BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

*WWW* **PRODUCTION: THE HUNT BEGINS**

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

(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

EM            ............          Electromagnetic Calorimeter

ID            ............          Inner Detector

LHC           ............          Large Hadron Collider

LO            ............          Leading-Order

MC            ............          Monte Carlo simulation

MS            ............          Muon Spectrometer

NLO           ............          Next-To-Leading-Order

QGC           ............          Quartic Gauge Coupling

SFOS          ............          Same-Flavor Opposite-Sign

SM            ............          Standard Model

PDF           ............          Parton Distribution Function

# Chapter 1

# Introduction

# Chapter 2

# Theory

We can describe the probability of a given process to be produced by calculating its cross-section, measured in barns (1 barn = $10^{-24}$ cm$^2$).

Since a $W$ boson is too short lived to ever be measured directly, we must instead study its decay products. The $W$ boson can decay into two different categories of particles: either "leptonically" into a charged lepton ($\ell$) and a neutrino ($\nu$) or "hadronically" into a pair of quarks. Quarks are particles that interact with the Strong force and are usually measured as complicated collections of particles called jets. The charged leptons include the electrons, $e^{\pm}$, familiar because of their presence in stable matter, and the less familiar muons, $\mu^{\pm}$, and taus, $\tau^{\pm}$, which themselves decay before forming stable matter. Muons live long enough to be measured in the ATLAS detector while tau leptons do not. Because of this, we only refer to electrons and muons when discussing the charged leptons. There are also three different types of neutrinos, named according to which charged lepton they pair with during $W$ decay; these are the electron neutrino ($\nu_e$), muon neutrino ($\nu_\mu$), and tau neutrino ($\nu_\tau$). Neutrinos are peculiar because they are notoriously difficult to measure. As a result, we can only infer their presence after measuring all of the particles in the detector and taking into account that the momentum is constrained at the collision point. If the momentum from all of the particles that were measured does not meet this constraint, the imbalance is attributed to the presence of neutrinos. We call this "missing energy". Thus, the leptonic $W$ decay channel is measured experimentally by the presence of a charged lepton (electron or muon) plus missing energy.

## 2.1 Standard Model

### 2.1.1 The Electroweak Theory

### 2.1.2 Electroweak Symmetry Breaking

### 2.1.3 The $W$ Boson

### 2.1.4 Quantum Chromodynamics

### 2.1.5 Parton Distribution Functions

## 2.2 Effective Field Theory

### 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: [28], [17] and [23]. As a benchmark, we choose two gauge invariant dimension-8 operators:

$$\mathcal{L}_{s,0} = [(\mathrm{D}_\mu\phi)^\dagger \mathrm{D}_\nu\phi] \times [(\mathrm{D}^\mu\phi)^\dagger \mathrm{D}^\nu\phi] \tag{2.1}$$

$$\mathcal{L}_{s,1} = [(\mathrm{D}_\mu\phi)^\dagger \mathrm{D}^\mu\phi] \times [(\mathrm{D}_\nu\phi)^\dagger \mathrm{D}^\nu\phi] \tag{2.2}$$

where $\phi$ is the Higgs field doublet, and $\mathrm{D}_\mu$ is the covariant derivative. The Lagrangian of the effective field theory is thus:

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \frac{f_{s0}}{\Lambda^4}\mathcal{L}_{s,0} + \frac{f_{s1}}{\Lambda^4}\mathcal{L}_{s,1} \tag{2.3}$$

The choice of the two operators is introduced to benchmark possible deviations from the Standard Model. 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

# The Large Hadron Collider

# Chapter 4

# The ATLAS Detector

[11]

## 4.1 Inner Detector

## 4.2 Calorimeters

## 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 [8]. 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 $\tau$ lepton decays). As can be seen in Fig. 5.1, this decay channel occurs only about 1% of the time, while 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 $p_{\mathrm{T}}$ leptons ($e$ or $\mu$), 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 while the backgrounds 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 are either
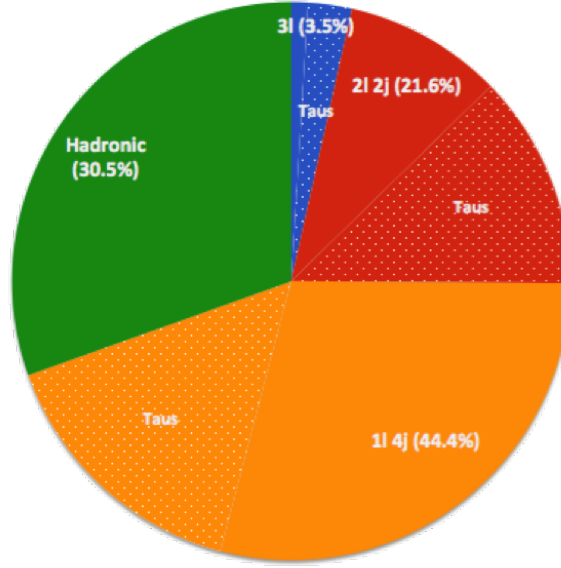
Figure 5.1: Pie chart showing the different decay modes contributing to the total cross-section for the $WWW$ process. The dotted areas indicate the portion of each decay mode which is due to the production of tau leptons.

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 [4]. 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 [5] 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.
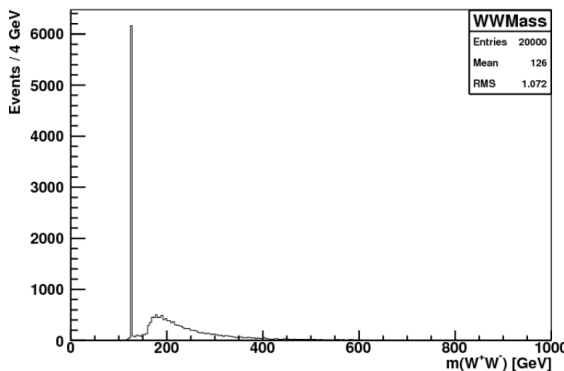


Figure 5.2:   Invariant mass distribution of two opposite-sign $W$ bosons in $WWW$ events generated with VBFNLO at LO. The Higgs mass peak is clearly visible at 126 GeV.

The SM signal processes are implemented in the Monte Carlo generator VBFNLO [9, 10], which can generate partonic events at leading-order (LO) with next-to-leading-order (NLO) cross-sections, and in MADGRAPH [8], which can generate partonic events at NLO with NLO cross-sections. The partonic events are further processed by PYTHIA8 [39] and PHOTOS [26] 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[3]. The MC generators must also be

provided an appropriate PDF. The PDF used in the LO VBFNLO generation is the LO CTEQ6L1 [36] PDF set while CT10 NLO [27] 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, while 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. Meanwhile, 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, $G_F$ | $1.16637 \times 10^{-5}$ GeV$^{-2}$ | $1.16637 \times 10^{-5}$ 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 | $W^+W^+W^- \to l\nu l\nu l\nu$ | $4.95 \pm 0.007$ | $0.2050 \pm 0.0070$ |
| | $W^-W^+W^- \to l\nu l\nu l\nu$ | $2.65 \pm 0.004$ | $0.0987 \pm 0.0037$ |
| | Sum | $7.60 \pm 0.008$ | $0.3037 \pm 0.0072$ |
| MadGraph | $W^+W^-W^+ \to$ Anything | $59.47 \pm 0.11$ | $0.0900 \pm 0.0048$ |
| | $W^-W^+W^- \to$ Anything | $28.069 \pm 0.076$ | $0.0476 \pm 0.0043$ |
| | $W^+H \to W^+W^+W^-(*) \to$ Anything | $99.106 \pm 0.019$ | $0.1114 \pm 0.0029$ |
| | $W^-H \to W^-W^+W^-(*) \to$ Anything | $54.804 \pm 0.010$ | $0.0603 \pm 0.0015$ |
| | Sum | $241.47 \pm 0.13$ | $0.3092 \pm 0.0072$ |

Table 5.2: Inclusive and common fiducial cross-sections at NLO for VBFNLO and Mad-Graph samples. The sum of the inclusive cross-sections are different because of the different branching fractions in the two cases. However, the sum of the fiducial cross-sections 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 [16] 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 [29], MSTW2008 NLO [32], and NNPDF 3.0 NLO [13]. 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 while 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 while the final PDF uncertainty on that estimate is taken as the envelope of the uncertainty bands for all three PDF sets.
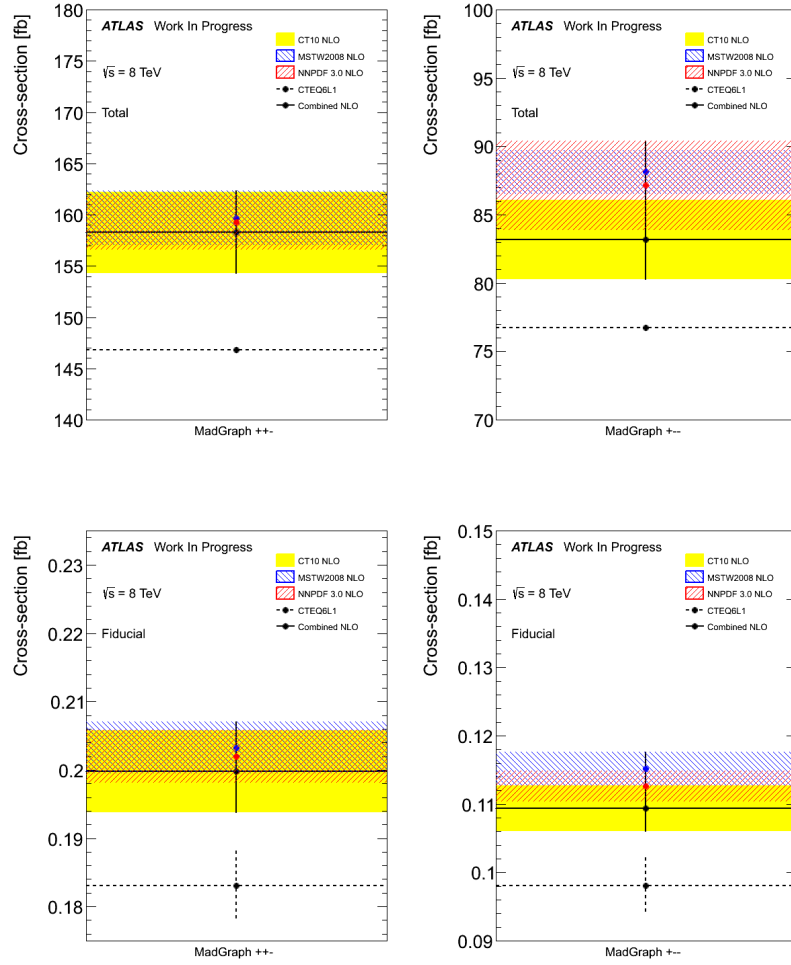
Figure 5.3: The signal cross-sections for different PDFs along with their uncertainties are shown on the MADGRAPH $WWW$ signal samples for the total $WWW$ phase space and branching fraction for for the $W^+W^+W^-$ (top left) and $W^+W^-W^-$ (top right) charge modes and in the fiducial region for $W^+W^+W^-$ (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) while 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 $W^-W^+W^-$.

