BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

WWW **PRODUCTION: THE HUNT BEGINS (VERSION 1.2)**

by

BRIAN ALEXANDER LONG

B.S., The University of North Carolina, 2010
M.A., Boston University, 2015

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2016

Approved by

First Reader ____________________

John M. Butler, PhD
Professor of Physics

Second Reader ____________________

Kevin M. Black, PhD
Assistant Professor of Physics

Acknowledgments

blank

WWW PRODUCTION: THE HUNT BEGINS (VERSION 1.2)

(Order No.)

BRIAN ALEXANDER LONG

Boston University, Graduate School of Arts and Sciences, 2016

Major Professor: John M. Butler, Professor of Physics

ABSTRACT

In 2012 a resonance with a mass of 125 GeV resembling the elusive Higgs boson was discovered simultaneously by the ATLAS and CMS experiments using data collected from the Large Hadron Collider (LHC) at CERN. With more data from the LHC, the evidence continues to mount in favor of this being the Higgs boson of the Standard Model. This would finally confirm the mechanism for Spontaneous Electroweak Symmetry Breaking (EWSB) necessary for describing the mass structure of the electroweak gauge bosons. In 2013, Peter Higgs and Francois Englert were awarded the Nobel Prize in physics for their work in developing this theory of EWSB now referred to as the Higgs mechanism. The explanation for EWSB is often referred to as the last piece of the puzzle required to build a consistent theory of the Standard Model. But does that mean that there are no new surprises to be found? Many electroweak processes have yet to be measured and are just starting to become accessible with the data collected at the LHC. Indeed, this unexplored region of electroweak physics may provide clues to as of yet unknown new physics processes at even higher energy scales. Using the 2012 LHC data recorded by the ATLAS experiment, we seek to make the first observation of one such electroweak process, the massive tri-boson final state: WWW. It represents one of the first searches to probe the Standard Model WWWW coupling directly at a collider. This search looks specifically at the channel where each W boson decays to a charged lepton and a neutrino, offering the best sensitivity for making such a measurement. In addition to testing the Standard Model directly, we also use an effective field theory approach to test for the existence of

anomalous quartic gauge couplings which could offer evidence for new physics at higher energies than those produced by the LHC.

Contents

List of Tables

List of Figures

List of Symbols

aQGC	....	anomalous Quartic Gauge Coupling
ATLAS	....	A Toroidal LHC ApparatuS
DPS	....	Double Parton Scattering
EMCAL	....	Electromagnetic Calorimeter
EW	....	Electroweak
EWSB	....	Electroweak Symmetry Breaking
FCAL	....	Forward Calorimeter
HCAL	....	Hadronic Calorimeter
HEC	....	Hadronic End-cap Calorimeter
ID	....	Inner Detector
LAr	....	Liquid Argon
LHC	....	Large Hadron Collider
LO	....	Leading-Order
MC	....	Monte Carlo simulation
MS	....	Muon Spectrometer
NLO	....	Next-To-Leading-Order
QCD	....	Quantum Chromodynamics
QGC	....	Quartic Gauge Coupling
SFOS	....	Same-Flavor Opposite-Sign
SM	....	Standard Model
PDF	....	Parton Distribution Function

Chapter 1

Introduction

blank

Chapter 2

Theory

2.1 The Standard Model

The Standard Model (SM) is a theory which describes all of the observed matter and interactions in the universe, except for gravity. It is built from a quantum field theory where the consitutent particles and interactions fit into a non-abelian SU(3) $\times$ SU(2) $\times$ U(1) gauge symmetry. From these symmetries come the matter fermions, split into the quarks and leptons, and the force-carrying bosons that mediate their interactions. The SU(3) symmetry describes the theory of Quantum Chromodynamics (QCD) which explains the interaction of the quarks via the gluons, the gauge bosons that mediate the strong force. The remaining SU(2) $\times$ U(1) symmetry describes the Electroweak (EW) theory which explains the interactions of the quarks and leptons via the electroweak gauge bosons that mediate the electroweak force: W, Z, and γ (i.e. the photon). The EW theory is itself a unified description of the weak force, involving the W and Z, and the electromagnetic force, involving just the photon. The W and Z gauge bosons (as well as the quarks and leptons) receive their non-zero masses through the process of electroweak symmetry breaking (EWSB). The simplest form of EWSB introduces an additional 'Higgs' field that predicts a single new fundamental scalar boson. This boson is the famous Higgs boson which was discovered recently at the LHC, thereby confirming this last component of the SM.

All of the observed fundamental matter particles in the universe are described by the quark and leptons of the SM. Their properties are listed in Table 2.1. The particles can be distinguished by their charges and their masses. The charges describe how (and if) the particles participate in different interactions. Those fermions with electric charge (all but the neutrinos) participate in the electromagnetic interactions. The quarks have color charge (sometimes just called color), which allows them to participate in the QCD interactions. All fermions also participate in the weak interactions. The types of allowed weak interactions are determined by a combination of the electric charge as well as the weak isospin and weak hypercharge, described later. Since all of the particles are fermions, they all have a spin of 1/2. The masses of the particles are not predicted by the theory, but are essential for understanding their stability and decay properties as well as their kinematic behavior. Each particle also has a corresponding anti-particle with the same mass but whose electric charge has opposite sign. The neutrinos, with zero electric charge, could possibly be their own anti-particle (so-called Majorana fermions), but this has yet to be confirmed.

	Generation	Name	Symbol	Charge	Mass [MeV]
Quarks	First	Up	u	2/3	$2.3^{+0.7}_{-0.5}$
		Down	d	$-1/3$	$4.8^{+0.5}_{-0.3}$
	Second	Charm	c	2/3	1275 ± 25
		Strange	s	$-1/3$	95 ± 5
	Third	Top	t	2/3	173210 ± 874
		Bottom	b	$-1/3$	4180 ± 30
Leptons	First	Electron	e	-1	$0.510998928 \pm 0.000000011$
		Electron Neutrino	ν_e	0	< 0.002
	Second	Muon	μ	-1	$105.6583715 \pm 0.0000035$
		Muon Neutrino	ν_μ	0	< 0.19
	Third	Tau	τ	-1	1776.86 ± 0.12
		Tau Neutrino	ν_τ	0	< 18.2

Table 2.1: Summary of the electric charge and measured masses of the SM fermions. Mass measurements are taken from the Particle Data Group [51] and are shown to the best precision available with their measured uncertainties. Particles are also organized by their generation. The bottom quark mass measurement is shown using the $\overline{\mathrm{MS}}$ renormalization scheme. The top quark mass uncertainty combines the reported statistical and systematic uncertainties in quadrature. The limits on the electron neutrino and muon neutrino masses are set at a 90% confidence level while the tau neturino limits are set at a 95% confidence level.

The quarks and leptons can each be divided up into three "generations" composed of pairs of particles with identical charges but whose masses increase with each generation.

The generations are labeled in Table 2.1. In the leptonic sector, leptons only interact exclusively with leptons of their own generation. However, in the quark sector, while there is a strong preference for the quarks to interact within the same generation, it is still possible for quarks to interact with quarks of the other generations. This is described by the CKM matrix... Even though there are three generations in both the lepton and quark sectors, the quarks and leptons are not observed to interact directly, thus the quark and lepton generations should be thought of as separate.

The SM can be written down using a lagrangian of the form

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{QCD}} + \mathcal{L}_{\mathrm{EW}} + \mathcal{L}_{\mathrm{EWSB}} \tag{2.1}$$

which is gauge invariant or something... From this, one can calculate all of the fundamental interactions of the SM. As written, the SM lagrangian can be split up into separate terms describing the QCD, EW, and EWSB behavior. The EWSB term includes the behavior related to fermion mass generation. The details of each term are described in more detail below.

CP violation?

2.1.1 Quantum Chromodynamics

The QCD term in the SM lagrangian can be written as

2.1.2 Parton Distribution Functions

2.1.3 The Electroweak Theory

The EW term in the SM lagrangian can be written as

2.1.4 Electroweak Symmetry Breaking

The EWSB term in the SM lagrangian can be written as

2.1.5 The W Boson

Of most interest to the topic of this thesis is the behavior and properties of the W boson, the charged gauge boson of the EW theory. The W was first discovered in 1983 via $p\overline{p}$ collisions at SPS Synchroton by looking at its decay to an electron and electron neutrino (cite). Its mass has been measured in $p\overline{p}$ collisions at the Tevatron and in e^+e^- collisions at LEP to give a world average of 80.385 ± 0.015 GeV [51]. More recently, measurements of the mass have also been performed in pp collisions at the LHC have been shown to be consistent with __ reported from ATLAS (cite) and ___ from CMS (cite). The width assuming a Breit-Wigner distribution has also been measured at LEP and the Tevatron with an average value of 2.085 ± 0.042 GeV [51].

refer to isospin and hypercharge? lepton universality?

The W decays into quarks roughly 2/3 of the time. The remaining third of the time the W decays approximately evenly into each of the three lepton generations. The measured branching fractions are summarized in Table 2.2. The leptonic decays of the W result in a charged lepton with the same charge as the parent W (as dictated by charge conservation) and a neutrino (or anti-neutrino if the parent W has negative charge). Thus, the allowed leptonic decays are of the form:

$$W^- \rightarrow l^- \overline{\nu}_l$$

$$W^+ \rightarrow l^+ \nu_l$$

blah

Table 2.2: Measured branching fractions of the W boson as reported by the Particle Data Group [51]

maybe show W mass plot from PDG or distribution from ATLAS?

The CKM...

2.2 Signal

2.2.1 WWW Signal

2.2.2 Effective Field Theory

2.2.2.1 Anomalous Quartic Gauge Couplings

A measurement of the production rate can be used to probe the gauge couplings, in particular, the process is sensitive to quartic gauge couplings. The VBFNLO code has implemented a list of higher order operators that parameterize the effects of new physics at energy scale beyond the reach of current collider experiments. The effective field theory approach is practical and widely used when there is no compelling specific model of new physics beyond the SM, see for example, discussions in Refs: [44], [29] and [36]. As a benchmark, we choose two gauge invariant dimension-8 operators:

$$\mathcal{L}_{s,0} = [(\mathrm{D}_\mu \phi)^\dagger \mathrm{D}_\nu \phi] \times [(\mathrm{D}^\mu \phi)^\dagger \mathrm{D}^\nu \phi] \tag{2.2}$$

$$\mathcal{L}_{s,1} = [(\mathrm{D}_\mu \phi)^\dagger \mathrm{D}^\mu \phi] \times [(\mathrm{D}_\nu \phi)^\dagger \mathrm{D}^\nu \phi] \tag{2.3}$$

where ϕ is the Higgs field doublet, and D_μ is the covariant derivative. The Lagrangian of the effective field theory is thus:

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \frac{f_{s0}}{\Lambda^4}\mathcal{L}_{s,0} + \frac{f_{s1}}{\Lambda^4}\mathcal{L}_{s,1} \tag{2.4}$$

The choice of the two operators is introduced to benchmark possible deviations from the SM. If a significant excess of events is observed in the data, the parameterization will be changed, to incorporate more operators, to investigate the nature of the observed new physics.

Chapter 3

Collider Physics and The Large Hadron Collider

blank

Chapter 4

The ATLAS Detector

The ATLAS detector [18] is designed to measure the products of the particle collisions produced by the LHC. In particular, the detector seeks to measure those stable (or meta-stable) particles whose decay lifetime is sufficiently long enough to interact with the detector. This includes a variety of fundamental particles (like muons) as well as composite particles (like neutrons). The wide variety of particles to be measured requires the implementation of several sub-detector systems that work in tandem to identify and measure the properties of these particles. The products of the LHC collisions travel in all directions, thus if there is any hope to measure all of the products of the collision the detector must completely surround the collision point. The symmetry of the LHC dictates a cylindrically symmetric shape for the ATLAS detector around this point. There is another reason to build the detector to enclose the collision point, since even with such a detector, the vanishingly small cross-section of the neutrinos, which are also products of the collision, means that these ghostly particles will inevitably go undetected. But there is still hope! If all other particles are measured, the presence of the neutrinos can be inferred using momentum constraints, as will be discussed later. In this way, the ATLAS detector is most generally suited for measuring the wide variety of processes produced by the LHC.

A diagram of the ATLAS detector can be seen in Fig. 4.1. It has a clearly cylindrical shape with a diameter of 25 meters and length of 44 meters. The detector is massive, weighing in at roughly 7000 tons; but it is also highly granular, with over 100 million electronic channels that are arranged very precisely, in many cases on the order of tens of microns. In the “opened” view of Fig. 4.1, the proton-proton collisions from the LHC

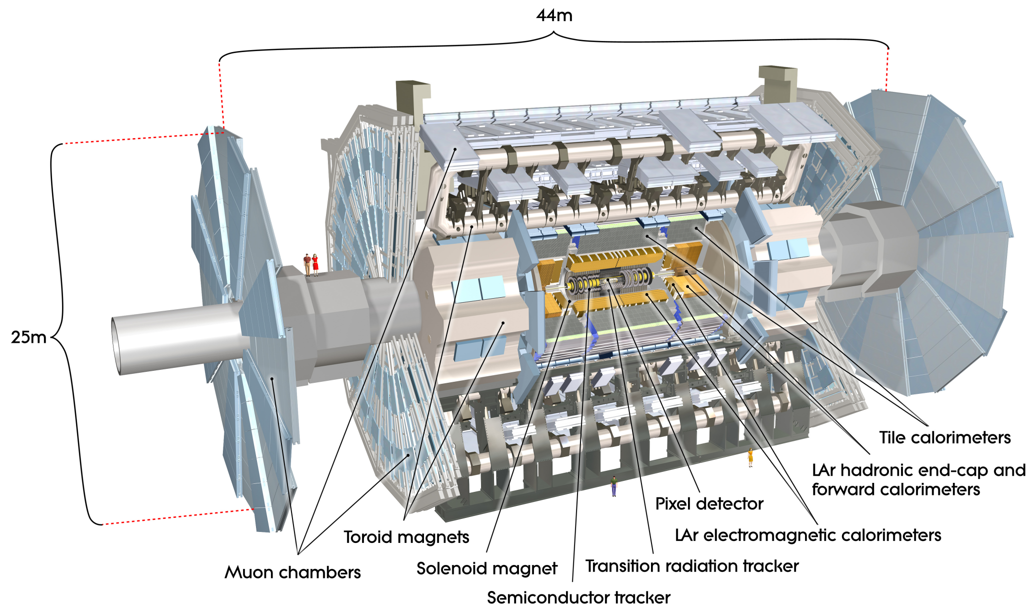

Figure 4.1: A diagram of the ATLAS detector where the detector has been artificially opened up to reveal the LHC beam line and the various sub-detector components within. The sub-detector components are labeled as such.

occur at the center core of the detector and the sub-detector components build up around this point. The detectable products of the collision pass outward from the collision point through the different components where their energy and momentum are measured. The way in which the particles interact with the various sub-detector systems helps to identify the types of particles produced. This can be more clearly seen in the diagram of Fig. 4.2, which shows how the most predominant products of the LHC collisions interact with the different components of the ATLAS detector. Nearest the collision point is the inner detector (ID), designed to measure the paths of charged particles passing through using several different subsystems. This is surrounded by a 2 Tesla solenoidal magnetic field which bends the trajectory of the charged particles by an amount inversely proportional to the particle momentum, thereby allowing for a momentum measurement. Beyond that is the calorimeter system which measures the energy deposits of all particles passing through (except for muons and neutrinos) by stopping them in their tracks. The calorimeter system itself is divided up into components which fall into two main categories: the electromagnetic (EMCAL) and hadronic calorimeter (HCAL) systems. The EMCAL is situated in front of the HCAL and is designed primarily to stop fundamental particles like electrons and photons. The HCAL then is designed to stop composite particles like protons and neutrons through. Surrounding the calorimeter system is the muon spectrometer (MS), which is the largest component of the ATLAS detector and the one that determines its size. It is designed to measure the trajectory of muons through ionization as they pass through and ultimately leave the detector. The MS is also composed of three large superconducting air-core toroid magnets with an average magnetic field of roughly 0.5 Tesla which allows for a measurement of the muon momentum. The neutrinos pass through undetected.

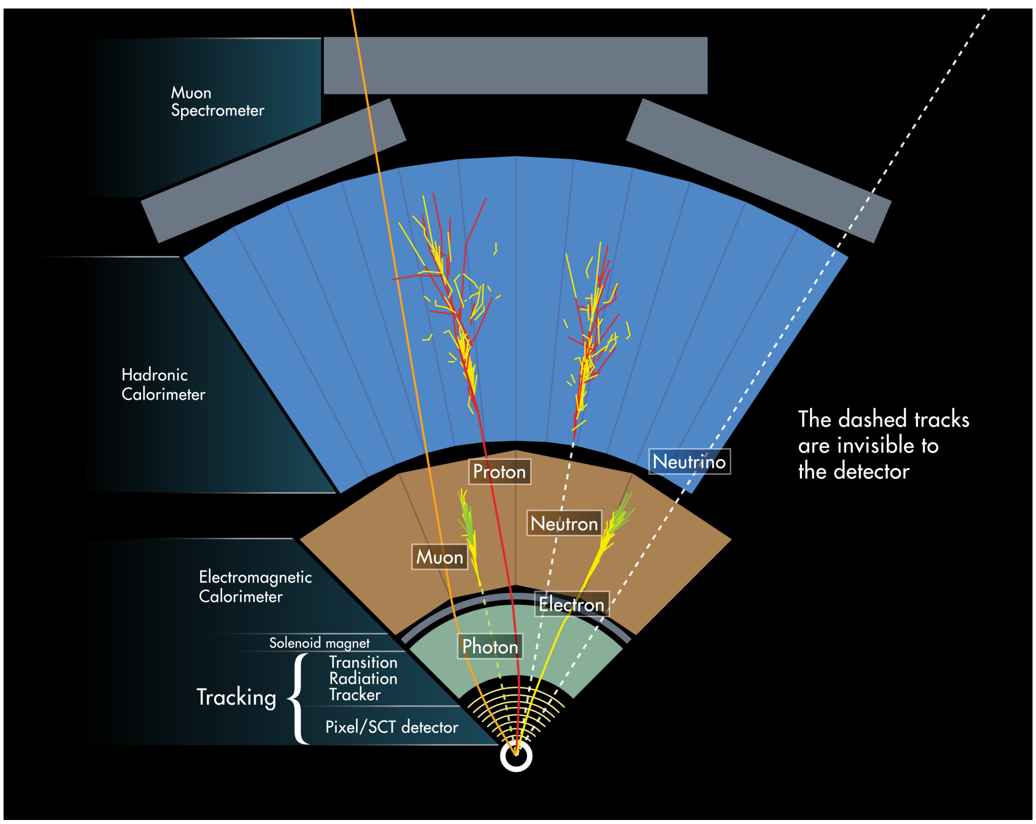

Figure 4.2: A diagram of one wedge of the ATLAS detector as viewed from looking down the beam line. The sub-detector components are shown along with the particles that most predominately come from the collision. The paths of the particles indicate how each particle typically interacts with the detector.

The geometry of the ATLAS detector is defined using a right-handed cylindrical coordinate system with the x-axis pointing inwards towards the center of the LHC ring, the y-axis point up, and the z-axis pointing along the beam-line. The x vs y plane, which is perpendicular to the beam-line, is referred to as the transverse plane and typically defined using cylindrical coordinates with r being the distance from the center and r being the azimuthal angle. For describing the direction of the particle with respect to the beam-line, a quantity called the pseudo-rapidity, η, is commonly used

$$\eta = -\ln\tan(\theta/2) \tag{4.1}$$

which is a function only of the polar angle, θ, which itself is defined as the direction of the particle with respect to the positive z-axis. Changes in the pseudo-rapidity are invariant under Lorentz transformations along the z-axis in the limit that the particle is approximately massless. As a result, its distribution is typically flat over a wide range of pseudo-rapidity. At the LHC, most stable particles are produced with energies much larger than their mass, making the massless approximation valid. The ATLAS detector has nearly uniform 2π coverage in ϕ, while in η the ID is restricted to $|\eta| < 2.5$, the MS to $|\eta| < 2.7$, and the calorimeter systems all the way to $|\eta| < 4.9$.

The total momentum of the proton-proton collision in the transverse plane is zero. Since the detector has full azimuthal coverage in the transverse plane, we can test this constraint by measuring the total momentum from the particles measured in the detector. The magnitude of a particle's momentum in the transverse plane is referred to as the transverse momentum, p_{T}. Thus, we may refer to this constraint as

$$\left| \sum_{i \in \text{All Particles}} \vec{p}_{\mathrm{T},i} \right| = 0 \tag{4.2}$$

where the transverse momentum is added vectorialy and then the magnitude is taken. After adding up the p_{T} of all of the particles to obtain the total transverse momentum, any imbalance with respect to this constraint is referred to as the missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, and is attributed to the neutrinos produced in the collision. There is no such constraint on the momentum along the z-direction because the momentum fraction of the partons as taken from the PDFs are not known with certainty. This is the case even though the momentum along the z-direction of the protons from which the partons are taken is, in fact, known. Thus there is no way of determining with certainty the momentum of the neutrinos in the z-direction.

4.1 Inner Detector

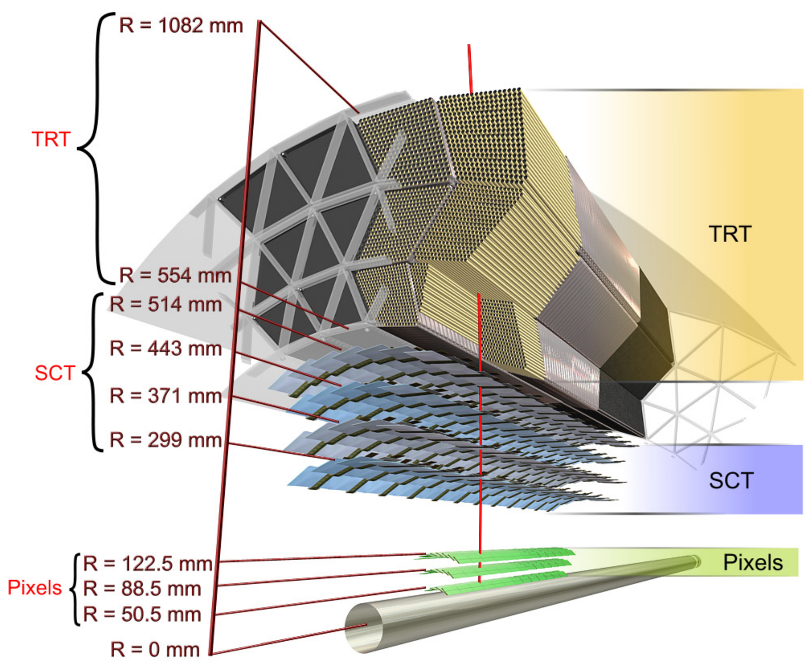

Figure 4.3: Diagram of the ATLAS Inner Detector (ID) system showing a wedge of the barrel system. The three detector systems are clearly labeled. The LHC beam pipe runs parallel to the system and is shown at the bottom of the diagram.

The inner detector (ID) is the detector system that is closest to the beam pipe and thus first system that the products of the LHC collisions encounter on their way from the collision point. Its primary role is to measure the trajectory and momentum of charged particles through ionization as they pass through the detector. It must be capable of measuring

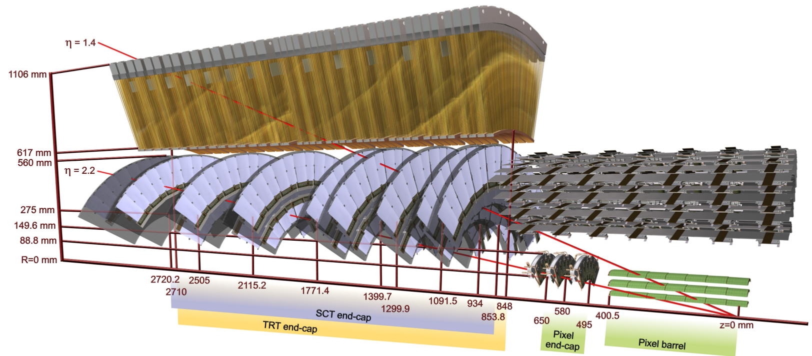

Figure 4.4: Diagram of the ATLAS Inner Detector (ID) system showing a wedge of the endcap system as well as a part of the SCT and Pixel barrel systems. The detector systems are clearly labeled. The LHC beam pipe runs parallel to the system but is not shown. Trajectories of two charged tracks with a $p_{\mathrm{T}} = 10$ GeV are shown along $\eta = 1.4$ and $\eta = 2.2$ are shown by the solid bright red lines.

these tracks with high precision in order to obtain precise momentum measurements and also to be able to accurately extrapolate the tracks back to the collision point to obtain primary and secondary interaction vertices. In addition, since the system is so close to the LHC beam line, it must be able to handle the high particle fluxes. This requires that the ID must have a very high granularity and fast electronics readouts such that the occupancy of the detector is small enough to distinguish individual tracks. In addition, the detector materials and electronics must be sufficiently radiation hard that they can withstand years of LHC exposure time. These tough requirements push the limits of available technology and thus make the ID the most sophisticated detector system in ATLAS.

There are three different detector subsystems within the ID, together immersed in a roughly uniform 2 Tesla axial magnetic field: the pixel detector, the silicon microstrip (SCT) detector, and the transition radiation tracker (TRT). These three detector systems can be seen in the barrel in Fig. 4.3 and from an alternate view also showing one of the end-caps in Fig. 4.4. The pixel detector is composed of more than seventeen hundred thin doped silicon sensors with dimension 19×64 mm^2. Each sensor has more than forty-six thousand readout channels, corresponding to the "pixels" which give the detector its name. A charged particle passing through an individual pixel can send a signal which identifies the

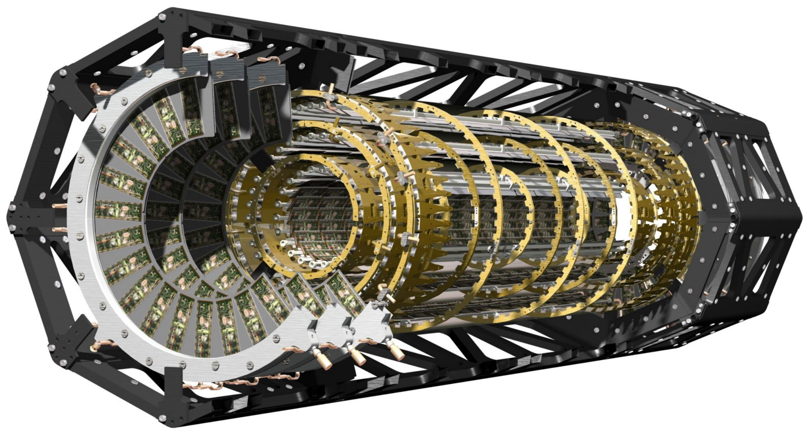

Figure 4.5: A cut-out diagram of the ATLAS pixel detector showing the arrangement of the pixel modules (green) in three layers of the barrel and three layers of one end-cap system. Some of the support structure is also shown.

location of a charged particle. The combination of several layers can thus be used to form the trajectory of the particle. Each sensor is attached to a single readout electronic board, which comprises one module. The modules are arranged into three cylindrical barrel layers and two end-caps each with three disk-shaped layers such that there is uniform azimuthal coverage. A cut-out diagram of the pixel detector structure with modules in place in both the barrel and end-caps is shown in Fig. 4.5. The barrel covers roughly $|\eta| < 1.7$ and the two end-caps roughly $1.7 < |\eta| < 2.5$. Test beam measurements show that the spatial resolution of the pixel detector is around 12 μm in the $R-\phi$ plane and is slightly degraded orthogonal to this plane.

The SCT uses almost sixteen thousand thin silicon strip sensors, though not of the same type as in the pixel detector. A barrel silicon strip sensor has dimension 6.36×6.40 cm^2 with 768 readout strips running along the longer dimension. The barrel strips are placed in four concentric cylindrical layers, uniformly in azimuth, with the strips aligned axially, and covering roughly $|\eta| < 1.4$, as can be seen in Fig. 4.3. In each of the two end-caps the sensors are made to form nine disks spaced apart along the axial direction, covering roughly $1.4 < |\eta| < 2.5$, seen in Fig. 4.4. The strips are similar to the barrel except that they are tapered along the strip direction. The sensors are then oriented such that the taper expands radially outward. In a test beam, the spatial resolution is found to be about 16 μm in the $R - \phi$ plane. Due to the length of the strips, the precision is considerably worse in the axial direction for the barrel and the radial direction for the end-caps, with a precision of roughly 580 μm.

The TRT uses a fundamentally different technology than the pixel and SCT. Drift tubes are used of 4 mm in diameter which are filled with a Xenon-based gas mixture and with an anode wire running through the center. The advantage of using these tubes is that they can be placed in close proximity such that many measurements, around 36, can be made on a single charged track. Another important feature of the TRT is its ability to identify electrons using transition radiation. This is achieved by surrounding the tubes in polypropylene material to induce transition radiation from incident electrons and taking advantage of the discrimination power of the Xenon-based gas between transition radiation and tracking signals. For electrons with $p_{\mathrm{T}} > 2$ GeV, usually 7 to 10 hits due to transition radiation will be measured. The barrel TRT runs from roughly $|\eta| < 0.7$ and is constructed from 144 cm long straws with two straws aligned axially back-to-back. Over fifty-two thousand straws are interleaved with polypropylene fibers to form 73 layers of straws surrounding the beam-pipe with a cylindrical symmetry and uniform coverage in azimuth, seen in Fig. 4.3. In each of the two end-caps, two wheels are formed from over seventy-three thousand straw tubes, 37 cm in length, oriented oriented and distributed uniformly in azimuth. The inner wheel is formed from twelve layers and the outer wheel from eight layers with 768 straws in each layer, seen in Fig. 4.4. The end-caps cover roughly $0.7 < |\eta| < 2.2$. An individual straw has a precision of about 170 μm along its diameter.

efficiency of electron identification?

large number of readouts? occupancy?? noise?? sagitta??? explain how the b-field is used? eta coverage? materials? electronics? more figures?

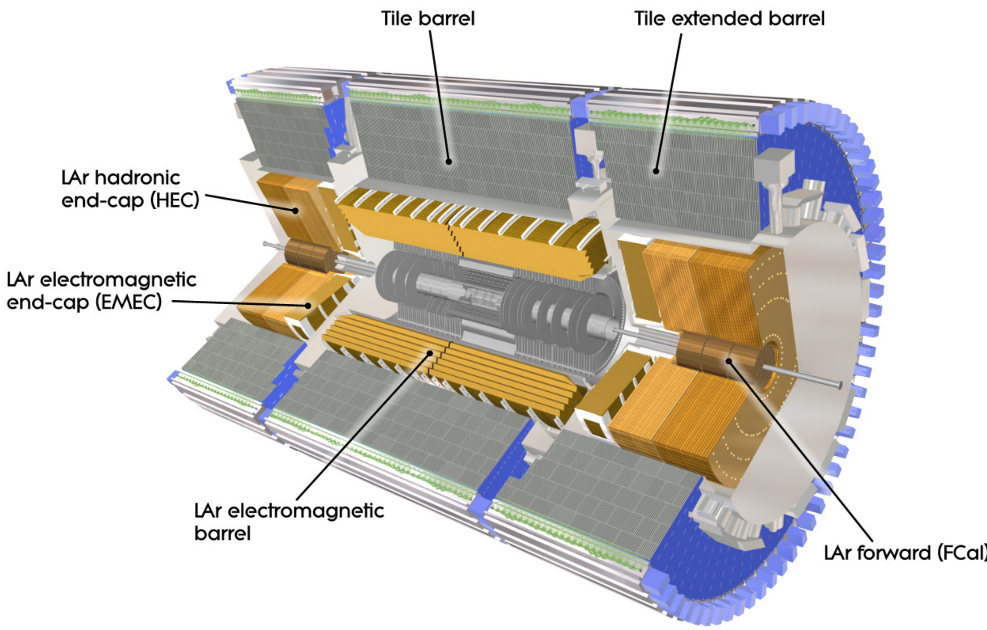

Figure 4.6: Diagram of ATLAS calorimeter system with cut-out portion to allow a view of the nested sub-components.

4.2 Calorimeters

The ATLAS calorimeter is designed to measure the energy deposits of the products of the LHC collisions which pass through it except for muons and neutrinos. A diagram of the calorimeter system can be seen in Fig. 4.6.

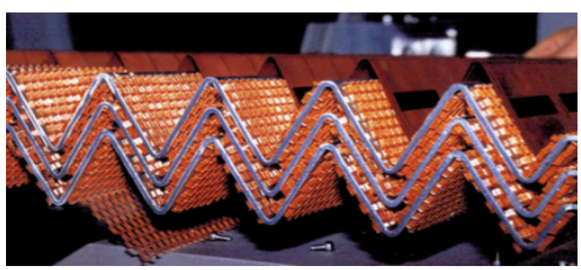

Figure 4.7: Photo of three EMCAL sampling layers showing the “accordion” structure. In the picture, the horizontal directions corresponds to the radial direction when the detector is in position, which is the direction the LHC products would follow.

The calorimeter system is split into three main systems, the electromagnetic calorimeter

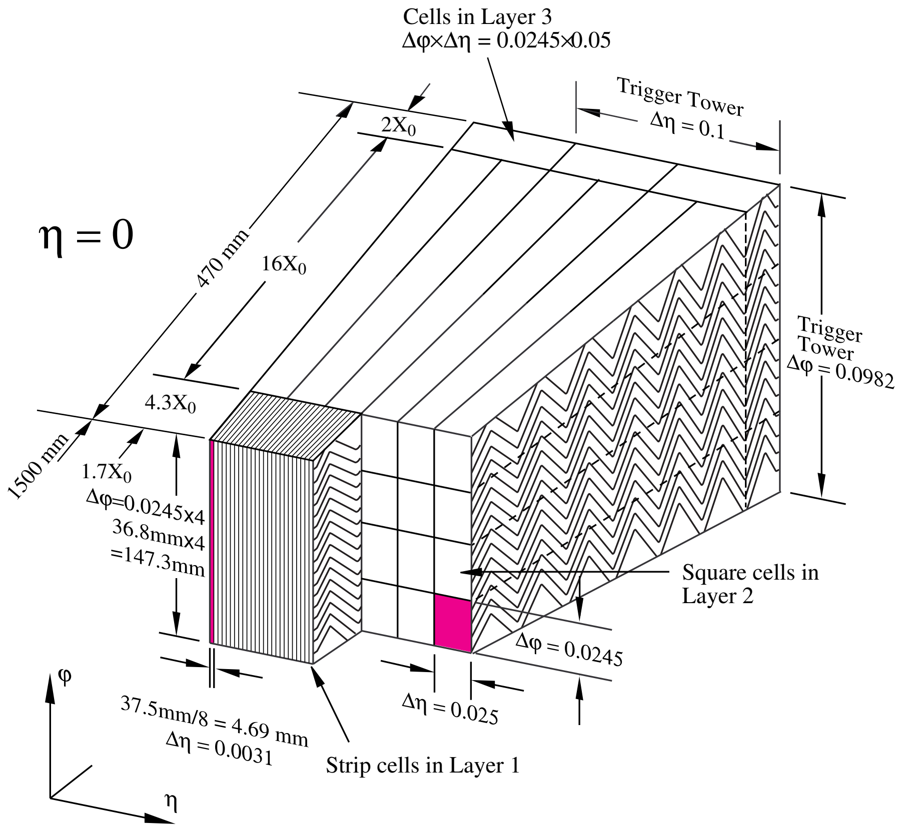

Figure 4.8: A diagram of one EMCAL barrel module covering 22.5° in azimuth.

