236 ± 0.375	4.225 ± 0.276	0.565 ± 0.104
Muon Fake-Factor	100.000 ± 5.193	57.732 ± 3.605	25.998 ± 2.012	3.451 ± 0.632	6.512 ± 0.888	6.306 ± 0.889
Muon Fake-Factor for $15 < p_T^{fake} < 20$ GeV	100.000 ± 6.836	66.192 ± 5.189	22.892 ± 2.400	0.990 ± 0.409	3.532 ± 0.821	6.394 ± 1.194
Muon Fake-Factor for $20 < p_T^{fake} < 25$ GeV	100.000 ± 12.304	58.955 ± 8.566	25.301 ± 4.831	6.501 ± 2.155	6.090 ± 2.016	3.154 ± 1.449
Muon Fake-Factor for $p_T^{fake} < 25$ GeV	100.000 ± 5.975	64.495 ± 4.449	23.457 ± 2.158	2.282 ± 0.586	4.131 ± 0.785	5.635 ± 0.973
Muon Fake-Factor for $p_T^{fake} > 25$ GeV	100.000 ± 10.499	36.322 ± 5.548	34.044 ± 4.898	7.153 ± 1.895	14.047 ± 2.812	8.434 ± 2.065

B.4.2 W+jets Sample

B.4.2.1 Fake Electron

Table B.10: The W+jets flavour composition cutflow in the $e\mu + \mu e$ channel showing all the fraction of the fake flavor in each column. The upper part of the cut-flow shows the common selection together with the electron fake selection, followed by the id and anti-id selections and the split in p_{T} binning.

$\sqrt{s} = 13TeV$, (Full Run 2)	Wjets Powheg	bottom	charm	strange	light	other
Channel Selection	703502.639 ± 2733.021	1407005.278 ± 3865.076	1407005.278 ± 3865.076	1407005.278 ± 3865.076	1407005.278 ± 3865.076	1407005.278 ± 3865.076
OS Leptons	442926.989 ± 2158.449	9236.935 ± 296.416	132144.576 ± 1150.535	67524.010 ± 861.371	186715.912 ± 1418.119	47305.555 ± 703.056
electron is fake	356116.088 ± 1971.555	3238.641 ± 190.055	63125.692 ± 842.061	62650.410 ± 835.287	181923.433 ± 1402.781	45177.912 ± 690.174
Fake El ID	9714.687 ± 323.922	596.111 ± 82.759	2052.586 ± 151.155	539.988 ± 77.424	3720.735 ± 199.512	2805.267 ± 171.549
Fake el pT cut	5557.599 ± 239.633	114.910 ± 35.454	864.744 ± 96.336	395.126 ± 65.858	2315.490 ± 154.137	1867.329 ± 137.080
Fake el eta cut	5516.318 ± 238.741	114.910 ± 35.454	864.744 ± 96.336	395.126 ± 65.858	2285.157 ± 153.138	1856.381 ± 136.642
Vgamma/Vjets Overlap	5189.827 ± 231.650	114.910 ± 35.454	864.744 ± 96.336	385.838 ± 65.200	2285.157 ± 153.138	1539.178 ± 124.184
prompt mu pT cut	4732.602 ± 220.676	114.910 ± 35.454	857.598 ± 96.071	332.683 ± 59.195	2131.974 ± 147.343	1295.438 ± 114.005
$M_{\ell\ell} > 10$	3896.668 ± 201.713	85.397 ± 30.985	857.598 ± 96.071	332.683 ± 59.195	2121.288 ± 146.955	499.703 ± 73.473
prompt mu ID	3424.942 ± 187.901	74.626 ± 29.053	758.083 ± 89.885	267.552 ± 52.384	1917.142 ± 138.297	407.539 ± 67.180
Fake el pT 15-20	902.145 ± 93.910	37.037 ± 18.689	256.698 ± 50.754	26.527 ± 15.324	445.872 ± 65.443	136.011 ± 37.096
Fake el pT 20-25	600.451 ± 78.352	21.435 ± 15.292	136.676 ± 35.849	40.103 ± 20.180	334.048 ± 57.303	68.190 ± 30.481
Fake el pT 25-35	859.111 ± 92.317	0.000	194.952 ± 45.083	113.024 ± 34.803	393.558 ± 61.557	157.576 ± 38.592
Fake el pT 35-1000	1063.235 ± 108.749	16.154 ± 16.154	169.758 ± 46.752	87.897 ± 29.846	743.664 ± 88.150	45.761 ± 26.806
Fake El AntiID	121945.371 ± 1149.643	1998.122 ± 147.395	33112.806 ± 606.908	14907.878 ± 402.832	58953.743 ± 795.736	12972.822 ± 368.979
Fake el pT cut	37588.956 ± 619.871	468.389 ± 70.473	8874.442 ± 304.507	4360.456 ± 212.502	18873.475 ± 437.126	5012.195 ± 224.307
Fake el eta cut	37308.873 ± 617.705	468.389 ± 70.473	8820.354 ± 303.671	4305.485 ± 211.295	18729.036 ± 435.493	4985.608 ± 223.781
Vgamma/Vjets Overlap	35875.004 ± 605.788	468.389 ± 70.473	8791.263 ± 303.147	4288.331 ± 210.932	18646.033 ± 434.478	3680.989 ± 191.967
prompt mu pT cut	35046.951 ± 598.303	468.389 ± 70.473	8722.058 ± 301.905	4180.747 ± 207.743	18282.546 ± 429.711	3393.211 ± 184.511
$M_{\ell\ell} > 10$	32591.253 ± 577.756	457.482 ± 69.624	8645.479 ± 300.680	4131.235 ± 206.555	18085.996 ± 427.524	1271.061 ± 114.475
prompt mu ID	28989.011 ± 543.471	410.290 ± 66.152	7796.808 ± 285.690	3627.976 ± 191.435	16148.428 ± 402.956	1005.509 ± 101.711
Fake el pT 15-20	13576.435 ± 371.848	323.221 ± 59.216	4543.287 ± 218.480	1409.880 ± 119.997	6800.346 ± 259.384	499.702 ± 73.155
Fake el pT 20-25	5915.422 ± 243.117	44.317 ± 20.010	1683.568 ± 132.246	671.021 ± 81.205	3295.912 ± 179.958	220.605 ± 47.299
Fake el pT 25-35	4815.645 ± 219.702	9.980 ± 9.980	976.250 ± 98.836	780.363 ± 87.909	2845.419 ± 169.591	203.634 ± 43.726
Fake el pT 35-1000	4681.508 ± 222.969	32.773 ± 19.223	593.703 ± 81.410	766.713 ± 89.028	3206.751 ± 184.249	81.569 ± 29.058

Table B.11: The W+jets flavour composition percentage cutflow in the $e\mu+\mu e$ channel showing all the fraction of the fake flavor in each column. The upper part of the cutflow shows the common selection together with the electron fake selection, followed by the id and anti-id selections and the split in p_{T} binning.