| | $\mu_R$ / $\mu_F$ | $\frac{1}{2}M_{WWW}$ | $M_{WWW}$ | $2M_{WWW}$ |
|---|---|---|---|---|
| $W^+W^+W^-$ | $\frac{1}{2}M_{WWW}$ | 2.62% | -0.14% | -2.11% |
| | $M_{WWW}$ | 2.13% | 0 | -2.41% |
| | $2M_{WWW}$ | 1.56% | 0.24% | -2.42% |
| | $\mu_R$ / $\mu_F$ | $\frac{1}{2}M_{WWW}$ | $M_{WWW}$ | $2M_{WWW}$ |
| $W^-W^+W^-$ | $\frac{1}{2}M_{WWW}$ | 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

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

for the inclusive cross-section and

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

for the fiducial cross-section.

### 5.1.2.2  aQGC signal

need to update

### 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 while the third type is estimated using the data itself. This will be described in more detail in Sec. ??. 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 [7, 34, 24, 6] 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 MADGRAPHwith the CTEQ6L1 PDF set and hadronized using PYTHIA6 [38] with the AUET2B [12] tune.

The second category of backgrounds to consider are those with two prompt leptons and a photon. We will call these the photon backgrounds. This background occurs 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 [25] 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 [31] generator with the CTEQ6L1 PDF set and hadronize it using JIMMY [19] with the AUET2C [12] tune.

Some of the di-boson and tri-boson processes just discussed can also be produced through loop induced processes or double parton scattering. The $WW$ and $ZZ$ loop induced processes are generated using the gg2ZZ [15] and gg2WW [14] generators with the CT10 PDF set and hadronized using JIMMY with the AU2 tunes. The double parton scattering processes are generated using PYTHIA8with the AU2 tunes and the CTEQ6L1 PDF set.

The last category of backgrounds are those with prompt leptons plus jets that mimic leptons. These are nominally estimated using the data as described in Sec. **??**. However, some of the contributions to this background can be simulated using MC as a cross-check of the estimate from data and for other studies. The main contributions to this 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 process of $Z$+jets are generated using SHERPA with the CT10 PDF set while 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. Meanwhile, the $t\bar{t}$ processes are generated using the MC@NLO [37] generator with the CT10 PDF set and hadronized in JIMMY. Finally, this 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 while 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. However, we also pay attention to the presence of hadronic activity and neutrinos, 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 be pointing within the boundaries of the MS and ID by requiring that $|\eta| < 2.5$. The muon $p_{\mathrm{T}}$ 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 $p_{\mathrm{T}}$ 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 $p_{\mathrm{T}}$. 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 $p_{\mathrm{T}}$ when $p_{\mathrm{T}} < 20$ GeV and less than 10% of the muon $p_{\mathrm{T}}$ 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 the same way as for 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 [30] using the anti-$k_t$ algorithm [20] with $\Delta R < 0.4$. The reconstructed jet objects are required to have a reconstructed $p_T > 25$ 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 the likelihood that they come 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" [33, 1] and is only used with jets with $p_T < 50$ 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 [2] is used with 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 $E_T^{\mathrm{miss}}$. The $E_T^{\mathrm{miss}}$ is calclulated 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): electron, photons, hadronically decaying $\tau$-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:

- Electron-Muon Overlap: If$|\Delta R(e,\mu)| < 0.1$ then the muon is kept while the electron is thrown away.

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

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

For electrons, the direction is taken only from the electron calorimeter information. Muons use the full combined track information while jets use the direction taken from the anti-$k_T$ algorithm. 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 data is expected to be very small compared to the total dataset. Fortunately, the three lepton signature of the signal allows us to quickly throw out many events which do not looke like the signal. Still, this signature is not unique enough that the background is small. 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; then we use more stringent cuts to discriminate between the signal and backgrounds, referred to as our signal region selection. The signal region selection is determined starting from the pre-selection stage by performing an optimization procedure that minimizes the uncertainty on the final measurement. This is described in Sec. ....

The signal region selection is further divided into different categories that are each used in the final measurement 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 determined as 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 above in Sec. 5.2, each with a $p_{\mathrm{T}}$ 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 know to occassionally produce artificial bursts of noise. These instances were tracked and events where this occured were thrown away. The event is also required to have a primary vertex with at least three associed tracks. Finally, the event is required to pass the single lepton trigger requirements already listed in Sec. 5.1.1 where at least one of the three leptons

selected caused the trigger to fire. The signal plus background model (described in detail
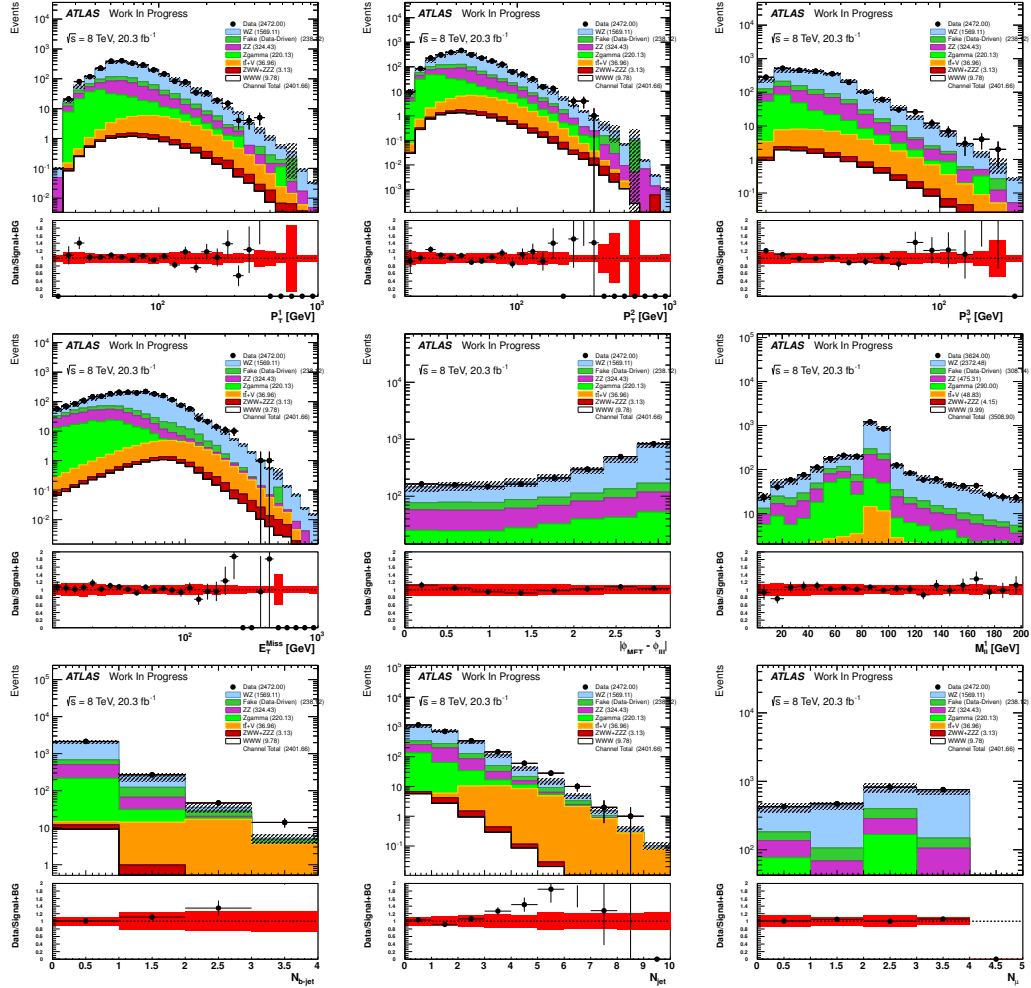


Figure 5.4: Distributions showing the observed data compared to the background estimate at event pre-selection.

in Sec. 5.4) is compared to data at pre-selection for a few different kinematic distributions of interest in Fig. 5.10. 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 into differnt categories based on the type of background and if it is the signal. Hashed bands are shown on the stacked histograms representing the size of the systematic uncertainties on the model, described later in Sec. **??**. 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:

- Lepton $P_T^1$: The momentum of the lepton in the transverse plane which has the largest transverse momentum out of all three leptons selected.

- Lepton $P_T^2$: The momentum of the lepton in the transverse plane which has the second largest transverse momentum out of all three leptons selected.

- Lepton $P_T^3$: The momentum of the lepton in the transverse plane which has the smallest transverse momentum out of all three leptons selected.

- $E_T^{\mathrm{miss}}$: The missing transverse energy

- $\Delta\varphi(lll, E_T^{\mathrm{miss}})$: The difference in the polar angle between the three lepton $p_T$, $\overrightarrow{p_T^{lll}} = \overrightarrow{p_T^1} + \overrightarrow{p_T^2} + \overrightarrow{p_T^2}$, and $E_T^{\mathrm{miss}}$. It can be expressed as $\Delta\varphi(lll, E_T^{\mathrm{miss}}) = \cos^{-1}\frac{\overrightarrow{p_T^{lll}} \cdot \overrightarrow{E_T^{\mathrm{miss}}}}{p_T^{lll} E_T^{\mathrm{miss}}}$

- SFOS Invariant Mass ($m_{\mathrm{SFOS}}$): The di-lepton invariant mass of each SFOS lepton pair in the event.