(EMCAL), the tile hadronic calorimeter (HCAL), the hadronic end-cap calorimeter (HEC), and the Forward Calorimeter (FCAL). The EMCAL is a sampling calorimeter that uses lead as the sampling medium and liquid Argon (LAr) as the active medium from which the charge of the electromagnetic shower produced by incident particles on the sampling medium can be collected. LAr is used as the active medium because of its radiation hardness and its linear response. The lead sampling medium alternates with the active LAr medium using lead plates 1-2 mm thick with an approximately 4 mm LAr gap between each sheet and electrodes placed in the middle of the gaps. The lead sheets are constructed using a unique "accordion" structure, as seen in Fig. 4.7. This is to provide a uniform resolution with no gaps in the azimuthal direction. The EMCAL itself can be split up into a barrel region ranging from $0 < |\eta| < 1.3$ and two end-cap regions ranging from

$1.5 < |\eta| < 3.2$. The thickness of the barrel region has a radiation length, X_0, ranging from 22 X_0 to 30 X_0 for $|\eta| < 0.8$ and from 24 X_0 to 33 X_0 for $0.8 < |\eta| < 1.3$. The barrel region is divided into individual modules which together surround the beam-line in a cylindrical shape. A diagram of one such module can be seen in Fig. 4.8. From this one can see that each module is segmented in η and ϕ, as well as into three layers in depth. The segmentation is applied to obtain pointing information, which aids in the identification and measurement of electromagnetic objects in conjunction with measurements from the ID. The very fine segmentation in η of the first layer in depth is important for precision tracking measurements. The second layer has a larger depth and thus collects most of the energy. There are two identical end-cap regions, one on each side of the collision point. Each end-cap region consists of two wheels: the outer wheel from $1.4756 < |\eta| < 2.5$, with a thickness ranging from 24 X_0 to 38 X_0, and the inner wheel from $2.5 < |\eta| < 3.2$, with a thickness ranging from 26 X_0 to 36 X_0. The regions from $1.5 < |\eta| < 2.5$ in the inner and outer wheels both have three layers, with the first being a finely segmented precision layer similar to the barrel regions. Outside this region there are only two layers with a coarser segmentation. The EMCAL also consists of a pre-sampler detector with a single layer of LAr in front of the full barrel EMCAL and in front of the end-cap EMCAL calorimeters from $1.5 < |\eta| < 1.8$; this aids in the measurement of the energy deposits prior to reaching the EMCAL and for a better understanding of the energy deposited in the transition region between the barrel and end-caps.

The tile HCAL is a steel sampling calorimeter with scintillating tiles used as the active material. Steel is chosen as the sampling material since it gives a good depth in interaction lengths, λ, with a maximum depth of 7.4 λ, while also having a low cost. It is split into a central barrel and two extended barrels which together cover a region from $|\eta| < 1.7$, as can be seen in Fig. 4.6. Same as in the EMCAL barrel, the tile HCAL is divided into individual modules that surround the collision point in azimuth; a diagram of one such module is shown in Fig. 4.9. The scintillating tiles alternate periodically with the self-supporting steel body and are oriented radially. The scintillation light is routed through

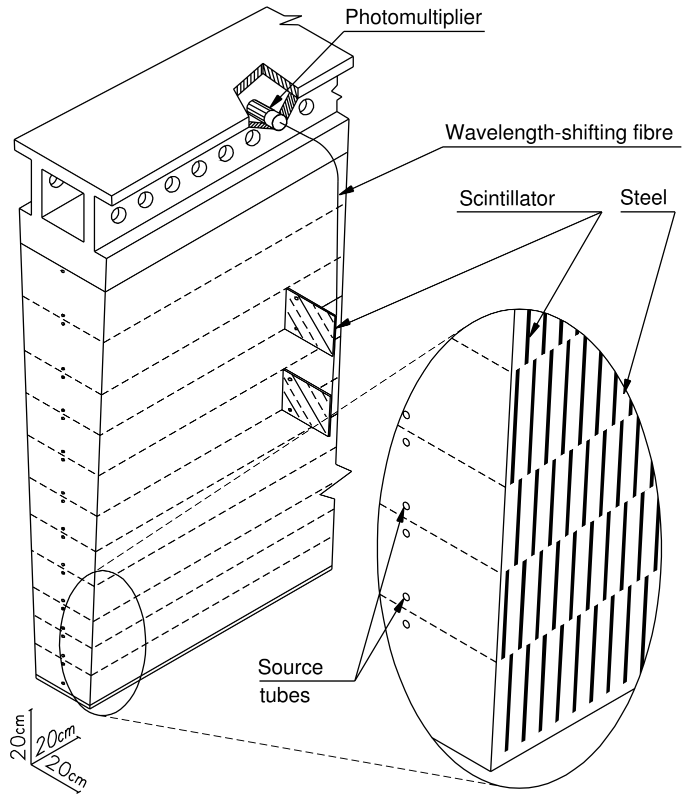

Figure 4.9: A diagram of one tile HCAL module covering 5.625° in azimuth. The radial direction when positioned in the detector corresponds to the vertical direction in the image.

wavelength-shifting fibers and collected at photo-multiplier tubes placed at the back of the module. This configuration allows for a near uniform coverage in azimuth. In the crack region from $1.2 < |\eta| < 1.6$ between the central barrel and extended barrels, special modules are placed to recover and correct for energy losses in this region.

The HEC is designed to measure hadronic energy deposits in the end-cap regions from $1.5 < |\eta| < 3.2$. It uses copper plates as the sampling material with LAr gaps for the active material. Two separate wheels are formed from flat plates of copper alternating with LAr gaps further divided by electrodes for collecting the ionization charge from the hadronic shower in the LAr. The rear wheel is more coarse than front wheel; This can be seen in the schematic of Fig. 4.10. The electronics readout is segmented such that pointing information can be obtained, as indicated by the dashed lines. The maximum radial depth

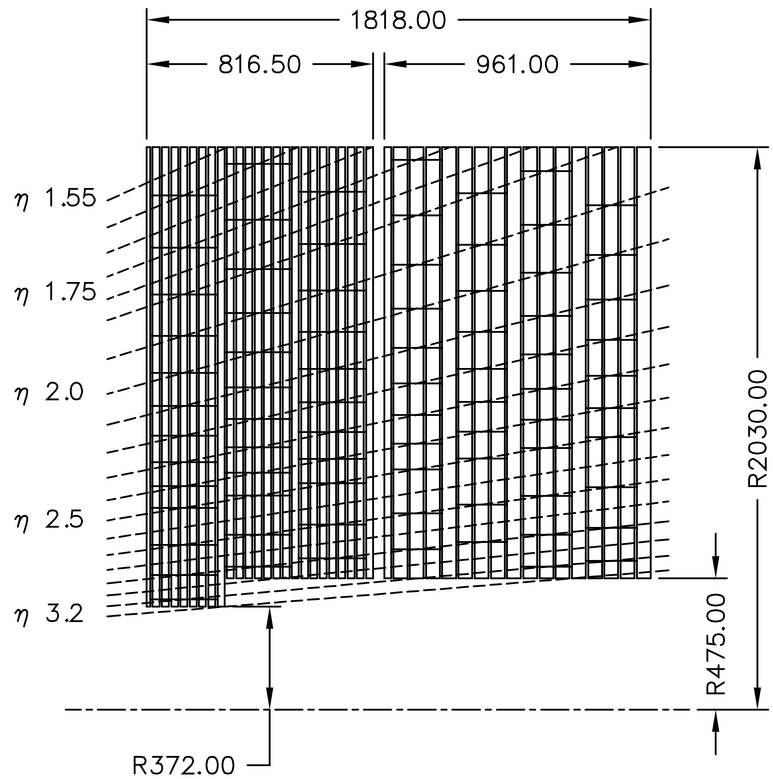

Figure 4.10: A schematic showing one quadrant of the HEC system in the R-z plane. The dashed lines indicate the pointing direction achieved by the segmentation of the readouts. Dimensions are in mm.

of the HEC is roughly 10 λ.

The FCAL is in the region of the detector nearest to the beam-line, where the radiation flux is highest, covering the range from $3.1 < |\eta| < 4.9$. It is split into three cylindrical modules, oriented as in Fig. 4.11, with the first being designed for measuring electromagnetic deposits and the other two for hadronic deposits. Each FCAL module is constructed from copper plates with roughly ten thousand uniformly spaced holes drilled in the direction parallel to the beam-line. The holes are filled with rods serving as the primary sampling material, with a thin LAr gap surrounding the rods serving as the active material. In the first FCAL, optimized for electromagnetic deposits, copper rods are used, while in the second and third FCAL modules, the rods are made from tungsten, which has a higher interaction length. The first FCAL has a radiation length of 27.6 X_0 and an

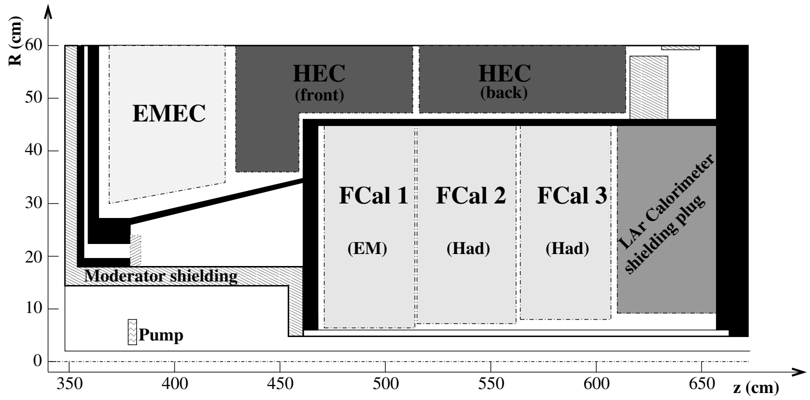

Figure 4.11: A schematic showing the end-cap of the EMCAL, the two HEC modules, and the three FCAL modules, as well as additional shielding, in one quadrant of the ATLAS detector as viewed in the R-z plane. The R-direction is shown with a larger scale than in the Z-direction.

interaction length of 2.66 λ. Meanwhile, the interaction length of the second and third modules is around 3.6 λ.

Shielding? Resolution and response?

4.3 Muon Spectrometer

4.4 Trigger

Chapter 5

The first search for $WWW \to \ell\nu\ \ell\nu\ \ell\nu$

The first measurement of the WWW production process is sought by using a dataset containing 20.3 fb^{-1} of integrated luminosity collected from the LHC at an energy of $\sqrt{s} = 8$ TeV in 2012. In addition to being the first study of this particular process, it is also the first study to search for a final state with more than two massive gauge bosons, and one of the first studies to search for aQGCs. The total cross-section for this process is expected to be roughly 224 femtobarns, as determined using MadGraph [13]. If measured, it would be one of the smallest cross-section measurements within ATLAS. For this search, the WWW process is studied in the so-called "fully leptonic" decay channel where each W boson decays leptonically (excluding τ lepton decays). As can be seen in Fig. 5.1, this decay channel occurs only about 1% of the time; the rest of the time at least one of the W bosons decays hadronically. While the branching fraction is small, this channel should have a smaller background than those that include hadronic W decays. As a result, the fully leptonic channel is one of the most sensible channels for obtaining sensitivity to this process.

The data is studied in a region where the signal is most prominent with respect to the background. This region is primarily characterized by having three high p_{T} leptons (e or μ), with additional requirements determined using an optimization procedure. To understand the data in this region we must model both the signal and the backgrounds that fall into it. The signal is modeled purely using Monte Carlo (MC) simulation. The backgrounds, however, are modeled using a combination of MC simulation and data-driven techniques. Prior to the measurement, each important background is studied in control regions which

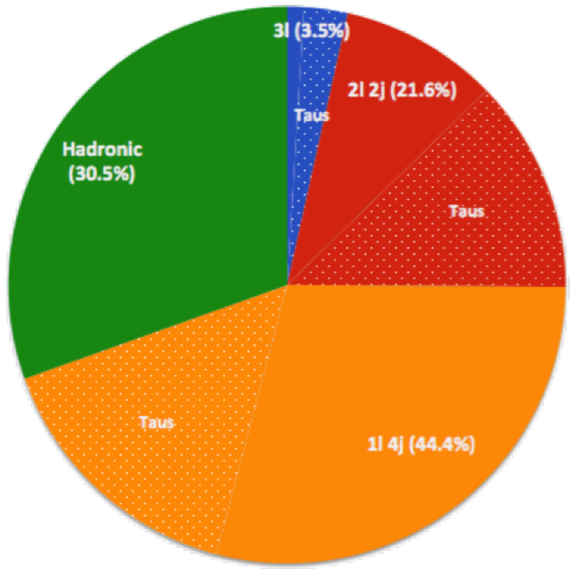

Figure 5.1: Pie chart showing the different decay modes contributing to the total cross-section for the WWW process. The dotted areas indicate the portion of each decay mode which is due to the production of tau leptons.

are either orthogonal to the signal region selection or where the signal is suppressed. This is to ensure that all backgrounds are described accurately. The agreement of the data with the signal plus background prediction is determined using a "cut-and-count" approach where the total number of data events observed in the signal regions is compared to the expected number of events from the model. A fit to the data is performed using a profile likelihood with the relative normalization of the signal as the parameter of interest and with statistical and systematic uncertainties treated as nuisance parameters. From this fit, the measured signal cross-section and uncertainty, the sensitivity of the data to the signal under the background only hypothesis, and limits on new physics in an effective field theory are extracted.

5.1 Data and Simulation Samples

5.1.1 Data

This analysis is based on the study of the full proton-proton collision data from the LHC in 2012. After quality requirements, the amount of data used in this analysis corresponds to an integrated luminosity of 20.3 fb^{-1}. The uncertainty on the integrated luminosity is 2.8% following the same methodology as in [8]. The data are selected after requiring that at least one of a series of single lepton triggers passed during data taking, specifically, one of the following: either an electron trigger requiring at least one isolated electron with $p_{\mathrm{T}} > 24$ GeV , an electron trigger requiring at least one (possibly non-isolated) electron with $p_{\mathrm{T}} > 60$ GeV, a muon trigger requiring at least one isolated muon with $p_{\mathrm{T}} > 24$ GeV, or a muon trigger requiring at least one (possibly non-isolated) muon with $p_{\mathrm{T}} > 36$ GeV.

5.1.2 Simulation samples

An important tool for the modeling of physics processes that are/could be produced at the LHC is Monte Carlo simulation (MC). MC relies on random sampling to connect the matrix element formulations derived from quantum mechanical perturbation theory into actual predictions for the results of proton-proton collisions at the LHC. The prediction of a single collision from the MC represents one possible outcome of the proton-proton collision, with all of the products of the hard-scattering and their four-momenta. This result can be passed through additional MC simulation to describe hadronization and the soft products of the collision e.g. photon radiation. Finally, these products are passed through a detailed simulation of the ATLAS detector built in GEANT4 [10] so that the same reconstruction algorithms can be applied as in the data. This sampling is repeated many times to populate the distribution of possible outcomes. Dedicated MC programs are provided by theorists for different processes and to different orders in perturbation theory, sometimes with different treatments. Details of the different processes simulated from MC and their treatment are presented below.

5.1.2.1 Signal Processes

The signal processes studied in this analysis are $pp \to W^+W^+W^- + X$ and $pp \to W^+W^-W^-$ $+ X$, where X is intended to refer to the fact that no requirements are placed on additional particles produced in the hard interaction. The process includes associated Higgs production, or "Higgsstrahlung", where a W boson radiates a Higgs boson, $pp \to WH$, and subsequently decays into a W^+W^- pair. The Higgs decay results in one W boson being produced off-shell, $H \to WW^*$, making this the leading contribution to off-shell production. The resonance from the Higgs can clearly be seen from the distribution of $m_{W^+W^-}$ taken from simulation of WWW events in Fig. 5.2. The WWW process also includes contributions from the $WWWW$ quartic coupling.

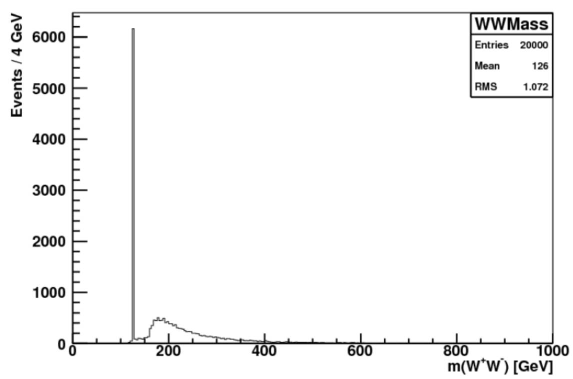

Figure 5.2: Invariant mass distribution of two opposite-sign W bosons in WWWevents generated with VBFNLO at LO. The Higgs mass peak is clearly visible at 126 GeV.

The SM signal processes are implemented in the Monte Carlo generator VBFNLO [16, 17], which can generate partonic events at leading-order (LO) with next-to-leading-order (NLO) cross-sections, and in MadGraph [13], which can generate partonic events at NLO with NLO cross-sections. The partonic events are further processed by Pythia8 [55] and Photos [41] to add effects of beam remnant interactions and initial and final state radiation. SM parameters must be provided to the MC generators as input. The most relevant input parameters are listed for the generators in Table 5.1. The parameters are set in Pythia8 using the ATLAS tune of AU2[6]. The MC generators must also be provided an appropriate PDF. The PDF used in the LO VBFNLO generation is the LO CTEQ6L1 [52] PDF set; CT10 NLO [43] is used in the NLO VBFNLO cross-section calculation. The PDF used in the NLO MadGraph generation and cross-section calculation is CTEQ6L1 but this is re-weighted to CT10 NLO using a k-factor. Since the MC generators are computed to finite order in perturbation theory, renormalization and factorization scales must be chosen. The renormalization and factorization scales are dynamically set to the WWW invariant mass in the VBFNLO samples; they are set to a fixed scale equal to the Z mass in MadGraph. The VBFNLO samples are restricted to leptonic decays of the W bosons where each lepton has a p_{T} of at least 5 GeV. The MadGraph samples include all decays of the W boson, with a requirement that jets have a a p_{T} of at least 10 GeV but with no requirement on the p_{T} of leptons. The VBFNLO and MadGraph samples handle interference between $WH \to WWW(*)$ and on-shell WWW production at LO, but MadGraph is not able to do this at NLO. As a result, the NLO MadGraph samples are split by on-shell WWW and $WH \to WWW(*)$ production. Both sets of samples are further split by the WWW charge mode. For each sample, the cross-sections are summarized in Table 5.2 in their full phase space and in a common fiducial phase space defined in Sec. 5.3.3. The fiducial cross-sections are observed to be nearly the same between the two generators, as expected. This serves as a good check of the understanding of the signal process. The MadGraph cross-sections are used throughout the remainder of the analysis.

	VBFNLO	MadGraph
Higgs mass, m_H	126.0 GeV	125.0 GeV
Top mass, m_t	172.4 GeV	172.5 GeV
Z mass, m_Z	91.1876 GeV	91.188 GeV
W mass, m_W	80.398 GeV	80.399 GeV
Fermi constant, G_F	1.16637×10^{-5} GeV^{-2}	1.16637×10^{-5} GeV^{-2}

Table 5.1: List of the most relevant SM parameters used as input to the signal MC generation.

The uncertainty on the PDF is derived for the MadGraph cross-sections following a

Sample		Cross-section [fb]	
		Inclusive	Fiducial
VBFNLO	$W^+W^+W^- \to l\nu l\nu l\nu$	4.95 ± 0.007	0.2050 ± 0.0070
	$W^-W^+W^- \to l\nu l\nu l\nu$	2.65 ± 0.004	0.0987 ± 0.0037
	Sum	7.60 ± 0.008	0.3037 ± 0.0072
MadGraph	$W^+W^-W^+ \to$ Anything	59.47 ± 0.11	0.0900 ± 0.0048
	$W^-W^+W^- \to$ Anything	28.069 ± 0.076	0.0476 ± 0.0043
	$W^+H \to W^+W^+W^-(*) \to$ Anything	99.106 ± 0.019	0.1114 ± 0.0029
	$W^-H \to W^-W^+W^-(*) \to$ Anything	54.804 ± 0.010	0.0603 ± 0.0015
	Sum	241.47 ± 0.13	0.3092 ± 0.0072

Table 5.2: Inclusive and common fiducial cross-sections at NLO for VBFNLO and MadGraph samples. The sum of the inclusive cross-sections are different because of the different branching fractions in the two cases. The sum of the fiducial cross-sections , however, are expected to be similar because they are computed for the same phase space, as described in Sec. ...

	PDF Uncertainty	
	$W^+W^+W^-$	$W^+W^-W^-$
Total	$+2.58\%\ -2.51\%$	$+8.69\%\ -3.47\%$
Fiducial	$+3.64\%\ -3.00\%$	$+7.57\%\ -3.08\%$

Table 5.3: Summary of PDF uncertainties estimated on NLO MadGraph cross-sections in both the fiducial and total phase space.

modified version of the pdf4lhc [28] recommendations. The resulting uncertainty is shown separately for the two different charge modes in both the fiducial and the inclusive phase space in Table 5.3. The uncertainty is determined by comparing three different PDFs; CT10 NLO [45], MSTW2008 NLO [48], and NNPDF 3.0 NLO [24]. This comparison is presented in Figure 5.3. Symmetric 68% CL uncertainties are determined for CT10 NLO and MSTW 2008 NLO using the 68% CL set provided for MSTW directly and the 90%CL set for CT10 after scaling down by a factor of 1.645. The uncertainty of the NNPDF 3.0 NLO PDF set is determined by using the standard deviation of the distribution of 101 MC PDFs provided in the PDF set; the nominal value is taken from the mean of the same PDFs. The CT10 NLO PDF central value is used as the nominal value of the final estimate. The final PDF uncertainty on that estimate is taken as the envelope of the uncertainty bands for all three PDF sets.

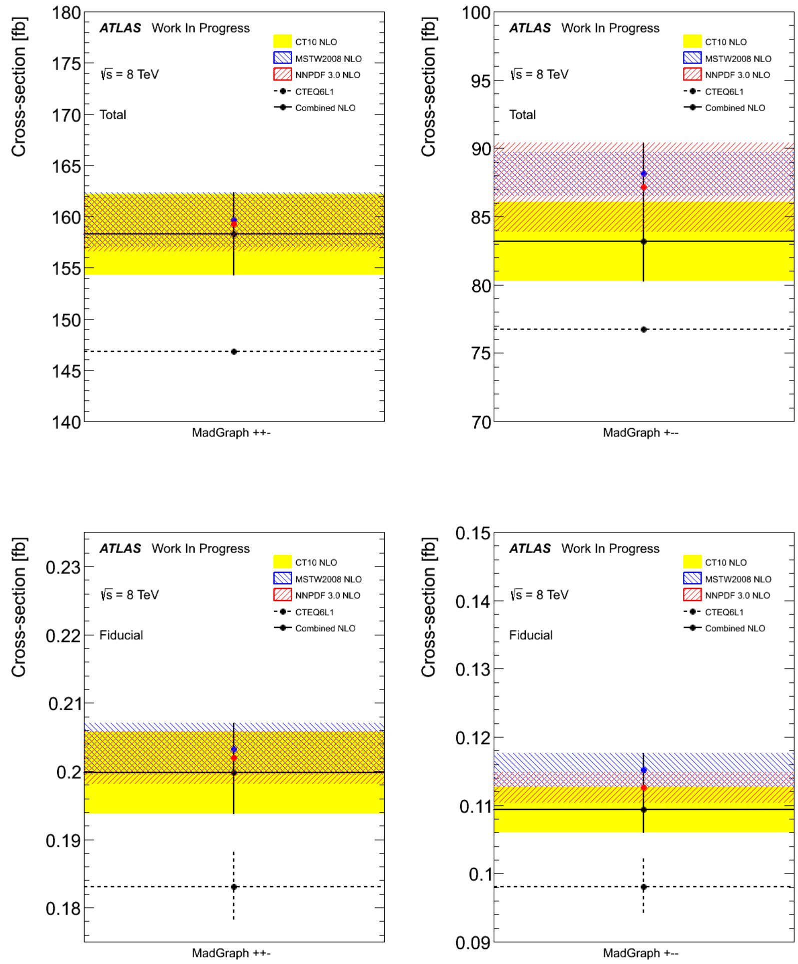

Figure 5.3: The signal cross-sections for different PDFs along with their uncertainties are shown on the MadGraph WWW signal samples for the total WWW phase space and branching fraction for for the $W^+W^+W^-$ (top left) and $W^+W^-W^-$ (top right) charge modes and in the fiducial region for $W^+W^+W^-$ (bottom left) and $W^+W^-W^-$ (bottom right). The bands show the PDF uncertainty for CT10 NLO (solid yellow), MSTW 2008 NLO (hashed blue), and NNPDF 3.0 NLO (hashed red). The solid line shows the envelope of all uncertainty bands used as the final PDF uncertainty estimate. The central value of CT10 NLO is taken as the central value of the estimate. The dashed-line shows the cross-section and statistical uncertainty for the CTEQ6L1 pdf sets used in the original generation step.

The uncertainty on the factorization and renormalization scales are determined by varying each of them independently up or down by a factor of two. The effect of these variations on the cross-sections as compared to the nominal are shown separately for the two different charge modes in Table 5.4. The uncertainty is then determined by taking the maximum variation for each charge mode, namely, 2.62% for $W^+W^+W^-$ and 2.53% for $W^-W^+W^-$.

	μ_F \ μ_R	$\frac{1}{2}M_{WWW}$	M_{WWW}	$2M_{WWW}$
$W^+W^+W^-$	$\frac{1}{2}M_{WWW}$	2.62%	-0.14%	-2.11%
	M_{WWW}	2.13%	0	-2.41%
	$2M_{WWW}$	1.56%	0.24%	-2.42%
	μ_F \ μ_R	$\frac{1}{2}M_{WWW}$	M_{WWW}	$2M_{WWW}$
$W^-W^+W^-$	$\frac{1}{2}M_{WWW}$	1.91%	1.38%	-2.00%
	M_{WWW}	1.61%	0	-2.53%
	$2M_{WWW}$	1.25%	-1.05%	-2.12%

Table 5.4: The relative variation of the NLO cross sections corresponding to different choices of factorization and renormalization scales for the $W^+W^+W^-$and $W^-W^+W^-$processes.

The signal cross-sections and uncertainties are thus determined to be

$$\sigma^{\text{Total}}_{\text{Theory}} = 241.47 \pm 0.13 \text{ (Stat.) } {}^{+10.33}_{-6.08} \text{ (PDF) } \pm 6.3 \text{ (Scale) fb} \tag{5.1}$$

for the inclusive cross-section and

$$\sigma^{\text{Fiducial}}_{\text{Theory}} = 309.2 \pm 7.2 \text{ (Stat.) } {}^{+15.05}_{-8.36} \text{ (PDF) } \pm 8.0 \text{ (Scale) ab} \tag{5.2}$$

for the fiducial cross-section.

5.1.2.2 aQGC signal

blank

5.1.2.3 Background samples

There are other processes produced in proton-proton collisions at the LHC which can mimic the signal processes. These are referred to as background processes. In many cases, the background processes are either more abundant than or of a similar abundance to the signal. As a result, they must be well understood if there is any hope of distinguishing between the two. The background processes to the signal are characterized by having either at least three prompt leptons, meaning they come directly from the hard scattering process; two prompt leptons and an isolated photon, which can mimic an electron; or two prompt leptons and a jet that mimics a lepton. The first two are estimated primarily using MC simulation; the third type is estimated using the data itself. This will be described in more detail in Sec. 5.4.3. For now, we will focus only on the processes estimated using MC simulation.

The most important backgrounds are those with at least three prompt leptons, hereby referred to as the prompt backgrounds. Of these prompt backgrounds, the WZ process is the most important since it has a large cross-section (compared to the signal) and results in a final state with exactly three leptons. Another important prompt background is the ZZ process, which has a similar cross-section to the WZ process, but is typically selected by producing four leptons and then not measuring one. Thus, this process is suppressed by the efficiency for not measuring the presence of a lepton. These are collectively referred to as the di-boson processes, sometimes indicated as VV where $V = W/Z$ (the WW process is also considered but can only produce at most two prompt leptons making it negligible). The di-boson processes are produced using the the POWHEG [12, 50, 37, 11] generator with the CT10 NLO PDF set and hadronized through Pythia8 using the AU2 tune, same as the signal.

Other prompt backgrounds include tri-boson processes like ZWW and ZZZ (typically referred to collectively as VVV) and $t\bar{t} + V$ production. Tri-boson processes have cross-sections of a similar size to the signal but are suppressed for a similar reason as the ZZ, since these can produce either four or six lepton final states. $t\bar{t}+V$ production is when a vector boson is produced in conjunction with a $t\bar{t}$ pair. Since the top quark almost always decays into a W-boson and a b-quark, $t\bar{t}+V$ production also results in an intermediate state of three vector bosons which ultimately results in a three to four lepton final state. The VVV and $t\bar{t}+V$ processes were generated using MadGraph with the CTEQ6L1 PDF set and hadronized using Pythia6 [54] with the AUET2B [19] tune.

The second category of backgrounds to consider are those with two prompt leptons and a photon. We will call these the photon backgrounds. The photon backgrounds occur entirely from the di-boson process $Z\gamma$ where the Z boson decays to two leptons and the photon mimics an electron. A photon is measured by observing an energy deposit in the electromagnetic calorimeter without any associated track in the inner detector. A photon can mimic an electron if it converts into an electron-positron pair while still inside the inner detector, thereby leaving a track in the inner detector while still leaving an energy deposit in the calorimeter, the tell-tale sign of an electron. The $Z\gamma$ samples were generated with the Sherpa [40] generator and the CT10 PDF set. In addition to this process, the $W\gamma$ process behaves similarly but only has one prompt lepton in addition to the photon, so it is negligible. Still, we generate it by using the AlpGen [47] generator with the CTEQ6L1 PDF set and hadronize it using JIMMY [31] with the AUET2C [19] tune.

Some of the di-boson and tri-boson processes just discussed can also be produced through loop induced processes or double parton scattering (DPS). The WW and ZZ loop induced processes are generated using the gg2ZZ [27] and gg2WW [26] generators with the CT10 PDF set and hadronized using JIMMY with the AU2 tunes. The DPS processes are generated using Pythia8 with the AU2 tunes and the CTEQ6L1 PDF set.

The last category of backgrounds are those with prompt leptons plus jets that mimic leptons, hereby referred to collectively as the fake background. The fake background is nominally estimated using the data as described in Sec. 5.4.3. Some of the contributions to this background, however, can be simulated using MC for cross-checks of the estimate from data. The main contributions to the fake background are the single boson processes

(V+jets) and $t\bar{t}$ production. These are processes with very large cross-sections so that even though the probability for a jet mimicking a lepton is small, the size of the cross-section means that their contribution is non-negligible. The single boson Z+jets processes are generated using SHERPA with the CT10 PDF set; the W+jets processes are generated using ALPGEN with the CTEQ6L1 PDF set and hadronized using JIMMY with the AUET2C tunes. For the Z+jets samples, special care must be taken to remove any overlap between with the $Z\gamma$ simulated samples described earlier. The $t\bar{t}$ processes are generated using the MC@NLO [53] generator with the CT10 PDF set and hadronized in JIMMY. Finally, the fake background also has contributions from single top production, though it is less important. Single top production is simulated separately for the s-channel, t-channel, and Wt-channel. The s-channel and Wt-channel are generated using MC@NLO with the CT10 PDF set and hadronized through JIMMY ; the t-channel is generated using MADGRAPH with the CTEQ6L1 PDF set and hadronized using PYTHIA6 with the AUET2B tunes.

5.2 Physics Object Definition and Selection

We attempt to identify the stable particles coming from the proton-proton collisions of the LHC by using the ATLAS detector. The most interesting physics objects to this analysis are the electrons and muons that come from the WWW decay. We also pay attention to the presence of hadronic activity and neutrinos, however, since these can help discriminate the signal from the backgrounds. Each type of particle has a unique signature in the detector that allows us to identify the particle and to reconstruct its properties, such as its charge and four-momentum. This reconstruction process does not guarantee 100% accuracy either in identifying the particle or measuring its properties. As such, the reconstruction process results in reconstructed "physics objects" that may or may not map accurately to the underlying particle or physics it is trying to describe. That being said, this mapping is usually very successful due to the high quality of the detector and the design of the reconstruction algorithms used. To maximize the success of reconstruction we look at physics objects selected only where the reconstruction is well understood. The selections used for the physics objects of interest are described below.

Muon objects are identified by the presence of tracks in both the ID and the MS that are shown to match using an extrapolation process through the gap between the two sub-detectors. To ensure that the track in the inner detector indeed comes from a muon, strict requirements are placed on the number of hits in the different sub-components of the inner detector. The track is extrapolated back to the primary vertex and is forced to point within the boundaries of the MS and ID by requiring that $|\eta| < 2.5$. The muon p_{T} at the primary vertex is chosen to be limited to $p_{\mathrm{T}} > 10$ GeV where there is adequate momentum resolution. We are not interested in muons coming from jets or other hadronic activity, therefore we ask that they be isolated. The isolation of the muon is evaluated in two ways: using tracks and using calorimeter deposits. The isolation determined using tracks is calculated by adding up the scalar sum of the p_{T} of all of the tracks (excluding the muon track) in a cone of $\Delta R < 0.2$ from the muon track. We ask that the isolation from tracks be less than 4% of the muon p_{T}. The isolation determined using calorimeter deposits is calculated in a similar way except that calorimeter deposits are used instead of tracks. We then ask that the isolation from calorimeter deposits be less than 7% of the muon p_{T} when $p_{\mathrm{T}} < 20$ GeV and less than 10% of the muon p_{T} otherwise. Additional requirements are placed upon the track extrapolation to ensure that it comes from the primary vertex.

The signature for electron objects are that they have a track in the inner detector that points to an energy deposit in the EM calorimeter. The electron at the primary vertex is expected to have $p_{\mathrm{T}} > 10$ GeV, similar to the muon objects. The direction of the electron energy deposits are also asked to fall within $|\eta| < 2.47$ and outside the transition region between the EM calorimeter barrel and endcap, $1.37 < |\eta| < 1.52$. The electron objects are required to be isolated and have additional requirements on the track extrapolation, similar to the muon objects.

