# Observation of $W^\pm W^\pm W^\mp$ Production in Proton-Proton Collisions at $\sqrt{s}$ = 13 TeV with the ATLAS Detector

by

Wenhao Xu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in the University of Michigan
2021

Doctoral Committee:

  Professor Junjie Zhu, Chair
  Professor Ratindranath Akhoury
  Associate Professor Thomas Schwarz
  Associate Professor Kevin Wood
  Professor Bing Zhou

Wenhao Xu

xuwenhao@umich.edu

ORCID iD: 0000-0001-5661-1917

# ACKNOWLEDGMENTS

Ph.D. degree in physics is a significant milestone in my entire life. I could never achieve this without the help given by so many great people. First and foremost, I would like to express my deepest gratitude to my advisor Junjie Zhu. I have been working with Junjie since 2014 and he never fails to support me in pursuing my Ph.D. degree. He opened the door of my research at ATLAS and is always available if I need his guidance. His patience, encouragement and insight in research helped me a lot to overcome obstacles I encountered during my study.

It is my honor to become a member in the University of Michigan ATLAS group, where I met many fabulous professors. I want to give special thanks for Bing Zhou, who not only delivered great lectures on Electromagnetism but also patiently answered my questions with clear explanations and gave me detailed comments on my thesis. I also would like to thank Jianming Qian for his lectures on experimental particle physics which helped me a lot in my research and Thomas Schwartz for his useful suggestions on my thesis. My thanks also go to post-docs and students in the group. I want to especially thank Matthew Klein for his help on the CxAOD framework, Ismet Siral for his guidance on the previous $W^{\pm}W^{\pm}W^{\mp}$ analysis framework, Siyuan Sun for the Monte Carlo samples, Jinhong Wang for the help on TDS chip hardware project, and Jake Searcy for his support at the beginning of my Ph.D. research with detailed guidance to help me get familiar with several research projects at ATLAS. I also want to thank all of my colleagues in the group: Tom Cheng, Cong Geng, Liang Guan, Yicheng Guo, Bing Li, Hao Liu, Yanlin Liu, Shawn Mckee, Zirui Wang, Aaron White, Yusheng Wu, Zhaoxu Xi, Lailin Xu, Dongliang Zhang, Zhi Zheng and many others. I want to thank Ratindranath Akhoury and Kevin Wood for becoming my doctoral committee members and reviewing my thesis.

Due to the collaborative nature of the research in experimental particle physics, not all of the work in analyses is done by myself. For the $W^{\pm}W^{\pm}W^{\mp}$ analysis, I want to thank Joseph Lambert for his help on the data-driven systematic uncertainties, the boosted decision tree and the fit setup. I also would like to thank Vallary Bhopatkar for her work on the theoretical uncertainties. My thanks go to also the remaining group members: Braden Abbott, Muhammad Alhroob, Walter Hopkins, Jessica Metcalfe, and Leonid Serkin for their inspirations and suggestions on this analysis. I also want to thank the ATLAS electroweak group conveners, Heather Russell and Joany Manjarres, for their trusts and inviting me to become an analysis contact of the $W^{\pm}W^{\pm}W^{\mp}$ analysis. For the same-

# TABLE OF CONTENTS

# LIST OF FIGURES

xii

# LIST OF TABLES

# LIST OF ABBREVIATIONS

**ALEPH**  Apparatus for LEP Physics

**ALICE**  A Large Ion Collider Experiment

**ATLAS**  A Toroidal LHC Apparatus

**AUC**  Area Under the Curve

**BDT**  Boosted Decision Tree

**BSM**  Beyond Standard Model

**CART**  Classification and Regression Tree

**CERN**  European Organization for Nuclear Research

**CMS**  Compact Muon Solenoid

**CP**  Combined Performance

**CR**  Control Region

**CSCs**  Cathode Strip Chambers

**CTP**  Central Trigger Processor

**DELPHI**  Detector with Lepton, Photon and Hadron Identification

**ECIDS**  Electron Charge Identification Selector

**EMEC**  Electromagnetic End-cap Calorimeter

**EM**  Electromagnetic

**EOM**  Equation of Motion

**FCal**  Forward Calorimeter

**FE**  Front-End

**GRL**  Good Run List

**HEC**  Hadronic End-cap Calorimeter

**HLT**  High Level Trigger

**IBL**  Insertable B-Layer

**JEP**  Jet/Energy-sum Processor

**JVT**  Jet Vertex Tagger

**L1Calo** L1 Calorimeter

**L1Muon** L1 Muon

**L1Topo** L1 Topological

**L1** Level-1

**L2** Level-2

**LAr** Liquid-Argon

**LEP** Large Electron-Positron Collider

**LHCb** Large Hadron Collider Beauty

**LHC** Large Hadron Collider

**Linac** Linear Accelerator

**LO** Leading-Order

**LUCID** Luminosity Cherenkov Integrating Detector

**MC** Monte Carlo

**MDTs** Monitored Drift Tubes

**MUCTPI** Muon Central Trigger Processor Interface

**NLO** Next-to-Leading Order

**NNLO** Next-to-Next-to-Leading Order

**NP** Nuisance Parameter

**OPAL** Omni Purpose Apparatus for LEP

**OS** Opposite-Sign

**PDF** Parton Distribution Function

**PLV** Prompt Lepton Veto

*pp* Proton-Proton

**PSB** Proton Synchrotron Booster

**PS** Proton Synchrotron

**QCD** Quantum Chromodynamics

**QED** Quantum Electrodynamics

**QGC** Quartic Gauge Coupling

**ROC** Receiver Operating Characteristic

**ROD** ReadOut Driver

**ROIs** Regions of Interest

**ROS** ReadOut System

**RPCs** Resistive Plate Chambers

**SCT**  Semiconductor Tracker

**SFOS**  Same-Flavor Opposite-Sign

**SF**  Same-Flavor

**SM**  Standard Model

**SPS**  Super Proton Synchrotron

**SR**  Signal Region

**SSB**  Spontaneous Symmetry Breaking

**SS**  Same-Sign

**TDAQ**  Trigger and Data Acquisition

**TGC**  Triple Gauge Coupling

**TRT**  Transition Radiation Tracker

**UA1**  Underground Area 1

**UA2**  Underground Area 2

**VBF**  Vector Boson Fusion

**VBS**  Vector Boson Scattering

**VEV**  Vacuum Expectation Value

**VR**  Validation Region

# ABSTRACT

This thesis presents the first observation of inclusive $W^{\pm}W^{\pm}W^{\mp}$ production using 139 fb$^{-1}$ of proton-proton collision data collected by the ATLAS detector at $\sqrt{s} = 13$ TeV. The analysis is performed in two decay channels: the $W^{\pm}W^{\pm}W^{\mp} \to \ell^{\pm}\nu\ell^{\pm}\nu jj$ channel where two same-sign $W$ bosons decay leptonically into leptons (electrons and muons only) and neutrinos while the remaining $W$ boson decays hadronically into jets, and the $W^{\pm}W^{\pm}W^{\mp} \to \ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ channel where all three $W$ bosons decay leptonically. Two signal regions, $\ell^{\pm}\nu\ell^{\pm}\nu jj$ signal region and $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ signal region, are defined. In the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ signal region, events are selected with two same-sign leptons and at least two jets, while in the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ signal region, events are required to have three leptons which do not form a same-flavor opposite-sign pair. The $WZ$ process is the dominant background and its contribution is constrained by the collision data in dedicated $WZ$ control regions. Two $WZ$ normalization factors are then derived for different jet multiplicities in two control regions: less than two jets and at least two jets. In each signal region, boosted decision tree is used to further enhance the separation between signal and background events. A binned maximum-likelihood fit is then performed with the distributions of boosted decision tree discriminant score in the signal regions and trilepton invariant mass in the $WZ$ control regions.

In the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ signal region, $1,545$ events are observed in data with approximately 231 signal and $1,284$ background events expected. In the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ signal region, 79 events are observed in data with approximately 37 signal and 54 background events expected. The background-only hypothesis is rejected with an observed (expected) significance of $7.8\sigma$ ($5.7\sigma$) where $\sigma$ stands for the standard deviation. The total inclusive $W^{\pm}W^{\pm}W^{\mp}$ cross section is measured to be $0.807^{+0.128}_{-0.123}$ pb and is found to be compatible with the Standard Model prediction within $2.5\sigma$.

# CHAPTER 1

# Introduction

Developed in the latter half of the 20<sup>th</sup> century, the Standard Model (SM) of particle physics is a great success as it makes many predictions which have been observed and verified experimentally. With the observation of the Higgs boson by A Toroidal LHC Apparatus (ATLAS) [1] and Compact Muon Solenoid (CMS) [2] at European Organization for Nuclear Research (CERN) on July 4<sup>th</sup>, 2012, the last missing elementary particle in the SM has been found. The SM is believed to be one of the most fundamental theories in nature.

The SM is a quantum field theory based on the $SU(3)_C \bigotimes SU(2)_L \bigotimes U(1)_Y$ gauge symmetry group and describes the strong and electroweak interactions among elementary particles. In the SM, electroweak gauge bosons, which includes $W^\pm$, $Z$, and $\gamma$, can interact with each other due to the non-Abelian gauge symmetry of the SM electroweak sector. The electroweak Lagrangian generates Triple Gauge Coupling (TGC) including $WW\gamma$ and $WWZ$ as wells as Quartic Gauge Coupling (QGC) including $WWWW$, $WW\gamma\gamma$, $WWZ\gamma$ and $WWZZ$. Other couplings including neutral TGC ($ZZZ$, $ZZ\gamma$, $Z\gamma\gamma$) and neutral QGC ($ZZZZ$, $Z\gamma\gamma\gamma$, $ZZ\gamma\gamma$, $Z\gamma\gamma\gamma$, $\gamma\gamma\gamma\gamma$) are prohibited in the SM. Therefore, the precise measurement on these couplings can provide a critical test of the SM as well as indicates new physics if there is a derivation from the SM prediction, such as anomalous QGC.

Recently, with $\sqrt{s} = 13$ TeV Proton-Proton ($pp$) collision data at the Large Hadron Collider (LHC), ATLAS saw the first evidence of $WVV$ production ($V = W^\pm$, $Z$) with a significance of 4.1 standard deviations ($\sigma$) using 80 fb$^{-1}$ of the Run 2 data [3], and CMS observed $VVV$ production with a significance of 5.7$\sigma$ using 137 fb$^{-1}$ of the Run 2 data [4]. As for $WWW$ production, ATLAS and CMS observed significances of 3.2$\sigma$ and 3.3$\sigma$ respectively.

This thesis presents the first observation of inclusive $W^\pm W^\pm W^\mp$ production using the $pp$ collision data taken by the ATLAS detector at $\sqrt{s} = 13$ TeV. The total integrated luminosity is 139 fb$^{-1}$, which corresponds to the full Run 2 dataset at the LHC. As shown in Figure 1.1, there are four types of processes that can produce $W^\pm W^\pm W^\mp$ events at Leading-Order (LO) Quantum Chromodynamics (QCD) at the LHC: three $W$ bosons produced from quarks directly, an associated $W$

1

boson production with an intermediate $Z/\gamma^*$ boson decaying to two $W$ bosons, three $W$ bosons produced from a $WWWW$ QGC vertex, and three $W$ bosons produced from a $WH$ process where two Opposite-Sign (OS) $W$ bosons are from the decay of the Higgs boson. Since experimentally we can not really distinguish $WH$ production from other production, all these productions are considered as the $W^\pm W^\pm W^\mp$ signal in this analysis.
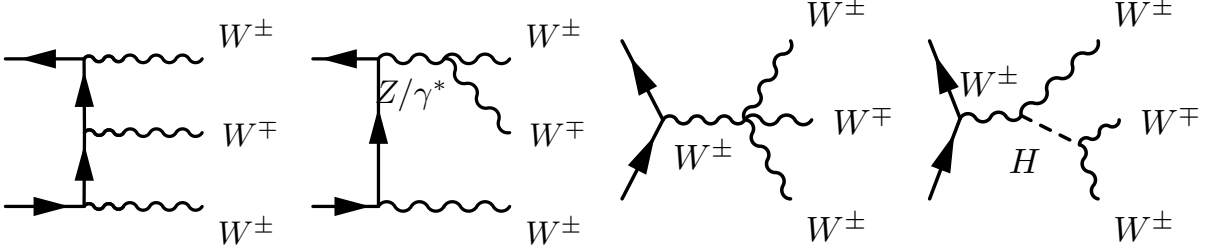


Figure 1.1: Representative Feynman diagrams for $W^\pm W^\pm W^\mp$ production at LO QCD at the LHC.

The study is performed in two decay channels: $W^\pm W^\pm W^\mp \rightarrow \ell^\pm \nu \ell^\pm \nu jj$ channel and $W^\pm W^\pm W^\mp \rightarrow \ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ channel, where $\ell$ stands for electron or muon and $j$ stands for jet which comes from the hadronization of quarks. Two Signal Regions (SRs), $\ell^\pm \nu \ell^\pm \nu jj$ SR and $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SR, are defined to study the $W^\pm W^\pm W^\mp$ production in each channel. In the $W^\pm W^\pm W^\mp \rightarrow \ell^\pm \nu \ell^\pm \nu jj$ channel, two Same-Sign (SS) $W$ bosons decay leptonically to neutrinos and electrons or muons whereas the other $W$ boson decays hadronically to two jets. Other major SM processes, such as the Drell-Yan and top pair production, are greatly reduced by the requirement of two SS leptons in the $\ell^\pm \nu \ell^\pm \nu jj$ SR. Additionally, the two leading jets have an invariant mass around the $W$ boson pole mass which also reduces the contribution from the SS $WW$ Vector Boson Scattering (VBS) process in the $\ell^\pm \nu \ell^\pm \nu jj$ SR. In the $W^\pm W^\pm W^\mp \rightarrow \ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ channel, all three $W$ bosons decay leptonically. The requirement of zero Same-Flavor Opposite-Sign (SFOS) lepton pair greatly suppresses contributions from $WZ$ process in the $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SR.

Monte Carlo (MC) simulation is used to estimate background processes with prompt leptons in the final state, while data-driven method is used to estimate background processes due to electron charge misidentification, jet-faked electrons, jet-faked muons, and photon-faked electrons. The $WZ$ process is the largest background in this analysis and its contribution is constrained by the data in dedicated $WZ$ Control Regions (CRs). Two $WZ$ normalization factors are derived with different jet multiplicities, less than two jets and at least two jets. The background modelings are cross-checked in various CRs and Validation Regions (VRs).

A machine learning algorithm, Boosted Decision Tree (BDT), is used to further enhance the separation between the signal and background processes in this analysis. Two BDTs models are trained for the $\ell^\pm \nu \ell^\pm \nu jj$ and $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SRs separately and then the distributions of the BDT

discriminant score in the SRs are combined with the trilepton invariant mass in the $WZ$ CRs to obtain the final significance and the signal strength using a binned maximum-likelihood fit. The signal strength is also used to determine the inclusive $W^\pm W^\pm W^\mp$ production cross section, $\sigma(pp \to W^\pm W^\pm W^\mp)$.

The organization of this thesis is as follows: Chapter 2 discusses elementary particles and the SM theory; Chapter 3 introduces the LHC and the ATLAS detector; Chapter 4 explains physics object reconstruction and identification procedures and summarizes the object-level selection cuts; Chapter 5 describes the data and MC samples used in the analysis; Chapter 6 summarizes the event-level selection cuts for the SRs; Chapter 7 explains the background modeling; Chapter 8 describes the BDT used for signal extraction; Chapter 9 lists all systematic uncertainties considered in this analysis; Chapter 10 explains the fitting procedure and presents the results; Chapter 11 summarizes the results of this analysis and draws the conclusion.

# CHAPTER 2

# Standard Model of Particle Physics

## 2.1 Elementary Particles

Elementary particles are the most fundamental building blocks of the universe. They are thought to have no internal structure. In the SM [5], all known elementary particles can be categorized into two types: elementary fermions and elementary bosons. Fermions are defined as particles that obey the Fermi-Dirac statistics and have odd half-integer spins (like 1/2, 3/2, 5/2, and so forth). For elementary fermions, they all have a spin of 1/2. On the other hand, bosons are defined as particles with integer spins (like 0, 1, 2, and so forth) and follow the Bose-Einstein statistics. For elementary bosons, they have a spin of 1 or 0. The summary of all elementary particles can be found in Figure 2.1.

Elementary fermions are matter particles as they are considered to be the ultimate building blocks of matter [5]. They are further categorized into quarks and leptons based on whether or not they carry color charge (red ($R$), green ($G$), blue ($B$)). Quarks have color charge and there are six flavors of quarks: up ($u$), charm ($c$) and top ($t$) quarks have an electric charge of $+\frac{2}{3}e$ ($e$ is the elementary charge) while down ($d$), strange ($s$), bottom ($b$) quarks have an electric charge of $-\frac{1}{3}e$. Leptons do not have color charge and there are six types of leptons: electron ($e$), muon ($\mu$) and tau ($\tau$) with an electric charge of $-e$ while electron neutrino ($\nu_e$), muon neutrino ($\nu_\mu$) and tau neutrino ($\nu_\tau$) with an electric charge of 0. Both quarks and leptons fall into three generations (also called families). Besides, each fermion also has its corresponding antiparticle which has the same mass but opposite charge compared to the original particle.

Elementary bosons can be categorized based on their spins. Gauge bosons are spin-1 force particles which can mediate fundamental interactions: photon ($\gamma$) is the mediator for the electromagnetic interaction between electrically-charged particles, gluon ($g$) is the mediator for the strong interaction that binds quarks inside the nucleus, $W^\pm$ and $Z$ bosons are the mediators for the weak interaction which participates in nuclear decays. Higgs boson ($H$) is a spin-0 scalar boson and is responsible for the intrinsic mass of all other elementary particles except for neutrinos.
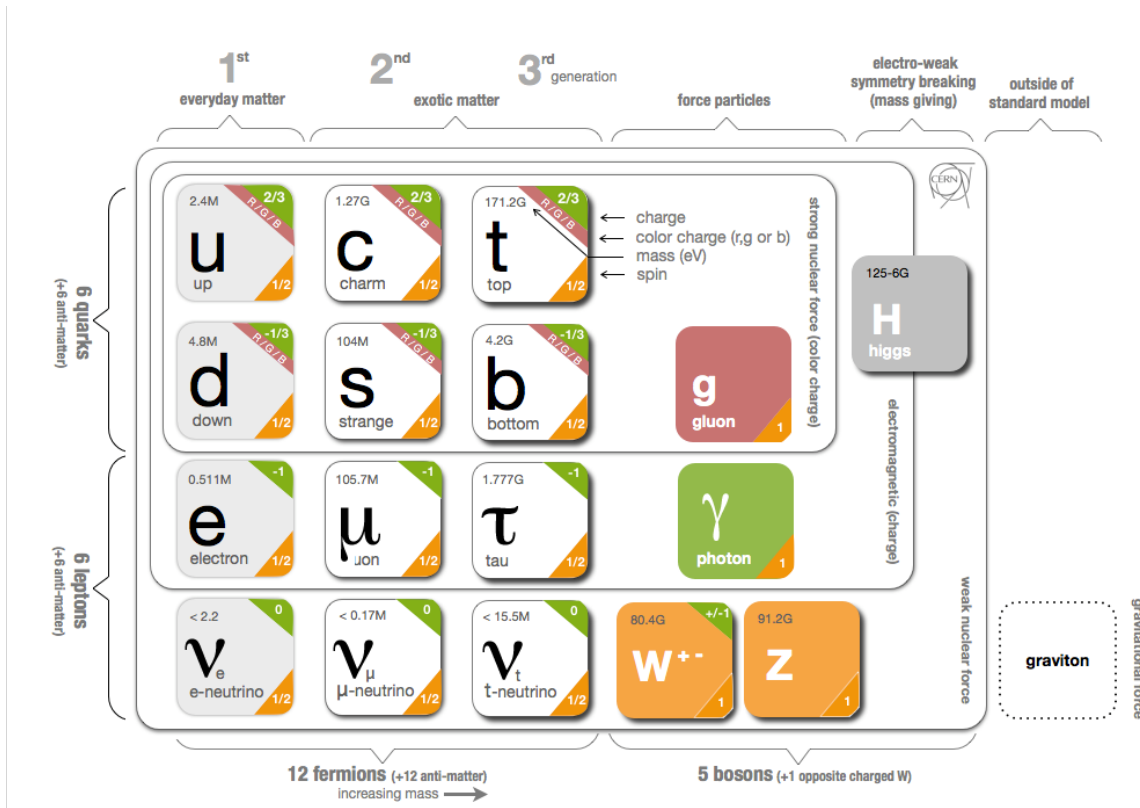
Figure 2.1: Summary of all known elementary particles in the SM plus a hypothetical graviton [6].

Additionally, there might be another elementary particle called the graviton which is hypothesized to be the mediator of the gravitational interaction [5]. If it exists, the graviton is expected to be massless and has a spin of 2. However, it is not included in the SM and has not been observed experimentally. In this thesis, the gravitational interaction is ignored as it is negligible at the atomic scale.

## 2.2 Lagrangian Formalism

In the SM, the electromagnetic, weak and strong interactions are explained through the exchange of corresponding gauge bosons using the Lagrangian formulation of a physical system. In the Lagrangian formalism, action is considered as an integration of four-dimensional space-time on the Lagrangian density $\mathcal{L}$ shown below [7]:

$$S = \int d^4x \mathcal{L}[\phi_i(x), \partial_\mu \phi_i(x)], \tag{2.1}$$

where $x^\mu$ is defined as $(t; x, y, z)$, $\phi_i(x)$ represents the field and $\partial_\mu \phi_i(x) = \partial \phi_i(x)/\partial x^\mu$ is the four-dimensional space-time derivative of the field. For simplicity, the Lagrangian density $\mathcal{L}$ is called as Lagrangian in the following text. Based on the principle of stationary action $\delta S = 0$, the Euler-Lagrange Equation of Motion (EOM) of the field is [7]:

$$\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \right) = 0. \tag{2.2}$$

Based on this EOM, Noether's theorem [8] can be generated to the four-dimensional space-time coordination: there is always a corresponding conserved quantity for every continuous symmetry transformation which leaves the Lagrangian stay the same.

In the SM, the symmetries are the global Poincaré symmetry from special relativity including translations, rotations and boosts as well as the local gauge symmetry based on symmetry groups $SU(3)_C \bigotimes SU(2)_L \bigotimes U(1)_Y$ where $C$ stands for the color of quarks, $L$ stands for the left-handed fields and $Y$ stands for the weak hypercharge [9]. The $SU(3)_C$ symmetry group governs the strong interaction whereas the $SU(2)_L \bigotimes U(1)_Y$ group governs the unified electroweak interaction. After the Spontaneous Symmetry Breaking (SSB), the electroweak group $SU(2)_L \bigotimes U(1)_Y$ becomes the electromagnetic subgroup $U(1)_{QED}$.

## 2.3 Quantum Electrodynamics Lagrangian

In the electromagnetic interaction, charged fermions interact with each other by exchanging photons. To produce the Dirac equation from the EOM in Eq. 2.2, the Lagrangian for a single fermion field ($\psi$) is [7]:

$$\mathcal{L}_0 = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi, \tag{2.3}$$

where $\gamma^\mu$ is the gamma matrix in the Dirac equation [10] and $m$ is the mass of the fermion. Under the $U(1)$ phase transformation we have:

$$\psi \xrightarrow{U(1)} \exp\{iQ\theta\}\psi, \tag{2.4}$$

where $Q\theta$ is an arbitrary real constant [7]. The Lagrangian $\mathcal{L}_0$ in Eq. 2.3 is invariant under the global $U(1)$ transformation where $\theta$ is a constant. However, $\mathcal{L}_0$ is not local $U(1)$ invariant if $\theta$ is a function of the space-time, i.e. $\theta = \theta(x)$. This is due to the extra term from the derivative of $\theta$:

$$\partial_\mu \psi \xrightarrow{U(1)} \exp\{iQ\theta\}(\partial_\mu + iQ\partial_\mu \theta)\psi. \tag{2.5}$$

To restore the local $U(1)$ phase invariance, an additional spin-1 photon field $A_\mu$ is introduced and transforms under the $U(1)$ transformation as:

$$A_\mu \xrightarrow{U(1)} A_\mu + \frac{1}{e}\partial_\mu\theta, \tag{2.6}$$

where $e$ is the electron charge [7]. An additional interaction term between the photon field and the fermion, $eQA_\mu\bar{\psi}\gamma^\mu\psi$, is added to the original Lagrangian to cancel out the term with the derivative of $\theta$ in Eq. 2.5, as well as the kinetic term of the photon field $A_\mu$:

$$\mathcal{L}_{\text{Kin}} = -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) \equiv -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.7}$$

where

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{2.8}$$

It can be shown that the kinetic term is gauge-invariant but the mass term of the field, $\frac{1}{2}m^2 A^\mu A_\mu$, is not. Therefore the photon field $A_\mu$ needs to be massless to keep the gauge invariance of the Lagrangian. The Lagrangian then becomes [7]:

$$\mathcal{L}_{\text{QED}} = \mathcal{L}_0 + eQA_\mu\bar{\psi}\gamma^\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} = i\bar{\psi}\gamma^\mu\mathcal{D}_\mu\psi - m\bar{\psi}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.9}$$

where the covariant derivative $\mathcal{D}_\mu$ is defined as:

$$\mathcal{D}_\mu \equiv \partial_\mu - ieQA_\mu, \tag{2.10}$$

and follows the local $U(1)$ transformation in the same way as $\psi$ in Eq. 2.4:

$$\mathcal{D}_\mu \xrightarrow{U(1)} \exp\{iQ\theta\}\mathcal{D}_\mu. \tag{2.11}$$

With this definition, the Lagrangian in Eq. 2.9 is local $U(1)$ invariant and is the Quantum Electrodynamics (QED) Lagrangian. The massless photon is introduced to satisfy the local gauge invariance.

## 2.4 Quantum Chromodynamics Lagrangian

In the strong interaction, quarks interact with each other through the exchange of gluons. If $q_f^\alpha$ is used to represent a quark field of color $\alpha$ ($\alpha = R, G, B$) and flavor $f$ ($f = u, d, c, s, t, b$), the $q_f$ vector

in the color space can then be written as:

$$q_f \equiv \begin{pmatrix} q_f^R \\ q_f^G \\ q_f^B \end{pmatrix}. \tag{2.12}$$

The Lagrangian of the quark field is expressed similarly to Eq. 2.3 [7] :

$$\mathcal{L}_0 = \sum_f (i\bar{q}_f \gamma^\mu \partial_\mu q_f - m\bar{q}_f q_f). \tag{2.13}$$

The $SU(3)_C$ transformation in the color space can be written as:

$$q_f \xrightarrow{SU(3)_C} \exp\{i\lambda_a \theta^a(x)/2\}q_f \qquad (a = 1, 2, ..., 8), \tag{2.14}$$

where $\lambda_a$ is one of the eight Gell-Mann matrices [7] and $\theta^a(x)$ is phase factor. Similar to Section 2.3, the Lagrangian is required to be invariant under local $SU(3)_C$ transformation. Eight spin-1 massless gluon fields, $G_\mu^1$, $G_\mu^2$, ..., $G_\mu^8$, have to be introduced. To restore the gauge invariance, the covariant derivative is constructed to transform as:

$$\mathcal{D}_\mu \equiv \partial_\mu - ig_s \frac{\lambda_a}{2} G_\mu^a \equiv \partial_\mu - ig_s G_\mu, \tag{2.15}$$

where $G_\mu \equiv \lambda_a G_\mu^a/2$ and $g_s$ is a unique $SU(3)_C$ coupling. The strong coupling constant $\alpha_S$ is defined as $\alpha_S \equiv g_s^2/(4\pi)$ and used commonly in experiments. The gauge-invariant kinetic term of the gluon fields can be express as:

$$\mathcal{L}_{\text{Kin}} = -\frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}, \tag{2.16}$$

where $G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f_{abc} G_\mu^b G_\nu^c$ and $f_{abc}$ is the $SU(3)_C$ structure constant from the commutation relation [7]:

$$\left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2}\right] = if_{abc}\frac{\lambda_c}{2}. \tag{2.17}$$

The $SU(3)_C$-invariant QCD Lagrangian thus can be written as:

$$\mathcal{L}_{\text{QCD}} = \sum_f (i\bar{q}_f \gamma^\mu \mathcal{D}_\mu q_f - m\bar{q}_f q_f) - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}. \tag{2.18}$$

With similar arguments as we have in Section 2.3, all eight gluons are required to be massless.

## 2.5  Weak Interaction Lagrangian

The local $SU(2)$-invariant Lagrangian can be derived following the same procedure described in Sections 2.3 and 2.4 by starting with the free Lagrangian for the weak field:

$$\mathcal{L}_0 = i\bar{\psi}\gamma^\mu \partial_\mu \psi, \tag{2.19}$$

where $\psi$ is a doublet of fermions:

$$\psi \equiv \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}. \tag{2.20}$$

Note that the mass terms of fermions are not included as they can spoil the procedure to restore the local $SU(2)$ invariance [7]. The $SU(2)$ transformation of the doublet can be written as:

$$\psi \xrightarrow{SU(2)} \exp\{i\sigma_a \theta^a(x)/2\}\psi \qquad (a = 1, 2, 3), \tag{2.21}$$

where $\sigma_a$ is one of the three Pauli matrices [7] and $\theta^a(x)$ is the phase factor. To satisfy local $SU(2)$ invariance, three massless spin-1 gauge bosons, $W_\mu^1$, $W_\mu^2$, $W_\mu^3$, are introduced with the gauge-invariant kinetic term written as:

$$\mathcal{L}_{\text{Kin}} = -\frac{1}{4} W_{\mu\nu}^a W_a^{\mu\nu}, \tag{2.22}$$

where

$$W_{\mu\nu}^a \equiv \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g\epsilon_{abc} W_\mu^b W_\nu^c \tag{2.23}$$

with $g$ as the $SU(2)$ coupling constant [7] and $\epsilon_{abc}$ as the Levi-Civita symbol from the $SU(2)$ commutation relation [10]:

$$\left[\frac{\sigma_a}{2}, \frac{\sigma_b}{2}\right] = i\epsilon_{abc}\frac{\sigma_c}{2}. \tag{2.24}$$

The covariant derivative is then constructed to have $SU(2)$ transformation in the same way as the doublet $\psi$ in Eq. 2.21 and can be expressed as:

$$\mathcal{D}_\mu \equiv \partial_\mu - ig\frac{\sigma_a}{2} W_\mu^a \equiv \partial_\mu - igW_\mu, \tag{2.25}$$

where $W_\mu \equiv \frac{\sigma_a}{2} W_\mu^a$. The $SU(2)$-invariant weak interaction Lagrangian (also called the Yang-Mills Lagrangian) [11] is the following:

$$\mathcal{L}_{\text{weak}} = i\bar{\psi}\gamma^\mu \mathcal{D}_\mu \psi - \frac{1}{4} W_{\mu\nu}^a W_a^{\mu\nu}. \tag{2.26}$$

Similarly, the gauge bosons, $W_\mu^1$, $W_\mu^2$, $W_\mu^3$, are required to be massless.

## 2.6 Electroweak Lagrangian

In the SM, electromagnetic interaction and weak interaction are unified as electroweak interaction which is mediated by three massive gauge bosons ($W^\pm$ and $Z$) and one massless photon ($\gamma$). Chirality is defined to split a fermion $\psi$ into left- and right-handed fields which can be express as:

$$\psi_L = P_L\psi = \frac{1-\gamma^5}{2}\psi,$$
$$\psi_R = P_R\psi = \frac{1+\gamma^5}{2}\psi,$$

(2.27)

where $\gamma^5$ is the fifth gamma matrix [10]. Based on experiments, $W^\pm$ bosons only couple to left-handed fermions (or right-handed antifermions) and $Z$ boson couples to left- and right-handed fermions differently, while photon couples to left- and right-handed fermions in the same way [7]. Therefore, the left-handed fermions are put as $SU(2)$ doublets:

$$\psi_1 = \begin{pmatrix} u \\ d \end{pmatrix}_L \quad \text{or} \quad \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L,$$

(2.28)

where $(u,d)$ can be $(u,d)$, $(c,s)$ or $(t,b)$ and $\ell$ can be $e$, $\mu$ or $\tau$. Using the same notation, the right-handed fermions are all singlets under $SU(2)$:

$$\psi_2 = u_R,$$
$$\psi_3 = d_R \quad \text{or} \quad \ell_R.$$

(2.29)

Hypothetical right-handed neutrinos (sterile neutrinos), $\nu_{eR}$, $\nu_{\mu R}$, $\nu_{\tau R}$, are not considered since they are not observed in experiments. From Section 2.5. the simplest group with a doublet is $SU(2)$. The electromagnetic interaction needs a $U(1)$ group based on Section 2.3. Hence the symmetry group considered for the electroweak Lagrangian is $SU(2)_L \bigotimes U(1)_Y$ where $L$ stands for left-handed fields as $SU(2)_L$ transformation is only applied on the left-handed doublet $\psi_1$, and $Y$ refers to the weak hypercharge which can be express as $Y = Q - T_3$ with $Q$ as the electromagnetic charge operator and $T_3$ as the third component of weak isospin, i.e. $T_3 \equiv \sigma_3/2$ [7].

For simplicity, only one generation of quarks or leptons is discussed here. Similar to Section 2.5, the free Lagrangian without mass term is proposed to be the electroweak Lagrangian:

$$\mathcal{L}_0 = \sum_{j=1}^{3} i\bar{\psi}_j \gamma^\mu \partial_\mu \psi_j.$$

(2.30)

The $SU(2)_L \otimes U(1)_Y$ transformation can be written by combining Eqs. 2.4 and 2.21:

$$\psi_1 \xrightarrow{SU(2)_L \otimes U(1)_Y} \exp\{iy_1\beta\}\exp\{i\sigma_a\alpha^a/2\}\psi_1 \qquad (a = 1,2,3),$$

$$\psi_2 \xrightarrow{SU(2)_L \otimes U(1)_Y} \exp\{iy_2\beta\}\psi_2, \qquad (2.31)$$

$$\psi_3 \xrightarrow{SU(2)_L \otimes U(1)_Y} \exp\{iy_3\beta\}\psi_3,$$

where $y_1$, $y_2$, $y_3$ are weak hypercharge analogous to $Q$ in Eq. 2.4, and $\beta$, $\alpha^a$ are phase factors which are the same as $\theta$, $\theta^a$ in Eqs. 2.4 and 2.21.

To make the Lagrangian local gauge-invariant under $SU(2)_L \otimes U(1)_Y$ with $\alpha^a = \alpha^a(x)$ and $\beta = \beta(x)$, four massless spin-1 gauge bosons, $W_\mu^1$, $W_\mu^2$, $W_\mu^3$, and $B_\mu$, are introduced as there are four phase parameters, $\alpha^a$ and $\beta$. By combining Eqs. 2.7 and 2.22, the total gauge-invariant kinetic term is:

$$\mathcal{L}_{\text{Kin}} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu}, \qquad (2.32)$$

where $W_{\mu\nu}^a$ is defined in Eq. 2.23 and $B_{\mu\nu}$ is defined similarly to $F_{\mu\nu}$ in Eq. 2.8, i.e. $B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu$. By combining Eqs. 2.10 and 2.25, the covariant derivatives can be summarized as:

$$\mathcal{D}_\mu\psi_1 \equiv \left[\partial_\mu - igW_\mu - ig'y_1B_\mu\right]\psi_1,$$

$$\mathcal{D}_\mu\psi_2 \equiv \left[\partial_\mu - ig'y_2B_\mu\right]\psi_2, \qquad (2.33)$$

$$\mathcal{D}_\mu\psi_3 \equiv \left[\partial_\mu - ig'y_3B_\mu\right]\psi_3,$$

where $g$ and $g'$ are coupling constants. The electroweak Lagrangian (on a single generation of fermions) that is invariant under both global and local $SU(2)_L \otimes U(1)_Y$ transformation can be written as:

$$\mathcal{L}_{\text{EWK}} = \sum_{j=1}^{3} i\bar{\psi}_j\gamma^\mu\mathcal{D}_\mu\psi_j - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu}. \qquad (2.34)$$

## 2.7 Higgs Mechanism

Although the gauge-invariant electroweak Lagrangian is constructed in Eq. 2.34, the fields of fermions and gauge bosons are massless in the Lagrangian which contradicts the experimental results. To fix this, Higgs mechanism [12, 13] is proposed by introducing a scalar Higgs field which is a complex $SU(2)_L$ doublet:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \qquad (2.35)$$

The $SU(2)_L \otimes U(1)_Y$ gauge-invariant Lagrangian is built to be [7]:

$$\mathcal{L}_{\text{Higgs}} = (\mathcal{D}_\mu \Phi)^\dagger \mathcal{D}^\mu \Phi - V(\Phi, \Phi^\dagger) = (\mathcal{D}_\mu \Phi)^\dagger \mathcal{D}^\mu \Phi - \mu^2 \Phi^\dagger \Phi - h(\Phi^\dagger \Phi)^2, \tag{2.36}$$

where $h > 0$ and $\mu^2 < 0$. The covariant derivative can be expressed as:

$$\mathcal{D}_\mu \equiv \partial_\mu - igW_\mu - ig'y_\Phi B_\mu, \tag{2.37}$$

with the weak hypercharge of the field as $y_\Phi$, similar to the first transformation in Eq. 2.31.

There are infinite degenerate states with the minimum Higgs potential energy $V(\Phi, \Phi^\dagger)$ as long as $\Phi^\dagger \Phi = -\mu^2/(2h)$. Without loss of generality, the Vacuum Expectation Value (VEV) of $\Phi$ is chosen as:

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \upsilon \end{pmatrix}, \tag{2.38}$$

where $\upsilon \equiv \sqrt{-\frac{\mu^2}{h}}$. In this way the symmetry $SU(2)_L \otimes U(1)_Y$ spontaneously breaks and the remaining gauge symmetry is the electromagnetic subgroup $U(1)_{\text{QED}}$. The doublet $\Phi$ can be written as its deviation from the VEV:

$$\Phi = \exp\{i\sigma_a \theta^a/2\} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \upsilon + H \end{pmatrix}, \tag{2.39}$$

where $H$ is the physical Higgs boson. By taking the unitary gauge, $\theta^a = 0$, the kinetic term in the Higgs Lagrangian shown in Eq. 2.36 can be written as

$$\begin{aligned} \mathcal{L}_{\text{Higgs-Kin}} &\equiv (\mathcal{D}_\mu \Phi)^\dagger \mathcal{D}^\mu \Phi \\ &= \frac{1}{2}(\partial_\mu H)(\partial^\mu H) + \frac{(\upsilon + H)^2}{4}\left[W_\mu^1 W^{1\mu} + W_\mu^2 W^{2\mu} + \frac{1}{2}(gW_\mu^3 - g'B_\mu)(gW^{3\mu} - g'B^\mu)\right]. \end{aligned} \tag{2.40}$$

This equation implies that the four gauge bosons, $W_\mu^1, W_\mu^2, W_\mu^3, B_\mu$, are not mass eigenstates. Therefore, the physical $W^\pm$ bosons ($W_\mu^\pm$) are introduced as the linear combinations of electrically-charged fields of $W_\mu^1$ and $W_\mu^2$:

$$W_\mu^\pm \equiv \frac{1}{2}(W_\mu^1 \mp iW_\mu^2). \tag{2.41}$$

The physical $Z$ boson ($Z_\mu$) and photon ($A_\mu$) are also introduced here by mixing electrically-neutral fields, $W_\mu^3$ and $B_\mu$, and can be expressed as a rotation with the weak mixing angle $\theta_W$:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}. \tag{2.42}$$

The weak mixing angle $\theta_W$ can be calculated from the coupling constants $g$ and $g'$:

$$\begin{aligned}
\sin\theta_W &= \frac{g'}{\sqrt{g^2 + g'^2}}, \\
\cos\theta_W &= \frac{g}{\sqrt{g^2 + g'^2}}.
\end{aligned} \tag{2.43}$$

Using Eqs. 2.41–2.43, the kinetic term in the Higgs Lagrangian in Eq. 2.40 can be rewritten as:

$$\mathcal{L}_{\text{Higgs-Kin}} = \frac{1}{2}(\partial_\mu H)(\partial^\mu H) + \frac{(v+H)^2}{4}\left[g^2 W_\mu^+ W^{-\mu} + \frac{1}{2}(g^2 + g'^2)Z_\mu Z^\mu\right]. \tag{2.44}$$

The masses of $W^\pm$ bosons and $Z$ boson are predicted from the terms proportional to $v^2$:

$$m_W = \frac{gv}{2}, \quad m_Z = \frac{v\sqrt{g^2 + g'^2}}{2}, \quad m_W/m_Z = \cos\theta_W. \tag{2.45}$$

Additionally, the absence of the photon field mass term in Eq. 2.44 indicates that the photon field remains massless.

Using the masses of $W^\pm$ bosons and $Z$ boson in Eq. 2.45, the kinetic term in the Higgs Lagrangian can be further expanded as:

$$\begin{aligned}
\mathcal{L}_{\text{Higgs-Kin}} =& \frac{1}{2}(\partial_\mu H)(\partial^\mu H) + m_W^2 W_\mu^+ W^{-\mu} + \frac{m_Z^2}{2}Z_\mu Z^\mu + \\
& \frac{2m_W^2}{v}W_\mu^+ W^{-\mu}H + \frac{m_Z^2}{v}Z_\mu Z^\mu H + \frac{m_W^2}{v}W_\mu^+ W^{-\mu}H^2 + \frac{m_Z^2}{2v}Z_\mu Z^\mu H^2.
\end{aligned} \tag{2.46}$$

The second line of the above equation are the couplings of the Higgs boson and gauge bosons ($W^+W^-H$, $ZZH$, $W^+W^-HH$ and $ZZHH$). The Feynman diagrams and coupling terms are summarized in Table 2.1 and Feynman rules can then be derived by taking the metric tensor into account.

By defining $M_H = \sqrt{-2\mu^2} = \sqrt{2}hv$, the Higgs Lagrangian in Eq. 2.36 can be written in terms of the Higgs boson $H$:

$$\begin{aligned}
\mathcal{L}_{\text{Higgs}} =& \frac{1}{4}hv^4 + \frac{1}{2}(\partial_\mu H)(\partial^\mu H) - \frac{M_H^2}{2}H^2 - \frac{M_H^2}{2v}H^3 - \frac{M_H^2}{8v^2}H^4 + \\
& M_W^2 W_\mu^+ W^{-\mu}\left(1 + \frac{H}{v}\right)^2 + \frac{M_Z^2}{2}Z_\mu Z^\mu\left(1 + \frac{H}{v}\right)^2.
\end{aligned} \tag{2.47}$$

Additionally, the fermion masses are generated from the coupling between fermions and the scalar Higgs filed $\Phi$ which are called as Yukawa coupling. The Lagrangian of Yukawa coupling is gauge-

| Vertex Types | Feynman Diagrams | Coupling Coefficients |
|:---:|:---:|:---:|
| $W^+W^-H$ | | $\dfrac{2m_W^2}{v}$ |
| $ZZH$ | | $\dfrac{m_Z^2}{v}$ |
| $W^+W^-HH$ | | $\dfrac{m_W^2}{v}$ |
| $ZZHH$ | | $\dfrac{m_Z^2}{2v}$ |

Table 2.1: The coupling terms of the Higgs boson and gauge bosons in the SM Lagrangian.

invariant after the SSB and can be summarized in the unitary gauge:

$$\mathcal{L}_{\text{Yukawa}} = -\sum_f m_f \left(1 + \frac{H}{v}\right)\bar{f}f, \tag{2.48}$$

where $f$ stands for a fermion and $m_f$ is the fermion mass.

## 2.8 Standard Model Lagrangian

Combining with Lagrangians introduced in Sections 2.4, 2.6 and 2.7, the completed SM Lagrangian can be expressed as follows [7, 9]:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EWK}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}. \tag{2.49}$$

$\mathcal{L}_{\text{QCD}}$ is the QCD sector of the SM Lagrangian and defined in Eq. 2.18 with mass terms of quarks as well as kinetic terms of both quarks and massless gluons. $\mathcal{L}_{\text{EWK}}$, $\mathcal{L}_{\text{Higgs}}$, and $\mathcal{L}_{\text{Yukawa}}$ are the electroweak sector of the SM Lagrangian and defined in Eq. 2.34, Eq. 2.47, and Eq. 2.48, respectively. $\mathcal{L}_{\text{EWK}}$ contains kinetic terms of fermions and electroweak gauge bosons ($W^{\pm}$, $Z$, and $\gamma$). $\mathcal{L}_{\text{Higgs}}$ has kinetic term and mass term of the Higgs boson as well as mass terms of $W^{\pm}$ bosons and $Z$ boson after the SSB. $\mathcal{L}_{\text{Yukawa}}$ contains the mass terms of fermions after the SSB except for neutrinos.

## 2.9 Triple and Quartic Gauge Couplings

The TGC and QGC terms can be generated by expanding the kinetic terms of the electroweak gauge bosons in Eq. 2.32 and expressed by using the physical $W^{\pm}$ bosons, $Z$ boson and photon fields defined in Eqs. 2.41 and 2.42.

The terms of TGC are [7]:

$$
\begin{aligned}
\mathcal{L}_3 = &-ie\cot\theta_W\left[(\partial^\mu W^{-\nu} - \partial^\nu W^{-\mu})W_\mu^+ Z_\nu - (\partial^\mu W^{+\nu} - \partial^\nu W^{+\mu})W_\mu^- Z_\nu + W_\mu^- W_\nu^+(\partial^\mu Z^\nu - \partial^\nu Z^\mu)\right] \\
&-ie\left[(\partial^\mu W^{-\nu} - \partial^\nu W^{-\mu})W_\mu^+ A_\nu - (\partial^\mu W^{+\nu} - \partial^\nu W^{+\mu})W_\mu^- A_\nu + W_\mu^- W_\nu^+(\partial^\mu A^\nu - \partial^\nu A^\mu)\right].
\end{aligned}
\tag{2.50}
$$

The terms of QGC are [7]:

$$
\begin{aligned}
\mathcal{L}_4 = &-\frac{e^2}{2\sin^2\theta_W}\left[W_\mu^+ W^{-\mu} W_\nu^+ W^{-\nu} - W_\mu^+ W^{\mu+} W_\nu^- W^{-\nu}\right] \\
&- e^2\cot^2\theta_W\left[W_\mu^+ W^{-\mu} Z_\nu Z^\nu - W_\mu^+ Z^\mu W_\nu^- Z^\nu\right] \\
&- e^2\cot\theta_W\left[2W_\mu^+ W^{-\mu} Z_\nu A^\nu - W_\mu^+ Z^\mu W_\nu^- A^\nu - W_\mu^+ A^\mu W_\nu^- Z^\nu\right] \\
&- e^2\left[W_\mu^+ W^{-\mu} A_\nu A^\nu - W_\mu^+ A^\mu W_\nu^- A^\nu\right].
\end{aligned}
\tag{2.51}
$$

The coupling terms of TGC and QGC are summarized in Table 2.2 and Feynman rules can be derived after including the metric tensor. The allowed TGC vertices are $W^+W^-Z$ and $W^+W^-\gamma$ and the allowed QGC vertices are $W^+W^-W^+W^-$, $W^+W^-ZZ$, $W^+W^-Z\gamma$, and $W^+W^-\gamma\gamma$. Neutral TGC vertices ($ZZZ$, $ZZ\gamma$, $Z\gamma\gamma$ and $\gamma\gamma\gamma$) and QGC vertices ($ZZZZ$, $ZZZ\gamma$, $ZZ\gamma\gamma$, $Z\gamma\gamma\gamma$ and $\gamma\gamma\gamma\gamma$) with only photons and $Z$ bosons do not appear in the SM Lagrangian and thus are prohibited in the SM.

The TGC and QGC are studied experimentally through diboson production, triboson production, VBS, and Vector Boson Fusion (VBF). Since the triboson production cross sections are small at low center-of-mass energies, they have only been studied in the $pp$ collisions at the LHC. These studies are categorized based on the types of the three bosons produced in the $pp$ collision. Recent studies on the production of three massive gauge bosons are published with the Run 2 dataset

while studies of electroweak triboson production including photon are published only with the Run 1 dataset.

For studies with three massive gauge bosons, different types of productions examine different combinations of TGC and QGC terms: $WWW$ analysis tests on $WWZ$, $WW\gamma$ and $WWWW$ vertices; $WWZ$ analysis tests on $WWZ$, $WW\gamma$ and $WWZZ$ vertices; and $WZZ$ analysis tests on $WWZ$ and $WWZZ$ vertices. There is no SM-allowed TGC or QGC vertex for $ZZZ$ analysis but anomalous TGC and QGC can be searched with this analysis. Using 80 fb$^{-1}$ of 13 TeV data taken in Run 2, ATLAS provided the first evidence for $WVV$ ($V = W$ or $Z$) production with a significance of $4.1\sigma$ and individual $WWW$ and $WVZ$ production each with a significance of $3.2\sigma$ [3]. Using 137 fb$^{-1}$ of data taken at 13 TeV, CMS first observed the combined production of $VVV$ with a significance of $5.7\sigma$ and evidences of individual $WWW$ and $WWZ$ with significances of $3.3\sigma$ and $3.4\sigma$ respectively, while $WZZ$ was found with a significance of $1.7\sigma$ and the significance of $ZZZ$ is $0.0\sigma$ [4].

Electroweak triboson production including photon are less sensitive to TGC and QGC as the photon can be easily generated in initial state and final state radiations. Similarly to three massive gauge boson production, different combinations of TGC and QGC terms are tested: $WW\gamma$ production studies on $WWZ\gamma$, $WW\gamma\gamma$, $WWZ$ and $WW\gamma$ vertices; $WZ\gamma$ production studies on $WWZ\gamma$, $WWZ$ and $WW\gamma$ vertices; $W\gamma\gamma$ production studies on $WW\gamma$, $WWZ\gamma$ and $WW\gamma\gamma$ vertices. There is no SM-allowed TGC or QGC vertex for $ZZ\gamma$, $Z\gamma\gamma$ or $\gamma\gamma\gamma$ production but similar to $ZZZ$ analysis, anomalous TGC and QGC can be searched with these analyses. Using 20.2 fb$^{-1}$ of data taken at 8 TeV in Run 1, ATLAS saw the $WV\gamma$ production with a significance of $1.4\sigma$ [14]. With 19.3 fb$^{-1}$ of data taken at 8 TeV in Run 1, CMS set an observed upper limit of 3.4 times larger than the SM prediction for $WV\gamma$ at 95% confidence level [15]. For $W\gamma\gamma$ production, ATLAS saw the evidence with a significance larger than $3\sigma$ using 20.3 fb$^{-1}$ of data taken at 8 TeV [16] while CMS saw it with a significance of $2.6\sigma$ using 19.4 fb$^{-1}$ of data taken at 8 TeV [17]. For $Z\gamma\gamma$ production, ATLAS observed the process with a significance of $6.3\sigma$ using 20.3 fb$^{-1}$ of data taken at 8 TeV [18] and CMS observed the process with a significance of $5.9\sigma$ using 19.4 fb$^{-1}$ of data taken at 8 TeV [17]. Limits are placed on parameters of anomalous TGC and QGC and no deviations from SM predictions are observed by both experiments. For $\gamma\gamma\gamma$ production, ATLAS measured the cross section using 20.2 fb$^{-1}$ of data taken at 8 TeV [19].
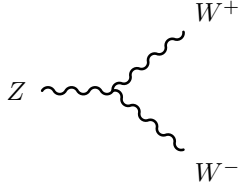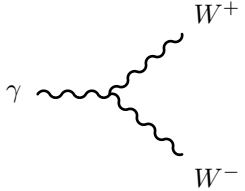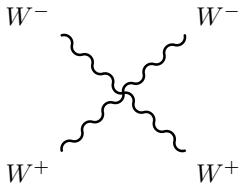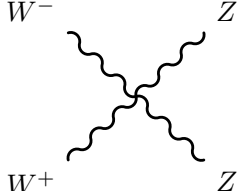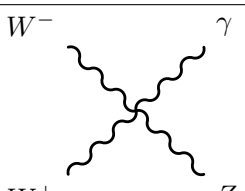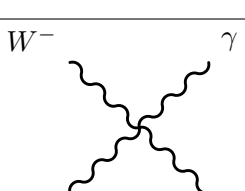
| Vertex Types | Feynman Diagrams | Coupling Coefficients |
|:---:|:---:|:---:|
| $W^+W^-Z$ |  | $-ie\cot\theta_W$ |
| $W^+W^-\gamma$ |  | $-ie$ |
| $W^+W^-W^+W^-$ |  | $-\dfrac{e^2}{2\sin^2\theta_W}$ |
| $W^+W^-ZZ$ |  | $-e^2\cot^2\theta_W$ |
| $W^+W^-Z\gamma$ |  | $-e^2\cot\theta_W$ |
| $W^+W^-\gamma\gamma$ |  | $-e^2$ |

Table 2.2: The TGC and QGC terms in the SM Lagrangian.

## 2.10   Effective Field Theory

Although the SM makes many predictions that have been verified experimentally, it is not a complete theory. Beyond Standard Model (BSM) theories are proposed to solve various problems in the SM, such as dark matter problem, neutrino mass problem, hierarchy problem, and so on. Effective field theory is an approximation of a theory at a given energy with appropriate degrees of freedom that can appear and ignoring the effect at higher energy. BSM theories can be studied using the SM effective field theory [20, 21] which can be expressed as:

$$\mathcal{L}_{\text{Eff}} = \mathcal{L}_{\text{SM}} + \sum_{d \geq 4} \sum_i \frac{f_i^{(d)}}{\Lambda^{d-4}} \mathcal{O}_i^{(d)}, \tag{2.52}$$

where $\mathcal{O}_i^{(d)}$ is the operator with dimension $d$, $f_i^{(d)}$ is the corresponding coupling, and $\Lambda$ is the energy scale of new physics.

Precise measurements on QGC vertices are important as derivations from the SM predictions, anomalous QGC, can indicate new physics. For simplicity, only interactions without derivatives of the gauge fields are considered. Therefore, the dimension of operators in Eq. 2.52 needs to be at least eight to only change QGC vertices without modifications on two or three weak gauge boson vertices. There are only two independent dimension-8 operators [22]:

$$\mathcal{L}_{S,0} = \frac{f_0}{\Lambda^4} [(D_\mu \Phi)^\dagger D_\nu \Phi] \times [(D^\mu \Phi)^\dagger D^\nu \Phi] \tag{2.53}$$

and

$$\mathcal{L}_{S,1} = \frac{f_1}{\Lambda^4} [(D_\mu \Phi)^\dagger D^\mu \Phi] \times [(D_\nu \Phi)^\dagger D^\nu \Phi], \tag{2.54}$$

where $\Phi$ is the Higgs field and $f_i$ is the anomalous QGC coefficient. By replacing $\Phi$ as its VEV ($v$), four gauge boson vertices can be generated from Eqs. 2.53 and 2.54. The $W^\pm W^\pm W^\mp$ analysis is sensitive to anomalous $WWWW$ QGC and the $WWWW$ terms can be expressed as:

$$\mathcal{L}^{WWW'W'} = g^2 \Delta c_0^{WW'} \mathcal{O}_0^{WW'} + g^2 \Delta c_1^{WW'} \mathcal{O}_1^{WW'}, \tag{2.55}$$

where $g$ is the $SU(2)_L$ coupling constant, $\Delta c_i^{WW'}$ is calculated as:

$$\Delta c_i^{WW'} = \frac{g^2 v^4 f_i}{8\Lambda^4}, \tag{2.56}$$

and $\mathcal{O}_i^{WW'}$ are the $WWWW$ operators:

$$
\begin{aligned}
\mathcal{O}_0^{WW'} &= g^{\alpha\beta} g^{\gamma\delta} \left[ W_\alpha^+ W_\beta^- W_\gamma^+ W_\delta^- \right], \\
\mathcal{O}_1^{WW'} &= g^{\alpha\beta} g^{\gamma\delta} \left[ W_\alpha^+ W_\beta^+ W_\gamma^- W_\delta^- \right].
\end{aligned}
\tag{2.57}
$$

However, unitary can be violated by adding operators to the SM Lagrangian, which can be solved with different methods. For example, a form factor can be introduced for anomalous QGC vertices [23, 24]:

$$
f_i \rightarrow \frac{f_i^0}{1 + s/\Lambda_{\mathrm{FF}}^2},
\tag{2.58}
$$

where $f_i^0$ is the anomalous QGC coefficient at the low energy, $s$ is the square of the center-of-mass energy, and $\Lambda_{\mathrm{FF}}$ is the cutoff scale for the form factor. In this way, new physics can be studied in the $W^\pm W^\pm W^\mp$ analysis and limit can be set for anomalous QGC vertices.

<div align="center">

# CHAPTER 3

# Large Hadron Collider and ATLAS Detector

</div>

## 3.1 Large Hadron Collider

Operated by CERN, the LHC [25] is currently the particle collider with the highest energy in the
world. It was built underground at the border of Switzerland and France starting from 1990s.
The LHC has an approximately 27-kilometer ring of superconducting magnets that accelerate two
beams of hadrons, either protons or lead ions, to a speed close to light and collide them in the
opposite directions [26]. There are four large particle detectors on the LHC ring: the ATLAS [27]
and CMS [28] detectors are two general-purpose detectors designed for discoveries in the energy
frontier including both the SM and the BSM physics; A Large Ion Collider Experiment (ALICE)
detector is optimized to study physics in heavy-ion collisions [29]; and Large Hadron Collider
Beauty (LHCb) detector is specialized to study $b$-physics [30].

Protons are accelerated by a series of accelerators before entering the LHC tunnel [32]. Fig-
ure 3.1 shows the layout of the LHC accelerator system. The protons are produced from hydrogen
gas by stripping electrons from hydrogen atoms with an electric field and are then injected into
Linear Accelerator (Linac). The Linac accelerates protons to an energy of 50 MeV. The proton
beam is then pushed to 1.4 GeV in Proton Synchrotron Booster (PSB) and then 25 GeV in Proton
Synchrotron (PS). The protons are accelerated to 450 GeV in Super Proton Synchrotron (SPS)
before they are finally injected into the LHC in two opposite directions.

The number of $pp$ collisions $N$ in the LHC can be expressed by:

$$N = \sigma \times \int \mathcal{L} dt = \sigma \times L, \tag{3.1}$$

where $\sigma$ is the cross section of $pp$ collisions at the collision energy, $\mathcal{L}$ is the instantaneous luminos-
ity, and $L$ is the integrated luminosity over time. Therefore luminosity is an important parameter
for accelerators in addition to the collision energy.