- Three Lepton $m_T$: The transverse mass of the three lepton system and the missing transverse energy, or $m_T^{lll} = \sqrt{2p_T^{lll} E_T^{\mathrm{miss}}(1 - \cos(\Delta\varphi(lll, E_T^{\mathrm{miss}})))}$

- $N_{\mathrm{Jet}}$: The number of jets selected in the event.

- $N_{b-\mathrm{Jet}}$: The number of $b$-jets tagged in the event.

- $N_\mu$: The number of selected leptons that are muons. The maximum possible is 3. If the lepton is not a muon it must be an electron.

In general, the signal plus background model is observed to be consistent with the data at the pre-selection, at least for those distributions considered here.

### 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 which 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 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.

Indeed, upon splitting the preselection region based on the number of SFOS pairs, we end up with signal and background predictions like in Fig. 5.11, where we can see difference in 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. Totalling about 10 signal events predicted at
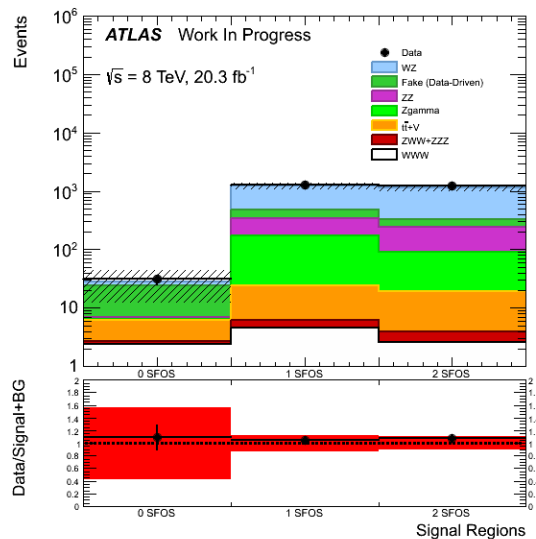
Figure 5.5: Yields at event pre-selection in the 0, 1 and 2 SFOS regions. The most important systematic uncertainties (discussed in section **??**) are shown, namely from the fake estimates and the uncertainties on the WZ and ZZ k-factors.

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, while only about 30 background events are predicted in the 0 SFOS region. Apparently 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 $\mu$. 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 the 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 $p_{\mathrm{T}}$). The large cross-section for theses processes means that they becomes

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 lepton charge, see section **??**. The 0 SFOS signal region is thus unique in having a small background which is almost entirely reducible and dominated instead by events where a jet is mis-measured as or overlaps with a lepton, called the fake lepton background, along with the aforementioned sub-dominant effect of lepton charge mis-identification described in Section **??**. From this, 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 while 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.

Within each signal region it may be that we can further reduce the background with respect to the signal region. Because the background composition in the 0 SFOS region is so different from the 1 and 2 SFOS regions, it is likely that the selection which does this will also differ between them. And even though the 1 and 2 SFOS regions are resonably similar, we should not rule out the possibility that they also have a uniquely optimial selection. Based on heuristic arguments, a list of reasonable physical quantities were considered. The most promising of these quantities are those listed previously in Sec. 5.3.1. By restricting these quantities to a region which removes a larger fraction of background than it does the signal, we can increase the signal prediction with respect to the background. The ratio of events which pass a selection with respect to the original number of events is refferred to as the efficiency. Thus, we refer to this as choosing a selection threshold where the signal efficiency is larger than the background efficiency. Also, because the predicted amount of signal is so precious, the signal efficiency should be kept as close to unity as possible. Consider the following selection choices using the quantities defined earlier:

- Lepton $P_T$: Require that exactly three leptons passing tight object quality require-

ments have a $P_T > X$.

- Missing $E_T$: Require that $E_{\mathrm{T}}^{\mathrm{miss}} > X$.

- $\Delta\varphi(lll, E_{\mathrm{T}}^{\mathrm{miss}})$: Require that $\Delta\varphi(lll, E_{\mathrm{T}}^{\mathrm{miss}}) = \cos^{-1}\frac{\overrightarrow{p_T^{lll}} \cdot \overrightarrow{E_{\mathrm{T}}^{\mathrm{miss}}}}{p_T^{lll} E_{\mathrm{T}}^{\mathrm{miss}}} > X$.

- Jet Veto: Require that $N_{Jet} \leq X$.

- b-Jet Veto: Require that $N_{b-Jet} \leq X$.

- Three Lepton $m_T$: Require that $m_T^{lll} > X$
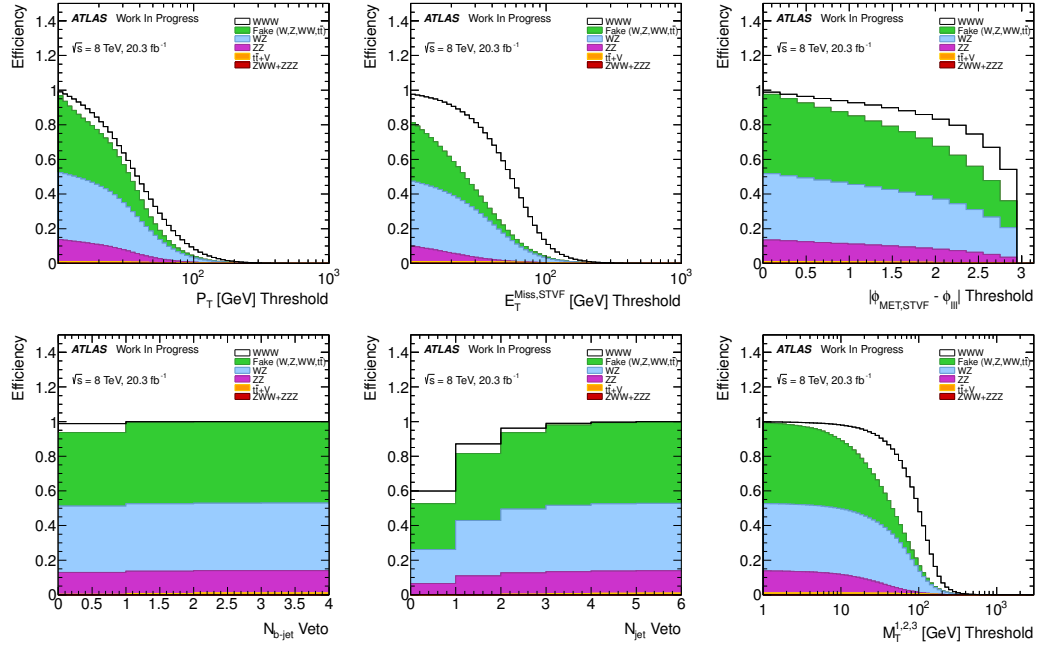


Figure 5.6: Signal and background efficiencies as a function of various cuts when starting from event pre-selection.

We can plot the efficiency for the signal as well as the background as a function of the threshold, $X$, for each of these quantities. The efficiencies are shown starting in the pre-selection region in Fig. 5.6 (which has a similar background composition to the 1 and 2 SFOS regions) and in the pre-selection + 0 SFOS region in Fig. 5.7. Clearly, the shape of the efficiencies between the signal and background is not always the same, which means
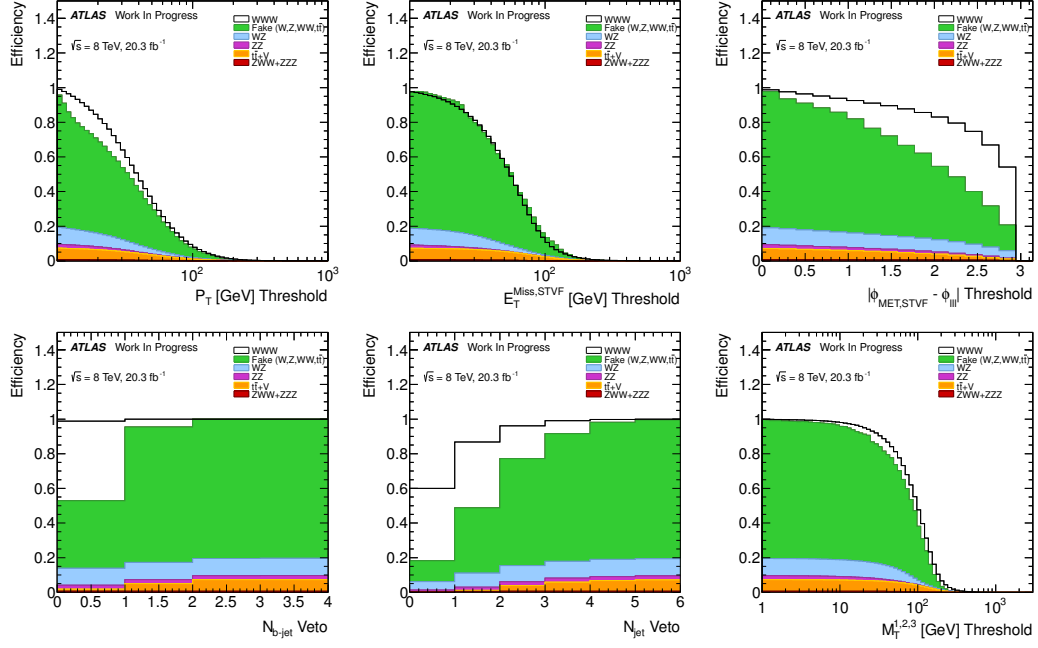
Figure 5.7: Signal and background efficiencies as a function of various cuts when starting from event pre-selection + 0 SFOS lepton pairs.

that we have some power to reduce the background with respect to the signal. Furthermore, these shapes are not the same in the pre-selection + 0 SFOS region as they are in just the pre-selection region. For instance, in the 0 SFOS region, the shape of the efficiencies for signal and background are very similar as a function of the $E_T^{\text{miss}}$ threshold, but in the pre-selection stage they are very differerent, with the background efficiency dropping much more rapidly than the signal efficiency. This suggests that a cut on the $E_T^{\text{miss}}$ distribution could be very useful in the 1 and 2 SFOS regions but not in the 0 SFOS region. Something similar may be true for a cut on $m_T^{lll}$. On the other hand, a cut on $\Delta\varphi(lll, E_T^{\text{Miss}})$ could prove useful in all signal regions. The lepton $p_T$ is frequently used as a useful discriminant but in this case it seems to rapidly remove the signal at about the same rate as the background, so a tighter cut on the lepton $p_T$ is likely to prove unuseful. Cuts on $N_{\text{Jet}}$ seem to remove a lot of the signal but cut tighter on the background in the pre-selection + 0 SFOS region, which could prove useful. Most impressive is the cut on $N_{b-\text{Jet}}$, which leaves the signal efficiency at almost exactly one, while having a large impact on the pre-selection + 0 SFOS

region where the fake contribution is large.