Jet objects are associated with energy deposits in multiple neighboring cells of the EM and hadronic calorimeter systems. Jet objects are reconstructed by grouping these cells as topological clusters [46] using the anti-k_t algorithm [32] with $\Delta R < 0.4$. The reconstructed jet objects are required to have a reconstructed $p_{\mathrm{T}} > 25$ GeV and to have $|\eta| < 4.5$ so that they are within the boundaries of the calorimeter systems. The reconstructed jets are furthermore selected to suppress contamination from pileup events. This selection is performed by requiring that the majority of the scalar sum of the p_{T} of the tracks associated with the jet are also matched to the primary vertex. This is referred to as the so-called "Jet Vertex Fraction" [49, 3] and is only used with having a $p_{\mathrm{T}} < 50$ GeV and $|\eta| < 2.4$, where the algorithm is shown to perform well. Jets without any associated tracks are always kept.

It is also possible to identify jets that come from heavy flavor decays, namely b quark and b-hadron decays. We refer to these as b-jets. b-jets can frequently be identified because of the relatively long lifetime of the b quark, which can result in a decay vertex that is displaced from the original primary vertex. This can be taken advantage of to "tag" jets as likely coming from b quarks. A multivariate b-tagging algorithm [4] is used with a working point determined to be 85% efficient at identifying b-jets. b-jets are associated with physics processes other than the signal and are helpful in identifying background processes. As a result, we choose to veto events where b-jets are present when looking in the signal regions.

The presence of neutrinos are inferred by a momentum imbalance in the transverse plane, referred to as the missing transverse energy or $E_{\mathrm{T}}^{\mathrm{miss}}$. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated by adding up all of the energy deposits from calorimeters cells within $|\eta| < 4.9$ and then calibrating them based on the the reconstructed physics object they are associated with. If the association is ambiguous then they are chosen based on the following preference (from most preferred to least): electrons, photons, hadronically decaying τ-leptons, jets, and muons. If the calorimeter deposit is not associated with any physics object they are still considered using their own calibration. The sum is modified to take into account the momentum of muons, which typically leave trace energy deposits in the calorimeter without being completely stopped.

It is possible that the reconstructed electrons, muons, and/or jets may overlap with each other inside the detector. This can occur because because of the same physics object being reconstructed as different objects in the ATLAS detector. We handle these occurrences using the following scheme in order of precedence:

1. Electron-Muon Overlap: If $|\Delta R(e, \mu)| < 0.1$, then keep the muon and throw away the electron.

2. Electron-Jet Overlap: If $|\Delta R(e, j)| < 0.2$, then keep the electron and throw away the jet.

3. Muon-Jet Overlap: If $|\Delta R(\mu, j)| < 0.2$, then keep the muon and throw away the jet.

The direction is taken from the calorimeter information for electrons, from the combined track information for muons, and from the anti-kT algorithm for jets. No momentum smearing or calibration corrections are applied to the reconstructed object directions. Using this scheme means that a precedence is set when reconstructed objects overlap such that $\mu > e > j$ where '>' should be interpreted to mean 'is kept instead of'.

The motivation for this scheme is as follows. Muons will frequently radiate photons which then can pair-produce to electrons. If the energy of one of the pair-produced electrons is large enough then this can be reconstructed as well and will likely be collimated with the muon. Since the electron comes from the muon radiation and since the reverse process with an electron having pair-produced muons is heavily suppressed, the muon is kept preferentially. The reconstruction of overlapping electrons and jets would rely on much of the same calorimeter energy deposits. But the electron reconstruction also relies on matching with a well defined inner detector track. It is thus assumed that if an electron overlaps with a reconstructed jet that this is more likely to be the signature of a high energy electron. Finally, if a muon overlaps with a jet, the muon could come from a heavy flavor decay. If this occurs, we choose to keep the event and consider only the muon.

5.3 Event Selection

The expected number of signal events in the total 2012 LHC dataset is expected to be very small compared to the background. Fortunately, the three lepton signature of the signal allows us to quickly throw out many events which do not look like the signal. Still, this signature is not so unique that it removes enough background to reveal the signal. Thus, we must devise a clever way to discriminate between the signal and these backgrounds. We select events in two stages: first we start by selecting events which have the general signature of the signal, this is referred to as the pre-selection stage; we then use more stringent cuts to discriminate between the signal and backgrounds, referred to as our signal region selection. The signal region selection is determined by performing an optimization procedure starting from the pre-selection stage that minimizes the uncertainty on the final measurement. This is described in Sec. 5.6.2. The signal region selection is further divided into different categories that are each used in the final measurement and which allows us to specially treat the different backgrounds in each category. The selections used are described in more detail below.

5.3.1 Pre-selection

The pre-selection is a broad selection which throws away backgrounds that do not at all resemble the signal process. It is mainly characterized by requiring the presence of exactly three leptons (electron or muon) following the requirements listed in Sec. 5.2, each with a p_{T} of at least 20 GeV. In addition, the events are required to be of good quality. This means that the events were collected under good conditions during data taking, both from the LHC operation and ATLAS detector operation. For instance, during the 2012 data collection, the LAr component of the EM calorimeter was known to occasionally produce artificial bursts of noise. These instances were tracked and events where this occurred were thrown away. The event is also required to have a primary vertex with at least three associated tracks. Finally, the event is required to pass the single lepton trigger requirements listed in Sec. 5.1.1 where at least one of the three leptons selected must have caused the trigger to fire.

5.3.2 Signal Region Selection

The signal regions used in this analysis are separated based on the number of Same-Flavor Opposite-Sign (SFOS) lepton pairs selected in the event. That is to say, the number of lepton pair combinations in the event which could feasibly come from the leptonic decay of a Z-boson. This results in three separate signal regions listed below with the lepton charge combinations that fall in each category:

- **0 SFOS**: $e^{\pm}e^{\pm}\mu^{\mp}$, $\mu^{\pm}\mu^{\pm}e^{\mp}$ ($e^{\pm}e^{\pm}\mu^{\pm}$, $\mu^{\pm}\mu^{\pm}e^{\pm}$, $e^{\pm}e^{\pm}e^{\pm}$, $\mu^{\pm}\mu^{\pm}\mu^{\pm}$)
- **1 SFOS**: $e^{\pm}e^{\mp}\mu^{\pm}$, $e^{\pm}e^{\mp}\mu^{\mp}$, $\mu^{\pm}\mu^{\mp}e^{\pm}$, $\mu^{\pm}\mu^{\mp}e^{\mp}$
- **2 SFOS**: $e^{\pm}e^{\pm}e^{\mp}$, $\mu^{\pm}\mu^{\pm}\mu^{\mp}$

Note that in the 2 SFOS region, one lepton is allowed to belong to both pair combinations. Those combinations listed in parentheses are not allowed for the signal based on charge conservation (neglecting charge mis-identification). The amount of the $W^{\pm}W^{\mp}W^{\pm}$ signal which falls into each category is purely combinatoric. From the above list one can thus see that there are twice as many ways for the signal combinations (again neglecting those in parentheses) to fall in the 1 SFOS regions as there are to fall in either the 0 SFOS or 2 SFOS regions. Absent possible differences in signal efficiencies based on the leptons in each signal region, one should expect branching fractions of 25%, 50% and 25% for the 0, 1, and 2 SFOS signal regions, respectively.

In each signal region, a unique selection is determined by an optimization procedure that minimizes the uncertainty on the expected SM measurement. The optimization procedure is described in detail in Sec. 5.6.2. The optimization considers many different physical quantities with which to perform a possible selection, comparing different thresholds for a given quantity and for different combinations of quantities. After optimization a few

	0 SFOS	1 SFOS	2 SFOS
Pre-selection	Exactly 3 leptons with $P_T > 20$ GeV where at least one is trigger matched. (See Section 5.3.1)		
b-tagged Jet Veto	$N_{b-jet} = 0$ (85 % b-tagging efficiency)		
Same-Flavor Mass	$m_{\mathrm{SF}} > 20$ GeV		
Z-Veto ($m_Z = 91.1876$ GeV	$\lvert m_{ee} - m_Z \rvert > 15$ GeV	$m_{\mathrm{SFOS}} < m_Z - 35$ GeV OR $m_{\mathrm{SFOS}} > m_Z + 20$ GeV	$\lvert m_{\mathrm{SFOS}} - m_Z \rvert > 20$ GeV
Missing E_T		$E_T^{Miss} > 45$ GeV	$E_T^{Miss} > 55$ GeV
Lepton-Missing E_T Angle	$\lvert \phi(3l) - \phi(E_T^{Miss}) \rvert > 2.5$		
Inclusive Jet veto	$N_{jet} \leq 1$		

Table 5.5: Optimized signal selection split by number of Same-Flavor Opposite-Sign (SFOS) lepton pairs.

different quantities are determined to be useful for selection. The final selection determined from the optimization is presented in Table 5.5. All cuts are decided from the optimization, and are motivated below.

Since the WWW process is a purely EW process, and since we are looking only at the fully leptonic channel, the signal is expected to have very little hadronic activity. Any observed hadronic activity should come exclusively from the momentum recoil of the WWW system with the incoming partons. Thus, the multi-jet contribution to the signal should be small. As a result, a selection of $N_{\mathrm{Jet}} \leq 1$ is applied in all signal regions. Further, the signal is expected to have negligible contributions from heavy flavor jets. As a result, vetoing events with jets tagged to come from b or b-hadron decays has little effect on the signal expectation. This is true even with the rate for heavy flavor jet mis-identification for the b-tagging algorithms. For the 85% b-tagging efficiency operating point described in Sec. 5.2, the heavy flavor mis-identification rate is measured to be about 1%.

Some of the backgrounds include the production of Z bosons. The invariant mass of the Z-boson can be reconstructed from the SFOS pair coming from the Z-boson decay. This will result in a peak from these backgrounds in the invariant mass distribution around the Z-mass ($m_Z = 91.1876$ GeV [51]). The signal, which does not include Z-bosons, will not have the same peak, but instead will be relatively flat around the region of the Z-peak. As a result, removing events within some window around the peak can do a good job of removing these backgrounds without having a large effect on the signal. For the 1 and 2 SFOS regions, the mass windows chosen for the veto are $m_Z - 35$ GeV $< m_{\text{SFOS}} < m_Z + 20$ GeV and $m_Z - 20$ GeV $< m_{\text{SFOS}} < m_Z + 20$ GeV, respectively. The windows are chosen differently based on the preferred window from the optimization, described in more detail in Sec. 5.6.2. In the 0 SFOS region, by definition, there are no SFOS pairs that could come from the decay of a Z-boson. The effect of electron charge mis-identification ,discussed in Sec. 5.4.2, however, means that a peak can show up in the background of the m_{ee} distribution for same-sign electron/positron pairs. Thus, a veto is performed in this distribution as well, with a mass window of $m_z - 15$ GeV $< m_{ee} < m_Z + 15$ GeV.

The presence of neutrinos in the signal mean that the signal should have a relatively large $E_{\text{T}}^{\text{miss}}$ compared to most of the backgrounds. Thus, cutting on the $E_{\text{T}}^{\text{miss}}$ distribution such that it is large can remove backgrounds expected to have small $E_{\text{T}}^{\text{miss}}$, like $Z\gamma$ production. Still, there are some large backgrounds with neutrinos, like WZ, and also backgrounds that have contributions to the $E_{\text{T}}^{\text{miss}}$ from objects that have missed reconstruction, like ZZ, which can also have a moderate to large $E_{\text{T}}^{\text{miss}}$. Thus, some care must be taken to choose a threshold to cut on the $E_{\text{T}}^{\text{miss}}$ and different thresholds are chosen for each signal region. In the 1 SFOS region the selection is $E_{\text{T}}^{\text{miss}} > 45$ GeV and in the 2 SFOS region the selection is $E_{\text{T}}^{\text{miss}} > 55$ GeV; in the 0 SFOS region, the $E_{\text{T}}^{\text{miss}}$ selection is kept inclusive.

The magnitude and direction of the missing E_T may be interpreted as coming from the vector sum of the neutrinos. By arguments of symmetry, one could then compare the azimuthal direction of the missing E_T to the azimuthal direction of the vector sum of the three charged leptons. When doing so, one finds that in the transverse plane, the direction of the three charged leptons tends to be back-to-back with the direction of the three neutrinos (missing E_T). The backgrounds also show this behavior, but it is less pronounced than it is for the signal. As a result, there is some discriminating power when cutting on the difference in the two angles:

$$\Delta\varphi(lll, E_T^{\text{Miss}}) = \phi(lll) - \phi(E_{\text{T}}^{\text{miss}}) = \cos^{-1}\frac{\overrightarrow{p_T^{lll}} \cdot \overrightarrow{E_{\text{T}}^{\text{miss}}}}{p_T^{lll} E_{\text{T}}^{\text{miss}}} \tag{5.3}$$

The behavior of this quantity for signal and background is similar in all three signal regions. As a result, based on the optimization it was chosen to apply the cut $|\Delta\varphi(lll, E_T^{\text{Miss}})| > 2.5$ everywhere.

5.3.3 Fiducial Region Selection

A fiducial phase space or fiducial region is the region the analysis is sensitive to, defined using purely truth information (generator information before being passed through ATLAS reconstruction). We define our fiducial region based on the optimized signal selection (defined at the reconstruction level) but using only truth information. For instance, the reconstructed lepton p_{T} requirement of $p_{\text{T}}^{\text{Reco}} > 20$ GeV is taken into account in the fiducial region selection by requiring $p_{\text{T}}^{\text{Truth}} > 20$ GeV. By applying this for all cuts in the reconstruction selection shown earlier in Table 5.5, one may compare the predicted signal yields after reconstruction using this selection to the one in the fiducial region selection using just truth information. Any differences are then attributed solely to effects from reconstruction. The fiducial selections are determined at truth level using Rivet [30], which allows for comparisons between different generators.

The chosen fiducial region selection is listed in Table 5.6. Only prompt leptons (those not originating from hadron decays) are used for lepton selections, and these leptons are dressed with prompt photons within a cone with $\Delta R = 0.1$. Generator-level jets are reconstructed by running the anti-kt algorithm with radius parameter $\Delta R = 0.4$ on all final-state particles after the parton showering and hadronization with the exception of prompt leptons, prompt photons, and neutrinos. The $E_{\text{T}}^{\text{miss}}$ variable is calculated using all generator-level neutrinos. As can be seen, the selection in Table 5.6 looks very similar to that in Table 5.5 except for the object definitions using truth information and that events

	0 SFOS	1 SFOS	2 SFOS
All	All		
Tau Veto	$N_\tau < 1$		
Fiducial Leptons	Exactly 3 leptons with $p_T > 20$ GeV and $\lvert\eta\rvert < 2.5$		
Lepton Overlap Removal	$\Delta R(\ell\ell) > 0.1$		
Same-Flavor Mass	$m_{\mathrm{SF}} > 20$ GeV		
Z-Veto ($m_Z = 91.1876$ GeV	$\lvert m_{ee} - m_Z\rvert > 15$ GeV	$m_{\mathrm{SFOS}} < m_Z - 35$ GeV OR $m_{\mathrm{SFOS}} > m_Z + 20$ GeV	$\lvert m_{\mathrm{SFOS}} - m_Z\rvert > 20$ GeV
Missing E_T		$E_T^{Miss} > 45$ GeV	$E_T^{Miss} > 55$ GeV
Lepton-Missing E_T Angle	$\lvert\phi(3l) - \phi(E_T^{Miss})\rvert > 2.5$		
Inclusive Jet veto	$N_{jet} \leq 1$ with fiducial jets of $p_T > 25$ GeV and $\lvert\eta\rvert < 4.5$		

Table 5.6: Fiducial regions based on optimized selection.

are removed if τ leptons are present from the W decays. Thus, the fiducial selection does not include the branching fraction for $W \to \tau\nu$ decay, even though there will be some contamination from this process in the final reconstruction level selection.

5.4 Background Estimates

In Sec. 5.1.2.3, three categories of backgrounds were listed based on the source of final state leptons: prompt, photon, and fake backgrounds. In this section, we will elaborate on how each of these backgrounds are determined as well as provide validation for each of these estimates using control regions. Control regions are regions of phase space that are selected to be enriched in a specific background or collection of backgrounds while at the same time being orthogonal to the signal regions of Sec. 5.3.2, or at least far enough removed so as not to bias the signal region estimate.

The prompt and photon backgrounds are estimated using the MC simulation samples listed earlier in Sec. 5.1.2.3. The most important of these backgrounds are the WZ, ZZ, and $Z\gamma$ backgrounds. The predictions for these backgrounds are studied in Sec. 5.4.1. Where appropriate, corrections to the normalization of these samples are applied to take into account higher order corrections; uncertainties on these corrections are also evaluated.

Even though the WZ and ZZ backgrounds predict at least one SFOS pair from Z-boson decay, they contaminate the 0 SFOS signal region, explained in Sec. 5.3.2, in part because of electron charge mis-identification. The effect of electron charge mis-identification is evaluated in the data and applied as a correction to the WZ and ZZ MC backgrounds in the 0 SFOS region. This is covered in Sec. 5.4.2.

Finally, the fake backgrounds are determined using the data as a model. The details of the fake background estimate and validation are presented in Sec. 5.4.3.

5.4.1 Monte Carlo Backgrounds

Several backgrounds to the signal are simulated purely using MC simulation. The details of these processes, like why they function as backgrounds to the signal and which MC generators are used in the simulation, have already been described in Sec. 5.1.2.3. In some cases, corrections and/or uncertainties on the normalization of these simulated samples are applied. The corrections are summarized in Table 5.7 and are described in more detail below. The performance of the final description for those simulated backgrounds which are most important have been checked in control regions and are also described below.

Background	Normalization Factor	Unceratinty
WZ	1.08	10 %
ZZ	1.05	15 %
$t\bar{t}+V$	1.0	30 %
$ZWW+ZZZ$	1.0	50 %

Table 5.7: Summary of normalizations and their uncertainties for the MC based background estimates used in the analysis.

5.4.1.1 $WZ \rightarrow lll\nu$

The $WZ \rightarrow lll\nu$ background is the most important prompt background to the WWW signal process. Thus, it must be studied carefully. The most recent measurements of the WZ process at the LHC [7, 14, 2] show some tension with the current NLO MC predictions for this process, with differences of about 10 to 15%. Studies of other di-boson processes [42, 33] suggest that this could be resolved by moving to a NNLO calculation. For the WZ process, however, this type of calculation is not yet available. As a result, we instead use the so-called “2D Sideband” method [21] to derive a correction using the data itself.

The 2D sideband method is able to determine an estimate for the process of interest using the data while also correcting for background contamination. To do this, first a signal region is chosen which is enriched in the process of interest. This signal region should have at least two independent selection requirements which when inverted suppress the signal and enhance the backgrounds to that signal. Next, by inverting one, the other, or both selection requirements, three different control regions can be formed where the signal is suppressed and the backgrounds are enhanced with respect to the signal region. These control regions are referred to as “sidebands”. The three sidebands and the signal region may be related to each other assuming independence of the two different selection requirements. If this assumption holds, then the relative change in the backgrounds should be the same when inverting one cut while keeping the other fixed, and vice-versa. In this way, one may solve algebraically for the background contamination in the signal region and subtract it out, resulting in a pure estimate of the signal from the data.

In this case, the signal region is chosen to be enhanced in the WZ process. The backgrounds to this process are from electroweak contributions (like ZZ, $t\bar{t}+V$, and VVV) and from backgrounds with fake leptons. The contributions to the signal region are thus parameterized as

$$N^{\text{Data}} = N^{WZ} + N^{\text{Fake}} + N^{\text{Electroweak}} \tag{5.4}$$

These backgrounds include processes without Z-bosons. Thus, the presence of the Z-boson in the signal means that applying a Z-veto of $|m_{\text{SFOS}} - m_Z| < 15$ GeV will suppress these contributions to the background. Also, requiring that the leptons be isolated does a good job of suppressing the fake background. Thus, the same track and calorimeter isolation requirements are applied to electrons and muons as in the WWW signal regions described in Sec. 5.2.

The Z-veto and the isolation requirements are independently inverted[1] to form the three sidebands. The expectation in each sideband can be parameterized in the same way as Eq. (5.4), resulting in one equation for each region. By specifying the Z-veto condition as A and the isolation condition as B, Eq. (5.4) can be rewritten as:

$$N_{A,B}^{\text{Data}} = N_{A,B}^{WZ} + N_{A,B}^{\text{Fake}} + N_{A,B}^{\text{Electroweak}} \tag{5.5}$$

representing the four different equations after varying A and B independently. For example, the signal region is when A = With Z-veto and B = Isolated. One more equation can be found by assuming that the effect of the isolation cut on the fake background is independent of the Z-veto. That is to say, it is assumed that:

$$R_{\text{With } Z\text{-veto}}^{\text{Fake}} = R_{\text{Without } Z\text{-veto}}^{\text{Fake}} \tag{5.6}$$

where

$$R_{A}^{\text{Fake}} = \frac{N_{A,\text{Isolated}}^{\text{Fake}}}{N_{A,\text{Non-Isolated}}^{\text{Fake}}} \tag{5.7}$$

This results in five equations: the expectations, Eq. (5.5), from varying the conditions A and B independently, and Eq. (5.6).

If we can solve the equations above for $N_{A,B}^{WZ}$ in the signal region, when A = With Z-veto and B = Isolated, then we have our estimate. This is 5 equations and 16 unknowns. The four unknowns, $N_{A,B}^{\text{Data}}$, are determined using the data directly while the electroweak backgrounds, $N_{A,B}^{\text{Electroweak}}$, and the WZ contributions in the sidebands, $N_{A,B}^{WZ}$ (when A = With Z-veto and B = Isolated are not both true) are determined using WZ MC. This reduces the problem to 5 equations and 5 unknowns. Thus, we can solve algebraically for the remaining unknowns including the desired value for the WZ estimate in the signal region.

[1]The thresholds are also slightly shifted so that there is a "dead" region between the signal regions and sidebands which is not used by either. This ensures separation between all regions.

	A \ B	Isolated	Non-Isolated
$N_{A,B}^{\text{Data}}$	With Z-veto	724 ± 27	272 ± 16
	Without Z-veto	67 ± 8	118 ± 11
$N_{A,B}^{\text{Electroweak}}$	With Z-veto	172 ± 3	7.7 ± 0.9
	Without Z-veto	29 ± 2	1.9 ± 0.6
$N_{A,B}^{WZ}$	With Z-veto	—	0.896 ± 0.050
	Without Z-veto	31.82 ± 0.35	0.095 ± 0.015

Table 5.8: All of the inputs used to constrain the system of five equations from Eq. (5.4) and Eq. (5.6). The values are derived in the signal region and three sideband regions described in the text. $N_{A,B}^{\text{Data}}$ are determined directly from the data; $N_{A,B}^{\text{Electroweak}}$ and $N_{A,B}^{WZ}$ are determined in MC. The value for $N_{\text{With }Z\text{-veto,Isolated}}^{WZ}$ is not used as an input and is instead solved for as the the main parameter of interest. Still, the value is determined in MC to be 498 ± 1. Only statistical uncertainties are shown.

	A \ B	Isolated	Non-Isolated
$N_{A,B}^{\text{Fake}}$	With Z-veto	14 ± 43	263 ± 16
	Without Z-veto	6.2 ± 8.3	116 ± 11
$N_{A,B}^{WZ}$	With Z-veto	537 ± 35	—
	Without Z-veto	—	—

Table 5.9: Outputs from the system of five equations from Eq. (5.4) and Eq. (5.6) after including the numbers from Table 5.8 as input. The value for $N_{\text{With }Z\text{-veto, Isolated}}^{WZ}$ is the value of primary interest. Only statistical uncertainties are shown.

The inputs to the system of equations are summarized in Table 5.8[2]. The derived values after solving the system of equations are summarized in Table 5.9. The derived estimate for the WZ contribution to the signal region is 537 ± 35 events, where the uncertainty is purely statistical. Compare this to the estimate from MC of 498 ± 1 events. The ratio of the two can be used to derive a k-factor of 1.08 ± 0.07 (stat.).

Systematic uncertainties are also derived on the method by varying the thresholds used to define the sideband regions, varying the normalization of the MC estimates in Table 5.8, and by varying the degree of equality in Eq. (5.6). The effect of each uncertainty is propagated to the estimate of the WZ normalization in the signal region and are combined in quadrature. The total systematic uncertainty is found to be 5.9%. The final k-factor is thus 1.08 ± 0.07 (stat.) $\pm\, 0.07$ (syst.).

The derived k-factor is applied to the MC estimate in another control region enhanced in the WZ process. This control region is determined using the pre-selection region as described in Sec. 5.3.1 plus an additional requirement that there be 2 SFOS lepton pairs. This gives a good test of the WZ normalization in a control region which is closer to the WWW signal regions. The comparison is shown in Fig. 5.4 where the data is shown to be in good agreement with the corrected WZ MC estimate, as desired.

As a further test of the method, a MC estimate which includes the WZ signal as well as the electroweak and fake backgrounds is used as input in place of $N_{A,B}^{\mathrm{Data}}$ to see if the MC estimate for the WZ contribution in the signal region can be recovered. This is referred to as a closure test. The measured value for the WZ normalization from the closure test is found to be 495 ± 39, which is indeed consistent with the estimate from pure MC of 498 ± 1. The closure test also shows consistent results when varying the normalizations of the different components in the MC independently.

[2]Note that the WZ MC prediction in the signal region is not used except as a comparison.

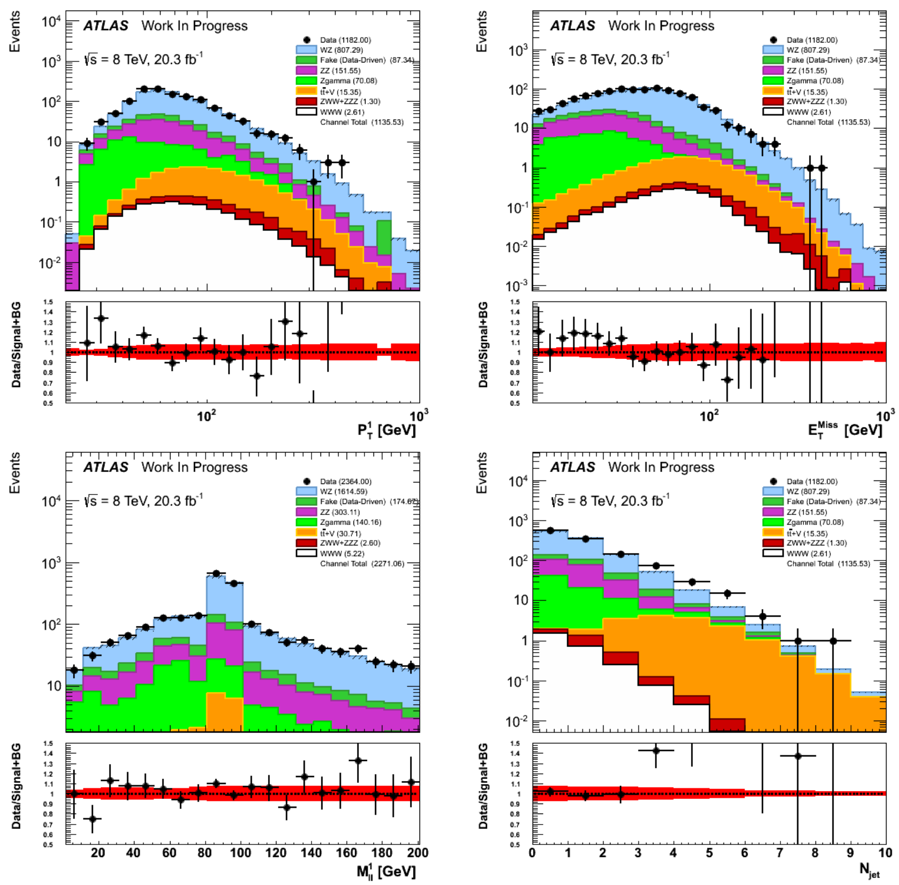

Figure 5.4: WZ control region with 3 lepton pre-selection plus 2 SFOS requirement. Distributions show leading lepton p_T, $E_{\mathrm{T}}^{\mathrm{miss}}$, m_{12}, and jet multiplicity. The systematic band shows the uncertainty on the WZ k-factor.

5.4.1.2 $ZZ \to llll$

The $ZZ \to llll$ process has a similar cross-section as the $WZ \to lll\nu$ process but is suppressed by the probability that exactly one lepton is not reconstructed. Still, this probability is large enough that the ZZ background is one of the largest in the 1 and 2 SFOS signal regions. Unlike the WZ process, NNLO predictions are available from [33, 23, 25] that suggest a k-factor of 1.05 on the overall ZZ prediction. The uncertainty on the prediction is determined to be 15% [33, 23, 25]. This correction is used instead of determining a correction in the data like in Sec. 5.4.1.1.

We may check how well the NLO $ZZ \to llll$ MC prediction and NNLO normalization correction describe the process in the data by looking in a four lepton control region. The leptons are required to have the same quality requirements as in Sec. 5.2. The leptons are sorted by p_{T} with the highest p_{T} lepton required to have $p_{\mathrm{T}} > 25$ GeV, the next two to have $p_{\mathrm{T}} > 15$ GeV, and the lowest p_{T} lepton to have $p_{\mathrm{T}} > 10$ GeV. From these leptons, two separate SFOS pairs are formed. If there is any ambiguity, first an SFOS pair is formed which gives the greatest possible di-lepton invariant mass and the remaining leptons form the other pair. This is a similar procedure to [22]. Finally, to suppress background contamination in the control region, the invariant mass of both SFOS pairs are required to be near the Z-mass, with $60 < m_{\mathrm{SFOS}} < 120$ GeV for both. The results of the comparison are summarized for a few different distributions in Fig. 5.5 and on the total yield in Table 5.10. The expectation is shown to agree well with the observed data within the stated systematic uncertainty on the k-factor of 15%.

5.4.1.3 $Z\gamma$

The $Z\gamma$ process can produce three leptons and thus fall into the signal regions. Measurements of this process within ATLAS have shown that this process is well described from MC simulation using the Sherpa generator at both 7 and 8 TeV [21, 20]. Thus, no further correction or uncertainty on the normalization is applied.

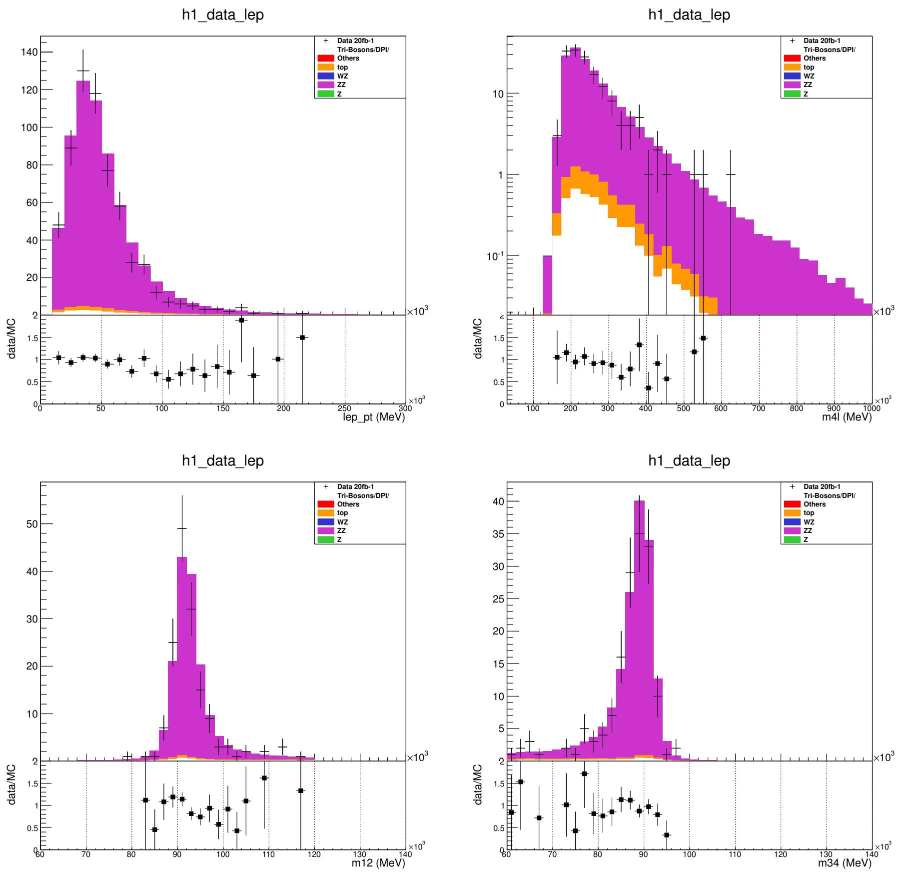

Figure 5.5: $ZZ \to lll$ control region with two separate SFOS pairs. Distribution are shown for the lepton p_{T}, the leading di-lepton mass (m_{12}), the minimum di-lepton mass (m_{34}), and the four lepton mass (m_{4l}).

	Event Yield
WZ	0.05 ± 0.01
ZZ	156.2 ± 0.3(stat) ± 22.3(syst)
$Z\gamma$	0.0 ± 0.0
Fake (MC)	3.6 ± 0.2
tri-boson and $t\bar{t} + V$	4.1 ± 0.2
Expected Signal + Background	164.0 ± 0.3 (stat) ± 22.3(syst)
Observed Data	155

Table 5.10: Number of data and predicted events in the ZZ control region. The error quoted on the MC samples represents only the statistical error. The systematic error due to the k-factor on the ZZ sample is also shown.

The description of the $Z\gamma$ process is tested in a three lepton control region starting from the pre-selection (described in Sec. 5.3.1) and with the same lepton quality requirements as in Sec. 5.2. One of the three leptons should be an electron while the remaining two are required to form a di-muon SFOS pair. For this final state to be produced by the $Z\gamma$ process, the electron should always come from pair production off of the photon, γ, which itself radiates off of the initial state Z boson. As a result, the invariant mass of the di-muon pair coming from the Z-decay will typically be shifted slightly below the Z-mass. However, the invariant mass of the three lepton system should restore this shift such that the mass peak is again centered on the Z-mass. Thus, in order to further suppress backgrounds to the $Z\gamma$ process, we also require that the three-lepton invariant mass, $m_{\mu\mu e}$, be within 15GeVof the Z-mass. The prediction after this selection is compared to data for a few different distributions in Fig. 5.6 and for the total yield in Table 5.11. The control region is clearly enhanced in the $Z\gamma$ process, and furthermore shows very good agreement. This is even true for distributions of the electron kinematics, such as η and p_{T}, which suggests that the photon conversion mechanism is being well modeled.

	Event Yield
WZ	7.47 ± 0.11
ZZ	9.116 ± 0.075
$Z\gamma$	80.3 ± 2.8
$ZWW + ZZZ$	0.0285 ± 0.0046
$t\bar{t} + V$	0.338 ± 0.012
Fake (data-driven)	21.9 ± 1.2
WWW	0.3142 ± 0.0072
Expected Background	119.2 ± 3.1
Expected Signal + Background	119.5 ± 3.1
Observed Data	119 ± 11

Table 5.11: Expected and observed event yields for the Zγ control region. Only statistical uncertainties are shown.