The LHC is designed to accelerate protons to an energy of 7 TeV and thus the center-of-

<div align="center">

20

</div>
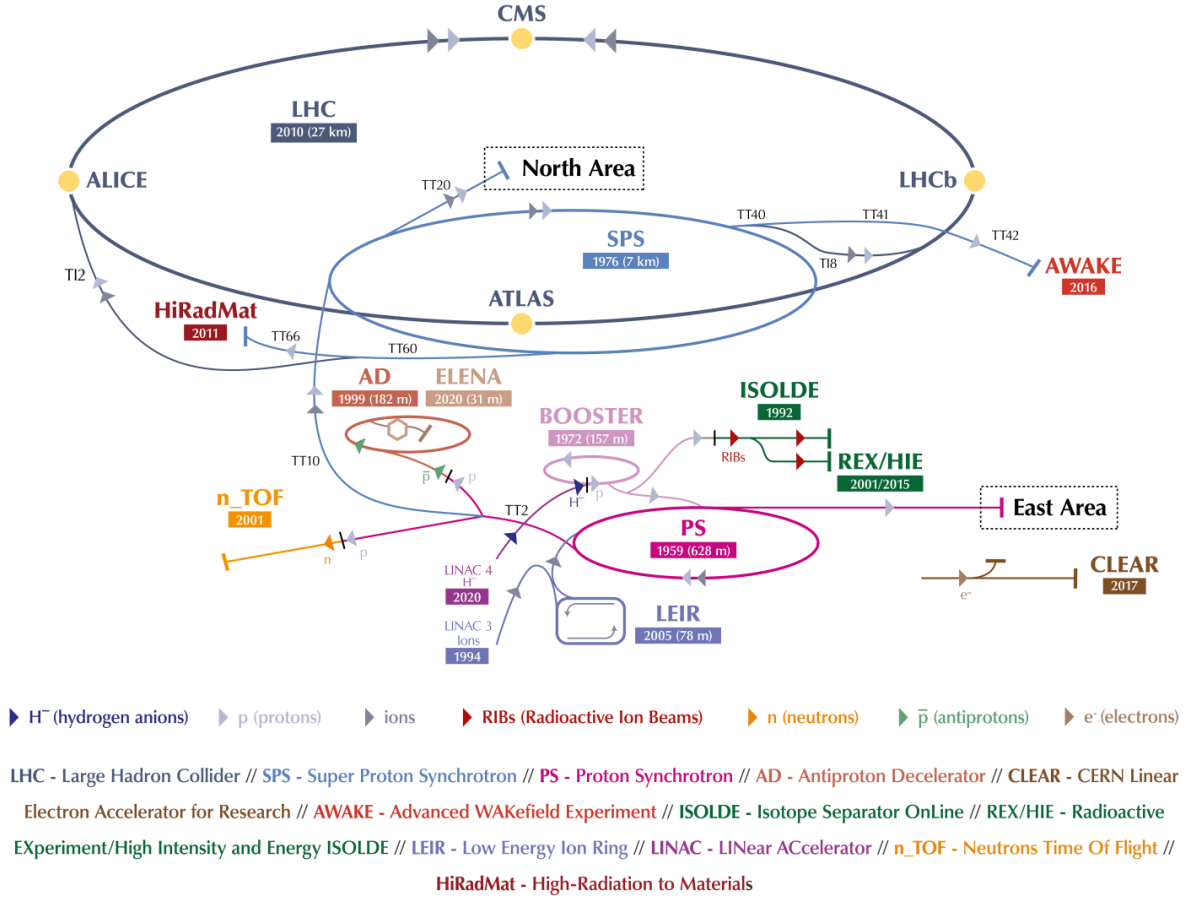
Figure 3.1: An illustration of the CERN accelerator complex where the LHC is the last ring in dark blue [31].

mass energy ($\sqrt{s}$) is 14 TeV. The maximum peak instantaneous luminosity is designed to be $1.0 \times 10^{34}$ cm$^{-2}$s$^{-1}$. Although the designed energy has not been reached, the maximum peak luminosity has reached $2.1 \times 10^{34}$ cm$^{-2}$s$^{-1}$. The $pp$ beam bunch crossing collisions happen at a rate of 40 MHz by design. During the first data taking period from 2010 to 2012 (Run 1), the LHC delivered an integrated luminosity of 0.0481 fb$^{-1}$ and 5.46 fb$^{-1}$ $pp$ collision data at $\sqrt{s} = 7$ TeV in 2010 and 2011 respectively, as well as 22.8 fb$^{-1}$ of data at $\sqrt{s} = 8$ TeV in 2012. During the second data taking period from 2015 to 2018 (Run 2), the LHC delivered in total 156 fb$^{-1}$ of $pp$ collision data at $\sqrt{s} = 13$ TeV. The distributions of the integrated luminosity versus time are shown in Figure 3.2. This thesis uses the full Run 2 data taken by the ATLAS detector, corresponding to an integrated luminosity of 139 fb$^{-1}$ that are good for physics analysis.
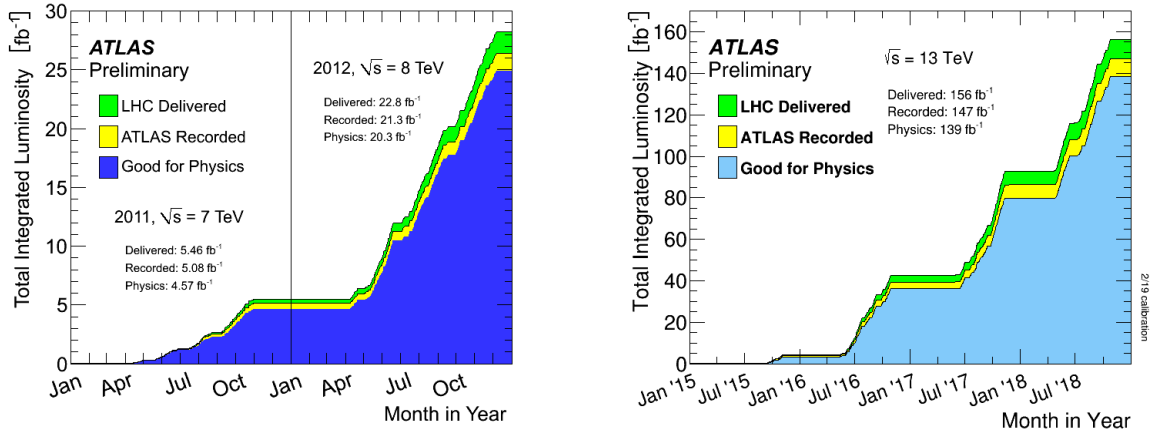
Figure 3.2: The distribution of the total integrated luminosity delivered (green) by the LHC, recorded by the ATLAS detector (yellow) and good for physics (blue) versus time in 2011 − 2012 (left) [33] and 2015 − 2018 (right) [34].

## 3.2 ATLAS Detector

The ATLAS detector is the largest detector on the LHC in terms of the detector volume. It is a 7000-ton cylinder with approximately 44 m in length and 25 m in diameter [27]. It consists of four major components: inner detector, calorimeters, muon spectrometer, and magnet system. Due to the high collision rate, trigger and data acquisition system at ATLAS only selects some collision events that are interesting in physics. These systems are briefly summarized in the following sections.

The cutaway diagram of the ATLAS detector is shown in Figure 3.3. Particles produced from the collision point are first identified by the inner detector, including pixel detector, Semiconductor Tracker (SCT) and Transition Radiation Tracker (TRT), which records particle tracks. The solenoid outside the inner detector provides a magnetic field that changes the moving directions of charged particles so that particles' charges and momenta can be measured. Electrons, photons, and hadrons are stopped and measured by calorimeters including Liquid-Argon (LAr) Electromagnetic (EM) calorimeter, tile calorimeter, LAr Hadronic End-cap Calorimeter (HEC), and LAr Forward Calorimeter (FCal). The toroid magnets are outside the calorimeters which further bend trajectories of muons. The outermost part of the detector is the muon spectrometer, also called muon chambers in the figure, that measures the momenta of muons.

### 3.2.1 Coordinate System

There are two coordinate systems, Cartesian coordinate system and spherical coordinate system, commonly used in the ATLAS detector as shown in the left plot of Figure 3.4 and both systems
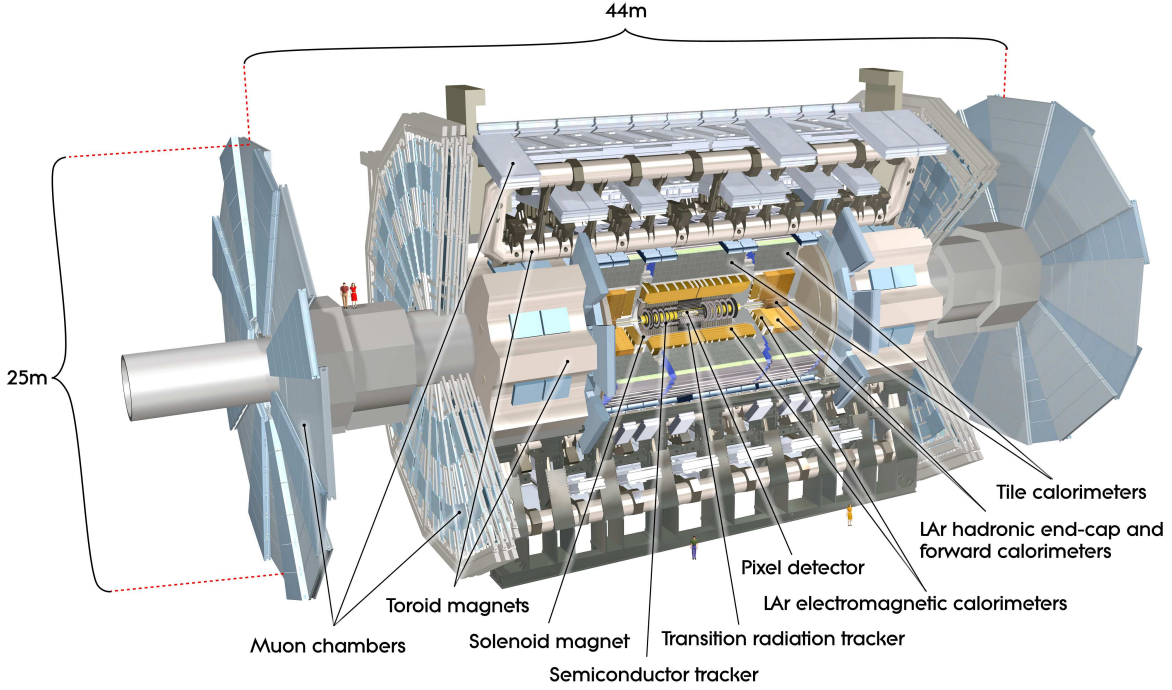
Figure 3.3: The cutaway diagram of the ATLAS detector [27].

use the primary collision vertex as the origin. In the Cartesian coordinate system, the $x$ axis points towards the center of the LHC ring, and the $y$ axis points to the sky. The $z$ axis is parallel to the beamline and its direction is determined by the right-hand rule. For the detector, the positive $z$ axis is defined as the A-side while the negative $z$ axis is defined as the C-side. The spherical coordinate system can be defined using the Cartesian coordinate system. In the spherical coordinate system, for a point $P$ shown in the left plot of Figure 3.4, the radius $r$ is defined as the distance between the point $P$ and the primary collision vertex $o$, and the polar angle $\theta$ is the angle between the directions of the $z$ axis and the point $P$, while the azimuthal angle $\phi$ is defined the angle between the directions of the $x$ axis and point $P'$ which is the projection of the point $P$ on the transverse $xy$-plane.

Since the longitudinal momenta of the two colliding partons (quarks or gluons inside the proton) along the $z$ axis are unknown before the collision, variables that are Lorentz-invariant under the longitudinal boost are preferred. Therefore, projected values of physics variables on the transverse plane are used more commonly, such as transverse momentum $p_T$ ($p_T = p \times \sin\theta$), transverse energy $E_T$ ($E_T = E \times \sin\theta$), and missing transverse momentum $E_T^{\mathrm{miss}}$, where $p$ and $E$ are particle's momentum and energy respectively. Instead of the polar angle $\theta$, rapidity $y$ and pseudorapidity $\eta$
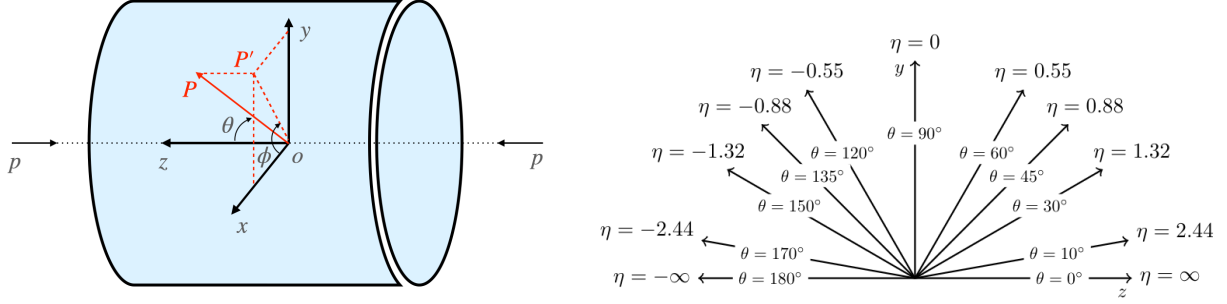
Figure 3.4: The illustration of the coordinate system used in the ATLAS detector. The left plot shows the right-handed Cartesian coordinate system and the spherical coordinate system used at ATLAS where the detector is represented by the blue cylinder and the black dash line stands for the $pp$ collision beam line with $o$ as the primary collision vertex. The right plot shows the relation between the polar angle $\theta$ and the pseudorapidity $\eta$.

are used. They are defined as:

$$
\begin{aligned}
y &\equiv \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right), \\
\eta &\equiv -\ln\left[\tan\left(\frac{\theta}{2}\right)\right],
\end{aligned}
\tag{3.2}
$$

where $p_z$ is the momentum along the $z$ axis. $y$ is Lorentz invariant under the longitudinal boost while $\eta$ is not. However $y$ can be approximated by $\eta$ when the particle's mass $m$ is much smaller than its momentum $|p|$ and thus $p_z \approx E \times \cos\theta$. The relation of the polar angle $\theta$ and the pseudorapidity $\eta$ can be seen on the right plot of Figure 3.4, and we have $\eta = 0$ when $\theta = 90$ ° and $\eta = \pm\infty$ when $\theta = 0$ ° and $180$ °. The separation between two particles in the $\eta - \phi$ plane, $\Delta R$, is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ with $\Delta\eta = \eta_1 - \eta_2$ and $\Delta\phi = \phi_1 - \phi_2$. $\eta_1$ ($\eta_2$) and $\phi_1$ ($\phi_2$) are the pseudorapidity and the azimuthal angle of the first (second) particle, respectively.

## 3.2.2  Magnet System

In order to bend tracks of charged particles to allow charge and momentum measurements, a magnet system [27, 36] is implemented in the ATLAS detector and it contains four major superconducting magnets as shown in Figure 3.5: the central solenoid, the barrel toroid and two end-cap toroids. Liquid helium is used to cool down the temperature so that the magnets enter the superconducting state which is critical to generate strong magnetic fields without loss of energy from electrical resistance.

The central solenoid is placed between the inner detector and the EM calorimeter and is designed to provide a peak axial magnetic field of 2.6 T for the inner detector. The nominal current

24
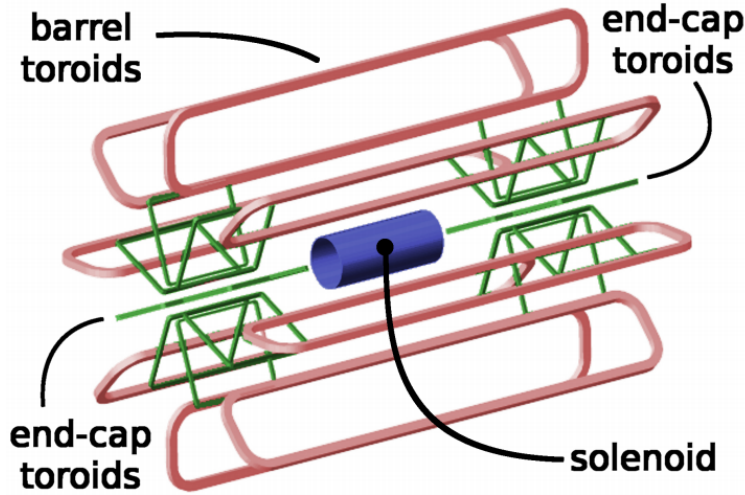
Figure 3.5: The spatial arrangement of the coil windings of the ATLAS magnet system [35].

that flows into the coil is 7.73 kA and the stored energy is 40 MJ. Its thickness in front of the barrel EM calorimeter has been minimized to reduce the influence on particle energy measurement. The central solenoid has an axial length of 5.8 m as well as inner and outer diameters of 2.46 m and 2.56 m, respectively.

The barrel toroid and the two end-cap toroids are used to bend muon tracks and they can provide up to 3.9 T and 4.1 T magnet fields in the barrel and end-cap regions of the muon spectrometer respectively. For the barrel toroid, the nominal current is 20.5 kA and the stored energy is 1080 MJ. The inner and outer diameters are 9.4 m and 20.1 m respectively and the axial length is 25.3 m. For each of the two end-cap toroids, the nominal current is also 20.5 kA and the stored energy is 250 MJ. The inner and outer diameters are 1.65 m and 10.7 m respectively and the axial length is 5.0 m.

### 3.2.3   Inner Detector

The inner detector [27, 38, 39] is the innermost part of the ATLAS detector and it has a cylinder shape with a diameter of 2.1 m and a length of 6.2 m, as shown in Figure 3.6. In order to reconstruct collision vertices and measure charges and momenta of particles with high precision and resolution in a large track density, the inner detector is made with fine granularity and can measure charged tracks with $p_T > 0.5$ GeV and $|\eta| < 2.5$. The detector has been designed to provide a transverse momentum resolution of $\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$ ($p_T$ in GeV) and a transverse impact parameter resolution of 10 μm for high momentum particles in the central $\eta$ region [27, 40]. As illustrated in Figure 3.7, a track satisfying the requirements will go through the pixel detector, the SCT and the
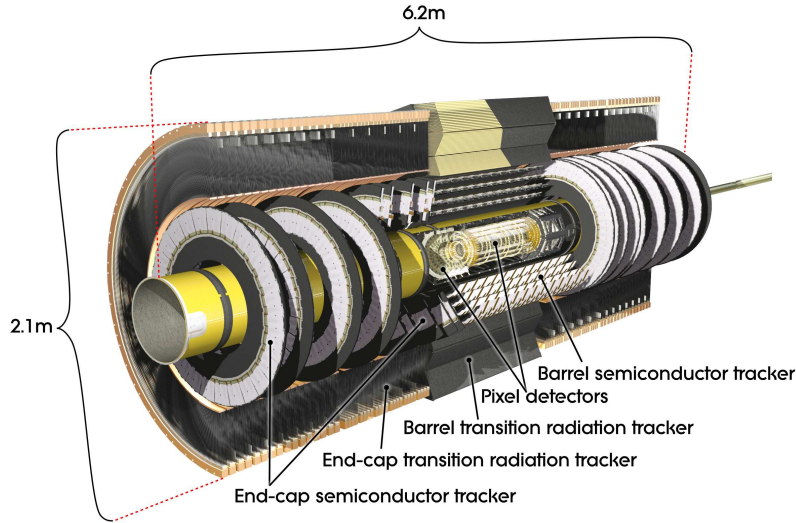
Figure 3.6: The cutaway diagram of the ATLAS inner detector [27].

TRT in order, which are three major components of the inner detector.

The pixel detector [41] is made of 1,744 pixel sensors that have a thickness of 250 μm and a dimension of $19 \times 63$ mm$^2$. The pixel sensor uses oxygenated $n$-type wafers with readout pixels on the $n^+$-implanted side of the detector. There are 46,080 readout channels on each sensor so in total the pixel detector has approximately 80 million channels, which provide an intrinsic accuracy of 10 μm × 115 μm. Pixel sensors are arranged in three coaxial cylindrical barrel layers: barrel layer 0 (b-layer), barrel layer 1, and barrel layer 2 as well as three disk layers in each of the two end-cap regions. In Run 2, an additional layer, Insertable B-Layer (IBL), is inserted between the new beam pipe with a reduced diameter and the b-layer [42, 43]. The IBL detector consists of 168 planar $n^+$-in-$n$ sensors and 112 3-D sensors and provides additional hit information so that track identification and vertex reconstruction can be further improved.

The SCT [44] is a silicon microstrip tracker with 4,088 modules and in total over 6 million channels on a 63 m$^2$ silicon surface. The SCT modules are distributed on four barrel layers (barrel layer $3 - 6$) and nine disk layers in each of the two end-cap regions. These sensors use a classic single-side $p$-in-$n$ technology with AC-coupled readout strips. The nominal position resolution is 17 μm × 580 μm.

The TRT [45, 46] is built with approximately 351,000 straw tubes with 4 mm in diameter. Each straw tube is filled with a xenon- or argon-based gas mixture. There is a gold-plated tungsten wire with a diameter of 0.03 mm in the center of the straw tube. When charged particles pass through these straw tubes, the gas atoms are ionized and electrons from the ionization are accelerated towards the center and create a signal current. In addition, straw tubes with a xenon-based gas mixture can detect photons from transition radiation, which provides powerful discrimination
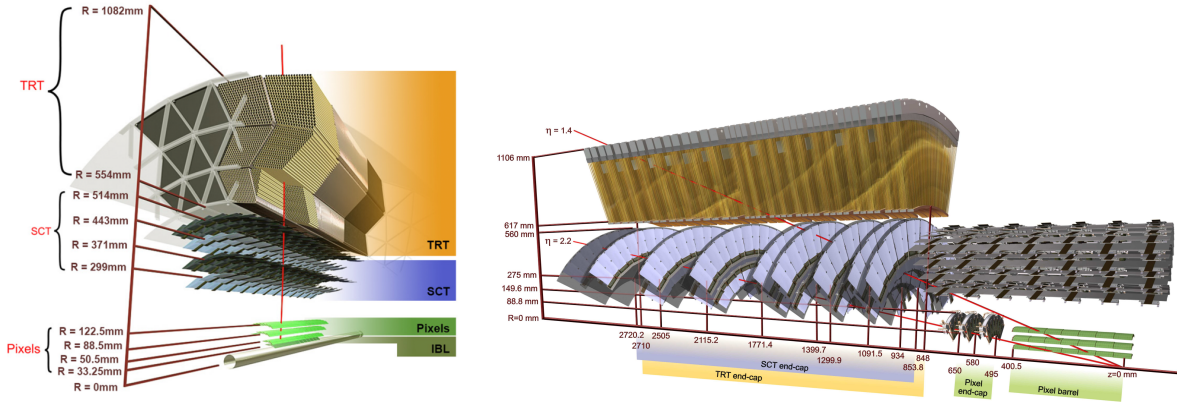
Figure 3.7: The illustration of the sensors and structural elements traversed by charged tracks with $p_T = 10$ GeV in the inner detector in the barrel region with $\eta = 0.3$ (left) [37] and the end-cap region with $\eta = 1.4$ and 2.2 (right) [27].

between electrons and charged pions. These straw tubes are distributed on 73 barrel layers separated with fibers in the barrel region and 160 end-cap planes separated with foils in each of the two end-cap regions. The TRT can reach a position resolution of 130 μm.

### 3.2.4 Calorimeters
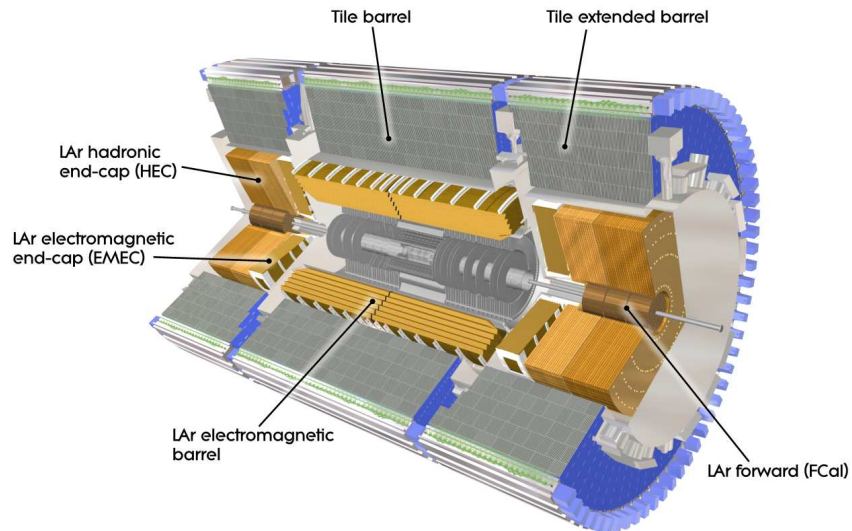


Figure 3.8: The cutaway diagram of the ATLAS calorimeters [27].

Placed outside the central solenoid and the inner detector, the calorimeters [27, 47–49] are used to measure the energy of particles, including electrons, photons, and hadrons within $|\eta| < 4.9$ by

stopping them from passing through and measuring their showers. In this way, they also leave a clean environment for the muon spectrometer. The calorimeters are all sampling calorimeters as they generate showers using an absorber with layers of high-density passive material and detect deposited particle energy using layers of active material. There are two types of calorimeters based on measured particles: the EM calorimeter, and the hadronic calorimeters including the tile calorimeter, the HEC, and the FCal. The spatial arrangement of the ATLAS calorimeters is shown in Figure 3.8.

Located right outside the central solenoid, the EM calorimeter [48] can be further divided based on the location: one EM barrel calorimeter which covers $|\eta| < 1.475$, and two Electromagnetic End-cap Calorimeters (EMECs) which cover $1.375 < |\eta| < 3.2$. The EM calorimeter is used to determine the energy of electrons or photons by measuring EM showers. It uses LAr as the active material due to its radiation hardness and lead as the passive material for the absorber. There is also a presampler detector with a layer of LAr that is used to correct the energy loss of electrons and photons before they enter the EM calorimeter in the region of $|\eta| < 1.8$. The EM calorimeter has a total thickness of $> 22X_0$ in the barrel region and $> 24X_0$ in the end-cap region, where $X_0$ is the radiation length. The energy resolution of the EM calorimeter is $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$ ($E$ in GeV) by design [27].

Hadronic calorimeters are placed outside the EM calorimeter and are used to measure the energy of hadrons using hadronic showers. The tile calorimeter [49] is outside the envelope of the EM calorimeter with three parts: one barrel part with $|\eta| < 1.0$ and two extended barrel parts with $0.8 < |\eta| < 1.7$. Steel is used for the absorber and scintillating tiles are used for the active material. The HEC [48] is located behind the EMEC and has two wheels in each side of the two end-cap regions, covering the region $1.5 < |\eta| < 3.2$. It uses the same active material (LAr) as the EM calorimeter while copper as the passive material for the absorber. The tile calorimeter has a total thickness of $7\lambda$ in the barrel region and $12\lambda$ in the end-cap region, where $\lambda$ is the nuclear interaction length. The energy resolution of the tile calorimeter is $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$ ($E$ in GeV) by design [27].

The FCal [48] is close to the beamline and covers the region $3.1 < |\eta| < 4.9$. It also uses LAr as the active material and can measure both hadronic and EM showers as the EM calorimeter does not cover this region. The FCal consists of three modules placed longitudinally in each of the two end-cap regions. In order to optimize the resolution and heat removal, the innermost module that is closest to the collision point (FCal1) uses copper as the absorber and is optimized for EM showers, while the other two modules (FCal2 and FCal3) use tungsten as the absorber and is optimized for hadronic showers. The energy resolution is $\sigma_E/E = 25\%/\sqrt{E} \oplus 3.8\%$ ($E$ in GeV) for electrons and $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$ ($E$ in GeV) for hadrons by design [27].

### 3.2.5 Muon Spectrometer
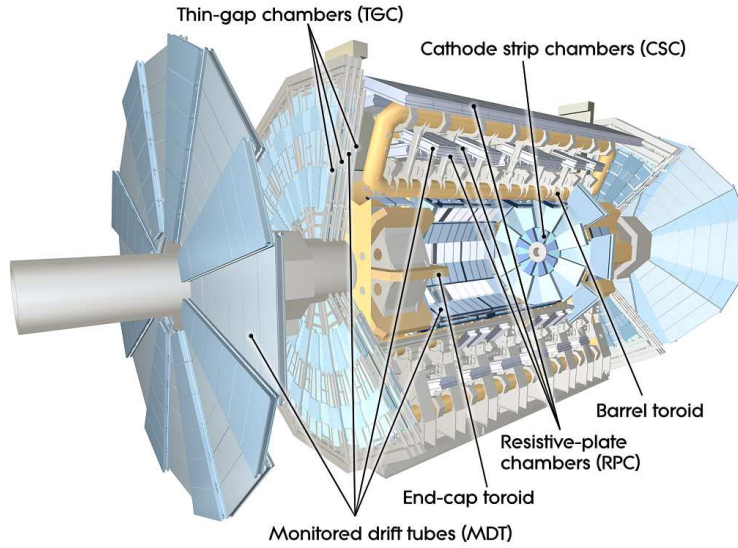


Figure 3.9: The cutaway diagram of the ATLAS muon spectrometer [27].

The muon spectrometer [27, 50] is the outermost detector and is used to detect and measure momentum and charge of muon as indicated by its name. Since all other particles have been stopped by the calorimeters except muon and neutrino, the muon spectrometer can perform the measurement in a relatively clean environment. The toroids provide the magnetic field for the muon spectrometer: the barrel toroid covers the region $|\eta| < 1.4$ while two end-cap toroids cover the region $1.6 < |\eta| < 2.7$. In the transition region of $1.4 < |\eta| < 1.6$, the magnetic field is provided by the combination of the barrel and end-cap toroids. There are three stations in the barrel region (called barrel inner, barrel middle, and barrel outer) and three stations in the end-cap region (called end-cap inner, end-cap middle, and end-cap outer). As shown in Figure 3.9, the muon spectrometer consists of two types of precision tracking chambers, Monitored Drift Tubes (MDTs) and Cathode Strip Chambers (CSCs), as well as two types of trigger chambers, Resistive Plate Chambers (RPCs) and thin gap chambers. The momentum resolution determined by the MDTs and CSCs is expected to be $\sigma_{p_T}/p_T = 10\%$ for 1 TeV muons ($p_T$ in GeV) [27].

The MDTs are used to measure the momentum of muon within the region $|\eta| < 2.7$, except for the innermost station, the small wheels, where the coverage is limited to $|\eta| < 2.0$, as shown in Figure 3.10. The MDTs consist of approximately 354,000 pressurized drift tubes which have a diameter of 29.970 mm and are filled with an argon-based gas mixture at 3 bar. There is a tungsten-rhenium wire in the center of each tube and the wire is held at a potential of 3,080 V. The maximum drift time is 780 ns. Similar to straw tubes in the TRT, muons passing the MDTs can ionize the gas atoms and electrons from ionization can be accelerated by the electric field
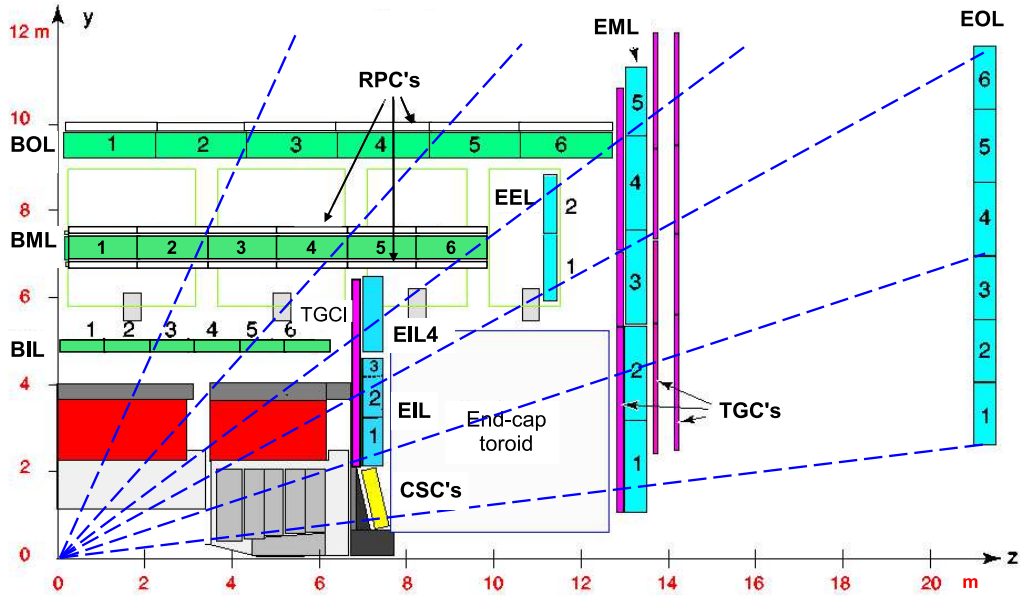
Figure 3.10: The detailed layout of chambers in one quarter of the ATLAS muon spectrometer: the barrel MDTs are shown in green; the end-cap MDTs are shown in light blue; the CSCs are shown in yellow and labeled; the RPCs are shown in white and labeled; the thin gap chambers are shown in pink and labeled [27].

and create signal currents on the central wire. The average resolution per tube for the MDTs is approximately 80 μm.

Due to higher rate capability and time resolution, the CSCs are used in the small wheels in the forward end-cap regions with $2 < |\eta| < 2.7$, as particle fluxes are high in the region close to the beam pipe. The locations of the CSCs are shown in Figure 3.10 and they are multiwire proportional chambers. The anode wires are oriented in the radial direction and held at a voltage of $1,900$ V. The gas filled in the chambers is also an argon-based gas mixture. Both cathodes are segmented strips with one in perpendicular to the wires while the other parallel to the wires so that the location on the transverse plane can be determined. Unlike the MDTs, the readout from the CSCs is from the cathode strips instead of the anode wires. The position resolution of the CSCs is 60 μm per plane which is comparable to the resolution of the MDTs.

In order to trigger on muon tracks, fast trigger chambers are implemented in the muon spectrometer with the RPCs covering the barrel region of $|\eta| < 1.05$ and the thin gap chambers covering two end-cap regions of $1.05 < |\eta| < 2.4$. Additionally, trigger chambers can measure muon tracks in both bending planes ($\eta$) and non-bending plane ($\phi$) while precision tracking chambers, the MDTs and CSCs, measure the tracks only in the bending plane.

The RPCs are gaseous electrode-plate detectors without wires inside. Each chamber has two re-

sistive plates placed in parallel and separated by a 2 mm gas gap that is filled with a $C_2H_2F_4$-based gas mixture. The RPCs are operated in the avalanche mode with a nominal voltage of $9,800$ V. The signal is read out from the longitudinal and transverse strips outside the two resistive plates. The RPCs can achieve a temporal resolution of 1.5 ns and a spatial resolution of 10 mm $\times$ 10 mm.

The thin gap chambers are also multiwire proportional chambers which are similar to the CSCs. However, unlike the CSCs, the thin gap chambers are made with finer granularity and as a result they have less response time to muon tracks. The wires are held at a much higher voltage of $2,900$ V and the gas filled is a mixture of $CO_2$ and n-pentane. The resolution of the thin gap chambers is 4 ns in time and 2-6 mm $\times$ 3-7 mm in space.

### 3.2.6 Trigger and Data Acquisition
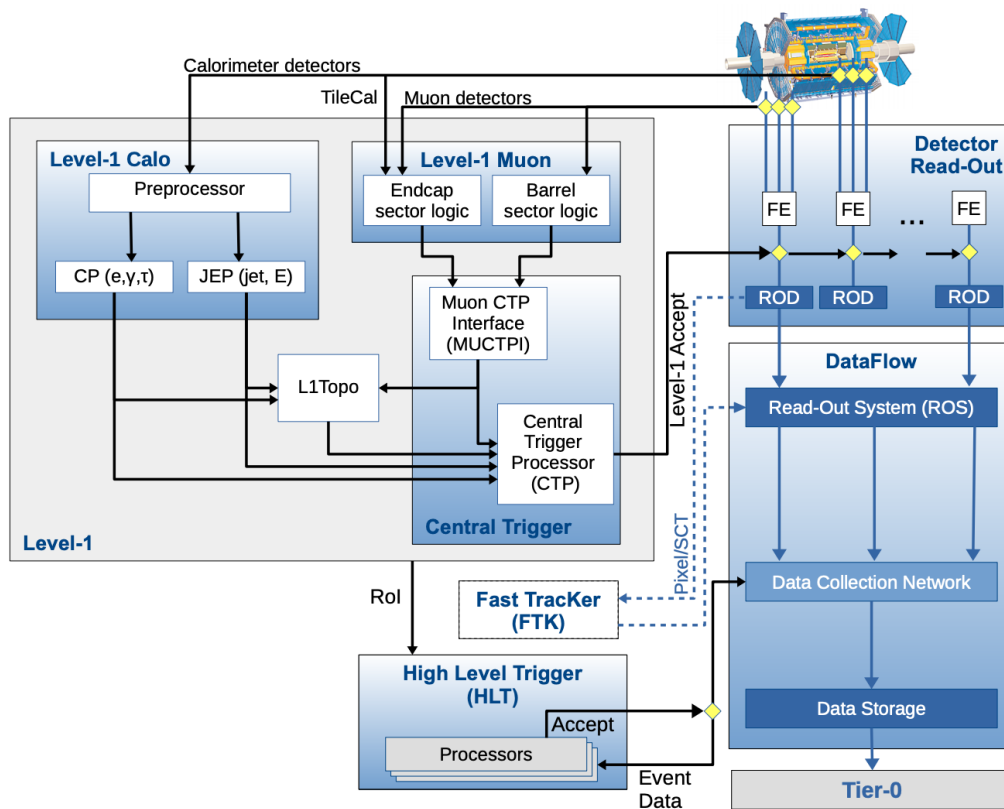


Figure 3.11: The flow chart of the ATLAS TDAQ system in Run 2 [51].

The ATLAS Trigger and Data Acquisition (TDAQ) system [51, 52] is responsible for selecting collision events of interest to record and is an essential part of the ATLAS detector. The LHC is designed to collide the $pp$ beam bunch crossings every 25 ns and the raw data takes up approximately 1.6 MB per event. It is impossible to save all data since the raw data rate is more than

50 TB per second, which is the storage of 100,000 CDs every second. The TDAQ system can select interesting events in physics which is only a tiny fraction of all events and thus reduce the data rate to approximately 1 kHz. As shown in the flow chart of the TDAQ system in Figure 3.11, the data rate in Run 2 is reduced in two major steps: the Level-1 (L1) trigger and the High Level Trigger (HLT).

The L1 trigger [53] is a hardware-based trigger system using information from the calorimeters and the muon spectrometer. The information from the calorimeters is processed in the L1 Calorimeter (L1Calo) trigger system. After prepossessing the calorimeter data, the cluster processor identifies electrons, photons and leptonically decayed taus ($\tau_{\text{lep}}$), while the Jet/Energy-sum Processor (JEP) identifies jets and computes the total transverse energy and missing transverse energy in parallel. The results from the cluster processor and the JEP are the output of the L1Calo. In the L1 Muon (L1Muon) trigger system, hit information from the RPCs is processed in the barrel sector logic while hit information from the thin gap chambers is processed together with the data from the tile calorimeter to reduce the rate of particles in the end-cap regions. The outputs from the L1Muon are then sent to the Muon Central Trigger Processor Interface (MUCTPI) for further identification of muons. Additionally, the L1 Topological (L1Topo) trigger system also applies some topological requirements by combining the outputs from the L1Calo and the MUCTPI to identify interesting events at the event level. Finally, the trigger outputs from the L1Calo, the MUCTPI, the L1Topo as well as some detector subsystems are passed to the Central Trigger Processor (CTP) where the final L1 trigger decision is made. Overall, the L1 trigger system reduces the data rate from 40 MHz to about 100 kHz within a latency of 2.5 µs.

The L1 trigger initiates the detector read-out process and provides the outputs to the HLT. If an event passes the L1 trigger requirement, the Front-End (FE) electronics of all sub-detectors then read out the data from their buffers and the data are passed to the ReadOut Drivers (RODs) for prepossessing. The output data is buffered in the ReadOut System (ROS) which provides information to the HLT using the data collection network. The L1 trigger also determines the Regions of Interest (ROIs) and passes the spacial information in $\phi$ and $\eta$ to the HLT.

The HLT [54] is a software-based trigger system. Unlike Run 1 when the Level-2 (L2) trigger and the event filter were separated, the HLT combines them in Run 2 and analyzes the detector data in the ROIs using approximately 40 thousand processing units. The typical procedure is to start with feature-extraction algorithms to request data in the ROIs and end with a hypothesis algorithm on whether the reconstructed information satisfies the trigger requirements or not. The HLT can further lower down the data rate to on average 1.2 kHz which corresponds to a writing speed of 1.2 GB/s. If the event is accepted by the HLT, the event data is then stored in the CERN Tier-0 facility [55] for further offline processing.

# CHAPTER 4

# Object Reconstruction and Selection

## 4.1 Overview

This chapter summarizes the reconstruction of physical objects as well as the object selection criteria used in this $W^\pm W^\pm W^\mp$ analysis. The $W$ boson has a very short lifetime of approximately $3 \times 10^{-25}$ s, and thus cannot be directly observed by the ATLAS detector. Instead, the detector reconstructs and identifies decayed particles from the $W^\pm W^\pm W^\mp$ production. These final products also need to satisfy the standard requirements recommended by ATLAS Combined Performance (CP) groups.

## 4.2 Vertices and Pileup Corrections
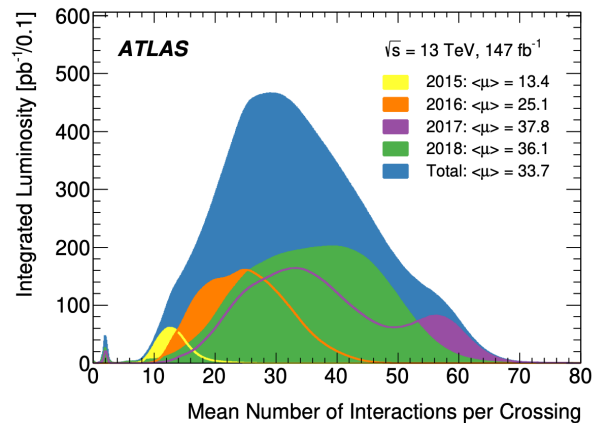


Figure 4.1: Luminosity-weighted distribution of the mean number of interactions per bunch crossing in each year during the Run 2 data taking period (2015 − 2018) [56].

At the LHC, multiple $pp$ collisions occur within each bunch crossing of 25 ns, and thus there are multiple vertices for each event. Since the total inelastic $pp$ cross section is measured to be

78.1 mb at $\sqrt{s} = 13$ TeV (1 b = $10^{-24}$ cm$^{-2}$) [57], the average number of inelastic $pp$ interactions per bunch crossing is about 19.5 for an instantaneous luminosity of $10^{34}$ cm$^{-2}$s$^{-1}$ (78.1 mb $\times$ $10^{34}$ cm$^{-2}$s$^{-1}$ $\times$ 25 ns $\approx$ 19.5). These additional collisions are uncorrelated with the hard scattering process, and the particles from the hard scattering process are recorded together with particles from these additional $pp$ interactions, so called pileup interactions. The average number of pileup collisions is about 33.7 per bunch crossing for the dataset used in this analysis and the number varied for different data taking years as shown in Figure 4.1.

Vertex candidates [58] are required to have at least two reconstructed tracks and the requirements for those tracks are summarized in Table 4.1 [59]. The tracks identified by the inner detector or the muon spectrometer need to pass the basic kinematic requirement of $p_T > 500$ MeV and be within detector geometry of $|\eta| < 2.5$. The tracks need to have at least nine hits if $|\eta| \leq 1.65$ and 11 hits if $|\eta| > 1.65$ in the pixel detector and the SCT. Additionally, there is at least one hit at the IBL or the b-layer in the pixel detector. The number of shared modules for selected tracks is at most 1, i.e. at most one shared hit in the pixel detector or two shared hits in the SCT. Holes on the detector are defined as expected hits on the detector that are not observed. There should not be any holes in the pixel detector and at most two holes in the SCT for selected tracks.

**Selected Tracks**

|  | Reconstructed track candidates |
| --- | --- |
| $p_T$ | > 500 MeV |
| $|\eta|$ | < 2.5 |
| Pixel and SCT hits | $\geq 9$ (for $|\eta| \leq 1.65$) or 11 (for $|\eta| > 1.65$) |
| IBL and b-layer hits | $\geq 1$ |
| Shared models | $\leq 1$ |
| Pixel holes | $= 0$ |
| SCT holes | $\leq 2$ |

Table 4.1: The requirements of tracks used in vertex reconstruction [58, 59].

After finding vertex candidates with selected tracks, vertex fitting is performed to reconstruct the primary vertex [58, 60]. Starting with a seed position for the first vertex, an iterative $\chi^2$ minimization is used to find the best vertex position with the seed position and parameters of tracks as inputs. After that, the procedure is repeated using the rejected tracks that are incompatible with the reconstructed vertex reconstructed earlier. However, only one primary vertex is considered in each event and it is the one with the largest sum in squared transverse momentum ($\sum p_T^2$) of all associated tracks. With the determination of the primary vertex, impact parameters can be defined by the distance between the primary vertex and the track: $d_0$ is the distance in the transverse plane, and $z_0$ is the distance along the $z$ axis. Those impact parameters are used to select electrons in

Section 4.3 and muons in Section 4.4.

The MC is produced usually before or during the data taking periods and is thus generated with only a guess on the pileup collisions expected in data. Therefore, a correction in pileup is needed for the MC simulation, which is called pileup reweighting. In the pileup reweighting, $\mu$ is defined as the number of pileup collisions in a bunch crossing, and $\langle\mu\rangle$ is defined as the average of $\mu$ over the colliding bunches. $\langle\mu\rangle$ and $\mu$ are the same in MC but are different in data. Three MC campaigns (MC16a, MC16d and MC16e) are made for the ATLAS Run 2 data collected between 2015 and 2018. MC16a is corrected to the $\langle\mu\rangle$ distribution in data taken in 2015-2016, while MC16d and MC16e are corrected to the $\mu$ distributions in data taken in 2017 and 2018 respectively [61].

## 4.3   Electrons

Electrons in the ATLAS detector are defined as objects with tracks in the inner detector matched to the energy deposits in the calorimeter (superclusters). Electron ($|\eta| < 2.5$) reconstruction [62–64] starts with the selection of reconstructed energy deposits in topologically-connected EM and hadronic calorimeter cells (topo-clusters) [65]. Proto-clusters are identified with high significance in the EM calorimeter and merged with neighboring cells as topo-clusters. These clusters need to have energy in the EM calorimeter greater than 400 MeV and the fraction of energy deposited in the EM calorimeter above 0.5 to reject pileup clusters. Topo-clusters are then matched with tracks reconstructed in the inner detector. The matching is performed using the requirements on $|\Delta\eta|$ and $q \times \Delta\phi$ ($q$ is the reconstructed charge of the track) between the cluster energy and the measured or rescaled track momentum with the consideration of energy loss due to bremsstrahlung. If multiple tracks are matched, tracks with hits in the pixel detector and smaller $\Delta R$ matched to the cluster are preferred over those with hits in the SCT and larger $\Delta R$.

After the preparation of tracks and clusters, the reconstruction of electron superclusters is performed in three steps. In the first step, a track-matched topo-cluster is selected as a seed cluster if it has $E_\mathrm{T} > 1$ GeV ($E_\mathrm{T}$ is defined as $E \times \sin\theta$ with $E$ as the energy of the cluster) and a matched track with at least four hits in the pixel detector and the SCT. Satellite clusters are then identified for the selected seed cluster. A satellite cluster can be either a cluster within a $0.075 \times 0.125$ window in $\Delta\eta \times \Delta\phi$ around the seed cluster or a cluster with a $0.125 \times 0.300$ window and has the same best-matched track as the seed cluster. The supercluster is then the combination of the seed cluster and its satellite clusters. The last step is to assign calorimeter cells for the reconstructed supercluster.

Initial energy calibration and position corrections are applied to the superclusters and the clusters are matched to tracks in the same way as the matching between topo-clusters and tracks mentioned above. Since photons are also reconstructed independently in a similar way, some electron superclusters may match with photons. In these cases, a procedure to discriminate electrons from

photons is performed using detailed information of hits in the inner detector and superclusters. The result of the procedure is then saved to a variable called author (also called ambiguity type) and the definition is as follows: author = 16 if it is not possible to determine if the object is an electron or a photon even after the extra procedure and both electron and photon objects are created from the supercluster; author = 0 if the object can be identified only as an electron which has a cluster with a good matched track and no good photon conversion vertex. In the end, the energies of electrons are re-calibrated and discriminating variables are calculated for electron identification.
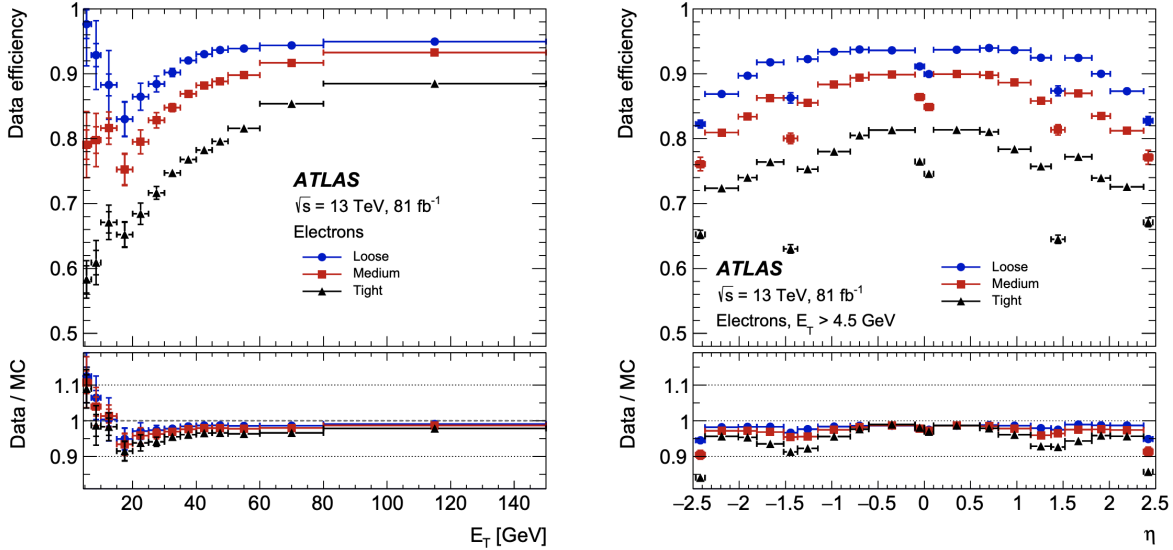


Figure 4.2: Electron identification efficiency in $Z \to ee$ events in data as a function of $E_T$ (left) and $\eta$ (right) for different working points: loose (blue), medium (red) and tight (black) [64].

Electron identification [62, 64] is an additional step after electron reconstruction to improve the purity of selected electron candidates. It is a likelihood discriminant based on measurements from the inner detector and the calorimeters. The inputs include variables from hadronic leakage, energies deposited in different layers of the EM calorimeter, track conditions, and track-cluster matching. There are three different working points based on the strictness of the cuts for electron identification: loose, medium and tight. There is a trade-off between electron identification efficiency and background rejection. The tight working point gives the purest prompt electrons but has the least identification efficiency. The identification efficiencies as a function of electron $E_T$ and $\eta$ measured in data and the ratios between data and MC efficiencies for different working points are shown in Figure 4.2. All three working points are used for different types of electrons selected in this analysis.

Electron isolation [62, 64] can be used to further separate prompt electrons from non-prompt electrons through tracks and energy deposits near electron candidates. In this analysis, the tight

Prompt Lepton Veto (PLV) working point [66] recommended by the ATLAS CP groups is used. It applies cuts on a track isolation variable $p_T^{\text{varcone30}}$ and a BDT discriminant score PLV. Electron $p_T^{\text{varcone30}}$ is defined as the sum of $p_T$ for all tracks with $p_T > 1$ GeV and loose vertex association within a varied-size cone of $\Delta R = \min(0.3, 10 \text{ GeV}/p_T)$ around the electron track direction. The electron track is excluded in the $p_T^{\text{varcone30}}$ calculation. The BDT of PLV [67] uses tracks and energy deposits in a cone around the electron direction and targets at separating prompt electrons and non-prompt electrons from heavy-flavor quark decays. The requirements of the tight PLV working point for electrons with $p_T > 20$ GeV are summarized as follows:

$$p_T^{\text{varcone30}} < 0.15 \times p_T \quad \text{and} \quad -1.1 < \text{PLV} < \max\left(-0.88, -0.881497 + 2.29469 \times e^{-\frac{p_T}{11.5776 \text{ GeV}}}\right).$$
(4.1)

A common problem for electrons is their electric charges could be misidentified either due to an incorrect measurement of the track curvature or a wrong choice of the associated track. Most charge misidentified electrons found in the ATLAS detector are due to the wrong choice of the associated track since electrons interact with detector material, generate photons and electron-position pairs, and the matching algorithm pick up the track with a wrong electric charge. Electron Charge Identification Selector (ECIDS) [64] is a BDT discriminant score used to suppress the electrons with a misidentified charge. The inputs of the BDT are electron $E_T$ and $|\eta|$, transverse impact parameter multiplied by its charge ($d_0 \times q$), average charge of matched tracks weighted by the hits in the SCT ($\bar{q}_{\text{SCT}}$), energy and momentum ratio ($E/p$), and $\Delta\phi_{res}$. A cut of ECIDS $> -0.337671$ is recommended by the ATLAS CP groups [68]. For electrons with medium or tight identification working point and tight isolation requirement found in $Z \rightarrow ee$ events, the efficiency of the recommended ECIDS cut is around 98% while approximately 90% of electrons with misidentified charges are removed.

There is an additional ambiguity-type variable, addAmbiguity [69], which is used to flag internal and material conversions that are missed due to conversions with radius $r < 20$ mm and conversion vertices with loose track quality. The values of addAmbiguity are defined as follows: addAmbiguity = $-1$ if no second track is found; addAmbiguity = 0 if a second track is found but no conversion is found; addAmbiguity = 1 if an internal conversion is found; addAmbiguity = 2 if a material conversion is found.

In the $W^\pm W^\pm W^\mp$ analysis, four types of electrons are defined and the selection criteria are listed in Table 4.2. Signal electrons, also called ID (identified) electrons, have the tightest definition and are likely prompt electrons. They need to pass basic kinematic requirement of $p_T > 20$ GeV and geometrical requirement of $|\eta| < 2.47$ and outside the crack region of $1.37 \leq |\eta| \leq 1.52$ due to the transition of EM barrel and end-cap calorimeters. Signal electrons also need to be produced from the primary vertices by requiring the impact parameters with $|d_0/\sigma_{d_0}| < 5$ and $|z_0 \sin(\theta)| < 0.5$ mm.

The tight working point is used for electron identification and the tight PLV working point is used for electron isolation. Signal electrons need to have author = 1 so the objects are exclusively reconstructed as electrons to suppress the photon conversion background. The photon conversion background is thus further suppressed by requiring addAmbiguity $\leq 0$. The recommended cut on ECIDS are applied to suppress charge-misidentified electrons.

**Selected Electrons**

| | ID | Anti-ID | Anti-BL | Veto |
|---|---|---|---|---|
| | Reconstructed electron candidates | | | |
| $p_T$ | > 20 GeV | | | > 7 GeV |
| $|\eta|$ | $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$ | | | < 2.47 |
| $|d_0/\sigma_{d_0}|$ | < 5 | | | - |
| $|z_0 \sin(\theta)|$ | < 0.5 mm | | | - |
| Identification | Tight | Medium | Tight + no b-layer | Loose |
| Isolation | TightPLV + ECIDS | - | TightPLV + ECIDS | - |
| Ambiguity | author = 1 and addAmbiguity $\leq 0$ | | - | - |
| Additional | - | !Signal | - | - |

Table 4.2: Summary of electron types used.

Anti-ID (anti-identified) electrons are electrons that are likely to be non-prompt electrons and are used in non-prompt background estimation described in Section 7.3. They do not have the isolation and ECIDS requirements while the medium electron identification working point is required. Anti-ID electrons are also required to be orthogonal to signal electrons and have the same requirements as signal electrons for $p_T$, $|\eta|$, impact parameters, and ambiguity.

Anti-BL (anti-b-layer) electrons are electrons that are likely to be electrons from photon conversion and are used in photon conversion background estimation described in Section 7.4. They do not have the ambiguity requirement and the tight electron identification working point is required but without hits in the b-layer of the pixel detector. Anti-BL electrons are orthogonal to signal electrons by construction and have the same requirements as signal electrons for $p_T$, $|\eta|$, impact parameters, electron isolation, and ECIDS.

Veto electrons are the loosest electrons in this analysis and are used to veto events with additional electrons. They are required to have $p_T > 7$ GeV, $|\eta| < 2.47$ and pass the loose electron identification working point.

MC events with electrons need to be corrected for efficiencies to match with what is observed in data. Based on the choice of electron identification and isolation working points as well as electron trigger, reconstruction, identification, and isolation scale factors are provided by the ATLAS CP groups and are included in the event weight calculation. The effects of ECIDS and tight PLV are

38

combined as a single isolation scale factor.

## 4.4 Muons

Unlike electrons, muons can penetrate through the ATLAS detector and leave tracks in both the inner detector and the muon spectrometer as well as characteristic energy deposits in the calorimeters. Muon reconstruction [70, 71] is performed with different strategies and leads to five types of muons: combined muons, muon-spectrometer extrapolated muons, inside-out combined muons, segment-tagged muons, and calorimeter-tagged muons.

Combined muons are the main type of muons used in the $W^\pm W^\pm W^\mp$ analysis. They are reconstructed by matching independently-reconstructed tracks in the muon spectrometer to those in the inner detector. A combined fit is then performed to update the tracks by taking in account momenta of tracks in the inner detector and the muon spectrometer as well as energy losses in the calorimeters. Silicon-associated forward muons are combined muons with $|\eta| > 2.5$ and they are reconstructed by matching tracks found in the muon spectrometer with short track segments formed from hits in the pixel detector and the SCT.

Muon-spectrometer extrapolated muons are reconstructed from extrapolating tracks in the muon spectrometer without matching tracks in the inner detector. A fit with only tracks in the muon spectrometer is then performed to determine the parameters of these muons. In this way, muon reconstruction can be extended into the region of $2.5 < |\eta| < 2.7$, where the inner detector is not covered.

Inside-out combined muons are reconstructed from extrapolating tracks in the inner detector to find hits in the muon spectrometer. A combined fit is performed with tracks in the inner detector and hits found in the muon spectrometer. In this way, muons with limited muon spectrometer coverage or low $p_T$ can be recovered.

Segment-tagged muons are reconstructed from extrapolating tracks in the inner detector to find segments in the muon spectrometer. If at least one segment is found, a fit with only tracks in the inner detector is performed to determine the parameters of these muons. This method can also be used to recover muons with low $p_T$ or reduced acceptance in the muon spectrometer.

Calorimeter-tagged muons are reconstructed from extrapolating tracks in the inner detector to find calorimeter energy deposits that are consistent with the minimum-ionizing signature of muons. If energy deposits are found, a fit with only tracks in the inner detector is performed to determine the parameters of these muons. While other types of muons can be reconstructed with $p_T > 2$ GeV, an increased muon $p_T$ threshold of 5 GeV is used for these muons due to high background contamination at low $p_T$. This method targets to reconstruct muons with $|\eta| < 0.1$ and $15\,\text{GeV} < p_T < 100\,\text{GeV}$.
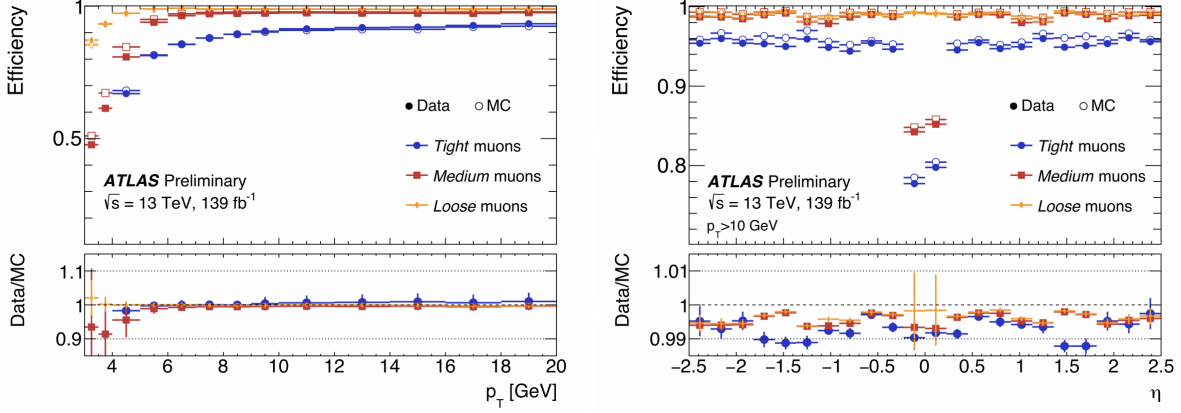
Figure 4.3: Muon reconstruction and identification efficiencies in data (filled dots) and MC (empty dots) as a function of $p_T$ in $J/\psi \to \mu^+\mu^-$ events (left) and $\eta$ in $Z \to \mu\mu$ events (right) for three identification working points: loose (yellow), medium (red) and tight (blue) [71].

Muon identification [71] can further improve the purity of prompt muons and suppress background mainly light-flavor quark decays. It uses detailed hit information on layers of both the inner detector and the muon spectrometer, track fit properties and variables related to the matching between the inner detector and the muon spectrometer. The requirements are different for different types of muons. There are three working points for most analyses: loose, medium, and tight. The reconstruction and identification efficiencies and also the ratios between data and simulation for the working points are shown in Figure 4.3. The tight working point has the purest prompt muons but with the lowest efficiency, while the medium working point is the recommendation from the ATLAS CP groups after considering the trade-off. Both medium and loose working points are used in the $W^\pm W^\pm W^\mp$ analysis. There are also two additional working points for extreme cases, low $p_T$ muons with $p_T \approx 3$ GeV and high $p_T$ muons with $p_T > 100$ GeV.

