Standard Model Precision Measurements with two Z bosons and two jets in ATLAS.

A Dissertation

Presented to

The Faculty of the Graduate School of Arts and Sciences
Brandeis University

Department of Physics

Professor Gabriella Sciolla, Advisor

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
Prajita Bhattarai

May 2023

This dissertation, directed and approved by Prajita Bhattarai's Committee, has been accepted and approved by the Faculty of Brandeis University in partial fulfillment of the requirements for the degree of:

DOCTOR OF PHILOSOPHY

Wendy Cadge, Dean
Graduate School of Arts and Sciences

Dissertation Committee:
Professor Gabriella Sciolla, Department of Physics, Brandeis University
Professor Aram Apyan, Department of Physics, Brandeis University
Dr. Alessandro Tricoli, Brookhaven National Laboratory

Acknowledgements

I am sincerely grateful to many people who have had a signifcant impact on my journey.

ABSTRACT

Standard Model Precision Measurements with two Z bosons and two jets in ATLAS.

A dissertation presented to the Faculty of the
Graduate School of Arts and Sciences of Brandeis University
Waltham, Massachusetts

By Prajita Bhattarai

This thesis presents the cross-section measurements of two Z bosons production in association with two jets in $ZZ^{*}(\rightarrow 4\ell)jj$ final state, differentially as a function of several kinematic observables, which are sensitive to the vector boson scattering. The electroweak $ZZjj$ production includes the rare triple and quartic self-couplings of gauge bosons, whose scattering amplitude at high energies is regularized by the Standard Model $H \rightarrow ZZ^{*}$ processes. The analysis is performed using the proton-proton collision data collected by the ATLAS experiment during LHC Run-2 at $\sqrt{s} = 13$ TeV center-of-mass collision energy, corresponding to the integrated luminosity of 139 fb^{-1}. Several beyond the Standard Model theories are expected to modify the electroweak $ZZjj$ cross-sections at high energies. Therefore, it is imperative to measure cross-section of electroweak $ZZjj$ processes differentially. Given the low statistics in Run-2, the detector effects corrected differential cross-sections are measured in electroweak-enhanced phase space and compared to the state-of-the-art Standard Model predictions. The differential cross-sections are also used to constrain the anomalous quartic gauge couplings using a dimension-8 Effective Field Theory formalism.

Table of Contents

List of Tables

List of Figures

- ATLAS: A Torodial LHC Apparatus
- BSM: Beyond the Standard Model
- C: Charge conjugation
- CERN: Conseil Européen pour la Recherche Nucléaire
- CR: Control Region
- CSC: Cathode Strip Chambers
- ECAL: Electromagnetic Calorimeter
- EFT: Effective Field Theory
- EWK: Electroweak
- fJVT: forward Jet Vertex Tagger
- FSR: Final State Radiation
- GRL: Good Run List
- H: Weak Hypercharge
- HCAL: Hadronic Calorimeter
- HF: Heavy Flavor
- HGTD: High Granularity Timing Detector
- HL-LHC: High Luminosity Large Hadron Collider
- HLT: High-level Trigger
- I: Weak Isospin
- IBL: Insertable B-Layer

- ID: Inner Detector
- IP: Interaction Point
- IFF:Isolation and Fake Forum
- JVT: Jet Vertex Tagger
- $\mathcal{L}_{\mathcal{SM}}$: Lagrangian
- LAr: Liquid Argon
- LB: Luminosity Block
- LEP: Large Electron Positron Collider
- LF: Light Flavor
- LH: Left Handed
- LHC: Large Hadron Collider
- MC: Monte Carlo
- MS: Muon Spectrometer
- P: Parity
- PS: Proton Synchroton
- PDF: Parton Distribution Function
- Q: Electric Charge
- QGC: Quartic Gauge Coupling
- QED: Quantum Electrodynamics
- QCD: Quantum Chromodynamics

- ($\mathcal{P}$): Poincare group
- RF: Radio frequency
- RH: RightHanded
- SCT: Semiconductor Tracker
- SF: Scale factors
- SF-OC: Same-flavor, Opposite-charged
- SM: Standard Model
- SR: Signal Region
- SPS: Super Proton Synchroton
- T: Time-reversal
- TGC: Triple Gauge Coupling
- TRT:Transition Radition Tracker
- TTVA: Track-to-vertex association
- VBS: Vector Boson Scattering
- VR: Validation Region
- VRSC: Same Charge Validation Region
- VRDF: Different Flavor Validation Region
- VEV: Vacuum Expectation Value

Chapter I: Introduction

Research in particle physics investigates the fundamental nature of the universe. The Standard Model (SM), the fundamental theory of particle physics, provides a theoretical formulation that explains all known elementary particles, their interactions, and three of the four fundamental forces observed in nature; strong, electromagnetic, and weak forces. Fifty years after its formulation, the SM predicted theory parameters are measured experimentally with high precision. The experimental discovery of the Higgs boson in 2012 established SM as a complete and highly successful theory. However, the lack of description of the fourth fundamental force, gravity, and other experimentally evident phenomena, such as the existence of dark matter, infer that the SM is an incomplete description of nature. Still, experimental evidence of new physics beyond the Standard Model (BSM) has yet to be observed. The current primary objective of the Large Hadron Collider (LHC) at CERN is to look for experimental evidence of new physics which might explain or resolve some of the shortcomings of the SM.

New physics searches can be broadly categorized into two types, direct and indirect. The direct search focuses on finding experimental evidence of new physics signatures directly. In contrast, the indirect approach focuses on precisely measuring the parameters of the SM-predicted processes, looking for deviations compared to the state-of-the-art theoretical predictions. One critical phenomenon of the SM is the vector boson scattering (VBS) in the final state involving multi vector-bosons, electroweak (electromagnetic & weak) force mediating particles. VBS processes consist of the rare triple and quartic self-couplings between the electroweak force's mediator, whose SM amplitude interferes destructively with the Higgs-mediated processes. Several BSM theories modify either strength of the self-couplings or that of the Higgs-mediated processes, thus, altering the extent of the interference and, consequently, the cross-sections from the predicted value. As many new physics particles are expected to exist at high energies, such deviations are expected to appear at higher energies that haven't been probed experimentally.

This thesis presents an indirect approach to new physics searches in one of the VBS-sensitive multiboson final states. The measurement analyzes the data collected by the ATLAS experiment of the LHC from 2015-2018 to measure the VBS-sensitive production of two Z bosons in association with two jets, where each Z boson decays into a pair of same-flavor opposite-charge (SF-OC) lepton pair. The quartic self-coupling of the vector bosons in $ZZ(\rightarrow 4\ell)jj$ final state has been experimentally accessible with the collected LHC dataset for the first time. Thus, measurements presented in this thesis are at the frontier of particle physics, pushing the boundaries of new physics searches through an indirect approach.

First, the theory of SM, its shortcomings, and the $ZZ(\rightarrow 4\ell)jj$ process are discussed in Chapter II. The LHC and ATLAS experiments are then introduced in Chapter III. Chapters IV and V discuss the details of the measurement whose results are presented in Chapter VI.

Chapter II: Theory

This chapter describes the theoretical framework of the experimental measurements discussed in this thesis. Section 1 introduces the Standard Model of particle physics and concepts relevant to the thesis. Section 2 discusses the outstanding problems with the SM, thus, motivating the experimental measurement. Section 3 discusses the phenomenology of the proton-proton collisions, and Section 4 discusses the physics of two Z bosons production in an association of two jets.

1 The Standard Model

The SM of particle physics is a mathematical framework based on quantum field theory, which incorporates quantum mechanics and special relativity. The SM describes all known fundamental particles in nature and their interactions. It consists of two sets of particles with intrinsic angular momentum, half-integer-spin fermions that are fundamental constituents of matter particles, and force-carrying integer-spin bosons. The seventeen fundamental particles of the SM and their properties, such as mass, charge, and intrinsic spin, are shown schematically by figure 1. Two textbooks on particle physics, Mark Thomson's Modern Particle Physics [20], and Halzen & Martin's Quarks & Leptons [2] guide the discussion written in this section.

1.1 Symmetries

The fundamental particles of the SM and their interactions can be described by constructing a general renormalizable Lagrangian ($\mathcal{L}_{SM}$) that respects certain sets of given symmetries. The Lagrangian of the SM is independent of the reference frame, naturally respecting the complete external symmetries of special relativity, the Poincare group ($\mathcal{P}$). Thus, the SM is invariant under spacetime translations, boosts, and rotations. Additionally, by construct of the Lagrangian, the SM respects an internal local gauge symmetry $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. The $SU(3)_C$ symmetry is associated with the Quantum Chromodynamics (QCD) discussed in detail in Section 1.3.3. The $SU(2)_L \otimes U(1)_Y$ gauge symmetry discussed in 1.3.4 is associated with the unified electroweak theory that combines Quantum Electrodynamics (QED) and the weak interactions.