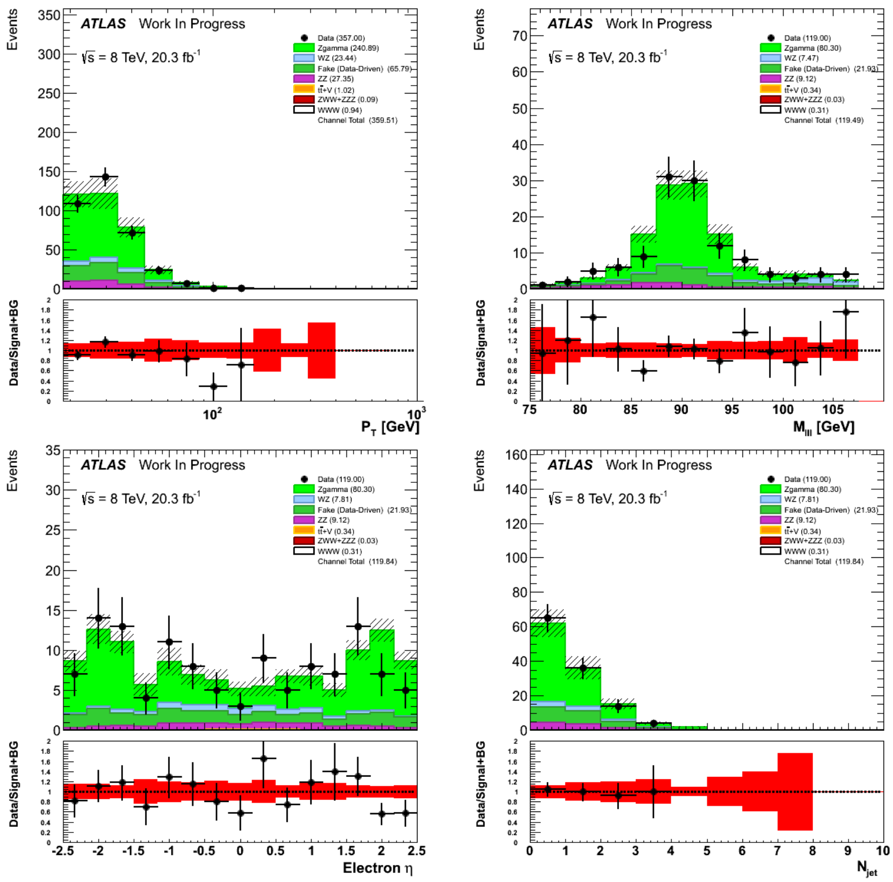

Figure 5.6: Three lepton $Z\gamma$ control region. Distribution are shown for the lepton p_T, three lepton invariant mass ($m_{\mu\mu e}$), electron η, and jet multiplicity.

5.4.1.4 Other Monte Carlo Backgrounds

Backgrounds due to DPS are generated using MC as described in Sec. 5.1.2.3. The cross-section of the DPS process is calculated assuming that the cross-sections of the two incoming processes can be factorized as in [38] using an effective proton cross-section measured in ATLAS at 7 TeV [9]. An overall 50% uncertainty is place on the normalization of these cross-sections. This is a conservative estimate of the uncertainty. However, the contributions of these processes are found to be negligible.

The remaining backgrounds evaluated using MC are those containing at least three real leptons but whose cross-sections are small or on the order of the signal process, namely $t\bar{t}+V$ and VVV processes. The theory uncertainties on the $t\bar{t}+V$ normalization have been found by ATLAS to be about 30% and have been shown to give a consistent prediction [5]. An uncertainty of 30% is also assigned to the normalization of the VVV samples.

5.4.2 Electron Charge Mis-identification

High energy electrons[3] produced from the hard scatter of the proton-proton collisions of the LHC will frequently radiate photons in the presence of the ATLAS detector material. Furthermore, it is also common for high energy photons to decay into an electron-positron pair. These two processes are shown as Feynman diagrams separately in Fig. **??**. Chaining these two processes together will cause an electron (positron) to radiate a photon which then produces an electron-positron pair, resulting in a three body final state with two electrons (positrons) and a positron (electron). Often, the energy difference between the products in the final state will be large, such that the most of the energy is carried away in only one product. It is thus possible that majority of the energy of the initial electron (positron) is carried away in the positron (electron), which has an opposite charge. If the energy imbalance is large enough, the other two final state electrons (positrons) may not have enough energy to be reconstructed. As a result, the initial electron (positron)

[3] Throughout this section we use electrons to collectively refer to both electrons and positrons unless otherwise specified.

will instead be measured as a positron (electron), and the charge of initial state electron (positron) will have effectively been mis-identified.

The probability of this to occur is non-negligible in the presence of the material from the ATLAS detector. This is due to bremstraahlung... Look in 'experimental foundations' and in 'particle detection'

While muons are also technically also capable of such a phenomenon, the energies required are too large, according to bethe-bloche. Indeed, we observe that the rate of charge mis-identification for muons is vanishingly small and so we neglect it.

The strong dependence upon the ATLAS material means that care must be taken when describing this process. In particular, the material description in MC, while sophisticated, is not perfect. Thus, the use of MC for determining the rate of electron charge mis-identification is inherently flawed. Instead, it would be better to use the data itself to determine a model for these rates. Thus, we extract the rates of electron charge mis-identification using the data and only use the rates determined in MC as a cross-check.

The background due to electron charge mis-identification is most important for this analysis in the 0 SFOS signal region, described in Sec. 5.3.2, where it is one of the only mechanisms by which the WZ and ZZ processes enter this region[4]. Without electron charge mis-identification, these events would fall equally in the 1 and 2 SFOS regions. As will be seen shortly, the overall rate of electron charge mis-identification is quite small (calculate???). Furthermore, it will be seen that the total background in the 0 SFOS region is a good deal smaller than the the 1 and 2 SFOS regions. Thus, the migration of events from the 1 and 2 SFOS regions to the 0 SFOS region, resulting from electron charge mis-identification, has a larger relative impact on the background in the 0 SFOS region[5]. As a result, we focus only on modeling the background due to electron charge mis-identification in the 0 SFOS region and assume that an out of the box estimate of this background from

[4]The WZ and ZZ processes can also enter in the 0 SFOS region if the Z bosons decay to τ leptons which then subsequently decay into either electrons or muons with the proper charge and flavor combination.

[5]There is also a migration from the 0 SFOS to the 1 and 2 SFOS regions, but the relative number of 0 SFOS events to 1 and 2 SFOS events before electron mis-identification is so small as to make this effect completely negligible.

MC is adequate for the 1 and 2 SFOS regions.

The electron charge mis-identification background is determined for the 0 SFOS signal region by first extracting the electron charge mis-identification rates using the data as a model described below. The extracted rates are compared to an alternative method using only MC. The difference between the two is used as a systematic on the rates. The rates are then used to re-weight the WZ and ZZ MC samples on an event-by-event basis according to the probability that electron charge mis-identification could cause the event to migrate into the 0 SFOS region. In this way, the full statistics of the MC samples can be utilized to get a model of the behavior of these processes in the 0 SFOS region, while also taking into account a more accurate material description. Other backgrounds due to electron charge mis-identification are assumed to be negligible. More details on the methods used to extract the rates and the re-weighting method are provided below.

5.4.2.1 Charge Mis-identification Rate Extraction

The rate of electron charge mis-identification is defined as the probability that an electron has it's charge mis-identified. These rates depend highly on the kinematics of the individual electrons. In particular, the sensitivity to material dependence described above means that the rate depends on where in the detector the electrons pass through. In general, the material density of the ATLAS detector increases for high η (i.e. as the electron gets closer to the beam pipe), as seen in Sec. 4.1. The rate also increases as a function of the electron energy, or p_{T}. These are the two most important kinematic variables for determining the rate[6], and so the rate extraction is binned as a function of both with nine η bins ranging from 0 to 2.5 and six p_{T} bins ranging from 15 to 120 GeV plus an additional overflow bin for $p_{\mathrm{T}} > 120$ GeV.

The rates are studied in a region with two electrons passing the object selection from Sec. **??** and that have a di-lepton mass within 10 GeVof the Zmass. No requirements are

[6]The material also varies as a function of the azimuthal angle, ϕ, in the detector. However, this is a sub-dominant effect. Furthermore, increasing the dimensionality further significantly harms the statistical power of the method. Thus it is ignored.

placed on their charge. Two different methods are used: one using purely MC and one using the data. The method using MC takes $Z \to ee$ MC simulation and relies on being able to determine the charge of each electron from the Zdecay by looking directly from at hard scattering process as provided by the generator. This is called "truth" information, at which point the processes of radiation and pair-production have not occurred. It then compares these truth electrons to the reconstructed electrons measured after all processes, including those of radiation and pair-production, have been simulated and have been reconstructed in the detector. The truth electrons and reconstructed electrons are matched by asking that they are nearby each other in η and ϕ. The charge of the matched truth and reconstruction electrons are then compared and it is recorded whether or not the charges agree in the appropriate p_{T} and η bin. Once all MC events have been recorded, the rate per bin may be determined simply by taking the ratio of the number of electrons where the truth and reconstructed electron charge disagreed per bin to the total number of electrons per bin.

The nominal method for extracting the electron charge mis-identification rates is instead one that uses the data exclusively. We attempt to subtract backgrounds to the $Z \to ee$ process using a method described later. It uses the same selection as in the MC method, with the events categorized based on whether the electrons from the Zdecay are of the same-sign or of opposite-sign. However, in this case there is no truth information to tell which electron's charge has been mis-identified. Instead, we assume that those events in the same-sign category are due purely to charge mis-identification and attempt to extract the rates by minimizing a likelihood. Refer to the rate for an electron in a particular p_{T} and η bin i as ε_i. Also, refer to the total number of events observed in data with one electron in bin i and the other in bin j as $N_{i,j}$. Given the rates, the expected number of same-sign events should be approximately $N_{i,j}(\varepsilon_i + \varepsilon_j)$, where we have ignored the probability for both electrons to have their charges flipped since it should be small. We do not know the rates *a priori*, but they should follow a Poisson likelihood given the observed total number of events, $N_{i,j}$, and the the observed number of same sign events, $N_{i,j}^{\mathrm{SS}}$, with the following form:

$$\mathcal{L}(\varepsilon_i, \varepsilon_j | N_{i,j}^{\mathrm{SS}}, N_{i,j}) = \frac{(N_{i,j}(\varepsilon_i + \varepsilon_j))^{N_{i,j}^{\mathrm{SS}}} e^{-N_{i,j}(\varepsilon_i+\varepsilon_j)}}{N_{i,j}^{\mathrm{SS}}!} \tag{5.8}$$

From this, we may construct a log likelihood which can be minimized as a function of ε_i and ε_j:

$$-\ln \mathcal{L}(\varepsilon_i, \varepsilon_j | N_{i,j}^{\mathrm{SS}}, N_{i,j}) = N_{i,j}(\varepsilon_i + \varepsilon_j) - N_{i,j}^{\mathrm{SS}} \ln(N_{i,j}(\varepsilon_i + \varepsilon_j)) \tag{5.9}$$

where the terms that are not dependent on ε_i and ε_j have been dropped. Thus, given the data, the values of ε_i and ε_j at the minimum value of the log likelihood are taken as the estimate of the rates.

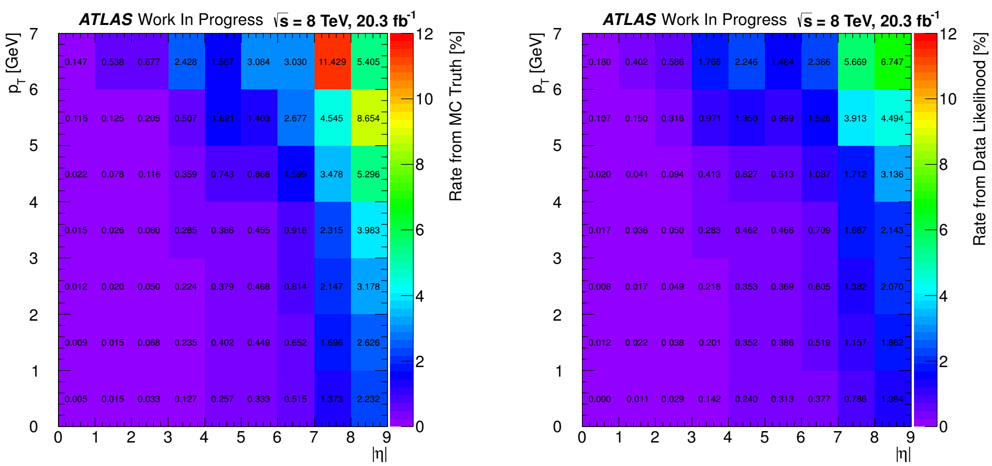

Figure 5.7: Electron charge mis-identification rates as a function of the electron $p_{\mathrm{T}}$and η extracted using the MC truth method (left) and the likelihood method in data (right).

The rates for the two different methods are shown in Fig. 5.7. For low values of $p_{\mathrm{T}}$and η, the rate is small enough to be negligible. The rate increases gradually along both dimensions, reaching as much as 6.7% in the region $p_{\mathrm{T}} > 120$ GeV and $2.4 < |\eta| < 2.5$ as measured in the data, which corresponds to the highest bin in both dimensions. The rates measured using MC truth information are systematically higher than those measured in data, almost by a factor of two. The MC simulation tends to overestimate the amount of material (figure?) actually in the detector, which could explain this difference....

Some additional alternatives to these two methods are also performed in order to better assess the performance of the methods and to determine systematic uncertainties. One alternative is to perform the same likelihood extraction as in the data, but using only reconstructed MC. This produces similar rates to the truth MC method, suggesting that the differences seen between the data likelihood method and the truth MC method are not due to the method itself.

Another method is to extract the rates from the data with the likelihood method but without performing the background subtraction mentioned earlier. In the original method, the background subtraction is performed by... ...with a template fit like in Fig. 5.8.

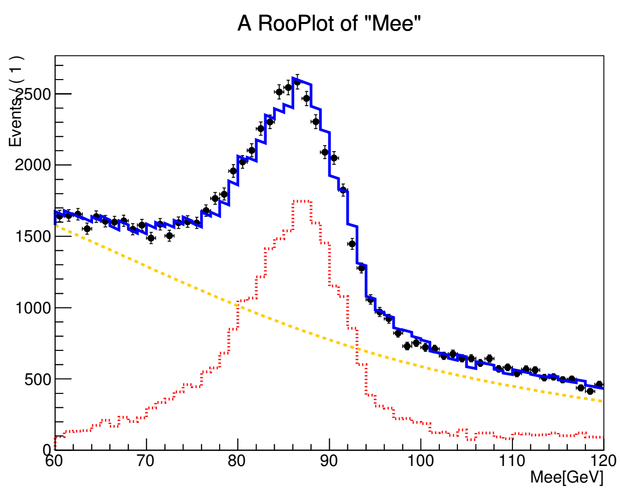

Figure 5.8: Plot of the di-lepton invariant mass in the region where one electron has $0 < |\eta_1| < 0.8$ and the other has $1.15 < |\eta_2| < 1.6$. The data (black points) are shown in a region where the electron isolation cuts are removed and the electron quality requirements are loosened. A template from $Z \to ee$ MC (red line) and a polynomial curve (orange line) are used to fit the data. The sum of the fit (blue line) is seen to fit the data well.

The different methods of rate estimation are compared to extract a final systematic. In Fig. 5.9, the two-dimensional rates are unfolded into one-dimension with the bins numbered counting from low values of η and p_{T} to high values. The nominal rate is the one shown as black points using a likelihood fit in data with background subtraction. The vertical bars on the black points show the statistical uncertainty on this method. The solid black line shows the same method but without background subtraction. These methods follow each other quite closely, uniformly differing by 5-6% throughout. The blue curve shows the rates evaluated using the MC truth method. As was already mentioned, these tend to be larger than the rates in data by about a factor of two. Finally, the red curve shows the rates evaluated using the same likelihood method applied to the data, but using only reconstructed MC. This is seen to follow closely the MC truth method closely, except in a few bins where the statistics are low. The relative difference between the MC truth and MC likelihood methods is transported to the nominal estimate in data and used as another systematic. The difference between the methods using data and those using MC is not used as a systematic since such a difference is expected. The systematic uncertainties are combined in quadrature with the statistical uncertainty on the nominal estimate to arrive at a final uncertainty on the rates, shown as a hashed band.

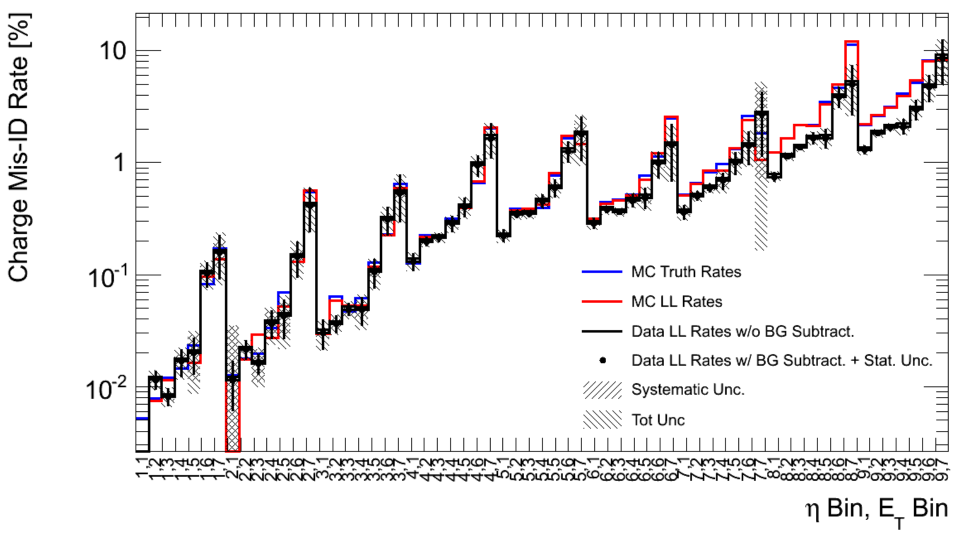

Figure 5.9: Summary of electron charge mis-identification rates using the likelihood method in data with background subtraction (black points) and without background subtraction (black line), the MC truth method (blue line), and the likelihood method in MC (red). Systematic uncertainties are extracted as described in the text and are shown in the gray hashed band pointing from bottom left to top right. The systematic uncertainties are combined with the statistical uncertainties on the black points to arrive at a total uncertainty on the rates, shown in the hashed band pointing from bottom right to top left.

5.4.2.2 Di-boson MC Re-weighting

The electron charge mis-identification rates are primarily important for the determination of the WZ and ZZ background contamination in the 0 SFOS region, as mentioned already. Once derived, the rates are applied to WZ and ZZ MC samples based on whether or not a charge flip could cause the event to appear in the 0 SFOS region. In particular, the following di-boson decays are considered:

- $WZ \rightarrow e^{\pm}\nu\ e^{+}e^{-}$
- $WZ \rightarrow \mu^{\pm}\nu\ e^{+}e^{-}$
- $WZ \rightarrow \tau^{\pm}\nu\ e^{+}e^{-}$
- $ZZ \rightarrow e^{+}e^{-}\ e^{+}e^{-}$
- $ZZ \rightarrow \mu^{+}\mu^{-}\ e^{+}e^{-}$

No other decay channels are considered. These all share in common that they have at least one electron-positron pair. Except for the $WZ \rightarrow \tau^{\pm}\nu\ e^{+}e^{-}$ decay channel, decay channels with tau leptons are not considered because they are suppressed by the tau branching fraction and are considered to be negligible.

The charge mis-identification rates are then applied to these channels on an event-by-event basis as follows. For each event that is processed, its decay channel is identified at truth level. Each reconstructed lepton is examined and assigned a rate, or a probability to charge flip, based on its reconstructed p_{T} and η values. The probability for a charge flip to occur in an event is then approximately the sum of rates for the individual electrons:

$$p(\text{Charge Mis-Identification in Event}) \approx \sum_{i \in \text{Electrons}} \text{Rate}(p_{\mathrm{T}}^{i}, \eta^{i}) \tag{5.10}$$

Higher order terms where multiple electrons are charge mis-identified is small and considered to be negligible. We are only concerned with the probability that a charge flip results in the event falling into the 0 SFOS region. Consider a step function, $\Theta(e)$, defined for an individual event:

$$\Theta(e) = \begin{cases} 1 & \text{if flipping charge of } e \text{ classifies event as 0 SFOS} \\ 0 & \text{if flipping charge of } e \text{ does NOT classify event as 0 SFOS} \end{cases}$$

Then the probability that a charge mis-identification occurs and results in the event falling in the 0 SFOS region is:

$$p(\text{Event is classified as 0 SFOS}) \approx \sum_{i \in \text{Electrons}} \text{Rate}(p_{\text{T}}^{i}, \eta^{i})\Theta(i) \tag{5.11}$$

Again, we ignore the case where multiple electrons have their charge mis-identified. This probability is then used as an event by event weight.

Once the weight has been determined, we then artificially flip the charge of one of the electrons/positrons in the event. If there is only one electron in the event that will lead the event to fall in the 0 SFOS region, its charge is flipped and one proceeds to the next event. However, if there are multiple electrons in the event, there is an ambiguity that must be resolved about which electron's charge should be flipped. One must then be careful in this case to not introduce any bias. We decided to choose a procedure where we pick a single electron from the event at random based on the charge flip rates of the individual electrons. Thus, for an individual electron in an event, the probability that it is chosen to have its charge flipped is:

$$p(e \text{ has been charge flipped}) = \text{Rate}(p_{\text{T}}^{e}, \eta^{e})\Theta(e) \;/ \sum_{i \in \text{Electrons}} \text{Rate}(p_{\text{T}}^{i}, \eta^{i})\Theta(i) \tag{5.12}$$

Consider an example where the event under consideration comes from the decay $WZ \rightarrow e^{+}\nu e^{+}e^{-}$. Assume all three charged leptons pass reconstruction and are selected then label them as: $e_1^{+}\ e_2^{+}e_3^{-}$. In this case, the only way that this event could be classified as 0 SFOS when flipping the charge of only one electron/positron is to flip the charge of the electron.

Thus, $\Theta(e_1^+) = \Theta(e_2^+) = 0$ and $\Theta(e_3^-) = 1$. The event weight will then be equal to the rate of charge mis-identification for e_3^- and it will have it's charge flipped to be positive.

Now consider an example of an event with the decay of $ZZ \rightarrow \mu^+\mu^-\ e^+e^-$. If all four leptons are reconstructed and selected, the event will not be considered at all in the three lepton selection of this analysis, so consider the case where the μ^+ is not selected leaving three leptons labeled as: $\mu_1^- e_2^+ e_3^-$. The probability for the muon to charge flip is negligible which leaves the electron and the positron. Flipping the charge of either one at a time will result in the event being classified as 0 SFOS. Thus, in this case $\Theta(\mu_1^-) = 0$ and $\Theta(e_2^+) = \Theta(e_3^-) = 1$. The event weight will then be the sum of the rates for e_2^+ and e_3^-. The probability that the electron has its charge flipped is then $\frac{\mathrm{Rate}(e_3^-)}{\mathrm{Rate}(e_2^+)+\mathrm{Rate}(e_3^-)}$ and similarly for the positron.

5.4.2.3 Validation

This procedure has been validated on the WZ and ZZ samples by comparing the predictions taken directly from MC to the predictions re-weighted in the 0 SFOS signal region using the procedure just described. This is done in Figure 5.10 for the WZ samples and on Figure 5.11 for the ZZ samples. It can be seen the agreement in the shape looks good for all the distributions. An offset between the two distributions is observed. This difference is covered partially by the systematic uncertainties of the method. Any remaining difference could be expected from the difference in rates observed at high η and high E_T as seen in Fig. 5.9 and serves as justification for using the data-driven method.

There is no special treatment of the charge mis-identification contribution to other background contributions in the 0 SFOS region or to any contributions to the 1 and 2 SFOS signal regions, including di-boson processes, as the effect is expected to be very small. Any charge mis-identification events are thus taken directly from MC in this case.

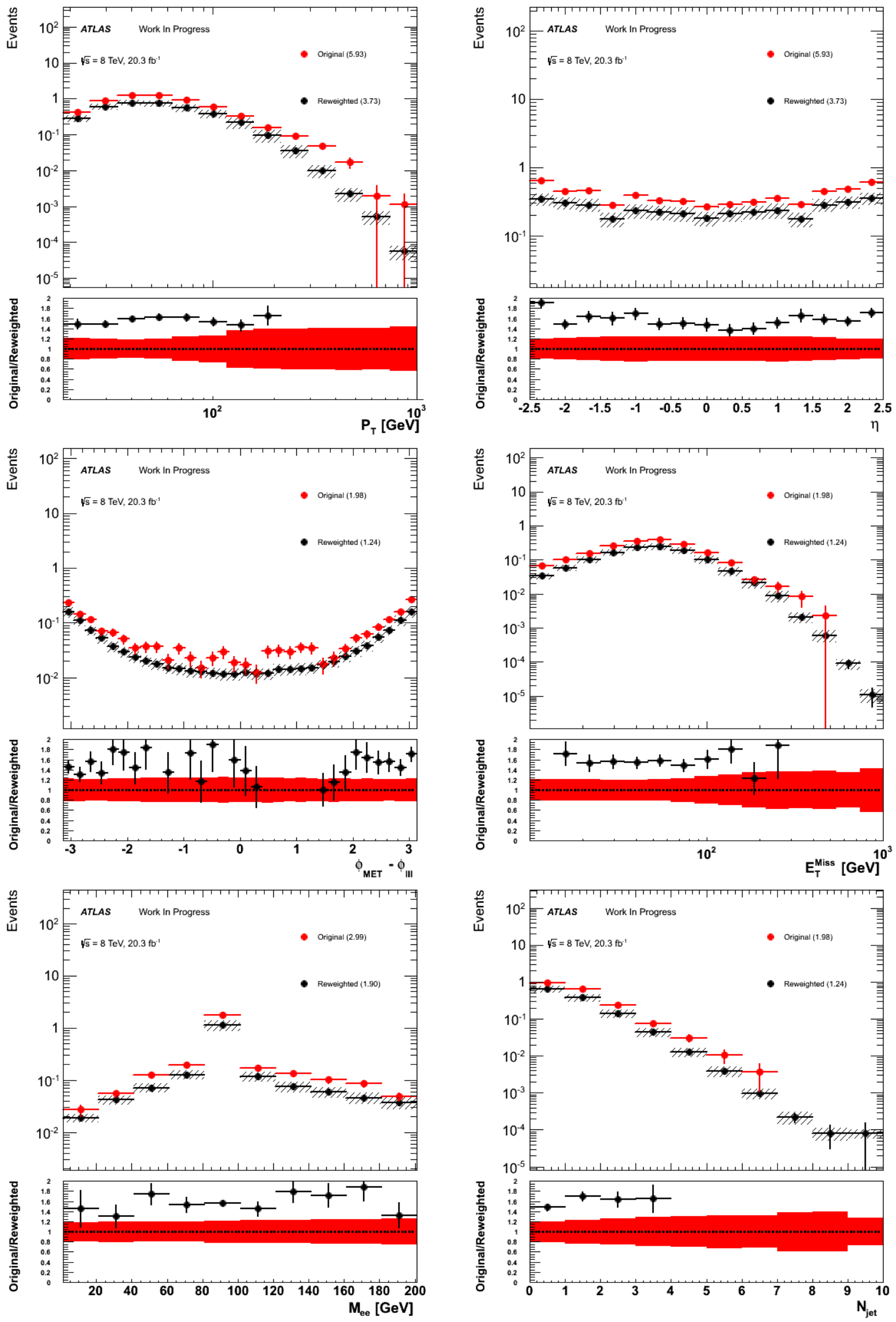

Figure 5.10: Validation of the charge mis-ID rates comparing MC $WZ \to \ell ee$ $(\ell = e, \mu)$ samples re-weighted with the charge mis-ID rates measured in the MC $Z \to ee$ sample to the original MC predictions. Distribution of lepton p_T, η, $\Delta\phi(3l, E_T^{Miss})$,$E_{\mathrm{T}}^{\mathrm{miss}}$, Same-sign di-electron invariant mass, and jet multiplicity.

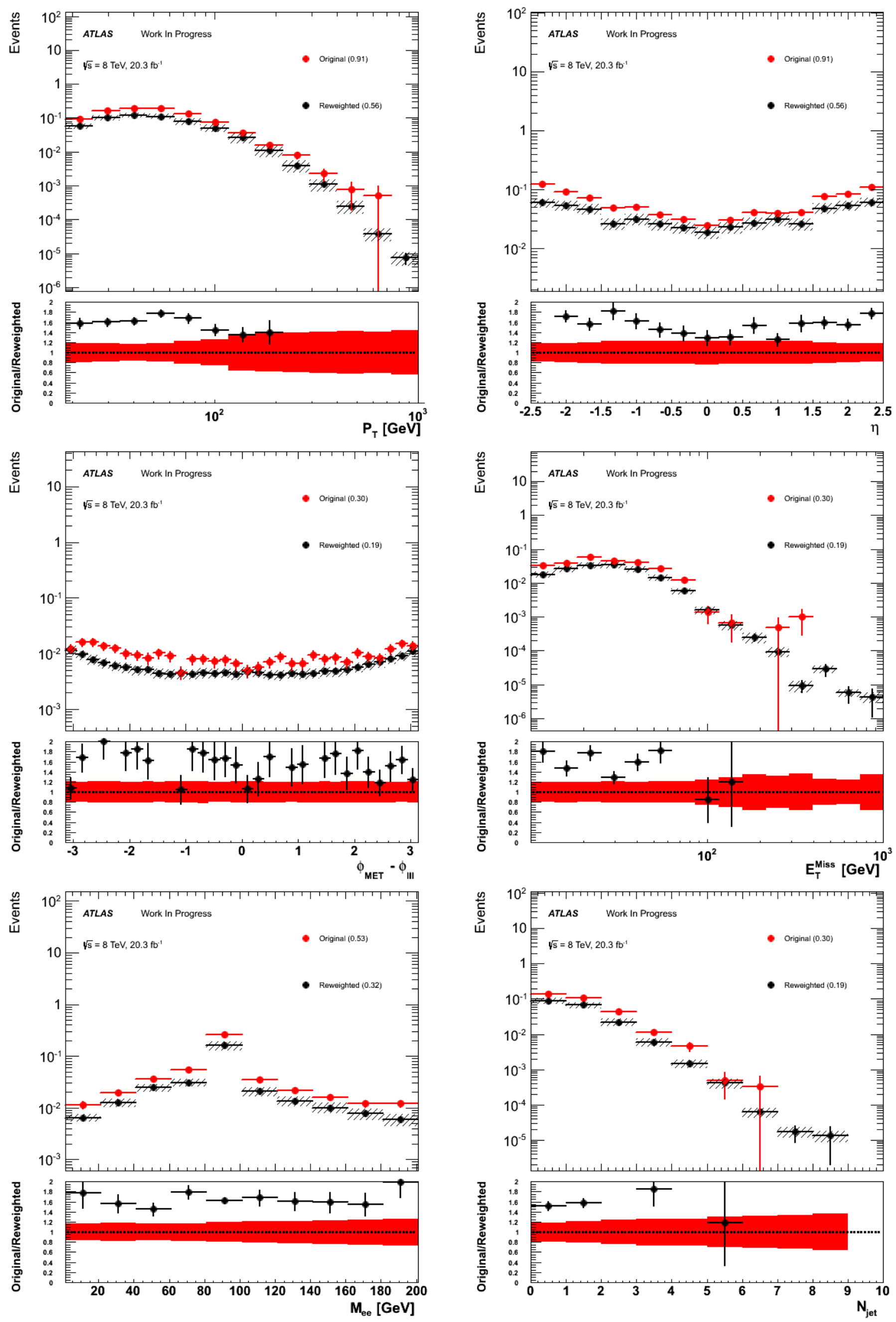

Figure 5.11: Validation of the charge mis-ID rates comparing MC $ZZ \to \ell\ell ee$ ($\ell = e, \mu$) samples re-weighted with the charge mis-ID rates measured in the MC $Z \to ee$ sample to the original MC predictions. Distribution of lepton p_T, η, $\Delta\phi(3l, E_T^{Miss})$,$E_{\mathrm{T}}^{\mathrm{miss}}$, Same-sign di-electron invariant mass, and jet multiplicity.

5.4.3 Fake lepton background

The physics objects reconstructed by the ATLAS detector are selected as described in Sec. 5.2 in order to identify "real" particles or physics processes produced in the collisions of the LHC. Still, this identification is not perfect. A jet, for instance, perhaps from a charged pion, could leave a single track in the inner detector along with a narrow energy deposit in the EM calorimeter; a very similar signature to an electron. Or, a b-hadron, could decay into a final state with a high energy muon, making it difficult to distinguish from a muon produced in the hard interaction. We call these mis-reconstructed objects, "fake" objects. In this analysis we are primarily interested in the subset of these which resemble leptons, like those described above; we refer to these as "fake" leptons.

Fake leptons might be considered physics objects which have "slipped through the cracks". In particle physics it is never the case that the features describing a given particle are completely separable from another, even hypothetically. Instead, the characteristic features for a type of particle will overlap with that of other particles. For example, both electrons and jets are characterized in part by the presence of calorimeter deposits in the EM calorimeter. The calorimeter deposits form a cone pointing back to the collision point, and the radius of this cone will follow some distribution. On average, the deposit from an electron will have a smaller radius than that of a jet. So, on average the radius of calorimeter deposits can be used to distinguish between the two physics processes. But the overlap of these two distributions is significant enough that using this radius alone will give an unsatisfactorily high error rate for identification. The error rate can be improved by adding information from the inner detector, and so on, further reducing the error rate but never reaching zero. So, while rare, the large number of collisions produced by the LHC means that the measurement of fake leptons will inevitably occur. Thus, we must take them into account.

The modeling of these fake leptons are in general heavily dependent upon the conditions of the detector. The detector is described in MC simulation using GEANT4, thus it is possible and relatively straightforward to model these processes using MC directly. However, in practice, this usually proves to be inadequate. While sophisticated, models of the detector are inevitably imperfect. Besides that, some of the effects which produce fake leptons are so rare that it may be difficult to generate enough MC collisions to obtain adequate statistics. The dataset from the LHC, however, has an extremely large sample size and uses the detector itself, not a model. Thus, it has all of the information we need. The trick is then how to extract this information for the signal regions of interest in an accurate and unbiased way.

We choose to use the Generalized Matrix Method [39], which may estimate from data the contribution of any combination fake and real leptons. It has been implemented previously in [15]. A simpler version of the matrix method, which is restricted to the estimation of events with exactly one fake lepton, was first implemented by ATLAS in [1], and variations of the method have been implemented in numerous publications by ATLAS ever since (though less often by CMS). In essence, the method release on the definition of two different selections, referred to as “tight” and “loose”, defined such that “real” leptons are more likely to pass the “tight” selection than “fake” leptons. If the probability of the “real” and “fake” leptons to pass these selections can be determined (typically in control regions), then in principle the easily defined “tight” and “loose” selections may be used as a proxy to extract an estimate of the “real” and “fake” lepton contributions in a region of one’s choosing. The method is described in more detail below.