Muon isolation [71] can be used to suppress background from heavy-flavor quark decays with track and energy isolation variables. Similar to the electron isolation, the tight PLV working point [66] is also defined for muon isolation with cuts applied on the track isolation variable $p_T^{\text{varcone30}}$ and the BDT discriminate PLV. Muon $p_T^{\text{varcone30}}$ is defined as the sum of $p_T$ for all tracks with $p_T > 1$ GeV and loose vertex association within a varied-size cone of $\Delta R = \min(0.3, 10\,\text{GeV}/p_T)$ around the muon track direction. The muon track is excluded in the $p_T^{\text{varcone30}}$ calculation. The requirements of the tight PLV working point for muons with $p_T > 20$ GeV are summarized as follows:

$$p_T^{\text{varcone30}} < 0.15 \times p_T \quad \text{and} \quad -1.1 < \text{PLV} < \max\left(-0.88, -0.929774 + 2.9159 \times e^{-\frac{p_T}{10.2339\,\text{GeV}}}\right).$$
(4.2)

Three types of muons are selected in the $W^\pm W^\pm W^\mp$ analysis and the selection criteria are listed

in Table 4.3. Signal muons, also called ID muons, are the tightest muons and likely prompt muons. They need to pass basic kinematic requirement of $p_T > 20$ GeV and geometrical requirement of $|\eta| < 2.5$. The muons also need to be produced from the primary vertices by requiring the impact parameters with $|d_0/\sigma_{d_0}| < 3$ and $|z_0 \sin(\theta)| < 0.5$ mm. The medium working point is used for muon identification and the tight PLV working point is used for muon isolation.

**Selected Muons**

|  | ID | Anti-ID | Veto |
|---|---|---|---|
|  | Reconstructed muon candidates | | |
| $p_T$ | > 20 GeV | | > 4.5 GeV |
| $|\eta|$ | < 2.5 | | < 2.7 |
| $|d_0/\sigma_{d_0}|$ | < 3 | < 10 | - |
| $|z_0 \sin(\theta)|$ | < 0.5 mm | | - |
| Identification | Medium | | Loose |
| Isolation | TightPLV | - | - |
| Additional | - | !Signal | - |

Table 4.3: Summary of muon types used.

Anti-ID muons are muons that are likely to be non-prompt muons and are used in non-prompt background estimation described in Section 7.3. They do not have the isolation requirement and the $d_0$ requirement is loosened to $|d_0/\sigma_{d_0}| < 10$ while the $z_0$ requirement is the same. Anti-ID muons also need to be orthogonal to signal muons and have the same requirements as signal muons for $p_T$, $|\eta|$, and identification.

Veto muons are the loosest muons defined and are used to veto events with additional muons. They are only required to have $p_T > 4.5$ GeV, $|\eta| < 2.7$ and pass the loose muon identification working point.

MC events with muons need to be corrected for efficiencies to match what is observed in data. Based on the choice of muon identification and isolation working points, reconstruction and isolation scale factors are provided by the ATLAS CP groups and are included in the event weight calculation. The effect of muon identification and muon reconstruction is combined as a single reconstruction scale factor. Since impact parameters are used in muon selection, there is also another track-to-vertex-association scale factor provided to match data and used as another term in the event weight product.

## 4.5 Jets

Jets are defined as narrow cones of hadrons from the hadronization of quarks or gluons. Particle flow jets are used in this analysis and jet reconstruction starts with applying the particle flow algorithm [72] to tracks in the inner detector and topo-clusters in the calorimeters to remove double counting of energy in these two detectors. Good-quality tracks are first sorted in descending order in $p_T$. Then the first track is matched to a single topo-cluster and the expected energy of the particle is calculated from the track momentum and the topo-cluster position. After that, additional topo-clusters are added if needed to recover full shower energy. The updated expected energy of the particle is then subtracted from the topo-clusters matched to the track cell by cell. The remaining energy in the cells or the matched topo-clusters after the subtraction is removed if it is consistent with the expected shower fluctuation of a single particle. The procedure is repeated with other selected tracks one-by-one and in the end, the remaining objects without overlaps in energies are tracks, modified topo-clusters with parts of energies removed, and unchanged topo-clusters without matched tracks.

Those surviving topo-clusters and selected tracks that are matched to the primary vertex are then passed to the anti-$k_t$ algorithm [73] with the radius parameter $R = 0.4$. In the anti-$k_t$ algorithm, a distance between an object $i$ and the beam $B$ ($d_{iB}$) and a distance between two objects $i, j$ ($d_{ij}$) are defined as follows:

$$d_{iB} = p_{T,i}^{-2},$$

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

$$(4.3)$$

where $p_{T,i}$ and $p_{T,j}$ are the $p_T$ of objects $i, j$ respectively, $\Delta R_{ij}$ is the separation of objects $i, j$ in the $\eta - \phi$ plane, and $R$ is the radius parameter and $R = 0.4$ in ATLAS. If the minimum of all possible distances for the object $i$ is $d_{ij}$, these two objects $i, j$ are merged and distances are recalculated. If the minimum of all possible distances for the object $i$ is $d_{iB}$, the object $i$ is defined as a jet and is removed from the input set of objects. This procedure is repeated until there are no remaining objects in the input set and jets are reconstructed.

Jet energy scale calibration [72, 74] is then performed for those reconstructed particle flow jets. First, the origins of jets are corrected by recalculating jet momenta to point to the primary vertices instead of the detector center without affecting jet energy. Then, the jets are corrected from pileup. A jet area-based pileup subtraction is applied as a function of the transverse energy density times the jet area ($\rho \times A$), and the residual pileup is corrected as a function of the number of vertices $N_{PV}$ and the number of pileup collisions in a bunch crossing $\mu$ derived from the MC. After that, the jet momentum is calibrated to the particle-level energy based on the MC, and $\eta$ is corrected from

the effect due to cracks in the calorimeters. Next is the global sequential calibration which is used to reduce energy leakage effects and flavor dependence with variables from the calorimeters, the muon spectrometer, and tracks. In the end, a residual in situ calibration is applied only to data to account for differences between data and the MC from previous steps by comparing $p_T$ and $\eta$ of other well-measured reference jets in dijet, $Z/\gamma^*$+jet and multi-jet events.

Jet cleaning [75] is needed to remove background jets from non-collision processes. Since background jets can influence calculations of other objects in the event such as $E_T^{miss}$, the event is discarded if it contains a background jet. In this analysis, the default loose working point is used following the recommendation from the ATLAS CP groups.

Jet Vertex Tagger (JVT) [76, 77] is applied on jet to further suppress jets likely to be arising from the pileup collisions. It is a discriminate based on a k-nearest-neighbor algorithm which calculates the fraction of the jet energy due to tracks originating from the primary vertex. In this analysis, the default tight working point is used which requires $JVT > 0.5$ for jets with $|\eta| < 2.4$ and $20\,\text{GeV} < p_T < 60\,\text{GeV}$.

The bottom quark has a relatively long lifetime of approximately $10^{-12}$ s so that it can travel a short distance before decaying. This special characteristic can be used to separate $b$-jets from jets from other flavor quark decays, which is called $b$-tagging. In this analysis, a high-level $b$-tagging discriminant DL1r [78, 79] is applied for jets within the coverage of the inner detector, $|\eta| < 2.5$. DL1r is based on a deep feed-forward neural network with inputs from jet kinematics, $p_T$ and $\eta$, as well as outputs from several low-level $b$-tagging discriminants including the IP3D tagger based on impact parameters of tracks in $b$-jets, the SV1 tagger that reconstructs secondary vertices in jets, the JetFitter that reconstructs topological $b$-jet decay, and the RNNIP tagger that is based on a recurrent neural network which is used to learn dependencies of tracks in the same jet. There are four working points with different $b$-jet efficiencies provided by the ATLAS CP groups for DL1r: 60%, 70%, 77%, and 85%. The 85% working point is used for this analysis to have the highest $b$-jet efficiency so that $b$-jet veto is powerful enough to suppress the large non-prompt background from $b$-jets.

There are three types of jets used in this analysis and the requirements of those jets are listed in Table 4.4. In most cases, jets used in this analysis are referred to central jets with $|\eta| < 2.5$. The jet-related variables such as the number of jets are defined only based on the number of central jets if otherwise specified. In addition to the $\eta$ requirement, the central jets also need to have $p_T > 20\,\text{GeV}$ and pass the tight JVT working point.

$b$-jets are a subset of central jets that pass the 85% $b$-tagging working point for the DL1r tagger. They are used to veto events containing $b$-jets to suppress non-prompt background or to create CRs to estimate non-prompt background, which is described in details in Section 7.3.

Forward jets are defined as jets with $|\eta| \geq 2.5$ and they are only used as an input of the BDT to

**Selected Jets**

| | Central jets (jets) | $b$-jets | Forward jets |
|---|---|---|---|
| | Reconstructed anti-$k_t$ ($R = 0.4$) particle flow jet candidates | | |
| $p_{\mathrm{T}}$ | > 20 GeV | | > 30 GeV |
| $|\eta|$ | < 2.5 | | $2.5 \leq |\eta| < 4.5$ |
| JVT | Tight | | - |
| $b$-tagging | - | DL1r@85% | - |

Table 4.4: Summary of jet selection criteria used.

further separate signal and background in the SRs. A tighter $p_{\mathrm{T}}$ cut of $p_{\mathrm{T}} > 30$ GeV is required for forward jets due to larger background contamination in the forward region. There is also an upper limit on the jet $\eta$ with $|\eta| < 4.5$ due to the acceptance of the hadronic calorimeters in ATLAS.

MC events with jets need to be corrected for $b$-tagging and JVT efficiencies to match with what is observed in data. The $b$-tagging and JVT scale factors provided by the ATLAS CP groups are used in the calculation of the event weight.

## 4.6 Missing Transverse Momentum

Missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$) is defined as the momentum not detected in the transverse plane. It is due to the presence of particles that do not interact with the detector such as neutrinos. Since the initial momentum is unknown along the $z$ direction, $E_{\mathrm{T}}^{\mathrm{miss}}$ is only calculated in the transverse plane. From the conservation of momentum and the total momentum $x$ and $y$ components are zero before the collision, $E_{\mathrm{T}}^{\mathrm{miss}}$ can be calculated as the negative vector sum of transverse momenta for all objects detected. $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction [80] is performed along both $x$ and $y$ axes and can be expressed as:

$$E_{x(y)}^{\mathrm{miss}} = -\left( \sum p_{x(y)}^{\text{hard term}} + \sum p_{x(y)}^{\text{soft term}} \right), \tag{4.4}$$

and $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as $E_{\mathrm{T}}^{\mathrm{miss}} \equiv \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2}$. The $E_{\mathrm{T}}^{\mathrm{miss}}$ hard terms are reconstructed objects which are electrons, muons, and jets used in this analysis. The $E_{\mathrm{T}}^{\mathrm{miss}}$ soft terms are detector signals not matched to these reconstructed objects. Since this analysis uses particle flow jets, the soft terms are based on charged particle flow tracks that are not matched to any reconstructed objects and come from the primary vertex.

An object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance [81] is also used for this analysis. The $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is defined to test the hypothesis of $E_{\mathrm{T}}^{\mathrm{miss}}$ coming from invisible particles against energy fluctuations

and is express as an event-by-event log-likelihood ratio:

$$E_T^{\text{miss}} \text{ sig} = \sqrt{2\ln\left[\frac{\max_{p_T^{\text{inv}} \neq 0} \mathcal{L}(E_T^{\text{miss}}|p_T^{\text{inv}})}{\max_{p_T^{\text{inv}} = 0} \mathcal{L}(E_T^{\text{miss}}|p_T^{\text{inv}})}\right]}, \tag{4.5}$$

where $p_T^{\text{inv}}$ is the total momentum of invisible particles and the likelihood function $\mathcal{L}$ takes into account expected resolutions of all hard terms and variance for soft terms to account for neutral particles.

## 4.7 Overlap Removal

Overlap removal [82] is performed to remove duplicated reconstructions with the same physics object. The standard working point is used and the details are summarized in Table 4.5. There are four types of overlap removals: $e/e$, $e/\mu$, $e/$jet, and $\mu/$jet.

| Reject | Keep | Criteria |
|--------|------|----------|
| electron | electron | shared track, $p_T(e_1) < p_T(e_2)$ |
| muon | electron | calorimeter-tagged muon and shared track |
| electron | muon | shared track |
| jet | electron | $\Delta R < 0.2$ |
| electron | jet | $\Delta R < 0.4$ |
| jet | muon | $N_{\text{Track}} < 3$ and (ghost-associated or $\Delta R < 0.2$) |
| muon | jet | $\Delta R < 0.4$ |

Table 4.5: Summary of the standard overlap removal working point. Steps performed in listed order and only surviving objects participate in subsequent steps.

First, electrons are required not to share tracks and if two electrons have a shared track, the one with a lower $p_T$ is discarded.

Second, electrons and muons are required not to have shared tracks. If an electron shares the same track with a muon in the inner detector, the muon is discarded if it is a calorimeter-tagged muon because the muon is likely formed from energy deposits of the electron in the calorimeter; otherwise, the electron is discarded since muons can radiate photons which can be converted into electron-positron pairs.

If a jet is within $\Delta R < 0.2$ from an electron, the jet is discarded since it is likely that the jet is misreconstructed from the electron energy deposit. If the separation is $0.2 \leq \Delta R < 0.4$, the electron is discarded since this could be a typical signature of a jet decay.

After that, a jet with less than three associated tracks is discarded if overlapping with a muon within $\Delta R < 0.2$ or ghost-associated to a muon since the jet is likely due to energy deposits left by the muon. If the separation is $\Delta R < 0.4$, the muon is discarded since the muon is likely due to a jet decay or a jet that punches through the calorimeters.

## 4.8 Triggers

As mentioned in Section 3.2.6, it is impossible to store all collision data and triggers must be used to select events of physics interest. Due to the characteristics of the final products in the $W^{\pm}W^{\pm}W^{\mp}$ analysis, unprescaled single electron and muon triggers in the HLT are used and they are summarized in Table 4.6.

There is a change in the trigger menu between 2015 and 2016-2018 as the instantaneous luminosity of the LHC increased significantly and the requirements on triggers must be tightened to keep the designed record rate. Events used in this analysis must pass one of the triggers in the list and at least one of the selected electrons or muons needs to match the trigger-level object that fires the trigger except for the trigger HLT_e300_etcut. An additional trigger scale factor provided by the ATLAS CP groups is applied on MC events to match the trigger efficiency we had in data.

| Data year | Electron triggers | Muon triggers |
|---|---|---|
| 2015 | HLT_e24_lhmedium_L1EM20VH HLT_e60_lhmedium HLT_e120_lhloose | HLT_mu20_iloose_L1MU15 HLT_mu50 |
| 2016 - 2018 | HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0 HLT_e300_etcut | HLT_mu26_ivarmedium HLT_mu50 |

Table 4.6: List of triggers used.

The naming of the triggers is based on the requirements to pass those triggers and all triggers are HLT mentioned in Section 3.2.6. As for the electron triggers [52, 83], each of them has a requirement on the electron $E_T$, such as $E_T > 24$ GeV (e24), 26 GeV (e26), 60 GeV (e60), 120 GeV (e120), 140 GeV (e140) or 300 GeV (e300). The triggers also require electrons passing a likelihood discriminant with four working points: very loose (lhvloose), loose (lhloose), medium (lhmedium) and tight (lhtight). There are some additional requirements for some electron triggers: the trigger that is alternatively seeded by the EM cluster in the L1Calo with $E_T > 20$ GeV and an hadronic veto which depends on the $\eta$-varied $E_T$ threshold (L1EM20VH); the trigger that satisfies the loose-track-only isolation (ivarloose); the trigger that does not include the transverse impact parameter

$d_0$ in the likelihood calculation (nod0); and the trigger that has the $E_T$ requirement but no tracking requirements (etcut).

Similarly, the muon trigger [52, 84] requirements are also encoded in its name. Each trigger needs to have a requirement on the muon $p_T$, such as $p_T > 20$ GeV (mu20), 26 GeV (mu26) or 50 GeV (mu50). Some triggers need to satisfy loose (iloose) or medium (ivarmedium) track isolation working point. Some triggers have an alternative seed, for example a L1Muon trigger with $p_T > 15$ GeV (L1MU15).

# CHAPTER 5

# Data and Monte Carlo Samples

## 5.1   Data

This analysis uses the $pp$ collision data taken by the ATLAS detector at $\sqrt{s} = 13$ TeV from 2015 to 2018, which corresponds to the full Run 2 dataset. Events are required to be on the Good Run List (GRL) which is a list of good lumi-blocks (about 1-2 minutes of data taking). Lumi-blocks not on the GRL are affected by various detector problems and thus all events in these lumi-blocks are removed. Additionally, some events in the good lumi-blocks may be recorded when the LAr system, the tile calorimeter, or the SCT is not fully functional or some event information due to detector problems is missing, while other events in these lumi-blocks are fine. Instead of discarding all events in these lumi-blocks, an event cleaning is performed for data after passing the GRL to only remove those problematic events. The total integrated luminosity of the full Run 2 dataset after the cleaning is 139 fb$^{-1}$ with an uncertainty of 1.7% [85], obtained using the upgraded Luminosity Cherenkov Integrating Detector (LUCID), LUCID-2 detector [86], for the primary luminosity measurements.

## 5.2   Monte Carlo Simulation

### 5.2.1   Event generation and simulation

MC is used to simulate both signal and some background processes. The MC samples can be generated in four steps. The fist step is generating Parton Distribution Function (PDF). The $pp$ collisions in fact are collisions of partons (quarks and gluons in protons). The PDF $f_A^h(x,\mu_f^2)$ describes the probability of finding a certain type of parton $A$ with four-momentum fraction $x$ inside a hadron $h$ at a factorization scale $\mu_f$ which is defined as the energy scale for the boundary between short-distance hard partonic process and long-distance physics from hadronization.

The second step is generating matrix elements. The matrix elements include hard scattering processes generated from parton interactions, such as the $W^\pm W^\pm W^\mp$ production from quarks as well as the leptonic and hadronic decays of the $W$ bosons. Additionally, initial and final state radiations, which are defined as the radiations of photons or gluons from initial or final state partons, can also be included in the matrix elements. The matrix elements are calculated with numerical integration using MC event generator such as SHERPA [87], MADGRAPH [88] and POWHEG [89–92] (Next-to-Leading Order (NLO) in QCD in most cases) based on the PDF. Since loop diagrams are included in the matrix elements, renormalization is performed with a renormalization scale $\mu_\mathrm{r}$ which is the cut-off mass set for the loop momentum and the contributions from the momentum higher than $\mu_\mathrm{r}$ are killed.

The third step is generating parton shower. The parton shower includes jet fragmentation and hadronization, initial and final state radiation (mostly for soft and collinear particles), and underlying event. The jet fragmentation and hadronization generate hadronic jets from the final state partons in the matrix elements. The underlying event includes interactions of proton remnants from the $pp$ collisions. Parton shower is simulated using MC event generator such as SHERPA, PYTHIA [93] and HERWIG [94, 95] based on the PDF with an underlying event tune which is used to estimate underlying event by fitting to previous collider data. A resummation scale $\mu_\mathrm{q}$ is defined in parton shower as an upper cutoff scale for the parton shower evolution. However, due to the overlap of parton shower and matrix elements in initial and final state radiations, a merging scheme is used to remove the overlap with a merging scale in general. Matrix elements are used for radiations above the merging scale while parton shower is used for those below the merging scale.

The last step is simulating the detector response. For the ATLAS detector, GEANT 4 [96] is used for the simulation and the MC output format is the same as that for data.

As mentioned in Section 4.2, MC-simulated events were produced with three campaigns to match the pileup profiles observed in data taken in different time periods. Events generated in each MC campaign also need to be scaled to the integrated luminosity of the data taken in the corresponding time period. The scale factor for luminosity scaling can be expressed as:

$$\text{scale factor} = \frac{\sigma \times \eta_\mathrm{filter} \times k \times \mathcal{L}}{\sum \text{weight}}, \tag{5.1}$$

where $\sigma$ is the production cross section of the specific physics process that the MC simulates, $\eta_\mathrm{filter}$ is the filter efficiency for the filters used in the MC generation to enhance a certain final state, $k$ is the $k$-factor representing the correction from LO to NLO in most cases, $\mathcal{L}$ is the integrated luminosity of the data, and weight is the event weight assigned by the MC generator when producing the event. The final MC event weight that used in the comparison with data is a product of differ-

ent scale factors not only from luminosity scaling but also from pileup, triggers, *b*-tagging, JVT, electron reconstruction, electron identification, election isolation, muon reconstruction, muon isolation, and muon track-to-vertex-association based on objects found in each event, as mentioned in Chapter 4.

## 5.2.2 MC Samples

There are three types of MC samples based on their usages in this analysis: signal MC samples for the signal processes, prompt background MC samples for the prompt background processes defined in Section 7.1, and data-driven background MC samples for subtracting various backgrounds in the data-driven estimations described in Sections 7.3 and 7.4 and for estimating data-driven background systematic uncertainties described in Section 9.2.2. All MC samples used in this analysis are listed in Table 5.1. In addition, this table summarizes the decay mode, category, matrix element calculation and parton shower generators, as well as QCD accuracy and PDF used for each physics process.

### 5.2.2.1 Signal Samples

As mentioned in Chapter 1, both *WWW* and $WH \rightarrow WWW$ are considered as signal processes. On-shell *WWW* production is simulated with SHERPA 2.2.2 at NLO in QCD using OPENLOOPS 1 library [97–99] and at LO in QCD for up to two additional partons for the fully-leptonic decay and up to three additional partons for the SS semi-leptonic decay. The NNPDF3.0NNLO PDF set [100] is used. Parton shower is performed using the default SHERPA tune set by the SHERPA authors and merged with the matrix element calculations based on the Catani-Seymour dipole factorization [101, 102] using the MEPS@NLO prescription [103–106]. The PDF, parton shower, and merging scheme are the same for all SHERPA samples used in this analysis.

On-shell *WH* process with one additional jet is simulated with POWHEG-BOX 2 at NLO in QCD using the MiNLO prescription [107] for both fully-leptonic and SS semi-leptonic final states, and NNPDF3.0NLO is the PDF used for the matrix element calculation. PYTHIA 8 is used for parton shower with the AZNLO tune [108] and the CTEQ6L1 PDF set [109]. The MC cross section is normalized to the production cross section calculated at Next-to-Next-to-Leading Order (NNLO) in QCD and NLO in electroweak correction.

### 5.2.2.2 Background Samples

As for background MC samples, diboson processes including *WZ*, *ZZ*, and OS *WW* are simulated off-shell with SHERPA 2.2.2 for the fully-leptonic final states and on-shell with SHERPA 2.2.1 for the semi-leptonic final states. Matrix elements are calculated at NLO in QCD for up to one

additional parton and at LO for up to three additional partons. The loop-induced processes $gg \rightarrow ZZ$ or $WW$ are also included and simulated with SHERPA 2.2.2 at LO in QCD for up to one additional parton. Electroweak productions of diboson with two jets are calculated at LO in QCD and the sixth order of the electroweak coupling $\alpha_{EW}$, $\mathcal{O}(\alpha_{EW}^6)$.

SS $WW$ production with two jets is simulated with SHERPA 2.2.2 at LO in QCD for up to one additional parton for both QCD production at $\mathcal{O}(\alpha_{EW}^4)$ and electroweak production at $\mathcal{O}(\alpha_{EW}^6)$. The $WWW$ contribution in the electroweak SS $WW$ process has been removed to avoid double counting.

Other triboson productions such as $WWZ$, $WZZ$, and $ZZZ$ with fully-leptonic decays are simulated on-shell with SHERPA 2.2.2 using factorized gauge boson decays at NLO in QCD. $VVVj$ and $VVVjj$ ($V = W, Z$) are also included in the matrix element calculation at LO. The semi-leptonic decays of triboson productions are included in the diboson electroweak productions and thus there is no need to include them again.

Rare top-quark MC samples used in this analysis are $t\bar{t}V$ ($V = W, Z$), $t\bar{t}H$ and $tZ$. $t\bar{t}V$ samples are generated with MADGRAPH5_aMC@NLO 2.3.3 at NLO in QCD using the NNPDF3.0NLO PDF set. $t\bar{t}H$ samples are modeled with POWHEG-BOX 2 at NLO in QCD and the NNPDF3.0NLO PDF set. $tZ$ sample uses MADGRAPH5_aMC@NLO 2.3.3 at LO in QCD nd the NNPDF3.0LO PDF set. All samples use PYTHIA 8 for parton shower with the A14 tune [110] and NNPDF2.3LO.

Other top-quark processes including $t\bar{t}$ and single top are generated with POWHEG-BOX 2 at NLO in QCD and the NNPDF3.0NLO PDF set. PYTHIA 8 is used for parton shower with A14 tune and NNPDF2.3LO. The diagram removal scheme [111] is performed for the $Wt$ channel in the single top production to avoid overlapping with $t\bar{t}$.

$V\gamma$ ($V = W, Z$) samples are modeled using SHERPA 2.2.8 at NLO in QCD for up to one additional parton and LO in QCD for up to three additional partons.

$V$+jets ($V = W, Z$) samples are modeled using SHERPA 2.2.1 at NLO in QCD for up to two additional partons and LO in QCD for up to four additional partons. The overlap between $V\gamma$ and $V$+jets samples is removed at the analysis level [112].

| Process | Decay Mode/ Final State | Category | Generator | Accuracy | PDF |
|---|---|---|---|---|---|
| $WWW$ | $\ell\nu\ell\nu\ell\nu$ | Signal | SHERPA 2.2.2 | 0j@NLO+1,2j@LO | NNPDF3.0NNLO |
| | $\ell\nu\ell\nu jj(SS)$ | Signal | SHERPA 2.2.2 | 0j@NLO+1,2,3j@LO | NNPDF3.0NNLO |
| $WH$ | $\ell\nu\ell\nu\ell\nu(H \to WW)$ | Signal | POWHEG+PYTHIA 8 | NLO | NNPDF3.0NLO |
| | $\ell\nu\ell\nu jj(SS, H \to WW)$ | Signal | POWHEG+PYTHIA 8 | NLO | NNPDF3.0NLO |
| $WZ$ | $\ell\ell\ell\nu$ | Prompt | SHERPA 2.2.2 | 0,1j@NLO+2,3j@LO | NNPDF3.0NNLO |
| | $\ell\ell\ell\nu jj(\alpha_{EW}^6)$ | Prompt | SHERPA 2.2.2 | 0j@LO | NNPDF3.0NNLO |
| | $W \to jj + Z \to \ell\ell$ | DataDriven | SHERPA 2.2.1 | 0,1j@NLO+2,3j@LO | NNPDF3.0NNLO |
| | $W \to \ell\nu + Z \to jj$ | DataDriven | SHERPA 2.2.1 | 0,1j@NLO+2,3j@LO | NNPDF3.0NNLO |
| $ZZ$ | $\ell\ell\ell\ell$ | Prompt | SHERPA 2.2.2 | 0,1j@NLO+2,3j@LO | NNPDF3.0NNLO |
| | $\ell\ell\ell\ell jj(\alpha_{EW}^6)$ | Prompt | SHERPA 2.2.2 | 0j@LO | NNPDF3.0NNLO |
| | $gg \to \ell\ell\ell\ell$ | Prompt | SHERPA 2.2.2 | 0,1j@LO | NNPDF3.0NNLO |
| | $Z \to \ell\ell + Z \to jj$ | DataDriven | SHERPA 2.2.1 | 0,1j@NLO+2,3j@LO | NNPDF3.0NNLO |
| SS $WW$ | $\ell\nu\ell\nu jj(SS)$ | Prompt | SHERPA 2.2.2 | 0,1j@LO | NNPDF3.0NNLO |
| | $\ell\nu\ell\nu jj(SS, \alpha_{EW}^6)$ | Prompt | SHERPA 2.2.2 | 0,1j@LO | NNPDF3.0NNLO |
| OS $WW$ | $\ell\nu\ell\nu$ | DataDriven | SHERPA 2.2.2 | 0,1j@NLO+2,3j@LO | NNPDF3.0NNLO |
| | $\ell\nu\ell\nu jj(\alpha_{EW}^6)$ | DataDriven | SHERPA 2.2.2 | 0j@LO | NNPDF3.0NNLO |
| | $gg \to \ell\nu\ell\nu$ | DataDriven | SHERPA 2.2.2 | 0,1j@LO | NNPDF3.0NNLO |
| | $\ell\nu jj$ | DataDriven | SHERPA 2.2.1 | 0,1j@NLO+2,3j@LO | NNPDF3.0NNLO |
| $VVV$ | $WWZ \to \ell\ell\ell\nu\ell\nu$ | Prompt | SHERPA 2.2.2 | 0j@NLO+1,2j@LO | NNPDF3.0NNLO |
| | $WWZ \to \ell\nu\ell\nu\nu\nu$ | DataDriven | SHERPA 2.2.2 | 0j@NLO+1,2j@LO | NNPDF3.0NNLO |
| | $WZZ \to \ell\ell\ell\nu\nu\nu$ | Prompt | SHERPA 2.2.2 | 0j@NLO+1,2j@LO | NNPDF3.0NNLO |
| | $ZZZ \to \ell\ell\ell\ell\nu\nu$ | Prompt | SHERPA 2.2.2 | 0j@NLO+1,2j@LO | NNPDF3.0NNLO |
| | $ZZZ \to \ell\ell\nu\nu\nu\nu$ | DataDriven | SHERPA 2.2.2 | 0j@NLO+1,2j@LO | NNPDF3.0NNLO |
| $t\bar{t}W$ | Inclusive | Prompt | aMC@NLO +PYTHIA 8 | NLO | NNPDF3.0NLO |
| $t\bar{t}Z$ | $Z \to \ell\ell$ | Prompt | aMC@NLO +PYTHIA 8 | NLO | NNPDF3.0NLO |
| $t\bar{t}H$ | $t\bar{t} \to \ell$+jets | Prompt | POWHEG+PYTHIA 8 | NLO | NNPDF3.0NLO |
| | $t\bar{t} \to \ell\ell$+jets | Prompt | POWHEG+PYTHIA 8 | NLO | NNPDF3.0NLO |
| $tZ$ | $\geq 1\ell$+jets | Prompt | aMC@NLO +PYTHIA 8 | LO | NNPDF3.0LO |
| SingleTop | s-channel | DataDriven | POWHEG+PYTHIA 8 | NLO | NNPDF3.0NLO |
| | t-channel | DataDriven | POWHEG+PYTHIA 8 | NLO | NNPDF3.0NLO |
| | $Wt$-channel | DataDriven | POWHEG+PYTHIA 8 | NLO | NNPDF3.0NLO |
| $t\bar{t}$ | $\geq 1\ell$+jets | DataDriven | POWHEG+PYTHIA 8 | NLO | NNPDF3.0NLO |
| $Z\gamma$ | $Z \to \ell\ell$ | DataDriven | SHERPA 2.2.8 | 0,1j@NLO+2,3j@LO | NNPDF3.0NNLO |
| $W\gamma$ | $W \to \ell\nu$ | DataDriven | SHERPA 2.2.8 | 0,1j@NLO+2,3j@LO | NNPDF3.0NNLO |
| $Z$+jets | $Z \to \ell\ell$ | DataDriven | SHERPA 2.2.1 | 0,1,2j@NLO+3,4j@LO | NNPDF3.0NNLO |
| $W$+jets | $W \to \ell\nu$ | DataDriven | SHERPA 2.2.1 | 0,1,2j@NLO+3,4j@LO | NNPDF3.0NNLO |

Table 5.1: Summary of all MC samples used in the $W^\pm W^\pm W^\mp$ analysis. $\ell$ stands for leptons including $e$, $\mu$ and $\tau$.

# CHAPTER 6

# Event Selection

## 6.1   Preselection

After selecting objects in the data and MC samples, preselection cuts are applied at the event level for all regions. Those cuts have already been explained in Chapters 4 and 5, and can be summarized as follows:

- Passing the GRL for data only, mentioned in Section 5.1;

- Requiring a primary vertex in the event, mentioned in Section 4.2;

- Passing event cleaning criteria for data only, mentioned in Section 5.1;

- Passing jet cleaning criteria, mentioned in Section 4.5;

- Passing trigger selection and trigger matching, mentioned in Section 4.8.

## 6.2   $\ell^{\pm}\nu\ell^{\pm}\nu jj$ Signal Region

In the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR, two $W$ bosons decay leptonically while the third one decays hadronically. Event-level cuts for the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR after the preselection are summarized in Table 6.1 under the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR column. In order to reduce the contributions from physics processes with two OS leptons, the two leptons are required to be SS. Based on the lepton flavor, there are three channels considered: $ee$, $e\mu$, and $\mu\mu$. Note that for the simplicity, leptonically-decayed tau $\tau_{\text{lep}}$ is treated as $e$ or $\mu$ depending on the tau decay products while hadronically-decayed tau $\tau_{\text{had}}$ is treated as a jet at the analysis level. Events are vetoed if there is a third veto lepton (veto electron or veto muon defined in Sections 4.3 and 4.4) to reduce the $WZ$ background. The leading lepton must have $p_{\text{T}} > 27$ GeV to satisfy the requirement of triggers. The dilepton invariant mass, $m_{\ell\ell}$, is required to be between 40 GeV and 400 GeV for all channels and the $Z$ mass window between 80 GeV

and 100 GeV is excluded only in the *ee* channel to reduce charge flip background that containing a leptonically-decayed *Z* boson.

At least two jets are required in the SR since the third *W* boson decays hadronically. Events must not have any *b*-jet to reduce the contributions from top quark processes. Dijet variables, the pseudorapidity difference $\Delta\eta_{jj}$ and the invariant mass $m_{jj}$ of the two leading jets, are required to have $|\Delta\eta_{jj}| < 1.5$ and $m_{jj} < 160$ GeV to reduce VBS processes such as SS *WW* electroweak and QCD production. $E_T^{\text{miss}}$ significance is required to be larger than 3 in the *ee* channel to further reduce the contributions from the Drell-Yan process.

In addition, MC samples used for the prompt background estimations are required to pass a truth-level cut [113] to avoid double counting with data-driven background estimations. For each event, the two leptons selected are required to be prompt at the truth level.

## 6.3 $\quad \ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ Signal Region

In the $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ SR, all three *W* bosons are decayed leptonically. The cuts applied are looser than the $\ell^\pm\nu\ell^\pm\nu jj$ SR due to less background contamination expected. Event-level cuts for the $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ SR after the preselection are summarized in Table 6.1 under the $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ SR column. The sum of the charges of three leptons is required to be $\pm1$. In order to reduce the *Z*-related process, events are required to have no SFOS lepton pair, and only *eeμ* and *μμe* channels are selected. The leading lepton $p_T$ must be greater than 27 GeV due to the requirement of the triggers used. The event is vetoed if there is a fourth veto lepton in order to reduce *ZZ* background. Events are also vetoed if there is any *b*-jet in order to reduce top processes contribution.

Similar to the $\ell^\pm\nu\ell^\pm\nu jj$ SR, MC samples used for prompt background estimations are also required to pass a truth-level cut. For each event, the three leptons selected are required to be prompt at the truth level.

| | $\ell^\pm\nu\ell^\pm\nu jj$ SR | | | $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ SR | |
|---|---|---|---|---|---|
| Channel | $ee$ | $e\mu$ | $\mu\mu$ | $ee\mu$ | $\mu\mu e$ |
| $N_{\text{lepton}}$ | $= 2$ | | | $= 3$ | |
| Veto lepton | veto $3^{\text{rd}}$ lepton | | | veto $4^{\text{th}}$ lepton | |
| $p_{\text{T}}(\ell_1)$ | $> 27$ GeV | | | | |
| $q_\ell$ | SS | | | no SFOS, $\sum q_\ell = \pm1$ | |
| $m_{\ell\ell}$ | $40 < m_{\ell\ell} \leq 80$ GeV $100 \leq m_{\ell\ell} < 400$ GeV | $40 < m_{\ell\ell} < 400$ GeV | | - | |
| $N_{b\text{-jet}}$ | $= 0$ | | | | |
| $N_{\text{jet}}$ | $\geq 2$ | | | - | |
| $|\Delta\eta_{jj}|$ | $< 1.5$ | | | - | |
| $m_{jj}$ | $< 160$ GeV | | | - | |
| $E_{\text{T}}^{\text{miss}}$ sig | $> 3$ | - | | - | |

Table 6.1: List of event-level selection cuts used in the SRs. Channel is defined by lepton flavor in the regions. $N_{\text{lepton}}$ stands for the number of leptons. Veto lepton is veto electron or veto muon. $p_{\text{T}}(\ell_1)$ is the $p_{\text{T}}$ of the leading lepton. $q_\ell$ stands for the lepton charge. $m_{\ell\ell}$ is the invariant mass of two leptons. $N_{b\text{-jet}}$ is number of $b$-jets. $N_{\text{jet}}$ is number of jets. $|\Delta\eta_{jj}|$ is the pseudorapidity difference of the two leading jets. $m_{jj}$ is the invariant mass of the two leading jets. $E_{\text{T}}^{\text{miss}}$ sig stands for $E_{\text{T}}^{\text{miss}}$ significance.

# CHAPTER 7

# Background Estimation

## 7.1 Overview

Besides the $W^\pm W^\pm W^\mp$ signal process, other SM processes can also have similar final states and thus pass the requirements in the $\ell^\pm \nu \ell^\pm \nu j j$ and $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SRs. The SM background processes can be summarized into four major categories: prompt background, non-prompt background, photon conversion background, and charge-flip background.

The prompt background is referred to those processes that have two SS leptons or three leptons without a SFOS lepton pair at the truth level. This is the largest background in this analysis and can be reduced by vetoing the presence of the third or forth veto lepton and cuts on $m_{jj}$ and $\Delta \eta_{jj}$ in the $\ell^\pm \nu \ell^\pm \nu j j$ SR. The left-over prompt background mainly comes from processes with additional prompt leptons not detected or not reconstructed. The background is estimated using the MC samples listed in Table 5.1 under the "prompt" category and it includes $WZ$, $ZZ$, SS $WW$, $VVV(WWZ, WZZ, ZZZ)$, $t\bar{t}V(V = W, Z)$, $t\bar{t}H$, and $tZ$ processes. $WZ$ is the dominate component of the prompt background and details of the $WZ$ background estimation are explained in Section 7.2.

The non-prompt background is defined as processes with leptons from jet decays. This is the second-largest background in this analysis. A data-driven non-prompt rate method is used to estimate this background in the SRs. Physics processes that contribute to this background mainly come from $t\bar{t}$ process dominated by heavy-flavor jet decays, and can also come from the $W+$jets process dominated by light-flavor jet decays in the $\ell^\pm \nu \ell^\pm \nu j j$ SR. The background has been reduced significantly by the tight PLV electron and muon isolation working points and the $b$-jet veto with the highest $b$-tagging efficiency working point (85%) of the DL1r tagger. The details of non-prompt background modeling are documented in Section 7.3 and "NonPrompt" is used to indicate this background in plots and tables.

The photon conversion background includes $W\gamma$ and $Z\gamma$ processes where photons are misre-constructed as electrons. This is another major background in the $\ell^\pm \nu \ell^\pm \nu j j$ SR. Cuts are applied on the values of the electron author variable and the addAmbiguity variable to reduce this back-

ground. This photon conversion background is estimated using a data-driven photon conversion rate method which is similar to the non-prompt rate method. The details of the photon conversion background modeling are included in Section 7.4 and "PhotonConv" is used to indicate this background in plots and tables.

The charge-flip background is the background with an electron having its electric charge misidentified. The muon charge can also be wrongly measured but the contribution is negligible and thus is ignored in this analysis. The background can come from OS processes such as $Z + \text{jets}$, $t\bar{t}$ or OS $WW$, and the OS lepton pair is reconstructed as a SS pair after the charge of one electron is flipped. This is a minor background due to the ECIDS requirement and cuts applied on the $E_\mathrm{T}^\mathrm{miss}$ significance and $m_{\ell\ell}$ in the $ee$ channel of the $\ell^\pm \nu \ell^\pm \nu j j$ SR. A fully data-driven method is used to estimate this background. The details of the charge-flip background are explained in Section 7.5 and "ChargeFlip" is used to indicate this background in plots and tables.

Several dedicated CRs and VRs are defined to study the modeling of each background. Selection criteria used for these regions after the preselection are summarized in Tables 7.1 and 7.2 and compared with those used for the $\ell^\pm \nu \ell^\pm \nu j j$ and $\ell^\pm \nu \ell^\pm \ell^\mp \nu$ SRs. These regions are orthogonal to the SRs by construction except for the charge-flip CR and the $W$ sideband CR. The charge-flip CR is a very loose region to calculate the charge-flip rate and has a small overlap with the $\ell^\pm \nu \ell^\pm \nu j j$ SR. It is dominated by the $Z + \text{jets}$ process and the signal contribution is negligible in this region. The $W$ sideband CR is used to check the overall background modeling and the $\ell^\pm \nu \ell^\pm \nu j j$ SR is thus partially unblinded. The signal contribution is small since the two leading jets from signal events are expected from the decay of a $W$ boson and thus the $m_{jj}$ should be around the $W$ boson mass window.

| | $\ell^\pm \nu \ell^\pm \nu j j$ SR | $W$ sideband CR | $\ell^\pm \nu \ell^\pm \nu j j$ $b$-tagging CR | $Z$ window VR | Charge-flip CR |
|---|---|---|---|---|---|
| Channel | $ee$, $e\mu$, $\mu\mu$ | | | $ee$ | |
| $N_\mathrm{lepton}$ | $= 2$ | | | | |
| Veto lepton | veto $3^\mathrm{rd}$ lepton | | | | |
| $p_\mathrm{T}(\ell_1)$ | $> 27$ GeV | | | | |
| $q_\ell$ | SS | | | | - |
| $m_{\ell\ell}$ | $40 < m_{\ell\ell} < 400$ GeV, exclude $80 < m_{\ell\ell} < 100$ GeV in $ee$ only | | | $80 < m_{\ell\ell} < 100$ GeV | $60 < m_{\ell\ell} < 120$ GeV |
| $N_{b\text{-jet}}$ | $= 0$ | | $= 1$ | $= 0$ | - |
| $N_\mathrm{jet}$ | $\geq 2$ | | | | - |
| $|\Delta \eta_{jj}|$ | $< 1.5$ | | - | $< 1.5$ | - |
| $m_{jj}$ | $< 160$ GeV | $\leq 50$ GeV, $120 \leq m_{jj} < 160$ GeV | - | $< 160$ GeV | - |
| $E_\mathrm{T}^\mathrm{miss}$ sig | $> 3$ in $ee$ only | | - | - | - |

Table 7.1: List of the event-level selection cuts used in the regions with two leptons. The definitions of the symbols are the same as those in Table 6.1.

| | $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ SR | $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ $b$-tagging CR | $WZ01j$ CR | $WZ2j$ CR | $Z\gamma$ CR |
|---|---|---|---|---|---|
| Channel | eeμ, μμe | | eee, eeμ, μμe, μμμ | | eee, μμe |
| $N_{\text{lepton}}$ | = 3 | | | | |
| Veto lepton | veto 4th lepton | | | | |
| $p_{\text{T}}(\ell_1)$ | > 27 GeV | | | | |
| $q_\ell$ | no SFOS, $\sum q_\ell = \pm 1$ | $\sum q_\ell = \pm 1$ | SFOS | | |
| $m_{\ell\ell}$ (SFOS) | - | < 70 or > 110 GeV or no SFOS | - | - | - |
| $m_{\ell\ell\ell}$ | - | - | > 110 GeV | | $80 < m_{\ell\ell\ell} < 100$ GeV |
| $N_{b\text{-jet}}$ | = 0 | = 1 | = 0 | | |
| $N_{\text{jet}}$ | - | - | < 2 | $\geq 2$ | - |
| $E_{\text{T}}^{\text{miss}}$ sig | - | - | > 3 | | - |

Table 7.2: List of the event-level selection cuts used in the regions with three leptons. $m_{\ell\ell}$ (SFOS) is the invariant mass of the SFOS lepton pair. $m_{\ell\ell\ell}$ is the invariant mass of three leptons. Other symbols have the same definitions as those in Table 6.1.

## 7.2  *WZ* Background

*WZ* background is the dominant background in this analysis and MC simulation is used to estimate its contribution. The *WZ*+jets events can pass the $\ell^\pm\nu\ell^\pm\nu jj$ SR selection criteria if one of the two leptons in the *Z* boson decay is not reconstructed while the other one forms a SS lepton pair with the lepton from the *W* boson decay. In the $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ SR, the *Z* boson in the *WZ* process decays into two taus and one tau decays into an electron while the other one decays into a muon. There is no SFOS lepton pair if the Same-Flavor (SF) leptons from the *W* and *Z* decays are SS. Due to different numbers of jets in the $\ell^\pm\nu\ell^\pm\nu jj$ and $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ SRs, two dedicated CRs, *WZ*01*j* CR and *WZ*2*j* CR, are defined to study the modeling of *WZ*+jets production. The event-level cuts after the preselection for these two CRs are list in Table 7.2. They are the same as the cuts applied on the $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ SR except the lepton flavor, lepton charge, trilepton invariant mass $m_{\ell\ell\ell}$, and $E_{\text{T}}^{\text{miss}}$ significance. All combinations of lepton flavor are included and a SFOS lepton pair is required together with a $E_{\text{T}}^{\text{miss}}$ significance > 3 to reconstruct the *WZ* events. The requirement of $m_{\ell\ell\ell}$ > 110 GeV is used to suppress Drell-Yan events and ensures the orthogonality to the $Z\gamma$ CR which is used to estimate photon conversion rate and described in Section 7.4. Additionally, less than two jets and at least two jets are required in the *WZ*01*j* CR and the *WZ*2*j* CR respectively, as indicated by their names.

The event yields in these two CRs are shown in Tables 7.3 and 7.4. The distributions of number of jets, $m_{\ell\ell\ell}$, and $E_{\text{T}}^{\text{miss}}$ significance are shown in Figure 7.1. Compared to the data, the MC prediction is slightly underestimated in the *WZ*01*j* CR while clearly overestimated in the *WZ*2*j* CR as indicated in these plots and tables. Therefore, two *WZ* normalization factors, $\mu_{WZ01j}$ and $\mu_{WZ2j}$, are derived based from the fit in the *WZ*01*j* CR and the *WZ*2*j* CR described in Section 10.2.

The *WZ* normalization factors obtained in Section 10.2 are then applied back to the *WZ* CRs

|  | $eee$ | $ee\mu$ | $\mu\mu e$ | $\mu\mu\mu$ | Inclusive |
|---|---|---|---|---|---|
| *WWW* | 2.14 ± 0.10 | 5.64 ± 0.20 | 7.19 ± 0.18 | 5.00 ± 0.15 | 19.97 ± 0.33 |
| *WH* | 1.58 ± 0.01 | 4.09 ± 0.02 | 5.62 ± 0.03 | 3.77 ± 0.02 | 15.07 ± 0.04 |
| *WZ* | 1018.4 ± 9.2 | 1493 ± 10 | 1970 ± 14 | 2831 ± 12 | 7312 ± 23 |
| NonPrompt | 8.30 ± 0.57 | 17.36 ± 0.72 | 20.05 ± 0.73 | 28.38 ± 0.84 | 74.1 ± 1.4 |
| PhotonConv | 19.3 ± 3.8 | 8.9 ± 3.0 | 34.6 ± 4.3 | - | 62.8 ± 6.4 |
| *ZZ* | 37.56 ± 0.61 | 70.11 ± 0.84 | 75.1 ± 1.2 | 144.1 ± 1.3 | 326.9 ± 2.1 |
| *VVV* | 2.03 ± 0.04 | 2.84 ± 0.05 | 3.71 ± 0.06 | 4.74 ± 0.07 | 13.32 ± 0.11 |
| $t\bar{t}W$ | 0.41 ± 0.05 | 1.12 ± 0.09 | 1.36 ± 0.10 | 0.98 ± 0.08 | 3.87 ± 0.16 |
| $t\bar{t}Z$ | 0.64 ± 0.05 | 0.86 ± 0.06 | 1.18 ± 0.07 | 1.48 ± 0.08 | 4.16 ± 0.13 |
| $t\bar{t}H$ | 0.05 ± 0.01 | 0.13 ± 0.01 | 0.16 ± 0.01 | 0.11 ± 0.01 | 0.44 ± 0.02 |
| *tZ* | 3.16 ± 0.18 | 3.77 ± 0.20 | 5.13 ± 0.23 | 7.14 ± 0.27 | 19.21 ± 0.44 |
| Total | 1094 ± 10 | 1608 ± 11 | 2124 ± 15 | 3027 ± 13 | 7852 ± 24 |
| Data | 1175 | 1641 | 2279 | 3193 | 8288 |

Table 7.3: Event yields with statistical uncertainties for data and estimated SM processes in the *WZ*01*j* CR. The normalization scale factors for the *WZ* process are not applied in this table.

|  | $eee$ | $ee\mu$ | $\mu\mu e$ | $\mu\mu\mu$ | Inclusive |
|---|---|---|---|---|---|
| *WWW* | 0.86 ± 0.06 | 2.65 ± 0.11 | 3.27 ± 0.12 | 1.90 ± 0.10 | 8.68 ± 0.21 |
| *WH* | 0.34 ± 0.01 | 0.88 ± 0.01 | 1.19 ± 0.01 | 0.80 ± 0.01 | 3.21 ± 0.02 |
| *WZ* | 416.3 ± 2.4 | 571.5 ± 2.9 | 762.1 ± 3.6 | 1038.2 ± 4.2 | 2788.0 ± 6.7 |
| NonPrompt | 2.18 ± 0.33 | 7.18 ± 0.45 | 7.72 ± 0.46 | 10.26 ± 0.46 | 27.34 ± 0.86 |
| PhotonConv | 8.1 ± 2.4 | 1.5 ± 1.6 | 8.1 ± 2.2 | - | 17.7 ± 3.6 |
| *ZZ* | 13.37 ± 0.27 | 17.81 ± 0.27 | 27.11 ± 0.38 | 36.67 ± 0.44 | 94.95 ± 0.70 |
| *VVV* | 1.22 ± 0.03 | 1.72 ± 0.04 | 2.17 ± 0.05 | 2.67 ± 0.05 | 7.78 ± 0.09 |
| $t\bar{t}W$ | 0.57 ± 0.07 | 1.37 ± 0.11 | 1.86 ± 0.13 | 1.44 ± 0.11 | 5.23 ± 0.21 |
| $t\bar{t}Z$ | 4.80 ± 0.16 | 6.82 ± 0.19 | 8.34 ± 0.21 | 10.23 ± 0.24 | 30.20 ± 0.40 |
| $t\bar{t}H$ | 0.28 ± 0.02 | 0.66 ± 0.02 | 0.85 ± 0.03 | 0.65 ± 0.03 | 2.44 ± 0.05 |
| *tZ* | 2.21 ± 0.15 | 2.88 ± 0.17 | 4.29 ± 0.21 | 5.18 ± 0.23 | 14.56 ± 0.38 |
| Total | 450.2 ± 3.4 | 615.0 ± 3.4 | 827.0 ± 4.3 | 1108.0 ± 4.3 | 3000.1 ± 7.7 |
| Data | 395 | 539 | 689 | 902 | 2525 |

Table 7.4: Event yields with statistical uncertainties for data and estimated SM processes in the *WZ*2*j* CR. The normalization scale factors for the *WZ* process are not applied in this table.

to check again the modeling of the *WZ* production. The event yields in these two CRs are shown in Tables 7.5 and 7.6. The distributions of number of jets, $m_{\ell\ell\ell}$, $E_T^{\mathrm{miss}}$ significance ratio (defined as $10 \times E_T^{\mathrm{miss}}$ Significance$/E_T^{\mathrm{miss}}$), and third lepton $p_T$ in the *WZ*01*j* CR are shown in Figure 7.2. The distributions of the number of jets , $m_{jj}$, $E_T^{\mathrm{miss}}$ significance, and leading forward jet $p_T$ in the *WZ*2*j* CR are shown in Figure 7.3. The distributions of those variables in each channel are attached in Appendices A.1 and A.2 for the *WZ*01*j* CR and the *WZ*2*j* CR respectively. Note that only statistical uncertainties of the MC predictions and data are included for tables and plots. Good agreement between the data and simulation is achieved within uncertainties. All plots and tables after this section have the *WZ* normalization factors applied by default otherwise specified.
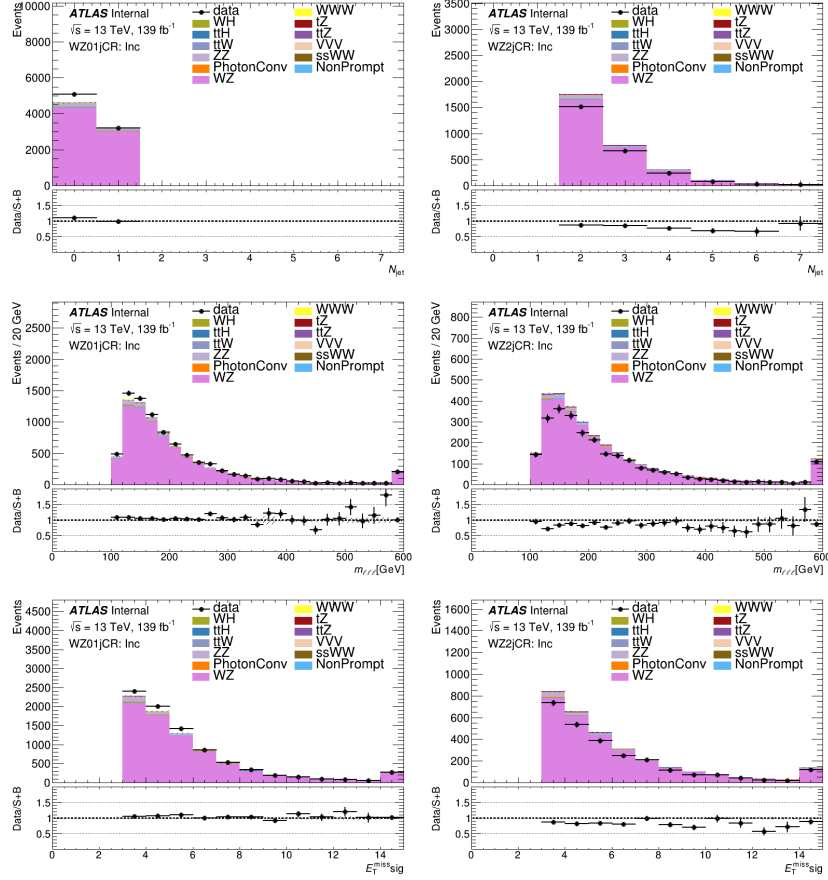
Figure 7.1: The number of jets (top), $m_{\ell\ell\ell}$ (middle), and $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (bottom) distributions with statistical uncertainties in the $WZ01j$ CR (left) and $WZ2j$ CR (right). The normalization scale factors for the $WZ$ process are not applied in these plots.

| | $eee$ | $ee\mu$ | $\mu\mu e$ | $\mu\mu\mu$ | Inclusive |
|---|---|---|---|---|---|
| $WWW$ | $2.14 \pm 0.10$ | $5.64 \pm 0.20$ | $7.19 \pm 0.18$ | $5.00 \pm 0.15$ | $19.97 \pm 0.33$ |
| $WH$ | $1.58 \pm 0.01$ | $4.09 \pm 0.02$ | $5.62 \pm 0.03$ | $3.77 \pm 0.02$ | $15.07 \pm 0.04$ |
| $WZ$ | $1071.3 \pm 9.6$ | $1571 \pm 11$ | $2072 \pm 15$ | $2978 \pm 13$ | $7692 \pm 24$ |
| NonPrompt | $8.54 \pm 0.65$ | $17.58 \pm 0.78$ | $20.94 \pm 0.81$ | $28.87 \pm 0.87$ | $75.9 \pm 1.6$ |
| PhotonConv | $18.7 \pm 3.8$ | $8.3 \pm 3.0$ | $34.4 \pm 4.3$ | - | $61.4 \pm 6.5$ |
| $ZZ$ | $37.56 \pm 0.61$ | $70.11 \pm 0.84$ | $75.1 \pm 1.2$ | $144.1 \pm 1.3$ | $326.9 \pm 2.1$ |
| $VVV$ | $2.03 \pm 0.04$ | $2.84 \pm 0.05$ | $3.71 \pm 0.06$ | $4.74 \pm 0.07$ | $13.32 \pm 0.11$ |
| $t\bar{t}W$ | $0.41 \pm 0.05$ | $1.12 \pm 0.09$ | $1.36 \pm 0.10$ | $0.98 \pm 0.08$ | $3.87 \pm 0.16$ |
| $t\bar{t}Z$ | $0.64 \pm 0.05$ | $0.86 \pm 0.06$ | $1.18 \pm 0.07$ | $1.48 \pm 0.08$ | $4.16 \pm 0.13$ |
| $t\bar{t}H$ | $0.05 \pm 0.01$ | $0.13 \pm 0.01$ | $0.16 \pm 0.01$ | $0.11 \pm 0.01$ | $0.44 \pm 0.02$ |
| $tZ$ | $3.16 \pm 0.18$ | $3.77 \pm 0.20$ | $5.13 \pm 0.23$ | $7.14 \pm 0.27$ | $19.21 \pm 0.44$ |
| Total | $1146 \pm 10$ | $1685 \pm 11$ | $2227 \pm 16$ | $3174 \pm 13$ | $8233 \pm 25$ |
| Data | 1175 | 1641 | 2279 | 3193 | 8288 |

Table 7.5: Event yields with statistical uncertainties for data and estimated SM processes in the $WZ01j$ CR. The normalization scale factors for the $WZ$ process are applied in this table.