According to Noether's theorem, a quantity is conserved for each continuous transformation that leaves the Lagrangian invariant [21]. Several interesting physical quantum numbers are conserved as a consequence of the symmetries respected by the SM. The $SU(3)_C$ in QCD conserves color charge. Weak isospin (I) and weak hypercharge (Y) are the quantum numbers associated with the $SU(2)_L$ and $U(1)_Y$ gauge groups, respectively. At low energies the $SU(2)_L \otimes U(1)_Y$ symmetry is spontaneously broken and will be discussed in Section 1.3.4. The $SU(2)_L$ group follows a chiral structure where the gauge fields couple explicitly to the left-handed (LH) chiral fermions states and the right-handed (RH) chiral anti-fermions states.

The SM also respects CPT symmetry, a combination of three additional discrete symmetries, charge conjugation (C), parity (P), and time-reversal (T). The charge-conjugation transformation transforms particles to anti-particles by reversing the quantum numbers, whereas the parity transformation transforms left-handed particles to right-handed particles and vice-versa.

1.2 Particles and Fields

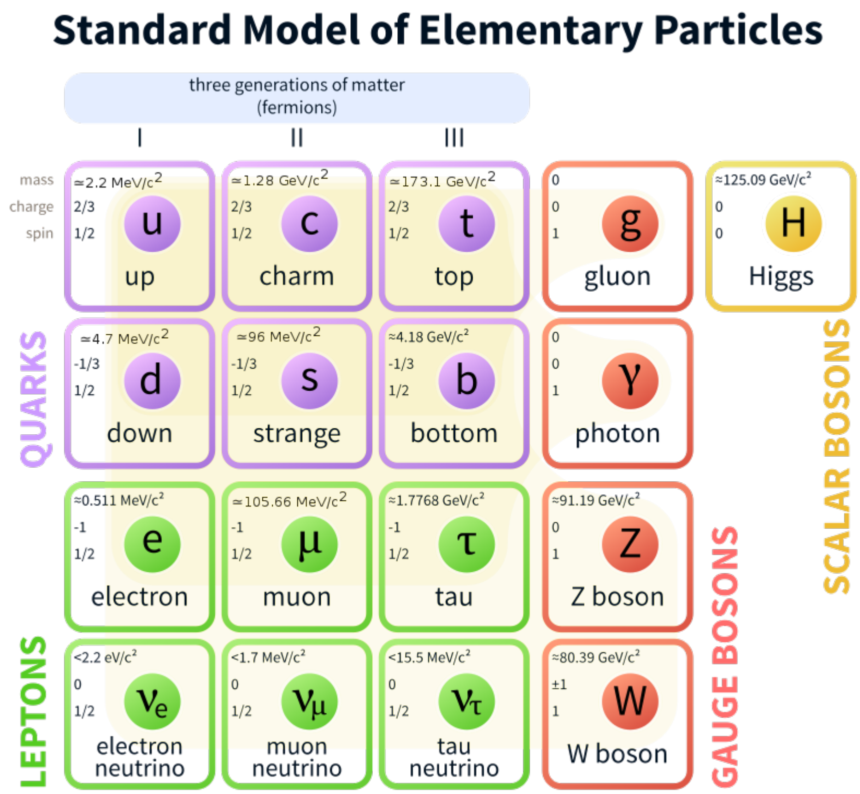


Figure 1: The seventeen fundamental particles of the SM include three generations of twelve fermions, four gauge bosons, and the scalar Higgs bosons. [3]

The twelve half-integer-spin fermions can be distinguished further into two categories, leptons and quarks, each having three generations of particles with similar properties as shown schematically by Figure 1. For each fermion, there exists an anti-fermion with the same additive quantum numbers but with opposite signs. Four spin 1 bosons shown in Table 1 are collectively called the gauge bosons. Quanta of these gauge fields mediate the electromagnetic, weak, and strong interactions and are invariant under various local gauge transformations [22]. As summarized in Table 2, different fermions participate in different interactions. A gauge coupling parameter governs the strength of the interaction.

Massless photon (γ) mediates the electromagnetic interaction, whereas the massive W

Table 1: Properties of SM gauge bosons. [1]

Interaction Type		Particle	Q	Mass $[GeV]$	Symmetry Group
Electroweak	Electromagnetic	Photon (γ)	0	0	$SU(2) \otimes U(1)$
	Weak	$W^{\pm}$	± 1	80.4	
		Z boson	0	91.2	
Strong		gluons (g)	0	0	$SU(3)$

Table 2: Summary of different interactions of fermions under different gauge theory. The check mark suggests that the fermions interact via associated force.

Particles		Strong $SU(3)$	Electromagnetic $U(1)$	Weak $SU(2)$
Quarks	u, c, t d, s, b	✓	✓	✓
Leptons	e, μ, τ	-	✓	✓
	ν_e, ν_μ, ν_τ	-	-	✓

and Z bosons mediate weak interaction. The electric charge (Q), which is conserved in all interactions, is related to the isospin and hypercharge by $Q = I_3 + \frac{Y}{2}$, where I_3 is the third component of the weak isospin. As a consequence of the chiral structure of $SU(2)_L$, each generation of fermion contains a left-handed doublet with $I_3 = \pm\frac{1}{2}$ and a right-handed singlet carrying $I_3 = 0$ as shown in Table 3.

Each generation of lepton, electron (e), muon (μ) and tau (τ) is accompanied by a neutral particle called neutrino (ν) with same lepton flavor (ν_e, ν_μ & ν_τ). In SM, anti-neutrinos have an opposite lepton flavor quantum number than neutrinos. The lepton flavor is conserved in all interactions.

The quarks are categorized further into two categories, the up-type quarks with $+\frac{2}{3}$ charge and the down-type quarks with $-\frac{1}{3}$ charge. Up (u), charm (c) , & top (t) are the first, second, and third generation of the up-type quarks, while the down (d), strange (s) & bottom (b) are the three generations of the down-type quarks. The quarks interact strongly with one another by strong interaction mediated by the massless neutral gluons, which follow from $SU(3)$ gauge symmetry by exchange of color charges. Each quark can have either one

Table 3: Electroweak quantum numbers of the SM half-integer spin fermions (quarks and leptons) arranged in a left-handed $SU(2)$ doublet and right-handed $SU(2)$ singlet. The down-type left-handed quarks in $SU(2)_L$ quark doublets $d^{'}$, $s^{'}$ & $b^{'}$ are linear combinations of d, s, b quarks [2].

Particle Types	First	Second	Third	I_3	Y	Q
Leptons	$\begin{pmatrix} e \\ \nu_e \end{pmatrix}_L$	$\begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}_L$	$\begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}_L$	$-\frac{1}{2}$ $\frac{1}{2}$	-1 -1	-1 0
	e_R	μ_R	τ_R	0	-2	-1
Quarks	$\begin{pmatrix} u \\ d^{'} \end{pmatrix}_L$	$\begin{pmatrix} c \\ s^{'} \end{pmatrix}_L$	$\begin{pmatrix} t \\ b^{'} \end{pmatrix}_L$	$\frac{1}{2}$ $-\frac{1}{2}$	$\frac{1}{3}$ $\frac{1}{3}$	$\frac{2}{3}$ $-\frac{1}{3}$
	u_R	c_R	t_R	0	$\frac{4}{3}$	$\frac{2}{3}$
	d_R	s_R	b_R	0	$-\frac{2}{3}$	$-\frac{1}{3}$

of the three color charges (red, blue &, green), whereas an anti-quark can have either an anti-red, anti-blue, or anti-green color charge. There are eight gluons in the SM with color charges formed by a linear combination of color with an anti-color charge excluding the colorless combination. Since gluons have a color charge, they interact with other gluons strongly. Only color-neutral hadronic states formed by a combination of quarks and gluons are observed experimentally.

Higgs boson is the only spin-0 scalar particle in the SM and has no charge. It gives masses to the three weak bosons ($W^{\pm}$ & Z), all quarks, and charged leptons through Spontaneous Symmetry Breaking, which is discussed in Section 1.3.4.

1.3 Theoretical Formulation of the Standard Model

Relativistic quantum field theory is the theoretical framework of the SM that describes elementary particles and their interactions. This section introduces the framework.