This type of heuristic study is useful for obtaining insight into the distributions that are useful for discriminating between signal and background. But using these distributions alone to determine a final selection is a little haphazard. For one, they do not tell us the impact that cutting on multiple distributions will have. These distributions are likely to be at least partially correlated, so that cutting on one distribution might enhance or deflate the impact of another cut. And two, it is not at all clear from these distributions what the best choice of cut is. We need a metric that will allow us to make decisions about which cuts to choose. The metric to use should be one that is related to the final prediction after all selection cuts are applied. Thus, the metric cannot be used to make a decision about each cut individually, but instead a decision about some combination of cuts.

To deal with the issues mentioned above, we choose to decide on the final signal region selection using an optimization procedure. The optimization takes as input a multi-dimensional space where each dimension is the selection threshold for one of the quantities listed above, plus some others not mentioned. The range of the multi-dimensinonal space that is interesting in the optimization 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. **??** 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 $N_{\text{Jet}}$, 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. However, looking at Fig. 5.6 and Fig. 5.7, the shape of the efficiencies tend to change relatively slowly from bin to bin. Thus, it should be acceptable to only sample discretely as long as they can capture the shape information of the efficiencies as they do above. Furthermore, this acknowledges the finite experimental resolution of these quantities. For example, the difference between $p_T > 20$ GeV and $p_T > 20.5$ GeV should not be taken too seriously because 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 sample space after discretizing, while 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 optimial choice for the selection.

The shape of the optimization can be seen in Fig. 5.8. *Figures need to be reproduced. Elaborate...*



Figure 5.8: Signal Yield vs Measurement Uncertainty for optimized points in the 0 SFOS (left), 1 SFOS (middle), and 2 SFOS (right) signal regions.

From the optimization, a final selection is chosen separately for each signal region. This selection is presented in Table 5.5. In all regions, a veto is applied on events with jets tagged to come from $b$ or $b$-hadron decays. Besides the nominal $b$-tagging operating point with an 85 % $b$-tagging efficiency (mentioned in Sec. 5.2), other $b$-tagging efficiency operating points were considered that had a lower efficiency (namely 80%, 70%, and 60%).

| | 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) | $|m_{ee} - m_Z| > 15$ GeV | No $m_{\mathrm{SFOS}}$ with $m_Z - 35\mathrm{GeV} < m_{\mathrm{SFOS}} < m_Z + 20$ GeV | $|m_{\mathrm{SFOS}} - m_Z| > 20$ GeV |
| Missing $E_T$ | | $E_T^{Miss} > 45$ GeV | $E_T^{Miss} > 55$ GeV |
| Lepton-Missing $E_T$ Angle | $|\phi(3l) - \phi(E_T^{Miss})| > 2.5$ | | |
| Inclusive Jet veto | $N_{jet} \leq 1$ | | |

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

These other operating points could be useful if the *b*-tagging mis-identification was high enough to remove a significant portion of the signal. However, even at the 85 % operating point the mis-identification efficiency remains manageable at about 1 %. Even with some jet mis-identification, the signal has a very high efficiency of passing this cut of $> 99$ % while offering some of the strongest reduction in the fake lepton background. Thus there is no benefit observed by going to a lower efficiency operating point. The analysis could potentially benefit from an even more stringent operating point, but the highest efficiency operating point that is supported by the ATLAS *b*-tagging group is the 85 % efficiency operating point. In addition to the the *b*-jet veto, there is an additional cut on the jet multiplicity, regardless of whether or not the jet is tagged. By only keeping events with no more than one jet, the signal efficiency is almost 90 % while reducing the background by about 50 %. Applying a veto on all jets does a very good job at removing the fake lepton background, but the signal efficiency is prohibitively small, at about 60 %.

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 [35]). 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. This can clearly be seen in the plot of $m_{\mathrm{SFOS}}$ in Fig. 5.10. As a result, removing events within some window around the peak can do a good job of removing these bacgkrounds without having a large effect on the signal. In the 0 SFOS region, by definition there

are no SFOS pairs that could come from the decay of a $Z$-boson. However, the effect of elecron charge mis-identification, discussed in Sec. **??**, means that a peak can show up in the background of the $m_{ee}$ distribution for same-sign electron/positron pairs. This can clearly be seen in Fig. **??**. Thus, a veto can be performed in this distribution as well.

The presence of neutrinos in the signal mean that the signal should have a relatively large $E_{\mathrm{T}}^{\mathrm{miss}}$ compared to most of the backgrounds. Thus, cutting on the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution such that it is large can remove backgrounds expected to have small $E_{\mathrm{T}}^{\mathrm{miss}}$, like $Z\gamma$ production. Still, there are some large backgrounds with neturinos, like $WZ$, and also backgrounds that have contributions to the $E_{\mathrm{T}}^{\mathrm{miss}}$ from objects that have missed reconstruction, like $ZZ$, which can also have a moderate to large $E_{\mathrm{T}}^{\mathrm{miss}}$. Thus, some care must be taken to choose a threshold to cut on in the $E_{\mathrm{T}}^{\mathrm{miss}}$.

The $Z\gamma$ background shows up in the low-shoulder of the $Z$ peak in the $m_{\mathrm{SFOS}}$ distribution and at low MET. This can be seen both for the 1 and 2 SFOS regions in Fig. 5.9. 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. Meanwhile, in the 2 SFOS region, the $Z\gamma$ contribution is not as prominent and the optimization happens to prefer a symmetric window of $\pm 20$ 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. In the 0 SFOS region there are no SFOS pairs by definition, but there is still a peak in the same-sign electron-electron mass distribution due to charge mis-identification. The optimization prefers a slightly narrower symmetric window of $\pm 15$ GeV around the $Z$-mass. Further, as already mentioned, this background turns out to have a similar missing $E_T$ distribution as the signal. As a result,
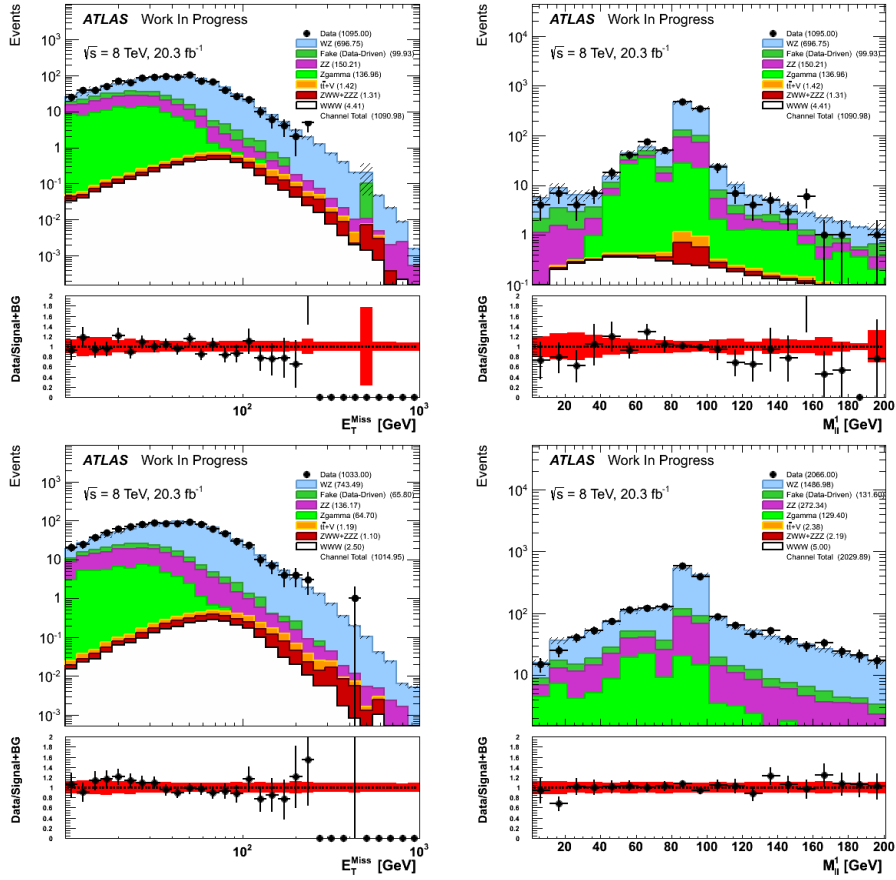
Figure 5.9: Plots of the $E_T^{miss}$(left) and $m_{SFOS}$ (right) distributions in the 1 SFOS (top) and 2 SFOS (bottom) regions after pre-selection plus the $b$-veto requirement.

cutting on the missing $E_T$ in this region offers little to no discriminating power between the signal and background so we have chosen not to apply any cut here in order to maximize the signal yield.