60

|  | *eee* | *eeμ* | *μμe* | *μμμ* | Inclusive |
|---|---|---|---|---|---|
| *WWW* | 0.86 ± 0.06 | 2.65 ± 0.11 | 3.27 ± 0.12 | 1.90 ± 0.10 | 8.68 ± 0.21 |
| *WH* | 0.34 ± 0.01 | 0.88 ± 0.01 | 1.19 ± 0.01 | 0.8 ± 0.01 | 3.21 ± 0.02 |
| *WZ* | 346.3 ± 2.0 | 475.5 ± 2.4 | 634.1 ± 3.0 | 863.8 ± 3.5 | 2319.6 ± 5.6 |
| NonPrompt | 3.21 ± 0.38 | 8.51 ± 0.49 | 9.19 ± 0.50 | 11.56 ± 0.48 | 32.47 ± 0.93 |
| PhotonConv | 8.9 ± 2.4 | 2.2 ± 1.6 | 8.5 ± 2.2 | - | 19.6 ± 3.6 |
| *ZZ* | 13.37 ± 0.27 | 17.81 ± 0.27 | 27.11 ± 0.38 | 36.67 ± 0.44 | 94.95 ± 0.70 |
| *VVV* | 1.22 ± 0.03 | 1.72 ± 0.04 | 2.17 ± 0.05 | 2.67 ± 0.05 | 7.78 ± 0.09 |
| $t\bar{t}W$ | 0.57 ± 0.07 | 1.37 ± 0.11 | 1.86 ± 0.13 | 1.44 ± 0.11 | 5.23 ± 0.21 |
| $t\bar{t}Z$ | 4.80 ± 0.16 | 6.82 ± 0.19 | 8.34 ± 0.21 | 10.23 ± 0.24 | 30.20 ± 0.40 |
| $t\bar{t}H$ | 0.28 ± 0.02 | 0.66 ± 0.02 | 0.85 ± 0.03 | 0.65 ± 0.03 | 2.44 ± 0.05 |
| *tZ* | 2.21 ± 0.15 | 2.88 ± 0.17 | 4.29 ± 0.21 | 5.18 ± 0.23 | 14.56 ± 0.38 |
| Total | 382.1 ± 3.1 | 521.0 ± 3.0 | 700.9 ± 3.8 | 934.9 ± 3.6 | 2538.7 ± 6.8 |
| Data | 395 | 539 | 689 | 902 | 2525 |

Table 7.6: Event yields with statistical uncertainties for data and estimated SM processes in the *WZ2j* CR. The normalization scale factors for the *WZ* process are applied in this table.
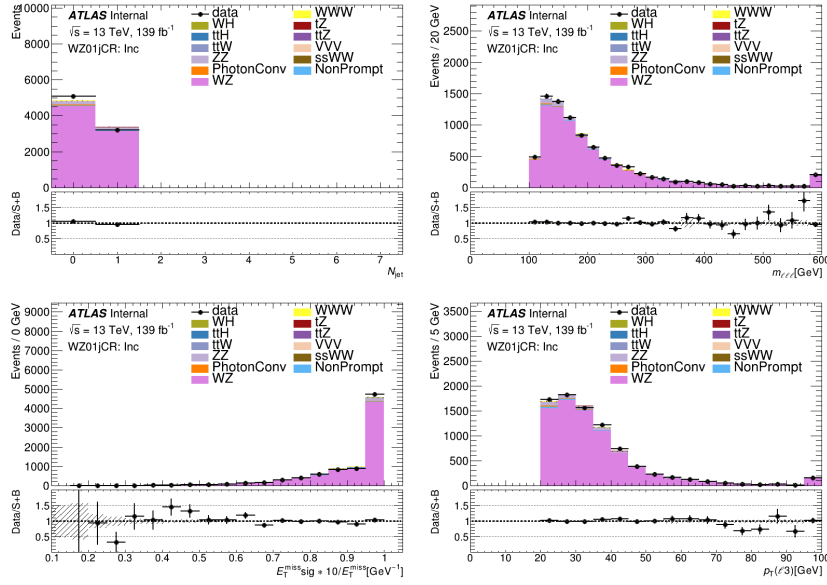


Figure 7.2: The distributions of the number of jets (top left), $m_{\ell\ell\ell}$ (top right), $E_T^{\text{miss}}$ significance ratio ($10 \times E_T^{\text{miss}}$ Significance$/E_T^{\text{miss}}$) (bottom left), and third lepton $p_T$ (bottom right) with statistical uncertainties in the *WZ01j* CR. The normalization scale factors for the *WZ* process are applied in these plots.
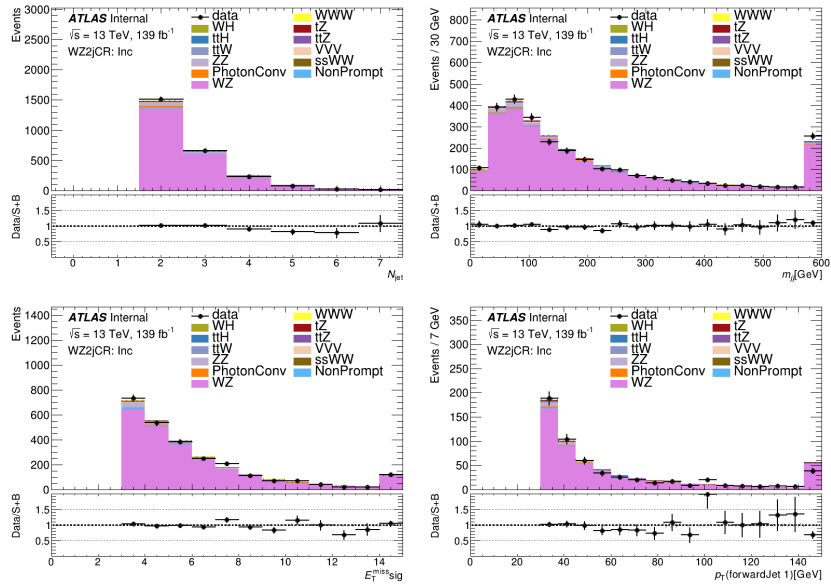
Figure 7.3: The distributions of the number of jets (top left), $m_{jj}$ (top right), $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (bottom left), and leading forward jet $p_{\mathrm{T}}$ (bottom right) with statistical uncertainties in the $WZ2j$ CR. The normalization scale factors for the $WZ$ process are applied in these plots.

## 7.3 Non-prompt Background

### 7.3.1 Non-prompt Background Compositions

The non-prompt background is the second-largest background in this analysis. It mainly comes from physics processes with leptons from jet decays. To investigate the composition of the non-prompt background in the SRs, two CRs, $\ell^\pm \nu \ell^\pm \nu jj$ anti-ID CR and $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ anti-ID CR, are defined with the same event-level cuts as the SRs. For object-level cuts, instead of all leptons in the event being ID leptons in the SRs, events in these CRs have one anti-ID lepton (anti-ID electron or anti-ID muon) while other leptons are ID leptons (ID electrons and ID muons).

We use MC events without the requirement of truth-level prompt leptons and compare the predictions to data in these CRs to understand the non-prompt background composition. The distributions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, anti-ID lepton $p_{\mathrm{T}}$, and number of jets in these CRs are shown in Figure 7.4. The event yield tables in these CRs are shown in Tables 7.7 and 7.8. The agreement between the data and MC simulation is found to be reasonable, thus the MC simulation can be used to indicate the background compositions in these two regions. Based on the MC simulation, the composition of the non-prompt background is dominated by $t\bar{t}$ process in both CRs while $W$+jets process contributes approximately one third of the non-prompt background in the $\ell^\pm \nu \ell^\pm \nu jj$ anti-ID CR. The non-prompt background composition in the SRs should be similar to the corresponding CRs due to the same event-level cuts applied.
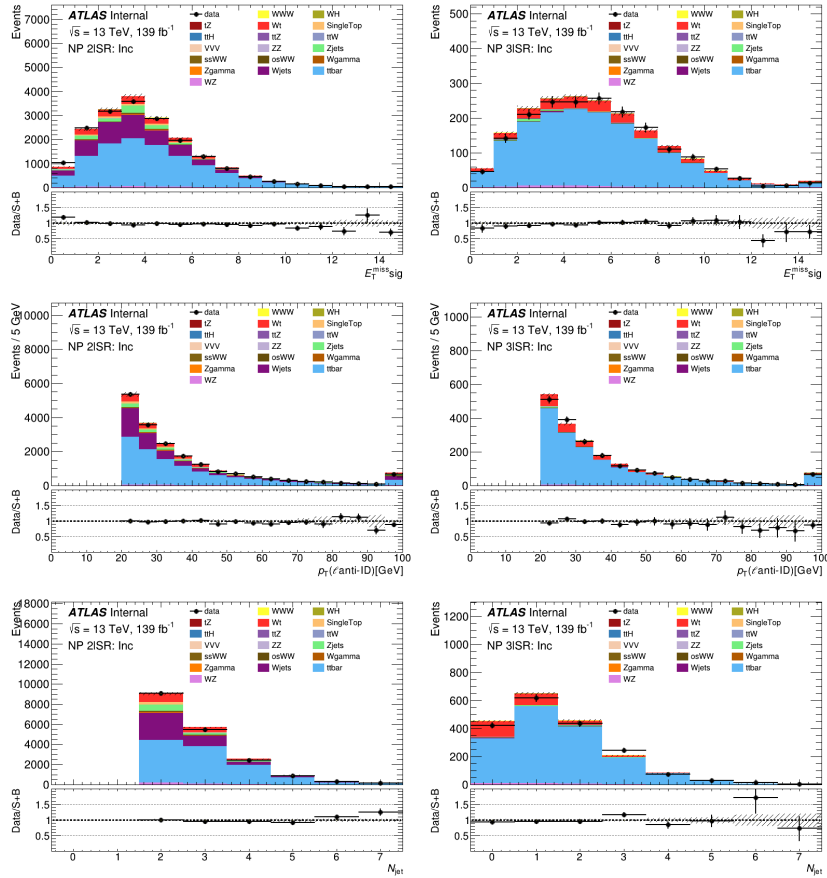
Figure 7.4: The $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (top), the anti-ID lepton $p_{\mathrm{T}}$ (middle), and the number of jets (bottom) distributions with statistical uncertainties in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ anti-ID CR (left) and the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ anti-ID CR (right). Only MC simulated events are used for the predictions.

|  | _ee_ | _eμ_ | _μμ_ | Inclusive |
|---|---|---|---|---|
| _WWW_ | 4.53 ± 0.15 | 13.77 ± 0.26 | 6.41 ± 0.17 | 24.71 ± 0.34 |
| _WH_ | 3.58 ± 0.03 | 14.13 ± 0.09 | 7.46 ± 0.10 | 25.17 ± 0.14 |
| _WZ_ | 45.8 ± 1.1 | 178.7 ± 2.7 | 88.4 ± 1.9 | 312.9 ± 3.4 |
| _tt̄_ | 760 ± 11 | 5589 ± 30 | 4520 ± 27 | 10869 ± 41 |
| _W_+jets | 340 ± 55 | 2380 ± 130 | 1526 ± 58 | 4240 ± 150 |
| _Zγ_ | 43 ± 12 | 22 ± 11 | 5.1 ± 2.5 | 70 ± 17 |
| _Wγ_ | 18.4 ± 4.1 | 70 ± 11 | 3.2 ± 8.5 | 92 ± 14 |
| OS _WW_ | 28.2 ± 1.7 | 82.4 ± 3.3 | 25.4 ± 2.3 | 135.9 ± 4.4 |
| SS _WW_ | 4.62 ± 0.09 | 14.84 ± 0.15 | 7.27 ± 0.11 | 26.73 ± 0.21 |
| _Z_+jets | 378 ± 58 | 338 ± 33 | 159 ± 15 | 874 ± 69 |
| _ZZ_ | 2.68 ± 0.20 | 14.2 ± 0.45 | 5.98 ± 0.29 | 22.87 ± 0.58 |
| _VVV_ | 0.08 ± 0.01 | 0.23 ± 0.02 | 0.08 ± 0.01 | 0.39 ± 0.02 |
| _tt̄W_ | 2.71 ± 0.15 | 11.03 ± 0.29 | 6.70 ± 0.23 | 20.44 ± 0.40 |
| _tt̄Z_ | 0.50 ± 0.06 | 1.80 ± 0.11 | 0.99 ± 0.08 | 3.29 ± 0.15 |
| _tt̄H_ | 0.72 ± 0.03 | 2.93 ± 0.07 | 1.94 ± 0.05 | 5.59 ± 0.09 |
| SingleTop | 19.7 ± 1.0 | 184.3 ± 3.1 | 158.6 ± 2.9 | 362.6 ± 4.4 |
| _Wt_ | 107.2 ± 3.7 | 842 ± 10 | 635.7 ± 9.0 | 1585 ± 14 |
| _tZ_ | 0.27 ± 0.05 | 1.53 ± 0.12 | 1.07 ± 0.10 | 2.87 ± 0.17 |
| Total | 1759 ± 82 | 9760 ± 140 | 7159 ± 67 | 18670 ± 170 |
| Data | 1545 | 9600 | 7095 | 18240 |

Table 7.7: Event yields with statistical uncertainties for data and SM predictions in the $\ell^\pm \nu \ell^\pm \nu j j$ anti-ID CR. Only MC simulated events are used for the predictions.

|  | _eeμ_ | _μμe_ | Inclusive |
|---|---|---|---|
| _WWW_ | 2.82 ± 0.11 | 3.07 ± 0.13 | 5.90 ± 0.17 |
| _WH_ | 2.58 ± 0.02 | 2.96 ± 0.03 | 5.54 ± 0.04 |
| _WZ_ | 20.91 ± 0.87 | 10.68 ± 0.61 | 31.6 ± 1.1 |
| _tt̄_ | 422.9 ± 8.0 | 1159 ± 13 | 1582 ± 16 |
| _W_+jets | - | 1.7 ± 1.7 | 1.7 ± 1.7 |
| _Zγ_ | 0.11 ± 0.06 | 0.09 ± 0.06 | 0.21 ± 0.09 |
| _Wγ_ | - | 0.01 ± 0.01 | 0.01 ± 0.01 |
| OS _WW_ | 3.38 ± 0.55 | 8.13 ± 0.66 | 11.51 ± 0.86 |
| SS _WW_ | 0.14 ± 0.01 | 0.07 ± 0.01 | 0.21 ± 0.02 |
| _Z_+jets | 7.4 ± 2.4 | 3.4 ± 4.1 | 10.7 ± 4.7 |
| _ZZ_ | 2.13 ± 0.13 | 1.02 ± 0.10 | 3.14 ± 0.16 |
| _VVV_ | 0.38 ± 0.02 | 0.32 ± 0.02 | 0.70 ± 0.03 |
| _tt̄W_ | 1.69 ± 0.12 | 2.46 ± 0.14 | 4.14 ± 0.18 |
| _tt̄Z_ | 0.45 ± 0.06 | 0.33 ± 0.05 | 0.78 ± 0.08 |
| _tt̄H_ | 0.71 ± 0.03 | 0.98 ± 0.03 | 1.68 ± 0.04 |
| SingleTop | 0.06 ± 0.06 | 0.27 ± 0.10 | 0.34 ± 0.12 |
| _Wt_ | 66.8 ± 2.9 | 174.0 ± 4.7 | 240.8 ± 5.5 |
| _tZ_ | 0.15 ± 0.04 | 0.07 ± 0.03 | 0.22 ± 0.05 |
| Total | 532.5 ± 8.9 | 1369 ± 15 | 1901 ± 17 |
| Data | 497 | 1341 | 1838 |

Table 7.8: Event yields with statistical uncertainties for data and SM predictions in the $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ anti-ID CR. Only MC simulated events are used for the predictions.

## 7.3.2 Non-prompt Rate Calculation

Since the dominate process in the non-prompt background is the $t\bar{t}$ process, two $t\bar{t}$-dominated CRs, $\ell^{\pm}\nu\ell^{\pm}\nu jj$ $b$-tagging CR and $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ $b$-tagging CR, are defined and the cuts after the preselection are listed in Tables 7.1 and 7.2. The $\ell^{\pm}\nu\ell^{\pm}\nu jj$ $b$-tagging CR has the same event-level cuts as the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR except for number of $b$-jets, $\Delta\eta_{jj}$, $m_{jj}$ and the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance. The number of $b$-jets in each event is required to be one to select the top-related processes. The cuts of $\Delta\eta_{jj}$, $m_{jj}$ and the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR are dropped to increase statistics in the CR since the non-prompt background has no dependence on these variables. The $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ $b$-tagging CR has the same event-level cuts as the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ SR except for number of $b$-jets, lepton charge and $m_{\ell\ell}$ for the SFOS pair. Similar to the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ $b$-tagging CR, the number of $b$-jets is required to be one. Events in the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ $b$-tagging CR can have a SFOS lepton pair to increase statistics since the non-prompt background has no dependence on that, but $m_{\ell\ell}$ of the SFOS lepton pair, if there is one, is required to be outside of the $Z$ mass window (between 70 GeV and 110 GeV) to reduce the contribution from the $WZ$ process. Similar to the anti-ID $\ell^{\pm}\nu\ell^{\pm}\nu jj$ CR and the anti-ID $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ CR, there are the anti-ID $\ell^{\pm}\nu\ell^{\pm}\nu jj$ $b$-tagging CR and the anti-ID $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ $b$-tagging CR which have the same event-level cuts but different object level cuts with one anti-ID lepton. Note that the anti-ID lepton in the anti-ID $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ $b$-tagging CR must be from the SS lepton pair due to the charge requirement of the $t\bar{t}$ process. In this thesis, ID regions are referring to those regions with only ID leptons while anti-ID regions are those with one anti-ID lepton and remaining leptons in the anti-ID regions are ID leptons.

A data-driven method is used to estimate the non-prompt background. The basic idea is to establish a map between number of events in the ID regions and number of events in the anti-ID regions with non-prompt rates. The electron and muon non-prompt rates are calculated in the ID and anti-ID $b$-tagging CRs by minimizing the negative log likelihood function ($\mathcal{L}$) defined as:

$$
\begin{aligned}
\mathcal{L}(\epsilon_e, \epsilon_\mu) &= \prod_{c\in\text{Channels}} \prod_{b_1\in p_{\mathrm{T}}(\ell_1)} \prod_{b_2\in p_{\mathrm{T}}(\ell_2)} \prod_{b_3\in p_{\mathrm{T}}(\ell_3)} \mathrm{Pois}\left(N_{\mathrm{ID},c,b_1,b_2,(b_3)}^{\mathrm{Data}} | N_{\mathrm{ID},c,b_1,b_2,(b_3)}^{\mathrm{Exp}}\right) \\
&= \prod_{c,b_1,b_2,(b_3)} \frac{(N_{\mathrm{ID},c,b_1,b_2,(b_3)}^{\mathrm{Exp}})^{N_{\mathrm{ID},c,b_1,b_2,(b_3)}^{\mathrm{Data}}} e^{-N_{\mathrm{ID},c,b_1,b_2,(b_3)}^{\mathrm{Exp}}}}{N_{\mathrm{ID},c,b_1,b_2,(b_3)}^{\mathrm{Data}}!},
\end{aligned}
\tag{7.1}
$$

where $\epsilon_e$ and $\epsilon_\mu$ are non-prompt rates for electron and muon respectively, while $N_{\mathrm{ID},c,b_1,b_2,(b_3)}^{\mathrm{Data}}$ and $N_{\mathrm{ID},c,b_1,b_2,(b_3)}^{\mathrm{Exp}}$ are number of data and number of expected events in a channel $c$ ($ee$, $e\mu$, $\mu\mu$, $ee\mu$, or $\mu\mu e$) in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ or $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ $b$-tagging CRs (ID region) with the leading lepton $p_{\mathrm{T}}$ in bin $b_1$, the subleading lepton $p_{\mathrm{T}}$ in bin $b_2$, and the third lepton $p_{\mathrm{T}}$ in bin $b_3$ for events with three leptons only. The number of expected events in a channel $c$ and lepton $p_{\mathrm{T}}$ bins $b_1$, $b_2$ (, $b_3$) in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$

or $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ b-tagging CR can be expressed as:

$$N^{\text{Exp}}_{\text{ID},c,b_1,b_2(,b_3)} = N^{\text{Prompt}}_{\text{ID},c,b_1,b_2(,b_3)} + N^{\text{ChargeFlip}}_{\text{ID},c,b_1,b_2(,b_3)} + N^{\text{PhotonConv}}_{\text{ID},c,b_1,b_2(,b_3)} + N^{\text{NonPrompt}}_{\text{ID},c,b_1,b_2(,b_3)}, \qquad (7.2)$$

where $N^{\text{Prompt}}_{\text{ID},c,b_1,b_2(,b_3)}$ is the prompt MC simulation with truth selection, $N^{\text{ChargeFlip}}_{\text{ID},c,b_1,b_2(,b_3)}$ and $N^{\text{PhotonConv}}_{\text{ID},c,b_1,b_2(,b_3)}$ are the data-driven estimations of the charge-flip background and the photon conversion background respectively, and $N^{\text{NonPrompt}}_{\text{ID},c,b_1,b_2(,b_3)}$ is the data-driven estimation of the non-prompt background that can be expressed as:

$$
\begin{aligned}
N^{\text{NonPrompt}}_{\text{ID},c,b_1,b_2(,b_3)} &= \sum_{\ell_{\text{Anti-ID}} \in \{\ell_1, \ell_2(,\ell_3)\}} \left( \epsilon^{b_{\text{Anti-ID}}}_{f_{\text{Anti-ID}}} \times N^{\text{NonPrompt},\ell_{\text{Anti-ID}}}_{\text{Anti-ID},c,b_1,b_2(,b_3)} \right) \\
&= \sum_{\ell_{\text{Anti-ID}} \in \{\ell_1, \ell_2(,\ell_3)\}} \left[ \epsilon^{b_{\text{Anti-ID}}}_{f_{\text{Anti-ID}}} \times \left( N^{\text{Data},\ell_{\text{Anti-ID}}}_{\text{Anti-ID},c,b_1,b_2(,b_3)} - N^{\text{MC},\ell_{\text{Anti-ID}}}_{\text{Anti-ID},c,b_1,b_2(,b_3)} \right) \right],
\end{aligned}
\qquad (7.3)
$$

where $f_{\text{Anti-ID}}$ is the flavor of the anti-ID lepton and $b_{\text{Anti-ID}}$ is the bin in the distribution of the anti-ID lepton $p_{\text{T}}$. The non-prompt rate is calculated for each anti-ID lepton flavor and $p_{\text{T}}$ bin. $\ell_{\text{Anti-ID}}$ is the anti-ID lepton in the event which can be the leading lepton ($\ell_1$), the subleading lepton ($\ell_2$) or the third lepton ($\ell_3$) for events with three leptons. $N^{\text{Data},\ell_{\text{Anti-ID}}}_{\text{Anti-ID},c,b_1,b_2(,b_3)}$ and $N^{\text{MC},\ell_{\text{Anti-ID}}}_{\text{Anti-ID},c,b_1,b_2(,b_3)}$ are numbers of data and MC events in the corresponding anti-ID b-tagging CR (anti-ID region) with the anti-ID lepton $\ell_{\text{Anti-ID}}$. The number of MC events in the anti-ID region is simulated using both prompt and data-driven MC samples excluding events with at least one non-prompt lepton at the truth level. Due to the limitation of statistics in the b-tagging CRs, the non-prompt rates are determined in two lepton $p_{\text{T}}$ bins and listed in Table 7.9. The non-prompt background in other ID regions can then be estimated with the non-prompt rates and numbers of data and MC events in the corresponding anti-ID regions using Eq. 7.3.

| Lepton $p_{\text{T}}$ | < 27 GeV | > 27 GeV |
|---|---|---|
| $\epsilon_e$ | $0.0158 \pm 0.0027$ | $0.0196 \pm 0.0048$ |
| $\epsilon_\mu$ | $0.0143 \pm 0.0017$ | $0.0138 \pm 0.0011$ |

Table 7.9: Non-prompt rates in different lepton flavors and $p_{\text{T}}$ ranges. Only statistical uncertainties for the non-prompt rates are included.

### 7.3.3 Non-prompt Background Validation

The non-prompt rates are applied back to the $\ell^\pm \nu \ell^\pm \nu j j$ and $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ b-tagging CRs to check the modeling of the non-prompt background. The event yield tables are shown in Tables 7.10 and 7.11. The distributions of the $m_{jj}, E^{\text{miss}}_{\text{T}}$ significance, subleading lepton $p_{\text{T}}$, and subleading

lepton $\eta$ in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ $b$-tagging CR are shown in Figure 7.5. The distributions of the third lepton $p_{\mathrm{T}}$, third lepton $\eta$, $E_{\mathrm{T}}^{\mathrm{miss}}$ significance ratio, and $m_{\ell\ell}$ of the SF lepton pair in the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ $b$-tagging CR are shown in Figure 7.6. The distributions of those variables in each channel are attached in Appendices A.3 and A.4 for the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ $b$-tagging CR and the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ $b$-tagging CR respectively. Note that only statistical uncertainties of the MC samples and data are included. Although the total number of predicted events is matched to that in data by construction, the good agreement in different variables in $b$-tagging CRs suggests that the non-prompt background is modeled well.

| | $ee$ | $e\mu$ | $\mu\mu$ | Inclusive |
|---|---|---|---|---|
| $WWW$ | 6.17 ± 0.16 | 18.43 ± 0.29 | 12.16 ± 0.23 | 36.75 ± 0.41 |
| $WH$ | 3.69 ± 0.02 | 12.32 ± 0.04 | 8.23 ± 0.03 | 24.24 ± 0.06 |
| $WZ$ | 56.58 ± 0.90 | 141.2 ± 1.4 | 67.11 ± 0.98 | 264.9 ± 2.0 |
| NonPrompt | 98.7 ± 1.9 | 444.9 ± 2.7 | 333.2 ± 2.2 | 876.8 ± 3.9 |
| PhotonConv | 77.3 ± 6.3 | 83.2 ± 6.7 | - | 160.5 ± 9.2 |
| SS $WW$ | 16.05 ± 0.17 | 46.60 ± 0.28 | 29.69 ± 0.23 | 92.34 ± 0.40 |
| ChargeFlip | 151.36 ± 0.99 | 124.64 ± 0.80 | - | 276.0 ± 1.3 |
| $ZZ$ | 4.43 ± 0.16 | 8.60 ± 0.22 | 3.16 ± 0.13 | 16.20 ± 0.30 |
| $VVV$ | 0.10 ± 0.01 | 0.20 ± 0.02 | 0.07 ± 0.01 | 0.38 ± 0.02 |
| $t\bar{t}W$ | 45.20 ± 0.59 | 141.7 ± 1.1 | 91.22 ± 0.84 | 278.1 ± 1.5 |
| $t\bar{t}Z$ | 10.43 ± 0.26 | 28.27 ± 0.42 | 14.28 ± 0.29 | 52.98 ± 0.57 |
| $t\bar{t}H$ | 10.45 ± 0.13 | 32.62 ± 0.22 | 20.15 ± 0.17 | 63.21 ± 0.31 |
| $tZ$ | 5.13 ± 0.23 | 11.86 ± 0.35 | 5.71 ± 0.24 | 22.70 ± 0.48 |
| Total | 485.6 ± 6.7 | 1094.5 ± 7.5 | 585.0 ± 2.6 | 2165 ± 10 |
| Data | 515 | 1106 | 594 | 2215 |

Table 7.10: Event yields with statistical uncertainties for data and estimated SM processes in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ $b$-tagging CR.

| | $ee\mu$ | $\mu\mu e$ | Inclusive |
|---|---|---|---|
| $WWW$ | 0.90 ± 0.07 | 1.23 ± 0.08 | 2.13 ± 0.10 |
| $WH$ | 0.50 ± 0.01 | 0.69 ± 0.01 | 1.19 ± 0.01 |
| $WZ$ | 18.99 ± 0.60 | 25.77 ± 0.76 | 44.76 ± 0.97 |
| NonPrompt | 47.21 ± 0.88 | 74.4 ± 1.1 | 121.6 ± 1.4 |
| PhotonConv | 3.9 ± 1.8 | 12.4 ± 2.5 | 16.3 ± 3.1 |
| $ZZ$ | 2.60 ± 0.13 | 5.91 ± 0.21 | 8.51 ± 0.24 |
| $VVV$ | 0.11 ± 0.01 | 0.11 ± 0.01 | 0.22 ± 0.02 |
| $t\bar{t}W$ | 17.24 ± 0.36 | 22.91 ± 0.42 | 40.15 ± 0.55 |
| $t\bar{t}Z$ | 8.63 ± 0.22 | 10.73 ± 0.25 | 19.36 ± 0.33 |
| $t\bar{t}H$ | 7.00 ± 0.08 | 9.24 ± 0.09 | 16.24 ± 0.12 |
| $tZ$ | 0.68 ± 0.08 | 1.30 ± 0.11 | 1.98 ± 0.14 |
| Total | 107.8 ± 2.2 | 164.7 ± 2.8 | 272.4 ± 3.6 |
| Data | 109 | 143 | 252 |

Table 7.11: Event yields with statistical uncertainties for data and estimated SM processes in the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ $b$-tagging CR.
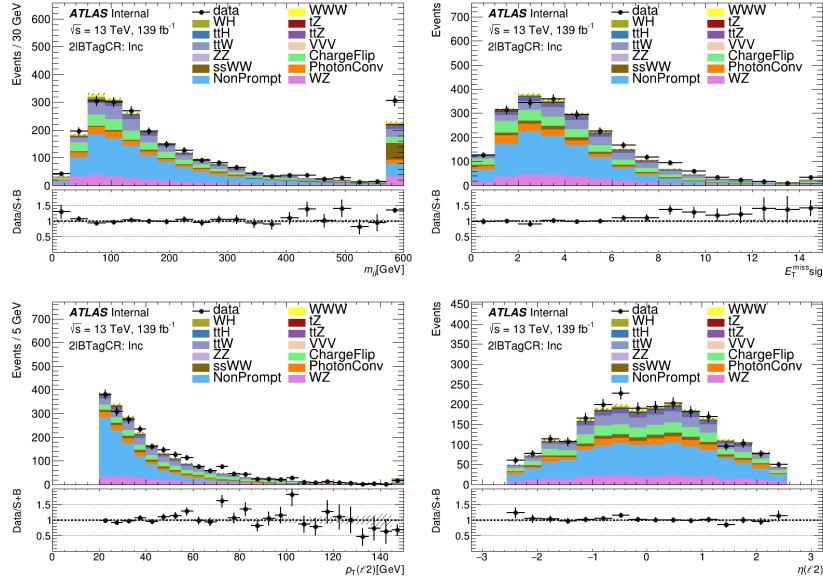
Figure 7.5: The distributions of the $m_{jj}$ (top left), $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (top right), subleading lepton $p_{\mathrm{T}}$ (bottom left), and subleading lepton $\eta$ (bottom right) with statistical uncertainties in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ $b$-tagging CR.
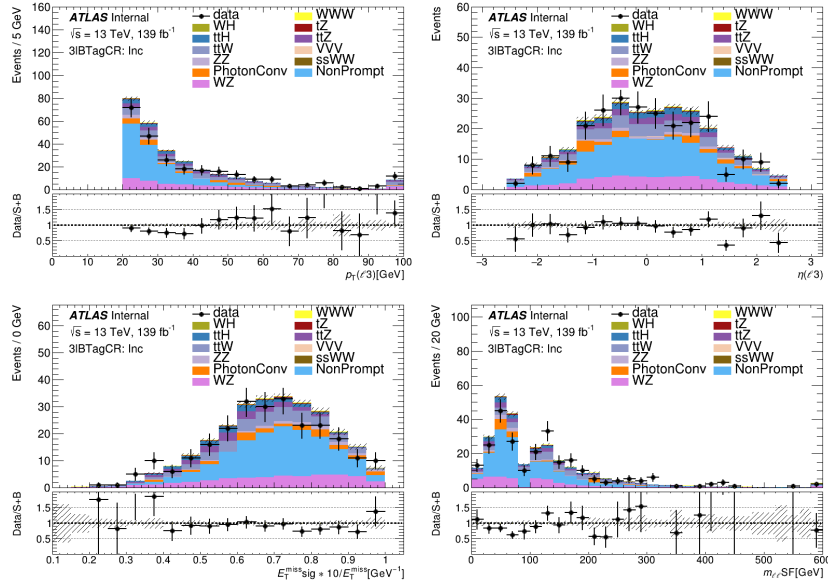


Figure 7.6: The distributions of the third lepton $p_{\mathrm{T}}$ (top left), third lepton $\eta$ (top right), $E_{\mathrm{T}}^{\mathrm{miss}}$ significance ratio (bottom left), and $m_{\ell\ell}$ (bottom right) of the SF lepton pair with statistical uncertainties in the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ $b$-tagging CR.

69

## 7.4  Photon Conversion Background

### 7.4.1  Photon Conversion Rate Calculation

Photon conversion background is defined as processes that have a photon misreconstructed as an electron. It mainly comes from $W\gamma$ and $Z\gamma$ processes. $W\gamma$+jets can enter the $\ell^\pm\nu\ell^\pm\nu jj$ SR with a leptonically-decayed $W$ boson and a misreconstructed electron from the photon if these two leptons are SS. In the $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ SR, the $Z$ boson from the $Z\gamma$ process decays into two taus and one tau decays into an electron while the other one decays into a muon. There is no SFOS lepton pair if the electron from $Z$ decay and the misreconstructed electron from the photon are SS.

The photon conversion background is studied in a dedicated $Z\gamma$ CR and the selection cuts after the preselection are listed in Table 7.2. It has the same cuts as the $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ SR except for the lepton flavor, lepton charge, and $m_{\ell\ell\ell}$ cuts. Only $eee$ and $\mu\mu e$ channels are considered and a SFOS lepton pair is required to reconstruct $Z\gamma$ process since $Z$ decays to a SFOS lepton pair while the other lepton must be an electron from a misreconstructed photon. The trilepton invariant mass $m_{\ell\ell\ell}$ is required to be inside the $Z$ mass window between 80 GeV and 100 GeV to further improve the purity of the $Z\gamma$ process. Additionally, similar to the anti-ID regions, the anti-BL $Z\gamma$ CR are defined with events having one anti-BL electron and passing the same event-level cuts as the $Z\gamma$ CR. In this thesis, anti-BL regions are referring to those regions with only one anti-BL electron while the remaining leptons are ID leptons.

The photon conversion background is estimated with a data-driven method similar to the one used for the non-prompt background estimation. Instead of anti-ID regions, a map between anti-BL regions and corresponding ID regions is constructed with a photon conversion rate. The photon conversion rate is calculated only with one bin in the $\mu\mu e$ channel of the $Z\gamma$ CR by minimizing the negative log-likelihood function. Since there is only one bin, one channel and one photon conversion rate, Eq. 7.1 can be simplified as:

$$\mathcal{L}(\epsilon_\gamma) = \mathrm{Pois}(N_{\mathrm{ID}}^{\mathrm{Data}}|N_{\mathrm{ID}}^{\mathrm{Exp}}) = \frac{(N_{\mathrm{ID}}^{\mathrm{Exp}})^{N_{\mathrm{ID}}^{\mathrm{Data}}}e^{-N_{\mathrm{ID}}^{\mathrm{Exp}}}}{N_{\mathrm{ID}}^{\mathrm{Data}}!}, \tag{7.4}$$

where $\epsilon_\gamma$ is the photon conversion rate. The number of expected events $N_{\mathrm{ID}}^{\mathrm{Exp}}$ in the $Z\gamma$ CR (ID region) can still be express as Eq. 7.2. Similar to Eq. 7.3, the data-driven estimation of the photon conversion background can be written as:

$$N_{\mathrm{ID}}^{\mathrm{PhotonConv}} = \epsilon_\gamma \times \left(N_{\mathrm{Anti\text{-}BL}}^{\mathrm{Data}} - N_{\mathrm{Anti\text{-}BL}}^{\mathrm{MC}}\right), \tag{7.5}$$

where $N_{\mathrm{Anti\text{-}BL}}^{\mathrm{Data}}$ and $N_{\mathrm{Anti\text{-}BL}}^{\mathrm{MC}}$ are the numbers of the data and MC events in the anti-BL $Z\gamma$ CR

(anti-BL region), respectively. The number of MC events in the anti-BL $Z\gamma$ CR is estimated using both prompt and data-driven MC samples excluding events with at least one electron from photon conversion in truth. The inclusive photon conversion rate is then determined with only statistical uncertainty as:

$$\epsilon_\gamma = 0.42 \pm 0.03. \tag{7.6}$$

The photon conversion background in other ID regions can then be estimated with the photon conversion rate and numbers of the data and MC events in the corresponding anti-BL regions using Eq. 7.5.

## 7.4.2 Photon Conversion Background Validation

The photon conversion background modeling is checked in both *eee* and *μμe* channels of the $Z\gamma$ CR. The event yield table is shown in Table 7.12 and the plots of $m_{\ell\ell\ell}$, smallest electron $p_T$ and $E_T^{\text{miss}}$ significance are shown in Figure 7.7. Note that only statistical uncertainties of the MC predictions and data are included for tables and plots. Although the total number of predicted events in *μμe* channel of the $Z\gamma$ CR is matched to that in data by construction, the event yield in *eee* channel as well as the distributions of different variables also show good agreement between data and prediction within uncertainties, which indicates that the modeling of the photon conversion background is reasonable.

|  | *eee* | *μμe* |
|---|---|---|
| *WWW* | $0.11 \pm 0.02$ | $0.37 \pm 0.04$ |
| *WH* | $0.27 \pm 0.01$ | $1.06 \pm 0.01$ |
| *WZ* | $12.00 \pm 0.66$ | $26.95 \pm 0.91$ |
| NonPrompt | $2.46 \pm 0.32$ | $5.35 \pm 0.54$ |
| PhotonConv | $92.7 \pm 6.3$ | $207.0 \pm 9.3$ |
| *ZZ* | $22.06 \pm 0.41$ | $40.67 \pm 0.54$ |
| *VVV* | $0.02 \pm 0.01$ | $0.03 \pm 0.01$ |
| $t\bar{t}W$ | $0.01 \pm 0.01$ | $0.15 \pm 0.03$ |
| $t\bar{t}Z$ | $0.09 \pm 0.02$ | $0.15 \pm 0.03$ |
| $t\bar{t}H$ | $0.02 \pm 0.01$ | $0.09 \pm 0.01$ |
| *tZ* | $0.03 \pm 0.02$ | $0.04 \pm 0.02$ |
| Total | $129.8 \pm 6.4$ | $281.8 \pm 9.4$ |
| Data | 150 | 282 |

Table 7.12: Event yields with statistical uncertainties for data and estimated SM processes in the $Z\gamma$ CR.
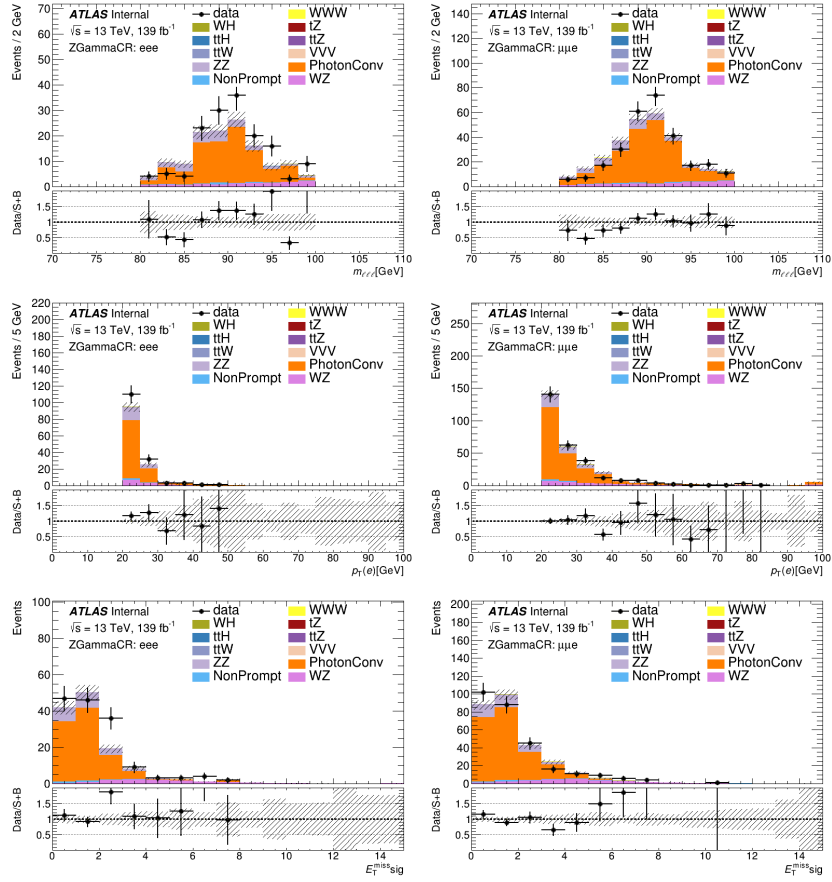
Figure 7.7: The $m_{\ell\ell\ell}$ (top), smallest electron $p_{\mathrm{T}}$ (middle) and $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (bottom) distributions with statistical uncertainties in the *eee* (left) and *μμe* (right) channels in the $Z\gamma$ CR.

## 7.5 Charge-flip Background

### 7.5.1 Charge-flip Rate Calculation

Charge-flip background is referring to those processes with the misidentified electric charge. Photons radiated from electrons decay into electron-positron pairs and the electron with the opposite charge from the original electron is picked up by the detector. The possibility to have one charge-flipped electron is rather small as the charge-flip rate shown later is in the range of $10^{-5}$ to $10^{-2}$. Therefore this background is only considered in the SS regions (regions with two SS leptons) where OS events are flipped into SS events but not in the OS regions (regions with two OS leptons) since cross sections of SS processes are much smaller than those for OS processes. Additionally, the number of charge-flipped muons is even small so it is ignored in this analysis. The dominate sources of the charge-flip background come from $Z+$jets, $t\bar{t}$, and OS $WW$ processes. This is a minor background in this analysis.

The charge-flip background is studied with $Z \to ee$ events and a dedicated charge-flip CR is defined using the cuts listed in Table 7.1 and the preselection cuts. Due to the low possibility of having one charge-flipped electron, the cuts of the charge-flip CR are very loose using only the same lepton cuts as the $\ell^\pm \nu \ell^\pm \nu jj$ SR except for the lepton flavor, lepton charge, and $m_{jj}$ cuts, and all cuts applied on jets are dropped. It only uses the $ee$ channel and includes both SS and OS events with $m_{\ell\ell}$ within the $Z$ mass window between 60 GeV and 120 GeV to select $Z \to ee$ events.

A fully data-driven method without MC samples is used for the charge-flip rate estimation. To further select $Z \to ee$ events, two regions, the sideband charge-flip CR and the central charge-flip CR, are defined as events with $m_{ee}$ outside of a tighter $Z$ mass window ($m_{\ell\ell} \leq 75$ GeV or $m_{\ell\ell} \geq 105$ GeV) and inside of the tighter $Z$ mass window (75 GeV $< m_{\ell\ell} < 105$ GeV) in addition to other cuts applied in the charge-flip CR. The number of non $Z \to ee$ events in the central charge-flip CR is estimated using the number of data in the sideband CR so that number of the $Z \to ee$ events is the difference in data between the central and sideband charge-flip CRs.

Similar to the non-prompt and photon conversion background estimations, the basic idea of the data-driven method to estimate the charge-flip background is to obtain a map between the SS events and the OS events with the charge-flip rate. The charge-flip rate is measured in each bin of the lepton $p_T$ and $|\eta|$ by minimizing the negative log-likelihood function. The charge-flip rate is defined as the possibility of an electron in a specific $p_T$ and $|\eta|$ range with its electric charge flipped. Similar to Eqs. 7.1 and 7.4, assuming the two electrons in the event are in bin $b_1$ and $b_2$

based on $p_T$ and $|\eta|$, the likelihood function can be expressed as:

$$\mathcal{L}(\epsilon_q) = \prod_{b_1 \in (p_T(\ell_1), |\eta(\ell_1)|)} \prod_{b_2 \in (p_T(\ell_2), |\eta(\ell_2)|)} \mathrm{Pois}(N_{\mathrm{SS},b_1,b_2}^{\mathrm{Data}} | N_{\mathrm{SS},b_1,b_2}^{\mathrm{Exp}})$$

$$= \prod_{b_1,b_2} \frac{(N_{\mathrm{SS},b_1,b_2}^{\mathrm{Exp}})^{N_{\mathrm{SS},b_1,b_2}^{\mathrm{Data}}} e^{-N_{\mathrm{SS},b_1,b_2}^{\mathrm{Exp}}}}{N_{\mathrm{SS},b_1,b_2}^{\mathrm{Data}}!}, \tag{7.7}$$

where $\epsilon_q$ is the charge-flip rate and $N_{\mathrm{SS},b_1,b_2}^{\mathrm{Data}}$ and $N_{\mathrm{SS},b_1,b_2}^{\mathrm{Exp}}$ are the measured and expected number of the $Z \to ee$ events with reconstructed SS electron pairs in the charge-flip CR, respectively. The likelihood function is a product looping all bins of two leptons. Based on the charge-flip rate definition, $N_{\mathrm{SS},b_1,b_2}^{\mathrm{Exp}}$ can be expressed as:

$$N_{\mathrm{SS},b_1,b_2}^{\mathrm{Exp}} = \left(N_{\mathrm{SS},b_1,b_2}^{\mathrm{Data}} + N_{\mathrm{OS},b_1,b_2}^{\mathrm{Data}}\right)\left[\epsilon_q^{b_1}(1 - \epsilon_q^{b_2}) + (1 - \epsilon_q^{b_1})\epsilon_q^{b_2}\right], \tag{7.8}$$

where $\epsilon_q^b$ is the charge-flip rate in a bin $b$ and $N_{\mathrm{OS},b_1,b_2}^{\mathrm{Data}}$ is the measured number of $Z \to ee$ events with the reconstructed OS lepton pair in the charge-flip CR. The measured charge-flip rates as functions of lepton $p_T$ and $\eta$ are shown in Figure 7.8.
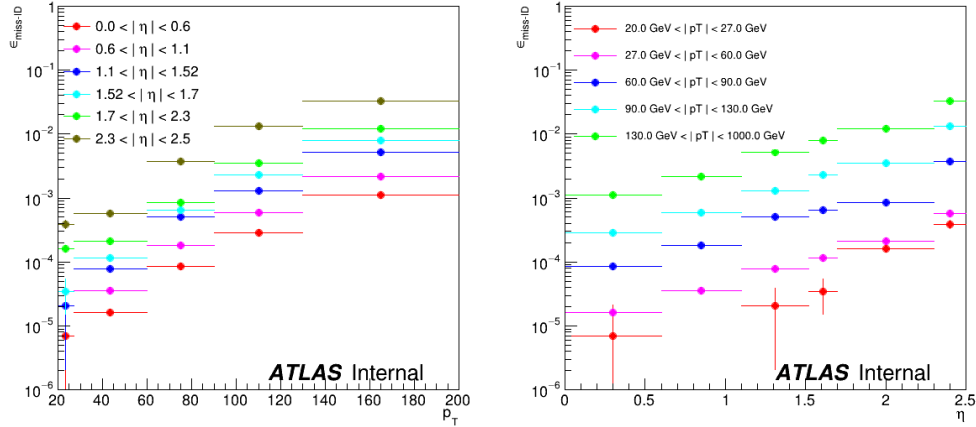


Figure 7.8: The measured charge-flip rates as a function of electron $p_T$ (left) and $|\eta|$ (right) with statistical uncertainties. There is a missing point for the charge-flip rate at $p_T$ between 20 GeV and 27 GeV and $|\eta|$ between 0.6 and 1.1, which is due to lack of statistics.

Since $N_{\mathrm{SS},b_1,b_2}^{\mathrm{Data}}$ is considered as unknown in other SS regions, by substituting $N_{\mathrm{SS},b_1,b_2}^{\mathrm{Data}}$ with

$N_{\text{SS},b_1,b_2}^{\text{Exp}}$, Eq. 7.8 can be rearranged as:

$$N_{\text{SS},b_1,b_2}^{\text{Exp}} = N_{\text{OS},b_1,b_2}^{\text{Data}} \times \left[ \frac{1}{1 - \epsilon_q^{b_1} - \epsilon_q^{b_2} + 2\epsilon_q^{b_1}\epsilon_q^{b_2}} - 1 \right],$$
(7.9)

which can be used to estimate the charge-flip background in other SS regions. The charge-flip rate for muon is set to zero as the charge-flip from muon is negligible and thus ignored in this analysis. For regions with three leptons, the charge-flip background is only considered in the regions with no SFOS for the SS electron pair in the events.

## 7.5.2 Charge-flip Background Validation

The modeling of the charge-flip background is checked in a dedicated $Z$ window VR with the preselection cuts and cuts listed in Table 7.1. It has the same cuts as the $\ell^\pm\nu\ell^\pm\nu jj$ SR except for lepton flavor, $m_{\ell\ell}$ and the $E_{\text{T}}^{\text{miss}}$ significance. To reconstruct the $Z \to ee$ process, only the $ee$ channel is selected and $m_{\ell\ell}$ needs to be inside the $Z$ mass window between 80 GeV and 100 GeV. The $E_{\text{T}}^{\text{miss}}$ significance cut in $ee$ channel is dropped to increase the statistics in this VR.

The event yield table is shown in Table 7.13 and the distributions of the $m_{\ell\ell}$, $m_{jj}$, $E_{\text{T}}^{\text{miss}}$ significance, leading forward jet $p_{\text{T}}$, subleading lepton $p_{\text{T}}$, and subleading lepton $\eta$ are shown in Figure 7.9. Note that only statistical uncertainties of the MC samples and data are included. The consistency between data and prediction within uncertainties in this VR indicates that the charge-flip background modeling is reasonable.

|  | $ee$ |
|---|---|
| $WWW$ | 2.11 ± 0.10 |
| $WH$ | 2.25 ± 0.02 |
| $WZ$ | 23.44 ± 0.67 |
| NonPrompt | 4.1 ± 7.6 |
| PhotonConv | 30 ± 14 |
| SS $WW$ | 2.43 ± 0.06 |
| ChargeFlip | 799.1 ± 1.9 |
| $ZZ$ | 1.86 ± 0.11 |
| $VVV$ | 0.02 ± 0.01 |
| $t\bar{t}W$ | 0.83 ± 0.08 |
| $t\bar{t}Z$ | 0.13 ± 0.03 |
| $t\bar{t}H$ | 0.20 ± 0.02 |
| $tZ$ | 0.13 ± 0.04 |
| Total | 867 ± 16 |
| Data | 858 |

Table 7.13: Event yields with statistical uncertainties for data and estimated SM processes in the $Z$ window CR.
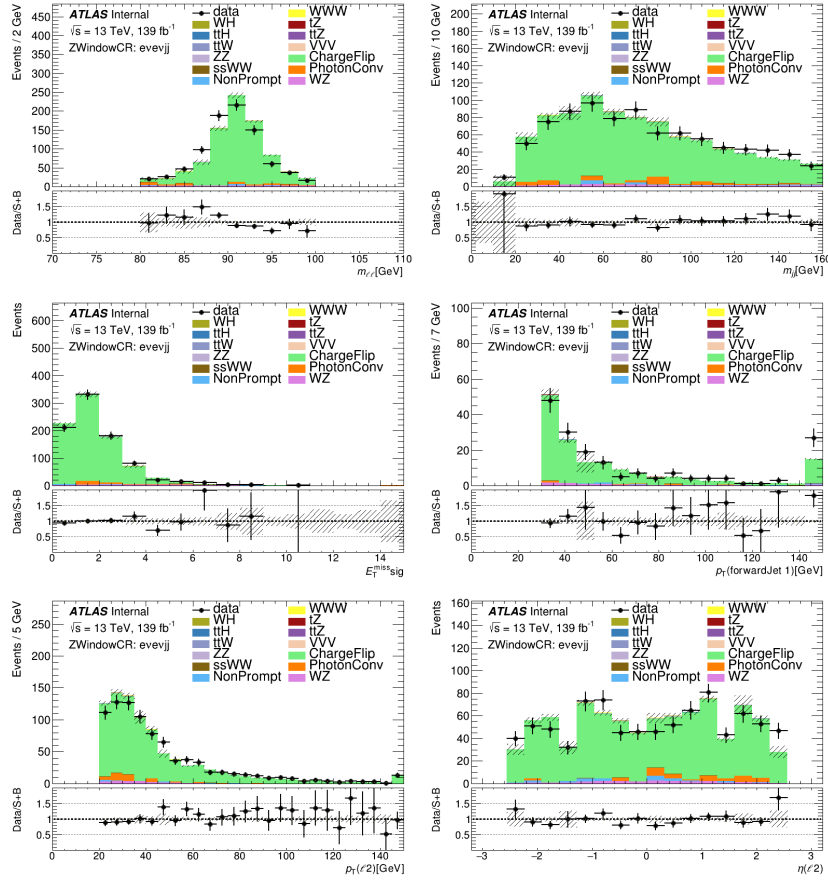
Figure 7.9: The $m_{\ell\ell}$ (top left), $m_{jj}$ (top right), $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (middle left), leading forward jet $p_{\mathrm{T}}$ (middle right), subleading lepton $p_{\mathrm{T}}$ (bottom left), and subleading lepton $\eta$ (bottom right) distributions with statistical uncertainties in the $Z$ window VR.

## 7.6 *W* Sideband Control Region

To check the overall modeling of all background processes, the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR is partially unblinded in the *W* sideband CR that is not very sensitive to the signal contribution and the selection cuts after the preselection are listed in Table 7.1. This CR has exactly the same cuts as the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR except that the $m_{jj}$ is required to be not only less than 160 GeV but also outside the *W* mass window between 50 GeV and 120 GeV since the two jets in the signal events are from a hadronically-decayed *W* boson and thus those events are more likely to have the $m_{jj}$ within the *W* mass window.

The event yield in this CR is shown in Table 7.14 and the distributions of the $m_{jj}$, $m_{\ell\ell}$, $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, leading forward jet $p_{\mathrm{T}}$, subleading lepton $p_{\mathrm{T}}$, and subleading lepton $\eta$ are shown in Figure 7.10. The distributions of those variables in each channel are attached in Appendix A.5. Note that only statistical uncertainties of the MC samples and data are included. The data and prediction agree within uncertainties, which indicates reasonable modelings of the background processes in this analysis.

|  | *ee* | *eμ* | *μμ* | Inclusive |
|---|---|---|---|---|
| *WWW* | 2.43 ± 0.11 | 9.73 ± 0.22 | 6.23 ± 0.18 | 18.39 ± 0.30 |
| *WH* | 2.84 ± 0.02 | 12.93 ± 0.04 | 8.58 ± 0.03 | 24.36 ± 0.06 |
| *WZ* | 33.89 ± 0.86 | 143.7 ± 1.8 | 74.3 ± 1.2 | 251.9 ± 2.4 |
| NonPrompt | 4.2 ± 1.0 | 45.81 ± 0.95 | 32.33 ± 0.68 | 82.3 ± 1.6 |
| PhotonConv | 28.4 ± 5.8 | 55.7 ± 6.8 | - | 84.1 ± 9.0 |
| SS *WW* | 5.08 ± 0.09 | 21.37 ± 0.18 | 14.16 ± 0.15 | 40.60 ± 0.25 |
| ChargeFlip | 13.11 ± 0.30 | 7.99 ± 0.23 | - | 21.10 ± 0.38 |
| ZZ | 1.10 ± 0.09 | 8.32 ± 0.23 | 3.76 ± 0.16 | 13.18 ± 0.29 |
| *VVV* | 0.05 ± 0.01 | 0.16 ± 0.01 | 0.06 ± 0.01 | 0.27 ± 0.02 |
| $t\bar{t}W$ | 1.30 ± 0.10 | 5.18 ± 0.21 | 2.99 ± 0.16 | 9.46 ± 0.28 |
| $t\bar{t}Z$ | 0.22 ± 0.04 | 0.98 ± 0.07 | 0.43 ± 0.05 | 1.63 ± 0.10 |
| $t\bar{t}H$ | 0.22 ± 0.02 | 0.92 ± 0.04 | 0.57 ± 0.03 | 1.72 ± 0.05 |
| *tZ* | 0.14 ± 0.04 | 0.55 ± 0.08 | 0.19 ± 0.04 | 0.88 ± 0.10 |
| Total | 93.0 ± 6.0 | 313.3 ± 7.1 | 143.6 ± 1.5 | 549.9 ± 9.4 |
| Data | 93 | 357 | 153 | 603 |

Table 7.14: Event yields with statistical uncertainties for data and estimated SM processes in the *W* sideband CR.
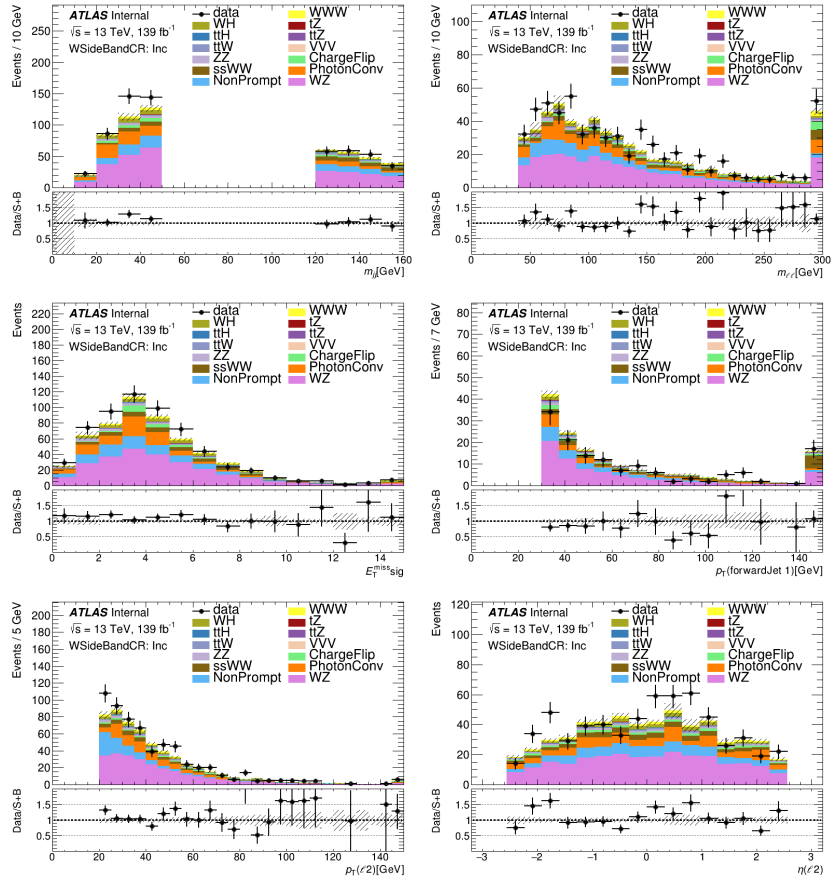
Figure 7.10: The distributions of the $m_{jj}$ (top left), $m_{\ell\ell}$ (top right), $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (middle left), leading forward jet $p_{\mathrm{T}}$ (middle right), subleading lepton $p_{\mathrm{T}}$ (bottom left) and subleading lepton $\eta$ (bottom right) with statistical uncertainties in the $W$ sideband CR.

# CHAPTER 8

# Boosted Decision Tree

## 8.1 Definition
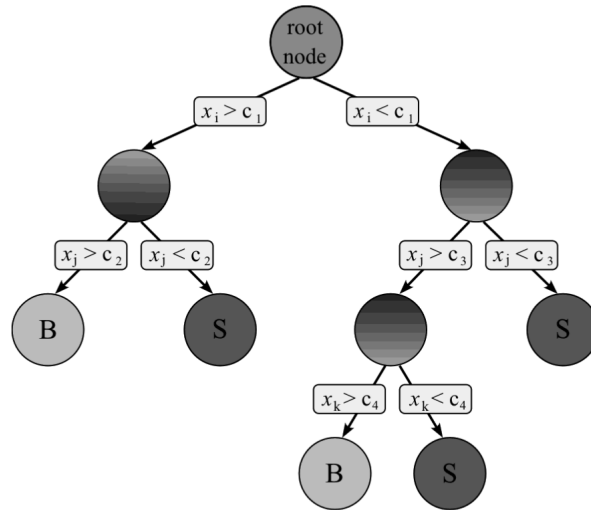


Figure 8.1: The illustration of an individual CART. B stands for the background node and S stands for the signal node [114].

To further separate signal and background in the SRs, the BDT model is implemented using XGBoost (Extreme Gradient Boosting) [115] in this analysis and the distributions of the output discriminant scores are used in the fit described in Section 10.3 to extract the $W^{\pm}W^{\pm}W^{\mp}$ signal. The BDT model consists of an ensemble of Classification and Regression Trees (CARTs). As is shown in Figure 8.1, the CART is a two-dimensional structure that starting from a single root node of all events and followed by a set of binary splits based on the comparisons between physical variables of events (called features) and corresponding thresholds. Parent nodes are the nodes before splitting and child nodes are the nodes after splitting. The nodes without any child nodes

are called leaf nodes. Depth of each node is defined as the number of splits from the root node, and the maximum depth is the maximum depth of all nodes in a tree. For example, the maximum depth of the tree in Figure 8.1 is 3. In the training dataset, signal events are assigned with a discriminant score of 1 and background events are assigned with a score of 0. The discriminant score for each leaf node is defined as the average of the scores of all events in that node in the training dataset. It can also be interpreted as purity and expressed as score $= \frac{N_s}{N_s+N_b}$ with $N_{s(b)}$ as number of signal (background) events in the training dataset. The output of the CART for each event during testing is the discriminant score of the corresponding leaf node of the event. The leaf nodes with scores greater than 0.5 are labeled as signal nodes while the remaining leaf nodes are background nodes. Events in the signal and background nodes are predicted to be signal and background respectively based on the CART model.

The input feature and corresponding threshold of a split in the CART are determined from searching through all possible combinations to minimize a cost function defined as:

$$C = \frac{N_{\text{Left}}}{N}G_{\text{Left}} + \frac{N_{\text{Right}}}{N}G_{\text{Right}}, \tag{8.1}$$

where $N$ is the number of events in the parent node, $N_{\text{Left(Right)}}$ is the number of events in the left (right) child node under a split, and $G_{\text{Left(Right)}}$ is the Gini impurity in the left (right) child node and defined as:

$$G_i = 1 - \sum_{k=\{s,b\}} \left(\frac{N_{i,k}}{N_i}\right)^2, \tag{8.2}$$

where $N_i$ is number of events in the node $i$ (left or right child node) and $N_{i,k}$ is the number of events labeled as $k$ (signal $s$ or background $b$) in the node $i$.

Since an individual CART is not strong enough to predict the classification of signal and background, the ensemble of CARTs is used in the BDT and gradient boosting [116, 117] is used in this analysis as the ensemble method to combine multiple CARTs. The gradient boosting adds CARTs to an ensemble sequentially and each CART corrects the BDT model by fitting to the residual error of the previous ensemble model. The BDT model can be expressed as:

$$\hat{y}_i = \sum_{t=1}^{T} f_t(x_i), \tag{8.3}$$

where $x_i$ is a set of the input features for event $i$, $\hat{y}_i$ is the output discriminant score for the event based on the BDT model, $T$ is the total number of CARTs, and $f_t$ stands for an individual CART $t$ and gives the discriminant score for the event based on the CART. In the gradient boosting, each CART $f_t(x_i)$ is trained with the output of $y_i - \hat{y}_i^{(t-1)}$, where $y_i$ is the truth discriminant score of the event in the training dataset (0 for background events and 1 for signal events) and $\hat{y}_i^{(t-1)}$ is the

output discriminant score of the event based on the previous ensemble model with $t-1$ CARTs. Additionally, a regularization technique to slow down the learning in the gradient boosting called shrinkage is used by scaling the contribution of each CART with a small learning rate between 0 and 1. The CARTs can be generated better with shrinkage.