1.3.1 Lagrangian of the Standard Model

The compact Lagrangian density given in equation 1.1 describes the dynamics of the SM and is invariant under the local gauge transformation of the $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ symmetry group.

$$\mathcal{L}_{\mathcal{SM}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \ + \ i\bar{\psi}\gamma^{\mu}D_{\mu}\psi \ + \ |D_{\mu}\phi|^2 \ - \ V(\phi) + \bar{\psi}_i y_{ij}\psi_j\phi \ + \ h.c. \tag{1.1}$$

The first term $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ describes the dynamics of the gauge boson interactions, the second term $i\bar{\psi}\gamma^{\mu}D_{\mu}\psi$ describes the interaction of the fermion fields. The third term $|D_{\mu}\phi|^2$ describes the couplings between the Higgs boson and gauge bosons, whereas the term $V(\phi)$ represents the Higgs potential and its self-interactions. The second last term $\bar{\psi}_i y_{ij}\psi_j\phi$ generates masses for fermions based on their Yukawa couplings y_{ij} to the Higgs field. Similarly, the last term $h.c.$ generates masses for anti-fermions. In reality, the SM lagrangian density, as shown by equation 1.1, is a combination of complex interactions that follow different types of local gauge symmetries, and concise derivations of different interactions are discussed individually in the following few sections.

1.3.2 Quantum Electrodynamics

Quantum electrodynamics describes the electromagnetic interaction. The Dirac Lagrangian density ($\mathcal{L}_{Dirac}$) describes the free propagation of a fermion in a vacuum as:

$$\mathcal{L}_{Dirac} = \bar{\psi}i\gamma^{\mu}\partial_{\mu}\psi \ - \ m\bar{\psi}\psi \tag{1.2}$$

where ψ is the fermionic spinor, γ^{μ} represents the Dirac matrices with μ being the Lorentz index running from 0 to 3, and m is the mass of the fermion.

The Lagrangian in equation 1.2 is invariant under a $U(1)$ global gauge transformation,

$$\psi \rightarrow \psi^{'} = e^{iq\alpha}\psi \tag{1.3}$$

where q is a parameter of the transformation itself and α is a real phase factor. However, under the local gauge transformation of form

$$\psi \rightarrow \psi^{'} = e^{iq\alpha(x)}\psi \tag{1.4}$$

where α depends on $x = (x_0, x_1, x_2, t)$ the Dirac Lagrangian in equation 1.2 is not invariant.

To make the Lagrangian of equation 1.2 invariant, a gauge field A_μ with the following transformation properties is introduced,

$$A_\mu \rightarrow A_\mu - \partial_\mu \alpha \tag{1.5}$$

A_μ couples to fermionic fields $\psi(x, t)$ with strength q. A covariant derivative specific to the local gauge transformation is defined as:

$$D_\mu = \partial_\mu - iqA_\mu \tag{1.6}$$

The quantity q can be interpreted as the electric charge $-e$ of fermion, which gives the coupling strength of QED. With these substitutions, the Dirac Lagrangian in equation 1.2 changes to the following

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu \ - \ m)\psi \tag{1.7}$$

which is invariant under $U(1)$ gauge transformation respecting the $U(1)$ gauge symmetry.

The gauge field A_μ can be interpreted as the photon field and is related to the electromagnetic field tensor by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{1.8}$$

The gauge invariant kinetic term of photon $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ can be inserted into the Lagrangian in equation 1.7 which gives us the full Lagrangian of QED invariant under $U(1)$ gauge transformation.

$$\mathcal{L}_{QED} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\gamma^{\mu}D_{\mu} \ - \ m)\psi \tag{1.9}$$

$\mathcal{L}_{QED}$ in equation 1.9 is the full Lagrangian for QED, and the electromagnetic phenomena can be derived by solving for the equations of motion applying the Lorentz gauge condition $\partial_{\mu}A^{\mu} = 0$. A small dimensionless coupling constant α determines the relative strength of the QED interactions. The probability of particle interactions and decay rates can be calculated as the square modulus of a perturbative series in powers of α. The value of α depends on the energy scale because of the additional contributions from the vacuum polarization of virtual quark and virtual lepton loop corrections. For a given momentum transfer squared q^2 of the exchanged photon, the effective coupling constant depends on the $\alpha_0 = \frac{e^2}{4\pi}$ with zero momentum transfer as,

$$\alpha(q^2) = \frac{\alpha_0(q^2 = 0)}{1 - \Delta\alpha(q^2)} \tag{1.10}$$

with the small value of $\alpha_0 \approx \frac{1}{137.036}$ [23].

1.3.3 Quantum Chromodynamics

Quantum Chromodynamics defines the interaction between the quarks, requiring $SU(3)$ gauge transformation on the quark field with color charge j (red, blue, or green). The Dirac Lagrangian for a quark can be written similarly to that of an electron with a small modification to include all possible colors of quark field q_j as

$$\mathcal{L} = \bar{q_j}(i\gamma^{\mu}\partial_{\mu} - m)q_j \tag{1.11}$$

The generators of the $SU(3)$ group are eight linearly independent traceless Gell-Mann matrices that do not commute with each other such that

$$[T_a, T_b] = if_{abc}T_c \tag{1.12}$$

where f_{abc} is the structure constant of $SU(3)$

The local $SU(3)$ gauge transformation is

$$q(x) \rightarrow e^{i\alpha_a(x)T_a} q(x) \tag{1.13}$$

where $T_a = \frac{\lambda_a}{2}$, and $a = 1, 2...8$. To understand the source of gauge invariance in the Lagrangian of equation 1.11, an infinitesimal transformation of the color field is considered as

$$q(x) \rightarrow [1 + i\alpha_a(x)T_a]q(x) \ni \ \partial_\mu q \rightarrow (1 + i\alpha_a T_a)\partial_\mu q + iT_a q\partial_\mu \alpha_a \tag{1.14}$$

The last term $iT_a q\partial_\mu \alpha_a$ breaks the gauge invariance. Similar to QED, eight gauge fields corresponding to each $a = 1, 2...8$ G^a_μ with following transformation properties are introduced

$$G^a_\mu \rightarrow G^a_\mu - \frac{1}{g_s}\partial_\mu \alpha_a - f_{abc}\alpha_b G^c_\mu \tag{1.15}$$

These gauge fields G^a_μ are the gluon fields. Similar to QED, the covariant derivative is defined as

$$D_\mu = \partial_\mu + ig_s\frac{\lambda_a}{2}G^a_\mu \tag{1.16}$$

where g_s is the coupling strength of the gluon fields to the quark fields.

The Lagrangian density in equation 1.11 is then

$$\mathcal{L} = \bar{q}_j(i\gamma^\mu D_\mu - m)q_j \tag{1.17}$$

Similar to QED, a gauge-invariant kinetic term $-\frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a$, dependent on the field strength tensor $G^a_{\mu\nu}$ is added to equation 1.17 to give the full QCD Lagrangian. The kinetic terms allow self-interaction within the gluon fields, which is an important feature of QCD. $G^a_{\mu\nu}$ is the field strength tensor defined as

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f_{abc} G^b_\mu G^c_\nu \tag{1.18}$$

Therefore, the complete $SU(3)$ gauge invariant Lagrangian describing the quarks and gluons interaction is

$$\mathcal{L}_{QCD} = \bar{q}_j (i\gamma^\mu D_\mu - m) q_j - \frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a \tag{1.19}$$

Similar to QED, the probability of particle interactions and decay rates in QCD can be calculated as the square-modulus of a perturbative series in powers of the strong coupling constant α_S $(= \frac{g_S^2}{4\pi})$ which depends on the energy scale. The energy-dependent effective strong coupling is,

$$\alpha_S(q^2) = \frac{12\pi}{(33 - 2n_f).\ln(q^2/\Lambda_{QCD})} \tag{1.20}$$

where n_f is the number of types of quarks with masses lower than the energy scale, and $\Lambda_{QCD} \approx 0.2$ GeV is the *Landau pole* below which the coupling constant increases and diverges [24]. There are two important experimental consequences of the running of the strong coupling constant; first, quarks and gluons cannot be isolated at low energies, leading to *color confinement*. Second, at high energies, the strong interaction reaches *asymptotic freedom*, and the exchange of gluons requires minimum energy, causing an abundance of gluon-induced QCD processes at hadron colliders [24].