5.4.3.1 Generalized Matrix Method

The Generalized Matrix Method allows one to extract from data the expected number of events with any combination of fake and real leptons. For any given selection, some fraction of the events will have real leptons, fake leptons, or some combination of the two. For a selection with exactly one lepton, the lepton can simply be either real or fake. Suppose one then defines two orthogonal single lepton selections with in general different combinations of real and fake leptons. Furthermore, design one of the selections to be much more likely to have real leptons than fake leptons, usually taken to be the signal region selection. We will call this the "tight" selection. We can measure directly the number of events in the data that pass this "tight" selection and call it n_T. Choose the other selection to have a different composition of real and fake leptons. Since the "tight" selection is enriched in real leptons, this can be achieved if this other selection has a larger proportion of fake leptons, though not necessarily dominated by fake leptons. We will call this the "loose" selection and designate the number of events measured in this selection as n_L. The total number of real leptons that fall in both regions can be called n_R. The probability that one of these real leptons passes the tight selection is called the real efficiency, or sometimes the real rate, and is denoted by ε_r. Similarly, the total number of fake leptons that fall in both regions is denoted n_F and the probability that one of these fake leptons passes the tight selection is called the fake efficiency, or fake rate, and is denoted by ε_f. The condition that more real leptons pass the tight selection can thus be summarized by saying that $\varepsilon_r >> \varepsilon_f$ be true. The expected values[7] of n_T and n_L, denoted $\langle n_T \rangle$ and $\langle n_L \rangle$, can be related to n_R and n_F using these rates via a system of equations:

$$\begin{pmatrix} \langle n_T \rangle \\ \langle n_L \rangle \end{pmatrix} = \begin{pmatrix} \varepsilon_r & \varepsilon_f \\ \bar{\varepsilon}_r & \bar{\varepsilon}_f \end{pmatrix} \begin{pmatrix} n_R \\ n_F \end{pmatrix} \tag{5.13}$$

where we have introduced the notation $\bar{\varepsilon}_r = 1 - \varepsilon_r$ and $\bar{\varepsilon}_f = 1 - \varepsilon_f$. Note that this equation is a function of the measured values of n_R and n_F which we are actually seeking to find in terms of the expectations of n_T and n_L. Thus, it is in fact more useful to solve for n_R and n_F by taking the inverse:

$$\begin{pmatrix} n_R \\ n_F \end{pmatrix} = \frac{1}{\varepsilon_r - \varepsilon_f} \begin{pmatrix} \bar{\varepsilon}_f & -\varepsilon_f \\ -\bar{\varepsilon}_r & \varepsilon_r \end{pmatrix} \begin{pmatrix} \langle n_T \rangle \\ \langle n_L \rangle \end{pmatrix} \tag{5.14}$$

So far everything is exact and as long as the condition that $\varepsilon_r >> \varepsilon_f$ is true, as it

[7]The issue of deriving an expected value instead from a measured value is an important one which shows how this method can be break down as will be discussed shortly.

should be by construction, then there is no risk of encountering the singular condition when $\varepsilon_r = \varepsilon_f$. But in the matrix method, Eq. (5.14) we wish to use the *measured* values of n_T and n_T to derive an *estimate* of the expectation for n_R and n_F, denoted $\hat{n}_R$ and $\hat{n}_F$. Thus, in a rather *ad hoc* way we interpret the equation as follows:

$$\begin{pmatrix} \langle n_R \rangle \\ \langle n_F \rangle \end{pmatrix} \approx \begin{pmatrix} \hat{n}_R \\ \hat{n}_F \end{pmatrix} = \frac{1}{\varepsilon_r - \varepsilon_f} \begin{pmatrix} \bar{\varepsilon}_f & -\varepsilon_f \\ -\bar{\varepsilon}_r & \varepsilon_r \end{pmatrix} \begin{pmatrix} n_T \\ n_L \end{pmatrix} \tag{5.15}$$

This equation solves for the estimators, $\hat{n}_R$ and $\hat{n}_F$, as a function of the measured values n_T and n_L, as well as the rates. The estimators are in general only approximately equal to the expected values, as discussed in [39]. This approximation can break down, sometimes even giving negative values for the estimate. Though it should be adequate if the number of events falling in the "tight" and "loose" selections are not too small. We will assume that the method holds, but these concerns are important to keep in mind whenever using this method.

We now have a way to approximately solve for the estimate of the real and fake lepton contributions to a single lepton selection in our data sample, but ultimately we are interested in an estimate of the number of fake leptons that fall into our tight selection, call this estimate $\hat{f}_T$ and the number of fake leptons that are loose, $\hat{f}_L$. This estimate can be solved for then in a straightforward way, by selecting only the estimated component of fakes.

$$\begin{pmatrix} \hat{f}_T \\ \hat{f}_L \end{pmatrix} = \begin{pmatrix} \varepsilon_r & \varepsilon_f \\ \bar{\varepsilon}_r & \bar{\varepsilon}_f \end{pmatrix} \begin{pmatrix} 0 \\ \hat{n}_F \end{pmatrix} = \begin{pmatrix} \varepsilon_r & \varepsilon_f \\ \bar{\varepsilon}_r & \bar{\varepsilon}_f \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{n}_R \\ \hat{n}_F \end{pmatrix} \tag{5.16}$$

Substituting in for equation Eq. (5.15) gives an expression for the expected number of tight and loose selected fake leptons as determined from the rates and the measured value of tight and loose leptons:

$$\begin{pmatrix} \hat{f}_T \\ \hat{f}_L \end{pmatrix} = \begin{pmatrix} \varepsilon_r & \varepsilon_f \\ \overline{\varepsilon}_r & \overline{\varepsilon}_f \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{\varepsilon_r - \varepsilon_f} \begin{pmatrix} \overline{\varepsilon}_f & -\varepsilon_f \\ -\overline{\varepsilon}_r & \varepsilon_r \end{pmatrix} \begin{pmatrix} n_T \\ n_L \end{pmatrix} \tag{5.17}$$

And then we may simply pluck out the estimated number of tight leptons from fakes, which is usually of interest:

$$\begin{pmatrix} \hat{f}_T \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \varepsilon_r & \varepsilon_f \\ \overline{\varepsilon}_r & \overline{\varepsilon}_f \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{\varepsilon_r - \varepsilon_f} \begin{pmatrix} \overline{\varepsilon}_f & -\varepsilon_f \\ -\overline{\varepsilon}_r & \varepsilon_r \end{pmatrix} \begin{pmatrix} n_T \\ n_L \end{pmatrix} \tag{5.18}$$

Evaluating the expression for $\hat{f}_T$ gives:

$$\begin{aligned} \hat{f}_T &= \frac{\varepsilon_f}{\varepsilon_r - \varepsilon_f}\Big(\varepsilon_r(n_T + n_L) - n_T\Big) && (5.19) \\ &= \Big(\frac{\varepsilon_f}{\varepsilon_r - \varepsilon_f} - \varepsilon_r\Big)n_T + \Big(\frac{\varepsilon_f}{\varepsilon_r - \varepsilon_f}\varepsilon_r\Big)n_L && (5.20) \\ &= w_T\ n_T + w_L\ n_L && (5.21) \end{aligned}$$

where in the last line we have reorganized the coefficients in front of n_T and n_T into parameters w_T and w_L which are dependent upon the rates. Practically, the final estimate of $\hat{f}_T$ can be determined by looping over each event in data, weighting each event using either w_T for those passing the tight selection and w_L for those passing the loose selection, and then summing up all of the weighted events. This is a very useful strategy since it allows one to compute the estimate on the fly using a setup similar to the one already processing the data itself. Note that since $\varepsilon_r >> \varepsilon_f$ and $0 < \varepsilon_r < 1$, w_T will always be negative. Thus the method will produce negative weights. This is not a concern as long as we keep in mind that the sum is the only thing that is ultimately of interest. However, it is worth noticing that the total estimate can itself be negative when $\varepsilon_r n_L < \overline{\varepsilon}_r n_T$. Though this can in general be avoided as long as ε_r is close to unity and if n_L is as large or larger than n_T, which should usually be the case anyway. In any case, it shows that it is possible to get negative results if the proper conditions are not met.

It will prove useful to rewrite Eq. (5.17) in a more general form:

$$\hat{F} = \mathbf{\Phi} \mathbf{W} \mathbf{\Phi}^{-1} N \tag{5.22}$$

where for the single lepton case $N^T = \begin{pmatrix} n_T & n_L \end{pmatrix}$ and $\hat{F}^T = \begin{pmatrix} \hat{f}_T & \hat{f}_L \end{pmatrix}$.[8] The quantity $\mathbf{\Phi}$ is the matrix from Eq. (5.13)

$$\mathbf{\Phi} = \begin{pmatrix} \varepsilon_r & \varepsilon_f \\ \overline{\varepsilon}_r & \overline{\varepsilon}_f \end{pmatrix} \tag{5.23}$$

and $\mathbf{\Phi}^{-1}$ is its inverse. Finally, $\mathbf{W}$ is the fake selection matrix which in this case is identified with

$$\mathbf{W} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{5.24}$$

And if we want only the estimate of the remaining tight leptons like in Eq. (5.18) then we can do

$$\hat{T} = \mathbf{M} \mathbf{\Phi} \mathbf{W} \mathbf{\Phi}^{-1} N \tag{5.25}$$

where $\hat{T}^T = \begin{pmatrix} \hat{f}_T & 0 \end{pmatrix}$ and $\mathbf{M}$ is the tight selection matrix:

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{5.26}$$

So far we have considered only the rates in a single category or bin and for a single lepton. But this process can be extended easily for different bins, the lepton p_{T} for instance, with different rates by simply keeping track of each bin using an index. For example, in bin i, one would measure the rates ε_r^i and ε_f^i as well as the values n_T^i and n_L^i to arrive at the expectations for in bin i of $\hat{f}_T^i$ and $\hat{f}_L^i$. Equation(5.22) then becomes $\hat{F}^i = \mathbf{\Phi}^i \mathbf{W} (\mathbf{\Phi}^{-1})^i N^i$. One may then sum over all the of the bins to get a total estimate if desired.

[8] A superscript T refers to the vector transpose.

The matrix method can be also extended to multiple leptons, resulting in the generalized matrix method. Consider the three lepton case, which is most relevant to this analysis. Equation (5.25) becomes

$$\hat{T}^{ijk} = \mathbf{M}\mathbf{\Phi}^{ijk}\mathbf{W}(\mathbf{\Phi}^{-1})^{ijk}N^{ijk} \tag{5.27}$$

where each of the three leptons can be in separate bins i, j, and k. The matrix $\mathbf{\Phi}^{ijk}$ can by constructed by taking the Kronecker product (cite) (denoted by $\otimes$) of the individual single lepton matrices of rates for each lepton:

$$\mathbf{\Phi}^{ijk} = \begin{pmatrix} \varepsilon_r^i & \varepsilon_f^i \\ \overline{\varepsilon}_r^i & \overline{\varepsilon}_f^i \end{pmatrix} \otimes \begin{pmatrix} \varepsilon_r^j & \varepsilon_f^j \\ \overline{\varepsilon}_r^j & \overline{\varepsilon}_f^j \end{pmatrix} \otimes \begin{pmatrix} \varepsilon_r^k & \varepsilon_f^k \\ \overline{\varepsilon}_r^k & \overline{\varepsilon}_f^k \end{pmatrix} \tag{5.28}$$

$$= \begin{pmatrix}
\varepsilon_r^i\varepsilon_r^j\varepsilon_r^k & \varepsilon_r^i\varepsilon_r^j\varepsilon_f^k & \varepsilon_r^i\varepsilon_f^j\varepsilon_r^k & \varepsilon_r^i\varepsilon_f^j\varepsilon_f^k & \varepsilon_f^i\varepsilon_r^j\varepsilon_r^k & \varepsilon_f^i\varepsilon_r^j\varepsilon_f^k & \varepsilon_f^i\varepsilon_f^j\varepsilon_r^k & \varepsilon_f^i\varepsilon_f^j\varepsilon_f^k \\
\varepsilon_r^i\varepsilon_r^j\overline{\varepsilon}_r^k & \varepsilon_r^i\varepsilon_r^j\overline{\varepsilon}_f^k & \varepsilon_r^i\varepsilon_f^j\overline{\varepsilon}_r^k & \varepsilon_r^i\varepsilon_f^j\overline{\varepsilon}_f^k & \varepsilon_f^i\varepsilon_r^j\overline{\varepsilon}_r^k & \varepsilon_f^i\varepsilon_r^j\overline{\varepsilon}_f^k & \varepsilon_f^i\varepsilon_f^j\overline{\varepsilon}_r^k & \varepsilon_f^i\varepsilon_f^j\overline{\varepsilon}_f^k \\
\varepsilon_r^i\overline{\varepsilon}_r^j\varepsilon_r^k & \varepsilon_r^i\overline{\varepsilon}_r^j\varepsilon_f^k & \varepsilon_r^i\overline{\varepsilon}_f^j\varepsilon_r^k & \varepsilon_r^i\overline{\varepsilon}_f^j\varepsilon_f^k & \varepsilon_f^i\overline{\varepsilon}_r^j\varepsilon_r^k & \varepsilon_f^i\overline{\varepsilon}_r^j\varepsilon_f^k & \varepsilon_f^i\overline{\varepsilon}_f^j\varepsilon_r^k & \varepsilon_f^i\overline{\varepsilon}_f^j\varepsilon_f^k \\
\varepsilon_r^i\overline{\varepsilon}_r^j\overline{\varepsilon}_r^k & \varepsilon_r^i\overline{\varepsilon}_r^j\overline{\varepsilon}_f^k & \varepsilon_r^i\overline{\varepsilon}_f^j\overline{\varepsilon}_r^k & \varepsilon_r^i\overline{\varepsilon}_f^j\overline{\varepsilon}_f^k & \varepsilon_f^i\overline{\varepsilon}_r^j\overline{\varepsilon}_r^k & \varepsilon_f^i\overline{\varepsilon}_r^j\overline{\varepsilon}_f^k & \varepsilon_f^i\overline{\varepsilon}_f^j\overline{\varepsilon}_r^k & \varepsilon_f^i\overline{\varepsilon}_f^j\overline{\varepsilon}_f^k \\
\overline{\varepsilon}_r^i\varepsilon_r^j\varepsilon_r^k & \overline{\varepsilon}_r^i\varepsilon_r^j\varepsilon_f^k & \overline{\varepsilon}_r^i\varepsilon_f^j\varepsilon_r^k & \overline{\varepsilon}_r^i\varepsilon_f^j\varepsilon_f^k & \overline{\varepsilon}_f^i\varepsilon_r^j\varepsilon_r^k & \overline{\varepsilon}_f^i\varepsilon_r^j\varepsilon_f^k & \overline{\varepsilon}_f^i\varepsilon_f^j\varepsilon_r^k & \overline{\varepsilon}_f^i\varepsilon_f^j\varepsilon_f^k \\
\overline{\varepsilon}_r^i\varepsilon_r^j\overline{\varepsilon}_r^k & \overline{\varepsilon}_r^i\varepsilon_r^j\overline{\varepsilon}_f^k & \overline{\varepsilon}_r^i\varepsilon_f^j\overline{\varepsilon}_r^k & \overline{\varepsilon}_r^i\varepsilon_f^j\overline{\varepsilon}_f^k & \overline{\varepsilon}_f^i\varepsilon_r^j\overline{\varepsilon}_r^k & \overline{\varepsilon}_f^i\varepsilon_r^j\overline{\varepsilon}_f^k & \overline{\varepsilon}_f^i\varepsilon_f^j\overline{\varepsilon}_r^k & \overline{\varepsilon}_f^i\varepsilon_f^j\overline{\varepsilon}_f^k \\
\overline{\varepsilon}_r^i\overline{\varepsilon}_r^j\varepsilon_r^k & \overline{\varepsilon}_r^i\overline{\varepsilon}_r^j\varepsilon_f^k & \overline{\varepsilon}_r^i\overline{\varepsilon}_f^j\varepsilon_r^k & \overline{\varepsilon}_r^i\overline{\varepsilon}_f^j\varepsilon_f^k & \overline{\varepsilon}_f^i\overline{\varepsilon}_r^j\varepsilon_r^k & \overline{\varepsilon}_f^i\overline{\varepsilon}_r^j\varepsilon_f^k & \overline{\varepsilon}_f^i\overline{\varepsilon}_f^j\varepsilon_r^k & \overline{\varepsilon}_f^i\overline{\varepsilon}_f^j\varepsilon_f^k \\
\overline{\varepsilon}_r^i\overline{\varepsilon}_r^j\overline{\varepsilon}_r^k & \overline{\varepsilon}_r^i\overline{\varepsilon}_r^j\overline{\varepsilon}_f^k & \overline{\varepsilon}_r^i\overline{\varepsilon}_f^j\overline{\varepsilon}_r^k & \overline{\varepsilon}_r^i\overline{\varepsilon}_f^j\overline{\varepsilon}_f^k & \overline{\varepsilon}_f^i\overline{\varepsilon}_r^j\overline{\varepsilon}_r^k & \overline{\varepsilon}_f^i\overline{\varepsilon}_r^j\overline{\varepsilon}_f^k & \overline{\varepsilon}_f^i\overline{\varepsilon}_f^j\overline{\varepsilon}_r^k & \overline{\varepsilon}_f^i\overline{\varepsilon}_f^j\overline{\varepsilon}_f^k
\end{pmatrix} \tag{5.29}$$

and we can solve for the inverse. We are only interested in the components with at least one fake lepton, thus we construct the matrix $\mathbf{W}$ such that:

$$\mathbf{W} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.30}$$

Furthermore, the vector

$$(N^{ijk})^T = \begin{pmatrix} n^{ijk}_{TTT} & n^{ijk}_{TTL} & n^{ijk}_{TLT} & n^{ijk}_{TLL} & n^{ijk}_{LTT} & n^{ijk}_{LTL} & n^{ijk}_{LLT} & n^{ijk}_{LLL} \end{pmatrix} \tag{5.31}$$

In this case, there is only one configuration that gives three tight leptons thus, the matrix $\mathbf{M}$ is constructed to have be an $8x8$ matrix with 1 in the first element and all other elements being zero. This gives

$(T^{ijk})^T = \begin{pmatrix} \hat{f}^{ijk}_{TTT} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ Putting everything together, we can solve for the estimated number of events with three tight leptons coming from three leptons events with at least one fake lepton with three leptons in bins i,j, and k:

$$\begin{aligned} \hat{f}^{ijk}_{TTT} = \; & w_{TTT}(i,j,k)\; n^{ijk}_{TTT} \\ & + (w_{TTL}(i,j,k)\; n^{ijk}_{TTL} + j \leftrightarrow k + i \leftrightarrow k) \\ & + (w_{LLT}(i,j,k)\; n^{ijk}_{LLT} + j \leftrightarrow k + i \leftrightarrow k) \\ & + w_{LLL}(i,k,j)\; n^{ijk}_{LLL} \end{aligned} \tag{5.32}$$

where the terms like $j \leftrightarrow k$ are intended to indicated a copy of the first term in parentheses but with the indices switched as shown. Each term has a w function that is a function of the three lepton indices. These are the weights extracted by the method and they end up taking a simple form:

$$w_{TTT}(i,j,k) = -\frac{\varepsilon_r^i \bar{\varepsilon}_f^i}{\varepsilon_r^i - \varepsilon_f^i}\frac{\varepsilon_r^j \bar{\varepsilon}_f^j}{\varepsilon_r^j - \varepsilon_f^j}\frac{\varepsilon_r^k \bar{\varepsilon}_f^k}{\varepsilon_r^k - \varepsilon_f^k} \tag{5.33}$$

$$w_{TTL}(i,j,k) = \frac{\varepsilon_r^i \bar{\varepsilon}_f^i}{\varepsilon_r^i - \varepsilon_f^i}\frac{\varepsilon_r^j \bar{\varepsilon}_f^j}{\varepsilon_r^j - \varepsilon_f^j}\frac{\varepsilon_r^k \varepsilon_f^k}{\varepsilon_r^k - \varepsilon_f^k} \tag{5.34}$$

$$w_{LLT}(i,j,k) = -\frac{\varepsilon_r^i \varepsilon_f^i}{\varepsilon_r^i - \varepsilon_f^i}\frac{\varepsilon_r^j \varepsilon_f^j}{\varepsilon_r^j - \varepsilon_f^j}\frac{\varepsilon_r^k \bar{\varepsilon}_f^k}{\varepsilon_r^k - \varepsilon_f^k} \tag{5.35}$$

$$w_{LLL}(i,j,k) = \frac{\varepsilon_r^i \varepsilon_f^i}{\varepsilon_r^i - \varepsilon_f^i}\frac{\varepsilon_r^j \varepsilon_f^j}{\varepsilon_r^j - \varepsilon_f^j}\frac{\varepsilon_r^k \varepsilon_f^k}{\varepsilon_r^k - \varepsilon_f^k} \tag{5.36}$$

One can see that for the case of zero or two loose leptons present, the magnitude of the weights are always negative (as long as $\varepsilon_r > \varepsilon_f$), while the for one and three loose leptons present, the magnitude is positive. As with the single lepton case this is not of a concern as long as we look only take seriously the final expectation and this remains positive. However, it might cause some concern to see that the magnitude of these weights decrease the more loose leptons are present, thus the highest magnitude weight will in general be $w(i,j,k)_{TTT}$, which is negative! Fortunately, in the sum this is balanced by the number of leptons observed, which tends to have the opposite trend. As a result, it is those terms with exactly one loose lepton observed that end up dominating the entire calculation, which has a positive weight. The generalized matrix method has been checked using the rates evaluated in Sec. **??** at the pre-selection stage and can be seen in Fig. 5.12. From this it is clear that the TTL term (which also includes the TLT and LTT terms) dominates the calculation, though the effects of the negative weights, in particular from the TTT term, can clearly be seen in the sum. Also, notice that the contribution from terms without three initial leptons is small. This includes contributions with greater than three leptons and can be explained by the rarity of processes that produce such high lepton multiplicities. Thus one could arrive at a good approximation to the full method by just using Eq. (5.32) along with just the weights in Eq. (5.33) and (5.34).

In the analysis, a specialized code is used to evaluate the Generalized Matrix Method on all possible combinations of input and output leptons and checks to see which leptons pass the final selection. It uses the on-the-fly weighting method described above and uses a tensor formulation that improves the computational efficiency of the method. This is also used described in [39]. Uncertainties are calculated by propagating through the uncertainties on the rates. Using the standard propagation of uncertainty, this relies on the derivative of the expectation with respect to the rates. Fortunately, this can be calculated in a straightforward way, though it will not be described here. Correlations between different bins are assumed to be negligible and are ignored. However, since the method relies on calculating multiple weights from the same event, there is a correlation in the uncertainty if these weights end up falling in the same bin. To handle this correlation the uncertainty for these weights are added linearly as opposed to in quadrature when extracting the final uncertainty on the method. The impact of this is seen to mostly negligible with respect to treating them as completely uncorrelated.

5.4.3.2 Rate Determination

The Generalized Matrix Method relies on being able to determine the real and fake rates to be used as inputs to the method. This is usually done by looking into control regions which are designed to be enhanced in sources of real and fake leptons. It is important to note that we can never know with certainty whether a lepton is real or fake. Instead we must be clever enough to find leptons that we are confident are of the appropriate type. One clever trick is to use a method called the tag-and-probe method to better identify real or fake leptons in the control regions; it will be described shortly. Once we have obtained our two separate collections of leptons, one we believe to be rich in real leptons and the other in fake leptons, we can use these leptons to extract the real and fake rates, respectively. The real rate, ε_r^i, in category (or bin) i, is simply defined as the ratio of tight candidate

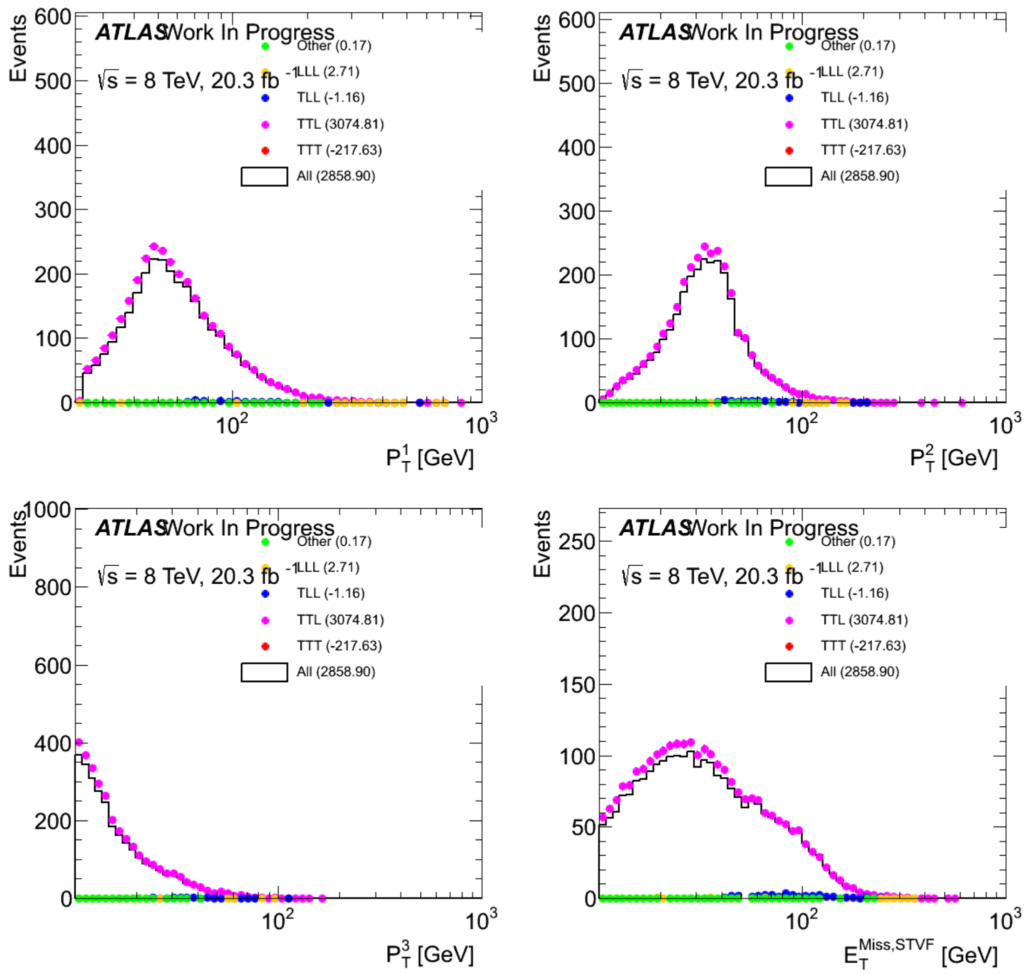

Figure 5.12: Fake background estimate at pre-selection broken into tight and loose lepton configuration components for the leading (top left), sub-leading (top right), and minimum lepton p_{T} (bottom left) as well as the $E_{\mathrm{T}}^{\mathrm{miss}}$ (bottom right). The three lepton TTT (orange dots), TTL (pink dots), TLL (blue dots), and LLL (yellow dots) components are shown along with any other components (green dots) such as those with four leptons initially. The sum of all these components is also shown (black line). (what if I showed the absolute value of these weights?)

real leptons over the number of tight plus loose candidate real leptons:

$$\varepsilon_r^i = \frac{r_T^i}{r_T^i + r_L^i} \tag{5.37}$$

where r_T^i and r_L^i are the number of tight and loose candidate real leptons in category i, to be distinguished from the n_T^i and n_L^i which are the number of candidate and loose real leptons in the signal regions and whose origin is unknown. Similarly, the fake rate, ε_f^i, in category i is, is defined as the ratio of tight candidate fake leptons over the number of tight plus loose candidate fake leptons:

$$\varepsilon_f^i = \frac{f_T^i}{f_T^i + f_L^i} \tag{5.38}$$

where f_T^i and f_L^i are the number of tight and loose candidate fake leptons in category i.

The real and fake rates are not universally the same for all leptons, and in fact can vary strongly. Thus, the choice of categories, i, is an important one. The rates are usually split by lepton flavor and also in bins of at least one kinematic quantity. The splitting of the categories by flavor is very important as the rates are typically very different for electrons and muons. This is in part because the loose and tight selections are usually chosen to be different by necessity. The tight selections are the same as in Sec. **??** for both electrons and muons. The loose selections, however, are similar to the tight selection except that the isolation requirements are removed and the object quality classification is loosened for electrons. Another reason why splitting the categories into lepton flavors is important is that the control regions which are enhanced in real and fake sources of leptons are typically different for electrons and muons. Thus, we choose to evaluate the rates separately for both.

The rates also tend to vary as a function of the lepton kinematics. In particular, both the real and fake rates tend to increase as a function of the lepton p_{T}. Thus, we further divide the electron and muon categories into sub-categories of mutually exclusive bins of p_{T}. The number of bins and the bin edges are determined to best capture the shape while also maintaining adequate statistics in each category. In practice it is usually not possible to subdivide the p_{T} by more than a 3-5 bins. For the same reason, while the rates also surely vary according to other kinematic criteria, like η, it is usually not possible to subdivide in more than one kinematic variable and still have good statistics.

The control regions are chosen so as to be dominated by a single physics process. For determining the real rates, the control region is chosen to be enhanced in $Z \to ll$ while the control region for determining the fake rates is chosen to be enhanced in $W \to l\nu +$ Jets. The reason for this choice is to allow for the application of the tag-and-probe method, which uses one well defined lepton, the "tag", to identify the process, and another lepton, the "probe", as the lepton under study. Both of these control regions have at least one lepton object.

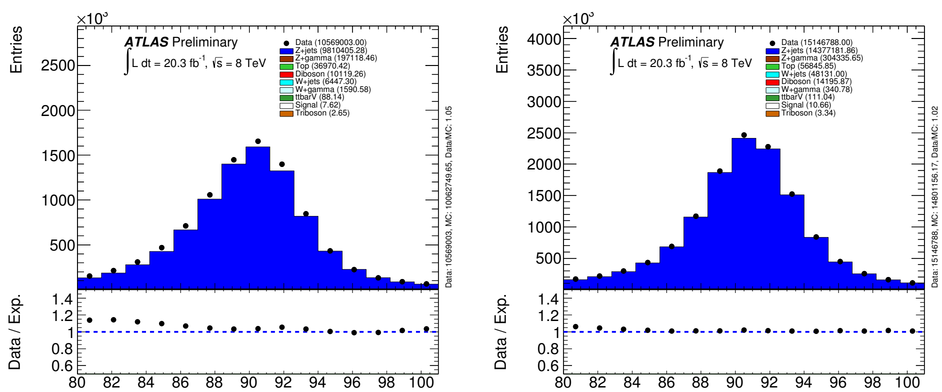

Figure 5.13: Invariant mass distribution of two opposite charge and same flavor di-lepton invariant mass electrons (left) and muons (right). Update figures!!

In the control region enhanced in $Z \to ll$, if one well reconstructed tag lepton passing the tight selection is found then the presence of an additional lepton will almost certainly be the other real lepton from the Z decay. Thus, this second "probe" lepton, which can pass either the loose or tight selection requirement is our candidate real lepton. Note that if the probe lepton also passes the tight selection then it could also be used as a tag. In fact, ignoring this possibility can introduce a bias. Thus, we consider both leptons as tag or probe and candidates. The control region is formed by a selection where the tag lepton passes the same single lepton triggers and trigger matching requirements as in Sec. **??** plus an additional probe lepton that forms an SFOS pair with the tag that has a

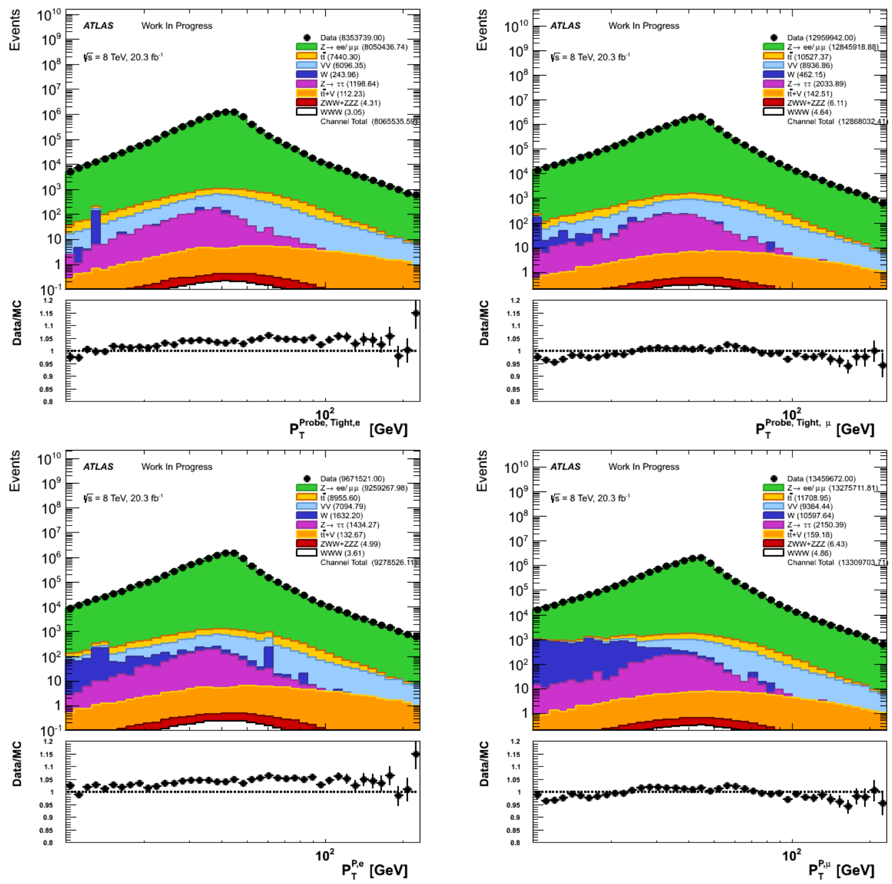

Figure 5.14: Probe lepton p_{T} distributions in SFOS tag and probe control regions used to derive real rates. Electron (left) and muon (right) are shown when the probe lepton is either tight (top) or no additional selection (besides the pre-selection) is required (bottom)

di-lepton mass within 10 GeVof the Z-mass. Two control regions are formed: one from e^+e^- tag-probe pairs for determining the electron real rates and another from $\mu^+\mu^-$ tag-probe pairs for determining the muon real rates. The Z-peak in the di-lepton invariant mass distribution for the two control regions are shown in Fig. 5.13 comparing the data to the model. Since the rates are also determined as a function of the lepton p_{T}, the lepton p_{T} distributions are shown in Fig. 5.14 for the data as well as the expectation separately for electrons and muons and also split based on whether the probe leptons pass just the tight selection (the top row of Fig. 5.14) or both the loose and tight selections (the bottom row of Fig. 5.14). The expectation clearly agrees well with the expectation, which is dominated by the $Z \to ll$ process, as expected. The ratio of the candidate real leptons passing just the tight selection over those passing the loose and tight selections determines the real rate according to Eq. (5.37). The real rates are shown separately for electrons and muons in Fig. 5.15 after adjusting to a coarser binning to improve the statistics. It is interesting to note that the real rates are uniformly lower for electrons than for muons, but both follow the same trend of increasing as a function of the lepton p_{T}, and are relatively high even for the lowest value of 81%. The rates are calculated both using the data or the MC exclusively and the difference is taken as a systematic uncertainty on the nominal estimate from the data. The rates and the systematic uncertainties are summarized for electrons in Table 5.12 and for muons in Table 5.13.