The cost function that the BDT model is trying to minimize can be written as:

$$\mathcal{C} = \sum_{i=1}^{N} l(y_i, \hat{y}_i) + \sum_{t=1}^{T} \Omega(f_t), \tag{8.4}$$

where $l$ is the loss function that measures the difference between $y_i$ and $\hat{y}_i$ and $\Omega$ is the regularization term that measures the model complexity of each CART $f_t$. $\Omega(f)$ in the XGBoost can be expressed as:

$$\Omega(f) = \gamma L + \frac{1}{2} \lambda \sum_{j=1}^{L} w_j^2, \tag{8.5}$$

where $L$ is the total number of leaf nodes in the CART $f$ and $W_j$ is the score of the leaf node $j$. $\gamma$ and $\lambda$ are constants and set to 0 and 1 respectively in this analysis as the L2 regularization is used.

## 8.2   Construction

In this analysis, two BDT models, $\ell^{\pm}\nu\ell^{\pm}\nu jj$ BDT and $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ BDT are trained separately in the SRs with the signal MC samples as well as the background MC samples and data-driven estimations. Some events have negative event weights so that they can not be directly used in the BDT training. To solve this problem, the absolute values of these weights are used for all events together with the ratio between the sum of original weights and the sum of absolute weights so that the total number of events does not change. The new event weight can be expressed as:

$$w_i^{\text{new}} = |w_i| \times \frac{\sum_j w_j}{\sum_k |w_k|}, \tag{8.6}$$

where $w_i$ stands for the original event weight for an event $i$.

In this analysis, Receiver Operating Characteristic (ROC) curve is used to evaluate the performance of each BDT model, which plots true positive rate versus false positive rate. The true positive rate is the ratio of positive (signal) events that are classified correctly by the model among all truth positive (signal) events and can be expressed as:

$$\text{True Positive Rate} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}. \tag{8.7}$$

The false positive rate is defined as the ratio of negative (background) events that are classified as positive (signal) by the model among all truth negative (background) events and can be expressed as:

$$\text{False Positive Rate} = \frac{\text{False Positive}}{\text{False Positive} + \text{True Negative}}. \tag{8.8}$$

The true positive rate and the false positive rate are calculated using different cut values on the output discriminant score from the BDT model so that a ROC curve is generated. The model performance is quantified by the Area Under the Curve (AUC) in the plot of the ROC curve. A random classifier has an AUC of 0.5 while a perfect classifier has an AUC of 1.

Only relevant input features to separate signal and background are selected for the BDT models. The feature selection is performed based on the feature importance determined from the BDT models. The feature importance is defined as the average of improvement in the Gini impurity for all nodes using that feature, weighted by the number of events in the training dataset associated with those nodes. Starting with a large set of features, each BDT model is trained with all background and signal samples in the corresponding SR. In each iteration, the feature with the lowest feature importance is dropped and then the BDT model is retrained until there is a significant drop in the performance of the BDT.

The lists of input features used in the $\ell^\pm \nu \ell^\pm \nu jj$ BDT and $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ BDT are shown in Table 8.1 with their importances. The $\ell^\pm \nu \ell^\pm \nu jj$ BDT uses the absolute difference between the mass of two leading jets and the $W$ boson pole mass, the $p_\text{T}$ of the leading forward jet, the $E_\text{T}^\text{miss}$ significance, the energy of the leading forward jet, the invariant mass of the subleading lepton and the leading jet, $\Delta R$ between the subleading lepton and the leading jet, the number of leptons that are overlapped with jets and thus removed by the overlap removal, $\Delta R$ between the leading lepton and the leading jet, the minimum invariant mass of a lepton and a jet, $\Delta R$ between the leading lepton and the subleading jet, the $p_\text{T}$ of the subleading jet, the $p_\text{T}$ of the subleading lepton, the absolute value of the leading lepton $\eta$, the number of jets, $\Delta R$ between the subleading lepton and the subleading jet, the total $p_\text{T}$ of all leptons and all jets, the absolute value of $\Delta\eta$ between two leptons, the absolute value of $\Delta\eta$ between two leading jets, the total invariant mass of two leptons, and the total invariant mass of the leading lepton and the leading jet.

The $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ BDT uses the $E_\text{T}^\text{miss}$ significance ratio, $\Delta R$ between two leading jets (if the number of jets is less than 2, $\Delta R_{jj} = -9,999$), the $p_\text{T}$ of the third lepton, the transverse mass $(m_T = \sqrt{m^2 + p_x^2 + p_y^2})$ of the $E_\text{T}^\text{miss}$ and all three leptons, the total invariant mass of the SF lepton pair, the $p_\text{T}$ of the subleading lepton, the transverse mass of the $E_\text{T}^\text{miss}$ and the subleading lepton, the minimum $\Delta R$ between two leptons, $\Delta\phi$ between all leptons and the $E_\text{T}^\text{miss}$, the $E_\text{T}^\text{miss}$ significance, the total mass of the subleading and the third leptons, and the energy of the subleading lepton. A good agreement between data and prediction is achieved for all features used in the BDTs in related

CRs and VRs. Distributions of some top features are shown already in Chapter 7 and Appendix A.

| $\ell^\pm\nu\ell^\pm\nu jj$ BDT | | $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ BDT | |
|---|---|---|---|
| Feature | Importance | Feature | Importance |
| $\|m_{jj}-m_W\|$ | 0.152 | $E_{\mathrm{T}}^{\mathrm{miss}}$ sig$\times 10/E_{\mathrm{T}}^{\mathrm{miss}}$ | 0.167 |
| $p_{\mathrm{T}}(j_{\mathrm{forward}})$ | 0.096 | $\Delta R_{jj}$ | 0.124 |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ sig | 0.065 | $p_{\mathrm{T}}(\ell_3)$ | 0.102 |
| $E(j_{\mathrm{forward}})$ | 0.064 | $m_{\mathrm{T}}(\ell_{\mathrm{all}},E_{\mathrm{T}}^{\mathrm{miss}})$ | 0.092 |
| $m_{\ell_2 j_1}$ | 0.061 | $m_{\ell\ell}(\mathrm{SF})$ | 0.08 |
| $\Delta R_{\ell_2 j_1}$ | 0.057 | $p_{\mathrm{T}}(\ell_2)$ | 0.077 |
| $N$(leptons in jets) | 0.049 | $m_{\mathrm{T}}(\ell_2,E_{\mathrm{T}}^{\mathrm{miss}})$ | 0.076 |
| $\Delta R_{\ell_1 j_1}$ | 0.047 | $\min(\Delta R_{\ell\ell})$ | 0.075 |
| $\min(m_{\ell j})$ | 0.044 | $\Delta\phi(\ell_{\mathrm{all}},E_{\mathrm{T}}^{\mathrm{miss}})$ | 0.071 |
| $\Delta R_{\ell_1 j_2}$ | 0.043 | $E_{\mathrm{T}}^{\mathrm{miss}}$ sig | 0.053 |
| $p_{\mathrm{T}}(j_2)$ | 0.041 | $m_{\ell_2\ell_3}$ | 0.052 |
| $p_{\mathrm{T}}(\ell_2)$ | 0.039 | $E(\ell_2)$ | 0.029 |
| $\|\eta(\ell_1)\|$ | 0.035 | | |
| $N_{\mathrm{jet}}$ | 0.035 | | |
| $\Delta R_{\ell_2 j_2}$ | 0.035 | | |
| $p_{\mathrm{T}}(\ell_{\mathrm{all}},j_{\mathrm{all}})$ | 0.031 | | |
| $\|\Delta\eta_{\ell\ell}\|$ | 0.029 | | |
| $\|\Delta\eta_{jj}\|$ | 0.028 | | |
| $m_{\ell\ell}$ | 0.026 | | |
| $m_{\ell_1 j_1}$ | 0.024 | | |

Table 8.1: Input features and their importances in the $\ell^\pm\nu\ell^\pm\nu jj$ BDT and $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ BDT.

Hyperparameters are parameters that control the learning process of an algorithm. They need to be set before training and remain constant during the training. After the feature selection with some default hyperparameters, the hyperparameter tuning needs to be performed with a trade-off between bias and variance as an over-simplified model can miss the relations between features and outputs called under-fitting while an over-complicated model can be very sensitive to capture small variations in the training data called over-fitting. In the BDT mode of this analysis, hyperparameters include maximum tree depth, number of trees (CARTs), learning rate, event ratio, and feature ratio. In order to avoid over-fitting, only random selected subsets of events and input features are used in the training for each CART. The event ratio and feature ratio are defined as the ratios of those subsets in all events and features. The remaining hyperparameters are defined already in Section 8.1.

In this analysis, $k$-fold cross-validation is used to determine the hyperparameters and balance the bias and variance. All training samples are split into $k$ subsets firstly. The BDT model is
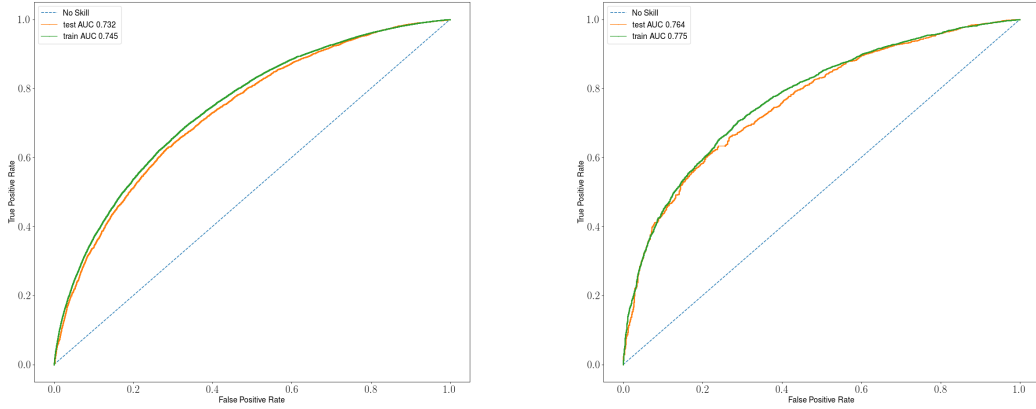
Figure 8.2: The ROC curves for the $\ell^\pm \nu \ell^\pm \nu jj$ BDT (left) and the $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ BDT (right) with the best performance hyperparameters for the first fold in the $k$-fold cross validation. The curves of the remaining folds are similar and thus omitted.

then trained with $k - 1$ subsets and tested with the remaining one subset, which is called one fold. The step of training and testing is repeated $k$ times with different choices of $k - 1$ training subsets and the test subset is always the remaining one. The performance of the BDT model is determined by averaging the performance in each step of training and testing. In the end, the set of hyperparameters with the best performance is chosen. Due to the limit statistics in the $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SR, a 4-fold cross-validation is used for the $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SR while a 5-fold cross-validation is used for the $\ell^\pm \nu \ell^\pm \nu jj$ SR. The values of hyperparameters with the best performance are summarized in Table 8.2. The corresponding ROC curves are shown in Figure 8.2 and based the plots, a good balancing between bias and variance is achieved without over-fitting or under-fitting. Then the BDT models are trained with all background and signal samples and applied to the data samples. The distributions of the output discriminant scores based on the $\ell^\pm \nu \ell^\pm \nu jj$ BDT and $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ BDT in the corresponding $\ell^\pm \nu \ell^\pm \nu jj$ and $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SR are shown in Figure 8.3.

| Hyperparameter | $\ell^\pm \nu \ell^\pm \nu jj$ BDT | $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ BDT |
|---|---|---|
| Max tree depth | 4 | 4 |
| Number of trees | 100 | 130 |
| Learning rate | 0.05 | 0.05 |
| Event ratio | 0.5 | 0.5 |
| Feature ratio | 0.5 | 0.5 |

Table 8.2: Input features and their importances in the $\ell^\pm \nu \ell^\pm \nu jj$ BDT and $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ BDT.
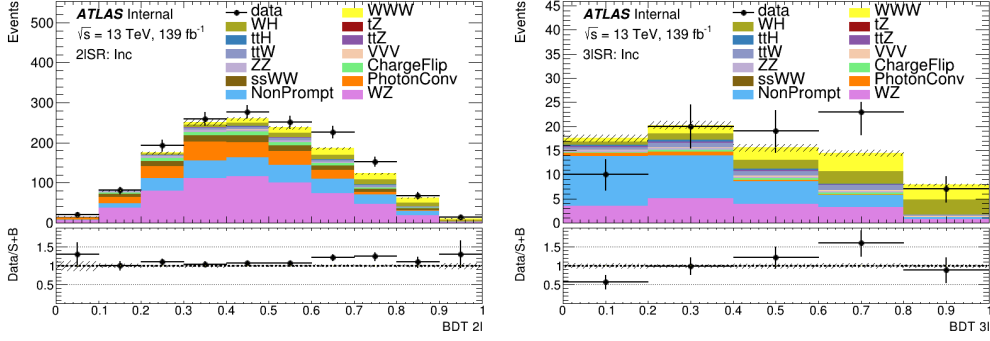
Figure 8.3: The output discriminant score distributions with statistical uncertainties based on the $\ell^\pm \nu \ell^\pm \nu jj$ BDT (left) and $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ BDT (right) in the corresponding $\ell^\pm \nu \ell^\pm \nu jj$ and $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SRs.

## 8.3 Validation

The BDT models are then applied and validated in the CRs and VRs with the data, signal, and background samples. The distributions of discriminant scores based on the $\ell^\pm \nu \ell^\pm \nu jj$ BDT in the $WZ2j$ CR, $\ell^\pm \nu \ell^\pm \nu jj$ $b$-tagging CR, $W$ sideband CR, $Z$ window VR, and $Z\gamma$ CR are shown in Figure 8.4, while the distributions in each channel in the $WZ2j$ CR, $\ell^\pm \nu \ell^\pm \nu jj$ $b$-tagging CR, and $W$ sideband CR are attached in Appendices A.2, A.3 and A.5. The distributions of discriminant scores based on the $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ BDT in the $WZ01j$ CR, $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ $b$-tagging CR, and $Z\gamma$ CR are shown in Figure 8.5, while the distributions in each channel in the $WZ01j$ CR and $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ $b$-tagging CR are attached in Appendices A.1 and A.4. The data agrees with the simulation within uncertainties for the discriminant scores based on both BDT models and thus the models do not have any bias on the data and simulation samples.
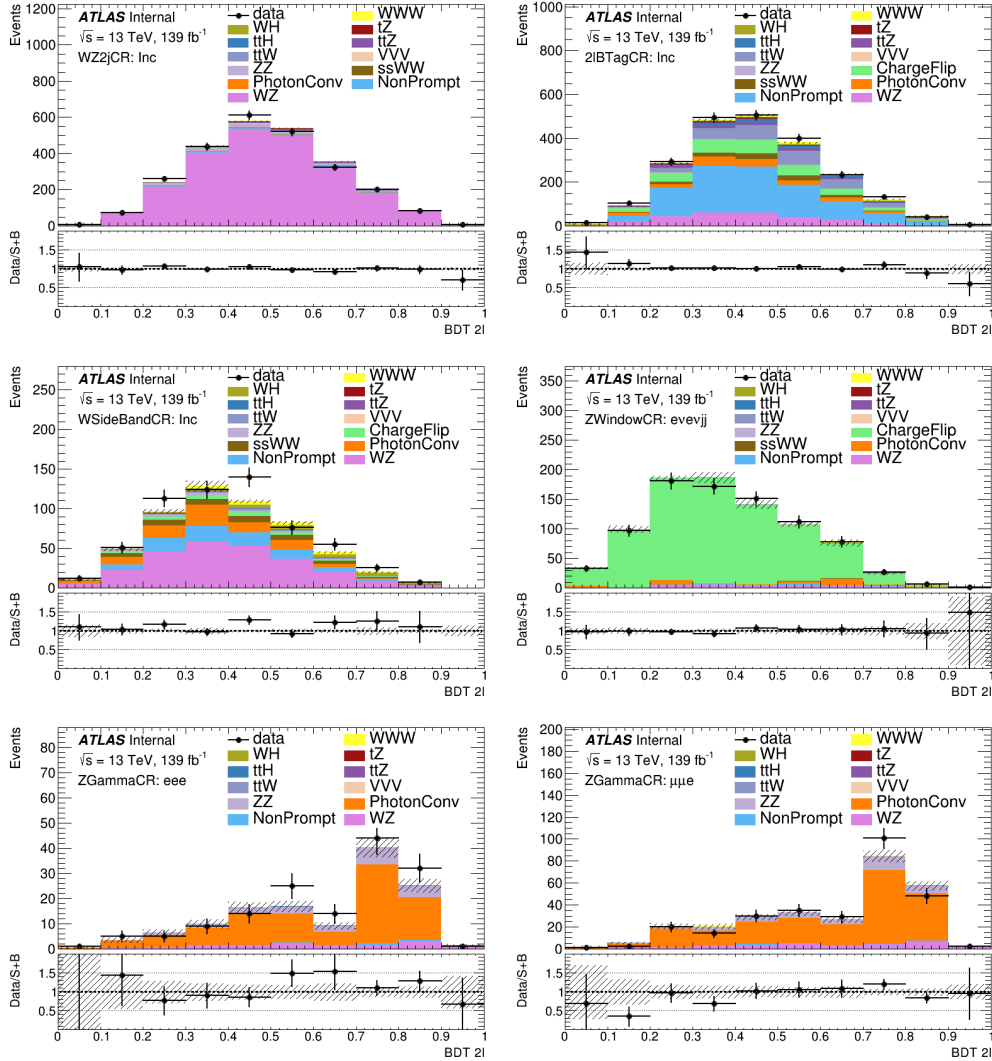
Figure 8.4: The discriminant score distributions with statistical uncertainties based on the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ BDT in the inclusive channels in the $WZ2j$ CR (top left), $\ell^{\pm}\nu\ell^{\pm}\nu jj$ $b$-tagging CR (top right), and $W$ sideband CR (middle left), as well as $ee$ channel in the $Z$ window VR (middle right), $eee$ channel in the $Z\gamma$ CR (bottom left), and $\mu\mu e$ channel in the $Z\gamma$ CR (bottom right).
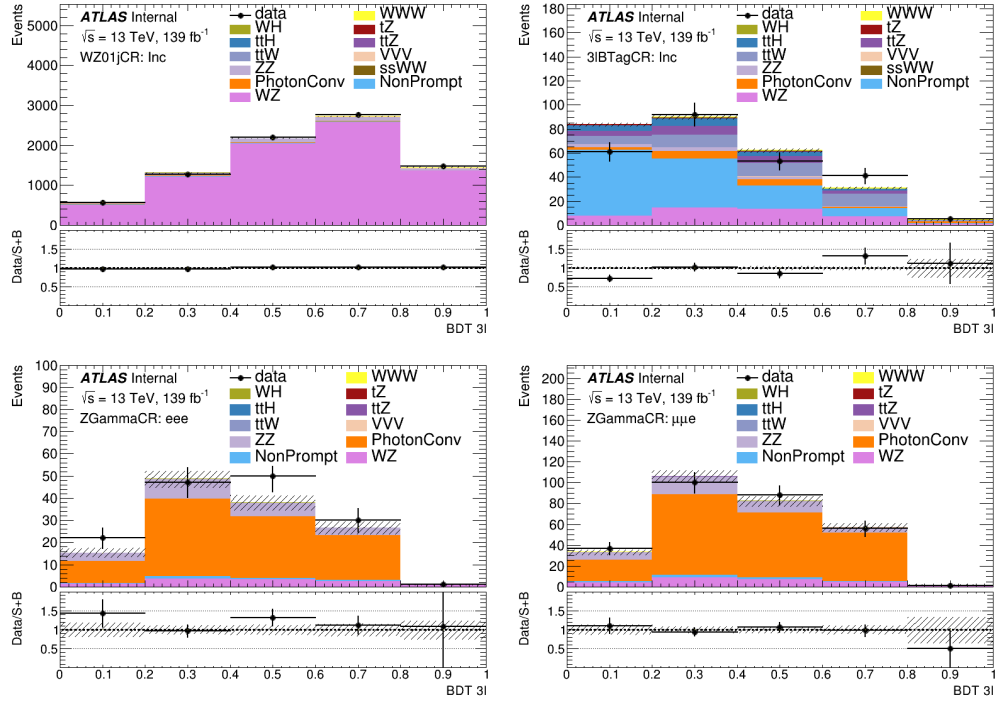
Figure 8.5: The discriminant score distributions with statistical uncertainties based on the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ BDT in the inclusive channels in the *WZ01j* CR (top left) and $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ *b*-tagging CR (top right), as well as *eee* (bottom left) and *μμe* (bottom right) channels in the *Zγ* CR (bottom left).

# CHAPTER 9

# Systematic Uncertainties

## 9.1 Overview

Besides statistical uncertainties from data and MC samples, systematic uncertainties also need to be considered in the $W^\pm W^\pm W^\mp$ analysis and influence the final results. Each systematic uncertainty can be interpreted as up and down variations or a one-side variation of one standard deviation ($\sigma$) with respect to the nominal value due to one source of error which can be either experimental or theoretical source. The systematic uncertainties can then be categorized into experimental systematic uncertainties and theoretical systematic uncertainties based on the types of sources they come from.

The effect of each systematic uncertainty in the $W^\pm W^\pm W^\mp$ cross section measurement is quantified by the shift of the signal strength defined in Section 10.1 ($\Delta\mu_s/\mu_s$) between the nominal fit described in Section 10.3 and another fit with the systematic uncertainty is fixed at up and down variations of one standard deviation ($\pm 1\sigma$) while remaining uncertainties are fixed still at their constrained values in the nominal fit. Table 9.1 shows the major uncertainties in the analysis and their effects. All systematic uncertainties studied in this analysis are listed and explained in details in Sections 9.2 and 9.3.

## 9.2 Experimental Systematic Uncertainties

### 9.2.1 Monte Carlo Uncertainties

Experimental systematic uncertainties for MC samples are provided by the ATLAS CP group and are listed as follows:

- Luminosity Uncertainty
  As mentioned in Section 5.1, the uncertainty of the total integrated luminosity used to scale

| Systematic uncertainty name | $\Delta\mu_s/\mu_s$ with up/down variations |
|---|---|
| Non-prompt MC closure ($\mu$) | $-12\%/18\%$ |
| $WZ2j$ scale | $4.0\%/-4.9\%$ |
| Luminosity | $-3.0\%/3.1\%$ |
| Jet energy scale pileup ($\mu$ offset) | $3.1\%/-3.1\%$ |
| Non-prompt rate stat. ($\mu$) | $-2.9\%/3.0\%$ |
| Electron isolation | $-2.5\%/2.5\%$ |
| Jet energy scale $\eta$-intercalibration (mismodeling) | $2.5\%/-2.5\%$ |
| Non-prompt rate stat. ($e$) | $-2.4\%/2.5\%$ |
| $WWW$ parton shower | $-2.0\%/2.0\%$ |
| Electron identification | $-2.0\%/2.0\%$ |
| Pileup reweighting | $1.8\%/-1.8\%$ |
| $b$-tagging $c$-jets (par. 0) | $-1.8\%/1.8\%$ |
| $WZ2j$ parton shower | $1.6\%/-1.7\%$ |
| Muon isolation (sys.) | $-1.7\%/1.7\%$ |
| Jet energy scale flavor composition | $1.6\%/-1.7\%$ |

Table 9.1: Major systematic uncertainties and their effects in the $W^\pm W^\pm W^\mp$ cross section measurement.

the MC samples is 1.7%. The uncertainty is not applied on $WZ$ samples as the normalization is determined from data in the $WZ$ CRs.

- Pileup Reweighting Uncertainty
  This is the uncertainty associated to the MC pileup reweighing mentioned in Section 4.2.

- Electron Systematic Uncertainties [52, 62–64, 83]
  Detailed uncertainties are listed below and they are variations in the electron reconstruction, identification, and isolation efficiencies mentioned in Section 4.3 and the electron trigger efficiency mentioned in Section 4.8. They also include energy resolution and energy scale uncertainties used for electron four-momentum (energy) corrections.

  – Electron Reconstruction Efficiency Uncertainty

  – Electron Identification Efficiency Uncertainty

  – Electron Isolation Efficiency Uncertainty

  – Electron Trigger Uncertainty
    It includes the effects from both trigger efficiency and trigger scale factor.

  – Electron Energy Resolution Uncertainty

  – Electron Energy Scale Uncertainty

89

- Muon Systematic Uncertainties [52, 70, 71, 84]

  Detailed uncertainties are listed below. Similar to electron systematic uncertainties, they include variations in muon reconstruction (combined with identification), isolation, track-to-vertex-association and trigger efficiencies mentioned in Sections 4.4 and 4.8. Additionally, they also include uncertainties in muon track resolution, momentum scale, and charge-dependent sagitta bias corrections for muon momentum scale.

  – Muon Reconstruction Efficiency Uncertainty

    It includes the effects from both statistical and systematic errors.

  – Muon Isolation Efficiency Uncertainty

    It includes the effects from both statistical and systematic errors.

  – Muon Track-to-vertex-association Efficiency Uncertainty

    It includes the effects from both statistical and systematic errors.

  – Muon Trigger Efficiency Uncertainty

    It includes the effects from both statistical and systematic errors.

  – Muon Track Resolution Uncertainty

    It includes the effects from muon track resolution in both the inner detector and the muon spectrometer.

  – Muon Momentum Scale Uncertainty

  – Muon Sagitta Bias Correction Uncertainty

    It includes the effects of the residual charge-dependent sagitta bias and combined correction at $Z$ scale on momentum measurements with both the inner detector and the muon spectrometer.

- Jet Systematic Uncertainties

  Detailed uncertainties are listed below and they include variations in jet energy scale calibration, JVT, and $b$-tagging mentioned in Section 4.5. They also include uncertainty in jet energy resolution.

  – Jet Energy Scale Uncertainty [74]

    It includes effects from calibrating the scale of forward jets with respect to central jets, calibrating with respect to well understood objects including $Z$ in $Z$+jet, $\gamma$ in $\gamma$+jet, and multijet, propagation of single particle and test beam uncertainties, pileup, jet flavor, and punch-through correction in the global sequential calibration.

  – Jet Energy Resolution Uncertainty [118]

    It includes effects from difference in data and MC, noise term from random cones in

zero bias data, and jet $p_T$ asymmetry in dijet balance.

– Jet JVT Efficiency Uncertainty [76, 77]

– $b$-tagging Efficiency Uncertainty [79, 119, 120]
  It includes effects from $b$-jets, $c$-jets, and light-flavor jets after eigenvector decomposition as well as high-$p_T$ jets extrapolation.

• $E_T^{miss}$ Track Soft Term Uncertainties [80]
  The detailed uncertainties are listed below and they include the variations in the $E_T^{miss}$ soft terms that are mentioned in Section 4.6.

– $E_T^{miss}$ Track Soft Term Scale Uncertainty
  It includes effect from scaling the magnitude of the soft terms in the direction of the $p_T$ of the hard terms.

– $E_T^{miss}$ Track Soft Term Resolution Uncertainty
  It includes effects from smearing the magnitude of the soft terms in the direction parallel and perpendicular to the direction of the $p_T$ of the hard terms.

### 9.2.2 Data-driven Background Uncertainties

Experimental systematic uncertainties of data-driven background estimation come from the method of calculating data-driven rates and the differences of data-driven background compositions in the CRs and the SRs. Data-driven background uncertainties studied are listed below. Most of them are included in the fits introduced in Sections 10.2 and 10.3 except for those specified in the descriptions as they are too small to be included.

• Non-prompt Background Uncertainties

– Non-prompt Rate Statistical Uncertainty
  As mentioned in Section 7.3.2, it comes from the rate calculation procedure and includes effects from both electron and muon non-prompt rates.

– $WZ$ Normalization Factor Uncertainty
  $WZ$ MC samples are used in non-prompt rate calculation and thus the uncertainties of the $WZ$ normalization factors mentioned in Section 10.2 are considered as an uncertainty of the non-prompt rates. The non-prompt rates calculated with up and down variations of one standard deviation ($\pm 1\sigma$) on the $WZ$ normalization factors are listed in Table 9.2 and the uncertainty includes effects from both electron and muon non-prompt rates with variations of both $\mu_{WZ01j}$ and $\mu_{WZ2j}$.

| $\mu_{WZ}$ | | $\mu_{WZ01j}$ | | $\mu_{WZ2j}$ | |
|---|---|---|---|---|---|
| Variation | | $-1\sigma$ | $+1\sigma$ | $-1\sigma$ | $+1\sigma$ |
| $\epsilon_e$ | $p_{\mathrm{T}} < 27$ GeV | $0.016 \pm 0.003$ | $0.016 \pm 0.003$ | $0.016 \pm 0.003$ | $0.015 \pm 0.003$ |
| | $p_{\mathrm{T}} > 27$ GeV | $0.019 \pm 0.005$ | $0.019 \pm 0.005$ | $0.022 \pm 0.005$ | $0.016 \pm 0.005$ |
| $\epsilon_\mu$ | $p_{\mathrm{T}} < 27$ GeV | $0.014 \pm 0.002$ | $0.014 \pm 0.002$ | $0.015 \pm 0.002$ | $0.014 \pm 0.002$ |
| | $p_{\mathrm{T}} > 27$ GeV | $0.014 \pm 0.001$ | $0.014 \pm 0.001$ | $0.014 \pm 0.001$ | $0.013 \pm 0.001$ |

Table 9.2: Non-prompt rates calculated with up and down variations of one standard deviation ($\pm 1\sigma$) on the $WZ$ normalization factors. The uncertainties on the rates are statistical uncertainties from the rate calculation.

    – Non-prompt Rate Binning Uncertainty

       The two-bin non-prompt rates are recalculated with different $p_{\mathrm{T}}$ bins and listed in Table 9.3. It includes the effects from both the electron and muon non-prompt rates.

| Lepton $p_{\mathrm{T}}$ | $< 25$ GeV | $> 25$ GeV | $< 30$ GeV | $> 30$ GeV |
|---|---|---|---|---|
| $\epsilon_e$ | $0.017 \pm 0.003$ | $0.018 \pm 0.004$ | $0.014 \pm 0.002$ | $0.025 \pm 0.006$ |
| $\epsilon_\mu$ | $0.015 \pm 0.002$ | $0.014 \pm 0.001$ | $0.014 \pm 0.001$ | $0.015 \pm 0.001$ |

Table 9.3: Non-prompt rates calculated with different lepton $p_{\mathrm{T}}$ bins. The uncertainties on the rates are statistical uncertainties from the rate calculation.

    – MC Closure Test Uncertainty

       As mentioned in Section 7.3.1, the non-prompt background in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR comes from both $W+$jets and $t\bar{t}$ processes while in the $b$-tagging CRs it mainly comes from $t\bar{t}$ process. Therefore, an uncertainty is needed to compensate the differences in non-prompt background compositions. All MC samples are used to calculate MC-based non-prompt rates in both $b$-tagging CRs and SRs by selecting events with at least one non-prompt lepton in truth. The calculated non-prompt rates are shown in Table 9.4. The difference between MC-based non-prompt rates in the $b$-tagging CRs and the SRs is quite significant in low $p_{\mathrm{T}}$ bins while it is not that much in high $p_{\mathrm{T}}$ bins. The reason is due to the $W+$jets events are dominated in low $p_{\mathrm{T}}$ bins in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR and this can been indicated from the anti-ID lepton $p_{\mathrm{T}}$ distributions in the anti-ID $\ell^{\pm}\nu\ell^{\pm}\nu jj$ CR in Figure 9.1. The difference between MC-based non-prompt rates in $b$-tagging CRs and SRs is applied to the nominal non-prompt rates as alternative rates used to estimate non-prompt background in the SRs. The uncertainty includes the effects from both the electron and muon non-prompt rates.

• Photon Conversion Background Uncertainties

| Region | | $b$-tagging CRs | SRs |
|---|---|---|---|
| $\epsilon_e$ | $p_T < 27$ GeV | $0.0133 \pm 0.0018$ | $0.0139 \pm 0.0031$ |
| | $p_T > 27$ GeV | $0.0147 \pm 0.0017$ | $0.0153 \pm 0.0029$ |
| $\epsilon_\mu$ | $p_T < 27$ GeV | $0.0104 \pm 0.0012$ | $0.0221 \pm 0.0031$ |
| | $p_T > 27$ GeV | $0.0082 \pm 0.0006$ | $0.0080 \pm 0.0010$ |

Table 9.4: MC-based non-prompt rates calculated in the $b$-tagging CRs and SRs using all MC samples. The uncertainties on the rates are statistical uncertainties from the rate calculation.

– Photon Conversion Rate Statistical Uncertainty

As mentioned in Section 7.4.1, it comes from the rate calculation procedure.

– *WZ* Normalization Factor Uncertainty

Similar to that in the non-prompt background uncertainties, the photon conversion rates are calculated with the variations of *WZ* normalization factors and the numbers are shown in Table 9.5. Since the changes in the photon conversion rates are very small compared with the statistical uncertainties, this uncertainty is not included in the fits.

| $\mu_{WZ}$ | $\mu_{WZ01j}$ | | $\mu_{WZ2j}$ | |
|---|---|---|---|---|
| Variation | $-1\sigma$ | $+1\sigma$ | $-1\sigma$ | $+1\sigma$ |
| $\epsilon_\gamma$ | $0.420 \pm 0.034$ | $0.416 \pm 0.034$ | $0.421 \pm 0.034$ | $0.415 \pm 0.034$ |

Table 9.5: Photon conversion rates calculated with up and down variations of one standard deviation ($\pm 1\sigma$) on the *WZ* normalization factors. The uncertainties on the rates are statistical uncertainties from the rate calculation.

– Electron $p_T$ Dependence Uncertainty

The nominal photon conversion rate can be considered as an average of electron $p_T$ and the $p_T$ dependence is studied. The $\mu\mu e$ channel of the $Z\gamma$ CR is further split into two regions with electron $p_T \leq 25$ GeV and $> 25$ GeV and the photon conversion rate is recalculated in each region. 25 GeV is chosen as the bin boundary for similar statistics in two $p_T$ bins as shown in Figure 9.2. The photon conversion rates in different electron $p_T$ ranges are listed in Table 9.6. Since the changes in the photon conversion rates are very small compared with the statistical uncertainties, this uncertainty is not included in the fits.

– MC Closure Test Uncertainty

Due to the requirement of the number of leptons, the photon conversion background in the $\ell^\pm \nu \ell^\pm \nu jj$ SR is mainly from the $W\gamma$ process while the photon conversion background in the $Z\gamma$ CR is mainly from the $Z\gamma$ process. Therefore the effect on the photon
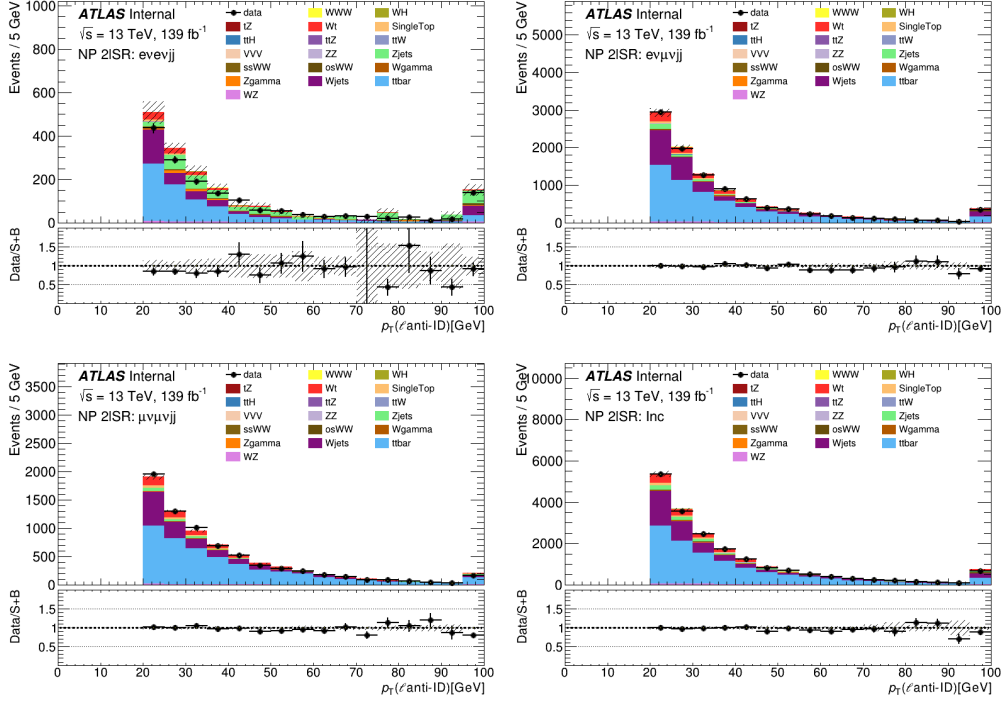
Figure 9.1: The anti-ID lepton $p_T$ distributions with statistical uncertainties in *ee* (top left), *eμ* (top right), *μμ* (bottom left), and inclusive (bottom right) channels in the anti-ID $\ell^{\pm}\nu\ell^{\pm}\nu jj$ CR. Only MC simulation is used for the predictions.

| Electron $p_T$ | $\leq 25$ GeV | $> 25$ GeV |
|---|---|---|
| $\epsilon_\gamma$ | $0.422 \pm 0.044$ | $0.414 \pm 0.052$ |

Table 9.6: Photon conversion rates in different electron $p_T$ ranges. The uncertainties on the rates are statistical uncertainties from the rate calculation.

conversion rate from different background compositions is studied. The photon conversion rates are recalculated with $Z\gamma$ MC samples in the *μμe* channel of the $Z\gamma$ CR and $W\gamma$ MC samples in the *ee* and *eμ* channels of the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR. The MC samples are selected with at least one photon conversion electron in truth. The numbers are shown Table 9.7. The difference between the MC-based photon conversion rates in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR and the $Z\gamma$ CR is applied to the nominal photon conversion rate as an alternative rate used to estimate photon conversion background in the SRs.

- Charge-flip Background Uncertainties

  - Charge-flip Rate Statistical Uncertainty
    As mentioned in Section 7.5.1, it comes from the rate calculation procedure.

94
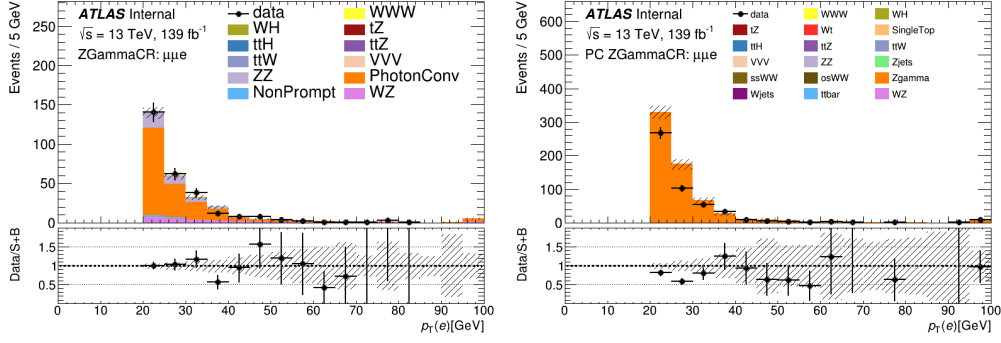
Figure 9.2: The electron $p_T$ distributions with statistical uncertainties in $\mu\mu e$ channel in the $Z\gamma$ CR (left) and the anti-BL $Z\gamma$ CR (right). Prompt MC and data-driven estimation are used for the prediction in the $Z\gamma$ CR while all MC samples are used for the prediction in the anti-BL $Z\gamma$ CR as the data-driven estimation is not available for the anti-BL region.

| MC sample | $W\gamma$ | | $Z\gamma$ |
|---|---|---|---|
| Region | $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR | | $Z\gamma$ CR |
| Channel | $ee$ | $e\mu$ | $\mu\mu e$ |
| $\epsilon_\gamma$ | $0.132 \pm 0.034$ | $0.176 \pm 0.026$ | $0.158 \pm 0.016$ |

Table 9.7: Photon conversion rates calculated with the $W\gamma$ MC and $Z\gamma$ MC in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR and $Z\gamma$ CR respectively. The uncertainties on the rates are statistical uncertainties from the rate calculation.

– *Z* Mass Window Uncertainty

As mentioned in Section 7.5.1, the charge-flip rate calculation is performed in the central charge-flip CR with 105 GeV $> m_{\ell\ell} > 75$ GeV and the sideband charge-flip CR with 120 GeV $> m_{\ell\ell} \geq 105$ GeV or 75 GeV $\geq m_{\ell\ell} > 60$ GeV. The *Z* mass window in the central charge-flip CR is varied for 6 GeV up and down while the cut of $m_{\ell\ell}$ is changed accordingly in the sideband charge-flip CR to match the *Z* mass window width. The up variation is defined with 108 GeV $> m_{\ell\ell} > 72$ GeV for the central charge-flip CR and 126 GeV $> m_{\ell\ell} \geq 108$ GeV or 72 GeV $\geq m_{\ell\ell} > 54$ GeV for the sideband charge-flip CR. The down variation is defined with 102 GeV $> m_{\ell\ell} > 78$ GeV for the central charge-flip CR and 114 GeV $> m_{\ell\ell} \geq 102$ GeV or 78 GeV $\geq m_{\ell\ell} > 66$ GeV for the sideband charge-flip CR. The recalculated charge-flip rates with different *Z* mass window are shown in Figure 9.3.

– MC Subtraction Uncertainty

As mentioned in Section 7.5.1, the nominal charge-flip rates are calculated from the $Z \to ee$ events which are selected by subtracting data events in the sideband charge-flip

95
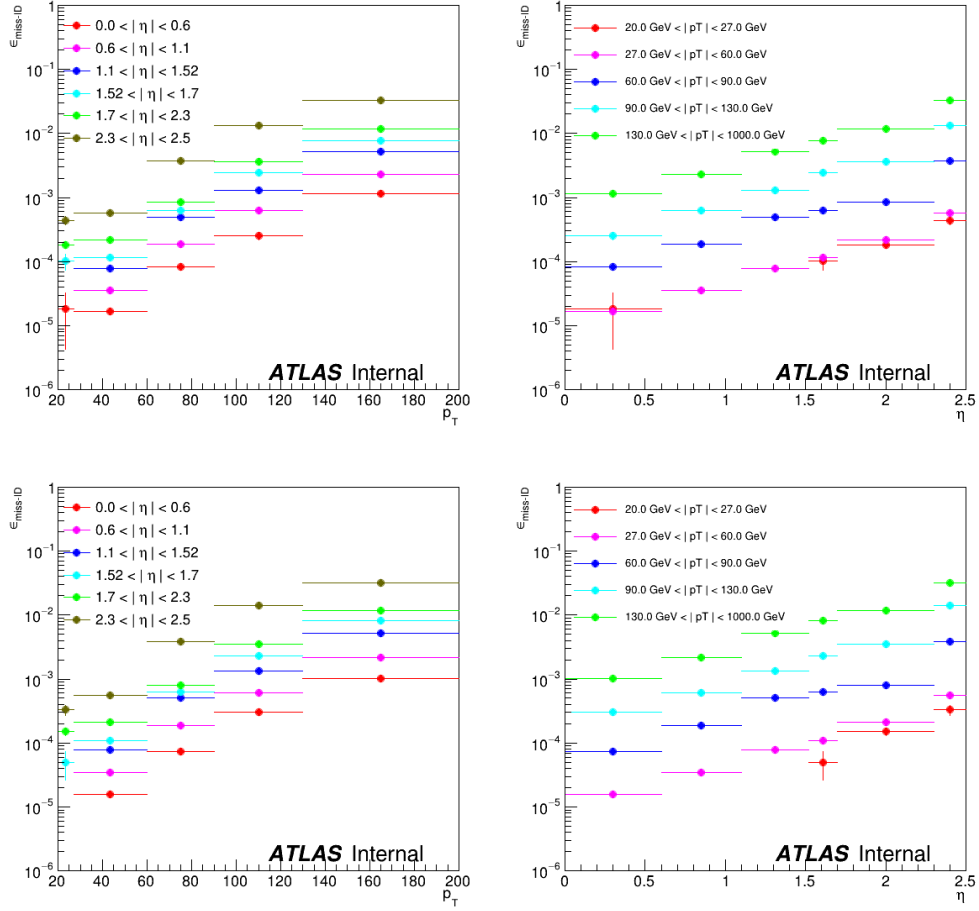
Figure 9.3: The charge-flip rates as a function of electron $p_T$ (left) and $|\eta|$ (right) calculated with the up (top) and down (bottom) variations of 6 GeV in the $Z$ mass window in the central charge-flip CR. The error bars on the rates are statistical uncertainties from the rate calculation.

CR from data events in the central charge-flip CR. For this uncertainty, the charge-flip rates are calculated alternatively by subtracting all MC samples excluding $Z$ + jets MC samples from the events in data in the central charge-flip CR and shown in Figure 9.4.

– $Z$ + jets MC Closure Test Uncertainty

For this uncertainty, MC-based charge-flip rates are derived just based on the $Z$ + jets MC samples in the central charge-flip CR and are shown in Figure 9.5. Then the MC-based charge-flip rates are applied to the OS $Z$ + jets MC events in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR and compared with the SS $Z$ + jets MC events in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR. The numbers are shown in Table 9.8. Since the difference in the numbers is within the statistical uncertainty, this uncertainty is not included in the fits.

– $t\bar{t}$ MC Closure Test Uncertainty
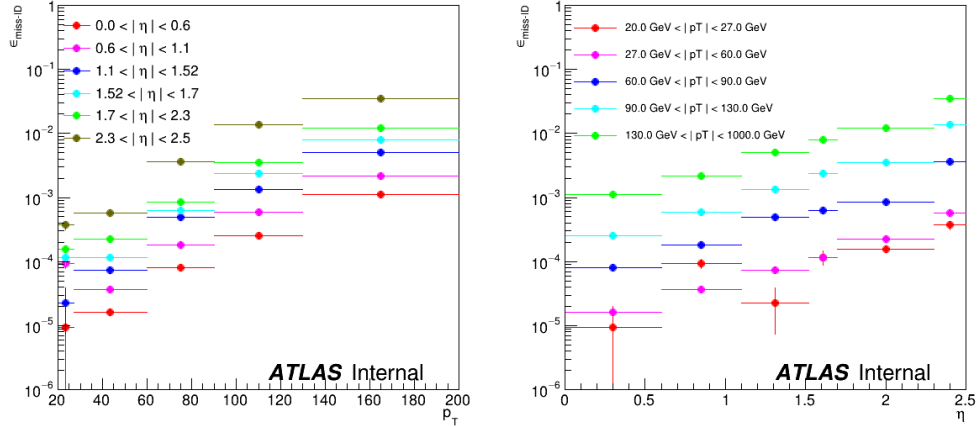
Figure 9.4: The measured charge-flip rates as a function of electron $p_T$ (left) and $|\eta|$ (right) by subtracting all MC samples excluding $Z+$jets MC samples from data events in the central charge-flip CR. The error bars on the rates are statistical uncertainties from the rate calculation.

| Estimation method | $Z+$jets MC | $Z+$jets MC-based charge-flip rate |
|---|---|---|
| $\ell^\pm \nu \ell^\pm \nu jj$ SR | $15.6 \pm 6.7$ | $20.99 \pm 0.75$ |

Table 9.8: The numbers of $Z+$jets events with statistical uncertainties in the $\ell^\pm \nu \ell^\pm \nu jj$ SR estimated from the $Z+$jets MC samples and $Z+$jets MC-based charge-flip rates calculated in the charge-flip CR.

Besides the $Z+$jets process, $t\bar{t}$ process can also contribute to the charge-flip background in the $\ell^\pm \nu \ell^\pm \nu jj$ SR. For this uncertainty, $t\bar{t}$ charge-flip CR is defined with the same cuts in the charge-flip CR except for cuts on $b$-jets and $m_{\ell\ell}$. The cuts after the preselection are listed in Table 9.9. Exactly two $b$-jets are required in this region and $m_{\ell\ell}$ cut is dropped as there is no $Z$ boson in the $t\bar{t}$ process. Similar to the $Z+$jets MC closure test, the same $Z+$jets MC-based charge-flip rates shown in Figure 9.5 are used but applied to the OS $t\bar{t}$ MC events in the $t\bar{t}$ charge-flip CR and compared with the SS $t\bar{t}$ MC events with at least one charge-flipped electron in truth in the $t\bar{t}$ charge-flip CR. The numbers are shown in Table 9.10.

## 9.3 Theoretical Systematic Uncertainties

Theoretical systematic uncertainties come from generating PDF, matrix elements and parton shower in the MC generation described in Section 5.2.1. Therefore, they are considered only for signal and prompt background processes, which are simulated with MC samples. Theoretical

97
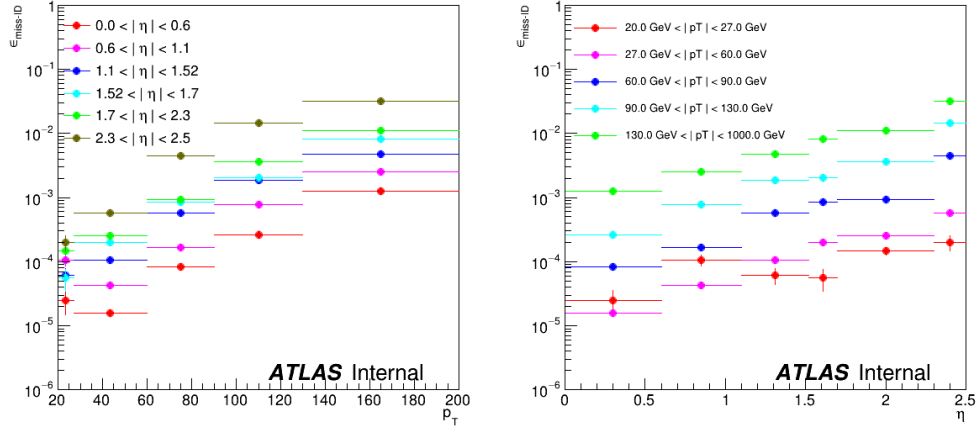
Figure 9.5: The measured MC-based charge-flip rates as a function of electron $p_T$ (left) and $|\eta|$ (right) with $Z$+jets MC in the central region of the charge-flip CR. The error bars on the rates are statistical uncertainties from the rate calculation.

| | Charge-flip CR | $t\bar{t}$ charge-flip CR |
|---|---|---|
| Channel | $ee$ | |
| $N_{\text{lepton}}$ | $= 2$ | |
| Veto lepton | veto $3^{\text{rd}}$ lepton | |
| $p_T(\ell_1)$ | $> 27$ GeV | |
| $m_{\ell\ell}$ | $60 < m_{\ell\ell} < 120$ GeV | - |
| $N_{b\text{-jet}}$ | - | $= 2$ |

Table 9.9: List of the event-level selection cuts used in the charge-flip CR and the $t\bar{t}$ charge-flip CR. The definitions of the symbols are the same as those in Table 6.1.

uncertainties of the signal processes $WWW$ and $WH$ as well as the dominant prompt background process $WZ$ are studied in details with the scale, PDF, and parton shower uncertainties, while an overall theoretical uncertainty is applied to the remaining prompt background processes. The details are listed as follows:

- $WWW$ Theoretical Uncertainties
  Detailed uncertainties are listed below. Note that $WWW$ scale and PDF uncertainties are considered for shape only in the fits described in Sections 10.2 and 10.3 by normalizing the total number of events to the nominal value. The normalization is considered as a theory uncertainty in the MC predicted signal cross section described in Section 10.4 since this is not an uncertainty from the $W^{\pm}W^{\pm}W^{\mp}$ cross section measurement.

  - $WWW$ Scale Uncertainty

| Estimation method | $t\bar{t}$ MC | $Z$ + jets MC-based charge-flip rate |
|---|---|---|
| $t\bar{t}$ charge-flip CR | $112.8 \pm 4.1$ | $118.72 \pm 0.29$ |

Table 9.10: The numbers of $t\bar{t}$ charge-flip events with statistical uncertainties in the $t\bar{t}$ charge-flip CR estimated with $t\bar{t}$ MC samples and $Z$ + jets MC-based charge-flip rate calculated in the charge-flip CR.

It consists of the effect with an envelope of seven-point variations of the renormalization scale ($\mu_{\rm r}$) and the factorization scale ($\mu_{\rm f}$) defined in Section 5.2.1, following the recommendation from the ATLAS CP group. The seven-point variations includes $(0.5\mu_{\rm r}, 0.5\mu_{\rm f})$, $(0.5\mu_{\rm r}, \mu_{\rm f})$, $(\mu_{\rm r}, 0.5\mu_{\rm f})$, $(2\mu_{\rm r}, \mu_{\rm f})$, $(\mu_{\rm r}, 2\mu_{\rm f})$, $(2\mu_{\rm r}, 2\mu_{\rm f})$, and the nominal $(\mu_{\rm r}, \mu_{\rm f})$. The variations are stored as event-by-event weights when generating MC samples.

– *WWW* PDF Uncertainty

As mentioned in Section 5.2.2, the nominal *WWW* sample is generated with a particular choice of the PDF NNPDF3.0NNLO and effects of alternative choices of the PDF need to be considered. Additionally, the strong coupling constant $\alpha_{\rm S}$ is determined experimentally as $\alpha_{\rm S} = 0.1180 \pm 0.0015$ [121] and its uncertainty also needs to be included. Following the recommendation from PDF4LHC [121], this uncertainty is evaluated with variations of alternative PDFs MMHT2014NNLO [122] and CT14NNLO [123] as well as up and down variations of the combined PDF and strong coupling constant (PDF+$\alpha_{\rm S}$) uncertainty. The variations are stored as event-by-event weights when generating MC samples and the uncertainty includes the effect with an envelope of all four variations.

– *WWW* Parton Shower Uncertainty

Following the recommendation from the ATLAS multi-boson simulation [124], the variations of the resummation scale ($\mu_{\rm q}$) and the merging scale defined in Section 5.2.1 are considered in this uncertainty. The variations of the resummation scale ($\mu_{\rm q}$) include $2\mu_{\rm q}$ and $0.5\mu_{\rm q}$. The variation on the CKKW [105, 106] merging scale is 15 GeV from the nominal value of 30 GeV. The Rivet routine [125] is used to evaluate these variations in acceptance for the SRs. The uncertainty is obtained by adding the larger uncertainty in two $\mu_{\rm q}$ variations and the CKKW merging scale uncertainty in quadrature. It is 2.3% in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR and 2.4% in the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ SR.

• *WH* Theoretical Uncertainties

Detailed uncertainties are listed below. Similar to *WWW*, only shape-only effects of the *WH*

scale and PDF uncertainties are included in the fits and the normalization is included in the MC predicted signal cross section.

- **WH Scale Uncertainty**

  It is calculated following the same procedure as the *WWW* scale uncertainty and includes the effect with an envelope of the seven-point variation of $\mu_r$ and $\mu_f$.

- **WH PDF Uncertainty**

  As mentioned in Section 5.2.2, the *WH* sample is generated with the NNPDF3.0NLO PDF. Therefore, MMHT2014NLO and CT14NLO are used as alternative PDFs following the recommendation from PDF4LHC. The PDF+$\alpha_S$ uncertainty is not included as this is only available for SHERPA samples but *WH* is generated with POWHEG. The uncertainty includes the effect with an envelope of these two variations.

- **WH Parton Shower Uncertainty**

  As mentioned in Section 5.2.2, the parton shower of the *WH* sample is generated with PYTHIA 8. For this uncertainty, HERWIG 7 is used as an alternative generator for the parton shower. The difference between samples obtained from these two generators is 2.1% in the $\ell^\pm \nu \ell^\pm \nu jj$ SR and 0.002% in the $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SR.

- *WZ* Theoretical Uncertainties

  Detailed uncertainties are listed below. Since *WZ* is considered as a free-floating background and the overall normalization is determined from the *WZ* CRs, the scale and PDF uncertainties on the *WZ* cross-section should not be considered. Therefore, for *WZ* scale and PDF uncertainties, only shape-only effects are included in the fits.

  - **WZ Scale Uncertainty**

    It is calculated following the same procedure as the *WWW* scale uncertainty and includes the effects with an envelope of the seven-point variation of $\mu_r$ and $\mu_f$ from *WZ* processes with less than two jets (*WZ01j*) and greater than two jets (*WZ2j*).

  - **WZ PDF Uncertainty**

    It is calculated following the same procedure as the *WWW* PDF uncertainty and includes the effects with an envelope of the four variations from both the *WZ01j* and *WZ2j* processes.

  - **WZ Parton Shower Uncertainty**

    It is calculated following the same procedure as the *WWW* parton shower uncertainty except that the CKKW merging scale variations are 15 GeV and 30 GeV from the nominal value of 20 GeV. The uncertainty is calculated by adding the larger uncertainty in

100

two $\mu_q$ variations and the larger uncertainty in two CKKW merging scales in quadrature. It is 5.2% in the $\ell^\pm \nu \ell^\pm \nu j j$ SR, 12.4% in the $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SR, 2.7% in the $WZ01j$ CR, and 7.8% in the $WZ2j$ CR from both the $WZ01j$ and $WZ2j$ processes.

- $ZZ$, SS $WW$, $t\bar{t}V(V = W,Z)$, $t\bar{t}H$, $VVV$(excluding $WWW$), and $tZ$ Theoretical Uncertainties Contributions from other prompt processes are minor in the SRs and an overall uncertainty of 10% is applied to each process as its theoretical uncertainty, which is at the similar scale of the uncertainty used in the previous $W^\pm W^\pm W^\mp$ analysis at ATLAS with 79.8 fb$^{-1}$ of the $pp$ collision data at $\sqrt{s} = 13$ TeV [3].

# CHAPTER 10

# Fitting Procedure and Signal Extraction

## 10.1   General Fitting Procedure

$W^{\pm}W^{\pm}W^{\mp}$ inclusive production cross section is measured with a binned maximum-likelihood fit using TRExFitter [126]. The measured $W^{\pm}W^{\pm}W^{\mp}$ inclusive production cross section $\sigma_s^{\text{Obs}}$, can be expressed as:

$$\sigma_s^{\text{Obs}} = \mu_s \times \sigma_s^{\text{MC}}, \tag{10.1}$$

where $\sigma_s^{\text{MC}}$ is the total cross section of the signal MC samples and $\mu_s$ is defined as signal strength. The fit is achieved equivalently by minimizing a negative log-likelihood function $-\ln\mathcal{L}$ and the likelihood function $\mathcal{L}$ is defined in the following way:

$$\mathcal{L}(\mu_s, \mu_{WZ}, \theta) = \prod_{r \in \text{Regions}} \prod_{c \in \text{Channels}} \prod_{b \in \text{Bins}} \text{Pois}(N_{r,c,b}^{\text{Data}} | N_{r,c,b}^{\text{Exp}}) \prod_{i \in \text{Systs}} \text{Gaus}(\theta_i | 0, 1), \tag{10.2}$$

which includes one Poisson probability density function for each bin in the distribution of the variable used for fitting in each channel and region, and one Gaussian probability density function for each Nuisance Parameter (NP) $\theta_i$ which is used to model systematic uncertainties mentioned in Chapter 9 with Eqs. 10.4 and 10.5 explained later. $N_{r,c,b}^{\text{Data}}$ and $N_{r,c,b}^{\text{Exp}}$ are defined as number of events in data and expectation in bin $b$ of the distribution in channel $c$ and region $r$. $\mu_{WZ01j}$ and $\mu_{WZ2j}$ are $WZ$ normalization factors and used in evaluating $N_{r,c,b}^{\text{Exp}}$ with the signal strength $\mu_s$:

$$N_{r,c,b}^{\text{Exp}} = \mu_s N_{r,c,b}^s(\theta_i^s) + \sum_{\alpha \in \{01j, 2j\}} \mu_{WZ\alpha} N_{r,c,b}^{WZ\alpha}(\theta_i^{WZ\alpha}) + \sum_{p \in \text{Bkgs}} N_{r,c,b}^p(\theta_i^p). \tag{10.3}$$

In this analysis, the signal ($s$) is the combined $WWW$ and $WH \rightarrow WWW$ process, and $WZ$ is treated as a free floating background and split into two independent processes ($WZ\alpha$, $\alpha = 01j, 2j$): $WZ$ with at least two jets ($WZ2j$) and with less than two jets ($WZ01j$). In Eqs. 10.2 and 10.3, $\theta$ stands for the complete set of NPs for all systematical uncertainties, $\theta_i$ stands for each individual NP, and

$\theta_i^p$ stands for the set of NPs related to process $p$. The NP follows the standard Gaussian distribution (mean= 0 and variance= 1) and therefore the Gaussian constrain for each NP is included in the likelihood function in Eq. 10.2. The one-sided systematical uncertainties are symmetrized based on the nominal values and the up and down ($\pm 1\sigma$) variations of systematical uncertainties corresponding to the NP with $\theta_i = \pm 1$. The number of expected events for process $p$ with one NP $\theta_i$ can be expressed as:

$$N_{r,c,b}^p(\theta_i) = N_{r,c,b}^{p,\text{Nom}} \delta_{r,c,b}^p(\theta_i), \tag{10.4}$$

where the function $\delta_{r,c,b}^p(\theta_i)$ satisfies:

$$\delta_{r,c,b}^p(\theta_i = \pm 1) = \frac{N_{r,c,b}^{p,\pm 1\sigma_i}}{N_{r,c,b}^{p,\text{Nom}}}, \tag{10.5}$$

$$\delta_{r,c,b}^p(\theta_i = 0) = 1.$$

The function $\delta_{r,c,b}^p(\theta_i)$ is obtained from quadratic interpolation for $\theta_i$ within $[-1,1]$ and linear extrapolation outside of $[-1,1]$ to make sure the function is differential at $-1$ and $1$. In this way, NPs model the effects of systematic uncertainties. After minimizing the negative log-likelihood function, the best fit values for $\mu_s$, $\mu_{WZ}$, and $\theta$ are obtained. Then the measured $W^\pm W^\pm W^\mp$ inclusive production cross section $\sigma_s^{\text{Obs}}$ can be calculated from Eq. 10.1.