$\sqrt{s} = 13TeV$, (Full Run 2)	Wjets Powheg %	bottom %	charm %	strange %	light %	other %
Channel Selection	100.000 ± 0.549	200.000 ± 0.952	200.000 ± 0.952	200.000 ± 0.952	200.000 ± 0.952	200.000 ± 0.952
OS Leptons	100.000 ± 0.689	2.085 ± 0.068	29.834 ± 0.298	15.245 ± 0.208	42.155 ± 0.380	10.680 ± 0.167
electron is fake	100.000 ± 0.783	0.909 ± 0.054	17.726 ± 0.256	17.593 ± 0.254	51.085 ± 0.485	12.686 ± 0.206
Fake El ID	100.000 ± 4.715	6.136 ± 0.876	21.129 ± 1.708	5.558 ± 0.818	38.300 ± 2.418	28.877 ± 2.011
Fake el pT cut	100.000 ± 6.098	2.068 ± 0.644	15.560 ± 1.859	7.110 ± 1.224	41.663 ± 3.304	33.600 ± 2.861
Fake el eta cut	100.000 ± 6.121	2.083 ± 0.649	15.676 ± 1.874	7.163 ± 1.233	41.425 ± 3.305	33.653 ± 2.874
Vgamma/Vjets Overlap	100.000 ± 6.312	2.214 ± 0.690	16.662 ± 2.000	7.435 ± 1.299	44.031 ± 3.545	29.658 ± 2.735
prompt mu pT cut	100.000 ± 6.594	2.428 ± 0.758	18.121 ± 2.199	7.030 ± 1.293	45.049 ± 3.756	27.373 ± 2.726
$M_{\ell\ell} > 10$	100.000 ± 7.321	2.192 ± 0.803	22.008 ± 2.716	8.538 ± 1.582	54.439 ± 4.708	12.824 ± 1.999
prompt mu ID	100.000 ± 7.759	2.179 ± 0.857	22.134 ± 2.892	7.812 ± 1.588	55.976 ± 5.073	11.899 ± 2.067
Fake el pT 15-20	100.000 ± 14.721	4.105 ± 2.115	28.454 ± 6.358	2.940 ± 1.726	49.424 ± 8.893	15.076 ± 4.401
Fake el pT 20-25	100.000 ± 18.454	3.570 ± 2.589	22.762 ± 6.668	6.679 ± 3.472	55.633 ± 11.991	11.356 ± 5.288
Fake el pT 25-35	100.000 ± 15.197	0.000	22.692 ± 5.787	13.156 ± 4.291	45.810 ± 8.693	18.342 ± 4.905
Fake el pT 35-1000	100.000 ± 14.465	1.519 ± 1.527	15.966 ± 4.691	8.267 ± 2.932	69.944 ± 10.951	4.304 ± 2.559
Fake El AntiID	100.000 ± 1.333	1.639 ± 0.122	27.154 ± 0.560	12.225 ± 0.350	48.344 ± 0.796	10.638 ± 0.319
Fake el pT cut	100.000 ± 2.332	1.246 ± 0.189	23.609 ± 0.899	11.600 ± 0.597	50.210 ± 1.428	13.334 ± 0.636
Fake el eta cut	100.000 ± 2.341	1.255 ± 0.190	23.641 ± 0.903	11.540 ± 0.598	50.200 ± 1.433	13.363 ± 0.639
Vgamma/Vjets Overlap	100.000 ± 2.388	1.306 ± 0.198	24.505 ± 0.941	11.954 ± 0.622	51.975 ± 1.496	10.261 ± 0.562
prompt mu pT cut	100.000 ± 2.414	1.336 ± 0.202	24.887 ± 0.961	11.929 ± 0.627	52.166 ± 1.515	9.682 ± 0.552
$M_{\ell\ell} > 10$	100.000 ± 2.507	1.404 ± 0.215	26.527 ± 1.036	12.676 ± 0.672	55.493 ± 1.640	3.900 ± 0.358
prompt mu ID	100.000 ± 2.651	1.415 ± 0.230	26.896 ± 1.107	12.515 ± 0.701	55.705 ± 1.739	3.469 ± 0.357
Fake el pT 15-20	100.000 ± 3.873	2.381 ± 0.441	33.465 ± 1.852	10.385 ± 0.929	50.089 ± 2.352	3.681 ± 0.548
Fake el pT 20-25	100.000 ± 5.812	0.749 ± 0.340	28.461 ± 2.523	11.344 ± 1.450	55.717 ± 3.808	3.729 ± 0.814
Fake el pT 25-35	100.000 ± 6.452	0.207 ± 0.207	20.272 ± 2.251	16.205 ± 1.970	59.087 ± 4.435	4.229 ± 0.928
Fake el pT 35-1000	100.000 ± 6.736	0.700 ± 0.412	12.682 ± 1.841	16.377 ± 2.055	68.498 ± 5.112	1.742 ± 0.626

B.4.2.2 Fake Muon

Table B.12: The W+jets flavour composition cutflow in the $e\mu + \mu e$ channel showing all the fraction of the fake flavor in each column. The upper part of the cut-flow shows the common selection together with the muon fake selection, followed by the id and anti-id selections and the split in p_{T} binning.