1.3.4 Electroweak Theory

Weak interactions describe the interactions mediated by massive gauge bosons, for example, radioactive beta decay, where a neutron turns into a proton by emitting an electron and an anti-electron-neutrino. The Dirac equation formulated in 1930s could explain the motions of electrons via relativistic quantum mechanics. However, the nuclear decay processes were still a mystery. Fermi developed the first theory of weak interaction to explain beta decay using four fermion interaction vertex. The formulation successfully describes the beta decay at low energies when the interaction energy is much smaller than the W boson mass. In the meantime, QED following the $U(1)$ symmetry was formulated to explain the electromagnetic interaction. Experimental evidence suggested that an exchange of spin-1 massive particles mediates the weak interactions. It was challenging to develop a local gauge invariant mathematical theory, including spin-1 massive gauge bosons, which conserves unitarity at high energies.

During the 1960s, Glashow, Salam, and Weinberg (GWS) worked independently and made different contributions to formulate a theoretical model of weak interactions following a local gauge invariance of $SU(2)_L \otimes U(1)_Y$ [25] [26] [27]. The theory postulates the existence of four massless gauge bosons, two electrically-charged and two electrically-neutral, which mediate unified electromagnetic and weak interactions. However, the observed short range of weak force could only be explained with massive gauge bosons of electroweak interactions. Therefore, implying the underlying symmetry of weak interactions is broken by some mechanism, which was later understood through the Higgs Mechanism discussed in Section 1.3.5.

Experimental observations suggest weak interactions violate parity by only affecting the left-handed fermion and right-handed anti-fermion fields. Thus, the unified electroweak theory is described by $SU(2)_L \otimes U(1)_Y$ gauge interactions. Similar to the electric charge Q conserved in QED by $U(1)$ symmetry, the weak hypercharge ($Y = 2(Q - I_3)$) related to the electric charge and the weak isospin (I_3) is conserved by the $U(1)_Y$ symmetry. The fermion fields are represented by the left-handed doublets χ_L and the right-handed singlets ψ_R, introduced in table 3. The doublet and singlet field for the first generation of leptons and quarks are,

$$\chi_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \quad \& \quad \chi_L = \begin{pmatrix} u \\ d \end{pmatrix}_L$$
$$\psi_R = e_R \quad \& \quad \psi_R = u_R \;\&\; d_R$$

The Lagrangian for these fermion fields should be invariant under local gauge transformation corresponding to both $SU(2)_L$ and $U(1)_Y$ symmetry as,

$$\chi_L \rightarrow e^{i\beta(x)Y + i\alpha_a(x)\tau_a}\chi_L \tag{1.21}$$

$$\psi_R \rightarrow e^{i\beta(x)Y}\psi_R \tag{1.22}$$

where, $\beta(x)$ and $\alpha(x)$ are the local phase transformation for $U(1)_Y$ and $SU(2)_L$ symmetry groups respectively. Weak hypercharge operator Y and Pauli matrices $\tau_{a=1,2,3}$ are generators of $U(1)_Y$ and $SU(2)_L$ symmetry groups respectively. Similar to the formulation in QED and QCD discussed in Section 1.3.2 and 1.3.3, four new field strength tensors $B_{\mu\nu}$ and $W^a_{\mu\nu}$ corresponding to respectively the $U(1)_Y$ and $SU(2)_L$ transformations are introduced. The $SU(2)_L \otimes U(1)_Y$ gauge-invariant Lagrangian for a massless fermion and massless gauge fields is:

$$\mathcal{L}_0 = \bar{\chi_L}\gamma^\mu[i\partial_\mu - g\frac{\tau_a}{2}W^a_\mu + \frac{g^{'}}{2}B_\mu]\chi_L + \bar{\psi_R}\gamma^\mu[i\partial_\mu + g^{'}B_\mu]\psi_R - \frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{1.23}$$

where similar to QED and QCD, field strength tensors are defined in terms of the covariant derivative to preserve gauge-invariance in kinetic terms as,

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.24}$$

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g\epsilon^{abc}W^b_\mu W^c_\nu \tag{1.25}$$

The non-Abelian part of the $SU(2)_L$ transformation is represented by the last term of equation 1.25, which gives the quartic and triple self-interactions between the gauge bosons with coupling strength g.

The electroweak Lagrangian in equation 1.23 contains two terms, one of which gives rise to the charged-current interaction with the two $SU(2)$ boson field

$$W_\mu^\pm \;=\; \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt(2)} \tag{1.26}$$

via exchange of the $W^\pm$ bosons and the neutral current interactions from the two neutral gauge boson fields W_μ^3 and B_μ.

The Lagrangian for the charged-current interaction for the first generation of quarks and leptons is,

$$\mathcal{L}_{CC} = \frac{g}{2\sqrt{2}}\{W_\mu^\dagger[\bar{u}\gamma^\mu(1-\gamma_5)d + \bar{\nu_e}\gamma^\mu(1-\gamma_5)e] \;+\; h.c\} \tag{1.27}$$

The $SU(2)_L$ charged-current interaction Lagrangian for the next two generations follows the same, establishing the universality of the quark and lepton interactions as a direct consequence of the gauge symmetry.

The neutral-current Lagrangian is given by,

$$\mathcal{L}_{NC} = \sum_j \bar{\psi_j}\gamma^\mu\{A_\mu[g\frac{\tau_3}{2}\sin\theta_W + g^{'}Y\cos\theta_W] + Z_\mu[\frac{\tau_3}{2}\cos\theta_W - g^{'}Y\sin\theta_W]\}\psi_j \tag{1.28}$$

where the two neutral gauge fields Z_μ and A_μ associated with Z boson and photon governing the weak neutral and electromagnetic interactions are obtained from an arbitrary linear combination of the W_μ^3 and B_μ fields as

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} cos\theta_W & sin\theta_W \\ -sin\theta_W & cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} \tag{1.29}$$

The following condition is imposed to obtain QED from A_μ:

$$g\sin\theta_W = g^{'}\cos\theta_W = e \;\&\; Y = Q - T_3 \tag{1.30}$$

where $T_3 = \frac{\tau_3}{2}$ is the weak isospin and θ_W is the Weinberg mixing angle, which has been measured experimentally with high precision and can be expressed in terms of the two

$SU(2)_L$ coupling $g^{'}$ and $U(1)_Y$ coupling g as:

$$\sin\theta_W = \frac{g^{'}}{\sqrt{g^2 + g^{'2}}} \;\;\&\;\; \cos\theta_W = \frac{g}{\sqrt{g^2 + g^{'2}}} \tag{1.31}$$

Similar to QED and QCD, the probability of particle interactions and decay rates in electroweak interactions can be calculated as the square-modulus of a perturbative series in powers of the weak coupling constant α_{EWK}. The value of α_{EWK} depends on the energy scale and the type of weak interaction, charged-current, neutral-current, or the non-Abelian self-couplings of the gauge bosons.

The Lagrangian in equation 1.23 describes the electroweak interactions only for massless fermions and massless gauge bosons, which contradicts the experimental observations. The mass origin of the fermions and gauge bosons is discussed below in Section 1.3.5.

1.3.5 Higgs Mechanism

Massive gauge bosons in the Lagrangian 1.23 can be accommodated through the Brout-Englert-Higgs (BEH) mechanism by introducing a complex scalar field ϕ in the spinor representation of $SU(2)_L$ doublet as [28],

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{1.32}$$

A new term in the SM Lagrangian $\mathcal{L}_{BEH}$ depending on this scalar field can be defined as,

$$\mathcal{L}_{BEH} = (D_\mu\phi)^\dagger(D^\mu\phi) - \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \tag{1.33}$$

The first term $(D_\mu\phi)^\dagger(D^\mu\phi)$ describes the kinematic of the new fields, and the BEH potential $V(\phi)$ is given by the second term as,

$$V(\phi) = \lambda(\phi^\dagger\phi)^2 - \mu^2\phi^\dagger\phi \tag{1.34}$$

where the term $\lambda(\phi^\dagger\phi)^2$ describes the quartic self-interactions of the scalar fields, and the vacuum stability imposes $\lambda > 0$.

For $\mu^2 > 0$, the scalar field develops a nonzero Vacuum Expectation Value (VEV) which spontaneously breaks the symmetry. Due to the symmetry of $V(\phi)$ an infinite number of degenerate states exists with the potential v only depending on the combination of $\phi^\dagger\phi$ [29] with minimum energy satisfying $\phi^\dagger\phi = \frac{v^2}{2}$. This minimum energy requirement reduces one of the four degrees of freedom of the complex scalar field ϕ. A gauge transformation can eliminate the three remaining degrees of freedom. We can choose ϕ by eliminating the upper component and the imaginary part of the lower component of the complex scalar field as,

$$< \phi >= \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \quad ; \; H(x) = H^*(x)\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.35}$$

where the Higgs field (H) emerges as the excitation from the vacuum state, this choice of the minimum spontaneously breaks the gauge symmetry [30].