The significance measures the incompatibility between the observed data and the background-only null hypothesis $H_0$ and is calculated using TRExFitter. The test statistic $t$ is chosen to have a clear separation between the null ($H_0$) and the alternative ($H_1$) hypotheses. It is constructed from a profile likelihood approach and can be expressed as:

$$t(\mu_s) = -2\ln\lambda(\mu_s) = -2\ln\frac{\mathcal{L}(\mu_s, \hat{\hat{\mu}}_{WZ}, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\mu}_s, \hat{\mu}_{WZ}, \hat{\theta})}, \tag{10.6}$$

where $\mathcal{L}(\mu_s, \hat{\hat{\mu}}_{WZ}, \hat{\hat{\theta}})$ is the likelihood conditionally maximized by $\hat{\hat{\mu}}_{WZ}$ and $\hat{\hat{\theta}}$ for a given $\mu_s$ and $\mathcal{L}(\hat{\mu}_s, \hat{\mu}_{WZ}, \hat{\theta})$ is the likelihood unconditionally maximized by $\hat{\mu}_s$, $\hat{\mu}_{WZ}$, and $\hat{\theta}$. Then a $p$-value is defined as the probability to observe the data at least as incompatible with the hypothesis as the current observed data if the hypothesis is true and can be expressed as:

$$p(\mu_s) = \int_{t_{\text{Obs}}}^{+\infty} f(t|\mu_s)dt, \tag{10.7}$$

where $f(t|\mu_s)$ is the probability density function of the test static given a signal strength $\mu_s$ and $t_{\text{Obs}}$ is the observed value of the test statistic with current data given a signal strength $\mu_s$. For the null hypothesis with $\mu_s = 0$, the test statistic $t_0 = t(\mu_s = 0) = -2\ln\lambda(\mu_s = 0)$ and the $p$-value

$p_0 = \int_{t_{0,\text{Obs}}}^{+\infty} f(t_0|\mu_s = 0)dt_0$. Then the significance $Z$ can be defined as:

$$Z = \Phi^{-1}(1 - p_0), \tag{10.8}$$

where $\Phi^{-1}$ is the cumulative distribution function of the standard Gaussian distribution. A $p$-value of $2.87 \times 10^{-7}$ corresponding to $Z = 5$ and the significance is $5\sigma$. In high energy physics, a significance above $5\sigma$ is used to claim as an observation while $3\sigma$ is an evidence.

## 10.2 *WZ* Normalization Factor

In order to fix the discrepancy between data and prediction in the *WZ* CRs mentioned in Section 7.2, a fit to determine the *WZ* normalization factors is performed following the same procedure as described in Section 10.1 except only in the *WZ* CRs. The regions, channels, distributions, and number of bins in each distribution are shown in Table 10.1. After maximizing the likelihood function, the best fit values for the *WZ* normalization factors are determined as:

$$\mu_{WZ01j} = 1.05^{+0.06}_{-0.06}, \\ \mu_{WZ2j} = 0.83^{+0.20}_{-0.15}. \tag{10.9}$$

Note that the *WZ* normalization factors determined here are not used in the final combined fit with data in the SRs to extract observed signal strength and significance. The final combined fit determines the *WZ* normalization factors by itself.

| Region | Channel | Distribution | $N_{\text{bin}}$ |
|---|---|---|---|
| $WZ01j$ CR | Inclusive | $m_{\ell\ell\ell}$ | 5 |
| $WZ2j$ CR | Inclusive | $m_{\ell\ell\ell}$ | 5 |

Table 10.1: Summary of regions, channels, distributions, and number of bins in each distribution used in the fit for the *WZ* normalization factors.

The plots of $m_{\ell\ell\ell}$ used in the fit are shown in Figure 10.1. The influence of each systematic uncertainty on $\mu_{WZ01j}$ and $\mu_{WZ2j}$ can be seen from the ranking plots in Figure 10.2. The ranking plots are generated after the nominal fit with four additional fits for each NP by fixing the NP $\theta$ to be up or down variation ($\pm\Delta\theta$ or $\pm\Delta\hat{\theta}$) of its best-fit value $\hat{\theta}$ based on the nominal fit while keeping the remaining NPs as their own best-fit values and calculating the corresponding pre-fit or post-fit changes in $\mu_{WZ01j}$ or $\mu_{WZ2j}$. The pre-fit impact is calculated with $\pm\Delta\theta = \pm1$ and the NP fixed to $\theta = \hat{\theta} \pm \Delta\theta$. The post-fit impact is calculated with $\pm\Delta\hat{\theta}$ as the uncertainty on the best fit value $\hat{\theta}$ and

the NP fixed to $\theta = \hat{\theta} \pm \Delta\hat{\theta}$. The dominate uncertainty on $\mu_{WZ01j}$ is from the $WZ01j$ parton shower uncertainty while the dominate uncertainty on $\mu_{WZ2j}$ is from the $WZ2j$ scale uncertainty.
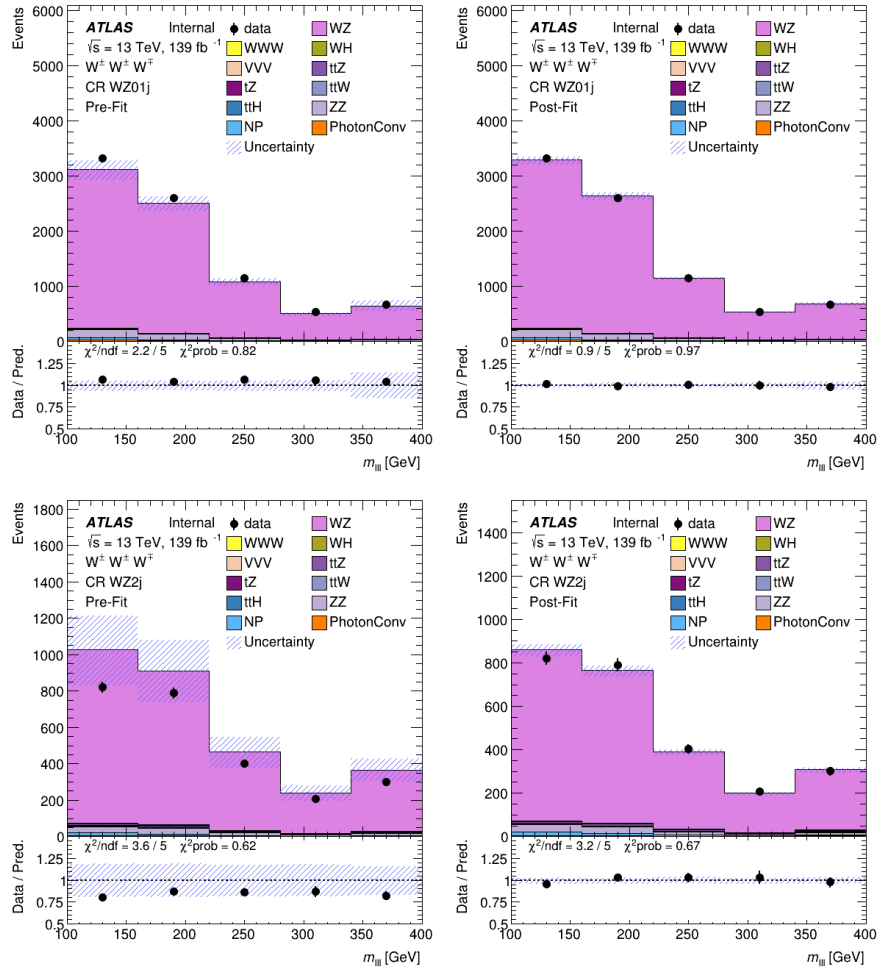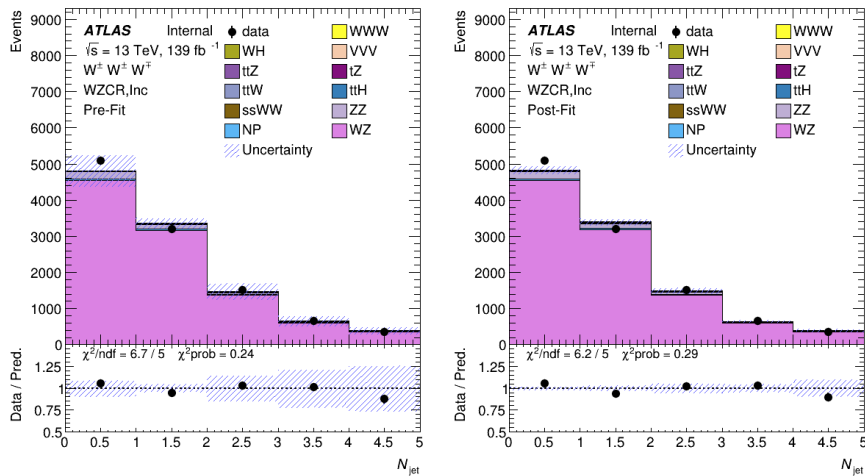


Figure 10.1: The pre-fit (left) and post-fit (right) $m_{\ell\ell\ell}$ distributions with statistical and systematical uncertainties in the $WZ01j$ CR (top) and the $WZ2j$ CR (bottom) in the $WZ$ normalization factor fit.
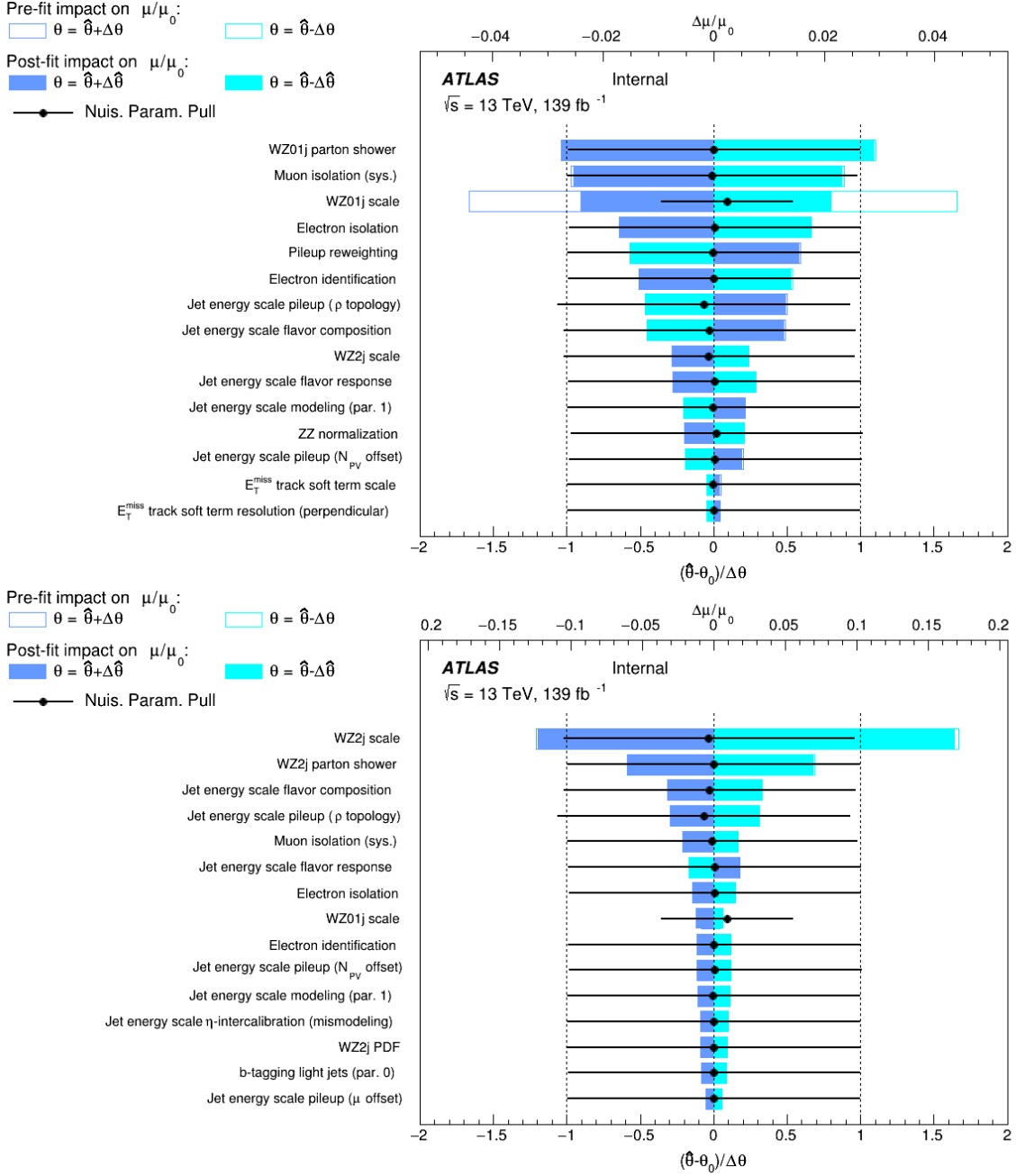
Figure 10.2: The rankings of NPs based on their post-fit impacts on $\mu_{WZ01j}$ (top) and $\mu_{WZ2j}$ (bottom). In each plot, the pull of each NP $(\hat{\theta} - \theta_0)/\Delta\theta$ is shown with black dot and solid black line as its error bar based on the lower axis. The impacts on $\Delta\mu/\mu_0$ are indicated based on the upper axis with empty dark (light) blue box for the pre-fit impact with up (down) variation of each NP and filled dark (light) blue box for the post-fit impact with up (down) variation of each NP.

## 10.3 Fit Results

In order to extract the signal strength and significance, a combined fit described in Section 10.1 is performed. The regions, channels, distributions, and number of bins in each distribution are summarized in Table 10.2. However, instead of unblinding data in the SRs, a fit with the Asimov data is performed first to extract the expected significance. The Asimov data is built from the expected background with the $WZ$ normalization factors determined in Section 10.2 and the expected signal yield with $\mu_s = 1$. Then the fit with real data is performed afterwards for the observed significance. Both the expected significance and the observed significance are shown in Table 10.3. The observed significance in the combined fit with both SRs is above $5\sigma$, so this is the first observation of the $W^\pm W^\pm W^\mp$ process. The fit is also performed in each individual decay channel for better comparison. The $\ell^\pm \nu \ell^\pm \nu jj$ SR and the $WZ$ CRs are used for the $\ell^\pm \nu \ell^\pm \nu jj$ decay channel fit while the $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SR and the $WZ$ CRs are used for the $\ell^\pm \nu \ell^\pm \nu jj$ decay channel fit. The channels, plots, and number of bins in each plot for the fits are the same as the combine fit.

| Region | Channel | Distribution | $N_{\text{bin}}$ |
|---|---|---|---|
| | $ee$ | $\ell^\pm \nu \ell^\pm \nu jj$ BDT discriminant score | 10 |
| $\ell^\pm \nu \ell^\pm \nu jj$ SR | $e\mu$ | $\ell^\pm \nu \ell^\pm \nu jj$ BDT discriminant score | 10 |
| | $\mu\mu$ | $\ell^\pm \nu \ell^\pm \nu jj$ BDT discriminant score | 10 |
| $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SR | Inclusive | $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ BDT discriminant score | 5 |
| $WZ01j$ CR | Inclusive | $m_{\ell\ell\ell}$ | 5 |
| $WZ2j$ CR | Inclusive | $m_{\ell\ell\ell}$ | 5 |

Table 10.2: Summary of regions, channels, distributions, and number of bins in each distribution used in the final combined fit for the signal strength and significance.

| Decay channel | Significance | |
|---|---|---|
| | Expected | Observed |
| $W^\pm W^\pm W^\mp \to \ell^\pm \nu \ell^\pm \nu jj$ | $4.2\sigma$ | $6.7\sigma$ |
| $W^\pm W^\pm W^\mp \to \ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ | $4.0\sigma$ | $4.5\sigma$ |
| $W^\pm W^\pm W^\mp$ combined | $5.7\sigma$ | $7.8\sigma$ |

Table 10.3: The expected and observed significance to reject background-only hypothesis for different decay channels of the $W^\pm W^\pm W^\mp$ signal.

The signal strengths are determined with different decay channels and the values are shown in Figure 10.3. The combined signal strength is found to be $1.58^{+0.25}_{-0.24}$. The $WZ$ normalization factors determined from the fits are shown in Figure 10.4 and they are applied in the post-fit figures

and tables while the *WZ* normalization factors determined from Section 10.2 are applied in the pre-fit figures and tables. The plot of the signal over background ratio is shown Figure 10.5. The influence of each systematic uncertainty on $\mu_{W^\pm W^\pm W^\mp}$ can be seen from the ranking plot in Figure 10.6. The dominate uncertainty comes from the muon non-prompt MC closure test. The pre-fit and post-fit distributions in the SRs are shown in Figures 10.8 and 10.9. Plots of other variables and distributions in each channel are attached in Appendices A.6 and A.7. The pre-fit and post-fit number of jets distributions in the inclusive *WZ* CR (combining *WZ01j* and *WZ2j* CRs) are shown in Figure 10.7 while the pre-fit and post-fit $m_{\ell\ell\ell}$ distributions in the *WZ* CRs are attached in Appendices A.1 and A.2. The pre-fit and post-ft event yield tables are shown in Tables 10.4 and 10.5. A good agreement between data and prediction is achieved in all regions.
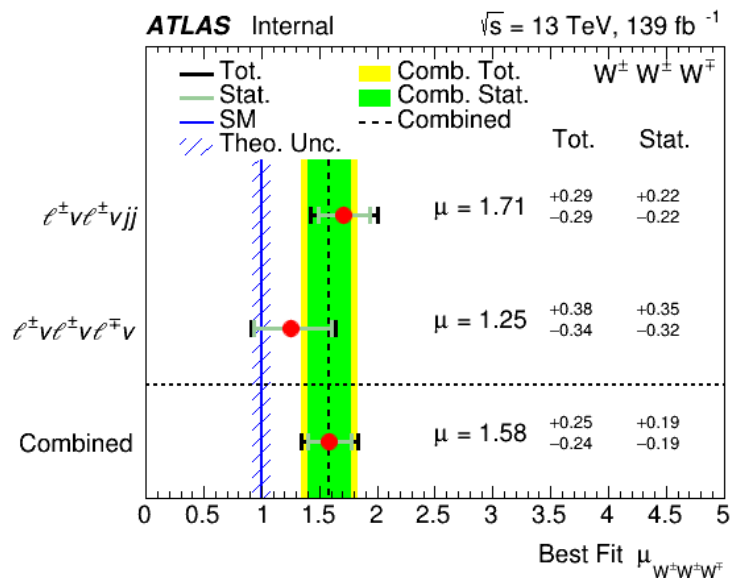


Figure 10.3: The signal strengths with total uncertainties (black solid line) and statistical only uncertainties (cyan solid line) in different decay channels and the combined channel. The SM prediction (blue solid line) is from the signal MC samples and the uncertainty (shaded blue box) only includes the combined PDF and scale uncertainty determined from Section 10.4. The best fit signal strength in the combined fit is also shown vertically (black dashed line) with its total uncertainty (yellow box) and statistical only uncertainty (green box) for better comparison.
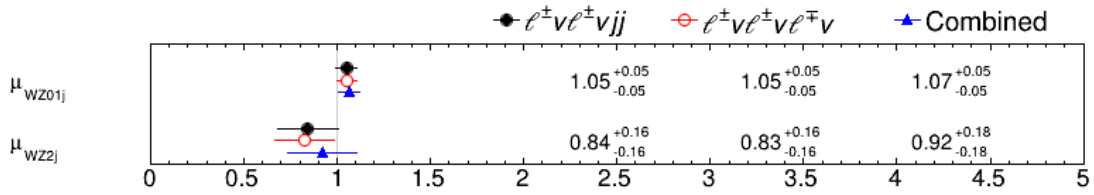
Figure 10.4: The *WZ* normalization factors determined from the fits in the $\ell^\pm \nu \ell^\pm \nu jj$ decay channel (black), the $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ decay channel (red) and the combined decay channel (blue).
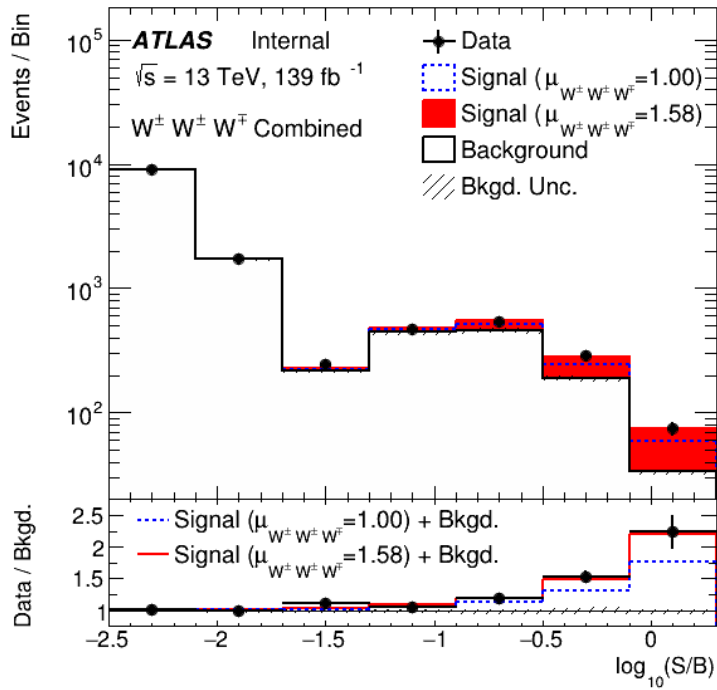


Figure 10.5: The post-fit plot of the signal over background ratio in logarithm for each bin in the plots used for the combined fit. Both expected (dashed blue box) and observed (filled red box) signal are shown with the post-fit background (black box) and data (black dot).
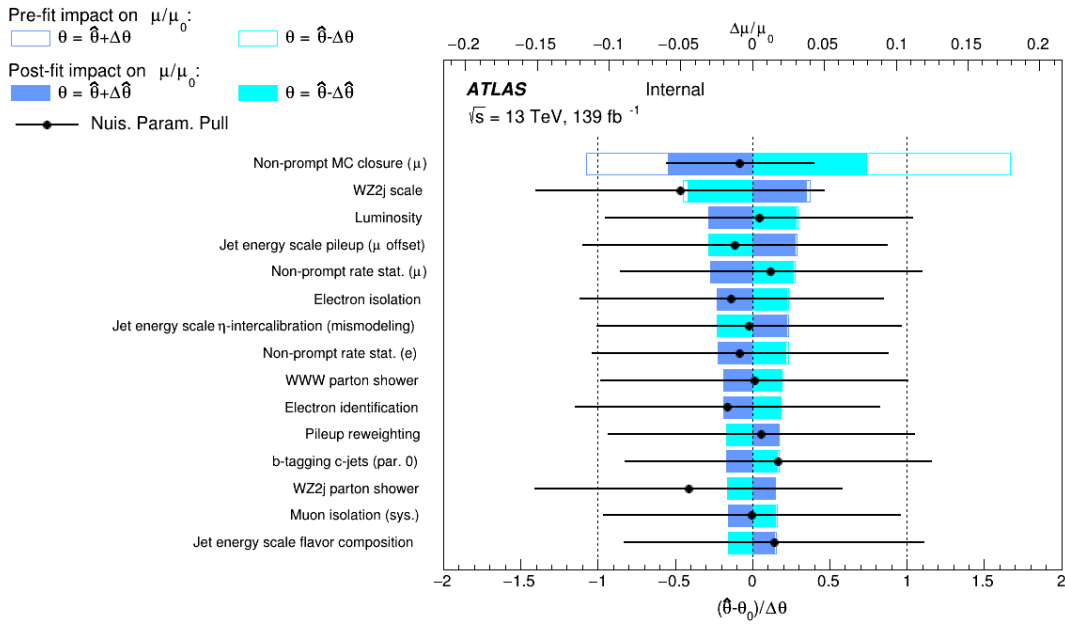
Figure 10.6: The rankings of NPs based on their post-fit impacts on the signal strength $\mu_{W^\pm W^\pm W^\mp}$. It can be read similarly as explained Figure 10.2.



Figure 10.7: The pre-fit (left) and post-fit (right) number of jets distributions with statistical and systematical uncertainties in the inclusive $WZ$ CR (combining $WZ01j$ and $WZ2j$ CRs) in the combined fit.
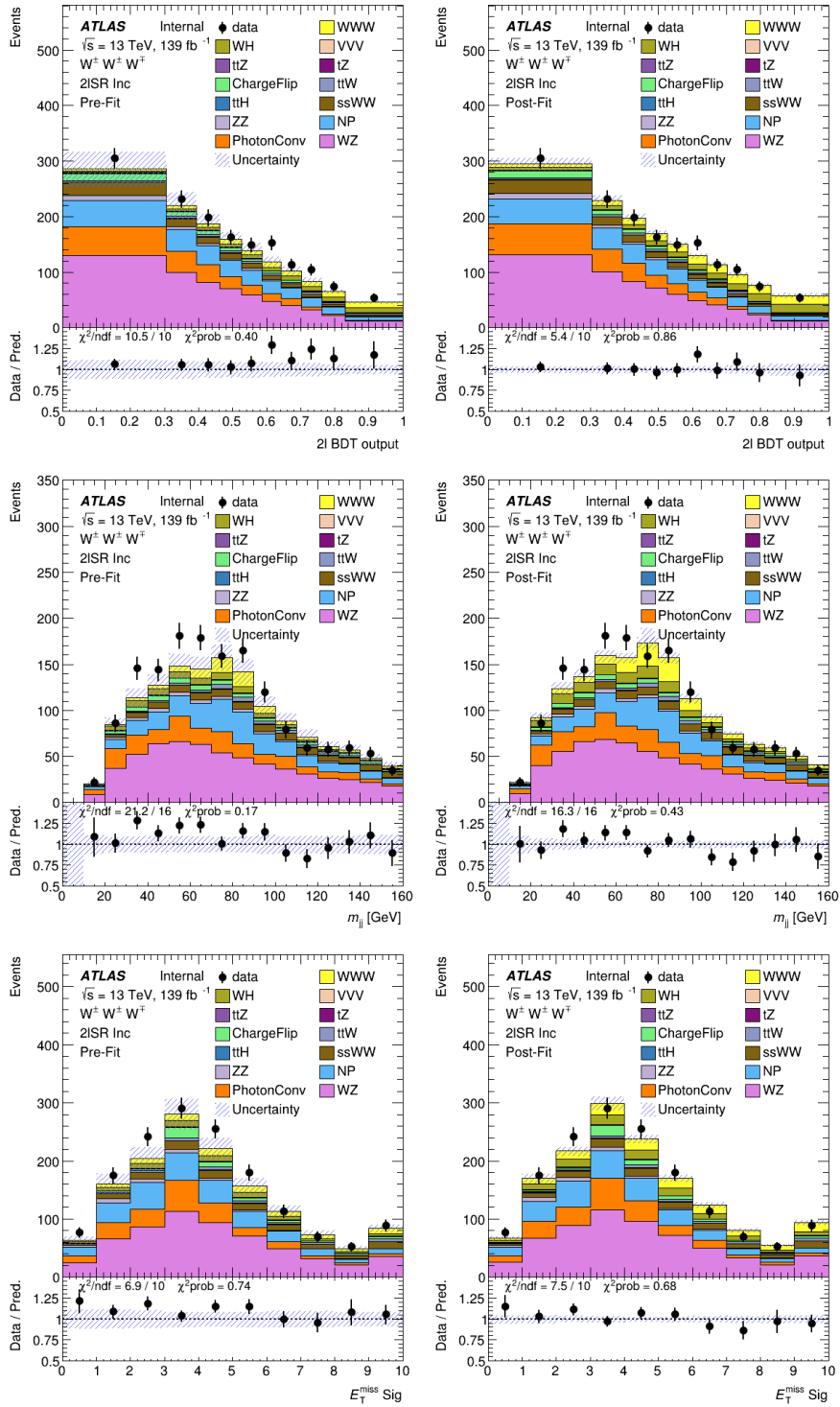
Figure 10.8: The pre-fit (left) and post-fit (right) distributions of the $\ell^\pm \nu \ell^\pm \nu jj$ BDT discriminant score (top), $m_{jj}$ (middle), and $E_T^{miss}$ significance (bottom) with statistical and systematical uncertainties in the $\ell^\pm \nu \ell^\pm \nu jj$ SR in the combined fit.
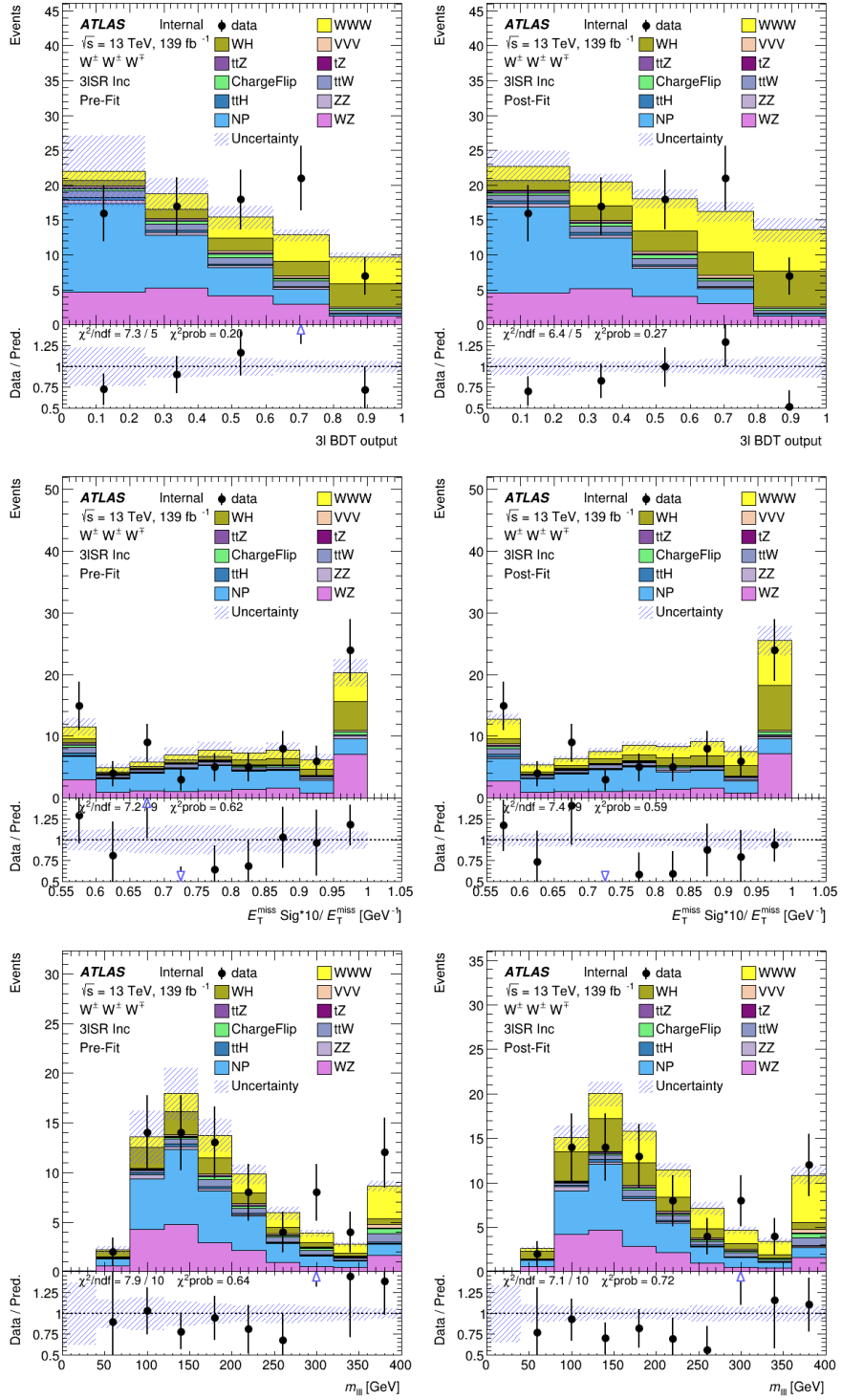
Figure 10.9: The pre-fit (left) and post-fit (right) distributions of the $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ BDT discriminant score (top), $E_T^{miss}$ significance ratio (middle), and $m_{\ell\ell\ell}$ (bottom) with statistical and systematical uncertainties in the $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ SR in the combined fit.

| | $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR | | | $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ SR |
| --- | --- | --- | --- | --- |
| | *ee* | *eμ* | *μμ* | |
| *WWW* | 10.74 ± 0.60 | 43.2 ± 1.9 | 28.1 ± 1.1 | 13.90 ± 0.84 |
| *WH* | 6.93 ± 0.51 | 33.4 ± 2.1 | 22.2 ± 1.4 | 9.56 ± 0.41 |
| *WZ* | 81 ± 13 | 342 ± 52 | 170 ± 28 | 18.3 ± 2.8 |
| NonPrompt | 17.8 ± 4.5 | 138 ± 38 | 97 ± 38 | 26.7 ± 8.6 |
| PhotonConv | 60 ± 13 | 139 ± 20 | - | - |
| SS *WW* | 11.9 ± 1.3 | 49.8 ± 5.3 | 32.3 ± 3.4 | - |
| ChargeFlip | 29.6 ± 6.4 | 18.3 ± 4.0 | - | 1.65 ± 0.37 |
| *ZZ* | 2.71 ± 0.42 | 20.2 ± 2.5 | 8.5 ± 1.1 | 1.48 ± 0.19 |
| *VVV* | 0.19 ± 0.03 | 0.54 ± 0.06 | 0.22 ± 0.03 | 1.23 ± 0.14 |
| $t\bar{t}W$ | 4.47 ± 0.65 | 17.1 ± 2.4 | 10.1 ± 1.4 | 3.91 ± 0.55 |
| $t\bar{t}Z$ | 0.80 ± 0.13 | 2.71 ± 0.40 | 1.45 ± 0.22 | 0.76 ± 0.12 |
| $t\bar{t}H$ | 0.68 ± 0.08 | 2.83 ± 0.28 | 1.74 ± 0.18 | 1.19 ± 0.11 |
| *tZ* | 0.34 ± 0.08 | 1.16 ± 0.19 | 0.41 ± 0.10 | 0.04 ± 0.02 |
| Total | 227 ± 20 | 808 ± 70 | 373 ± 48 | 78.8 ± 9.3 |
| Data | 242 | 885 | 418 | 79 |

Table 10.4: The pre-fit event yields with statistical and systematical uncertainties in the SRs.

| | $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR | | | $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ SR |
| --- | --- | --- | --- | --- |
| | *ee* | *eμ* | *μμ* | |
| *WWW* | 17.0 ± 2.6 | 69 ± 10 | 44.7 ± 6.8 | 21.9 ± 3.4 |
| *WH* | 11.1 ± 1.8 | 53.5 ± 8.6 | 35.8 ± 5.8 | 15.0 ± 2.3 |
| *WZ* | 82.3 ± 5.5 | 351 ± 22 | 174.1 ± 9.6 | 18.1 ± 1.4 |
| NonPrompt | 17.4 ± 4.0 | 135 ± 19 | 96 ± 17 | 26.0 ± 4.2 |
| PhotonConv | 62.4 ± 8.6 | 144 ± 15 | - | - |
| SS *WW* | 12.1 ± 1.4 | 51.1 ± 5.3 | 33.3 ± 3.4 | - |
| ChargeFlip | 30.0 ± 6.1 | 18.6 ± 3.8 | - | 1.67 ± 0.34 |
| *ZZ* | 2.76 ± 0.40 | 20.6 ± 2.5 | 8.7 ± 1.1 | 1.47 ± 0.16 |
| *VVV* | 0.19 ± 0.02 | 0.54 ± 0.06 | 0.23 ± 0.03 | 1.22 ± 0.13 |
| $t\bar{t}W$ | 4.44 ± 0.62 | 17.1 ± 2.3 | 10.1 ± 1.3 | 3.88 ± 0.52 |
| $t\bar{t}Z$ | 0.80 ± 0.11 | 2.70 ± 0.37 | 1.45 ± 0.20 | 0.75 ± 0.10 |
| $t\bar{t}H$ | 0.67 ± 0.07 | 2.81 ± 0.27 | 1.74 ± 0.17 | 1.18 ± 0.11 |
| *tZ* | 0.34 ± 0.05 | 1.18 ± 0.16 | 0.42 ± 0.07 | 0.036 ± 0.005 |
| Total | 242 ± 10 | 867 ± 22 | 406 ± 15 | 91.2 ± 5.7 |
| Data | 242 | 885 | 418 | 79 |

Table 10.5: The post-fit event yields with statistical and systematical uncertainties in the SRs.

## 10.4 Signal Cross Section Measurement

The total cross section of the $W^\pm W^\pm W^\mp$ signal MC samples is $0.511 \pm 0.039$ pb. The uncertainty is the combined PDF and scale uncertainty on the inclusive $WWW$ and $WH$ cross section determined from the same prescription described in Section 9.3. Note that the uncertainty is not included in the measured cross section since the theoretical uncertainty on the predicted cross section does not affect the actual cross section measurement. As mentioned in Section 10.3, the signal strength is determined as:

$$\mu_{W^\pm W^\pm W^\mp} = 1.58^{+0.25}_{-0.24}. \tag{10.10}$$

The measured cross section for the $W^\pm W^\pm W^\mp$ signal can then be determined from Eq. 10.1:

$$\sigma^{\text{Obs}}_{W^\pm W^\pm W^\mp} = 0.807^{+0.128}_{-0.123} \text{ pb}. \tag{10.11}$$

From the SM $WWW$ cross section based on the theory paper [127] and the SM $WH$ cross section based on the CERN report [128], the SM inclusive $W^\pm W^\pm W^\mp$ cross section is:

$$\sigma^{\text{SM}}_{W^\pm W^\pm W^\mp} = 0.505 \pm 0.016 \text{ pb}, \tag{10.12}$$

where the uncertainty includes the PDF and scale uncertainties. Therefore, the measured cross section is compatible with the SM prediction within $2.5\sigma$.

# CHAPTER 11

# Conclusion

This thesis presents the first observation of inclusive $W^\pm W^\pm W^\mp$ production using 139 fb$^{-1}$ of $pp$ collision data collected by the ATLAS detector at $\sqrt{s} = 13$ TeV. The background-only hypothesis is rejected with an observed (expected) significance of $7.8\sigma$ ($5.7\sigma$). This $W^\pm W^\pm W^\mp$ analysis currently is still an ongoing study at ATLAS and will be finalized with official results in a few months. It is interesting to note that the productions of a single $W$ boson, two $W$ bosons, and three $W$ bosons are all first observed at CERN: Underground Area 1 (UA1) and Underground Area 2 (UA2) discovered the $W$ boson at the SPS in 1983 [129, 130]; Apparatus for LEP Physics (ALEPH), Detector with Lepton, Photon and Hadron Identification (DELPHI), L3, and Omni Purpose Apparatus for LEP (OPAL) observed the $W$ boson pair production for the first time at the Large Electron-Positron Collider (LEP) in 1996 [131–134]; ATLAS observed the production of three $W$ bosons for the first time at the LHC in 2021, which is covered in this thesis.

The analysis is performed in two decay channels of the inclusive $W^\pm W^\pm W^\mp$ production: $W^\pm W^\pm W^\mp \rightarrow \ell^\pm \nu \ell^\pm \nu jj$ and $W^\pm W^\pm W^\mp \rightarrow \ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$. In the $\ell^\pm \nu \ell^\pm \nu jj$ SR, $1,545$ events are observed in data with approximately 231 signal and $1,284$ background events expected. In the $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SR, 79 events are observed in data with 37 signal and 54 background events expected. The total inclusive $W^\pm W^\pm W^\mp$ cross section is measured to be $0.807^{+0.128}_{-0.123}$ pb, which is found to be compatible with the SM prediction within $2.5\sigma$.

However, the current measurement on the inclusive $W^\pm W^\pm W^\mp$ production cross section is still limited by statistics and a more accurate measurement on the inclusive $W^\pm W^\pm W^\mp$ production can be performed in the future with the Run 3 data from the ATLAS detector. The SM is not perfect and a lot of BSM theories are proposed. With precise measurement on the $W^\pm W^\pm W^\mp$ production, new physics with anomalous $WWWW$ QGC, such as the process with the extra-dimensional Kaluza-Klein excitations of SM electroweak gauge bosons and the radion mentioned in Ref. [135], can be tested with the effective field theory described in Section 2.10, since the $W^\pm W^\pm W^\mp$ production is very sensitive to the anomalous $WWWW$ QGC.

# APPENDIX A

# Supplementary Figures

Due to the length limitation, not all figures are placed in the main content of this thesis. Instead, the remaining figures are put in this appendix and grouped by regions.

## A.1  *WZ01j* Control Region
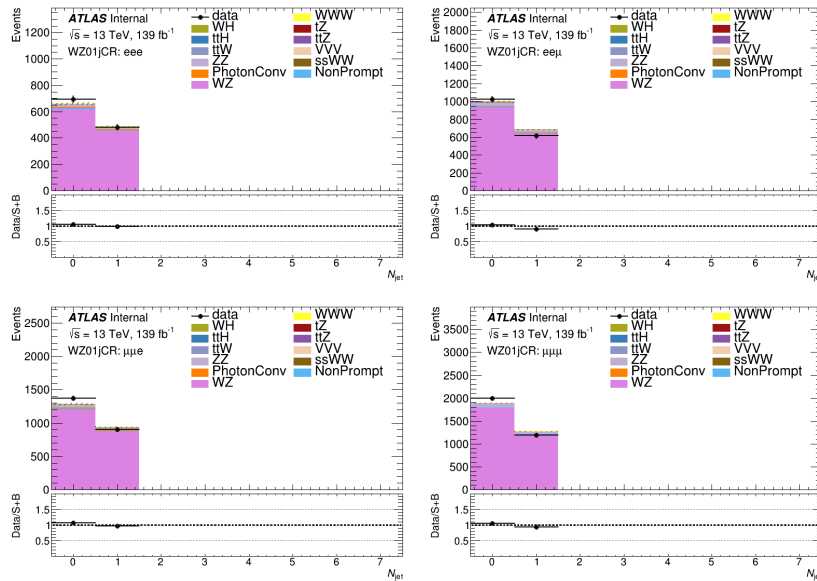


Figure A.1: The number of jets distributions with statistical uncertainties in *eee* (top left), *eeμ* (top right), *μμe* (bottom left), and *μμμ* (bottom right) channels in the *WZ01j* CR. The normalization scale factors for the *WZ* process are applied in these plots.
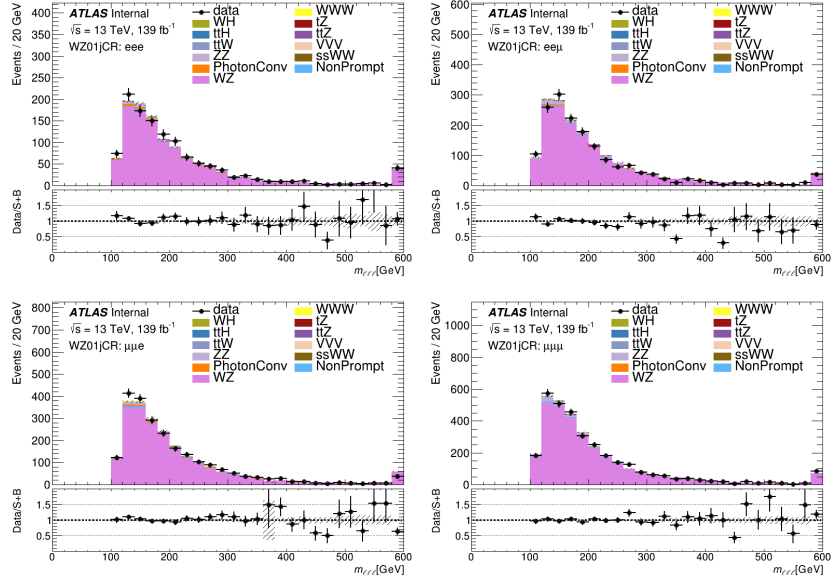
Figure A.2: The $m_{\ell\ell\ell}$ distributions with statistical uncertainties in *eee* (top left), *eeµ* (top right), *µµe* (bottom left), and *µµµ* (bottom right) channels in the *WZ*01*j* CR. The normalization scale factors for the *WZ* process are applied in these plots.
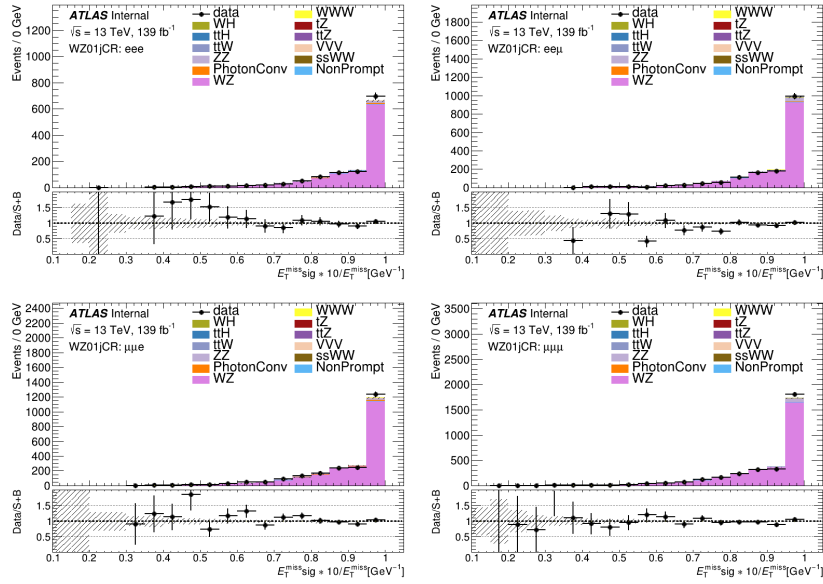


Figure A.3: The $E_T^{\mathrm{miss}}$ significance ratio distributions with statistical uncertainties in *eee* (top left), *eeµ* (top right), *µµe* (bottom left), and *µµµ* (bottom right) channels in the *WZ*01*j* CR. The normalization scale factors for the *WZ* process are applied in these plots.
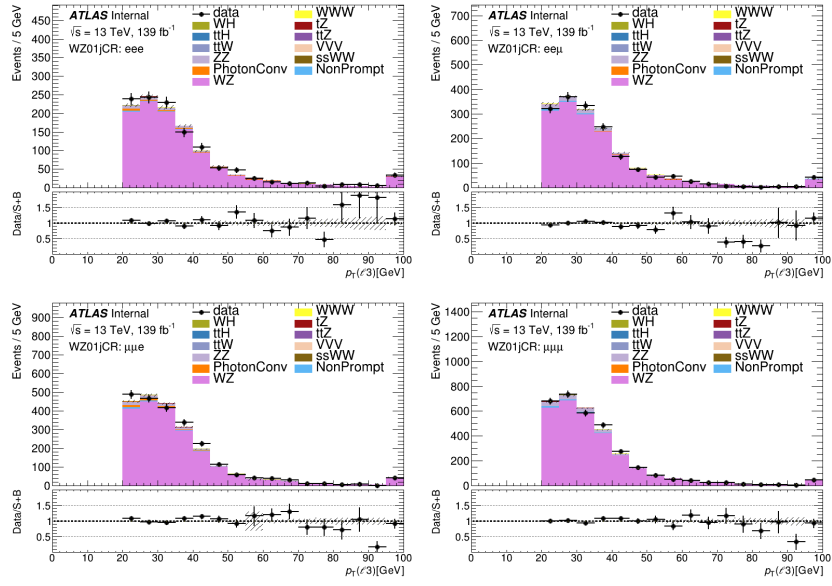
Figure A.4: The third lepton $p_T$ distributions with statistical uncertainties in *eee* (top left), *eeμ* (top right), *μμe* (bottom left), and *μμμ* (bottom right) channels in the *WZ*01*j* CR. The normalization scale factors for the *WZ* process are applied in these plots.
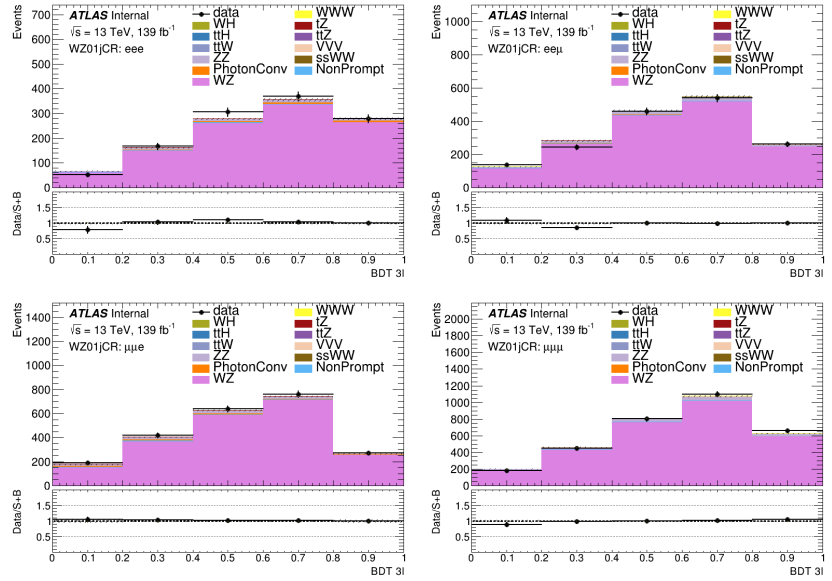


Figure A.5: The discriminant score distributions with statistical uncertainties based on the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ BDT in *eee* (top left), *eeμ* (top right), *μμe* (bottom left), and *μμμ* (bottom right) channels in the *WZ*01*j* CR. The normalization scale factors for the *WZ* process are applied in these plots.
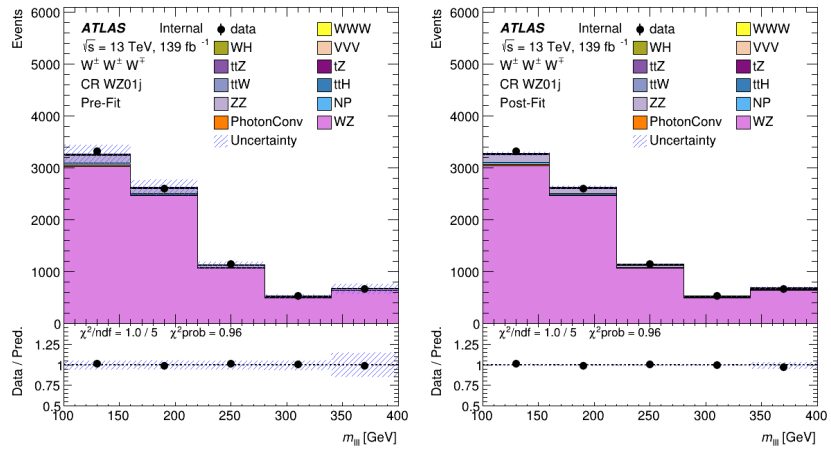
Figure A.6: The pre-fit (left) and post-fit (right) $m_{\ell\ell\ell}$ distributions with statistical and systematical uncertainties in the $WZ01j$ CR in the combined fit.

# A.2 *WZ2j* Control Region
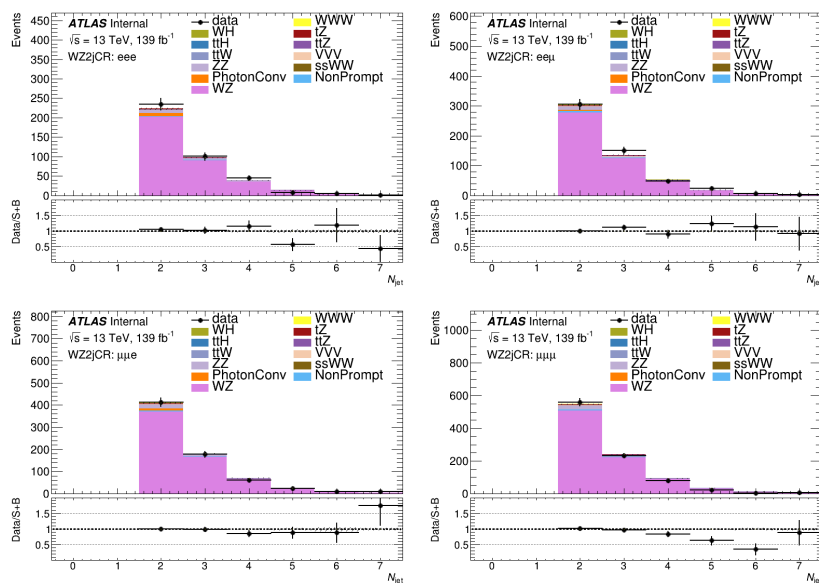


Figure A.7: The number of jets distributions with statistical uncertainties in *eee* (top left), *eeμ* (top right), *μμe* (bottom left), and *μμμ* (bottom right) channels in the *WZ2j* CR. The normalization scale factors for the *WZ* process are applied in these plots.
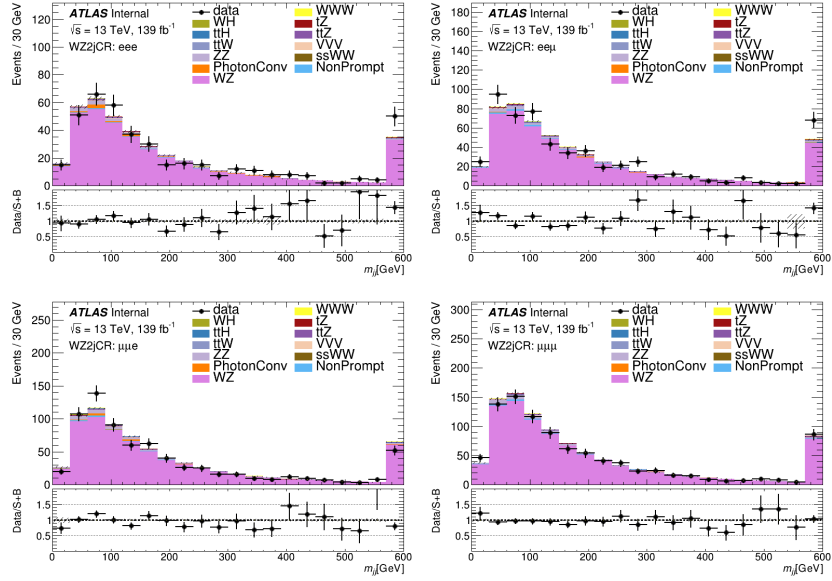
Figure A.8: The $m_{jj}$ distributions with statistical uncertainties in *eee* (top left), *eeµ* (top right), *µµe* (bottom left), and *µµµ* (bottom right) channels in the *WZ2j* CR. The normalization scale factors for the *WZ* process are applied in these plots.
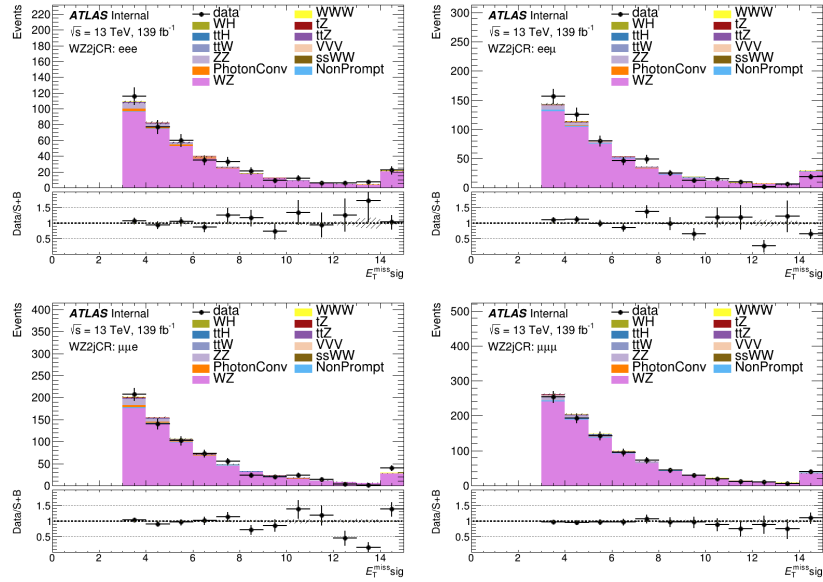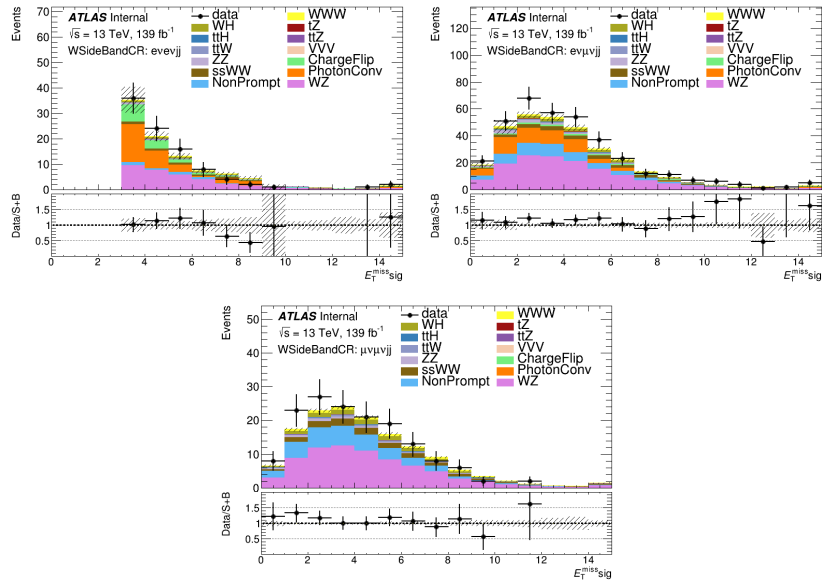


Figure A.9: The $E_T^{miss}$ significance distributions with statistical uncertainties in *eee* (top left), *eeµ* (top right), *µµe* (bottom left), and *µµµ* (bottom right) channels in the *WZ2j* CR. The normalization scale factors for the *WZ* process are applied in these plots.
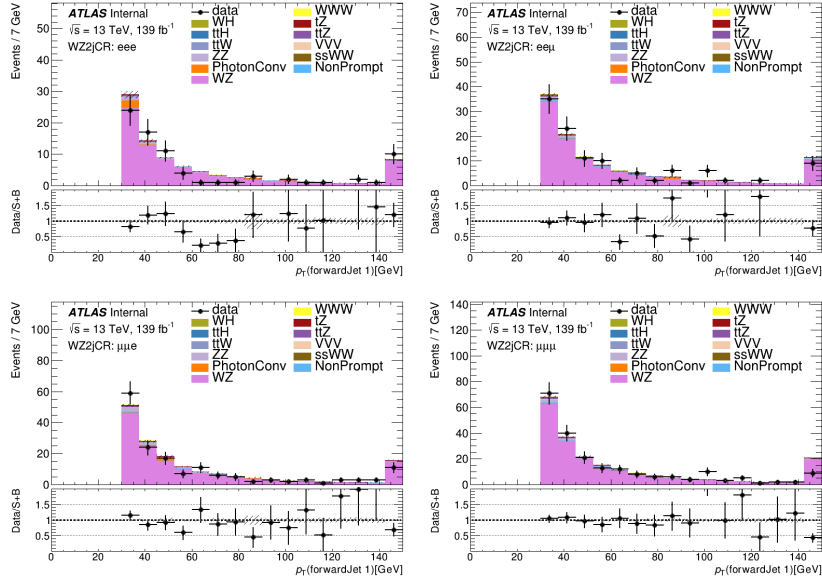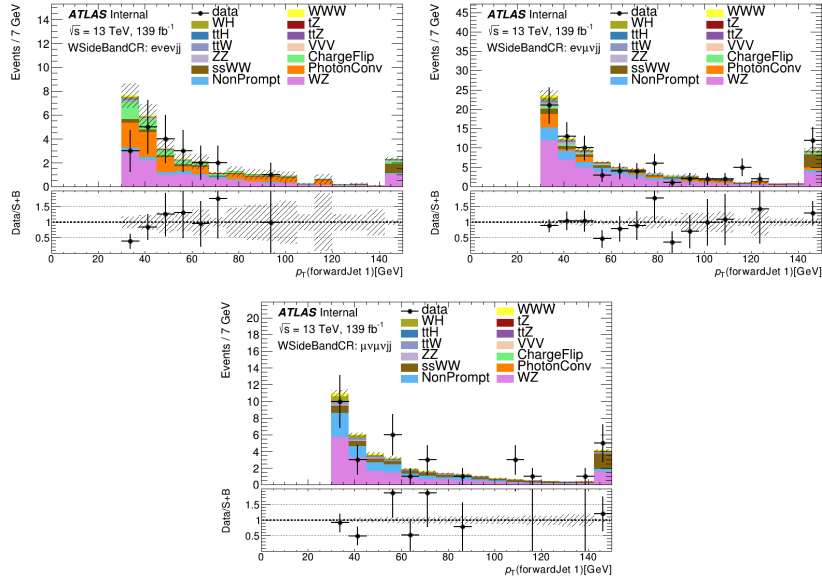
Figure A.10: The distributions of the leading forward jet $p_T$ with statistical uncertainties in *eee* (top left), *eeμ* (top right), *μμe* (bottom left), and *μμμ* (bottom right) channels in the *WZ2j* CR. The normalization scale factors for the *WZ* process are applied in these plots.
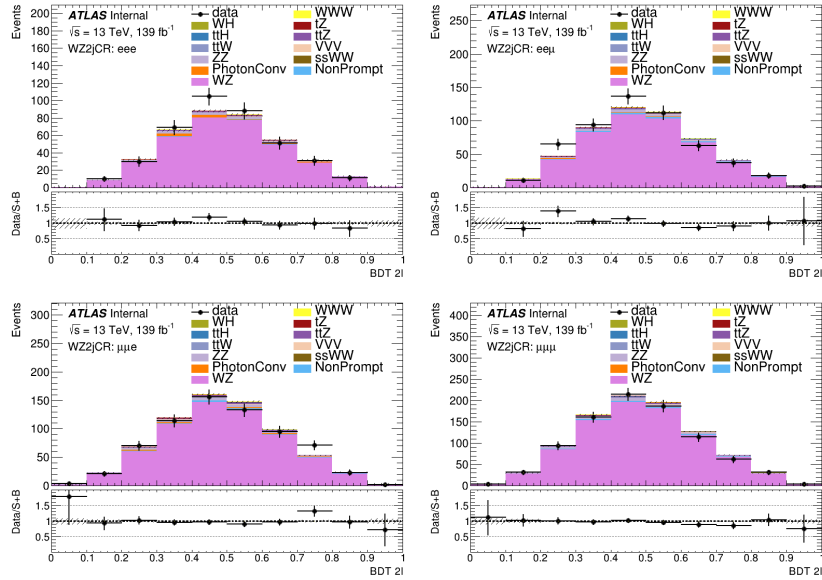


Figure A.11: The discriminant score distributions with statistical uncertainties based on the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ BDT in *eee* (top left), *eeμ* (top right), *μμe* (bottom left), and *μμμ* (bottom right) channels in the *WZ2j* CR. The normalization scale factors for the *WZ* process are applied in these plots.
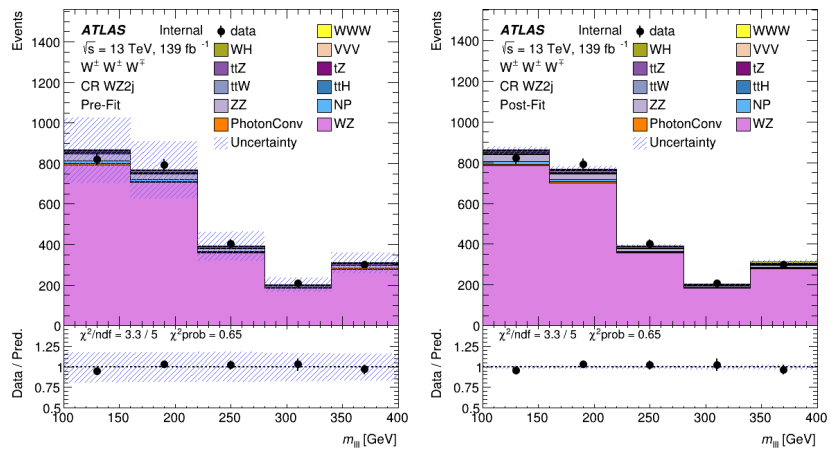
Figure A.12: The pre-fit (left) and post-fit (right) $m_{\ell\ell\ell}$ distributions with statistical and systematical uncertainties in the the $WZ2j$ CR in the combined fit.