$\sqrt{s} = 13TeV$, (Full Run 2)	Wjets Powheg	bottom	charm	strange	light	other
Channel Selection	703502.639 ± 2733.021	1407005.278 ± 3865.076	1407005.278 ± 3865.076	1407005.278 ± 3865.076	1407005.278 ± 3865.076	1407005.278 ± 3865.076
OS Leptons	442926.989 ± 2158.449	9236.935 ± 296.416	132144.576 ± 1150.535	67524.010 ± 861.371	186715.912 ± 1418.119	47305.555 ± 703.056
muon is fake	85456.366 ± 870.821	5856.652 ± 224.248	68382.358 ± 780.307	4784.898 ± 208.525	4525.461 ± 199.537	1906.997 ± 125.955
Fake Mu ID	4157.535 ± 187.250	696.767 ± 77.023	2834.468 ± 155.724	220.726 ± 40.821	226.149 ± 44.045	179.424 ± 35.693
Fake mu pT cut	1548.908 ± 116.425	133.450 ± 33.822	1178.923 ± 102.195	99.732 ± 28.541	86.690 ± 27.910	50.114 ± 19.327
Fake mu η cut	1548.908 ± 116.425	133.450 ± 33.822	1178.923 ± 102.195	99.732 ± 28.541	86.690 ± 27.910	50.114 ± 19.327
Vgamma/Vjets Overlap	1548.908 ± 116.425	133.450 ± 33.822	1178.923 ± 102.195	99.732 ± 28.541	86.690 ± 27.910	50.114 ± 19.327
prompt el pT cut	1404.110 ± 109.642	116.884 ± 31.489	1080.145 ± 96.894	93.537 ± 27.861	72.160 ± 23.830	41.384 ± 17.243
$M_{\ell\ell} > 10$	1404.110 ± 109.642	116.884 ± 31.489	1080.145 ± 96.894	93.537 ± 27.861	72.160 ± 23.830	41.384 ± 17.243
prompt el ID	1247.274 ± 103.183	105.578 ± 29.389	950.041 ± 90.935	93.537 ± 27.861	72.160 ± 23.830	25.958 ± 13.027
Fake mu pT 15-20	584.626 ± 70.838	79.137 ± 26.014	440.286 ± 61.974	36.115 ± 16.584	16.768 ± 12.207	12.319 ± 8.748
Fake mu pT 20-25	290.716 ± 49.851	26.441 ± 13.675	215.046 ± 43.066	21.604 ± 15.482	27.624 ± 14.275	0.000
Fake mu pT 25-1000	371.933 ± 56.067	0.000	294.709 ± 50.733	35.818 ± 16.171	27.767 ± 14.665	13.639 ± 9.652
Fake Mu AntiID	79526.107 ± 841.830	5024.997 ± 208.169	64884.025 ± 760.844	4536.147 ± 203.986	4275.466 ± 194.248	805.473 ± 84.372
Fake mu pT cut	33722.137 ± 552.401	1557.671 ± 115.341	28225.080 ± 505.906	1855.999 ± 130.085	1731.129 ± 126.016	352.258 ± 55.681
Fake mu η cut	33722.137 ± 552.401	1557.671 ± 115.341	28225.080 ± 505.906	1855.999 ± 130.085	1731.129 ± 126.016	352.258 ± 55.681
Vgamma/Vjets Overlap	33646.526 ± 551.735	1557.671 ± 115.341	28149.469 ± 505.178	1855.999 ± 130.085	1731.129 ± 126.016	352.258 ± 55.681
prompt el pT cut	31093.074 ± 528.619	1340.177 ± 105.550	26152.052 ± 485.486	1701.328 ± 123.581	1567.383 ± 120.185	332.135 ± 53.705
$M_{\ell\ell} > 10$	31006.573 ± 528.009	1334.203 ± 105.380	26087.183 ± 484.985	1692.778 ± 123.285	1567.383 ± 120.185	325.025 ± 53.232
prompt el ID	27938.913 ± 502.093	1158.078 ± 98.213	23596.485 ± 461.914	1556.480 ± 118.365	1368.268 ± 112.955	259.602 ± 48.142
Fake mu pT 15-20	13296.852 ± 344.994	707.113 ± 77.119	11126.733 ± 315.685	682.723 ± 78.011	625.290 ± 76.956	154.992 ± 37.532
Fake mu pT 20-25	6416.369 ± 240.897	211.377 ± 40.519	5469.764 ± 223.072	365.788 ± 57.703	327.748 ± 53.904	41.692 ± 19.832
Fake mu pT 25-1000	8225.692 ± 273.943	239.588 ± 45.350	6999.987 ± 252.877	507.969 ± 67.786	415.230 ± 62.698	62.919 ± 22.709

Table B.13: The W+jets flavour composition percentage cutflow in the $e\mu + \mu e$ channel showing all the fraction of the fake flavor in each column. The upper part of the cutflow shows the common selection together with the muon fake selection, followed by the id and anti-id selections and the split in p_{T} binning.

$\sqrt{s} = 13TeV$, $\mathcal{L} = 139fb^{-1}$ (Full Run 2)	Wjets Powheg %	bottom %	charm %	strange %	light %	other %
Channel Selection	100.000 ± 0.549	200.000 ± 0.952	200.000 ± 0.952	200.000 ± 0.952	200.000 ± 0.952	200.000 ± 0.952
OS Leptons	100.000 ± 0.689	2.085 ± 0.068	29.834 ± 0.298	15.245 ± 0.208	42.155 ± 0.380	10.680 ± 0.167
muon is fake	100.000 ± 1.441	6.853 ± 0.272	80.020 ± 1.224	5.599 ± 0.251	5.296 ± 0.240	2.232 ± 0.149
Fake Mu ID	100.000 ± 6.369	16.759 ± 2.000	68.177 ± 4.843	5.309 ± 1.011	5.439 ± 1.087	4.316 ± 0.880
Fake mu pT cut	100.000 ± 10.630	8.616 ± 2.278	76.113 ± 8.733	6.439 ± 1.905	5.597 ± 1.850	3.235 ± 1.271
Fake mu eta cut	100.000 ± 10.630	8.616 ± 2.278	76.113 ± 8.733	6.439 ± 1.905	5.597 ± 1.850	3.235 ± 1.271
Vgamma/Vjets Overlap	100.000 ± 10.630	8.616 ± 2.278	76.113 ± 8.733	6.439 ± 1.905	5.597 ± 1.850	3.235 ± 1.271
prompt el pT cut	100.000 ± 11.043	8.324 ± 2.335	76.927 ± 9.149	6.662 ± 2.051	5.139 ± 1.744	2.947 ± 1.249
$M_{\ell\ell} > 10$	100.000 ± 11.043	8.324 ± 2.335	76.927 ± 9.149	6.662 ± 2.051	5.139 ± 1.744	2.947 ± 1.249
prompt el ID	100.000 ± 11.699	8.465 ± 2.458	76.169 ± 9.636	7.499 ± 2.318	5.785 ± 1.970	2.081 ± 1.058
Fake mu pT 15-20	100.000 ± 17.136	13.536 ± 4.742	75.311 ± 13.987	6.178 ± 2.934	2.868 ± 2.117	2.107 ± 1.518
Fake mu pT 20-25	100.000 ± 24.251	9.095 ± 4.956	73.971 ± 19.502	7.431 ± 5.476	9.502 ± 5.174	0.000
Fake mu pT 25-1000	100.000 ± 21.319	0.000	79.237 ± 18.131	9.630 ± 4.584	7.466 ± 4.100	3.667 ± 2.653
Fake Mu AntiID	100.000 ± 1.497	6.319 ± 0.270	81.588 ± 1.289	5.704 ± 0.264	5.376 ± 0.251	1.013 ± 0.107
Fake mu pT cut	100.000 ± 2.317	4.619 ± 0.350	83.699 ± 2.032	5.504 ± 0.396	5.134 ± 0.383	1.045 ± 0.166
Fake mu eta cut	100.000 ± 2.317	4.619 ± 0.350	83.699 ± 2.032	5.504 ± 0.396	5.134 ± 0.383	1.045 ± 0.166
Vgamma/Vjets Overlap	100.000 ± 2.319	4.630 ± 0.351	83.662 ± 2.034	5.516 ± 0.397	5.145 ± 0.384	1.047 ± 0.166
prompt el pT cut	100.000 ± 2.404	4.310 ± 0.347	84.109 ± 2.117	5.472 ± 0.408	5.041 ± 0.396	1.068 ± 0.174
$M_{\ell\ell} > 10$	100.000 ± 2.408	4.303 ± 0.348	84.134 ± 2.121	5.459 ± 0.408	5.055 ± 0.397	1.048 ± 0.173
prompt el ID	100.000 ± 2.541	4.145 ± 0.359	84.457 ± 2.244	5.571 ± 0.435	4.897 ± 0.414	0.929 ± 0.173
Fake mu pT 15-20	100.000 ± 3.669	5.318 ± 0.596	83.679 ± 3.217	5.134 ± 0.602	4.703 ± 0.591	1.166 ± 0.284
Fake mu pT 20-25	100.000 ± 5.310	3.294 ± 0.643	85.247 ± 4.725	5.701 ± 0.924	5.108 ± 0.862	0.650 ± 0.310
Fake mu pT 25-1000	100.000 ± 4.710	2.913 ± 0.560	85.099 ± 4.181	6.175 ± 0.849	5.048 ± 0.781	0.765 ± 0.277