After substituting the ϕ in the Lagrangian in equation 1.33, the kinetic term takes the form

$$\begin{aligned} \mathcal{L}_{BEH\ Kinetic} = & \frac{\lambda}{2}v^4 \\ & + \frac{1}{2}\partial_\mu H \partial^\mu H - \lambda v^2 H^2 + \frac{\lambda}{\sqrt{2}} v H^3 + \frac{\lambda}{8}H^4 \\ & + \frac{1}{4}(v + \frac{1}{\sqrt{2}}H)^2 (W_\mu^1 \quad W_\mu^2 \quad W_\mu^3 \quad B_\mu) \begin{pmatrix} g^2 & 0 & 0 & 0 \\ 0 & g^2 & 0 & 0 \\ 0 & 0 & g^2 & gg' \\ 0 & 0 & gg' & g^2 \end{pmatrix} \begin{pmatrix} W^{1\mu} \\ W^{2\mu} \\ W^{3\mu} \\ B^\mu \end{pmatrix} \end{aligned} \tag{1.36}$$

where the first line is the vacuum energy density and can be ignored in the case of QFT. The second line describes the triple and quartic self-interactions of the Higgs field as well as the mass term of the real scalar field H as $m_H = 2\lambda v^2$. The last line contains the mass term for the vector bosons.

From equations 1.36 and 1.26 is evident the mass of the two charged vector bosons $W^{\pm}$ is $m_W = \frac{1}{2}g^2v^2$. Similarly, from equations 1.36 and 1.29, mass of the Z boson is $m_Z = \frac{1}{2}(g^2 + g^{'})v^2$ and mass of the photon is $m_\gamma = 0$.

The initial $SU(2)_L$ Lagrangian in equation 1.33 started with four gauge symmetries, which is reduced to a single $U(1)_Q$ gauge symmetry associated with the massless vector field in equation 1.36. This phenomenon in the Higgs mechanism is called the Electroweak Symmetry Breaking (EWSB) mechanism. As discussed above, the EWSB mechanism is at the heart of the SM by which the gauge boson gets the mass which also gives rise to the longitudinal polarization of the massive vector bosons. This thesis summarizes a measurement with an experimental sensitivity to a such important property of the theory.

The last remaining piece in the SM is adding the fermion mass to the Lagrangian. A simple Lagrangian with the fermion mass can be written as,

$$\mathcal{L}_{mass\ fermion} = -m(\bar{\chi_L}\psi_R + \bar{\psi}_R\chi_L) \tag{1.37}$$

This term violates $SU(2)_L$ gauge symmetry because the left-handed fermions are doublets, and the right-handed are singlets. Adding a scalar complex field $\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}$ in the Lagrangian becomes,

$$\mathcal{L}_{Yukawa,\ \ell} = \frac{G_\ell v}{\sqrt{2}}(\bar{\chi_L}\psi_R + \bar{\psi}_R\chi_L) - \frac{G_\ell}{\sqrt{2}}(\bar{\chi_L}\psi_R + \bar{\psi}_R\chi_L)H \tag{1.38}$$

with arbitrary parameters $G_{\ell=e,\mu,\tau}$. The constant in the first term $\frac{G_\ell v}{\sqrt{2}}$ represents the mass of the fermions, whereas the second term gives the interaction of fermions with the Higgs field.

Similarly, the mass terms for quarks follow but including the down-type quarks, the parameters corresponding to G_ℓ are matrices G_q^{ij} for the quark families i, j and up-type and down-type quarks as:

$$\mathcal{L}_{Yukawa,\ Q} = -G_d^{ij}(\bar{u}_i, \bar{d}_i)_L \phi d_{jR} - G_u^{ij}(\bar{u}_i, \bar{d}_i)_L \phi u_{jR} + h.c. \tag{1.39}$$

The final Standard Model Lagrangian is the sum of the QED (equation 1.9), QCD (equation 1.19), electroweak interactions including the self-interactions of vector bosons (equation 1.23), Higgs potential, Higgs self-interactions, and Higgs-vector boson interactions (equation 1.33), and the Higgs-fermion Yukawa coupling (equations 1.38 & 1.39), which in compact form is written as equation 1.1. The final electroweak theory of SM with massive gauge bosons tightly constrains the W, Z, and Higgs interactions, their masses, and the self-couplings of gauge bosons in terms of a few parameters. Any deviations from predictions indicate the presence of physics beyond the SM, thus, losing the mathematical underpinnings of the theory. Therefore, this thesis targets precision measurement of the electroweak sector to detect possible deviations caused by BSM effects.

2 Limitations of the Standard Model

Many discoveries have experimentally validated the Standard Model's predictions since the 20^{th} century. The breakthrough discovery of the Higgs boson in 2012 at the LHC validated the last piece of the theory [31]& [32]. Many predicted parameters, such as production cross-sections and decay branching ratios for several processes, have been measured with high precision. No, statistically significant discrepancy from theory has been observed except for the controversial $W^{\pm}$ boson mass measurement from the CDF II Collaboration [33]. Despite the incredible success of the theory, experimental evidence suggests that the theory is incomplete. SM has the following limitations:

- SM fails to explain the gravitational force.
- SM only describes 5% of the total energy density of the universe. It fails to explain dark matter whose existence is experimentally supported by astrophysical observations such as galactic rotation curves and gravitational lensing [34]. It also does not explain dark energy.
- The CP violation allowed in SM cannot explain the amount of matter/anti-matter asymmetry observed in the universe.
- The strengths of the four fundamental forces are different by many orders of magnitude. It has yet to be understood the hierarchy of such interactions.
- The simplest formulation of SM discussed in this thesis assumes the neutrino to be massless left-handed particles. However, recent experimental results suggest that the neutrino masses must be non-zero to generate the observed neutrino oscillation [35]. There are two possibilities to accommodate the neutrino masses in theory by adding either a Dirac or Majorana mass term, which remains an open question to the SM.

These limitations suggest that the SM is an effective field theory, only describing an approximation of our universe, thus, motivating the experimental searches for new physics beyond the Standard Model. Experimentally there are two ways to look for BSM physics, direct searches and indirect precision measurements. Experimental signatures of BSM-predicted particles, such as their invariant mass, are searched directly by direct searches. The thesis focuses on the indirect approach, where precisely measured SM-predicted differential cross-sections are compared with state-of-the-art theoretical predictions looking for hints of deviation caused by the BSM physics.

3 Phenomenology of Proton-Proton Collisions

The main results discussed in this thesis are differential cross-sections for di-Z boson production in association with two jets in a proton-proton collider at the center of mass energy of $\sqrt{s} = 13\ TeV$. Protons are composite particles made up of quarks and gluons. Thus, the theoretical formalism discussed above does not provide all the necessary tools for experimental cross-section measurements in hadron colliders. The differential cross-section $d\sigma$ for two particles is given by,

$$d\sigma = \frac{|\mathcal{M}|^2}{F} dQ \tag{3.1}$$

where F is the incident flux, and dQ represents the Lorentz invariant phase space factor. The scattering amplitude $\mathcal{M}$ is the matrix element calculated from the Lagrangian density of the SM using a perturbative expansion [36].

The cross-section of a process with two initial-state protons p_1 and p_2 producing the final state X is given by:

$$d\sigma_{p_1 p_2 \to X} = \int dx_1 dx_2 \sum_{q_1, q_2} f_{q_1}(x_1, \mu_F) f_{q_2}(x_2, \mu_F) d\sigma_{q_1 q_2 \to X}(x_1, x_2, \mu_F, \mu_R) \tag{3.2}$$

where, q_1, q_2 are the partons of the protons, and $d\sigma_{q_1 q_2 \to X}(x_1, x_2, \mu_F, \mu_R)$ is the partonic cross-section. The functions $f_{q_1}(x_1, \mu_F)$ & $f_{q_2}(x_2, \mu_F)$ are the parton distribution functions (PDF) representing the density of the partons q inside a proton carrying the longitudinal momentum fraction x. The PDFs are determined experimentally using data from deep-inelastic-scattering, jets production, and Drell Yan events [37] [38]. As shown by figure 2, a PDF of a parton depends on the reference value of the momentum transfer Q_0^2. The differences are driven by modifications of partons' momenta resulting from the emission of gluons off of quarks and the splitting of gluons to $q\bar{q}$ pairs. A PDF at any value of Q^2 can be calculated using the PDF at reference scale Q_0^2. The factorization scale μ_F determines