# A.3 $\ell^\pm \nu \ell^\pm \nu j j$ $b$-tagging Control Region
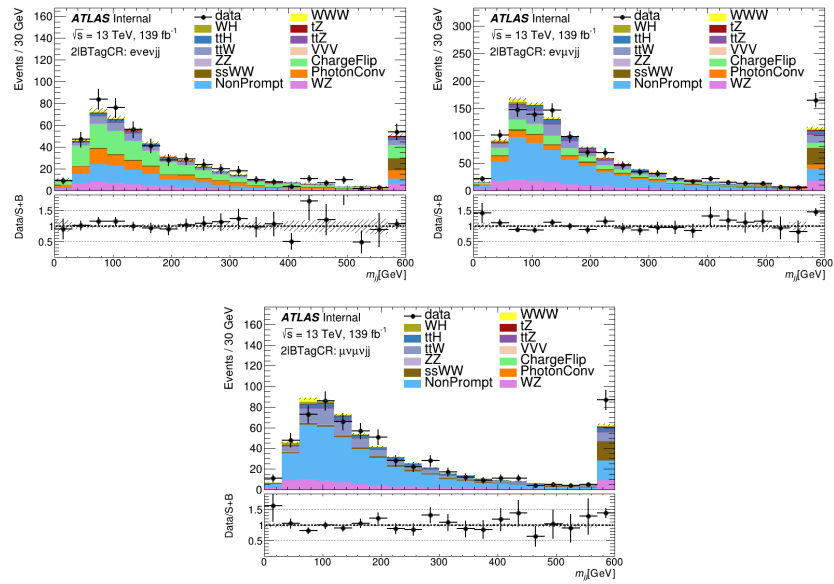


Figure A.13: The $m_{jj}$ distributions with statistical uncertainties in *ee* (top left), *eμ* (top right), and *μμ* (bottom) channels in the $\ell^\pm \nu \ell^\pm \nu j j$ $b$-tagging CR.
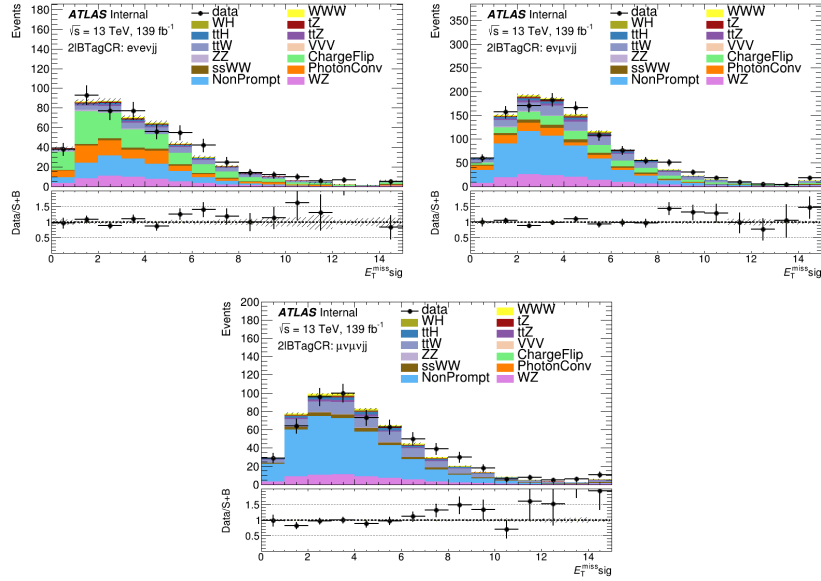
Figure A.14: The $E_T^{miss}$ significance distributions with statistical uncertainties in *ee* (top left), *eµ* (top right), and *µµ* (bottom) channels in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ *b*-tagging CR.
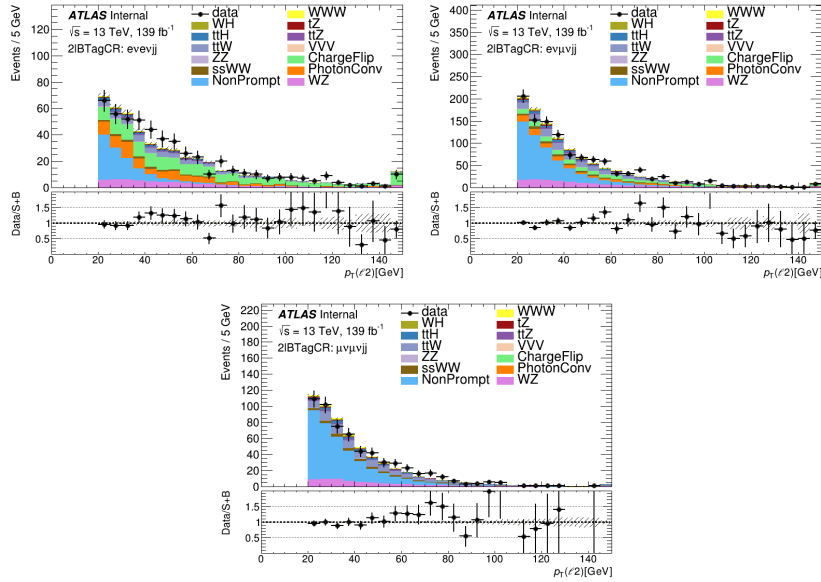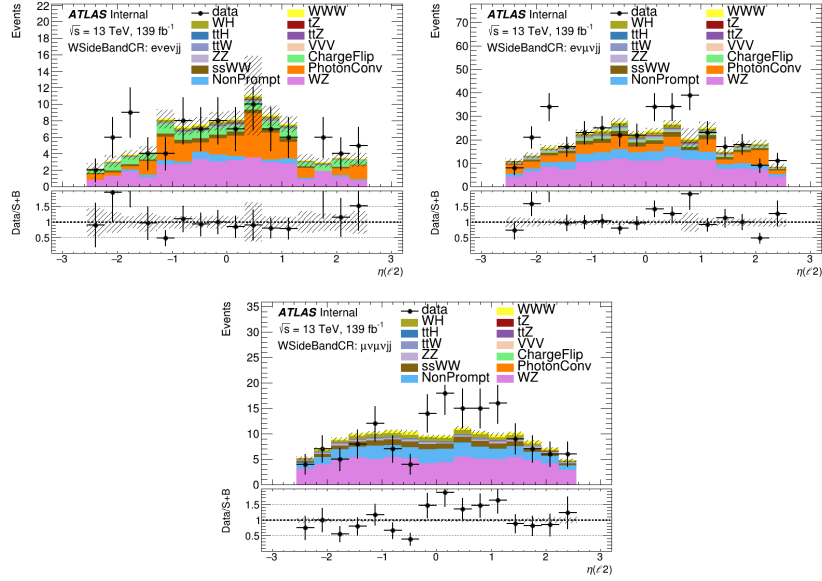


Figure A.15: The subleading lepton $p_T$ distributions with statistical uncertainties in *ee* (top left), *eµ* (top right), and *µµ* (bottom) channels in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ *b*-tagging CR.
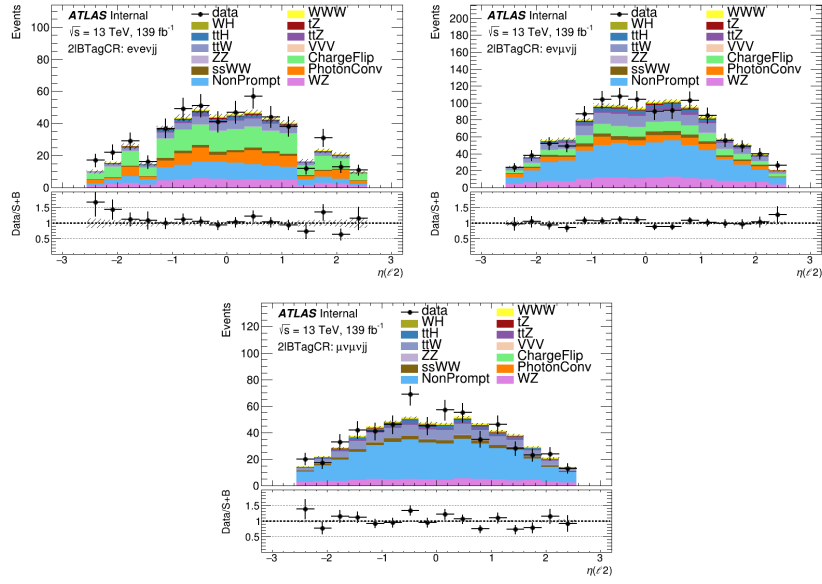
Figure A.16: The subleading lepton $\eta$ distributions with statistical uncertainties in *ee* (top left), *eμ* (top right), and *μμ* (bottom) channels in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ *b*-tagging CR.
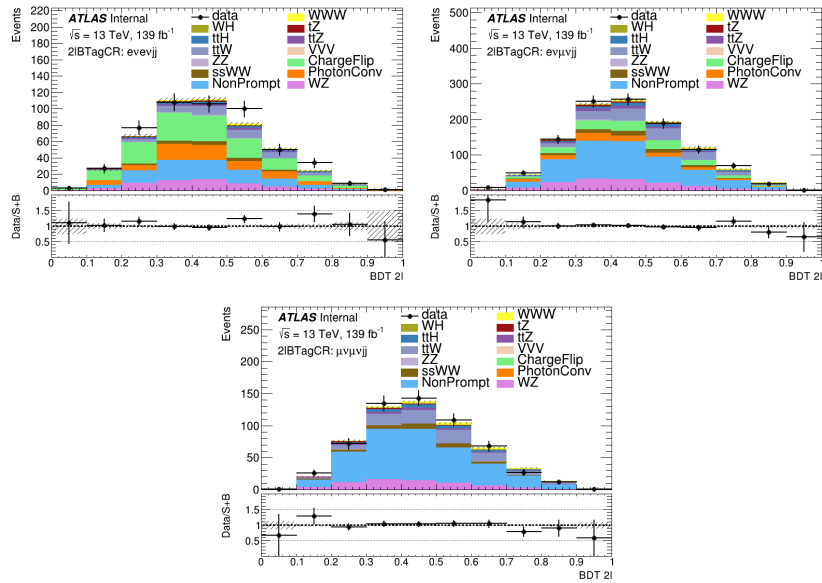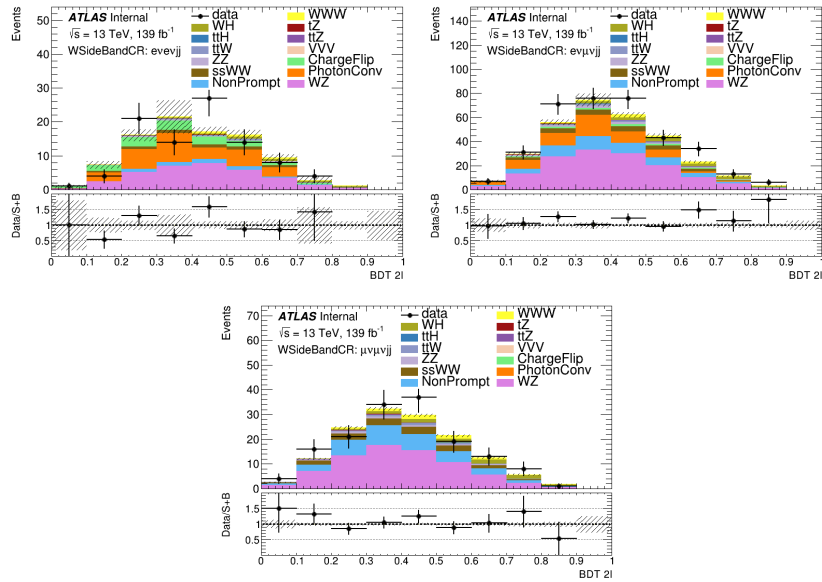


Figure A.17: The discriminant score distributions with statistical uncertainties based on the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ BDT in *ee* (top left), *eμ* (top right), and *μμ* (bottom) channels in the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ *b*-tagging CR.

# A.4 $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ *b*-tagging Control Region
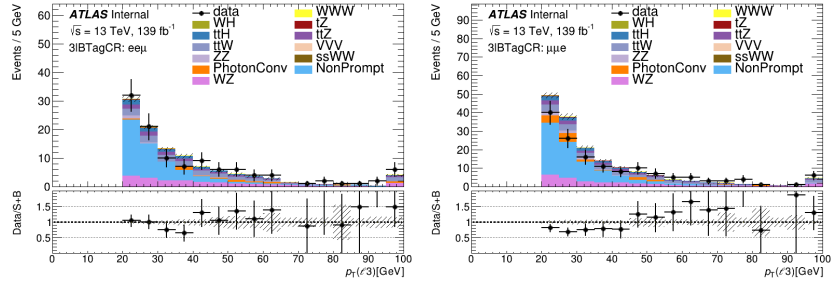


Figure A.18: The third lepton $p_{\mathrm{T}}$ distributions with statistical uncertainties in *eeμ* (left) and *μμe* (right) channels in the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ *b*-tagging CR.
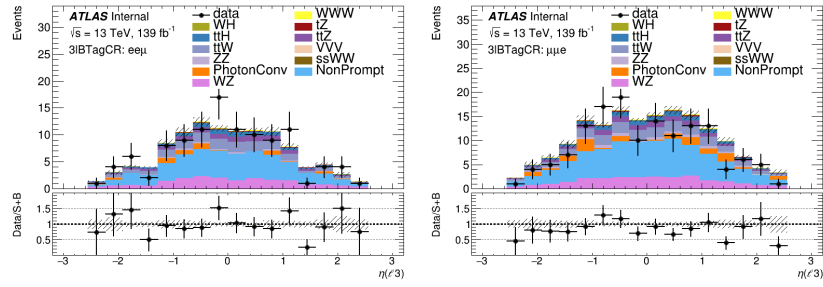


Figure A.19: The third lepton $\eta$ distributions with statistical uncertainties in *eeμ* (left) and *μμe* (right) channels in the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ *b*-tagging CR.
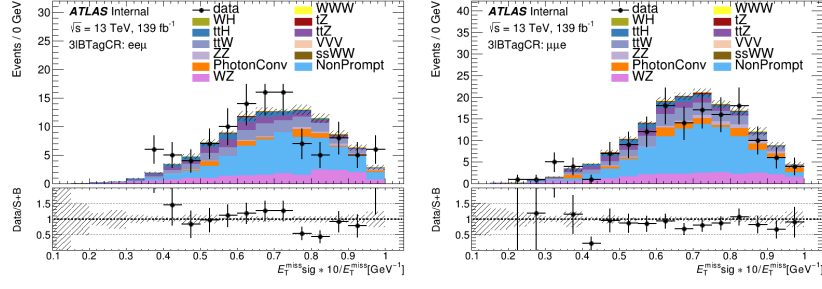
Figure A.20: The $E_T^{miss}$ significance ratio distributions with statistical uncertainties in *eeμ* (left) and *μμe* (right) channels in the $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ *b*-tagging CR.
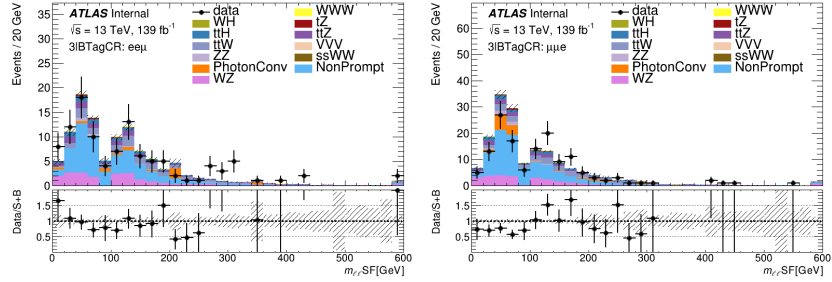


Figure A.21: The $m_{\ell\ell}$ of the SF lepton pair distributions with statistical uncertainties in *eeμ* (left) and *μμe* (right) channels in the $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ *b*-tagging CR.
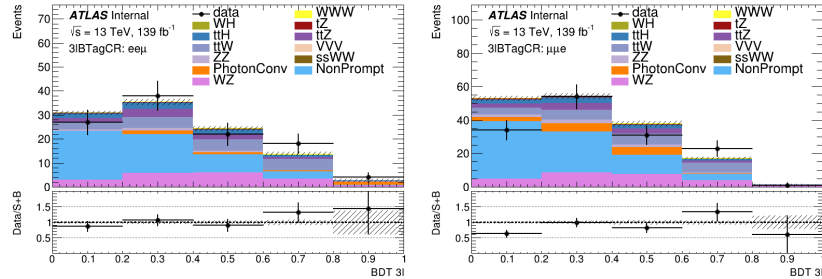


Figure A.22: The discriminant score distributions with statistical uncertainties based on the $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ BDT in *eeμ* (left) and *μμe* (right) channels in the $\ell^\pm\nu\ell^\pm\nu\ell^\mp\nu$ *b*-tagging CR.

## A.5  *W* Sideband Control Region
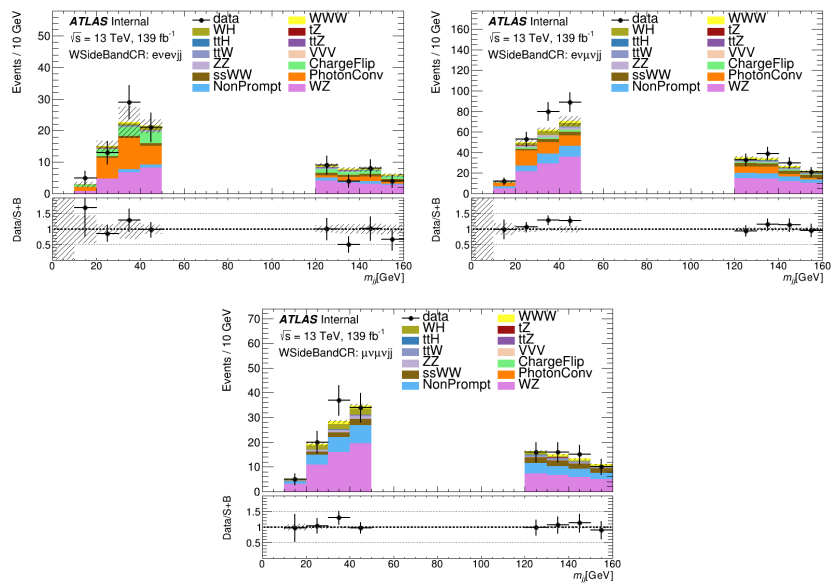


Figure A.23: The $m_{jj}$ distributions with statistical uncertainties in *ee* (top left), *eμ* (top right), and *μμ* (bottom) channels in the *W* sideband CR.
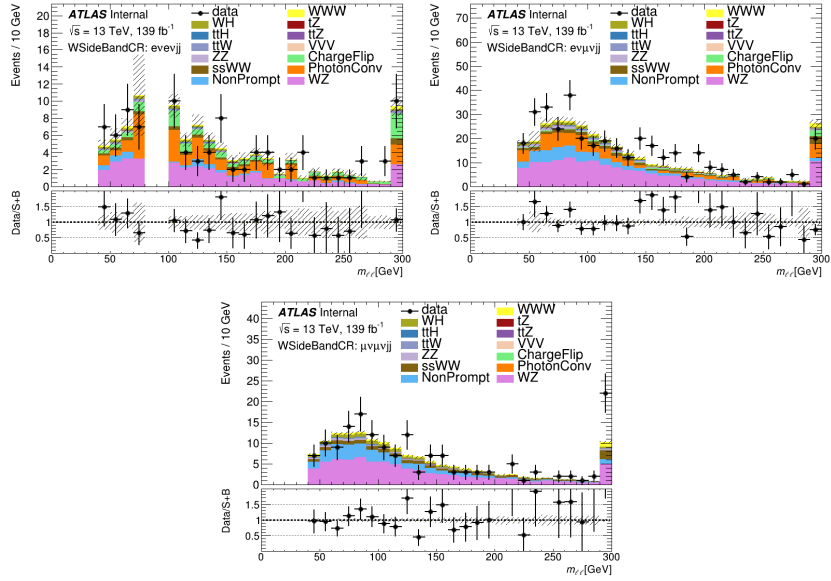
Figure A.24: The $m_{\ell\ell}$ distributions with statistical uncertainties in *ee* (top left), *eμ* (top right), and *μμ* (bottom) channels in the *W* sideband CR.



Figure A.25: The $E_{\mathrm{T}}^{\mathrm{miss}}$ significance distributions with statistical uncertainties in *ee* (top left), *eμ* (top right), and *μμ* (bottom) channels in the *W* sideband CR.

Figure A.26: The distributions of the leading forward jet $p_T$ with statistical uncertainties in *ee* (top left), *eμ* (top right), and *μμ* (bottom) channels in the *W* sideband CR.
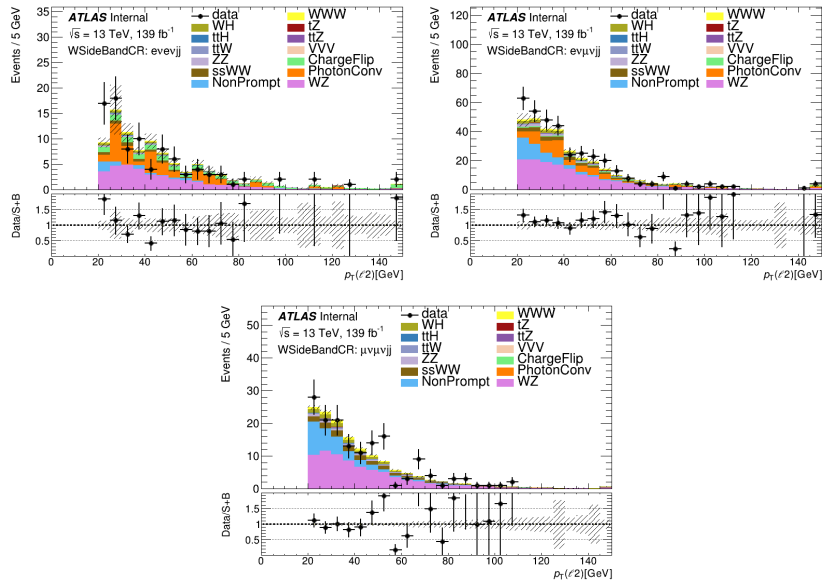


Figure A.27: The subleading lepton $p_T$ distributions with statistical uncertainties in *ee* (top left), *eμ* (top right), and *μμ* (bottom) channels in the *W* sideband CR.

Figure A.28: The subleading lepton $\eta$ distributions with statistical uncertainties in *ee* (top left), *eμ* (top right), and *μμ* (bottom) channels in the *W* sideband CR.



Figure A.29: The discriminant score distributions with statistical uncertainties based on the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ BDT in *ee* (top left), *eμ* (top right), and *μμ* (bottom) channels in the *W* sideband CR.

# A.6 $\ell^\pm\nu\ell^\pm\nu jj$ Signal Region
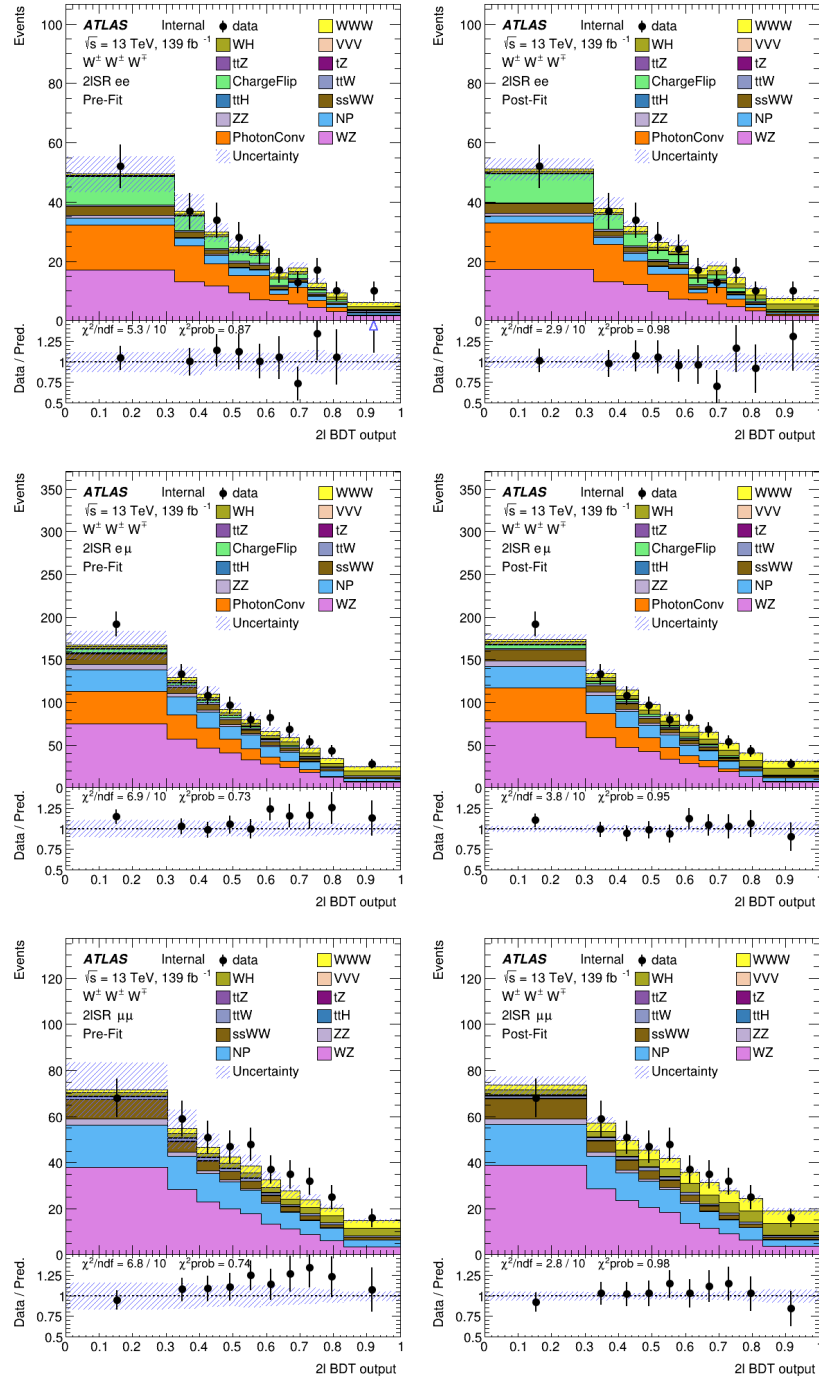


Figure A.30: The pre-fit (left) and post-fit (right) distributions of the $\ell^\pm\nu\ell^\pm\nu jj$ BDT discriminant score with statistical and systematical uncertainties in the *ee* (top), *eμ* (middle), and *μμ* (bottom) channels of the $\ell^\pm\nu\ell^\pm\nu jj$ SR in the combined fit.
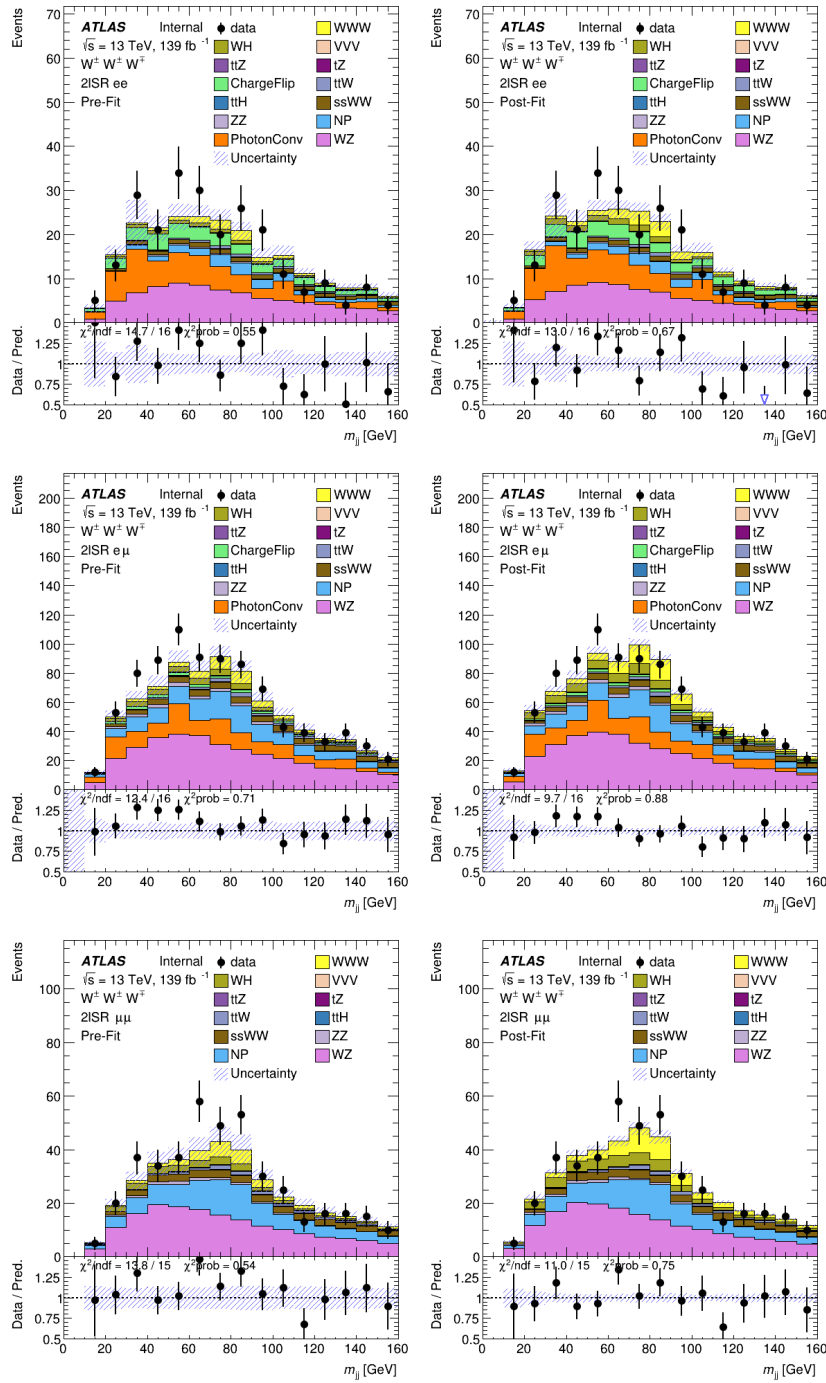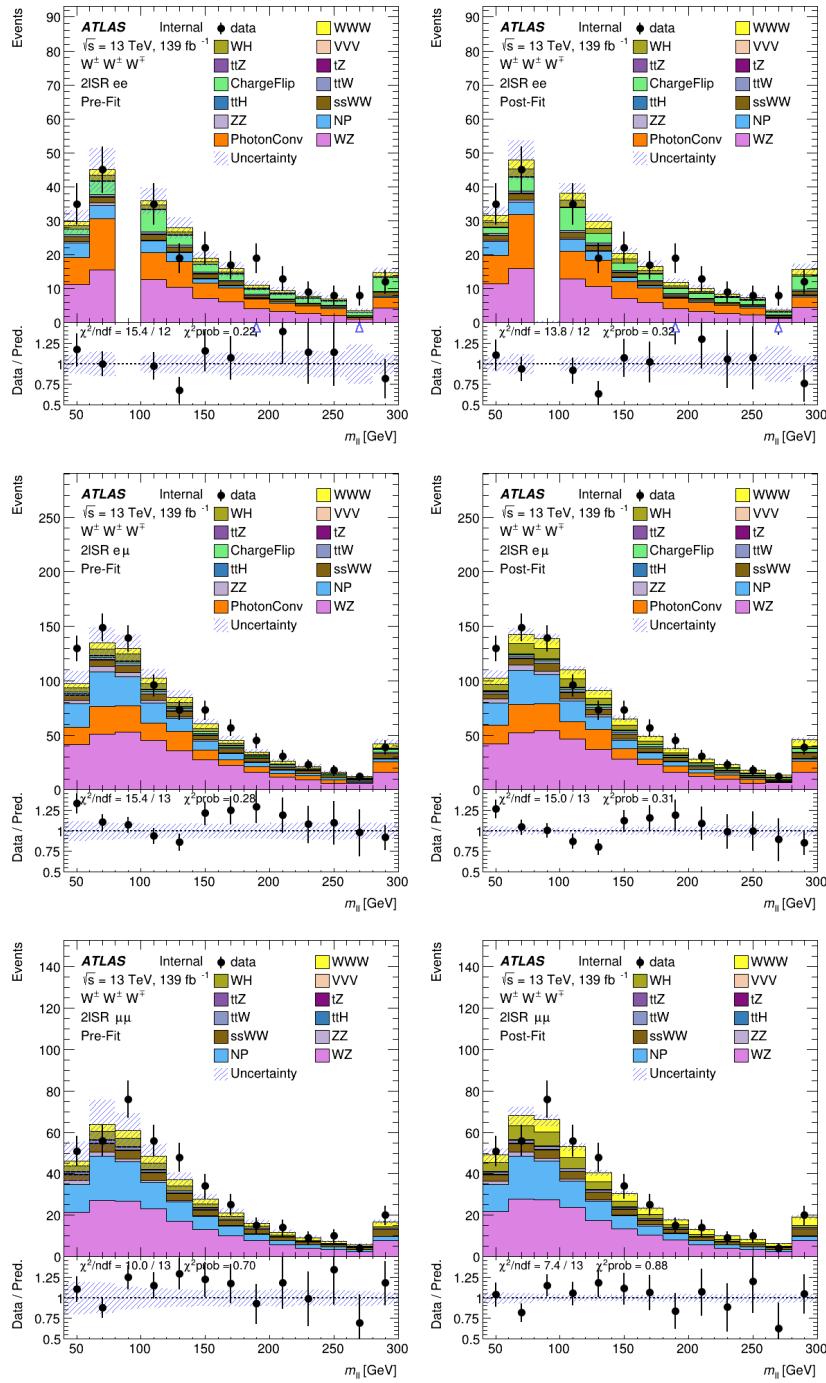
Figure A.31: The pre-fit (left) and post-fit (right) $m_{jj}$ distributions with statistical and systematical uncertainties in the *ee* (top), *eμ* (middle), and *μμ* (bottom) channels of the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR in the combined fit.
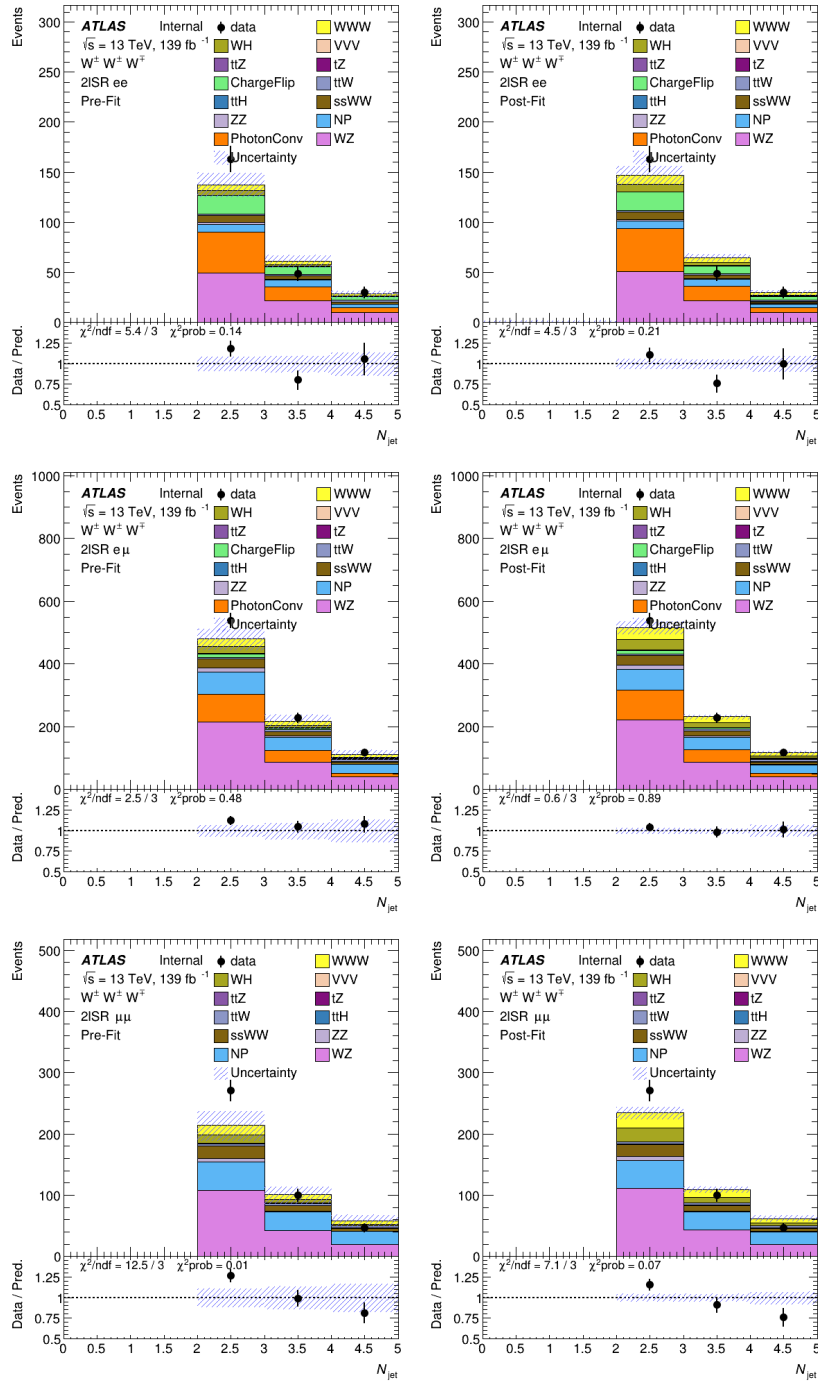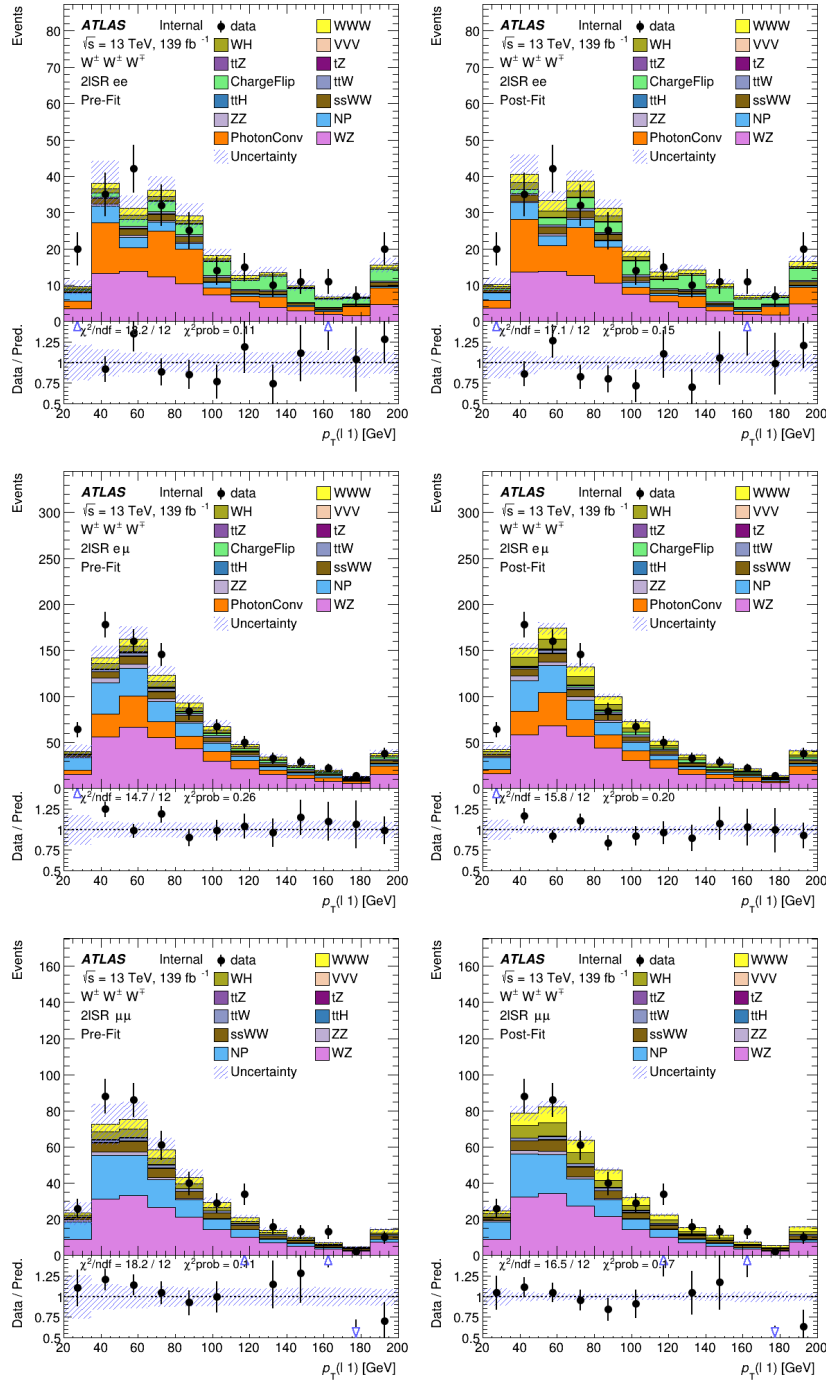
Figure A.32: The pre-fit (left) and post-fit (right) number of jets distributions with statistical and systematical uncertainties in the *ee* (top), *eμ* (middle), and *μμ* (bottom) channels of the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR in the combined fit.
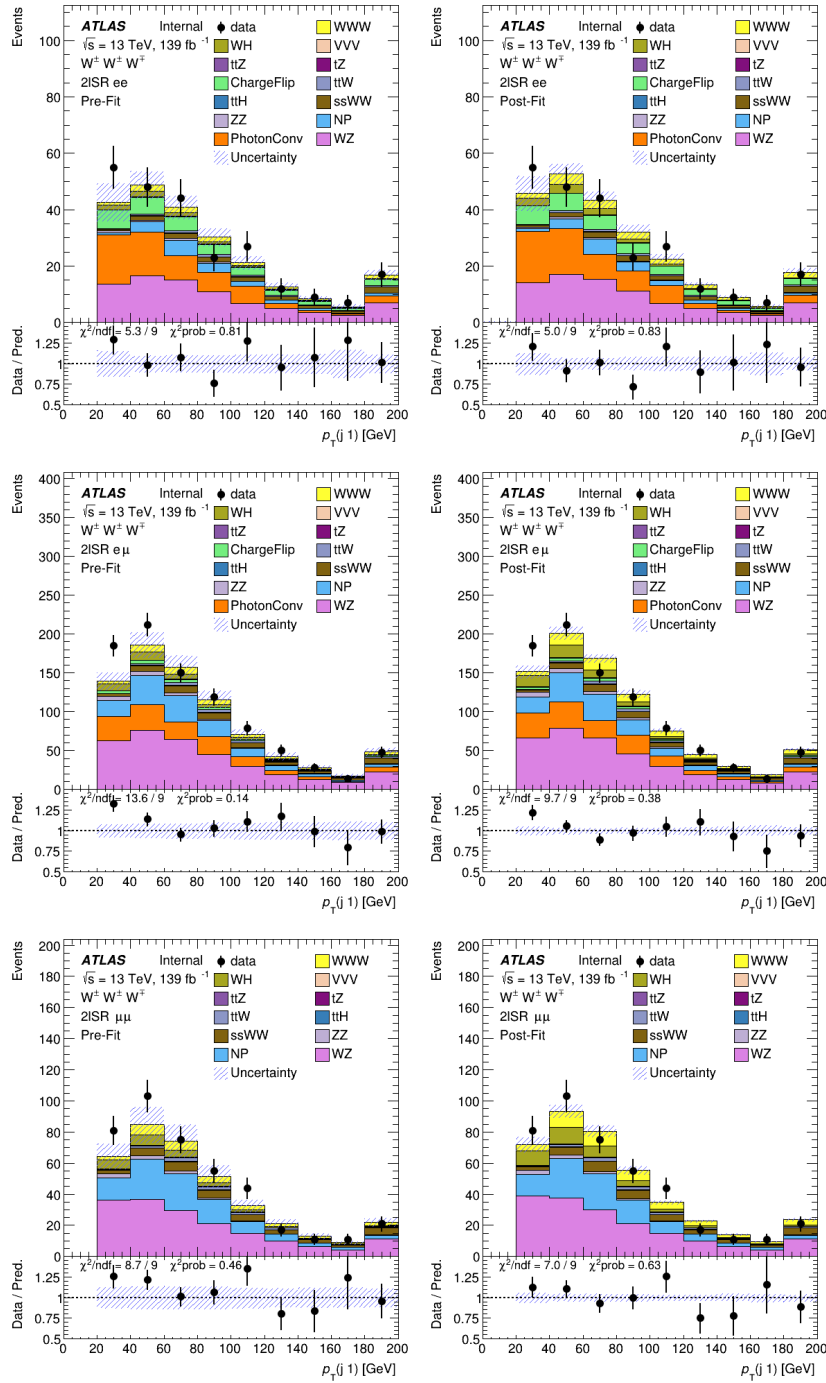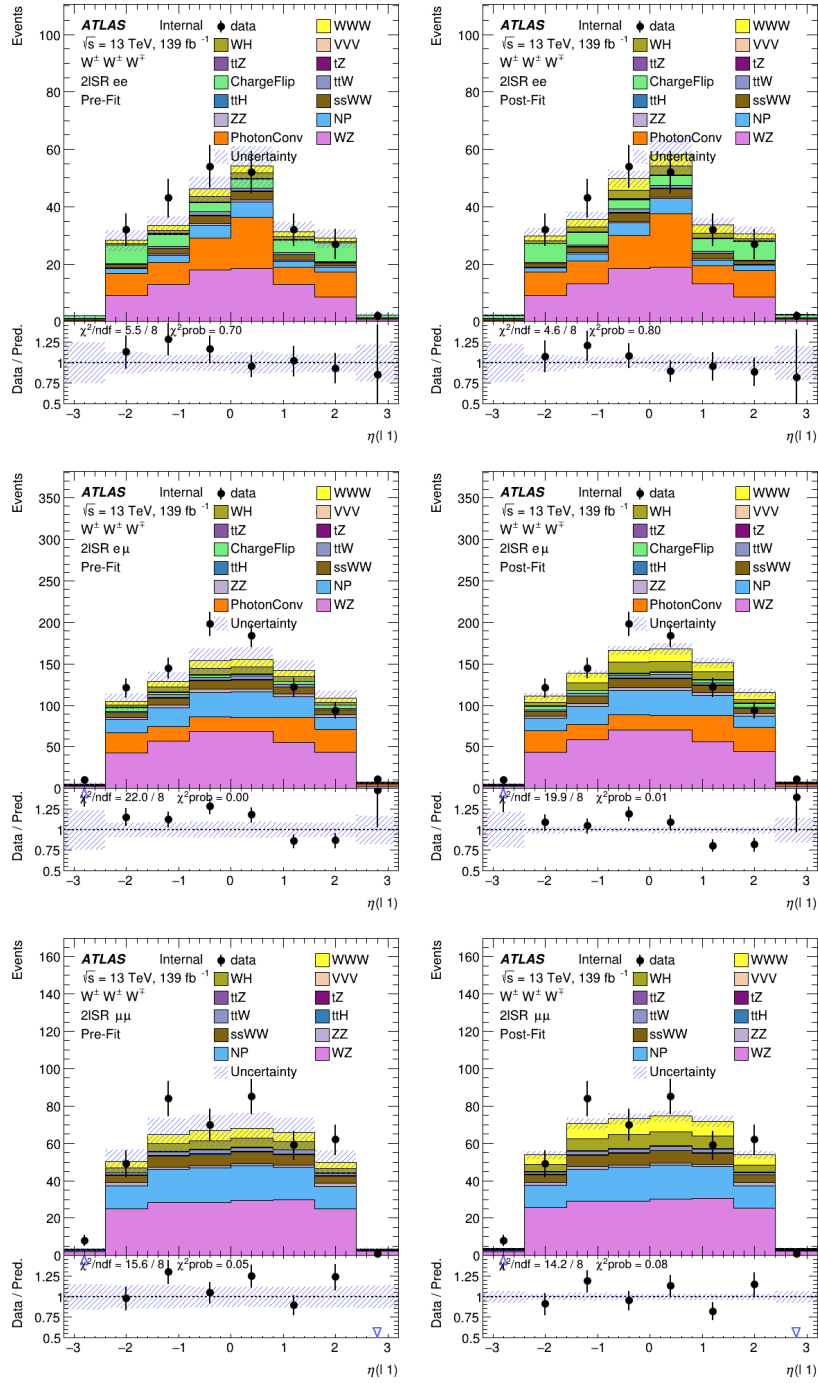
Figure A.33: The pre-fit (left) and post-fit (right) the leading jet $p_T$ distributions with statistical and systematical uncertainties in the *ee* (top), *eμ* (middle), and *μμ* (bottom) channels of the $\ell^\pm \nu \ell^\pm \nu j j$ SR in the combined fit.
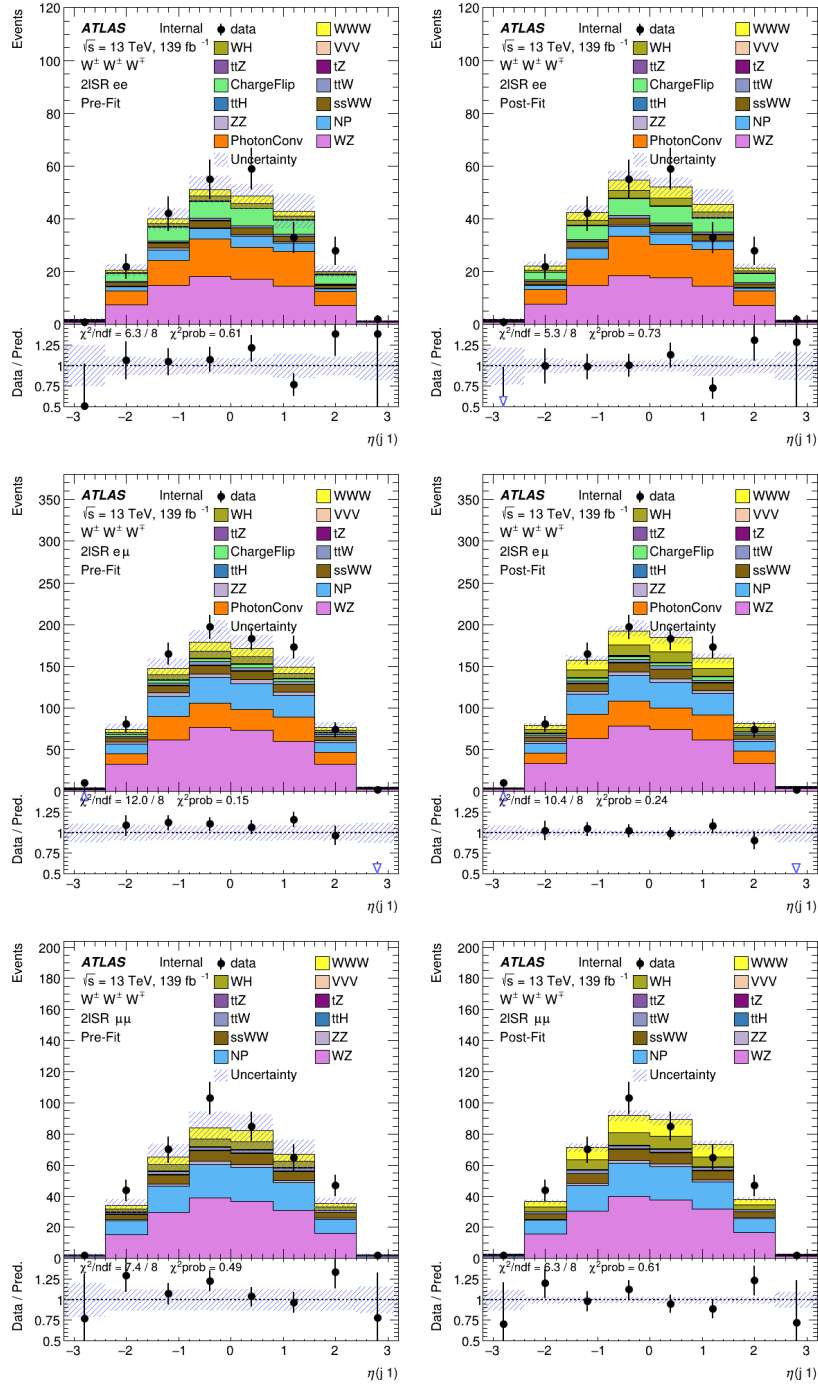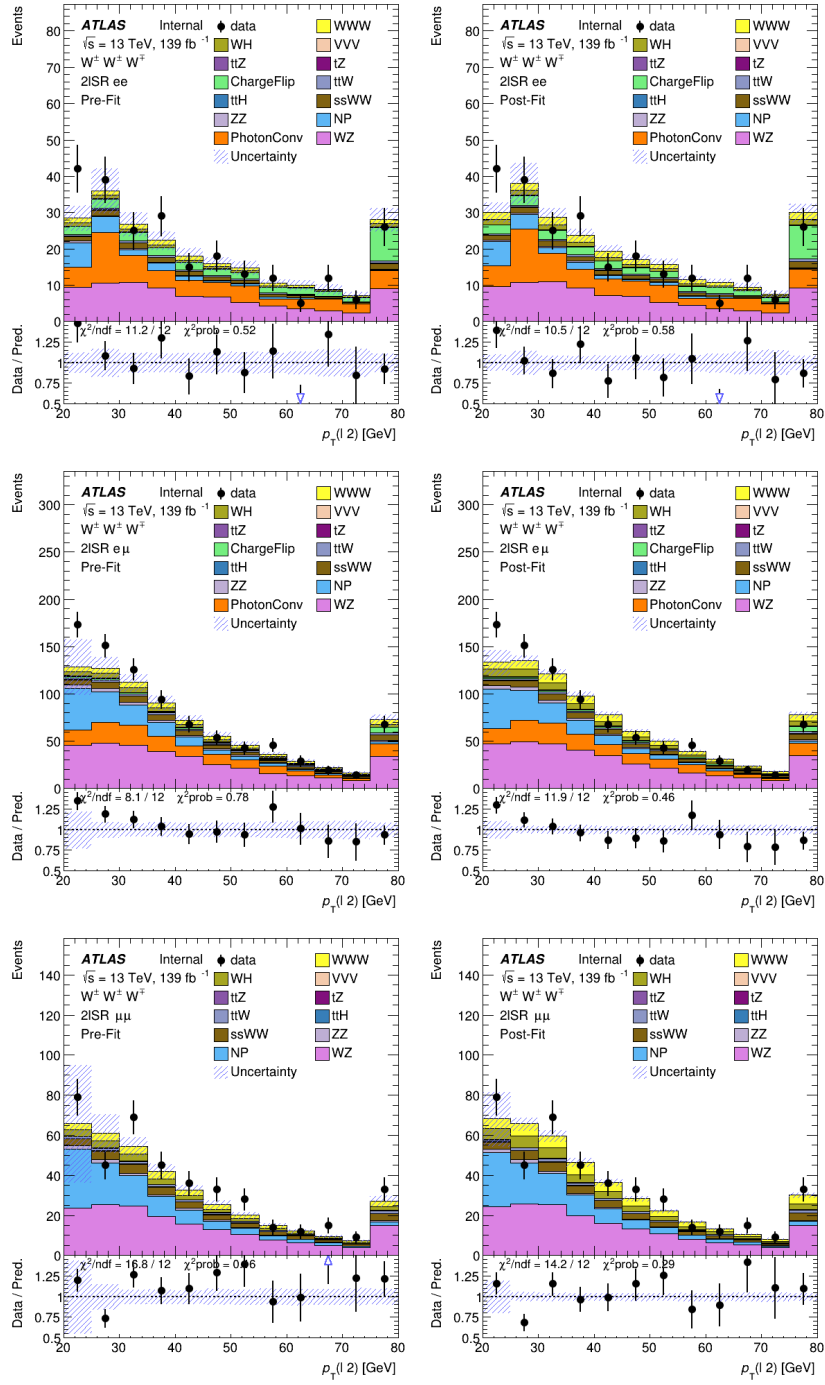
Figure A.34: The pre-fit (left) and post-fit (right) the leading jet $\eta$ distributions with statistical and systematical uncertainties in the *ee* (top), *eμ* (middle), and *μμ* (bottom) channels of the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR in the combined fit.
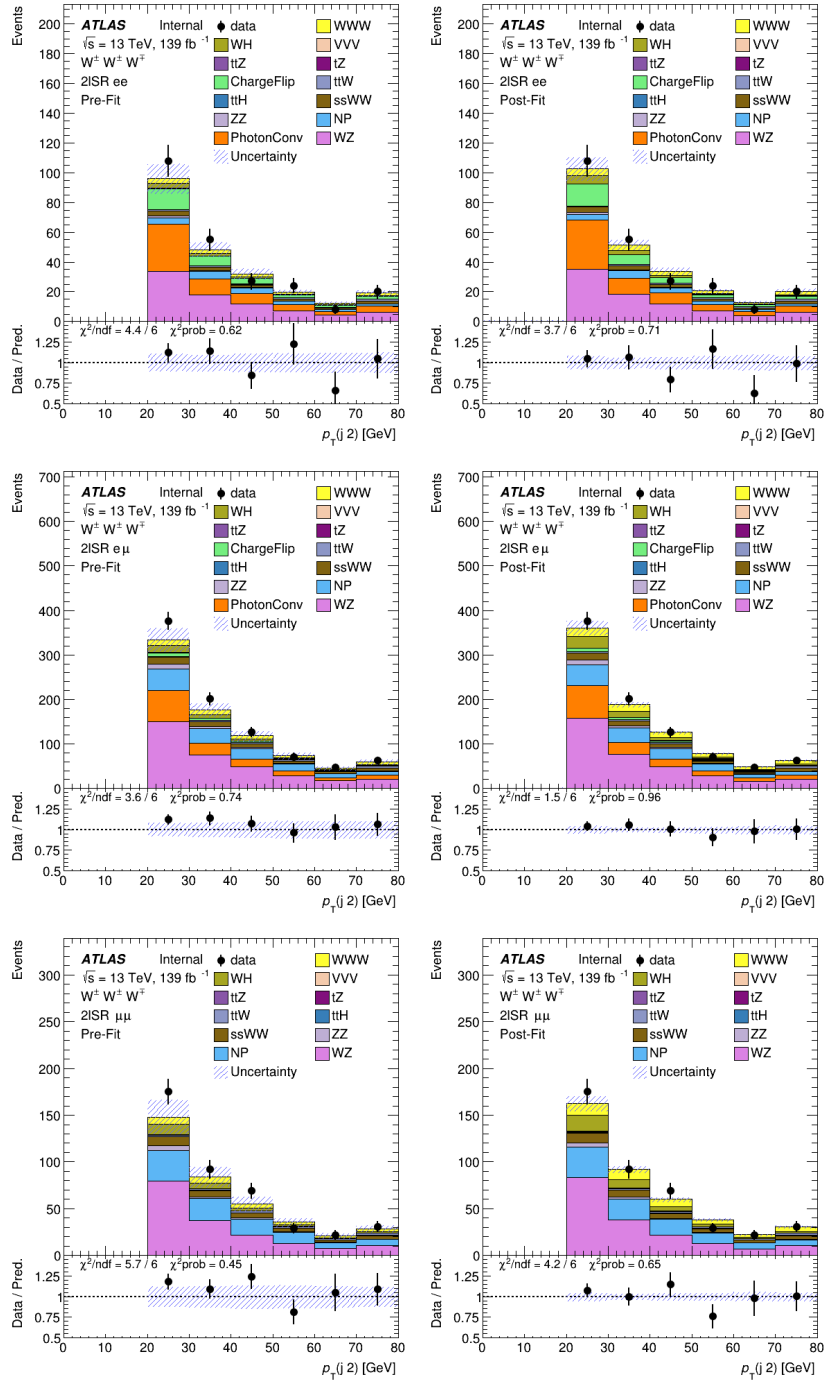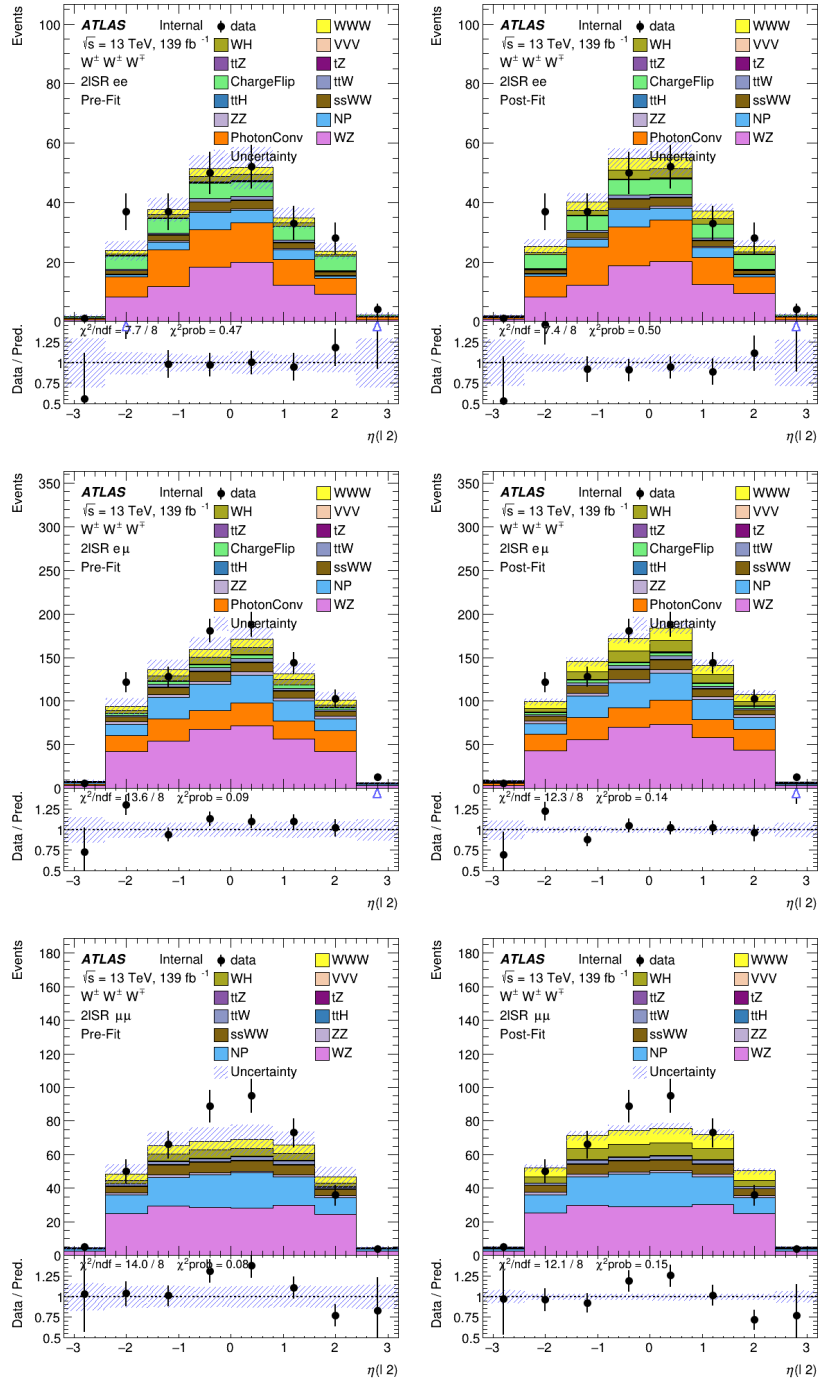
Figure A.35: The pre-fit (left) and post-fit (right) the subleading jet $p_T$ distributions with statistical and systematical uncertainties in the *ee* (top), *eμ* (middle), and *μμ* (bottom) channels of the $\ell^\pm \nu \ell^\pm \nu j j$ SR in the combined fit.
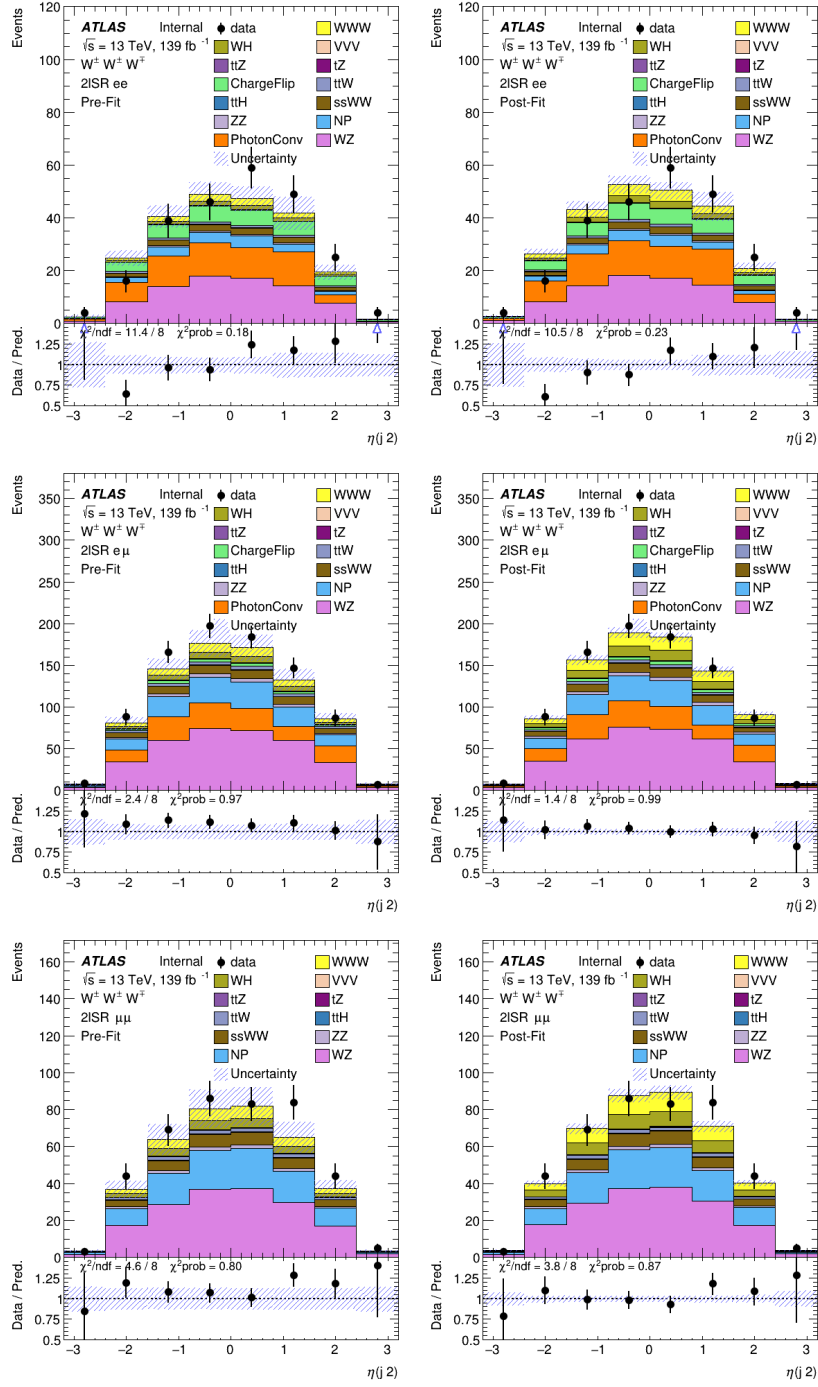
Figure A.36: The pre-fit (left) and post-fit (right) the subleading jet $\eta$ distributions with statistical and systematical uncertainties in the *ee* (top), *eμ* (middle), and *μμ* (bottom) channels of the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR in the combined fit.
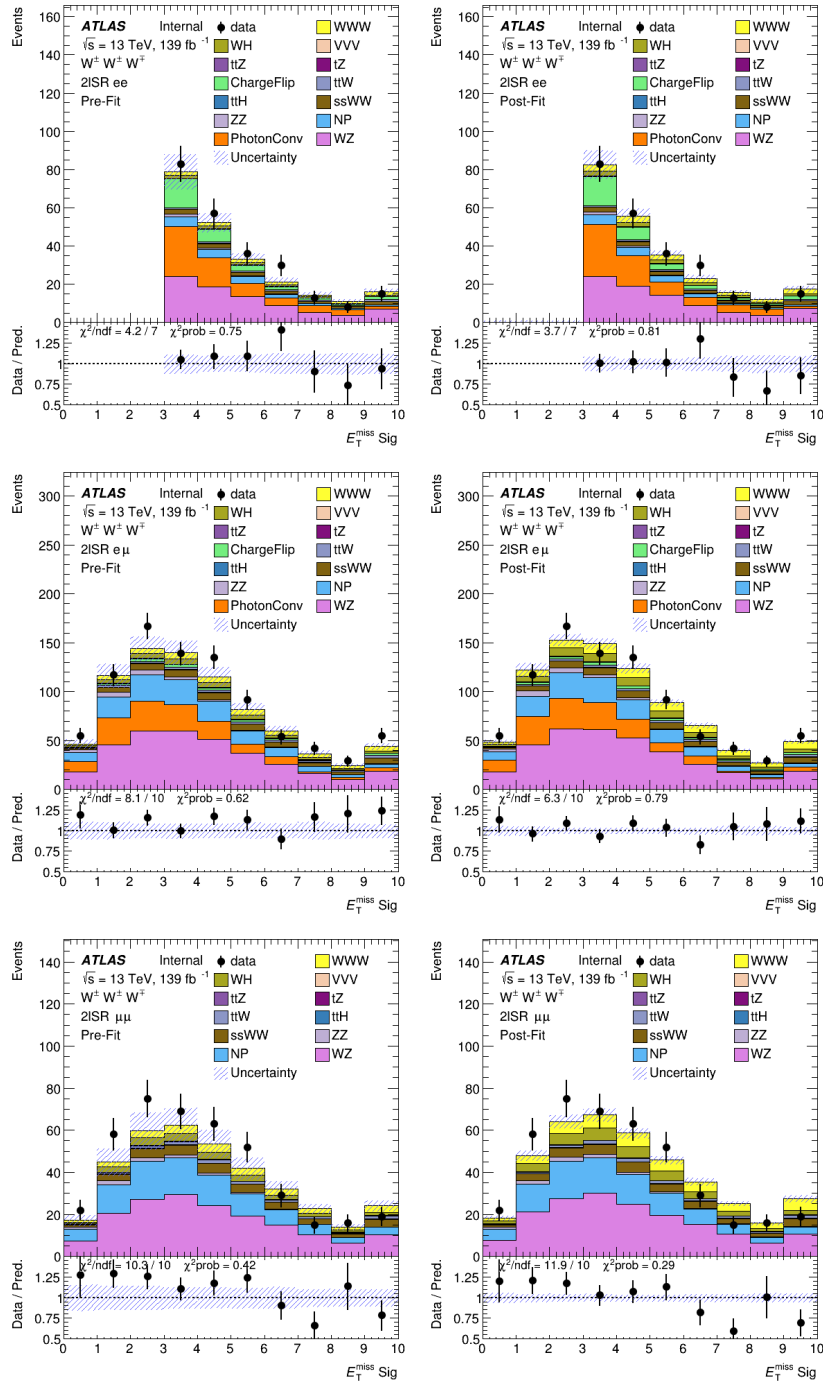
Figure A.37: The pre-fit (left) and post-fit (right) $m_{\ell\ell}$ distributions with statistical and systematical uncertainties in the *ee* (top), *eμ* (middle), and *μμ* (bottom) channels of the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR in the combined fit.
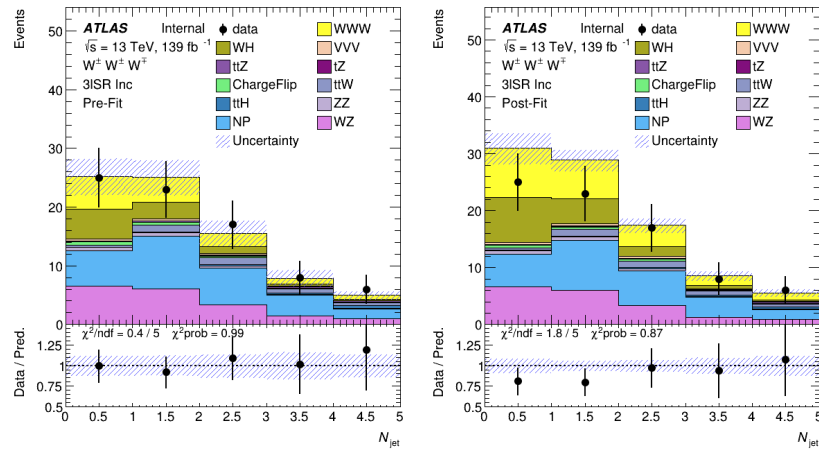
Figure A.38: The pre-fit (left) and post-fit (right) the leading lepton $p_T$ distributions with statistical and systematical uncertainties in the *ee* (top), *eμ* (middle), and *μμ* (bottom) channels of the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR in the combined fit.