Table B.14: The SR cutflow in the $e\mu+\mu e$ channel. The upper part of the cutflow shows the common selection together with applied electron and muon fake factor selection, followed by the SR baseline selection. The split in electron and muon fake events is shown in column, two and three.

$\sqrt{s} = 13TeV, \mathcal{L} = 140 fb^{-1}$ (Full Run 2)	Higgs	WW	Other VV	$V\gamma$	triboson	$t\bar{t}$	Single Top	Z/γ^*	eFakes	μFakes	AllBkg
Channel Selection	16043.40 ± 12.56	205503.72 ± 149.30	104007.47 ± 244.63	313121.58 ± 855.55	483.65 ± 1.39	1830525.52 ± 363.92	186873.78 ± 169.54	1821464.44 ± 3443.21	6448387.99 ± 4873.42	6448387.99 ± 4873.42	17374799.54 ± 7767.34
W+jets flavour split muon	16043.40 ± 12.56	205503.72 ± 149.30	104007.47 ± 244.63	313121.58 ± 855.55	483.65 ± 1.39	1830525.52 ± 363.92	186873.78 ± 169.54	1821464.44 ± 3443.21	6448387.99 ± 4873.42	4512133.06 ± 2868.74	15438544.60 ± 6694.10
W+jets flavour split electron	16043.40 ± 12.56	205503.72 ± 149.30	104007.47 ± 244.63	313121.58 ± 855.55	483.65 ± 1.39	1830525.52 ± 363.92	186873.78 ± 169.54	1821464.44 ± 3443.21	3066314.75 ± 2739.30	4512133.06 ± 2868.74	12056471.36 ± 5344.58
Lepton ID	8780.22 ± 9.27	140700.39 ± 124.61	53556.89 ± 78.14	15044.12 ± 164.98	345.73 ± 1.18	1166332.26 ± 291.51	123686.49 ± 138.35	426375.89 ± 702.29	670164.52 ± 2046.16	708948.04 ± 1439.25	3313934.54 ± 2627.64
Electron FF, $15 < p_T^{e-fake} < 20$ GeV	8780.22 ± 9.27	140700.39 ± 124.61	53556.89 ± 78.14	15044.12 ± 164.98	345.73 ± 1.18	1166332.26 ± 291.51	123686.49 ± 138.35	426375.89 ± 702.29	574934.13 ± 1887.11	708948.04 ± 1439.25	3218704.16 ± 2505.77
Electron FF, $20 < p_T^{e-fake} < 25$ GeV	8780.22 ± 9.27	140700.39 ± 124.61	53556.89 ± 78.14	15044.12 ± 164.98	345.73 ± 1.18	1166332.26 ± 291.51	123686.49 ± 138.35	426375.89 ± 702.29	525787.81 ± 1762.58	708948.04 ± 1439.25	3169557.84 ± 2413.38
Electron FF, $25 < p_T^{e-fake} < 35$ GeV	8780.22 ± 9.27	140700.39 ± 124.61	53556.89 ± 78.14	15044.12 ± 164.98	345.73 ± 1.18	1166332.26 ± 291.51	123686.49 ± 138.35	426375.89 ± 702.29	488731.86 ± 1660.05	708948.04 ± 1439.25	3132501.89 ± 2339.55
Electron FF, $p_T^{e-fake} > 35$ GeV	8780.22 ± 9.27	140700.39 ± 124.61	53556.89 ± 78.14	15044.12 ± 164.98	345.73 ± 1.18	1166332.26 ± 291.51	123686.49 ± 138.35	426375.89 ± 702.29	458261.04 ± 1575.35	708948.04 ± 1439.25	3102031.06 ± 2280.23
Muon FF, $15 < p_T^{m-fake} < 20$ GeV	8780.22 ± 9.27	140700.39 ± 124.61	53556.89 ± 78.14	15044.12 ± 164.98	345.73 ± 1.18	1166332.26 ± 291.51	123686.49 ± 138.35	426375.89 ± 702.29	458261.04 ± 1575.35	559653.56 ± 1286.09	2952736.59 ± 2186.78
Muon FF, $20 < p_T^{m-fake} < 25$ GeV	8780.22 ± 9.27	140700.39 ± 124.61	53556.89 ± 78.14	15044.12 ± 164.98	345.73 ± 1.18	1166332.26 ± 291.51	123686.49 ± 138.35	426375.89 ± 702.29	458261.04 ± 1575.35	486549.44 ± 1182.37	2879632.46 ± 2127.44
Muon FF, $p_T^{m-fake} > 25$ GeV	8780.22 ± 9.27	140700.39 ± 124.61	53556.89 ± 78.14	15044.12 ± 164.98	345.73 ± 1.18	1166332.26 ± 291.51	123686.49 ± 138.35	426375.89 ± 702.29	458261.04 ± 1575.35	394940.70 ± 946.48	2788023.72 ± 2005.95
$p_T^{lead} > 27$ GeV	8079.25 ± 8.88	136979.51 ± 122.92	50880.68 ± 76.32	12195.16 ± 132.26	341.28 ± 1.17	1146388.21 ± 289.08	121977.84 ± 137.42	334395.42 ± 639.99	402152.55 ± 1468.27	364294.84 ± 894.29	2577684.73 ± 1873.99
$p_T^{subl} > 15$ GeV	6661.59 ± 8.04	129829.53 ± 119.53	45062.63 ± 69.23	8622.47 ± 122.91	325.64 ± 1.14	1081801.37 ± 280.97	115493.76 ± 133.81	259491.74 ± 558.19	26198.68 ± 206.33	16902.22 ± 72.58	1690389.63 ± 700.37