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 $E_T$). To some extent, 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: $|\Delta\phi(3l, E_{\mathrm{T}}^{\mathrm{miss}})| = |\phi(3l) - \phi(E_{\mathrm{T}}^{\mathrm{miss}})|$. The behavior of this quantity for signal and background is similar in all three signal regions so based on the optimization it was chosen to apply the cut $|\Delta\phi(3l, E_{\mathrm{T}}^{\mathrm{miss}})| > 2.5$ everywhere. We also considered taking the difference in angle between the missing $E_T$ and individual leptons (e.g. the highest $p_{\mathrm{T}}$ lepton) but this was shown to be not nearly as effective.

Finally, the distribution of the invariant mass from same-flavor lepton pairs (regardless of sign) was considered in the 0 SFOS region to remove any low-mass contamination from processes like from QCD. This was shown to offer some modest discriminating power and a threshold of $m_{SF} > 20$ GeV was chosen only for the 0 SFOS region.

### 5.3.3  Fiducial Region Selection

A fiducial phase space or fiducial region is defined purely using truth information (generator information before being passed through ATLAS reconstruction) as the region the analysis is sensitive to. 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 $p_{\mathrm{T}}$  requirement of $p_{\mathrm{T}}^{\mathrm{Reco}} > 20$  GeV is taken into account in the fiducial region selection by requiring $p_{\mathrm{T}}^{\mathrm{Truth}} > 20$  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 [18], which allows for comparisons between different generators.

| | 0 SFOS | 1 SFOS | 2 SFOS |
|---|---|---|---|
| All | All | | |
| Tau Veto | $N_\tau < 1$ | | |
| Fiducial Leptons | Exactly 3 leptons with $p_T > 20$ GeV and $|\eta| < 2.5$ | | |
| Lepton Overlap Removal | $\Delta R(\ell\ell) > 0.1$ | | |
| Same-Flavor Mass | $m_{\mathrm{SF}} > 20$ GeV | | |
| Z-Veto ($m_Z = 91.1876$ GeV) | $|m_{ee} - m_Z| > 15$ GeV | No $m_{\mathrm{SFOS}}$ with $m_Z - 35\mathrm{GeV} < m_{\mathrm{SFOS}} < m_Z + 20$ GeV | $|m_{\mathrm{SFOS}} - m_Z| > 20$ GeV |
| Missing $E_T$ | | $E_T^{Miss} > 45$ GeV | $E_T^{Miss} > 55$ GeV |
| Lepton-Missing $E_T$ Angle | $|\phi(3l) - \phi(E_T^{Miss})| > 2.5$ | | |
| Inclusive Jet veto | $N_{jet} \leq 1$ with fiducial jets of $p_T > 25$ GeV and $|\eta| < 4.5$ | | |

Table 5.6: Fiducial regions based on optimized selection.

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

The fiducial cross-sections are evaluated in both the VBFNLO and MADGRAPH samples mentioned in Section 5.1.2.1. The predicted fiducial cross-sections at each stage of the selection are listed for both generators in Table ??. The fiducial cross-sections and their statistical uncertainties after the entire fiducial selection are summarized in each channel for both generators in Table 5.7. They are seen to be in good agreement between both generators.

A summary of the signal sample phase space and cross-sections is presented in Table ??. The derived fiducial cross-sections using the two generators are shown in Table 5.7. The cross-sections are observed to be in good agreement between the two generators. The MadGraph NLO fiducial cross-sections are used in the final measurement. The PDF and scale uncertainties on the fiducial cross-sections are determined after summing over all three signal regions as described earlier in Sec. 5.1.2.1.

| Channel | Fiducial Cross-section [ab] | |
| --- | --- | --- |
| | MadGraph | VBFNLO |
| 0 SFOS | $114.7 \pm 4.3$ | $126.9 \pm 1.0$ |
| 1 SFOS | $126.6 \pm 4.3$ | $126.1 \pm 1.0$ |
| 2 SFOS | $50.2 \pm 2.7$ | $50.62 \pm .66$ |

Table 5.7: Fiducial cross-sections derived in each signal region for the two generators. Production modes are summed together to get one fiducial cross-section per channel per generator. The cross-sections are seen to be in good agreement between the two generators.

## 5.4  Background Estimates

## 5.5  Event Yield in event Pre-selection and Signal Regions

The background estimation described in Section **??** and the data are passed through the three-lepton selections described in Section **??**.

### 5.5.1  Event Pre-selection

The three lepton event pre-selection region is used as a starting point for all three signal regions and follows the selection described in Section 5.3.1. Kinematic distributions in this region are shown in Figure 5.10 while the yields are shown binned by the lepton flavor combinations in Table 5.9. The yields in the individual signal channels just after pre-selection are shown in Figure 5.11 with systematic uncertainties. In general, good agreement is observed.
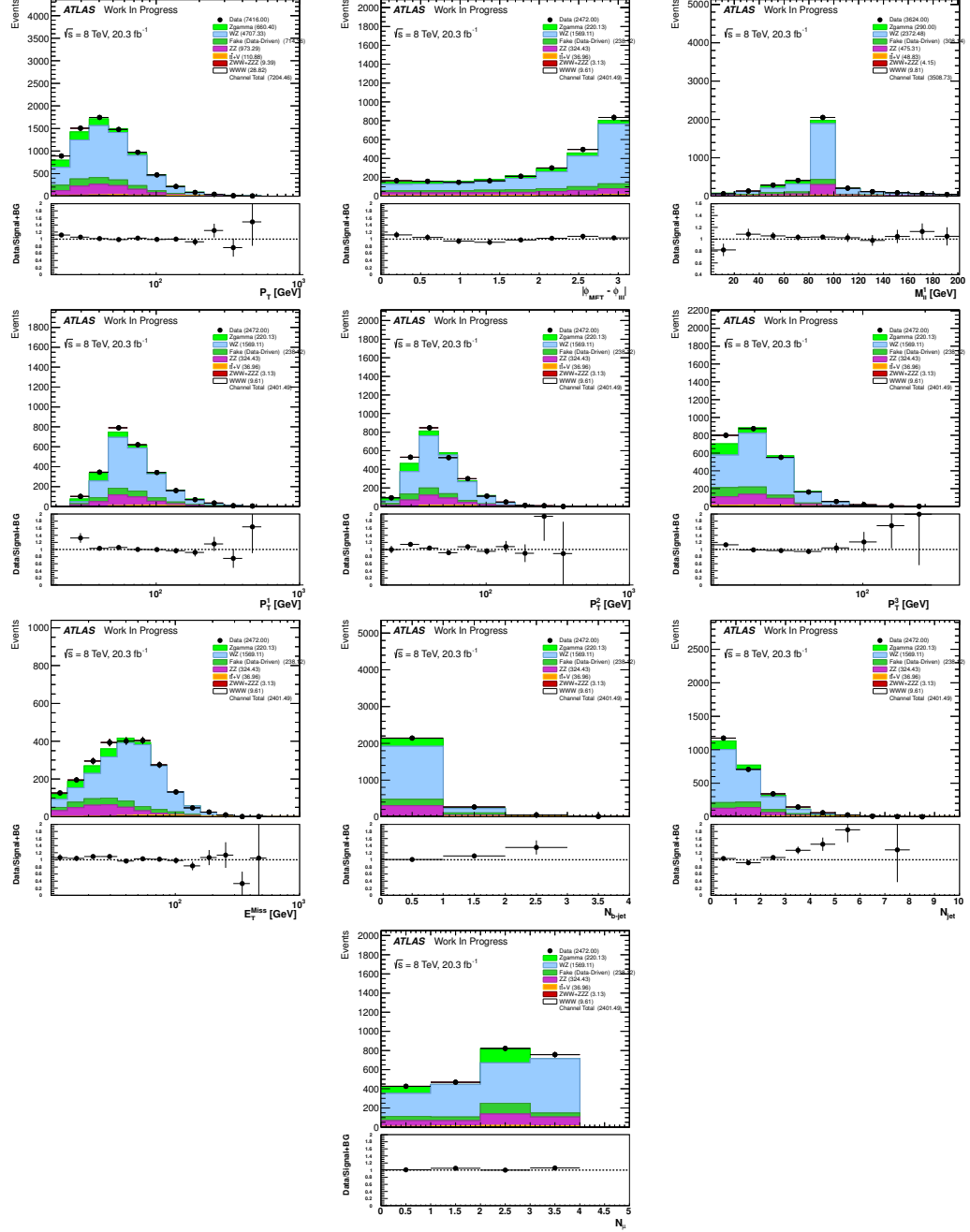
Figure 5.10: Distributions showing the observed data compared to the background estimate at event pre-selection.

|  | $eee$ | $ee\mu$ | $e\mu\mu$ | $\mu\mu\mu$ |
|---|---|---|---|---|
| $WZ$ | $240.85 \pm 0.67$ | $339.17 \pm 0.82$ | $422.07 \pm 0.87$ | $567.0 \pm 1$ |
| $ZZ$ | $60.21 \pm 0.13$ | $54.1 \pm 0.2$ | $118.60 \pm 0.31$ | $91.48 \pm 0.17$ |
| $Z\gamma$ | $70.1 \pm 2.7$ | $0.47 \pm 0.22$ | $149.4 \pm 3.9$ | $0.17 \pm 0.12$ |
| $ZWW + ZZZ$ | $0.436 \pm 0.019$ | $0.834 \pm 0.027$ | $1.00 \pm 0.03$ | $0.864 \pm 0.028$ |
| $t\bar{t} + V$ | $4.854 \pm 0.044$ | $9.549 \pm 0.064$ | $12.047 \pm 0.072$ | $10.510 \pm 0.066$ |
| Fake (data-driven) | $45.1 \pm 2.2$ | $37.8 \pm 1.6$ | $112.7 \pm 2.8$ | $42.5 \pm 1.2$ |
| $WWW$ | $0.770 \pm 0.011$ | $3.023 \pm 0.023$ | $3.970 \pm 0.026$ | $1.843 \pm 0.018$ |
| Expected Background | $421.6 \pm 3.5$ | $441.9 \pm 1.8$ | $815.8 \pm 4.9$ | $712.5 \pm 1.6$ |
| Expected Signal + Background | $422.4 \pm 3.5$ | $444.9 \pm 1.8$ | $819.8 \pm 4.9$ | $714.4 \pm 1.6$ |
| Observed Data | $426 \pm 21$ | $468 \pm 22$ | $821 \pm 29$ | $757 \pm 28$ |