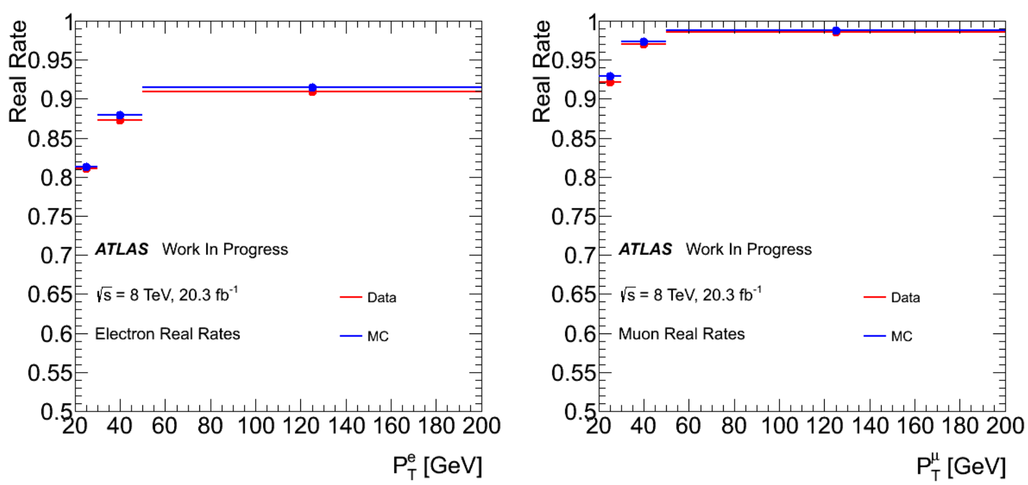

Figure 5.15: Real lepton efficiency as a function of p_{T} and measured in data (red) and MC (blue) for electrons (left) and muons (right).

	Data		MC		
	ε	σ_{stat}	ε	σ_{stat}	σ_{sys}
$p_T \in [20, 30]$ GeV	0.8105	0.0011	0.8134	0.0013	0.0028
$p_T \in [30, 50]$ GeV	0.8732	0.0005	0.8794	0.0006	0.0062
$p_T > 50$ GeV	0.9097	0.0012	0.9150	0.0012	0.0053

Table 5.12: Measured real efficiencies for electrons including statistical and systematic absolute uncertainties. Systematic is calculated by taking the difference between the efficiencies measured in data and MC. The efficiency measured in data is used as the nominal central value.

	Data		MC		
	ε	σ_{stat}	ε	σ_{stat}	σ_{sys}
$p_T \in [20, 30]$ GeV	0.9217	0.0010	0.9291	0.0012	0.0074
$p_T \in [30, 50]$ GeV	0.9700	0.0004	0.9737	0.0006	0.0038
$p_T > 50$ GeV	0.9862	0.0011	0.9878	0.0011	0.0017

Table 5.13: Measured real efficiencies for muons including statistical and systematic absolute uncertainties. Systematic is calculated by taking the difference between the efficiencies measured in data and MC. The efficiency measured in data is used as the nominal central value.

On the other hand, in the $W \to l\nu$ + Jets control region there is only one real lepton being produced by the process. If a well reconstructed tag lepton passing the tight selection is found in this control region it is most likely coming from the W decay. In this case, if we measure a second "probe" lepton it is most likely a jet faking a lepton. Thus, we have found a candidate fake lepton. The control regions are formed by requiring the presence of one tag lepton passing the tight selection plus trigger requirements of Sec. 5.3 with a $p_{\mathrm{T}} > 40$ GeV and a probe lepton passing either the loose or tight selection. The leptons are required to have the same sign, since on average a fake lepton will have equal probably of a positive or negative charge, while background processes like WW,$t\bar{t}$, and Z production produce opposite-sign lepton pairs. Only muons are used as tag leptons. The reason for this is that the chance of an electron passing tight selection to be a jet fake is higher than that for muons. It is also possible for electrons to come from photon conversion while for muons this is very unlikely. Thus, using only muons as tag leptons further reduces contamination from backgrounds in this control region. The control region is then split based on whether the probe lepton is an electron or a muon in order to determine the electron and muon fake rates separately. Events with additional leptons are thrown away. To suppress contamination from multi-jet background processes to the $W \to l\nu$ + Jets process, like QCD, a cut of $E_{\mathrm{T}}^{\mathrm{miss}} > 10$ GeV is also applied. The fake rate that is determined depends upon the source of fake leptons. One way to assess this sensitivity is to consider the number of b-jets present in the event. We consider two different cases regarding the b-jet multiplicity: inclusive and exclusive. The inclusive case makes no requirement on the number of b-jets while the exclusive case asks that at least one b-jet is present. These two different scenarios are ultimately compared in order to assess a final systematic on the fake rate. The exclusive case is used as the nominal estimate.

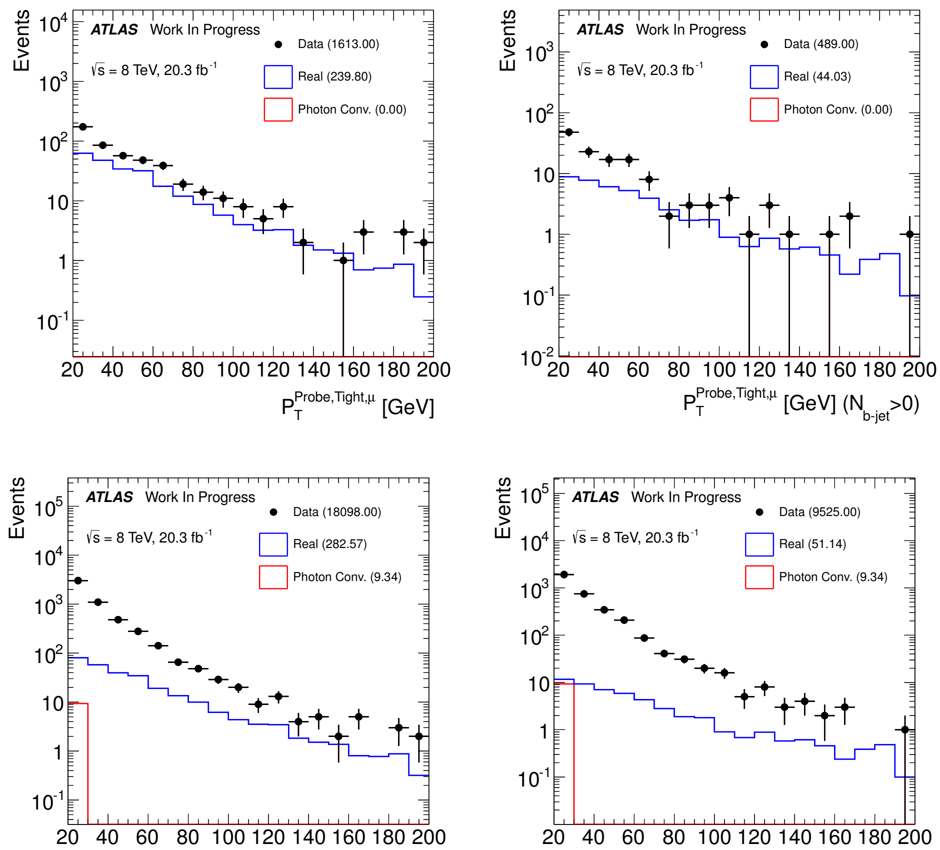

Figure 5.16: Transverse momentum distributions p_T of tight probe muons (top) and loose OR tight probe muons (bottom) passing signal selection criteria in the control Same-Sign $\mu - \mu$ control region without any additional requirement on b-jets in the event (left) and at least one b-jet (right). The amount observed in data (black points) corresponds to n (bottom) and n_{Tight} (top) in Eq. 5.38. Meanwhile, the contribution determined in MC to come from real leptons (blue line) and from photon conversion (red line) are shown separately; they are not stacked. The real lepton contribution corresponds to $n_{\text{Tight}}^{\text{Real}}$ (top) and n^{Real} (bottom) and the photon conversion contribution corresponds to $n_{\text{Tight}}^{\text{PC}}$ (top) and n^{PC} (bottom) in Eq. 5.38. The photon conversion is observed to be negligible for muons.

The processes contributing to the fake rate are known to not be well modeled by MC, as is discussed in Sec. **??**. This is the primary reason for attempting to estimate the fake lepton contribution from data in the first place. Thus, we do not seek to describe the data

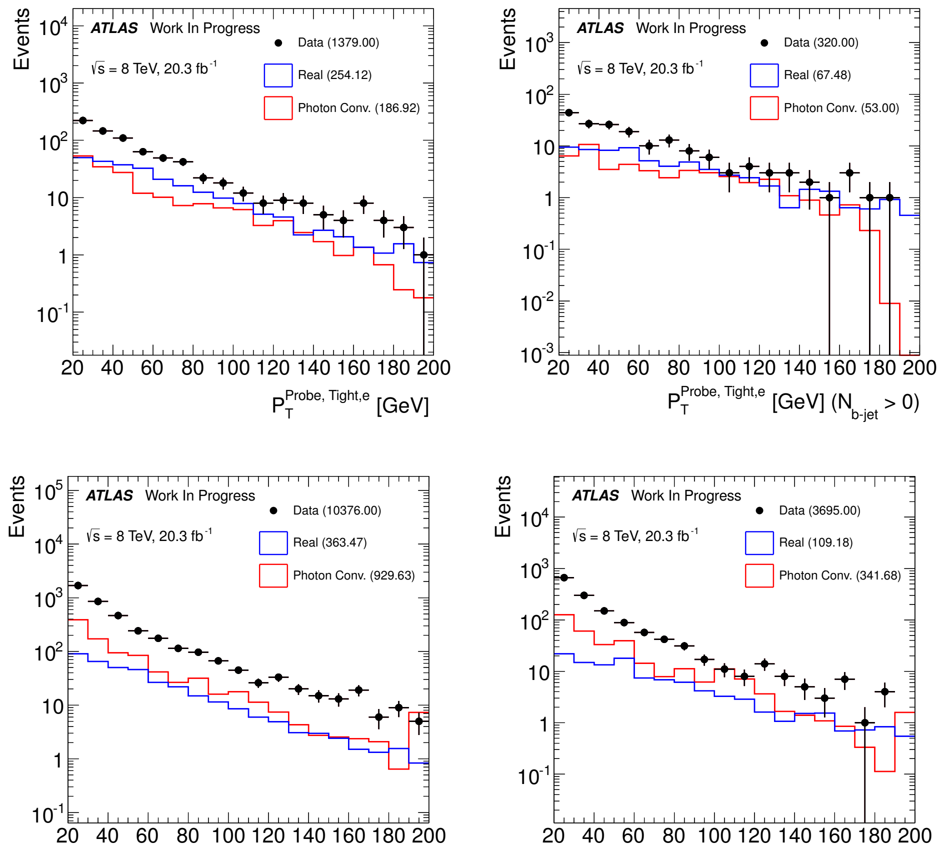

Figure 5.17: Transverse momentum distributions p_{T} of tight probe electrons (top) and loose or tight probe electrons (bottom) passing signal selection criteria in the Same-Sign $e-\mu$ control region without any additional requirement on b-jets in the event (left) and at least one b-jet (right). The amount observed in data (black points) corresponds to n (bottom) and n_{Tight} (top) in Eq. 5.38. Meanwhile, the contribution determined in MC to come from real leptons (blue line) and from photon conversion (red line) are shown separately; they are not stacked. The real lepton contribution corresponds to $n^{\mathrm{Real}}_{\mathrm{Tight}}$ (top) and n^{Real} (bottom) and the photon conversion contribution corresponds to $n^{\mathrm{PC}}_{\mathrm{Tight}}$ (top) and n^{PC} (bottom) in Eq. 5.38.

using MC. However, this control region is also not as pure with sources of fake leptons as the real lepton control region is for real leptons. In particular, there is a significant contamination from processes with real leptons, like WW, $t\bar{t}$, and Z processes as well as processes from photon conversion sources, even after the attempts at reducing these backgrounds in the control region selection described above. These backgrounds can be modeled using MC and so we attempt to subtract the MC estimates of these backgrounds from the data before extracting the fake rates. In effect, this means that the values of f_T^i and f_T^i in Eq. (5.38) are not taken directly from the data but are instead corrected by the subtraction

$$f_{T/L}^i = N_{T/L}^{\mathrm{Data},i} - N_{T/L}^{\mathrm{Real},i} - N_{T/L}^{\mathrm{PC},i} \tag{5.39}$$

where $N_{T/L}^{\mathrm{Data},i}$ is the number of tight or loose probe leptons selected from data in bin i of the fake lepton control region, while $N_{T/L}^{\mathrm{Real},i}$ and $N_{T/L}^{\mathrm{PC},i}$ are the number of tight or loose probe leptons estimated from MC to fall in bin i of the fake lepton control region for real lepton and photon conversion background sources, respectively. The separate contributions to these terms are shown as a function of the lepton p_{T} for split both by whether the lepton passes just the tight selection or both the tight and loose selections and for the inclusive and exclusive b-jet multiplicity categories in Fig. 5.16 for muons and in Fig. 5.17 for electrons. These are then used to calculate the fake rate as in Eq. (5.38). A detailed breakdown of the numbers going into the fake rate calculation are summarized in the exclusive case for electrons in Table 5.14 and for muons in Table 5.15.

Remaining details of fake rate calculation, estimate and validation...

5.5 Systematic Uncertainties

blank

5.6 Event Yields

5.6.1 Event Pre-selection

The signal plus background model (described in detail in Sec. 5.4) is compared to data at pre-selection, defined in Sec. 5.3.1, for a few different kinematic distributions in Fig. 5.18. In the upper plot of each distribution, the colored histograms represent the different categories contributing to the signal plus background model and are split by color based on the category. Hashed bands are shown on the stacked histograms representing the size of the

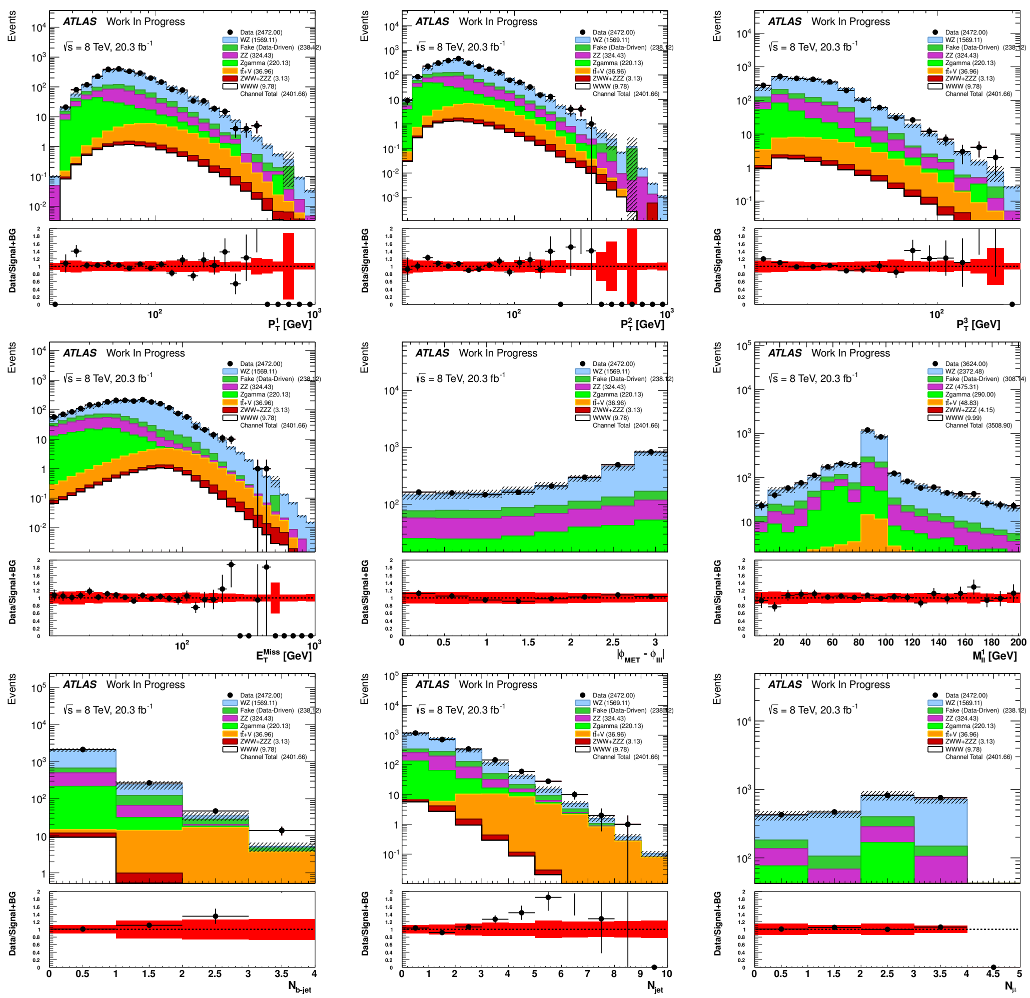

Figure 5.18: Distributions showing the observed data compared to the background estimate at event pre-selection.

systematic uncertainties on the model, described in Sec. 5.5. The data is shown in the black points where the bars on the points represent the statistical uncertainty on the data. The lower plot shows the ratio of the data over the model. In this case, the error bars correspond to the statistical uncertainty on the ratio due to both the data and the model. The red band shows the size of the systematic uncertainties with respect to the model. The model is said to be consistent with the data if the ratio is consistent with unity after considering statistical and systematic uncertainties. The different distributions are chosen primarily because of their potential to discriminate between signal and background. From top to bottom and left to right, these distributions are: the leading, sub-leading, and minimum lepton p_{T} (ordered by their p_{T}), $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\varphi(lll, E_T^{\mathrm{Miss}})$, m_{SFOS}, N_{Jet}, $N_{b-\mathrm{Jet}}$, and N_μ. In general, the signal plus background model is observed to be consistent with the data at pre-selection, at least for those distributions considered here.

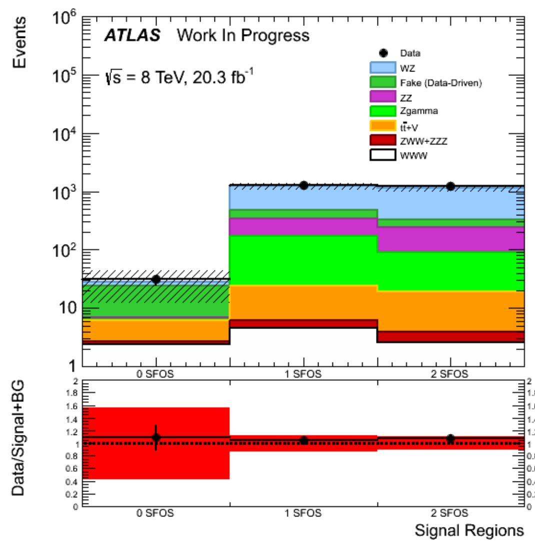

Figure 5.19: Yields at event pre-selection in the 0, 1 and 2 SFOS regions. The most important systematic uncertainties (discussed in section 5.5) are shown, namely from the fake estimates and the uncertainties on the WZ and ZZ k-factors.

Upon splitting the pre-selection region based on the number of SFOS pairs, we end up with signal and background predictions like in Fig. 5.19, where we can see differences in the branching fraction for the signal to each of the three signal regions. In the 0 and 2 SFOS regions, roughly 2.5 signal events are predicted whereas closer to 5 signal events are predicted in the 1 SFOS region, totaling about 10 signal events predicted at the pre-selection stage. Shifting to looking at the background, perhaps the most striking feature of this plot is the clear difference in background yield and background composition between the 0 SFOS region and the 1 and 2 SFOS regions. More than 1000 background events are predicted in both the 1 and the 2 SFOS regions, but only about 30 background events are predicted in the 0 SFOS region. Clearly then, the advantage of splitting the signal region based on this classification comes when looking at the background, specifically the electroweak WZ and ZZ backgrounds where SFOS lepton pairs may be produced from the decay of the Z boson(s). Consider only the case where the WZ and ZZ decay to either e or μ. The WZ production process is thus characterized by 3 leptons with at least 1 SFOS lepton pair which comes from the Z. If all three leptons from the WZ decay have been reconstructed, then there is a 50 % chance the third lepton will also be able to form a SFOS pair with one of the leptons from the Z decay. Thus, the WZ background will split evenly between the 1 and 2 SFOS classification. Something similar occurs for the ZZ background except that the fourth lepton in the decay must be lost (usually due to possessing a low p_{T}). The large cross-section for theses processes means that they become the dominant backgrounds in the 1 and 2 SFOS regions. The 0 SFOS signal region is mostly spared from contamination by these large processes but still includes both the WZ and ZZ processes as background due to the non-negligible (albeit small) effect of mis-measurement of the electron charge described in Sec. 5.4.2. The 0 SFOS signal region is thus unique in having a small background which is almost entirely reducible and dominated instead by the fake background, described in Sec. 5.4.3, along with the aforementioned sub-dominant effect of electron charge mis-identification.

5.6.2 Optimization

From the above discussion, one can clearly see that it is advantageous to split these signal regions so that the dominant backgrounds in each region may be targeted individually. Furthermore, note that even though the 1 SFOS region contains more of the signal than the 0 and 2 SFOS regions, it is the 0 SFOS region which is most likely to have the best sensitivity due to the smaller background contribution. In Sec. 5.3.2 it was already shown that a selection was chosen based on an optimization procedure designed to further reduce the background with respect to the signal region.

The optimization takes as input a multi-dimensional space where each dimension is the selection threshold for one of the quantities listed in Table 5.5, plus some others. The range of the multi-dimensional space is restricted so that the predicted signal remains finite i.e. non-zero. At an individual point in this space, the optimization computes the expected signal and background events after the selection along with the size of statistical uncertainties and systematic uncertainties on the model. These are then used as input to the measurement extraction framework described in Sec. 5.7 to determine the width of the precision on the final measurement. This width is used as the metric to minimize in the optimization. By considering a metric like this, we are optimizing directly the quantity of interest to the final measurement, and taking into account not just the individual predictions, but also their uncertainties. This is important because it can more stringently remove backgrounds that have large uncertainties.

We choose to treat the sample space as being discrete as opposed to continuous. For some dimensions of the space, such as the threshold on N_{Jet}, this is manifestly true, as there can only be an integer number of observed jets. For other dimensions, such as the threshold on the lepton p_{T}, these quantities are real valued and thus continuous. It should be acceptable to only sample discretely, however, as long as they can capture the shape information of the efficiencies. Furthermore, this acknowledges the finite experimental resolution of these quantities. For example, the difference between $p_{\mathrm{T}} > 20$ GeV and $p_{\mathrm{T}} > 20.5$ GeV should not be taken too seriously because of the effects of limited track and energy resolution used to derive the muon and electron p_{T}. Treating the sample space as discrete means that the optimization function is not smooth and so cannot readily take into account derivative information to be used for instance in some sophisticated minimization algorithm. Fortunately, the number of points in the sample space after discretizing, though large, is small enough that it can be evaluated in its entirety using a brute force approach. Thus, we choose to evaluate the optimization in the restricted and discretized sample space in order to find an optimal choice for the selection.

The shape of the optimization can be seen in Fig. 5.20. *Figures need to be reproduced. Elaborate...*

The final selection is presented in Table 5.5. Details of the specific cut thresholds that are chosen can be understood by looking closer at some of the quantities used as input to

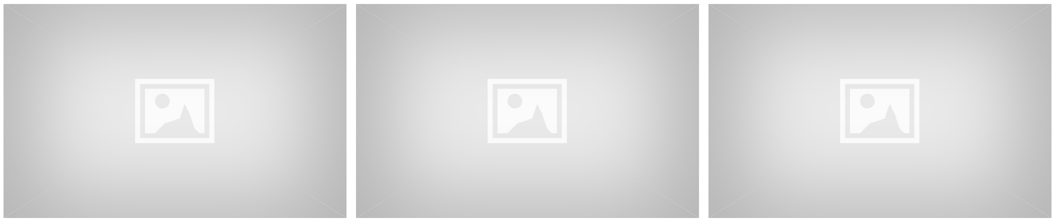

Figure 5.20: Signal Yield vs Measurement Uncertainty for optimized points in the 0 SFOS (left), 1 SFOS (middle), and 2 SFOS (right) signal regions.

the optimization. For instance, it is observed that different $E_{\mathrm{T}}^{\mathrm{miss}}$ and Z-veto thresholds are chosen for the 1 and 2 SFOS regions. This can be understood to come from a correlation between these two quantities due to their ability to isolate the $Z\gamma$ background. The $Z\gamma$ background shows up in the low-shoulder of the Z-peak in the m_{SFOS} distribution and at low MET. This can be seen both for the 1 and 2 SFOS regions in Fig. 5.21. As a result, the $Z\gamma$ background can be removed either by tuning the Z-mass window used in the veto above, or by removing events with low $E_{\mathrm{T}}^{\mathrm{miss}}$. Thus, the optimization shows that there is some correlation between the Z-veto window and the $E_{\mathrm{T}}^{\mathrm{miss}}$ selection threshold. In the 1 SFOS region, there is a larger contribution from $Z\gamma$ processes than in the 2 SFOS region. This process mostly shows up in the low shoulder of the Z peak. The optimization prefers removing this $Z\gamma$ contribution by setting an asymmetric Z-window in the 1 SFOS region, with the boundaries being 35 GeV below the Z-pole and 20 GeV above and then keeping the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut a little loose, with a threshold of $E_{\mathrm{T}}^{\mathrm{miss}} > 45$ GeV. In the 2 SFOS region, however, the $Z\gamma$ contribution is not as prominent and the optimization happens to prefer a symmetric window of ± 20 GeV around the Z-pole. The looser Z-veto then allows for a tighter missing E_T cut with a threshold of $E_{\mathrm{T}}^{\mathrm{miss}} > 55$ GeV.

The absence of any cut on the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the 0 SFOS region can be better understood by looking at the the efficiency for selection between the signal and the background as a function of the $E_{\mathrm{T}}^{\mathrm{miss}}$ selection threshold. This is shown in Fig. 5.22 both after pre-selection and in the 0 SFOS region. Clearly, the signal efficiency closely follows the background efficiency in the 0 SFOS region. Thus, there is no change in the signal-to-

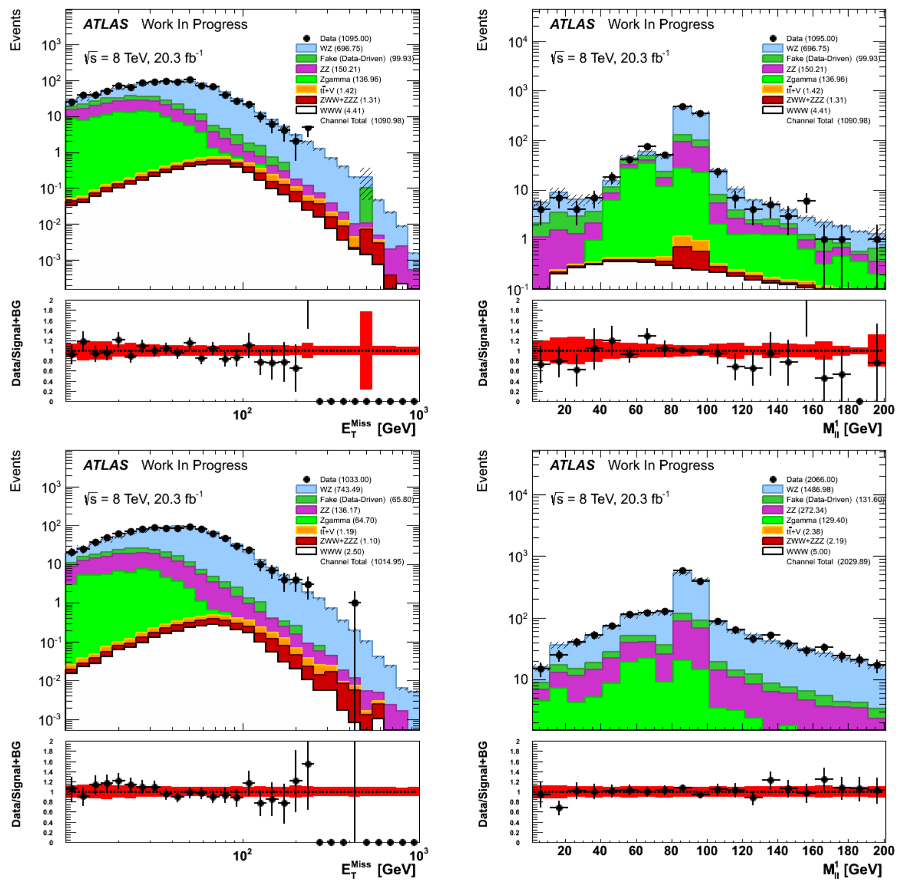

Figure 5.21: Plots of the $E_{\mathrm{T}}^{\mathrm{miss}}$(left) and m_{SFOS} (right) distributions in the 1 SFOS (top) and 2 SFOS (bottom) regions after pre-selection plus the b-veto requirement.

background ratio when cutting on the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the 0 SFOS region and thus no improvement in the sensitivity. On the other hand, there are large shape differences between the signal and background efficiencies at pre-selection, with the signal efficiency remaining flatter at low values of the $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold. So, from this one would expect a selection on the $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold to be useful in the 1 and 2 SFOS regions which have a similar background composition. Indeed, this is what we observe.

The threshold for the jet multiplicity cut of $N_{\mathrm{Jet}} \leq 1$ applied in all signal regions is also determined from the optimization. One might expect that a different value for the threshold, such as a complete veto on the presence of jets, would perform better. Indeed, looking at the efficiency for selection on the jet multiplicity in Fig. 5.23 does show a much

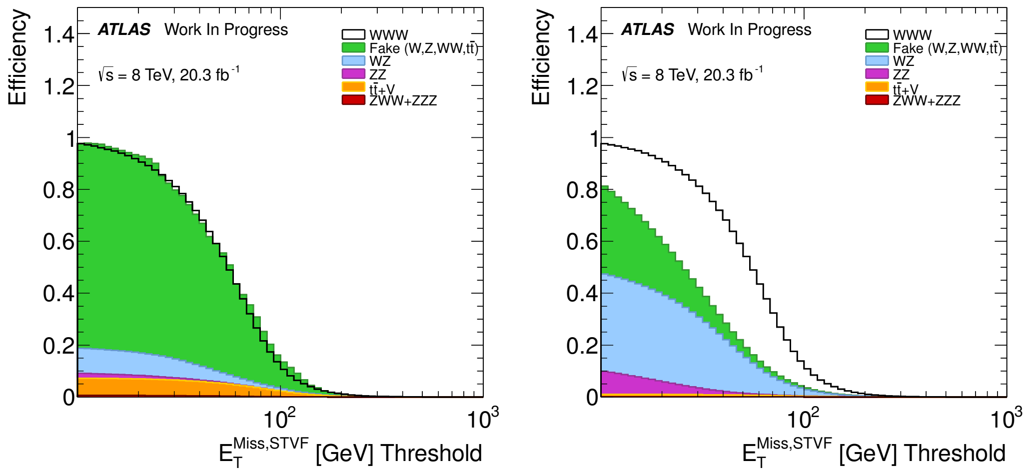

Figure 5.22: Signal and background efficiencies for the selection $E_{\mathrm{T}}^{\mathrm{miss}} > X$ as a function of the $E_{\mathrm{T}}^{\mathrm{miss}}$ selection threshold, X, in both the 0 SFOS (left) and pre-selection (right) regions.

stronger background rejection when applying a veto in both the pre-selection region and especially in the 0 SFOS region where there is a larger contribution from fakes due to hadronic activity. The signal rejection, however, of about 40% observed in both regions, is prohibitive. Loosening the selection to the nominal threshold of $N_{\mathrm{Jet}} \leq 1$ instead preserves 90% of the signal, which is quite precious. We are still able to remove much of the fake background in the 0 SFOS region by vetoing events with b-tagged jets as can be seen in Fig. 5.24. It is possible that using a b-tagging operating point with an even higher b-tagging efficiency would further improve the sensitivity in the 0 SFOS region. The nominal operating point used here, however, is the highest efficiency operating point supported by ATLAS. Clearly, there is no advantage gained from using a looser operating point as this would only cut less on the background without having an impact on the signal.

The $\Delta\varphi(lll, E_T^{\mathrm{Miss}})$ distribution for the signal is observed to be more back-to-back (i.e. closer to π) than that for the background. This is especially true in the 0 SFOS region, as can be seen from the efficiencies plotted as a function of the $\Delta\varphi(lll, E_T^{\mathrm{Miss}})$ selection threshold shown in Fig. 5.25. The selection efficiency for the signal is relatively flat for most of the range up to about a threshold of $|\Delta\varphi(lll, E_T^{\mathrm{Miss}})| > 2.5$ in both the pre-selection

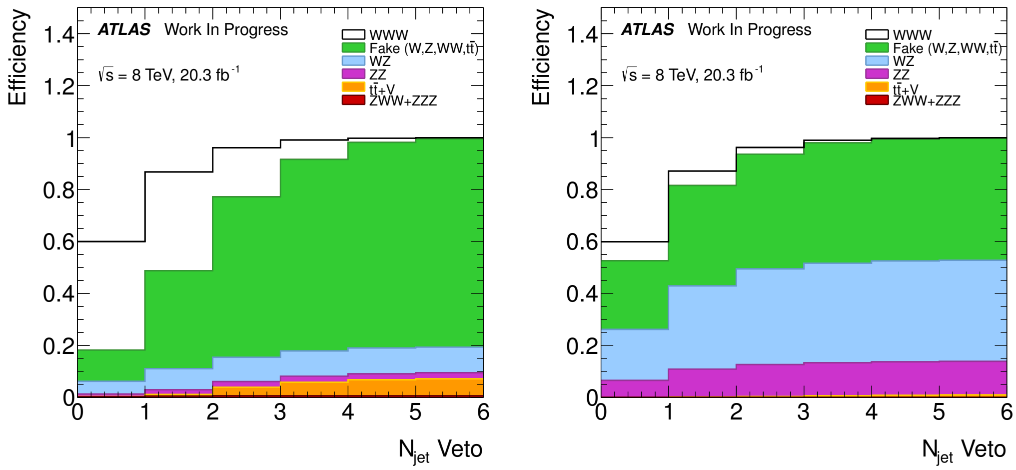

Figure 5.23: Signal and background efficiencies for the selection $N_{\mathrm{Jet}} \leq X$ as a function of the N_{Jet} selection threshold, X, in both the 0 SFOS (left) and pre-selection (right) regions.