GRL Selection	6661.59 ± 8.04	129829.53 ± 119.53	45062.63 ± 69.23	8622.47 ± 122.91	325.64 ± 1.14	1081801.37 ± 280.97	115493.76 ± 133.81	259491.74 ± 558.19	26198.68 ± 206.33	16902.22 ± 72.58	1690389.63 ± 700.37
Jet Cleaning	6661.59 ± 8.04	129829.53 ± 119.53	45062.63 ± 69.23	8622.47 ± 122.91	325.64 ± 1.14	1081801.37 ± 280.97	115493.76 ± 133.81	259491.74 ± 558.19	26198.68 ± 206.33	16902.22 ± 72.58	1690389.63 ± 700.37
Overlap: Vgamma/Vjets	6661.59 ± 8.04	129829.53 ± 119.53	45062.63 ± 69.23	8622.47 ± 122.91	325.64 ± 1.14	1081801.37 ± 280.97	115493.76 ± 133.81	253521.25 ± 503.15	28771.90 ± 189.59	17058.00 ± 72.11	1687148.14 ± 652.23
Trigger Selection	6196.99 ± 7.74	125520.62 ± 117.54	43713.92 ± 68.01	7866.56 ± 118.03	320.55 ± 1.13	1048903.36 ± 276.60	112347.27 ± 132.00	230773.05 ± 484.83	25263.12 ± 178.50	13510.67 ± 67.56	1614416.12 ± 630.75
Trigger Matching	5988.63 ± 7.57	123460.76 ± 116.26	42134.96 ± 67.18	7411.30 ± 112.98	312.65 ± 1.12	1033705.49 ± 273.45	110829.40 ± 130.75	219825.10 ± 466.78	24305.08 ± 172.38	13594.77 ± 66.60	1581568.14 ± 612.17
Only two Leptons	5830.10 ± 7.56	123448.66 ± 116.26	20120.56 ± 56.87	6492.66 ± 101.83	159.69 ± 0.85	1032387.58 ± 273.28	110493.08 ± 130.69	218703.45 ± 454.71	24207.29 ± 160.69	13474.98 ± 64.52	1555318.05 ± 596.40
OS Leptons	5671.73 ± 7.52	121888.99 ± 115.85	10533.96 ± 49.44	4330.15 ± 83.03	112.98 ± 0.72	1026352.84 ± 272.47	109773.89 ± 130.30	216636.14 ± 434.47	15604.56 ± 141.31	9511.17 ± 59.30	1520416.41 ± 571.21
3rd lepton veto + OS	5671.73 ± 7.52	121888.99 ± 115.85	10533.96 ± 49.44	4330.15 ± 83.03	112.98 ± 0.72	1026352.84 ± 272.47	109773.89 ± 130.30	216636.14 ± 434.47	15604.56 ± 141.31	9511.17 ± 59.30	1520416.41 ± 571.21
$M_{\ell\ell} > 10$ GeV	5633.09 ± 7.50	121799.26 ± 115.81	10503.07 ± 49.39	4291.53 ± 82.69	112.93 ± 0.72	1025690.43 ± 272.38	109712.43 ± 130.27	216553.79 ± 432.84	15529.51 ± 141.19	8468.58 ± 58.24	1518294.61 ± 569.72
>=2jets	1099.15 ± 3.13	17016.44 ± 21.28	2502.96 ± 18.51	618.49 ± 25.33	35.57 ± 0.40	769726.34 ± 235.51	44922.19 ± 83.19	15928.47 ± 92.88	3805.39 ± 80.78	2518.69 ± 36.13	858173.69 ± 283.35
nBjets=0	770.52 ± 2.78	14149.91 ± 19.66	1985.72 ± 17.95	503.39 ± 22.97	27.21 ± 0.36	50605.63 ± 61.33	5524.96 ± 29.48	12681.77 ± 86.32	1294.88 ± 33.23	694.10 ± 14.64	88238.11 ± 121.01
$M_{\mathrm{T}} < 70$ GeV	664.94 ± 2.59	7942.84 ± 14.92	1010.03 ± 9.19	256.61 ± 16.88	15.63 ± 0.27	29918.15 ± 47.08	3190.02 ± 22.36	4288.00 ± 58.18	746.70 ± 24.34	419.22 ± 10.64	48452.14 ± 86.06
$M_T < 300$ GeV	663.17 ± 2.59	7217.20 ± 14.23	944.17 ± 9.10	246.74 ± 15.92	11.77 ± 0.24	27272.75 ± 44.91	2861.64 ± 21.14	4285.40 ± 58.18	701.11 ± 23.68	414.28 ± 10.47	44618.21 ± 84.05

Appendix C

Additional Material

C.1 Fake Factor Dependency on b-jets

In this section, the Z+jets sample is divided into a selection with or without b-jets. The aim is to attempt to confirm the results from the simulation-based sample composition studies described in Section 6.6 in the data. The category of the b-jets jets is applied in addition to the selection of Z+jets as shown in Section 6.4.1. In Table C.1, fake yields for the b-jet selection and b-veto are shown in MC SHERPA and data. The ratio of the number of Z+jets and fakes in the id and anti-id region after the b-veto and b-jet selection is taken. For electrons, in MC Z+jets(fakes), the ratio between these two categories is 7% (20%) and 9% (11%) respectively for "id" and "anti-id" categories. In Table C.2, fake yields for the b-jet selection and b-veto are shown for Z+jets MC SHERPA and data. For muons, in MC Z+jets(fakes), the ratio between these two categories is 14% (37%) and 18% (24%) respectively for id and anti-id categories.