Table 5.8: Expected and observed event yields binned by lepton flavor combination at event pre-selection. Only statistical uncertainties are shown.
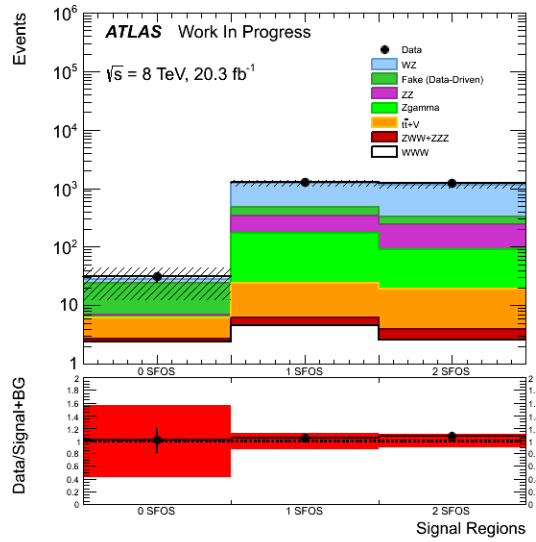


Figure 5.11: Yields at event pre-selection in the 0, 1 and 2 SFOS regions. The most important systematic uncertainties (discussed in section ??) are shown, namely from the fake estimates and the uncertainties on the WZ and ZZ k-factors.

| | $eee$ | $ee\mu$ | $e\mu\mu$ | $\mu\mu\mu$ |
|---|---|---|---|---|
| $WZ$ | $240.85 \pm 0.67$ | $339.17 \pm 0.82$ | $422.07 \pm 0.87$ | $567.0 \pm 1$ |
| $ZZ$ | $60.21 \pm 0.13$ | $54.1 \pm 0.2$ | $118.60 \pm 0.31$ | $91.48 \pm 0.17$ |
| $Z\gamma$ | $70.1 \pm 2.7$ | $0.47 \pm 0.22$ | $149.4 \pm 3.9$ | $0.17 \pm 0.12$ |
| $ZWW + ZZZ$ | $0.436 \pm 0.019$ | $0.834 \pm 0.027$ | $1.00 \pm 0.03$ | $0.864 \pm 0.028$ |
| $t\bar{t} + V$ | $4.854 \pm 0.044$ | $9.549 \pm 0.064$ | $12.047 \pm 0.072$ | $10.510 \pm 0.066$ |
| Fake (data-driven) | $45.1 \pm 2.2$ | $37.8 \pm 1.6$ | $112.7 \pm 2.8$ | $42.5 \pm 1.2$ |
| $WWW$ | $0.770 \pm 0.011$ | $3.023 \pm 0.023$ | $3.970 \pm 0.026$ | $1.843 \pm 0.018$ |
| Expected Background | $421.6 \pm 3.5$ | $441.9 \pm 1.8$ | $815.8 \pm 4.9$ | $712.5 \pm 1.6$ |
| Expected Signal + Background | $422.4 \pm 3.5$ | $444.9 \pm 1.8$ | $819.8 \pm 4.9$ | $714.4 \pm 1.6$ |
| Observed Data | $426 \pm 21$ | $468 \pm 22$ | $821 \pm 29$ | $757 \pm 28$ |

Table 5.9: Expected and observed event yields binned by lepton flavor combination at event pre-selection. Only statistical uncertainties are shown.

## 5.5.2 Signal Regions

The selection used in the final signal regions is determined as described in Section 5.3.2 and is summarized in Table 5.5. Kinematic distributions are shown for the distribution that is cut on just before the cut is applied for each stage of the selection in Figures 5.12, 5.13, and 5.14 for the 0, 1, and 2 SFOS regions, respectively. A more detailed set of kinematic distributions at each stage of the selection is shown in appendix **??**. Cut-flows showing the weighted cut-flows at each stage of the signal selection are shown for the 0, 1, and 2 SFOS regions in Tables 5.10, 5.11, 5.12. Unweighted cut-flows are shown in appendix **??**.

### 5.5.2.1 0 SFOS Signal Region

The cut-flows for the selection in the 0 SFOS region are shown in Table 5.10 while the distributions at each stage of the selection are shown in Figure 5.12. This is the most sensitive channel with an expected signal to background ratio of 56%. The data is observed to agree with the expectation within statistical uncertainties. The background is ultimately dominated by fake background contributions, which themselves have a large systematic uncertainty that is similar in size to the statistical uncertainty. This is described in more detail in section **??**. The Poisson probability of observing 5 events with 3.67 events expected from the signal plus background prediction is 14.1%.

| | Signal | | Background | | | | | | | | | |
| | | | $WZ$ | | $ZZ$ | | $t\bar{t}+V$ | | $ZZZ+ZWW$ | | $Z\gamma$ | |
| | Yield | Eff. | Yield | Eff. | Yield | Eff. | Yield | Eff. | Yield | Eff. | Yield | Eff. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Pre-selection | 9.61 | — | 1566.91 | — | 323.60 | — | 36.93 | — | 3.12 | — | 219.80 | — |
| 2. 0 SFOS | 2.27 | 0.24 | 2.84 | 0.002 | 0.50 | 0.002 | 0.26 | 0.01 | 0.25 | 0.08 | 0.20 | 0.001 |
| 3. Charge Sum = ±1 | 2.26 | 1.00 | 1.92 | 0.68 | 0.33 | 0.65 | 0.26 | 0.99 | 0.25 | 1.00 | 0.00 | 0.00 |
| 4. $N_{\mathrm{b-jet}}=0$ | 2.25 | 0.99 | 1.91 | 0.99 | 0.33 | 0.99 | 0.25 | 0.98 | 0.25 | 0.99 | 0.00 | 0.00 |
| 5. $m_{SF}>20$ GeV | 2.21 | 0.98 | 1.88 | 0.98 | 0.32 | 0.98 | 0.25 | 0.98 | 0.24 | 0.98 | 0.00 | 0.00 |
| 6. $|m_{ee}-m_Z|>15$ GeV | 2.02 | 0.91 | 1.27 | 0.68 | 0.21 | 0.66 | 0.22 | 0.90 | 0.22 | 0.90 | 0.00 | 0.00 |
| 7. $|\Delta\phi(3l,E_T^{Miss})|>2.5$ | 1.39 | 0.69 | 0.65 | 0.51 | 0.07 | 0.34 | 0.09 | 0.38 | 0.13 | 0.59 | 0.00 | 0.00 |
| 8. $N_{\mathrm{Jet}}\leq1$ | 1.32 | 0.95 | 0.62 | 0.95 | 0.07 | 0.91 | 0.04 | 0.45 | 0.11 | 0.86 | 0.00 | 0.00 |

Table 5.10: 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 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 5.12: 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 5.10, the top left plot is shown before cut #4 is applied, top middle is before cut #5, and so on until the bottom right which is after all cuts are applied.

### 5.5.2.2    1 SFOS Signal Region

The cut-flows for the selection in the 1 SFOS region are shown in Table 5.11 while the distributions at each stage of the selection are shown in Figure 5.13. This 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. The observed data is slightly discrepant from the overall prediction, but is about 1 sigma if considering systematics, as demonstrated later in Table ??. The difference is observed to come almost entirely from the $ee\mu$ region, while the $e\mu\mu$ region shows good agreement, as can be seen in the bottom right of Figure 5.13. From Figure 5.13 we can also see that the agreement is quite good at each stage of the selection until the cut on the jet multiplicity (bottom middle of Figure 5.13) where the 0 and 1 jet bins are kept but the 1 jet bin is discrepant. This suggests that the discrepancy comes from this cut. This was investigated further for the different lepton flavor combinations in appendix ??. The efficiencies for the predictions are very similar for the jet multiplicity cut between the two lepton flavor bins. However, there is a smaller efficiency observed in the data in the $ee\mu$ bin. This suggests that the difference is most likely due to a statistical fluctuation. The Poisson probability of observing 13 events with 16.14 events expected from the signal plus background prediction is 7.9%.

| | Signal | | Background | | | | | | | | | | |
| | | | $WZ$ | | $ZZ$ | | $t\bar{t}+V$ | | $ZZZ+ZWW$ | | $Z\gamma$ | | |
| | Yield | Eff. | Yield | Eff. | Yield | Eff. | Yield | Eff. | Yield | Eff. | Yield | Eff. | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Pre-selection | 9.61 | — | 1566.91 | — | 323.60 | — | 36.93 | — | 3.12 | — | 219.80 | — | 2 |
| 2. 1 SFOS | 4.59 | 0.48 | 757.38 | 0.48 | 171.39 | 0.53 | 18.10 | 0.49 | 1.55 | 0.50 | 149.60 | 0.68 | 1 |
| 3. $N_{\text{b}-\text{jet}}=0$ | 4.34 | 0.94 | 696.90 | 0.92 | 150.14 | 0.88 | 1.42 | 0.08 | 1.31 | 0.84 | 136.96 | 0.92 | 9 |
| 4. NOT $m_Z - 35$ GeV $< m_{\text{SFOS}} < m_Z + 20$ GeV | 2.71 | 0.63 | 44.30 | 0.06 | 13.79 | 0.09 | 0.37 | 0.26 | 0.34 | 0.26 | 22.44 | 0.16 | 1 |
| 5. $E_T^{Miss} > 45$ GeV | 1.88 | 0.69 | 21.38 | 0.48 | 1.46 | 0.11 | 0.29 | 0.78 | 0.24 | 0.71 | 1.36 | 0.06 | |
| 6. $|\Delta\phi(3l, E_T^{Miss})| > 2.5$ | 1.45 | 0.77 | 13.07 | 0.61 | 0.71 | 0.49 | 0.11 | 0.39 | 0.17 | 0.69 | 0.20 | 0.15 | |
| 7. $N_{\text{Jet}} \le 1$ | 1.37 | 0.94 | 11.90 | 0.91 | 0.58 | 0.82 | 0.05 | 0.45 | 0.14 | 0.84 | 0.20 | 1.00 | |