Figure A.39: The pre-fit (left) and post-fit (right) the leading lepton $\eta$ distributions with statistical and systematical uncertainties in the *ee* (top), *eμ* (middle), and *μμ* (bottom) channels of the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR in the combined fit.

Figure A.40: The pre-fit (left) and post-fit (right) the subleading lepton $p_\mathrm{T}$ distributions with statistical and systematical uncertainties in the *ee* (top), *eμ* (middle), and *μμ* (bottom) channels of the $\ell^\pm\nu\ell^\pm\nu jj$ SR in the combined fit.

Figure A.41: The pre-fit (left) and post-fit (right) the subleading lepton $\eta$ distributions with statistical and systematical uncertainties in the *ee* (top), *eµ* (middle), and *µµ* (bottom) channels of the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR in the combined fit.

144

Figure A.42: The pre-fit (left) and post-fit (right) $E_T^{\mathrm{miss}}$ significance distributions with statistical and systematical uncertainties in the *ee* (top), *eμ* (middle), and *μμ* (bottom) channels of the $\ell^{\pm}\nu\ell^{\pm}\nu jj$ SR in the combined fit.

# A.7  $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ Signal Region



Figure A.43: The pre-fit (left) and post-fit (right) number of jets distributions with statistical and systematical uncertainties in the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ SR in the combined fit.
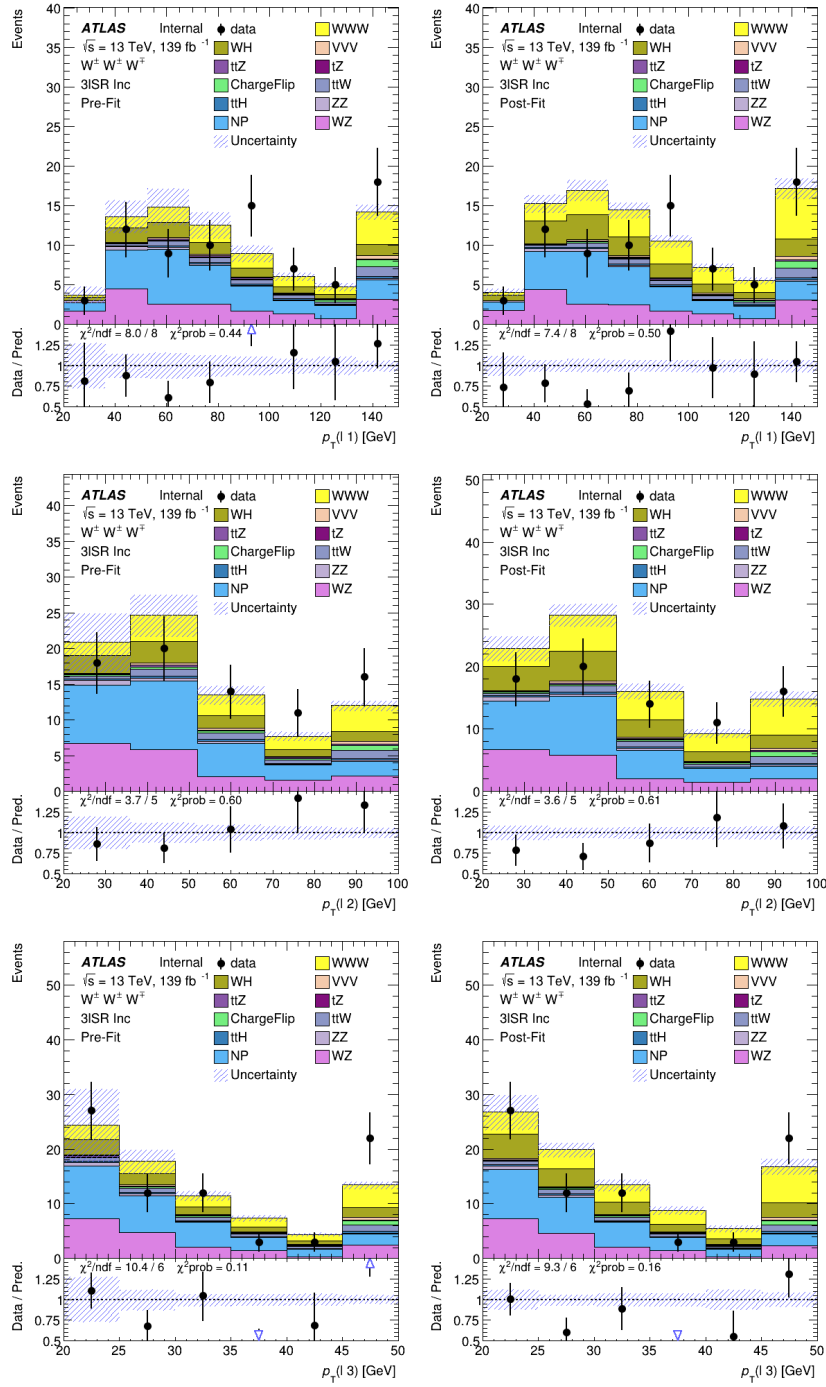
Figure A.44: The pre-fit (left) and post-fit (right) leading (top), subleading (middle), and third (bottom) lepton $p_T$ distributions with statistical and systematical uncertainties in the $\ell^\pm \nu \ell^\pm \nu \ell^\mp \nu$ SR in the combined fit.
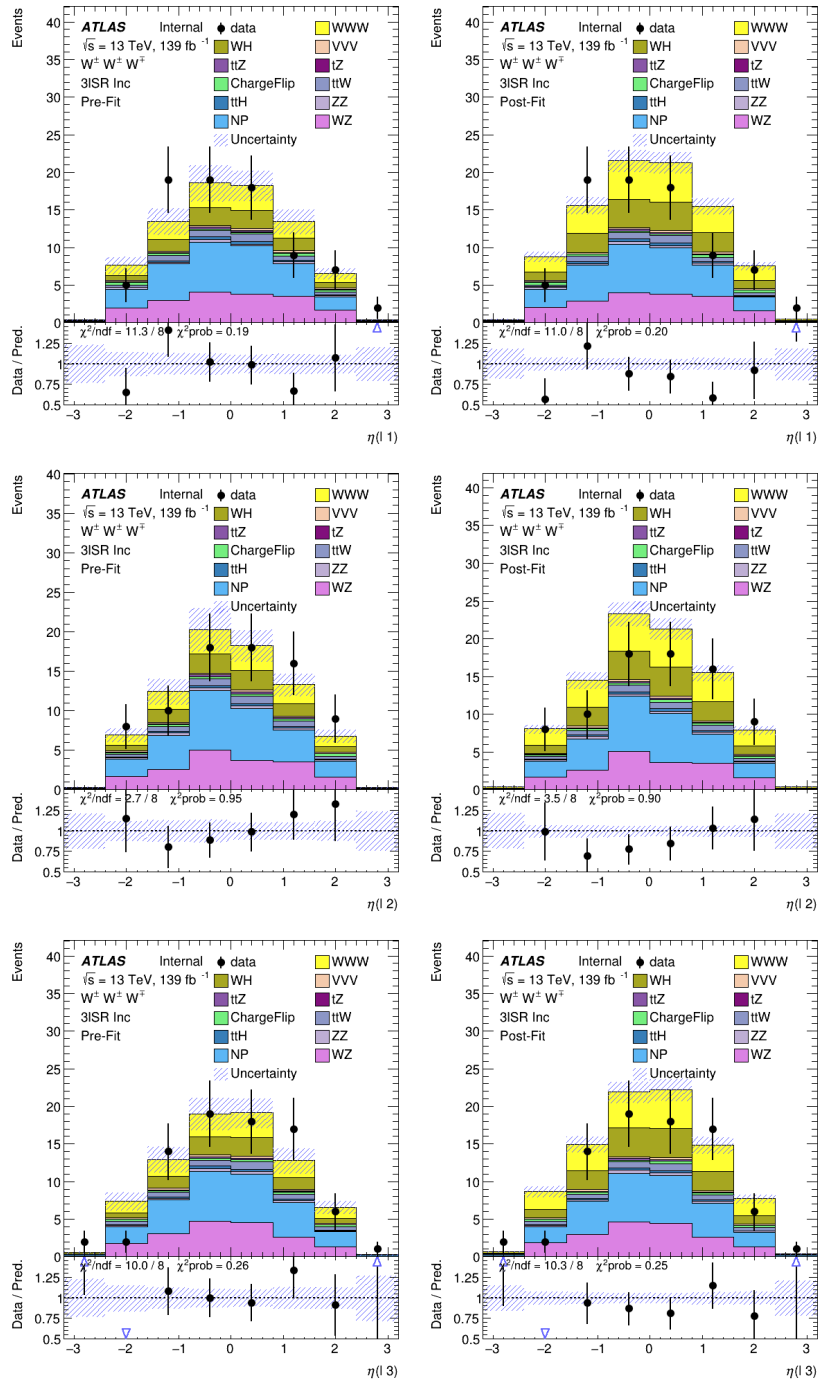
Figure A.45: The pre-fit (left) and post-fit (right) leading (top), subleading (middle), and third (bottom) lepton $\eta$ distributions with statistical and systematical uncertainties in the $\ell^{\pm}\nu\ell^{\pm}\nu\ell^{\mp}\nu$ SR in the combined fit.

# BIBLIOGRAPHY

[1]  ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". *Phys. Lett. B* 716 (2012), p. 1. DOI: `10.1016/j.physletb.2012.08.020`. arXiv: `1207.7214 [hep-ex]` (cited on page 1).

[2]  CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". *Phys. Lett. B* 716 (2012), p. 30. DOI: `10.1016/j.physletb.2012.08.021`. arXiv: `1207.7235 [hep-ex]` (cited on page 1).

[3]  ATLAS Collaboration. "Evidence for the production of three massive vector bosons with the ATLAS detector". *Phys. Lett. B* 798 (2019), p. 134913. DOI: `10.1016/j.physletb.2019.134913`. arXiv: `1903.10415 [hep-ex]` (cited on pages 1, 16, 101).

[4]  CMS Collaboration. "Observation of the Production of Three Massive Gauge Bosons at $\sqrt{s}$ =13 TeV". *Phys. Rev. Lett.* 125.15 (2020), p. 151802. DOI: `10.1103/PhysRevLett.125.151802`. arXiv: `2006.11191 [hep-ex]` (cited on pages 1, 16).

[5]  S. Braibant, G. Giacomelli, and M. Spurio. *Particles and Fundamental Interactions: An Introduction to Particle Physics*. Undergraduate Lecture Notes in Physics. Springer Netherlands, 2011. ISBN: 9789400724648. URL: `https://books.google.com/books?id=e8YUUG2pGeIC` (cited on pages 4, 5).

[6]  Andrew Purcell. *Go on a particle quest at the first CERN webfest*. Aug. 2012, p. 10. URL: `https://cds.cern.ch/record/1473657` (cited on page 5).

[7]  Antonio Pich. "The Standard model of electroweak interactions" (2008), pp. 1–49. arXiv: `0705.4264 [hep-ph]` (cited on pages 5–10, 12, 14, 15).

[8]  Emmy Noether. "Invariant Variation Problems". *Gott. Nachr.* 1918 (1918), pp. 235–257. DOI: `10.1080/00411457108231446`. arXiv: `physics/0503066` (cited on page 6).

[9]     Paul Langacker. *The Standard Model and Beyond*. Series in High Energy Physics, Cosmology and Gravitation. CRC Press, 2017. ISBN: 9781498763240. URL: `https://books.google.com/books?id=dKnuDwAAQBAJ` (cited on pages 6, 14).

[10]    David Griffiths. *Introduction to Elementary Particles*. Physics textbook. Wiley, 2008. ISBN: 9783527618477. URL: `https://books.google.com/books?id=Wb9DYrjcoKAC` (cited on pages 6, 9, 10).

[11]    Chen-Ning Yang and Robert L. Mills. "Conservation of Isotopic Spin and Isotopic Gauge Invariance". *Phys. Rev.* 96 (1954). Ed. by Jong-Ping Hsu and D. Fine, pp. 191–195. DOI: `10.1103/PhysRev.96.191` (cited on page 9).

[12]    Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 508–509. DOI: `10.1103/PhysRevLett.13.508` (cited on page 11).

[13]    F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 321–323. DOI: `10.1103/PhysRevLett.13.321` (cited on page 11).

[14]    ATLAS Collaboration. "Study of $WW\gamma$ and $WZ\gamma$ production in $pp$ collisions at $\sqrt{s} = 8\,\text{TeV}$ and search for anomalous quartic gauge couplings with the ATLAS experiment". *Eur. Phys. J. C* 77 (2017), p. 646. DOI: `10.1140/epjc/s10052-017-5180-3`. arXiv: `1707.05597 [hep-ex]` (cited on page 16).

[15]    CMS Collaboration. "A search for $WW\gamma$ and $WZ\gamma$ production and constraints on anomalous quartic gauge couplings in $pp$ collisions at $\sqrt{s} = 8\,\text{TeV}$". *Phys. Rev. D* 90 (2014), p. 032008. DOI: `10.1103/PhysRevD.90.032008`. arXiv: `1404.4619 [hep-ex]` (cited on page 16).

[16]    ATLAS Collaboration. "Evidence of $W\gamma\gamma$ Production in $pp$ Collisions at $\sqrt{s} = 8\,\text{TeV}$ and Limits on Anomalous Quartic Gauge Couplings with the ATLAS Detector". *Phys. Rev. Lett.* 115 (2015), p. 031802. DOI: `10.1103/PhysRevLett.115.031802`. arXiv: `1503.03243 [hep-ex]` (cited on page 16).

[17]    CMS Collaboration. "Measurements of the $pp \to W\gamma\gamma$ and $pp \to Z\gamma\gamma$ cross sections and limits on anomalous quartic gauge couplings at $\sqrt{s} = 8\,\text{TeV}$". *JHEP* 10 (2017), p. 072. DOI: `10.1007/JHEP10(2017)072`. arXiv: `1704.00366 [hep-ex]` (cited on page 16).

[18]    ATLAS Collaboration. "Measurements of $Z\gamma$ and $Z\gamma\gamma$ production in $pp$ collisions at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS detector". *Phys. Rev. D* 93 (2016), p. 112002. DOI: `10.1103/PhysRevD.93.112002`. arXiv: `1604.05232 [hep-ex]` (cited on page 16).

[19] ATLAS Collaboration. "Measurement of the production cross section of three isolated photons in $pp$ collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector". *Phys. Lett. B* 781 (2018), p. 55. DOI: `10.1016/j.physletb.2018.03.057`. arXiv: `1712.07291 [hep-ex]` (cited on page 16).

[20] Steven Weinberg. "Baryon and Lepton Nonconserving Processes". *Phys. Rev. Lett.* 43 (1979), pp. 1566–1570. DOI: `10.1103/PhysRevLett.43.1566` (cited on page 18).

[21] W. Buchmuller and D. Wyler. "Effective Lagrangian Analysis of New Interactions and Flavor Conservation". *Nucl. Phys. B* 268 (1986), pp. 621–653. DOI: `10.1016/0550-3213(86)90262-2` (cited on page 18).

[22] O. J. P. Eboli, M. C. Gonzalez-Garcia, and J. K. Mizukoshi. "$pp \to jje^{\pm}\mu^{\pm}\nu\nu$ and $jje^{\pm}\mu^{\mp}\nu\nu$ at $\mathcal{O}(\alpha_{\text{em}}^6)$ and $\mathcal{O}(\alpha_{\text{em}}^4\alpha_s^2)$ for the study of the quartic electroweak gauge boson vertex at LHC". *Phys. Rev. D* 74 (2006), p. 073005. DOI: `10.1103/PhysRevD.74.073005`. arXiv: `hep-ph/0606118` (cited on page 18).

[23] O. J. P. Eboli, M. C. Gonzalez-Garcia, and S. M. Lietti. "Bosonic quartic couplings at CERN LHC". *Phys. Rev. D* 69 (2004), p. 095005. DOI: `10.1103/PhysRevD.69.095005`. arXiv: `hep-ph/0310141` (cited on page 19).

[24] ATLAS Collaboration. "Search for triboson $W^{\pm}W^{\pm}W^{\mp}$ production in $pp$ collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector". *Eur. Phys. J. C* 77 (2017), p. 141. DOI: `10.1140/epjc/s10052-017-4692-1`. arXiv: `1610.05088 [hep-ex]` (cited on page 19).

[25] Lyndon Evans and Philip Bryant. "LHC Machine". *JINST* 3 (2008), S08001. DOI: `10.1088/1748-0221/3/08/S08001` (cited on page 20).

[26] CERN. *LHC Season 2 facts and figures*. URL: `https://home.cern/sites/home.web.cern.ch/files/2018-07/factsandfigures-en_0.pdf`, last accessed on 01/21/2021 (cited on page 20).

[27] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". *JINST* 3 (2008), S08003. DOI: `10.1088/1748-0221/3/08/S08003` (cited on pages 20, 22–30).

[28] CMS Collaboration. "The CMS experiment at the CERN LHC". *JINST* 3 (2008), S08004. DOI: `10.1088/1748-0221/3/08/S08004` (cited on page 20).

[29] ALICE Collaboration. "The ALICE experiment at the CERN LHC". *JINST* 3 (2008), S08002. DOI: `10.1088/1748-0221/3/08/S08002` (cited on page 20).

[30] LHCb Collaboration. "The LHCb Detector at the LHC". *JINST* 3 (2008), S08005. DOI: `10.1088/1748-0221/3/08/S08005` (cited on page 20).

[31] Esma Mobs. "The CERN accelerator complex - 2019" (July 2019). General Photo. URL: https://cds.cern.ch/record/2684277 (cited on page 21).

[32] CERN. *LHC the guide*. URL: https://home.cern/resources/brochure/cern/lhc-guide, last accessed on 01/21/2021 (cited on page 20).

[33] ATLAS Collaboration. *Luminosity Public Results*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults, last accessed on 01/21/2021 (cited on page 22).

[34] ATLAS Collaboration. *Luminosity Public Results Run2*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2, last accessed on 01/21/2021 (cited on page 22).

[35] Konstantinos Ntekas. "Performance characterization of the Micromegas detector for the New Small Wheel upgrade and Development and improvement of the Muon Spectrometer Detector Control System in the ATLAS experiment". PhD thesis. Natl. Tech. U., Athens, 2016. URL: http://cds.cern.ch/record/2143887 (cited on page 25).

[36] ATLAS Collaboration. *ATLAS Magnet System: Magnet Project Technical Design Report, Volume 1*. ATLAS-TDR-6; CERN-LHCC-97-018. 1997. URL: https://cds.cern.ch/record/338080 (cited on page 24).

[37] ATLAS Collaboration. *Experiment Briefing: Keeping the ATLAS Inner Detector in perfect alignment*. General Photo. July 2020. URL: https://cds.cern.ch/record/2723878 (cited on page 27).

[38] ATLAS Collaboration. *ATLAS Inner Detector: Technical Design Report, Volume 1*. ATLAS-TDR-4; CERN-LHCC-97-016. 1997. URL: https://cds.cern.ch/record/331063 (cited on page 25).

[39] ATLAS Collaboration. *ATLAS Inner Detector: Technical Design Report, Volume 2*. ATLAS-TDR-5, CERN-LHCC-97-017. 1997. URL: https://cds.cern.ch/record/331064 (cited on page 25).

[40] ATLAS Collaboration. "The ATLAS Inner Detector commissioning and calibration". *Eur. Phys. J. C* 70 (2010), p. 787. DOI: 10.1140/epjc/s10052-010-1366-7. arXiv: 1004.5293 [hep-ex] (cited on page 25).

[41] ATLAS Collaboration. *ATLAS Pixel Detector: Technical Design Report*. ATLAS-TDR-11; CERN-LHCC-98-013. 1998. URL: https://cds.cern.ch/record/381263 (cited on page 26).

[42] ATLAS Collaboration. *ATLAS Insertable B-Layer Technical Design Report*. ATLAS-TDR-19; CERN-LHCC-2010-013. 2010. URL: https://cds.cern.ch/record/1291633 (cited on page 26).

[43] B. Abbott et al. "Production and integration of the ATLAS Insertable B-Layer". *JINST* 13 (2018), T05008. DOI: 10.1088/1748-0221/13/05/T05008. arXiv: 1803.00844 [physics.ins-det] (cited on page 26).

[44] ATLAS Collaboration. "Operation and performance of the ATLAS semiconductor tracker". *JINST* 9 (2014), P08009. DOI: 10.1088/1748-0221/9/08/P08009. arXiv: 1404.7473 [hep-ex] (cited on page 26).

[45] ATLAS Collaboration. *ATLAS Transition Radiation Tracker (TRT): Straw Tube Gaseous Detectors at High Rates*. Apr. 2013. URL: http://cds.cern.ch/record/1537991 (cited on page 26).

[46] ATLAS Collaboration. *ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider*. Mar. 2016. DOI: 10.1016/j.nima.2016.04.026. URL: https://cds.cern.ch/record/2139567 (cited on page 26).

[47] ATLAS Collaboration. *ATLAS Calorimeter Performance: Technical Design Report*. ATLAS-TDR-1; CERN-LHCC-96-040. 1996. URL: https://cds.cern.ch/record/331059 (cited on page 27).

[48] ATLAS Collaboration. *ATLAS Liquid Argon Calorimeter: Technical Design Report*. ATLAS-TDR-2; CERN-LHCC-96-041. 1996. URL: https://cds.cern.ch/record/331061 (cited on pages 27, 28).

[49] ATLAS Collaboration. *ATLAS Tile Calorimeter: Technical Design Report*. ATLAS-TDR-3; CERN-LHCC-96-042. 1996. URL: https://cds.cern.ch/record/331062 (cited on pages 27, 28).

[50] ATLAS Collaboration. *ATLAS Muon Spectrometer: Technical Design Report*. ATLAS-TDR-10; CERN-LHCC-97-022. CERN, 1997. URL: https://cds.cern.ch/record/331068 (cited on page 29).

[51] ATLAS Collaboration. "Operation of the ATLAS trigger system in Run 2". *JINST* 15.10 (2020), P10004. DOI: 10.1088/1748-0221/15/10/P10004. arXiv: 2007.12539 [physics.ins-det] (cited on page 31).

[52] ATLAS Collaboration. "Performance of the ATLAS trigger system in 2015". *Eur. Phys. J. C* 77 (2017), p. 317. DOI: 10.1140/epjc/s10052-017-4852-3. arXiv: 1611.09661 [hep-ex] (cited on pages 31, 46, 47, 89, 90).

[53]   ATLAS Collaboration. *ATLAS Level-1 Trigger: Technical Design Report*. ATLAS-TDR-12; CERN-LHCC-98-014. 1998. URL: https://cds.cern.ch/record/381429 (cited on page 32).

[54]   ATLAS Collaboration. *ATLAS High-Level Trigger, Data Acquisition and Controls: Technical Design Report*. ATLAS-TDR-16; CERN-LHCC-2003-022. 2003. URL: https://cds.cern.ch/record/616089 (cited on page 32).

[55]   ATLAS Collaboration. *ATLAS Computing: Technical Design Report*. ATLAS-TDR-17; CERN-LHCC-2005-022. 2005. URL: https://cds.cern.ch/record/837738 (cited on page 32).

[56]   ATLAS Collaboration. "ATLAS data quality operations and performance for 2015–2018 data-taking". *JINST* 15 (2020), P04003. DOI: 10.1088/1748-0221/15/04/P04003. arXiv: 1911.04632 [physics.ins-det] (cited on page 33).

[57]   ATLAS Collaboration. "Measurement of the Inelastic Proton–Proton Cross Section at $\sqrt{s} =$ 13 TeV with the ATLAS Detector at the LHC". *Phys. Rev. Lett.* 117 (2016), p. 182002. DOI: 10.1103/PhysRevLett.117.182002. arXiv: 1606.02625 [hep-ex] (cited on page 34).

[58]   ATLAS Collaboration. *Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} =$ 13 TeV*. ATL-PHYS-PUB-2015-026. 2015. URL: https://cds.cern.ch/record/2037717 (cited on page 34).

[59]   ATLAS Collaboration. *Final Tracking CP Recommendations for Run 2*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TrackingCPRecsRun2Final, last accessed on 01/24/2021 (cited on page 34).

[60]   ATLAS Collaboration. "Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC". *Eur. Phys. J. C* 77 (2017), p. 332. DOI: 10.1140/epjc/s10052-017-4887-5. arXiv: 1611.10235 [hep-ex] (cited on page 34).

[61]   Will Buttinger and Peter Onyisi. *Pileup Reweighting for MC16d*. URL: https://indico.cern.ch/event/712774/contributions/2928042/attachments/1614637/2565496/prw_mc16d.pdf, last accessed on 01/24/2021 (cited on page 35).

[62]   ATLAS Collaboration. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} =$ 13 TeV". *Eur. Phys. J. C* 79 (2019), p. 639. DOI: 10.1140/epjc/s10052-019-7140-6. arXiv: 1902.04655 [hep-ex] (cited on pages 35, 36, 89).

[63] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data". *JINST* 14 (2019), P03017. DOI: `10.1088/1748-0221/14/03/P03017`. arXiv: `1812.03848 [hep-ex]` (cited on pages 35, 89).

[64] ATLAS Collaboration. "Electron and photon performance measurements with the AT-LAS detector using the 2015–2017 LHC proton–proton collision data". *JINST* 14 (2019), P12006. DOI: `10.1088/1748-0221/14/12/P12006`. arXiv: `1908.00005 [hep-ex]` (cited on pages 35–37, 89).

[65] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". *Eur. Phys. J. C* 77 (2017), p. 490. DOI: `10.1140/epjc/s10052-017-5004-5`. arXiv: `1603.02934 [hep-ex]` (cited on page 35).

[66] ATLAS Collaboration. *Definition of Lepton PLV (isolation) working points*. URL: `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/RecommendedIsolationWPs#Definition_of_Lepton_PLV_isolati`, last accessed on 01/24/2021 (cited on pages 37, 40).

[67] ATLAS Collaboration. "Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector". *Phys. Rev. D* 97 (2018), p. 072003. DOI: `10.1103/PhysRevD.97.072003`. arXiv: `1712.08891 [hep-ex]` (cited on page 37).

[68] ATLAS Collaboration. *Electron Charge ID Selector Tool*. URL: `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronChargeFlipTaggerTool`, last accessed on 01/24/2021 (cited on page 37).

[69] Henri Bachacou and Jean-Baptiste de Vivie. *Additional y* and material conversion veto*. URL: `https://indico.cern.ch/event/862748/contributions/3635039/attachments/1946503/3235131/conversionsIFF2019.11.18.pdf`, last accessed on 01/24/2021 (cited on page 37).

[70] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13\,\mathrm{TeV}$". *Eur. Phys. J. C* 76 (2016), p. 292. DOI: `10.1140/epjc/s10052-016-4120-y`. arXiv: `1603.05598 [hep-ex]` (cited on pages 39, 90).

[71] ATLAS Collaboration. *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13\,\mathrm{TeV}$*. ATLAS-CONF-2020-030. 2020. URL: `https://cds.cern.ch/record/2725736` (cited on pages 39, 40, 90).

[72] ATLAS Collaboration. "Jet reconstruction and performance using particle flow with the ATLAS Detector". *Eur. Phys. J. C* 77 (2017), p. 466. DOI: `10.1140/epjc/s10052-017-5031-2`. arXiv: `1703.10485 [hep-ex]` (cited on page 42).

[73] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-$k_t$ jet clustering algorithm". *JHEP* 04 (2008), p. 063. DOI: `10.1088/1126-6708/2008/04/063`. arXiv: `0802.1189 [hep-ph]` (cited on page 42).

[74] ATLAS Collaboration. "Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". *Phys. Rev. D* 96 (2017), p. 072002. DOI: `10.1103/PhysRevD.96.072002`. arXiv: `1703.09665 [hep-ex]` (cited on pages 42, 90).

[75] ATLAS Collaboration. *Selection of jets produced in* 13 *TeV proton–proton collisions with the ATLAS detector*. ATLAS-CONF-2015-029. 2015. URL: `https://cds.cern.ch/record/2037702` (cited on page 43).

[76] ATLAS Collaboration. *Tagging and suppression of pileup jets*. ATL-PHYS-PUB-2014-001. 2014. URL: `https://cds.cern.ch/record/1643929` (cited on pages 43, 91).

[77] ATLAS Collaboration. "Performance of pile-up mitigation techniques for jets in *pp* collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector". *Eur. Phys. J. C* 76 (2016), p. 581. DOI: `10.1140/epjc/s10052-016-4395-z`. arXiv: `1510.03823 [hep-ex]` (cited on pages 43, 91).

[78] ATLAS Collaboration. *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. ATL-PHYS-PUB-2017-013. 2017. URL: `https://cds.cern.ch/record/2273281` (cited on page 43).

[79] ATLAS Collaboration. "ATLAS *b*-jet identification performance and efficiency measurement with $t\bar{t}$ events in *pp* collisions at $\sqrt{s}$ = 13 TeV". *Eur. Phys. J. C* 79 (2019), p. 970. DOI: `10.1140/epjc/s10052-019-7450-8`. arXiv: `1907.05120 [hep-ex]` (cited on pages 43, 91).

[80] ATLAS Collaboration. $E_T^{miss}$ *performance in the ATLAS detector using 2015–2016 LHC pp collisions*. ATLAS-CONF-2018-023. 2018. URL: `https://cds.cern.ch/record/2625233` (cited on pages 44, 91).

[81] ATLAS Collaboration. *Object-based missing transverse momentum significance in the ATLAS Detector*. ATLAS-CONF-2018-038. 2018. URL: `https://cds.cern.ch/record/2630948` (cited on page 44).

[82]  ATLAS Collaboration. *Recommendations of the Physics Objects and Analysis Harmon-isation Study Groups 2014*. July 2014. URL: https://cds.cern.ch/record/1743654 (cited on page 45).

[83]  ATLAS Collaboration. "Performance of electron and photon triggers in ATLAS during LHC Run 2". *Eur. Phys. J. C* 80 (2020), p. 47. DOI: 10.1140/epjc/s10052-019-7500-2. arXiv: 1909.00761 [hep-ex] (cited on pages 46, 89).

[84]  ATLAS Collaboration. "Performance of the ATLAS muon triggers in Run 2" (2020). arXiv: 2004.13447 [hep-ex] (cited on pages 47, 90).

[85]  ATLAS Collaboration. *Luminosity determination in pp collisions at $\sqrt{s} = 13\,TeV$ using the ATLAS detector at the LHC*. ATLAS-CONF-2019-021. 2019. URL: https://cds.cern.ch/record/2677054 (cited on page 48).

[86]  G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". *JINST* 13.07 (2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017 (cited on page 48).

[87]  Enrico Bothmann et al. "Event generation with Sherpa 2.2". *SciPost Phys.* 7.3 (2019), p. 034. DOI: 10.21468/SciPostPhys.7.3.034. arXiv: 1905.09127 [hep-ph] (cited on page 49).

[88]  J. Alwall et al. "The automated computation of tree-level and next-to-leading order differ-ential cross sections, and their matching to parton shower simulations". *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph] (cited on page 49).

[89]  Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. "A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction". *JHEP* 09 (2007), p. 126. DOI: 10.1088/1126-6708/2007/09/126. arXiv: 0707.3088 [hep-ph] (cited on page 49).

[90]  Paolo Nason. "A new method for combining NLO QCD with shower Monte Carlo algo-rithms". *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146 (cited on page 49).

[91]  Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with parton shower simulations: the POWHEG method". *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph] (cited on page 49).

[92] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". *JHEP* 06 (2010), p. 043. DOI: `10.1007/JHEP06(2010)043`. arXiv: `1002.2581 [hep-ph]` (cited on page 49).

[93] T. Sjöstrand, S. Mrenna, and P. Skands. "A brief introduction to PYTHIA 8.1". *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: `10.1016/j.cpc.2008.01.036`. arXiv: `0710.3820 [hep-ph]` (cited on page 49).

[94] M. Bähr et al. "Herwig++ physics and manual". *Eur. Phys. J. C* 58 (2008), p. 639. DOI: `10.1140/epjc/s10052-008-0798-9`. arXiv: `0803.0883 [hep-ph]` (cited on page 49).

[95] Johannes Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: `10.1140/epjc/s10052-016-4018-8`. arXiv: `1512.01178 [hep-ph]` (cited on page 49).

[96] S. Agostinelli et al. "GEANT4 – a simulation toolkit". *Nucl. Instrum. Meth. A* 506 (2003), p. 250. DOI: `10.1016/S0168-9002(03)01368-8` (cited on page 49).

[97] Federico Buccioni et al. "OpenLoops 2". *Eur. Phys. J. C* 79.10 (2019), p. 866. DOI: `10.1140/epjc/s10052-019-7306-2`. arXiv: `1907.13071 [hep-ph]` (cited on page 50).

[98] Fabio Cascioli, Philipp Maierhöfer, and Stefano Pozzorini. "Scattering Amplitudes with Open Loops". *Phys. Rev. Lett.* 108 (2012), p. 111601. DOI: `10.1103/PhysRevLett.108.111601`. arXiv: `1111.5206 [hep-ph]` (cited on page 50).

[99] Ansgar Denner, Stefan Dittmaier, and Lars Hofer. "COLLIER: A fortran-based complex one-loop library in extended regularizations". *Comput. Phys. Commun.* 212 (2017), pp. 220–238. DOI: `10.1016/j.cpc.2016.10.013`. arXiv: `1604.06792 [hep-ph]` (cited on page 50).

[100] Richard D. Ball et al. "Parton distributions for the LHC run II". *JHEP* 04 (2015), p. 040. DOI: `10.1007/JHEP04(2015)040`. arXiv: `1410.8849 [hep-ph]` (cited on page 50).

[101] Tanju Gleisberg and Stefan Höche. "Comix, a new matrix element generator". *JHEP* 12 (2008), p. 039. DOI: `10.1088/1126-6708/2008/12/039`. arXiv: `0808.3674 [hep-ph]` (cited on page 50).

[102] Steffen Schumann and Frank Krauss. "A parton shower algorithm based on Catani–Seymour dipole factorisation". *JHEP* 03 (2008), p. 038. DOI: `10.1088/1126-6708/2008/03/038`. arXiv: `0709.1027 [hep-ph]` (cited on page 50).

[103]   Stefan Höche et al. "A critical appraisal of NLO+PS matching methods". *JHEP* 09 (2012), p. 049. DOI: `10.1007/JHEP09(2012)049`. arXiv: `1111.1220 [hep-ph]` (cited on page 50).

[104]   Stefan Höche et al. "QCD matrix elements + parton showers. The NLO case". *JHEP* 04 (2013), p. 027. DOI: `10.1007/JHEP04(2013)027`. arXiv: `1207.5030 [hep-ph]` (cited on page 50).

[105]   S. Catani et al. "QCD Matrix Elements + Parton Showers". *JHEP* 11 (2001), p. 063. DOI: `10.1088/1126-6708/2001/11/063`. arXiv: `hep-ph/0109231` (cited on pages 50, 99).

[106]   Stefan Höche et al. "QCD matrix elements and truncated showers". *JHEP* 05 (2009), p. 053. DOI: `10.1088/1126-6708/2009/05/053`. arXiv: `0903.1219 [hep-ph]` (cited on pages 50, 99).

[107]   Gionata Luisoni et al. "$HW^{\pm}/HZ + 0$ and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO". *JHEP* 10 (2013), p. 083. DOI: `10.1007/JHEP10(2013)083`. arXiv: `1306.2542 [hep-ph]` (cited on page 50).

[108]   ATLAS Collaboration. "Measurement of the $Z/\gamma^*$ boson transverse momentum distribution in $pp$ collisions at $\sqrt{s} = 7\,\text{TeV}$ with the ATLAS detector". *JHEP* 09 (2014), p. 145. DOI: `10.1007/JHEP09(2014)145`. arXiv: `1406.3660 [hep-ex]` (cited on page 50).

[109]   J. Pumplin et al. "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis". *JHEP* 07 (2002), p. 012. DOI: `10.1088/1126-6708/2002/07/012`. arXiv: `hep-ph/0201195` (cited on page 50).

[110]   ATLAS Collaboration. *ATLAS Pythia 8 tunes to 7 TeV data*. ATL-PHYS-PUB-2014-021. 2014. URL: `https://cds.cern.ch/record/1966419` (cited on page 51).

[111]   Stefano Frixione et al. "Single-top hadroproduction in association with a *W* boson". *JHEP* 07 (2008), p. 029. DOI: `10.1088/1126-6708/2008/07/029`. arXiv: `0805.3067 [hep-ph]` (cited on page 51).

[112]   ATLAS Collaboration. *VGammaORTool*. URL: `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/VGammaORTool`, last accessed on 01/24/2021 (cited on page 51).

[113]   ATLAS Collaboration. *IFFTruthClassifer*. URL: `https://gitlab.cern.ch/ATLAS-IFF/IFFTruthClassifier/-/tree/master`, last accessed on 01/24/2021 (cited on page 54).

[114] O. Behnke et al. *Data Analysis in High Energy Physics: A Practical Guide to Statistical Methods*. Wiley, 2013. ISBN: 9783527410583. URL: https://books.google.com/books?id=U77IAgAAQBAJ (cited on page 79).

[115] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". *CoRR* abs/1603.02754 (2016). arXiv: 1603.02754. URL: http://arxiv.org/abs/1603.02754 (cited on page 79).

[116] Leo Breiman. "Arcing the edge". *Technical Report 486* (1997). URL: https://www.stat.berkeley.edu/~breiman/arcing-the-edge.pdf (cited on page 80).

[117] Jerome H. Friedman. "Greedy Function Approximation: A Gradient Boosting Machine". *The Annals of Statistics* 29.5 (2001), pp. 1189–1232. ISSN: 00905364. URL: http://www.jstor.org/stable/2699986 (cited on page 80).

[118] ATLAS Collaboration. "Jet energy resolution in proton–proton collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ recorded in 2010 with the ATLAS detector". *Eur. Phys. J. C* 73 (2013), p. 2306. DOI: 10.1140/epjc/s10052-013-2306-0. arXiv: 1210.6210 [hep-ex] (cited on page 90).

[119] ATLAS Collaboration. *Measurement of b-tagging efficiency of c-jets in $t\bar{t}$ events using a likelihood approach with the ATLAS detector*. ATLAS-CONF-2018-001. 2018. URL: https://cds.cern.ch/record/2306649 (cited on page 91).

[120] ATLAS Collaboration. *Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV*. ATLAS-CONF-2018-006. 2018. URL: https://cds.cern.ch/record/2314418 (cited on page 91).

[121] Jon Butterworth et al. "PDF4LHC recommendations for LHC Run II". *J. Phys. G* 43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph] (cited on page 99).

[122] L.A. Harland-Lang et al. "Parton distributions in the LHC era: MMHT 2014 PDFs". *Eur. Phys. J. C* 75.5 (2015), p. 204. DOI: 10.1140/epjc/s10052-015-3397-6. arXiv: 1412.3989 [hep-ph] (cited on page 99).

[123] Sayipjamal Dulat et al. "New parton distribution functions from a global analysis of quantum chromodynamics". *Phys. Rev. D* 93.3 (2016), p. 033006. DOI: 10.1103/PhysRevD.93.033006. arXiv: 1506.07443 [hep-ph] (cited on page 99).

[124] ATLAS Collaboration. *Multi-boson simulation for 13 TeV ATLAS analyses*. ATL-PHYS-PUB-2016-002. 2016. URL: https://cds.cern.ch/record/2119986 (cited on page 99).

[125] Christian Bierlich et al. "Robust Independent Validation of Experiment and Theory: Rivet version 3". *SciPost Phys.* 8 (2020), p. 026. DOI: `10.21468/SciPostPhys.8.2.026`. arXiv: `1912.05451 [hep-ph]` (cited on page 99).

[126] ATLAS Collaboration. *TRExFitter*. URL: `https://trexfitter-docs.web.cern.ch/trexfitter-docs/`, last accessed on 02/01/2021 (cited on page 102).

[127] Stefan Dittmaier, Alexander Huss, and Gernot Knippen. "Next-to-leading-order QCD and electroweak corrections to WWW production at proton-proton colliders". *JHEP* 09 (2017), p. 034. DOI: `10.1007/JHEP09(2017)034`. arXiv: `1705.03722 [hep-ph]` (cited on page 114).

[128] D. de Florian et al. "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector" (2016). DOI: `10.23731/CYRM-2017-002`. arXiv: `1610.07922 [hep-ph]` (cited on page 114).

[129] UA1 Collaboration. "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $S^{(1/2)}$ = 540 GeV". *Phys. Lett. B* 122 (1983), pp. 103–116. DOI: `10.1016/0370-2693(83)91177-2` (cited on page 115).

[130] UA2 Collaboration. "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN $\bar{p}p$ Collider". *Phys. Lett. B* 122 (1983), pp. 476–485. DOI: `10.1016/0370-2693(83)91605-2` (cited on page 115).

[131] ALEPH Collaboration. "Measurement of the $W$ mass in $e^+e^-$ collisions at production threshold". *Phys. Lett. B* 401 (1997), pp. 347–362. DOI: `10.1016/S0370-2693(97)00460-7` (cited on page 115).

[132] DELPHI Collaboration. "Measurement and interpretation of the $W$ pair cross-section in $e^+e^-$ interactions at 161 GeV". *Phys. Lett. B* 397 (1997), pp. 158–170. DOI: `10.1016/S0370-2693(97)00226-8` (cited on page 115).

[133] L3 Collaboration. "Pair production of $W$ bosons in $e^+e^-$ interactions at $\sqrt{s}$ = 161 GeV". *Phys. Lett. B* 398 (1997), pp. 223–238. DOI: `10.1016/S0370-2693(97)00253-0` (cited on page 115).

[134] OPAL Collaboration. "Measurement of the mass of the $W$ boson in $e^+e^-$ collisions at $S^{(1/2)}$ = 161 GeV". *Phys. Lett. B* 389 (1996), pp. 416–428. DOI: `10.1016/S0370-2693(96)01452-9` (cited on page 115).

[135] Kaustubh Agashe et al. "Dedicated Strategies for Triboson Signals from Cascade Decays of Vector Resonances". *Phys. Rev. D* 99.7 (2019), p. 075016. DOI: `10.1103/PhysRevD.99.075016`. arXiv: `1711.09920 [hep-ph]` (cited on page 115).