The FFs values in "id" and "anti-id" region from Table C.1 are taken for comparison with the FFs obtained from b-jet section and b-veto. Most of the events are lost when the b-jet selection is applied especially in the "id" region, making it difficult to make a meaning full comparison. Nonetheless, we note that the b-jet selection category

Table C.1: Electron fake yields for the Z+jets selection for the two subcategories with b-jets and without b-jets. The MC Z+jets yields are taken from SHERPA. The uncertainties are statistical only.

Selections	Z+jets	Fakes (Data - Prompt)
ID cuts	989.65 ± 113.03	2109.63 ± 79.06
b-veto	923.74 ± 112.61	1759.90 ± 73.98
b-selec	66.55 ± 9.50	347.40 ± 27.90
Anti-ID cuts	16911.80 ± 429.46	17777.70 ± 137.41
b-veto	15552.97 ± 425.92	15949.80 ± 130.19
b-selec	1364.99 ± 54.95	1827.29 ± 43.94

Table C.2: Muon fake yields for the Z+jets selection for the two subcategories with b-jets and without b-jets. The MC Z+jets yields are taken from SHERPA. The uncertainties are statistical only.

Selections	Z+jets	Fakes (Data - Prompt)
ID cuts	537.19 ± 58.14	728.09 ± 73.92
b-veto	471.45 ± 57.54	453.69 ± 68.78
b-selec	66.30 ± 8.39	272.69 ± 27.09
Anti-ID cuts	9973.61 ± 218.99	12128.29 ± 117.11
b-veto	8449.16 ± 214.03	9770.06 ± 105.84
b-selec	1553.49 ± 47.07	2357.83 ± 50.11

is consistent with FFs obtained from data for both electrons and muons within the broad uncertainties, see Table C.3, giving confidence that muon fakes are dominated by heavy flavour objects and electrons mostly come from photons as seen in Section 6.6.

Table C.3: Electron and muon FF comparison from data with b-section, and b-veto.

Fake type	Fakes (Data - Prompt) FF	b-veto FF	b-selection FF
Electron	0.12 ± 0.01	0.11± 0.01	0.19 ± 0.02
Muon	0.06 ± 0.01	0.05 ± 0.01	0.12 ± 0.01

References

[1] P.W. Higgs, *Broken symmetries, massless particles and gauge fields,* Physics Letters, **12**(2):132–133, 1964.
https://doi.org/10.1016/0031-9163(64)91136-9.

[2] P. W. Higgs, *Broken symmetries and the masses of gauge bosons,* Phys. Rev. Lett. **13**:508–509, Oct 1964.
https://doi.org/10.1103/PhysRevLett.13.508.

[3] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons,* Phys. Rev. Lett. **13**:321–323, Aug 1964.
https://doi.org/10.1103/PhysRevLett.13.321.

[4] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,* Phys. Lett. **B716** (2012) 1–29.
https://doi.org/10.1016/j.physletb.2012.08.020.

[5] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,* Phys. Lett. **B716** (2012) 30–61.
https://doi.org/10.1016/j.physletb.2012.08.021.

[6] S. von Buddenbrock et al., *The compatibility of LHC Run 1 data with a heavy scalar of mass around 270 GeV,*
arXiv1506.00612 [hep-ph].

[7] S. von Buddenbrock et al., *Phenomenological signatures of additional scalar bosons at the LHC,* Eur. Phys. J. **C76** (2016) no.10, 580, arXiv:1606.01674 [hep-ph].

[8] S. von Buddenbrock, A. S. Cornell, A. Fadol, M. Kumar, B. Mellado and X. Ruan, *Multi-lepton signatures of additional scalar bosons beyond the Standard Model at the LHC,* J. Phys. **G45** (2018) no.11, 115003, arXiv:1711.07874 [hep-ph].

[9] S. von Buddenbrock, A. S. Cornell, Y. Fang, A. Fadol Mohammed, M. Kumar, B. Mellado, and K. G. Tomiwa, *The emergence of multi-lepton anomalies at the LHC and their compatibility with new physics at the EW scale,* arXiv:1901.05300 [hep-ph].

[10] The ATLAS Collaboration, *Measurement of the WW cross section in $\sqrt{s} = 7$ TeV pp collisions with ATLAS,* Phys. Rev. Lett. **107** (2011) 041802, arXiv:1104.5225. 6.1, 6.9.1.

[11] D. J. Griffiths, *Introduction to elementary particles;* 2nd rev. version. Physics textbook. Wiley, New York, NY, 2008. https://cds.cern.ch/record/111880.

[12] G. Shaw B. R. Martin, *Particle Physics,* 3rd Edition. Physics textbook. Wiley, New York, NY, 1997.

[13] S. L. Glashow, *Partial-symmetries of weak interactions,* Nucl. Phys. **22** (1961) no. 4, 579. 1.1 https://doi.org/10.1016/0029-5582(61)90469-2.

[14] S. Weinberg, *A Model of Leptons,* Phys. Rev. Lett. **19**(1967) 1264. 1.1 https://doi.org/10.1103/PhysRevLett.19.1264.

[15] A. Salam, *Elementary Particle Theory,* p. 367. Almqvist and Wiksell, Stockholm, 1968. 1.1

[16] G. 't Hooft and M. Veltman, *Regularization and Renormalization of Gauge Fields,* Nucl. Phys. **B44** (1972) 189. 1.1. https://doi.org/10.1016/0550-3213(72)90279-9.

[17] S. Weinberg, *The Quantum Theory of Fields,* Cambridge Univ. Press, Cambridge, 1995. 1.1

[18] C. Rovelli, *Notes for a brief history of quantum gravity. In Recent developments in theoretical and experimental general relativity, gravitation and relativistic field theories*, Proceedings, 9th Marcel Grossmann Meeting, MG'9, Rome, Italy, July 2-8, 2000. Pts. A-C, pages 742–768, 2000.

[19] C. Rubbia, *Experimental Observation of the Intermediate Vector Bosons W^+, W^-, and Z^0,* Rev.Mod.Phys. **57** (1985) 699-722. https://doi.org/10.1103/RevModPhys.57.699.

[20] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons,* Phys. Rev.**145** (1966) 1156–1163. https://doi.org/10.1103/PhysRev.145.1156.

[21] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables,* http://cds.cern.ch/record/1318996.

[22] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions,* http://cds.cern.ch/record/1416519.