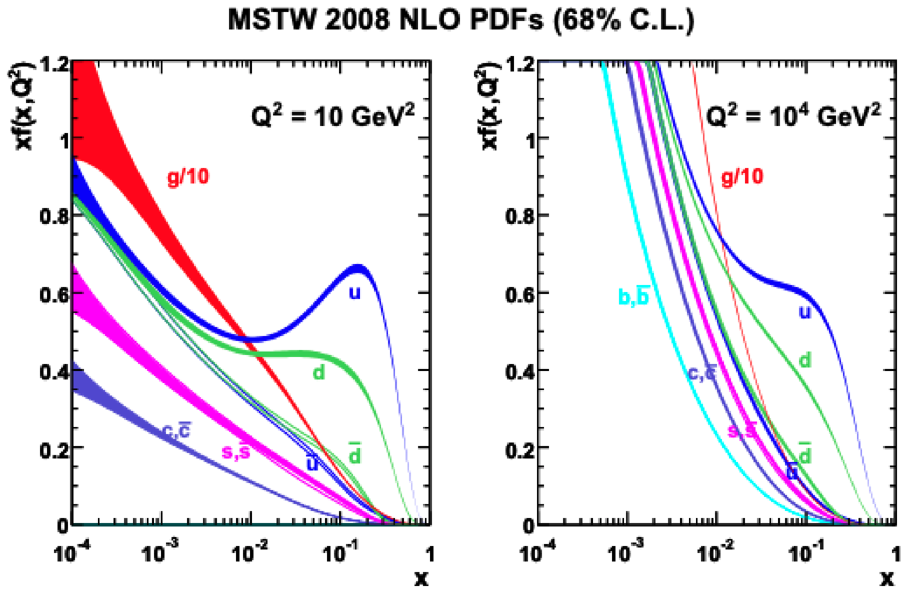


Figure 2: Parton distribution functions $xf_q(x,Q^2)$ for reference momentum transfer $Q_0^2 = 10\ GeV^2$ (left) and $Q_0^2 = 10^4\ GeV^2$ (right). [4].

the threshold whether the perturbative corrections modify the PDF or are included in the partonic cross-sections $d\sigma_{q_1 q_2}$ [36].

The partonic cross-section is calculated perturbatively as an expansion in terms of the strong coupling constant α_S as,

$$d\sigma_{q_1 q_2 \to X} = \alpha_S^k \sum_{m=0}^{n} c_m \alpha_S^m \tag{3.3}$$

The coefficient c_m depends on the center-of-mass energy, and theoretical calculations usually contain a finite number of coefficients. Leading order (LO) calculations include one term ($n = 0$), whereas next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) contains two ($n = 1$) and three ($n = 2$) terms, respectively. The theoretical calculations relevant to the thesis are generally calculated at NLO. The higher-order terms in the series contain additional virtual loop contributions and real emissions of quarks and gluons. The presence of virtual loops beyond the LO introduces singularities in the calculation of scattering amplitudes. The divergences are controlled via the renormalization procedure, where the singularities are absorbed by reparametrization of coupling and mass parameters. The renormalization process is energy-dependent. Therefore, the predicted cross-sections from theoretical calculations depend on the renormalization scale μ_R and the factorization scale μ_F. The scale dependence is varied in Monte Carlo simulations to derive uncertainties on the predicted cross-section due to missing higher-order contributions.

The additional partons of the two protons that interact in the hard interaction process lead to minor energy deposits in the detector referred to as an underlying event. Any outgoing partons from the interaction emit multiple QCD radiation via the parton showering process, where the energy of each parton is split among an increasing number of other elementary particles. Due to the color confinement nature of QCD, at lower energies of the order of the pole of the QCD running coupling (λ_{QCD}), the partons are bound into stable and unstable hadrons. This process is named *hadronization* and leads to the formation of collimated sprays of charged and neutral hadrons in the detector called *jets*. Figure 3 schematically shows the phenomenology of di-Z boson production in association with two jets in the proton-proton collider. The theoretical predictions of such events are calculated using Monte Carlo (MC) simulations which include matrix element calculations for two partons giving two Z bosons, the parton showering, the effect of the underlying events, hadronizations, and pile-up. A comprehensive overview of the methods used in MC simulation is discussed in Ref [39].

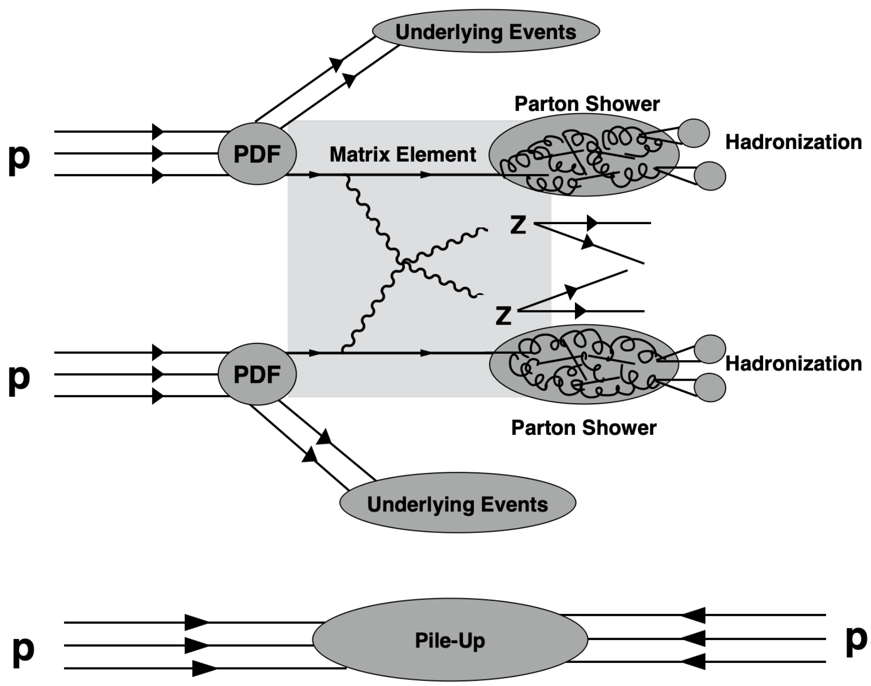


Figure 3: Phenomenology of di-Z boson production in association with two jets in proton-proton collider

4 Electroweak Diboson Physics

In LHC, two types of physics processes, the QCD production at the order $\alpha_S^{\geq 2}\alpha_{EWK}^4$ and the EWK production at order $\alpha_{EWK}^{\geq 6}$ contribute to the production of di-Z bosons in an association of two jets [$ZZ(\rightarrow 4\ell)jj$] [40]. Figures 4 and 5 show the Feynman diagrams at leading order for parton-initiated and gluon loop-initiated QCD $ZZ(\rightarrow 4\ell)jj$ process, respectively, whereas figure 6 shows the Feynman diagrams at leading order for the EWK production of $ZZ(\rightarrow 4\ell)jj$ [41]. The EWK production consists of two sets of interactions, first, the Vector Boson Scattering processes involving either triple (Figure 6a) or quartic (Figure 6b) self-interactions of the gauge-bosons, and second the diagrams featuring the Higgs bosons (Figure 6c & 6d). The scattering amplitudes of the VBS processes involving longitudinally polarized vector bosons grow quadratically with the center of mass energy ($\sqrt{s}$), eventually violating the unitarity bounds. The precise SM interference between the Higgs-featured and VBS processes restores the unitarity [42].

As discussed in Section 1.3.5, the massive W and Z bosons get their masses via the BEH mechanism through EWSB. As a consequence of EWSB, the W and Z bosons acquire an additional degree of freedom (the longitudinal polarization mode) whose scattering interfere with the Higgs-featured processes. Thus, measuring the cross-sections of electroweak production of the di-Z bosons in association with two jets provides a direct probe of the EWSB, which is at the heart of the SM [40]. As the unitarity is restored at high energies, the cross-section of the electroweak $ZZ(\rightarrow 4\ell)jj$ process is sensitive to possible BSM deviations at high energies. Therefore, measuring the cross-sections of the electroweak $ZZ(\rightarrow 4\ell)jj$ processes differentially as a function of kinematically sensitive observables is essential.