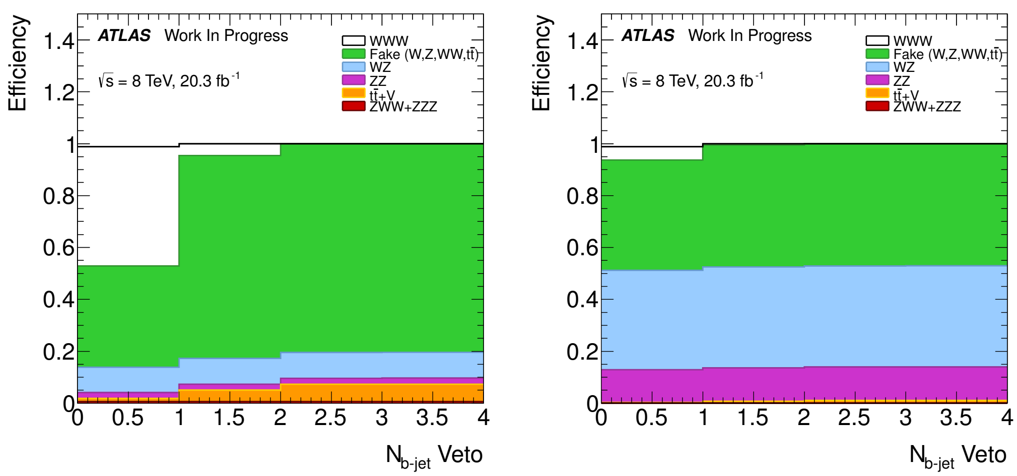

Figure 5.24: Signal and background efficiencies for the selection $N_{b-\mathrm{Jet}} \leq X$ as a function of the $N_{b-\mathrm{Jet}}$ selection threshold, X, in both the 0 SFOS (left) and pre-selection (right) regions.

and 0 SFOS regions. At this threshold the signal selection efficiency is about 80%. The optimization prefers a selection around this range for all signal regions. The optimization also considered selecting on alternative definitions of $\Delta\phi$ that only considered one of the three leptons but this was observed to not offer as strong of a separation between the signal and background.

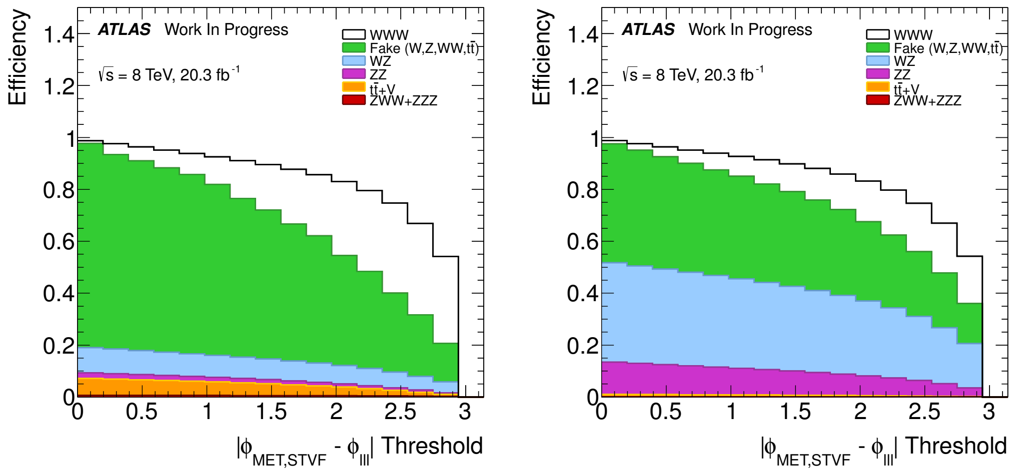

Figure 5.25: Signal and background efficiencies for the selection $|\Delta\varphi(lll, E_T^{\text{Miss}})| > X$ as a function of the $\Delta\varphi(lll, E_T^{\text{Miss}})$ selection threshold, X, in both the 0 SFOS (left) and pre-selection (right) regions.

The efficiencies as a function of the lepton p_{T} threshold are shown in Fig. 5.26. The signal efficiency is observed to be slightly flatter than the background efficiency. The signal efficiency, however, still falls fairly rapidly as a function of the lepton p_{T} threshold. Thus, a tighter selection on the lepton p_{T} is not preferred by the optimization. We also considered applying different p_{T} thresholds to the leptons based on their p_{T} order and other criteria, but this did not show any increased performance.

Finally, we considered other quantities like the transverse mass of the $E_{\text{T}}^{\text{miss}}$ and three lepton system:

$$m_T^{lll} = \sqrt{2p_T^{lll} E_{\text{T}}^{\text{miss}}(1 - \cos(\Delta\varphi(lll, E_{\text{T}}^{\text{miss}})))} \tag{5.40}$$

as well as vetoes on additional leptons with lower p_{T}, and various di-lepton mass selections. None of these, however, were preferred by the optimization.

5.6.3 Signal Region Yields

The optimized signal region selection described in Sec. 5.6.2 and Sec. 5.3.2 and listed in Table 5.5 is applied to the data as well as the signal plus background model. A plot of

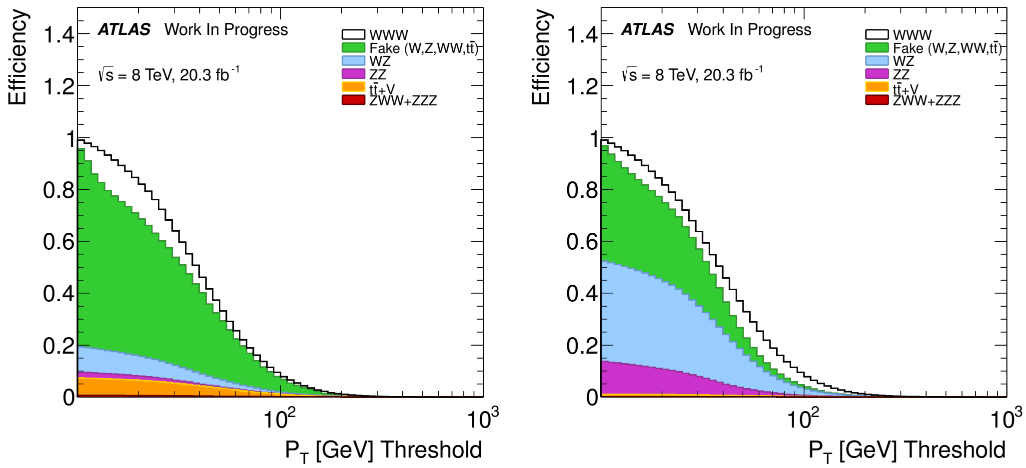

Figure 5.26: Signal and background efficiencies for the selection Lepton $p_{\mathrm{T}} > X$ as a function of the p_{T} selection threshold, X, in both the 0 SFOS (left) and pre-selection (right) regions.

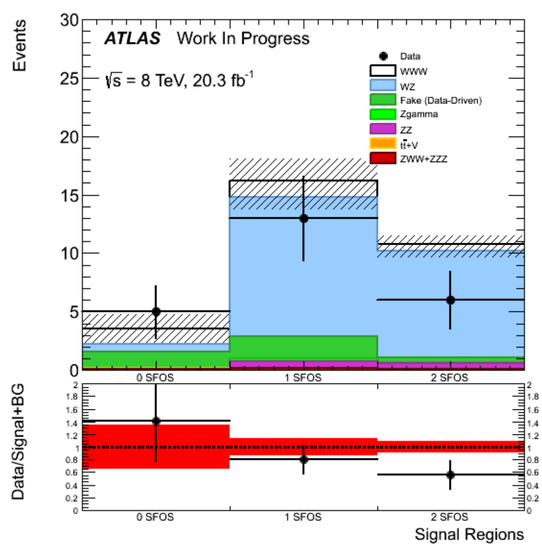

Figure 5.27: Yields after full selection in the 0, 1 and 2 SFOS regions. The most important systematic uncertainties are shown, namely from the fake estimates and the uncertainties on the WZ and ZZ k-factors.

the predicted yields for the signal plus background, along with systematic uncertainties, is compared to the data for each signal region in Fig. 5.27. A detailed breakdown of the predicted yields and overall uncertainties on each background as well as the signal prediction and observed data are presented in Table 5.16. A breakdown of the systematic uncertainty contributions to the signal and the backgrounds in each signal region are summarized in Table 5.17. More details are presented about each signal region below.

			$p_T \in [20, 30]$ GeV	$p_T \in [30, 50]$ GeV	$p_T > 50$ GeV
Tight	Data		44.0 ± 6.6	53.0 ± 7.3	77.0 ± 8.8
	Real	WWW	0.0422 ± 0.0028	0.0642 ± 0.0035	0.1008 ± 0.0045
		VVV	0.0064 ± 0.0024	0.0208 ± 0.0044	0.0320 ± 0.0055
		Top	4.39 ± 0.59	8.91 ± 0.86	25.3 ± 1.8
		ZZ	0.380 ± 0.034	0.605 ± 0.042	0.932 ± 0.053
		WZ	4.70 ± 0.48	7.02 ± 0.59	12.92 ± 0.82
		WW	0.0 ± 0	0.171 ± 0.066	0.46 ± 0.13
		Z	0.0 ± 0	0.0 ± 0	0.0 ± 0
		Ztautau	0.0 ± 0	0.42 ± 0.42	0.0063 ± 0.0063
		Wgamma	0.0 ± 0	0.0 ± 0	0.0 ± 0
		W+Jets	0.0 ± 0	0.0 ± 0	0.0 ± 0
		Sum	9.52 ± 0.76	17.2 ± 1.1	39.7 ± 2
	PC	WWW	0.0 ± 0	0.0 ± 0	0.00013 ± 0.00013
		VVV	0.0 ± 0	0.0 ± 0	0.0012 ± 0.0012
		Top	2.3 ± 0.5	5.79 ± 0.86	19.7 ± 1.7
		ZZ	0.0 ± 0	0.0029 ± 0.0029	0.00096 ± 0.00096
		WZ	0.0 ± 0	0.0 ± 0	0.0 ± 0
		WW	0.027 ± 0.027	0.016 ± 0.016	0.145 ± 0.074
		Z	1.26 ± 0.38	0.65 ± 0.25	0.7 ± 0.3
		Ztautau	0.0 ± 0	0.73 ± 0.73	0.0 ± 0
		Wgamma	2.80 ± 0.71	3.40 ± 0.76	6.0 ± 1.1
		W+Jets	0.0 ± 0	4.3 ± 4.3	0.0 ± 0
		Sum	6.37 ± 0.95	14.9 ± 4.5	26.6 ± 2
	Data - Real - PC		28.1 ± 6.7	20.9 ± 8.7	10.6 ± 9.2
All	Data		662 ± 26	450 ± 21	297 ± 17
	Real	WWW	0.0737 ± 0.0038	0.0907 ± 0.0041	0.1231 ± 0.0049
		VVV	0.0064 ± 0.0024	0.0235 ± 0.0046	0.0322 ± 0.0055
		Top	9.99 ± 0.98	14.9 ± 1.2	32.8 ± 2.2
		ZZ	0.682 ± 0.049	0.86 ± 0.05	1.073 ± 0.056
		WZ	10.17 ± 0.71	10.99 ± 0.75	16.46 ± 0.94
		WW	0.76 ± 0.17	1.31 ± 0.22	1.58 ± 0.24
		Z	0.0 ± 0	0.0 ± 0	0.14 ± 0.14
		Ztautau	0.0 ± 0	1.6 ± 1.3	0.0063 ± 0.0063
		Wgamma	0.31 ± 0.24	0.0 ± 0	0.0 ± 0
		W+Jets	0.0 ± 0	0.0 ± 0	5.0 ± 5
		Sum	22.0 ± 1.3	29.8 ± 1.9	57.2 ± 5.5
	PC	WWW	0.0008 ± 0.0004	0.00013 ± 0.00013	0.00084 ± 0.00039
		VVV	0.0 ± 0	0.0 ± 0	0.0026 ± 0.0019
		Top	14.7 ± 2.3	21.3 ± 1.9	35.0 ± 2.9
		ZZ	0.0038 ± 0.0038	0.0059 ± 0.0036	0.0015 ± 0.0011
		WZ	0.061 ± 0.052	0.078 ± 0.069	0.06 ± 0.03
		WW	0.243 ± 0.082	0.54 ± 0.13	0.60 ± 0.14
		Z	22.0 ± 2.9	9.3 ± 2.2	7.6 ± 1.8
		Ztautau	2.1 ± 1.3	2.7 ± 2.1	1.4 ± 1.4
		Wgamma	41.0 ± 2.8	29.7 ± 2.3	25.2 ± 2.1
		W+Jets	48 ± 19	33 ± 14	40 ± 14
		Sum	128 ± 19	97 ± 14	109 ± 15
	Data - Real - PC		512 ± 32	324 ± 26	130 ± 23
Rate			0.055 ± 0.014	0.065 ± 0.027	0.082 ± 0.072

Table 5.14: Calculation of fake rate for electrons when $N_{b-jet} > 0$. Goes outside margins!

		$p_T \in [20, 30]$ GeV	$p_T \in [30, 40]$ GeV	$p_T > 40$ GeV
Tight	Data	48.0 ± 6.9	23.0 ± 4.8	63.0 ± 7.9
	Real WWW	0.0704 ± 0.0037	0.0610 ± 0.0034	0.1991 ± 0.0063
	Real VVV	0.0092 ± 0.0029	0.010 ± 0.003	0.0301 ± 0.0053
	Real Top	3.96 ± 0.55	4.00 ± 0.46	15.67 ± 0.88
	Real ZZ	0.34 ± 0.03	0.191 ± 0.022	0.493 ± 0.037
	Real WZ	4.5 ± 0.4	3.51 ± 0.37	10.03 ± 0.63
	Real WW	0.0 ± 0	0.0 ± 0	0.0 ± 0
	Real Z	0.0 ± 0	0.0 ± 0	0.0 ± 0
	Real Ztautau	0.0 ± 0	0.0 ± 0	0.0 ± 0
	Real Wgamma	0.0 ± 0	0.0 ± 0	0.0 ± 0
	Real W+Jets	0.0 ± 0	0.0 ± 0	0.0 ± 0
	Real Sum	8.85 ± 0.68	7.78 ± 0.59	26.4 ± 1.1
	PC WWW	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC VVV	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC Top	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC ZZ	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC WZ	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC WW	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC Z	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC Ztautau	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC Wgamma	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC W+Jets	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC Sum	0.0 ± 0	0.0 ± 0	0.0 ± 0
	Data - Real - PC	39.1 ± 7	15.2 ± 4.8	36.6 ± 8
All	Data	1910 ± 44	750 ± 27	774 ± 28
	Real WWW	0.0790 ± 0.0039	0.0646 ± 0.0035	0.2074 ± 0.0064
	Real VVV	0.0127 ± 0.0034	0.0118 ± 0.0032	0.0306 ± 0.0053
	Real Top	4.53 ± 0.57	4.49 ± 0.49	16.77 ± 0.91
	Real ZZ	0.385 ± 0.033	0.220 ± 0.024	0.517 ± 0.037
	Real WZ	6.62 ± 0.51	4.52 ± 0.43	11.60 ± 0.69
	Real WW	0.0 ± 0	0.0 ± 0	0.0 ± 0
	Real Z	0.0 ± 0	0.0 ± 0	0.0 ± 0
	Real Ztautau	8.2 ± 2.8	4.2 ± 1.9	1.4 ± 1
	Real Wgamma	0.0 ± 0	0.0 ± 0	0.0 ± 0
	Real W+Jets	0.0 ± 0	0.0 ± 0	0.0 ± 0
	Real Sum	19.8 ± 2.9	13.5 ± 2	30.5 ± 1.5
	PC WWW	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC VVV	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC Top	0.094 ± 0.094	0.0 ± 0	0.0 ± 0
	PC ZZ	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC WZ	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC WW	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC Z	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC Ztautau	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC Wgamma	0.0 ± 0	0.0 ± 0	0.0 ± 0
	PC W+Jets	9.2 ± 9.2	0.0 ± 0	0.0 ± 0
	PC Sum	9.3 ± 9.2	0.0 ± 0	0.0 ± 0
	Data - Real - PC	1881 ± 45	737 ± 27	744 ± 28
Rate		0.0208 ± 0.0037	0.0207 ± 0.0066	0.049 ± 0.011

Table 5.15: Calculation of fake rate for muons when $N_{b-jet} > 0$. Goes outside margins!

	0 SFOS	1 SFOS	2 SFOS
WZ	$0.6176 \pm 0.0043\ ^{+0.0699}_{-0.0701}$	$11.89 \pm 0.14\ ^{+1.32}_{-1.29}$	$9.05 \pm 0.13\ ^{+0.99}_{-1.00}$
ZZ	$0.0658 \pm 0.0039\ ^{+0.0112}_{-0.0112}$	$0.581 \pm 0.016\ ^{+0.106}_{-0.105}$	$0.477 \pm 0.011\ ^{+0.095}_{-0.086}$
$WWZ + WZZ$	$0.1126 \pm 0.0099\ ^{+0.0146}_{-0.0117}$	$0.140 \pm 0.011\ ^{+0.015}_{-0.013}$	$0.0785 \pm 0.0080\ ^{+0.0097}_{-0.0106}$
$t\bar{t}$+V	$0.0388 \pm 0.0043\ ^{+0.0061}_{-0.0077}$	$0.0503 \pm 0.0048\ ^{+0.0074}_{-0.0089}$	$0.0239 \pm 0.0033\ ^{+0.0074}_{-0.0058}$
DPS	$0.0 \pm 0.0\ ^{+0.0}_{-0.0}$	$0.0088 \pm 0.0080\ ^{+0.0080}_{-0.0084}$	$0.023 \pm 0.016\ ^{+0.019}_{-0.029}$
$Z\gamma$	$0.0 \pm 0.0\ ^{+0.0}_{-0.0}$	$0.20 \pm 0.13\ ^{+0.29}_{-0.13}$	$0.110 \pm 0.096\ ^{+0.163}_{-0.288}$
Fake	$1.51 \pm 0.26\ ^{+1.40}_{-1.29}$	$1.90 \pm 0.34\ ^{+1.90}_{-1.77}$	$0.49 \pm 0.16\ ^{+0.47}_{-0.46}$
Signal	$1.344 \pm 0.015\ ^{+0.073}_{-0.079}$	$1.394 \pm 0.016\ ^{+0.073}_{-0.082}$	$0.611 \pm 0.010\ ^{+0.032}_{-0.036}$
Total Background	$2.35 \pm 0.26\ ^{+1.40}_{-1.30}$	$14.77 \pm 0.39\ ^{+2.36}_{-2.22}$	$10.25 \pm 0.23\ ^{+1.15}_{-1.22}$
Total Predicted	$3.69 \pm 0.26\ ^{+1.41}_{-1.30}$	$16.16 \pm 0.39\ ^{+2.33}_{-2.18}$	$10.86 \pm 0.23\ ^{+1.12}_{-1.19}$
Data	5	13	6

Table 5.16: A summary of the expected yields compared to data for all three signal regions. Statistical uncertainties are shown as a symmetric uncertainty on the central value. Systematic uncertainties are shown as an asymmetric uncertainty and are shown after taking the quadrature sum of all individual uncertainties. In the actual analysis, each systematic uncertainty is treated as an individual nuisance parameter and are NOT added in quadrature. The presentation here serves only as a demonstration of the overall size of the systematic uncertainties for each source in the individual signal regions.

Source of Uncertainty	Signal			Backgound		
	0 SFOS	1 SFOS	2 SFOS	0 SFOS	1 SFOS	2 SFOS
Electron	$^{+1.56}_{-1.47}$	$^{+1.66}_{-1.61}$	$^{+1.02}_{-1.06}$	$^{+0.68}_{-0.69}$	$^{+2.34}_{-1.49}$	$^{+1.05}_{-1.54}$
Muon	$^{+0.56}_{-0.54}$	$^{+0.54}_{-0.54}$	$^{+0.74}_{-0.83}$	$^{+0.19}_{-0.19}$	$^{+1.09}_{-0.48}$	$^{+0.81}_{-0.80}$
MET	$^{+1.38}_{-1.75}$	$^{+0.71}_{-0.89}$	$^{+0.23}_{-0.35}$	$^{+0.79}_{-0.73}$	$^{+1.38}_{-0.11}$	$^{+2.12}_{-2.66}$
Jet	$^{+2.36}_{-2.26}$	$^{+2.06}_{-2.34}$	$^{+1.56}_{-2.22}$	$^{+1.10}_{-1.06}$	$^{+2.74}_{-2.03}$	$^{+2.94}_{-4.41}$
Trigger	$^{+0.09}_{-0.09}$	$^{+0.09}_{-0.09}$	$^{+0.20}_{-0.20}$	$^{+0.06}_{-0.06}$	$^{+0.09}_{-0.09}$	$^{+0.21}_{-0.21}$
Matrix Method	—	—	—	$^{+58.56}_{-53.98}$	$^{+12.64}_{-11.78}$	$^{+4.34}_{-4.23}$
Charge Mis-ID	—	—	—	$^{+0.45}_{-0.44}$	—	—
Pileup	$^{+0.92}_{-0.77}$	$^{+1.10}_{-1.30}$	$^{+1.50}_{-1.24}$	$^{+0.52}_{-0.42}$	$^{+0.22}_{+0.00}$	$^{+1.39}_{-1.40}$
Luminosity	$^{+2.80}_{-2.80}$	$^{+2.80}_{-2.80}$	$^{+2.80}_{-2.80}$	$^{+2.80}_{-2.80}$	$^{+2.80}_{-2.80}$	$^{+2.80}_{-2.80}$
Theory	$^{+5.55}_{-3.75}$	$^{+5.55}_{-3.75}$	$^{+5.55}_{-3.75}$	$^{+2.66}_{-2.66}$	$^{+8.07}_{-8.07}$	$^{+8.85}_{-8.85}$
Statistical	$^{+1.14}_{-1.14}$	$^{+1.12}_{-1.12}$	$^{+1.70}_{-1.70}$	$^{+10.99}_{-10.99}$	$^{+2.67}_{-2.67}$	$^{+2.20}_{-2.20}$

Table 5.17: Categorized systematic uncertainties for signal and background predictions in all three signal regions. All uncertainties are shown as a percentage of the nominal prediction.

5.6.3.1 0 SFOS Signal Region

	Signal		Background		Data	
	Yield	Eff.	Yield	Eff.	Yield	Eff.
1. Pre-selection	9.78	—	2388.48	—	2472	—
2. 0 SFOS	2.31	0.24	21.36	0.0089	30	0.01
3. Charge Sum $= \pm 1$	2.30	1.00	19.55	0.92	27	0.90
4. $N_{\text{b-jet}} = 0$	2.29	0.99	8.59	0.44	10	0.37
5. $m_{SF} > 20$ GeV	2.25	0.98	8.32	0.97	10	1.00
6. $\lvert m_{ee} - m_Z \rvert > 15$ GeV	2.06	0.91	7.09	0.85	9	0.90
7. $\lvert \Delta\phi(3l, E_T^{Miss}) \rvert > 2.5$	1.41	0.69	2.51	0.35	6	0.67
8. $N_{\text{Jet}} \leq 1$	1.34	0.95	2.35	0.94	5	0.83

Table 5.18: Cut-flows showing the event yields and efficiencies for each cut in the 0 SFOS signal region starting from event pre-selection separately for the total signal and total background predictions, along with the observed data. Event yields for MC backgrounds and signal include all weights and are normalized to an integrated luminosity of 20.3 fb^{-1}. The fake lepton background only includes the matrix method weights. The data is unweighted. Efficiencies show the ratio of the yield with respect to the previous cut. The efficiency is first calculated at the first cut after event pre-selection.

The prediction from the 0 SFOS signal region at each stage of the selection is summarized in Table 5.18 for the signal and background predictions, as well as for the data. There is also a more detailed set of predictions at each stage for the different background sources in Table 5.19. From this, we can clearly see the enormous impact of the 0 SFOS cut on removing the backgrounds, for the WZ background in particular. We can also see the strong impact that the $N_{b-\text{Jet}}$ and $\Delta\varphi(lll, E_T^{\text{Miss}})$ cuts have without removing much of the signal. We can also see the signal plus background predictions as compared to the data for the distribution just before each cut is applied in Fig. 5.28. From this we can clearly see that the data seems to be well modeled at each stage of the selection.

After the full selection is applied, the 0 SFOS signal region is found to be the most sensitive of the three channels, as expected, with a predicted signal to background ratio of 56%. This can be seen from Table 5.16, where the expected signal is 1.344 compared to an expected background of 2.35. Together they combine to give a total prediction of 3.69 signal plus background events. The Poisson probability of observing ≤ 5 events with 3.69 events expected from the signal plus background prediction is 30.7%. Thus, we can

	Background					
	WZ		ZZ		$t\bar{t}+V$	
	Yield	Eff.	Yield	Eff.	Yield	Eff.
1. Pre-selection	1566.91	—	323.60	—	36.93	—
2. 0 SFOS	2.84	0.002	0.50	0.002	0.26	0.01
3. Charge Sum = ±1	1.92	0.68	0.33	0.65	0.26	0.99
4. $N_{\text{b-jet}} = 0$	1.91	0.99	0.33	0.99	0.25	0.98
5. $m_{SF} > 20$ GeV	1.88	0.98	0.32	0.98	0.25	0.98
6. $\lvert m_{ee} - m_Z\rvert > 15$ GeV	1.27	0.68	0.21	0.66	0.22	0.90
7. $\lvert\Delta\phi(3l, E_T^{Miss})\rvert > 2.5$	0.65	0.51	0.07	0.34	0.09	0.38
8. $N_{\text{Jet}} \leq 1$	0.62	0.95	0.07	0.91	0.04	0.45
	Background					
	$ZZZ+ZWW$		$Z\gamma$		Fake	
	Yield	Eff.	Yield	Eff.	Yield	Eff.
1. Pre-selection	3.12	—	219.80	—	238.12	—
2. 0 SFOS	0.25	0.08	0.20	0.001	17.31	0.07
3. Charge Sum = ±1	0.25	1.00	0.00	0.00	16.79	0.97
4. $N_{\text{b-jet}} = 0$	0.25	0.99	0.00	0.00	5.85	0.35
5. $m_{SF} > 20$ GeV	0.24	0.98	0.00	0.00	5.63	0.96
6. $\lvert m_{ee} - m_Z\rvert > 15$ GeV	0.22	0.90	0.00	0.00	5.17	0.92
7. $\lvert\Delta\phi(3l, E_T^{Miss})\rvert > 2.5$	0.13	0.59	0.00	0.00	2.17	0.42
8. $N_{\text{Jet}} \leq 1$	0.11	0.86	0.00	0.00	1.51	0.70

Table 5.19: Cut-flows showing the event yields and efficiencies for each cut in the 0 SFOS signal region starting from event pre-selection and binned by background category. Event yields for MC backgrounds and signal include all weights and are normalized to an integrated luminosity of 20.3 fb^{-1}. The fake lepton background only includes the matrix method weights. The data is unweighted. Efficiencies show the ratio of the yield with respect to the previous cut. The efficiency is first calculated at the first cut after event pre-selection.

see that this is in good agreement with the observed 5 events in data from the statistical uncertainty alone.

The fake background makes up more than half of the total expected background prediction, with 1.51 background events predicted from fakes compared to 2.35 events expected from the total background. The systematic uncertainty on the fake background is approaching 100%. As can be seen in Table 5.17, this results in the systematic uncertainty on the total background estimate that is approaching 60%, or roughly the size of the fake background contribution. This further increases the compatibility of the data with the expectation, but also reduces the sensitivity.

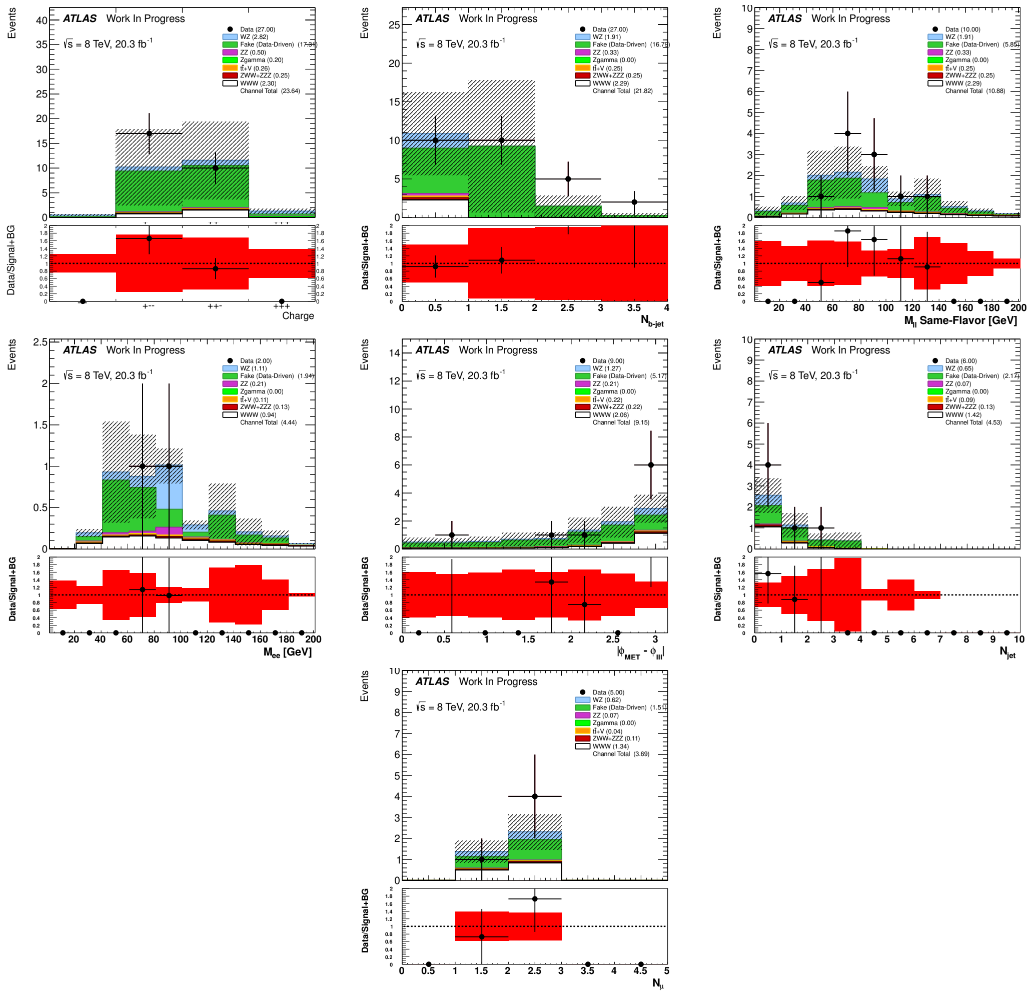

Figure 5.28: Distributions showing data compared to the signal plus background estimate in the 0 SFOS region at each stage of the selection before the cuts are applied to the given distribution. Plots should be read sequentially from left to right and from top to bottom. Referring to Table 5.18, the top left plot is shown before cut #3 is applied, top middle is before cut #5, and so on until the bottom right which is after all cuts are applied.

The other backgrounds are less important. The WZ background is the second largest, coming from charge mis-identification, with 0.6176 events predicted. The uncertainty on the WZ background is dominated by that from the WZ normalization uncertainty, which is 10%, and also has a small contribution from the charge mis-identification estimate uncertainty. The VVV contributions is the third largest, predicting 0.1126 with a small uncertainty. The ZZ background has a similar source and uncertainty as the WZ, but is about 10 times smaller in size. The $t\bar{t}+V$ background contributes even less and the DPS and $Z\gamma$ backgrounds have 0 contribution within the statistical uncertainties of the MC.

5.6.3.2 1 SFOS Signal Region

	Signal		Background		Data	
	Yield	Eff.	Yield	Eff.	Yield	Eff.
1. Pre-selection	9.78	—	2388.48	—	2472	—
2. 1 SFOS	4.67	0.48	1231.49	0.52	1260	0.51
3. $N_{\mathrm{b-jet}} = 0$	4.42	0.94	1086.66	0.88	1095	0.87
4. NOT $m_Z - 35$ GeV $<$ $m_{\mathrm{SFOS}} < m_Z + 20$ GeV	2.76	0.63	97.96	0.090	93	0.08
5. $E_T^{Miss} > 45$ GeV	1.91	0.69	29.83	0.30	27	0.29
6. $\lvert\Delta\phi(3l, E_T^{Miss})\rvert > 2.5$	1.48	0.77	16.73	0.56	16	0.59
7. $N_{\mathrm{Jet}} \leq 1$	1.39	0.94	14.77	0.88	13	0.81

Table 5.20: Cut-flows showing the event yields and efficiencies for each cut in the 1 SFOS signal region starting from event pre-selection separately for the total signal and total background predictions, along with the observed by data. Event yields for MC backgrounds and signal include all weights and are normalized to an integrated luminosity of 20.3 fb^{-1}. The fake lepton background only includes the matrix method weights. The data is unweighted. Efficiencies show the ratio of the yield with respect to the previous cut. The efficiency is first calculated at the first cut after event pre-selection.

The 1 SFOS signal region is not as sensitive as the 0 SFOS region, with a signal to background ratio of about 9.2%. The background is overwhelmingly dominated by WZ contributions. Similar to the 0 SFOS region, the predictions and data at each stage of the 1 SFOS signal region selection are shown in Table 5.20 and Table 5.21. The 1 SFOS requirement leaves much of the WZ and ZZ backgrounds, but the Z-veto and $E_{\mathrm{T}}^{\mathrm{miss}}$ cuts are very effective at removing most of this while keeping the signal.

Again, we can also see the signal plus background predictions as compared to the data for the distribution just before each cut is applied in the 1 SFOS region by looking at Fig. 5.29. Here, the distributions again appear to be well modeled at each stage of the selection. Looking closer at the N_{Jet} distribution, we can see that there is a deficit of data in the $N_{\mathrm{Jet}} = 1$ bin which is kept in the selection and results in a slight deficit in the prediction. Further, if we look at the N_{μ} distribution we see that this deficit seems to

	Background					
	WZ		ZZ		$t\bar{t}+V$	
	Yield	Eff.	Yield	Eff.	Yield	Eff.
1. Pre-selection	1566.91	—	323.60	—	36.93	—
2. 1 SFOS	757.38	0.48	171.39	0.53	18.10	0.49
3. $N_{\text{b-jet}} = 0$	696.90	0.92	150.14	0.88	1.42	0.08
4. NOT $m_Z - 35$ GeV $<$ $m_{\text{SFOS}} < m_Z + 20$ GeV	44.30	0.06	13.79	0.09	0.37	0.26
5. $E_T^{Miss} > 45$ GeV	21.38	0.48	1.46	0.11	0.29	0.78
6. $\lvert\Delta\phi(3l, E_T^{Miss})\rvert > 2.5$	13.07	0.61	0.71	0.49	0.11	0.39
7. $N_{\text{Jet}} \leq 1$	11.90	0.91	0.58	0.82	0.05	0.45

	Background					
	$ZZZ + ZWW$		$Z\gamma$		Fake	
	Yield	Eff.	Yield	Eff.	Yield	Eff.
1. Pre-selection	3.12	—	219.80	—	238.12	—
2. 1 SFOS	1.55	0.50	149.60	0.68	133.47	0.56
3. $N_{\text{b-jet}} = 0$	1.31	0.84	136.96	0.92	99.93	0.75
4. NOT $m_Z - 35$ GeV $<$ $m_{\text{SFOS}} < m_Z + 20$ GeV	0.34	0.26	22.44	0.16	16.72	0.17
5. $E_T^{Miss} > 45$ GeV	0.24	0.71	1.36	0.06	5.10	0.31
6. $\lvert\Delta\phi(3l, E_T^{Miss})\rvert > 2.5$	0.17	0.69	0.20	0.15	2.47	0.48
7. $N_{\text{Jet}} \leq 1$	0.14	0.84	0.20	1.00	1.90	0.77

Table 5.21: Cut-flows showing the event yields and efficiencies for each cut in the 1 SFOS signal region starting from event pre-selection and binned by background category. Event yields for MC backgrounds and signal include all weights and are normalized to an integrated luminosity of 20.3 fb^{-1}. The fake lepton background only includes the matrix method weights. The data is unweighted. Efficiencies show the ratio of the yield with respect to the previous cut. The efficiency is first calculated at the first cut after event pre-selection.

fall exclusively in the $N_\mu = 1$ bin. A more detailed investigation of the cut-flows in the individual $N_\mu = 1$ and $N_\mu = 2$ bins suggests that this is most likely a statistical fluctuation. Overall, the deficit is not very significant, with the Poisson probability of observing 13 or less events with 16.16 expected being 26.2%.