Table 5.11: 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 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.
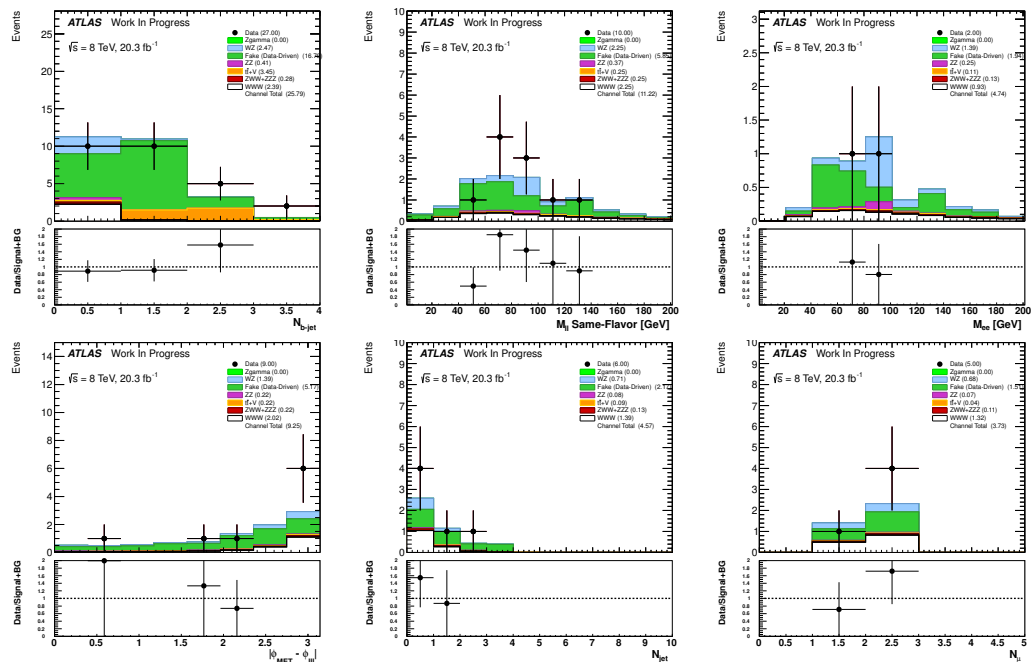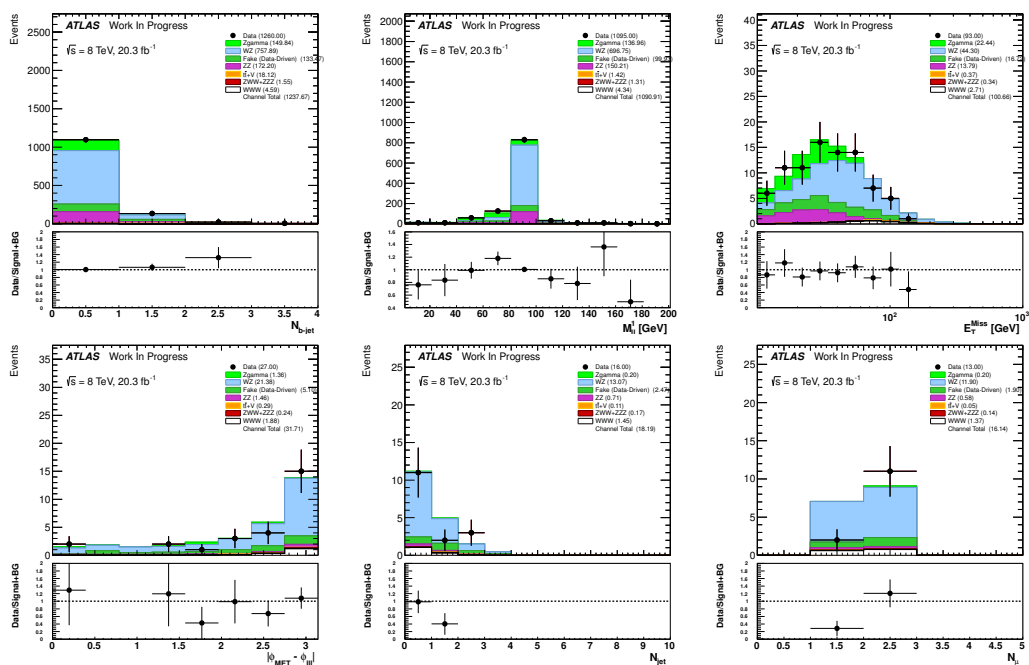
Figure 5.13: Distributions showing data compared to the signal plus background estimate in the 1 SFOS region at each stage of the selection before the cuts are applied to the given distribution. Plots should be read sequentially from left to right and from top to bottom. Referring to Table 5.11, the 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.

### 5.5.2.3 2 SFOS Signal Region

The cut-flows for the selection in the 2 SFOS region are shown in Table 5.12 while the distributions at each stage of the selection are shown in Figure 5.14. Even though, the expected background is lower in this region than in the 1 SFOS region, this is the least sensitive signal region due to a similar signal efficiency in this region as compared to the 1 SFOS region. The signal to background ratio is 5.8% with the dominant background being due to $WZ$, similar to the 1 SFOS region. There is a fairly large discrepancy observed between the total prediction and the data in this region that is (possibly) about 2-3 sigma if considering systematic uncertainties reported in Table **??**. The discrepancy is similar for both the *eee* and $\mu\mu\mu$ lepton flavor combinations as can be seen in the bottom right of Figure 5.14. If we examine the other distributions of Figure 5.14 we see that the discrepancy begins to appear at the cut on $E_T^{Miss} > 55$ GeV. The $E_T^{Miss}$ distribution is shown in the top right of Figure 5.14 before the cut is applied. Here we can see that the prediction shows reasonably good agreement with the data, with only 2 bins showing a discrepancy larger than 1 sigma from the statistical uncertainty. However, after the cut on $E_T^{Miss}$, one of these discrepant bins ends up being the highest contribution to the overall estimate, thus enhancing the disagreement. This suggests that the disagreement may just be due to a statistical fluctuation. The Poisson probability of observing 6 events with 10.86 events expected from the signal plus background prediction is 5.7%.

### 5.5.2.4 Signal Efficiency

The signal efficiency, $\varepsilon_i$, is defined for each channel, $i$, as the ratio of the number of expected signal events measured at the reconstruction level, $N_i^{\text{Signal}}$, over the fiducial cross-section, $\sigma_i^{\text{Fiducial}}$, times the integrated luminosity:

$$\varepsilon_i = \frac{N_i^{\text{Signal}}}{\sigma_i^{\text{Fiducial}} \cdot \int \mathscr{L} \, \mathrm{d}t} \tag{5.3}$$

| | Signal | | Background | | | | | | | | |
| | Yield | Eff. | WZ | | ZZ | | $t\bar{t}+V$ | | $ZZZ+ZWW$ | | $Z\gamma$ | |
| | | | Yield | Eff. | Yield | Eff. | Yield | Eff. | Yield | Eff. | Yield | Eff. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Pre-selection | 9.61 | — | 1566.91 | — | 323.60 | — | 36.93 | — | 3.12 | — | 219.80 | — |
| 2. 2 SFOS | 2.61 | 0.27 | 807.27 | 0.52 | 151.28 | 0.47 | 15.35 | 0.42 | 1.30 | 0.41 | 69.99 | 0.32 |
| 3. $N_{\text{b-jet}}=0$ | 2.46 | 0.94 | 743.12 | 0.92 | 136.16 | 0.90 | 1.19 | 0.08 | 1.10 | 0.85 | 64.70 | 0.92 |
| 4. $|m_{\text{SFOS}}-m_Z|>20$ GeV | 1.43 | 0.58 | 44.95 | 0.06 | 21.13 | 0.16 | 0.22 | 0.18 | 0.19 | 0.17 | 29.52 | 0.46 |
| 5. $E_T^{Miss}>55$ GeV | 0.82 | 0.57 | 15.86 | 0.35 | 0.97 | 0.05 | 0.14 | 0.65 | 0.12 | 0.63 | 0.43 | 0.01 |
| 6. $|\Delta\phi(3l,E_T^{Miss})|>2.5$ | 0.64 | 0.78 | 10.09 | 0.64 | 0.55 | 0.57 | 0.07 | 0.49 | 0.10 | 0.82 | 0.11 | 0.25 |
| 7. $N_{\text{Jet}}\leq1$ | 0.60 | 0.94 | 9.07 | 0.90 | 0.48 | 0.86 | 0.02 | 0.35 | 0.08 | 0.82 | 0.11 | 1.00 |

Table 5.12: 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 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.
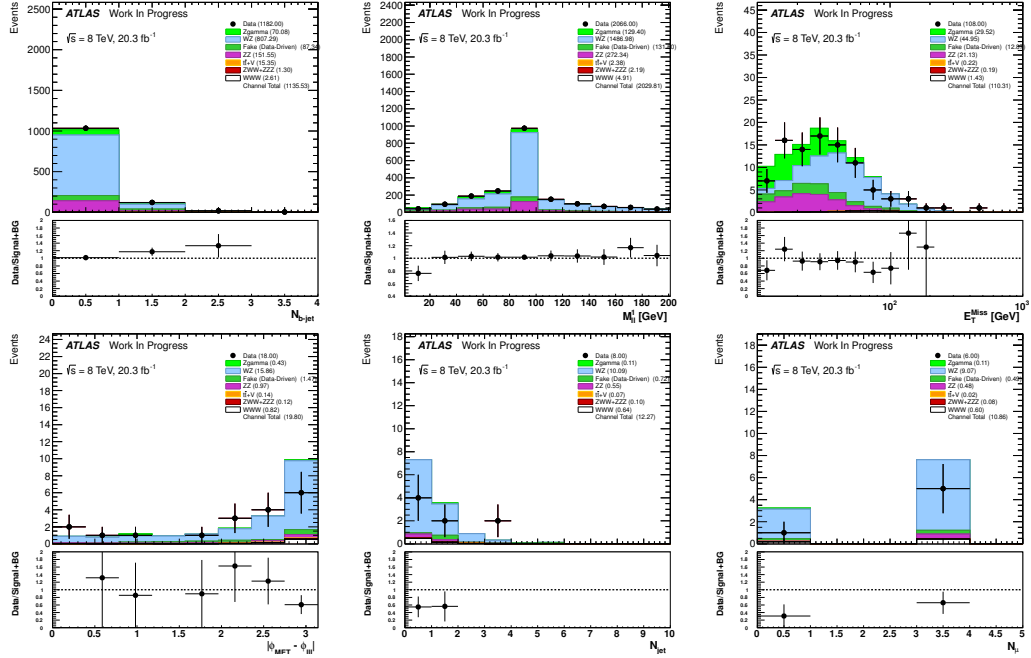


Figure 5.14: Distributions showing data compared to the signal plus background estimate in the 2 SFOS region at each stage of the selection before the cuts are applied to the given distribution. Plots should be read sequentially from left to right and from top to bottom. Referring to Table 5.12, the 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.