The triple and quartic self-interactions of the gauge bosons arise from the square of the non-Abelian structure of $SU(2)$ in the kinetic term $\frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu}$ of the EWK Lagrangian in equation 1.23. Implementing the values of the field strength tensor $W_{\mu\nu}^a$ from equation 1.25, the relations of $W_\mu^\pm$ fields in equation 1.26, and the relations of neutral gauge fields in

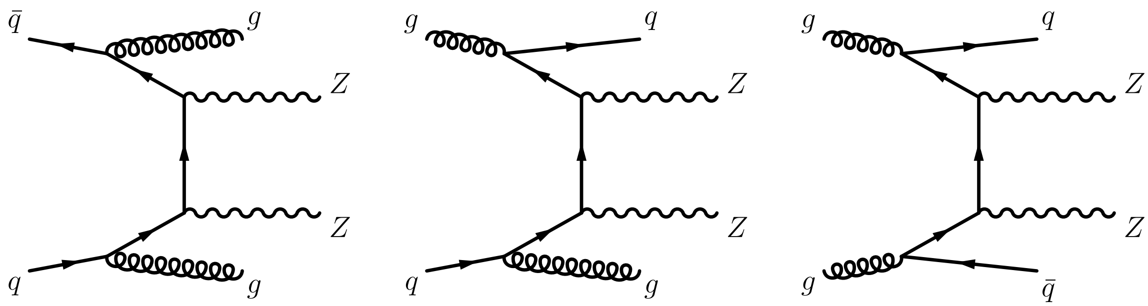


Figure 4: Typical diagrams of LO qq and gg induced QCD $\alpha_S^2\alpha_{EWK}^2$ production of $ZZjj$. The two $Z \to \ell\ell$ vertices each contribute an additional electroweak coupling of α_{EWK}.

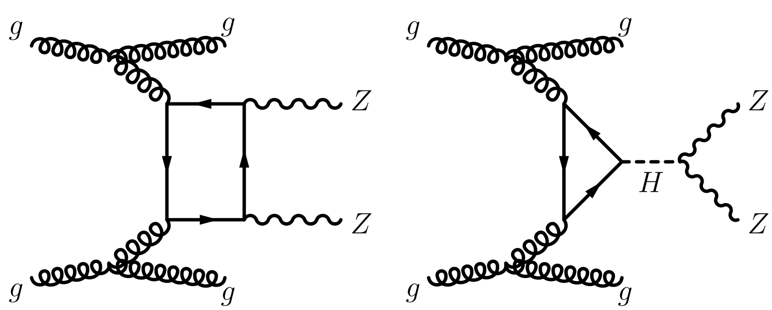
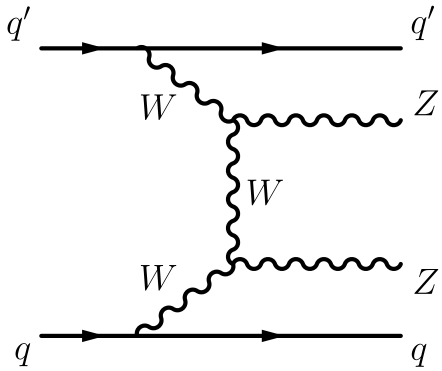
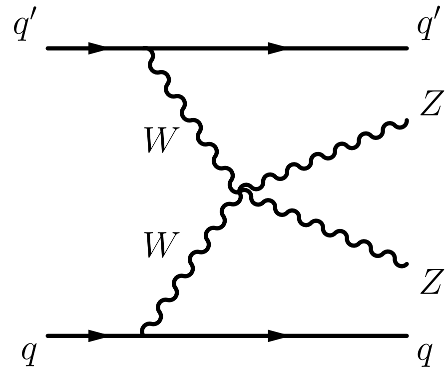


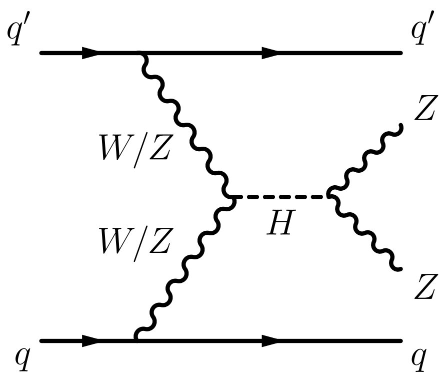
Figure 5: Typical diagrams for LO gg loop induced the QCD $\alpha_S^4\alpha_{EWK}^2$ production of $ZZjj$. The two $Z \to \ell\ell$ vertices each contribute an additional electroweak coupling of α_{EWK}.



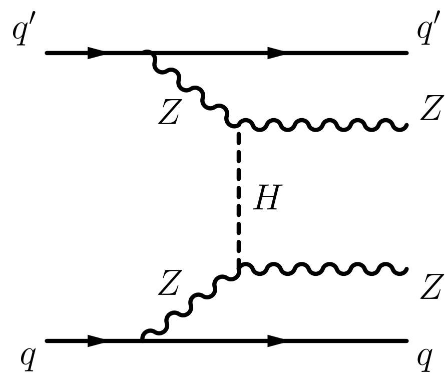
(a) ZZjj production with two triple gauge coupling (TGC) vertices.



(b) ZZjj production with a quartic gauge coupling (QGC) vertex.



(c) s-channel Higgs ZZjj Production.



(d) t-channel Higgs ZZjj Production.

Figure 6: Feynman diagrams at LO for the EWK α_{EWK}^4 production of $ZZjj$. The two $Z \to \ell\ell$ vertices each contribute an additional electroweak coupling of α_{EWK}.

equation 1.29, the triple and quartic self interaction terms become,

$$\mathcal{L}_3 = ie_{V=\gamma,Z}[W^+_{\mu\nu}W^{-\mu}V^{\nu} - W^-_{\mu\nu}W^{+\mu}V^{\nu} + W^+_{\mu}W^-_{\nu}V^{\mu\nu}] \tag{4.1}$$

$$\begin{aligned}\mathcal{L}_4 = {} & e_W^2[W^-_{\mu}W^{+\mu}W^-_{\nu}W^{+\nu} - W^-_{\mu}W^{-\mu}W^+_{\nu}W^{+\nu}] \\ & +e^2_{V=\gamma,Z}[W^-_{\mu}W^{+\mu}V_{\nu}V^{\nu} - W^-_{\mu}V^{\mu}W^+_{\nu}Z^{\nu}] \\ & +e_{\gamma}e_Z[2W^-_{\mu}W^{+\mu}Z_{\nu}A^{\nu} - W^-_{\mu}Z^{\mu}W^+_{\nu}A^{\nu} - W^-_{\mu}A^{\mu}W^+_{\nu}Z^{\nu}]\end{aligned} \tag{4.2}$$

where, $e_\gamma = g\sin\theta_W$; $e_W = \frac{e_\gamma}{2\sqrt{2}\sin\theta_W}$ & $e_Z = e_\gamma\cot\theta_W$ are the precise coupling strengths for vector boson self-interaction. Both triple and quartic neutral couplings, such as ZZZ or $ZZZZ$ are absent in the SM.

Similarly, the couplings of Higgs to vector bosons are also predicted precisely by the BEH mechanism in equation 1.36 as:

$$\mathcal{L}_{HVV} = \frac{m_W^2}{v^2} W_\mu^+ W^{-\mu} h^2 + \frac{m_Z^2}{v^2} Z_\mu Z^\mu h^2 \tag{4.3}$$

The EWK production of $ZZ(\rightarrow 4\ell)jj$ is extremely sensitive to any possible anomalous triple gauge couplings (aTGC), anomalous quartic gauge couplings (aQGC), or anomalous Higgs to vector boson coupling [43] [44] [45]. Therefore, it is imperative to probe the high energy behavior of the EWK production of $ZZ(\rightarrow 4\ell)jj$ to seek possible deviations from the physics processes beyond the Standard Model (BSM).

The EWK $ZZ(\rightarrow 4\ell)jj$ production with each Z boson decaying to a pair of SF-OC lepton pair is an extremely rare process. Moreover, with limited statistics in Run-2, the QCD background processes dominate the $ZZ(\rightarrow 4\ell)jj$ final state [17]. Therefore, the differential cross-sections discussed in this thesis are measured in a VBS-enhanced phase space with a high fraction of events resulting from the EWK $ZZ(\rightarrow 4\ell)jj$ process. The enhanced phase space relies on the characteristic feature of the EWK process with two jets (jj) in the forward-backward region originating from the scattered initial-state quarks. These jets have significant rapidity separation, large invariant mass, and no additional hadronic activity from the hard scattering between the two jets [46]. The decay of the two Z bosons into SF-OC muons or electron pairs defines the final signature of the VBS-$ZZ(\rightarrow 4\ell)jj$-like event.