[23] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections:*

3. Inclusive Observables,
http://cds.cern.ch/record/1559921.

[24] LHC Higgs Cross Section Working Group. *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties (2013),* Ed. by S. Heinemeyer, C. Mariotti, G. Passarino and R. Tanaka.
arXiv: 1307.1347 [hep-ph].

[25] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions,*
http://cds.cern.ch/record/1416519, arXiv:1201.3084 [hep-ph].

[26] K. A. Olive et al. *Review of Particle Physics,* Chin. Phys., **C38**:090001, 2014.

[27] ATLAS Collaboration, *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector,* Phys. Lett. **B 786** (2018) 59,
arXiv:1808.08238 [hep-ex].

[28] ATLAS Collaboration, *Measurements of fiducial and differential cross sections for Higgs boson production in the diphoton decay channel at $\sqrt{s}$ = 8 TeV with ATLAS,* JHEP **09** (2014) 112,
arXiv:1407.4222 [hep-ex].

[29] ATLAS Collaboration, *Fiducial and differential cross sections of Higgs boson production measured in the four-lepton decay channel in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector,* Phys. Lett. **B738** (2014) 234–253,
arXiv:1408.3226 [hep-ex].

[30] CMS Collaboration, *Measurement of differential cross sections for Higgs boson production in the diphoton decay channel in pp collisions at $\sqrt{s}$ = 8 TeV,* Eur. Phys.J. **C76** no. 1, (2016) 13,
arXiv:1508.07819 [hep-ex].

[31] CMS Collaboration, *Measurement of differential and integrated fiducial cross sections for Higgs boson production in the four-lepton decay channel in pp collisions at $\sqrt{s}$ = 7 and 8 TeV,* JHEP 04 (2016) 005, arXiv:1512.08377 [hep-ex].

[32] ATLAS Collaboration, *Searches for Higgs boson pair production in the hh → bbττ, γγ WW*,γγ bb and bbbb channels with the ATLAS detector,* Phys. Rev. **D92** (2015) 092004, arXiv:1509.04670 [hep-ex].

[33] CMS Collaboration, *Search for two Higgs bosons in final states containing two photons and two bottom quarks in proton-proton collisions at 8 TeV,* Phys. Rev. **D94** no. 5, (2016) 052012, arXiv:1603.06896 [hep-ex].

[34] CMS Collaboration, *Searches for a heavy scalar boson H decaying to a pair of 125 GeV Higgs bosons hh or for a heavy pseudo-scalar boson A decaying to Zh, in the final states with, h → ττ,* Phys. Lett. **B755** (2016) 217–244, arXiv:1510.01181 [hep-ex].

[35] CMS Collaboration, *Searches for heavy Higgs bosons in two-Higgs-doublet models and for t → ch decay using multilepton and diphoton final states in pp collisions at 8 TeV,* Phys. Rev. **D90** (2014) 112013. https://link.aps.org/doi/10.1103/PhysRevD.90.112013.

[36] ATLAS Collaboration, *Search for H → γγ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector,* Phys. Lett. **B740** (2015) 222–242, arXiv:1409.3122 [hep-ex].

[37] ATLAS Collaboration, *Search for the associated production of the Higgs boson with a top quark pair in multilepton final states with the ATLAS detector,* Phys. Lett. **B749** (2015) 519–541,
arXiv:1506.05988 [hep-ex].

[38] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector,* Eur. Phys. J. **C75** no. 7, (2015) 349,
arXiv:1503.05066.

[39] CMS Collaboration, *Search for the associated production of the Higgs boson with a top-quark pair,* JHEP 09 (2014) 087,
arXiv:1408.1682 [hep-ex]. [hep-ex].

[40] ATLAS Collaboration, *Search for a high-mass Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector,* JHEP 01 (2016) 032,
arXiv:1509.00389 [hep-ex].

[41] ATLAS Collaboration, *Search for an additional, heavy Higgs boson in the H decay channel at $\sqrt{s}$ = 8 TeV in pp collision data with the ATLAS detector,* Eur. Phys. J. **C76** no. 1, (2016) 45,
arXiv:1507.05930 [hep-ex].

[42] CMS Collaboration, *Search for a Higgs boson in the mass range from 145 to 1000 GeV decaying to a pair of W or Z bosons,* JHEP 10 (2015) 144,
arXiv:1504.00936 [hep-ex].

[43] L. Evans and P. Bryant. *LHC Machine.* In: JINST **3** (2008), S08001.
DOI :10.1088/1748-0221/3/08/S08001.

[44] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider,* In: JINST **3** (2008), S08003. DOI : 10.1088/1748-0221/3/08/S08003.

[45] ATLAS Collaboration, *ATLAS: letter of intent for a general-purpose pp experiment at the large hadron collider at CERN,* http://cds.cern.ch/record/291061.

[46] ATLAS Collaboration, *ATLAS detector and physics performance:* Technical Design Report, 1, https://cdsweb.cern.ch/record/391176.

[47] ATLAS Collaboration, *ATLAS detector and physics performance:* Technical Design Report, 2, https://cdsweb.cern.ch/record/391177.

[48] CMS Collaboration, *The CMS experiment at the CERN LHC, Journal of Instrumentation* **3** (2008) S08004, http://stacks.iop.org/1748-0221/3/i=08/a=S08004.

[49] ALICE Collaboration, *The ALICE experiment at the CERN LHC, Journal of Instrumentation* **3** (2008) S08002, http://stacks.iop.org/1748-0221/3/i=08/a=S08002.

[50] LHCb Collaboration, *The LHCb Detector at the LHC, Journal of Instrumentation* **3** (2008) S08005, http://stacks.iop.org/1748-0221/3/i=08/a=S08005.

[51] CERN, *CERN Accelerator Complex,* https://beams.web.cern.ch/sites/beams.web.cern.ch/files/.

[52] Z. Marshall, *Simulation of Pile-up in the ATLAS Experiment,* J. Phys. Conf. Ser. (2014) 513: 022024.

[53] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1997.
https://cds.cern.ch/record/331063.

[54] S. Haywood, L. Rossi, R. Nickerson, and A. Romaniouk, *ATLAS inner detector: Technical Design Report, 2.* Technical Design Report ATLAS. CERN, Geneva, 1997.
https://cds.cern.ch/record/331064.

[55] ATLAS Collaboration, *ATLAS calorimeter performance: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1996.
https://cds.cern.ch/record/331059.

[56] ATLAS Collaboration, *Track Reconstruction Performance of the ATLAS Inner Detector at* $\sqrt{s} = 13$ *TeV,* Tech. Rep. ATL-PHYS-PUB-2015-018, CERN, Geneva, Jul, 2015.
https://cds.cern.ch/record/2037683.