Recall that the fiducial cross-sections are presented in Section ??. Further, the fiducial cross-section definition does not include the branching fraction from $W\to\tau\nu$ decays. We observe by looking at truth information that about 20% of signal events reconstructed

at truth level contain at least one $W \to \tau\nu$ decay. Thus, the signal efficiency definition includes this level of contamination from these events even though they do not belong explicitly to our signal definition. Using the reconstructed signal yields listed above and the fiducial cross-sections generated using MadGraph from Table 5.7, we arrive at the signal efficiencies listed in Table 5.13.

| Channel | Signal Efficiency |
|---------|-------------------|
| 0 SFOS | $0.567 \pm .022$ |
| 1 SFOS | $0.533 \pm .019$ |
| 2 SFOS | $0.589 \pm .033$ |

Table 5.13: Signal efficiencies derived separately for each signal region. Only statistical uncertainties are shown.

## 5.6  Standard Model Measurement

### 5.6.1  Introduction

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

$$\sigma^{\text{Observed}} = \mu \sum_{i \in \text{Channels}} \sigma_i^{\text{Fiducial}} \tag{5.4}$$

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

$$N_i^{\text{exp}}(\mu, \boldsymbol{\theta}) = N_i^{\text{exp}}(\mu, \mathcal{L}_0, \Delta_{\mathcal{L}}, \boldsymbol{\theta}_s, \boldsymbol{\theta}_b) = \mu \cdot \left( \mathcal{L}(\mathcal{L}_0, \Delta_{\mathcal{L}}) \cdot \sigma_i^{\text{Fiducial}} \cdot \varepsilon_i(\boldsymbol{\theta}_s) \right) + \sum_{\text{bkg}} N_i^{\text{bkg}}(\boldsymbol{\theta}_b) \tag{5.5}$$

where $\varepsilon_i$ is the efficiency measured in each bin as discussed in section 5.5.2.4 and $\sigma_i^{\text{Fiducial}}$ is the fiducial cross-section in each bin discussed in section **??**. The individual background expectations in a given bin/channel, $i$, are expressed simply by the number of events for

a given background as $N_i^{\text{bkg}}$. 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 signal efficiency shape parameters are referred to as $\boldsymbol{\theta}_s$ while 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 $\mu$ 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 [21]. The measurement and uncertainty are evaluated by using the shape of the profile likelihood ratio [35] which is a function of the data and the signal strength.

### 5.6.2   Profile Likelihood Ratio

The likelihood used is constructed as follows:

$$L(\mu, \boldsymbol{\theta}) = \text{Gaus}(\mathcal{L}_0 | \mathcal{L}, \Delta_{\mathcal{L}}) \prod_{i \in \text{Chan}} \text{Pois}(N_i^{obs} | N_i^{exp}(\mu, \boldsymbol{\theta})) \prod_{j \in \text{Sys}} \text{Gaus}(\theta_j^0 | \theta_j, 1) \qquad (5.6)$$

using the HistFactory tool developed within ATLAS [22]. 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:

$$-2 \ln \lambda(\mu) = -2 \ln \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \qquad (5.7)$$

Note that it no longer depends on the nuisance parameters, $\boldsymbol{\theta}$, and instead depends only on $\mu$. 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. Meanwhile, the numerator is the conditional ML which depends on $\mu$ and evaluated at the conditional ML estimator for the set of nuisance parameters, $\hat{\hat{\boldsymbol{\theta}}}$, which itself depends on $\mu$. 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, $\mu$. 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, while 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.

### 5.6.3 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:

$$q_0 = \begin{cases} -2\ln\lambda(0), & \hat{\mu} \geq 0 \\ 0, & \hat{\mu} < 0 \end{cases} \tag{5.8}$$

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:

$$p_0 = \int_{q_{0,\text{obs}}}^{\infty} f(q_0 | \mu = 0) \, \mathrm{d}q_0 \tag{5.9}$$

where $q_{0,\text{obs}}$ 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:

$$Z = \Phi^{-1}(1 - p_0) \tag{5.10}$$

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

The distribution of $q_0$ is shown in Fig. 5.15 for the combination. The observed null p-value is found to be 0.27 for the combination which corresponds to a significance of $0.63\sigma$. One may compare to this to an expected p-value of 0.26 corresponding to a significance of $0.65\sigma$.

### 5.6.4 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 [40] can be used [35] to determine that the range of the uncertainty for a given
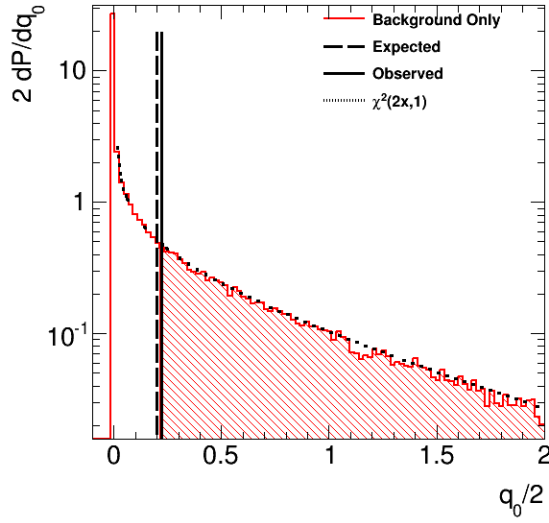
Figure 5.15: PDF of the background only hypothesis as a function of $q_0$ for the combination of all three channels. PDFs are determined using toy MC. The solid black line represents the observed value of $q_0$ seen in the data. The shaded area above this line represents the null p-value or the integral of the background hypothesis in the signal-like region. The dotted black curve shows a $\chi^2$ distribution for 1 degree of freedom with which it can be seen is a good approximation of the the background only PDF.

number of Gaussian $\sigma$ can be related directly to the negative profile log likelihood. In particular, for a $1\sigma$ 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 $\mu$ 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. 5.16. The expected value and uncertainties for the fiducial cross-section is:

$$\sigma^{\text{Expected}} = 291.5^{+416}_{-322}(\text{stat})^{+303}_{-328}(\text{sys})\text{ab} \tag{5.11}$$

while the observed fiducial cross-section is:

$$\sigma^{\text{Observed:}} = 302.2^{+333}_{-320}(\text{stat})^{+313}_{-333}(\text{sys})\text{ab} \tag{5.12}$$
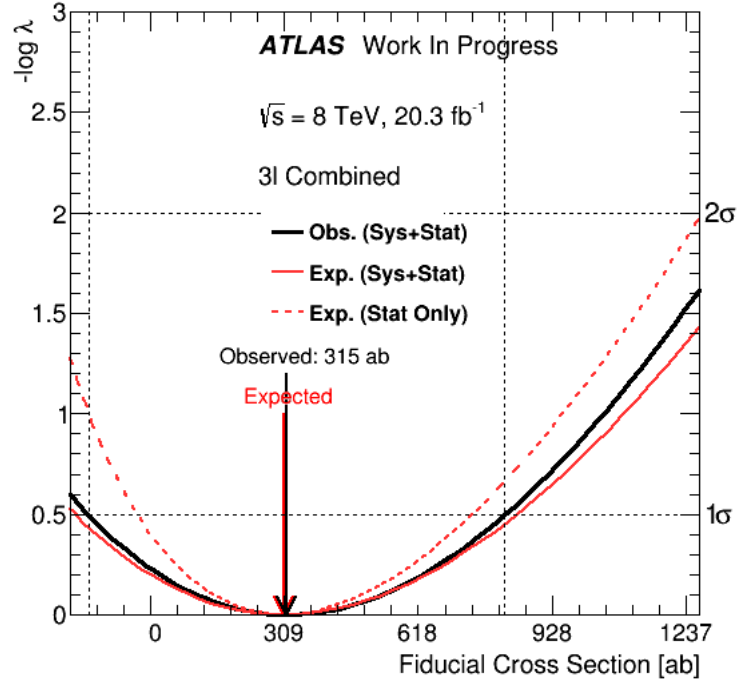


Figure 5.16: The profile likelihood contours evaluated as a function of the signal strength for the combination of all three channels. The observed (black) and expected (red) contours are shown when considering only statistical uncertainty (dashed line) and when considering both statistical and systematic uncertainties (solid line). The dotted black lines pinpoint the location of the 1 $\sigma$ and 2 $\sigma$ total Gaussian uncertainties on the measurement of the signal strength which corresponds to the minimum value of the contour.

## 5.7 Limits on anomalous Quartic Gauge Couplings

Chapter 6

# Conclusions

# List of Journal Abbreviations

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

# Bibliography

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

[2] 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.

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

[4] 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.

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

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

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

[8] 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.

[9] 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.

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

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

[12] ATLAS Collaboration. ATLAS tunes of PYTHIA6 and PYTHIA8 for MC11. ATL-PHYS-PUB-2011-009, 2011.

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

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

[15] 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.

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

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

[18] 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.

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

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

[21] 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.

[22] 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.

[23] 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.

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

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

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

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

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

[33] 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.

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

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

[36] 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.

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

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

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

[40] 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