The fake background is only the second largest background in this region, making up about 13% of the total. Still, even with the 10% uncertainty on the normalization of the dominant WZ background, the fake background uncertainty is the largest uncertainty on the background estimation, approaching 13%, as can be seen in Table 5.17. The $t\bar{t}+V$ and VVV backgrounds are of a similar absolute size as in the 0 SFOS region, but the larger

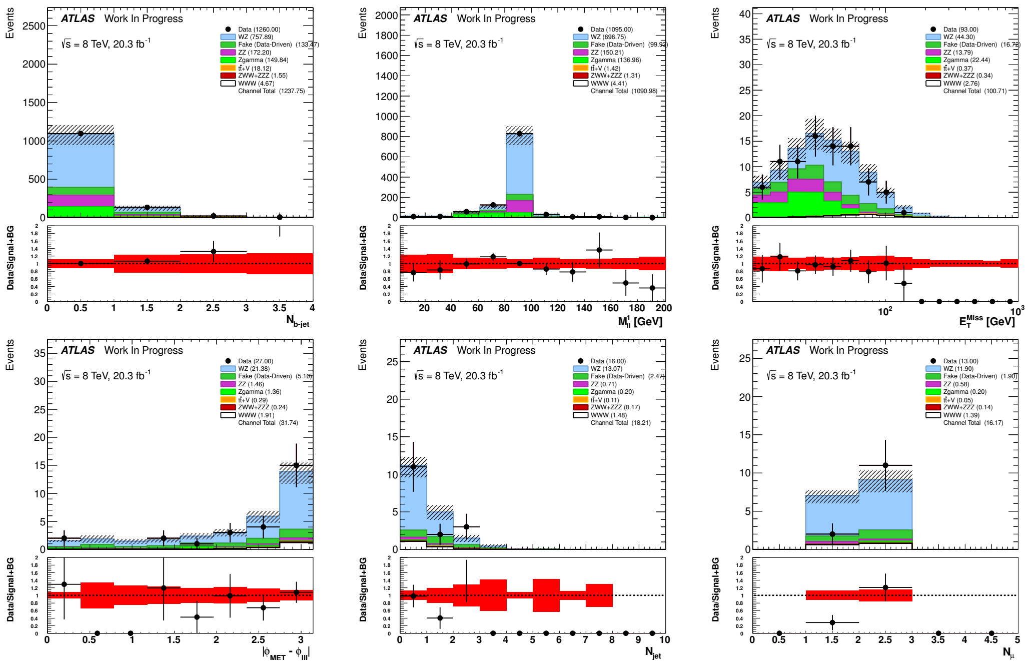

Figure 5.29: Distributions showing data compared to the signal plus background estimate in the 1 SFOS region at each stage of the selection before the cuts are applied to the given distribution. Plots should be read sequentially from left to right and from top to bottom. Referring to Table 5.20, the top left plot is shown before cut #3 is applied, top middle is before cut #4, and so on until the bottom right which is after all cuts are applied.

overall background makes them even less important. The DPS and $Z\gamma$ uncertainties contribute a finite amount to the background within the statistical uncertainties, but remain negligible.

5.6.3.3 2 SFOS Signal Region

The 2 SFOS signal region has a similar background composition as the 1 SFOS signal regions, since it is also dominated by the WZ background. As a result, the systematic uncertainties on the signal and background are very similar to the 1 SFOS region. As can be seen in Table 5.20 and Table 5.21, however, the overall background prediction is slightly smaller than the 1 SFOS signal region This is mainly because the tighter $E_{\mathrm{T}}^{\mathrm{miss}}$ cut removes more of the WZ background. The signal and the fake background also contributes slightly less to the total background but this is true immediately after applying the SFOS

	Signal Yield	Signal Eff.	Background Yield	Background Eff.	Data Yield	Data Eff.
1. Pre-selection	9.78	—	2388.48	—	2472	—
2. 2 SFOS	2.66	0.27	1132.53	0.47	1182	0.48
3. $N_{\text{b-jet}} = 0$	2.50	0.94	1012.07	0.89	1033	0.87
4. $\|m_{\text{SFOS}} - m_Z\| > 20$ GeV	1.46	0.58	108.88	0.11	108	0.10
5. $E_T^{Miss} > 55$ GeV	0.83	0.57	18.99	0.17	18	0.17
6. $\|\Delta\phi(3l, E_T^{Miss})\| > 2.5$	0.65	0.78	11.64	0.61	8	0.44
7. $N_{\text{Jet}} \leq 1$	0.61	0.94	10.25	0.88	6	0.75

Table 5.22: Cut-flows showing the event yields and efficiencies for each cut in the 2 SFOS signal region starting from event pre-selection separately for the total signal and total background predictions, along with the observed data. Event yields for MC backgrounds and signal include all weights and are normalized to an integrated luminosity of 20.3 fb^{-1}. The fake lepton background only includes the matrix method weights. The data is unweighted. Efficiencies show the ratio of the yield with respect to the previous cut. The efficiency is first calculated at the first cut after event pre-selection.

	Background					
	WZ		ZZ		$t\bar{t}+V$	
	Yield	Eff.	Yield	Eff.	Yield	Eff.
1. Pre-selection	1566.91	—	323.60	—	36.93	—
2. 2 SFOS	807.27	0.52	151.28	0.47	15.35	0.42
3. $N_{\text{b-jet}} = 0$	743.12	0.92	136.16	0.90	1.19	0.08
4. $\|m_{\text{SFOS}} - m_Z\| > 20$ GeV	44.95	0.06	21.13	0.16	0.22	0.18
5. $E_T^{Miss} > 55$ GeV	15.86	0.35	0.97	0.05	0.14	0.65
6. $\|\Delta\phi(3l, E_T^{Miss})\| > 2.5$	10.09	0.64	0.55	0.57	0.07	0.49
7. $N_{\text{Jet}} \leq 1$	9.07	0.90	0.48	0.86	0.02	0.35

	Background					
	$ZZZ+ZWW$		$Z\gamma$		Fake	
	Yield	Eff.	Yield	Eff.	Yield	Eff.
1. Pre-selection	3.12	—	219.80	—	238.12	—
2. 2 SFOS	1.30	0.41	69.99	0.32	87.34	0.37
3. $N_{\text{b-jet}} = 0$	1.10	0.85	64.70	0.92	65.80	0.75
4. $\|m_{\text{SFOS}} - m_Z\| > 20$ GeV	0.19	0.17	29.52	0.46	12.87	0.20
5. $E_T^{Miss} > 55$ GeV	0.12	0.63	0.43	0.01	1.47	0.11
6. $\|\Delta\phi(3l, E_T^{Miss})\| > 2.5$	0.10	0.82	0.11	0.25	0.72	0.49
7. $N_{\text{Jet}} \leq 1$	0.08	0.82	0.11	1.00	0.49	0.69

Table 5.23: Cut-flows showing the event yields and efficiencies for each cut in the 2 SFOS signal region starting from event pre-selection and binned by background category. Event yields for MC backgrounds and signal include all weights and are normalized to an integrated luminosity of 20.3 fb^{-1}. The fake lepton background only includes the matrix method weights. The data is unweighted. Efficiencies show the ratio of the yield with respect to the previous cut. The efficiency is first calculated at the first cut after event pre-selection.

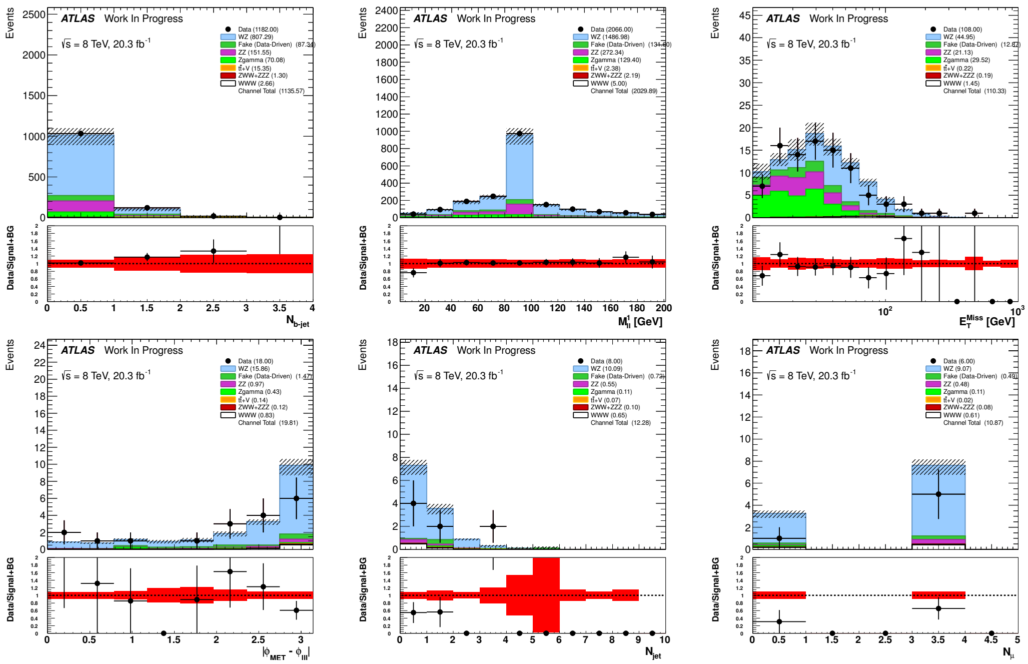

Figure 5.30: Distributions showing data compared to the signal plus background estimate in the 2 SFOS region at each stage of the selection before the cuts are applied to the given distribution. Plots should be read sequentially from left to right and from top to bottom. Referring to Table 5.22, the top left plot is shown before cut #3 is applied, the top middle is before cut #4, and so on until the bottom right which is after all cuts are applied.

requirement, and not due to the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut. The reason can be understood as described in Sec. 5.3.2: there are twice many charge and flavor combinations to produce 1 SFOS pairs as there are 2 SFOS pairs.

From the cut-flow tables we can also see that there is a deficit in the data compared to the prediction which appears after the $\Delta\varphi(lll, E_T^{\mathrm{Miss}})$ selection. Looking at the distributions at each cut for the 2 SFOS region in Fig. 5.30, one can clearly see the deficit occurring in the bin furthest to the right in the $|\Delta\varphi(lll, E_T^{\mathrm{Miss}})|$ distribution. The deficit then propagates through uniformly in the N_{Jet} and N_μ distributions until the final estimate. Note that the bin where the deficit occurs in the $|\Delta\varphi(lll, E_T^{\mathrm{Miss}})|$ distribution is also dominated by the WZ background. We have verified the modeling of the WZ background as a function of this quantity in control regions. Furthermore, the $|\Delta\varphi(lll, E_T^{\mathrm{Miss}})|$ distribution shows good agreement in the 1 SFOS region at this stage where it is also dominated by the WZ

background. We have no reason to believe that the modeling of the WZ background should be very different or should break down in the 2 SFOS region as compared to elsewhere. Thus, the deficit is most likely a statistical fluctuation and not due to a problem in the modeling of the background. The Poisson probability of observing ≤ 6 events when 10.86 events are expected is 8.5%. Thus, even though this is the largest deviation observed in the signal regions, it is still within 2 standard deviations (5%).

5.6.4 Correction Factors and Fiducial Cross-sections

The correction factor, C_i, is defined for each channel, i, as the ratio of the number of expected signal events measured at the reconstruction level, N_i^{Reco}, over the number expected from truth information, N_i^{Truth}.

$$\varepsilon_i = \frac{N_i^{\text{Reco}}}{N_i^{\text{Truth}}} \tag{5.41}$$

N_i^{Reco} is determined using the reconstruction level selection described in Sec. 5.3.2 and listed in Table 5.5; N_i^{Truth} is determined using the fiducial selection described in Sec. 5.3.3 and listed in Table 5.6. The same generator, VBFNLO, is used for both to remove any dependence on the cross-section or other generator specific effects.

The fiducial cross-sections are calculated also using the selection from the Sec. 5.3.2 after weighting to the cross-section for the given sample. Recall also from Sec. 5.1.2.1 and Table 5.2 that the fiducial cross-sections were generated using both MadGraph and VBFNLO and were shown to be in good agreement. The fiducial cross-sections from MadGraph are used in the final estimates.

Channel	C_i	Fiducial Cross-section [ab]
0 SFOS	$0.534 \pm .021$	123.6 ± 4.7
1 SFOS	$0.500 \pm .018$	136.9 ± 4.7
2 SFOS	$0.615 \pm .038$	48.8 ± 2.9

Table 5.24: Correction factors, C_i, and fiducial cross-sections derived separately for each signal region. Correction factors are determined using VBFNLO ; fiducial cross-sections are determined using MadGraph.

The correction factors and fiducial cross-sections are summarized separately for each signal region in Table 5.24. Note that the sum of the fiducial cross-sections in each signal region gives the combined fiducial cross-section which was reported in Eq. 5.2 along with PDF and scale uncertainties.

5.7 Standard Model Measurement

In this analysis we seek to measure the fiducial cross-section, σ^{Observed}, for the WWW production process in the fully-leptonic channel (e,μ). The observed cross-section is parameterized by looking at the signal strength, μ, which is related to the expected fiducial cross-sections from section 5.6.4 by the relation:

$$\sigma^{\text{Observed}} = \mu \sum_{i \in \text{Channels}} \sigma_i^{\text{Fiducial}} \tag{5.42}$$

Assuming a counting experiment in each bin i, the expected event count is given by:

$$N_i^{\text{exp}}(\mu, \boldsymbol{\theta}) = N_i^{\text{exp}}(\mu, \mathcal{L}_0, \Delta_{\mathcal{L}}, \boldsymbol{\theta}_s, \boldsymbol{\theta}_b) = \mu \cdot \left(\mathcal{L}(\mathcal{L}_0, \Delta_{\mathcal{L}}) \cdot \sigma_i^{\text{Fiducial}} \cdot C_i(\boldsymbol{\theta}_s) \right) + \sum_{\text{bkg}} N_i^{\text{bkg}}(\boldsymbol{\theta}_b) \tag{5.43}$$

where C_i is the correction factor measured in each bin as discussed in section 5.6.4 and $\sigma_i^{\text{Fiducial}}$ is the fiducial cross-section in each bin. The individual background expectations in a given bin/channel, i, are expressed simply by the number of events for a given background as N_i^{bkg}. The signal efficiencies and background expectations are assumed to follow probability distributions described by shape parameters determined from dedicated measurements of the background normalizations and systematic uncertainties. The set of correction factor shape parameters are referred to as $\boldsymbol{\theta}_s$; the set of normalization and shape parameters on the background expectations are referred to as $\boldsymbol{\theta}_b$. The integrated luminosity, $\mathcal{L}$, is assumed to follow a Gaussian distribution with nominal integrated luminosity, $\mathcal{L}_0$, and width, $\Delta_{\mathcal{L}}$. Collectively, we refer to all of these parameters, except for μ as the set of nuisance parameters, $\boldsymbol{\theta} = (\mathcal{L}_0, \Delta_{\mathcal{L}}, \boldsymbol{\theta}_s, \boldsymbol{\theta}_b)$.

The discovery significance is tested using frequentist statistics to estimate the degree of compatibility with the background only hypothesis [34]. The measurement and uncertainty are evaluated by using the shape of the profile likelihood ratio [51] which is a function of the data and the signal strength.

5.7.1 Profile Likelihood Ratio

The likelihood used is constructed as follows:

$$L(\mu, \boldsymbol{\theta}) = \mathrm{Gaus}(\mathcal{L}; \mathcal{L}_0, \Delta_{\mathcal{L}}) \prod_{i \in \mathrm{Chan}} \mathrm{Pois}(N_i^{obs} | N_i^{exp}(\mu, \boldsymbol{\theta})) \prod_{j \in \mathrm{Sys}} \mathrm{Gaus}(\theta_j; \theta_j^0, 1) \tag{5.44}$$

using the HistFactory tool developed within ATLAS [35]. Note that the systematic uncertainties are given Gaussian constraints with $\pm 1\sigma$ uncertainties.

The basic form of the test statistic used for comparing hypotheses is called the profile likelihood ratio, $\lambda(\mu)$ and is defined as:

$$-2 \ln \lambda(\mu) = -2 \ln \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{5.45}$$

Note that it no longer depends on the nuisance parameters, $\boldsymbol{\theta}$, and instead depends only on μ. The negative of twice the logarithm of the profile likelihood ratio is used because the logarithm is monotonic and typically easier to work with. The presence of the nuisance parameters are handled in the profiling step when constructing the profile likelihood ratio, which results in a smearing of the profile likelihood ratio contour. During profiling, the systematic uncertainties are interpolated using a piecewise linear function for shape uncertainties and a piecewise exponential function for the normalization uncertainties in order to maintain a normalization that is greater than zero. The denominator is the unconditional maximum likelihood (ML) evaluated at the ML estimators $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$. This quantity is a unique constant when specified for a given likelihood and set of nuisance parameters. The numerator is the conditional ML which depends on μ and evaluated at the conditional

ML estimator for the set of nuisance parameters, $\hat{\hat{\boldsymbol{\theta}}}$, which itself depends on μ. Clearly, the profile likelihood ratio runs from $0 < \lambda(\mu) < 1$ with values close to 0 showing more agreement with the background only hypothesis and values closer to 1 showing more agreement with the signal hypothesis, μ. When taking the negative log likelihood, the range is mapped to the entire positive axis and inverted. This means that values close to 0 are more background-like and larger values are more-signal like.

The minimum of the negative log of the profile likelihood is taken as the measurement of the signal strength; the uncertainty on the measurement is taken from the shape of the negative log profile likelihood assuming the behavior in the asymptotic limit can be used. The asymptotic behavior of the profile likelihood is used to evaluate the final confidence interval.

5.7.2 Testing for Discovery Significance

The rejection of the background-only hypothesis ($\mu = 0$) is used to estimate the significance of a possible observation of the signal. For the purposes of this test, the following test statistic is used:

$$q_0 = \begin{cases} -2\ln\lambda(0), & \hat{\mu} \geq 0 \\ 0, & \hat{\mu} < 0 \end{cases} \tag{5.46}$$

The test statistic is set to 0 when $\hat{\mu} < 0$ to enforce the notion that an observation which is less than the background expectation should not be treated as signal like. The p-value in this case tells us the degree of incompatibility with the background only hypothesis and is defined as:

$$p_0 = \int_{q_{0,\mathrm{obs}}}^{\infty} f(q_0|\mu = 0)\,\mathrm{d}q_0 \tag{5.47}$$

where $q_{0,\mathrm{obs}}$ is the observed value of q_0 and $f(q_0|\mu = 0)$ is the probability density of the test statistic q_0 under the background only hypothesis which is evaluated using toy MC. By examining the p-value one can say what the probability is that the deviation away from the background only hypothesis is due to chance. A small probability suggests that such a fluctuation is unlikely. Frequently one refers to the significance:

$$Z = \Phi^{-1}(1 - p_0) \tag{5.48}$$

where Φ^{-1} is the inverse of the Gaussian cumulative distribution function. In this way, one may refer to $Z\sigma$ significance of a measurement where usually 3σ is considered to constitute 'evidence' and 5σ constitutes discovery.

The distribution of q_0 is shown in Fig. 5.31 for the combination. The observed null p-value is found to be 0.24 for the combination which corresponds to a significance of 0.70σ. One may compare to this to an expected p-value of 0.25 corresponding to a significance of 0.66σ.

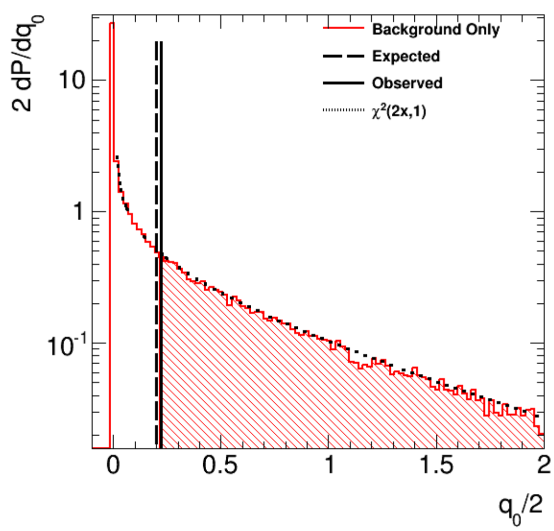

Figure 5.31: PDF of the background only hypothesis as a function of q_0 for the combination of all three channels. PDFs are determined using toy MC. The solid black line represents the observed value of q_0 seen in the data. The shaded area above this line represents the null p-value or the integral of the background hypothesis in the signal-like region. The dotted black curve shows a χ^2 distribution for 1 degree of freedom with which it can be seen is a good approximation of the the background only PDF.

5.7.3 Measurement and Uncertainty using Profile Likelihood Interval

The measured value of the signal strength is determined by looking at the minimum of the negative log profile likelihood for each channel separately and also for the combination of all channels. The size of the uncertainty on the measurement is taken by looking at the shape of the negative log profile likelihood contour which in general should follow a parabolic shape centered about the minimum in the asymptotic limit. In this limit, Wilk's theorem [56] can be used [51] to determine that the range of the uncertainty for a given number of Gaussian σ can be related directly to the negative profile log likelihood. In particular, for a 1σ uncertainty, where 68.3% of experiments will fall, one expects that $|-\ln\lambda(\mu)| \leq 1/2$. Note that even if the contour is not distributed symmetrically about the minimum value, invariance of the likelihood under transformations like $g(\hat{\mu}, \hat{\boldsymbol{\theta}})$ where g is some function, means the same conclusion still holds. The value of μ is not forced to be only positive and is left unrestricted.

The profile likelihood contour is evaluated once without systematic uncertainties included as nuisance parameters in order to estimate the size of the measurement uncertainty purely from statistical effects and then a second time with the systematic uncertainties included as nuisance parameters whose errors are constrained to be Gaussian and then profiled out. The contour with systematic uncertainties included represent the total uncertainty and the systematic uncertainty is determined by assuming that the total uncertainty is formed from the statistical and systematic uncertainties being added in quadrature. The negative log likelihood contour is for the combination of all three channels in Fig. 5.32. The expected value and uncertainties for the fiducial cross-section is:

$$\sigma^{\text{Expected}} = 309.2^{+434}_{-338}(\text{stat})^{+316}_{-342}(\text{sys})\text{ab} \tag{5.49}$$

and the observed fiducial cross-section is:

$$\sigma^{\text{Observed:}} = 315.1^{+347}_{-334}(\text{stat})^{+326}_{-348}(\text{sys})\text{ab} \tag{5.50}$$

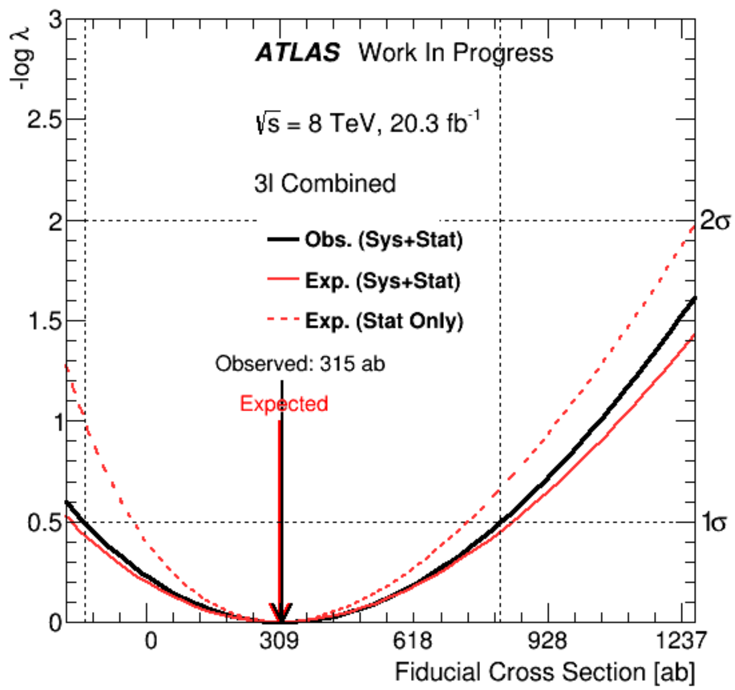

Figure 5.32: The profile likelihood contours evaluated as a function of the signal strength for the combination of all three channels. The observed (black) and expected (red) contours are shown when considering only statistical uncertainty (dashed line) and when considering both statistical and systematic uncertainties (solid line). The dotted black lines pinpoint the location of the 1 σ and 2 σ total Gaussian uncertainties on the measurement of the signal strength which corresponds to the minimum value of the contour.

5.8 Limits on anomalous Quartic Gauge Couplings

blank

Chapter 6

Conclusions

blank

List of Journal Abbreviations

Nucl. Phys. B	Nuclear Physics B: Particle physics, field theory and statistical systems, physical mathematics

Bibliography

[1] Background studies for top-pair production in lepton plus jets final states in $\sqrt{s} = 7TeV$ ATLAS data. Technical Report ATLAS-CONF-2010-087, CERN, Geneva, Oct 2010.

[2] Measurement of WZ production rate. Technical Report CMS-PAS-SMP-12-006, CERN, Geneva, 2013.

[3] Pile-up subtraction and suppression for jets in ATLAS. Technical Report ATLAS-CONF-2013-083, CERN, Geneva, Aug 2013.

[4] Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment. Technical Report ATLAS-CONF-2014-004, CERN, Geneva, Feb 2014.

[5] Measurement of the $t\bar{t}W$ and $t\bar{t}Z$ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. Technical Report ATLAS-CONF-2015-032, CERN, Geneva, Jul 2015.

[6] ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11. 2011.

[7] Georges Aad et al. Measurement of WZ production in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Eur.Phys.J.*, C72:2173, 2012.

[8] Georges Aad et al. Improved luminosity determination in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector at the LHC. *Eur.Phys.J.*, C73(8):2518, 2013.

[9] Georges Aad et al. Measurement of hard double-parton interactions in $W(\rightarrow l\nu)$+ 2 jet events at $\sqrt{s}$=7 TeV with the ATLAS detector. *New J.Phys.*, 15:033038, 2013.

[10] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl.Instrum.Meth.*, A506:250–303, 2003.

[11] Simone Alioli et al. A general framework for implementing NLO calculations in shower Monte Carlo programs: the Powheg BOX. *JHEP*, 1006:043, 2010.

[12] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. NLO vector-boson production matched with shower in POWHEG. *JHEP*, 0807:060, 2008.

[13] Johan Alwall, Pavel Demin, Simon De Visscher, Rikkert Frederix, Michel Herquet, Fabio Maltoni, Tilman Plehn, David L. Rainwater, and Tim Stelzer. Madgraph/madevent v4: The new web generation. *JHEP 0709*, page 07062334.

[14] P Anger, K Bachas, S Barnes, S Cole, L Di Ciaccio, S Gkaitatzis, S Hassani, D Iliadis, H Keoshkerian, M Kobel, J F Laporte, M Lefebvre, J Manjarres, A Melzer, A Oh, A Ouraou, C Petridou, T Sandmann, E Sauvan, U Schnoor, F Socher, W Taylor, and A Vest. A Measurement of WZ Production in Proton-Proton Collisions at $\sqrt{\mathbf{s}} = \mathbf{8}$ TeV with the ATLAS Detector. Technical Report ATL-COM-PHYS-2014-144, CERN, Geneva, Feb 2014.

[15] J-F Arguin, D Cote, O A Ducu, M Fiascaris, L Gauthier, T P S Gillam, G Herten, M Kruse, J Maurer, T J Mueller, M Pagacova, A Paramonov, J Pilcher, R Povey, T Rave, J-E Sundermann, and C Zhou. Search for strongly-produced superpartners in final states with two same-sign leptons or three leptons at sqrt(s) = 8 TeV. Technical Report ATL-COM-PHYS-2013-887, CERN, Geneva, Jun 2013.

[16] K. Arnold, J. Bellm, G. Bozzi, M. Brieg, F. Campanario, et al. VBFNLO: A Parton Level Monte Carlo for Processes with Electroweak Bosons – Manual for Version 2.5.0. 2011.

[17] K. Arnold, J. Bellm, G. Bozzi, F. Campanario, C. Englert, et al. Release Note – Vbfnlo-2.6.0. 2012.

[18] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *J. Instrum.*, 3:S08003. 437 p, 2008.

[19] ATLAS Collaboration. ATLAS tunes of Pythia6 and Pythia8 for MC11. ATL-PHYS-PUB-2011-009, 2011.

[20] ATLAS Collaboration. Measurement of the $Z\gamma$ and $Z\gamma\gamma$ Production Cross Section in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector at the LHC. Technical Report ATL-COM-PHYS-2013-1573, CERN, Geneva, Nov 2013.

[21] ATLAS Collaboration. Measurements of $W\gamma$ and $Z\gamma$ production in pp collisions at $\sqrt{s}$=7TeV with the ATLAS detector at the LHC. *Phys.Rev.*, D87(11):112003, 2013.

[22] ATLAS Collaboration. Measurements of Four-Lepton Production at the Z Resonance in pp Collisions at $\sqrt{s}$ =7 and 8 TeV with ATLAS. *Phys.Rev.Lett.*, 112(23):231806, 2014.

[23] Julien Baglio, Le Duc Ninh, and Marcus M. Weber. Massive gauge boson pair production at the LHC: a next-to-leading order story. *Phys.Rev.*, D88:113005, 2013.

[24] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.

[25] Anastasiya Bierweiler, Tobias Kasprzik, and Johann H. Khn. Vector-boson pair production at the LHC to $\mathcal{O}(\alpha^3)$ accuracy. *JHEP*, 1312:071, 2013.

[26] T. Binoth, M. Ciccolini, N. Kauer, and M. Kramer. Gluon-induced W-boson pair production at the LHC. *JHEP*, 0612:046, 2006.

[27] T. Binoth, N. Kauer, and P. Mertsch. Gluon-induced QCD corrections to pp —¿ ZZ —¿ l anti-l l-prime anti-l-prime. page 142, 2008.

[28] Michiel Botje, Jon Butterworth, Amanda Cooper-Sarkar, Albert de Roeck, Joel Feltesse, et al. The PDF4LHC Working Group Interim Recommendations. 2011.

[29] W. Buchmuller and D. Wyler. Effective Lagrangian Analysis of New Interactions and Flavor Conservation. *Nucl.Phys.*, B268:621–653, 1986.

[30] Andy Buckley, Jonathan Butterworth, Leif Lonnblad, David Grellscheid, Hendrik Hoeth, James Monk, Holger Schulz, and Frank Siegert. Rivet user manual. *Comput. Phys. Commun.*, 184:2803–2819, 2013.

[31] J. M. Butterworth, Jeffrey R. Forshaw, and M. H. Seymour. Multiparton interactions in photoproduction at hera. *Z. Phys. C*, 72:637, 1996.

[32] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 04:063, 2008.

[33] F. Cascioli, T. Gehrmann, M. Grazzini, S. Kallweit, P. Maierhfer, et al. ZZ production at hadron colliders in NNLO QCD. *Phys.Lett.*, B735:311–313, 2014.

[34] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71(arXiv:1007.1727):1554, Jul 2010.

[35] Kyle Cranmer, George Lewis, Lorenzo Moneta, Akira Shibata, and Wouter Verkerke. HistFactory: A tool for creating statistical models for use with RooFit and RooStats. Technical Report CERN-OPEN-2012-016, New York U., New York, Jan 2012.

[36] O.J.P. Eboli, M.C. Gonzalez-Garcia, and J.K. Mizukoshi. p p -¿ j j e+/- mu+/- nu nu and j j e+- mu-+ nu nu at O(alpha(em)**6) and O(alpha(em)**4 alpha(s)**2) for the study of the quartic electroweak gauge boson vertex at CERN LHC. *Phys.Rev.*, D74:073005, 2006.

[37] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 0711:070, 2007.

[38] Jonathan R. Gaunt, Chun-Hay Kom, Anna Kulesza, and W. James Stirling. Same-sign W pair production as a probe of double parton scattering at the LHC. *Eur.Phys.J.*, C69:53–65, 2010.

[39] Thomas P. S. Gillam and Christopher G. Lester. Improving estimates of the number of 'fake' leptons and other mis-reconstructed objects in hadron collider events: BoB's your UNCLE. *JHEP*, 11:031, 2014.

[40] T. Gleisberg et al. Event generation with SHERPA 1.1. *JHEP*, 0902:007, 2009.

[41] Piotr Golonka and Zbigniew Was. PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays. *Eur.Phys.J.*, C45:97–107, 2006.

[42] Massimiliano Grazzini, Stefan Kallweit, and Dirk Rathlev. $W\gamma$ and $Z\gamma$ production at the LHC in NNLO QCD. 2015.

[43] Marco Guzzi, Pavel Nadolsky, Edmond Berger, Hung-Liang Lai, Fredrick Olness, et al. CT10 parton distributions and other developments in the global QCD analysis. 2011.

[44] Kaoru Hagiwara, S. Ishihara, R. Szalapski, and D. Zeppenfeld. Low-energy effects of new interactions in the electroweak boson sector. *Phys.Rev.*, D48:2182–2203, 1993.

[45] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, et al. New parton distributions for collider physics. *Phys.Rev.*, D82:074024, 2010.

[46] W Lampl, S Laplace, D Lelas, P Loch, H Ma, S Menke, S Rajagopalan, D Rousseau, S Snyder, and G Unal. Calorimeter Clustering Algorithms: Description and Performance. Technical Report ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008.

[47] M. Mangano et al. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003.

[48] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63:189–285, 2009.

[49] D W Miller, A Schwartzman, and D Su. Pile-up jet energy scale corrections using the jet-vertex fraction method. Technical Report ATL-PHYS-INT-2009-090, CERN, Geneva, Sep 2009.

[50] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.

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

[52] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, Pavel M. Nadolsky, et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 0207:012, 2002.

[53] S. Frixione and B. R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 06:029, 2002.

[54] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *JHEP*, 05:026, 2006.

[55] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A brief introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852, 2008.

[56] S. S. Wilks. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, 9(1):pp. 60–62, 1938.

Curriculum Vitae

blank