[57] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report.* Technical Design Report ATLAS. Geneva: CERN, 2010.
https://cds.cern.ch/record/1291633.

[58] G. Ripellino, *The alignment of the ATLAS Inner Detector in Run-2,*
PoS, LHCP2016:196, 2016.

[59] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector,* Technical Design Report ATLAS. Geneva: CERN.
https://cds.cern.ch/record/2285585?ln=en.

[60] A. Abdesselam et al., *The ATLAS semiconductor tracker end-cap module, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* **575** no. 3, (2007) 353–389. 21

[61] A. Ahmad et al., *The silicon microstrip sensors of the ATLAS semiconductor tracker, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* **578** no. 1, (2007) 98–118. 21

[62] The ATLAS TRT Collaboration, *The ATLAS TRT end-cap detectors, Journal of Instrumentation,* **3** no. 10, (2008) P10003–P10003. 21

[63] The ATLAS TRT Collaboration, *The ATLAS TRT Barrel Detector," Journal of Instrumentation,* **3** (2008) P02014. 21

[64] ATLAS Collaboration. *The ATLAS Experiment at the CERN Large Hadron Collider,* In: JINST **3** (2008), S08003. DOI : 10.1088/1748-0221/3/08/S08003.

[65] ATLAS Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1996. https://cds.cern.ch/record/331061.

[66] ATLAS Collaboration, *ATLAS tile calorimeter: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1996. https://cds.cern.ch/record/331062.

[67] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1997. https://cds.cern.ch/record/331068.

[68] ATLAS Collaboration, *ATLAS Run-2, TDAQ Summary,* https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsDAQ

[69] ATLAS Collaboration, *ATLAS Run II, Data Summary,* https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2

[70] T. Sjostrand, S. Mrenna, and PZ. Skands. *PYTHIA 6.4 Physics and Manual.* In: JHEP 05 (2006), p. 026. DOI : 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph].

[71] Z. Was. TAUOLA, *The library for tau lepton decay, and KKMC / KORALB / KORALZ /... status report.* In: Nucl. Phys. Proc. Suppl. **98** (2001), p. 96. DOI : 10.1016/S0920-5632(01)01200-2. arXiv: hep-ph/0011305 [hep-ph].

[72] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms,* JHEP 11 (2004) 040, arXiv:hep-ph/0409146.

[73] H.-L. Lai et al., *New parton distributions for collider physics,* Phys. Rev. D **82** (2010) 074024, arXiv: 1007.2241 [hep-ph].

[74] ATLAS Collaboration, *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 [hep-ex].

[75] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis,* JHEP **09** (2002) 012, arXiv: hep-ph/0201195.

[76] T. Gleisberg et al., *Event generation with SHERPA 1.1,* JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph].

[77] R. D. Ball, NNPDF Collaboration, et al., *Parton distributions for the LHC Run II,* JHEP **02** (2015) 040, arXiv:1410.8849 [hep-ph].

[78] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data,* ATL-PHYS-PUB-2014-021, 2014, url:https://cds.cern.ch/record/1966419.

[79] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes,* Phys. Rev. D **82** (2010) 074018, arXiv:1005.3457 [hep-ph].

[80] F. Cascioli et al., *Precise Higgs-background predictions: merging NLO QCD and squared quark-loop corrections to four-lepton + 0,1 jet production,* JHEP 01 (2014) 046, arXiv: 1309.0500 [hep-ph].

[81] ATLAS Collaboration, *Combined Performance(CP) Groups,* https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AtlasPhysics.

[82] ATLAS Collaboration, *Electron identification measurements in ATLAS using* $\sqrt{s}$ *= 13 TeV data with 50 ns bunch spacing,* tech. rep. ATL-PHYS-PUB-2015-041, CERN, 2015, url: https://cds.cern.ch/record/2048202

[83] *Muon reconstruction performance in early* $\sqrt{s}$ *= 13 TeV data,* tech. rep. ATL-PHYS-PUB-2015-037, CERN, 2015, url: https://cds.cern.ch/record/2047831.

[84] K. Becker et al., *Optimisation note of the ggF+VBF analysis in* $H \to WW$ *using 36* fb^{-1} *of data collected with the ATLAS detector at* $\sqrt{s}$ *= 13 TeV,* tech. rep. ATL-COM-PHYS-2017-1089,CERN, 2017, url: https://cds.cern.ch/record/2276101

[85] M. Cacciari, G. P. Salam and G. Soyez, *The Anti-k(t) jet clustering algorithm,* JHEP 04 (2008) 063, arXiv: 0802.1189 [hep-ph].

[86] G. Aad et al., *Jet energy measurement with the ATLAS detector in proton-proton collisions at* $\sqrt{s}$ *= 7 TeV,* Eur. Phys. J. **C73** (2013) 2304, arXiv: 1112.6426 [hep-ex].

[87] Eur. Phys. J. **C73** (2013) 2304, arXiv: 1112.6426 [hep-ex]. *Tagging and suppression of pileup jets with the ATLAS detector,* tech. rep. ATLAS-CONF-2014-018,

CERN, 2014, url:
https://cds.cern.ch/record/1700870.

[88] M. H. Klein, F. Rubbo and A. Schwartzman, *Forward Jet Vertex Tagging: A new technique for the identification and rejection of forward pileup jets,* tech. rep. ATL-COM-PHYS-2015-723, CERN, 2015, url:
https://cds.cern.ch/record/2034507.

[89] G. Aad et al., *Performance of b-jet Identification in the ATLAS Experiment,* (2015), arXiv:
1512.01094 [hep-ex].

[90] *Higgs decay to $\tau^+\tau$ a possible signature of intermediate mass higgs bosons at high energy hadron colliders,* Nuclear Physics **B297**(2):221 – 243, 1988.

[91] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment Detector, Trigger and Physics,*
arXiv:0901.0512 [hep-ex].

[92] ATLAS Collaboration. *Observation and measurement of higgs boson decays to WW^* with the ATLAS detector,* Phys. Rev. D, 92:012006, Jul 2015.

[93] ATLAS Collaboration, *Measurement of the $W^{\pm}Z$ boson pair-production cross section in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS Detector,* Phys. Lett. **B762** 1 (2016),
arXiv: 1606.04017 [hep-ex].

[94] G. Artoni et al., *Study of $\ell\ell$+MET with proton-proton collisions at $\sqrt{s}$ = 13 TeV in the HZZ group: introduction of common analysis strategies,* ATL-COM-PHYS-2016-1600, url:
https://cds.cern.ch/record/2231582.