Chapter III: Experimental Setup

The European Organization for Nuclear Research, CERN, in Geneva, Switzerland, is home to the world's largest particle accelerator, the Large Hadron Collider. The measurements presented in this thesis correspond to the processes at the frontier of high-energy collisions. The relevant energy scale is only possible through large particle accelerators, giant detectors, and international collaboration. There are currently eight experiments analyzing the data from the LHC, among which ATLAS and CMS are the two largest multipurpose experiments. They analyze the collected data to perform SM precision measurements and direct searches for new physics. This thesis analyzes the data collected by the ATLAS experiment between 2015-2018.

This chapter gives a description of the LHC in Section 5, the ATLAS experiment in Section 6, details on physics object reconstruction in Section 7, and plans for future upgrades in Section 8.

5 The Large Hadron Collider

The LHC was approved in 1994 by the CERN Council and was designed to collide hadrons, either protons or heavy ions, at instantaneous luminosity up to $10^{34} cm^{-2} s^{-1}$ and center-of-mass energies up to $\sqrt{s} = 14$ TeV [47]. LHC, the most powerful particle accelerator, is the successor of the Large Electron Positron Collider (LEP), which collided electrons and positrons at center-of-mass energies up to $\sqrt{s} = 250$ GeV [48]. LHC reuses the same tunnel system as LEP with a circumference of 27 km and lies 50 to 175 meters underground at the French-Swiss border outside Geneva, Switzerland.

The first run of the LHC, Run-1, started in 2010 at a center-of-mass energy of $\sqrt{s} = 7$ TeV, which was later increased to 8 TeV. Run-1 lasted until 2013, after which LHC was shut down for two years of planned upgrades to enhance the accelerator and the detectors. LHC resumed its operation for Run-2 from 2015 to 2018 at the center-of-mass energy of $\sqrt{s} = 13$ TeV. The data collected during the Run-2 period is analyzed for this thesis.

Accelerating the proton to the desired high center-of-mass energy is a multi-step process shown schematically in figure 7. First, the protons are created from hydrogen gas by removing the electrons through ionization with an intense electric field. A proton beam is then formed in LINAC 2 linear accelerator and injected into the circular PS Booster, which accelerates the 50 MeV proton to energies of 1.4 GeV [49]. The beams are then injected into the Proton Synchroton (PS), accelerating them to energies of 25 GeV [49]. The proton beams are then injected into the Super Proton Synchroton (SPS) to further accelerate at 450 GeV energies and injected into the main LHC rings. The two opposite proton beams reach the desired final energy of 6.5 TeV using radio frequency (RF) acceleration cavities [49]. The accelerated proton beams are maintained for several hours of data-taking.

The final RF accelerated proton beams are in evenly-spaced discrete bunches, each consisting of 10^{11} protons. The bunches are separated by 25 ns spacing [49]. The proton beams are guided in the tunnel by a magnetic field using superconducting dipole and quadruplet

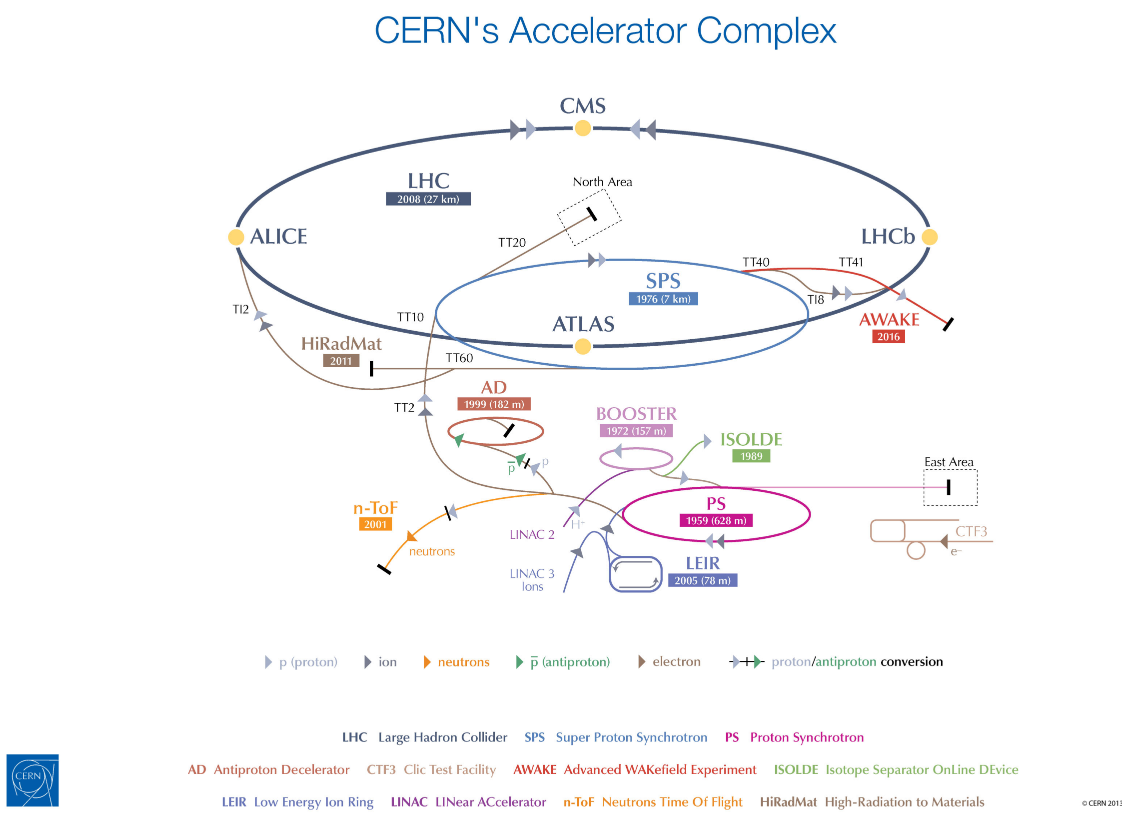


Figure 7: A detailed layout of multiple-steps that goes into proton acceleration before entering the main LHC ring [5].

magnets. The main LHC ring comprises 1232 dipole magnets that provide a strong magnetic field of 8 T to bend the beams and about 392 quadruplet magnets to focus the beams in the transverse plane [49]. The superconducting magnets are cooled down to 1.9 K, which requires two vacuum systems to hold the cryomagnets and the helium distribution lines. To avoid unnecessary interactions, the beams are accelerated and maintained in an ultra-high vacuum of 10^{-13} atm [49].

The two opposite beam lines meet in four interaction points, thus, creating proton-proton collisions. The collisions are recorded by the LHC's four main detectors: ATLAS, CMS, LHCb, and ALICE. ATLAS and CMS are the two multipurpose experiments designed to perform new physics searches and precision SM measurements. ALICE experiment is designed to investigate the heavy-ion collisions and the quark-gluon plasma, whereas the LHCb experiment is designed to study flavor physics [49].

6 ATLAS Detector

A Toroidal LHC ApparatuS (ATLAS) is a general-purpose detector of LHC that detects events from proton-proton and heavy ion collisions [6]. It is a 44 meters long, and 25 meters wide cylindrical-shaped detector built around LHC Interaction Point 1 [6]. ATLAS has multiple concentric sub-detectors layered around the beamline, providing forward-backward symmetric coverage. The two proton beams collide at the center of the detector producing outgoing particles from hard scattering, underlying events, and pile-up. The outgoing particles interact with the detector material leaving tracks and energy deposits in several layers of the sub-detectors. The sub-detector closest to the beamline is called *Inner Detector (ID)*, which measures the trajectories of the charged particle and plays a crucial role in identifying the physical position of hard-scattering, also known as the *interaction point (IP)*. ID is surrounded by a solenoid magnet that provides a 2 T magnetic field to bend the particle trajectories for momentum measurements [6]. After the solenoid magnet lies the *electromagnetic calorimeter (ECAL)* and then the *hadronic calorimeter (HCAL)*, which measure the energy of electromagnetic and hadronic physic objects, respectively. The outermost layer of the ATLAS detector is the *Muon Spectrometer(MS)* that provides a secondary measure of muon trajectories for momentum measurement. MS is embedded inside a toroidal magnetic field that provides a magnetic field up to 3.5 T [6]. Figure 8 shows a schematic of the ATLAS detector with all its sub-detectors.

6.1 ATLAS Coordinate System

ATLAS measurements use a right-handed coordinate system with the nominal interaction point as the origin. The beamline is along the cylindrical symmetry axis of the detector, which defines the longitudinal z-axis. The transverse xy-plane is perpendicular to the beam direction, where x-axis points to the center of the LHC ring and y-axis points upwards towards the surface. Figure 9 shows a schematic of the ATLAS coordinate system. The

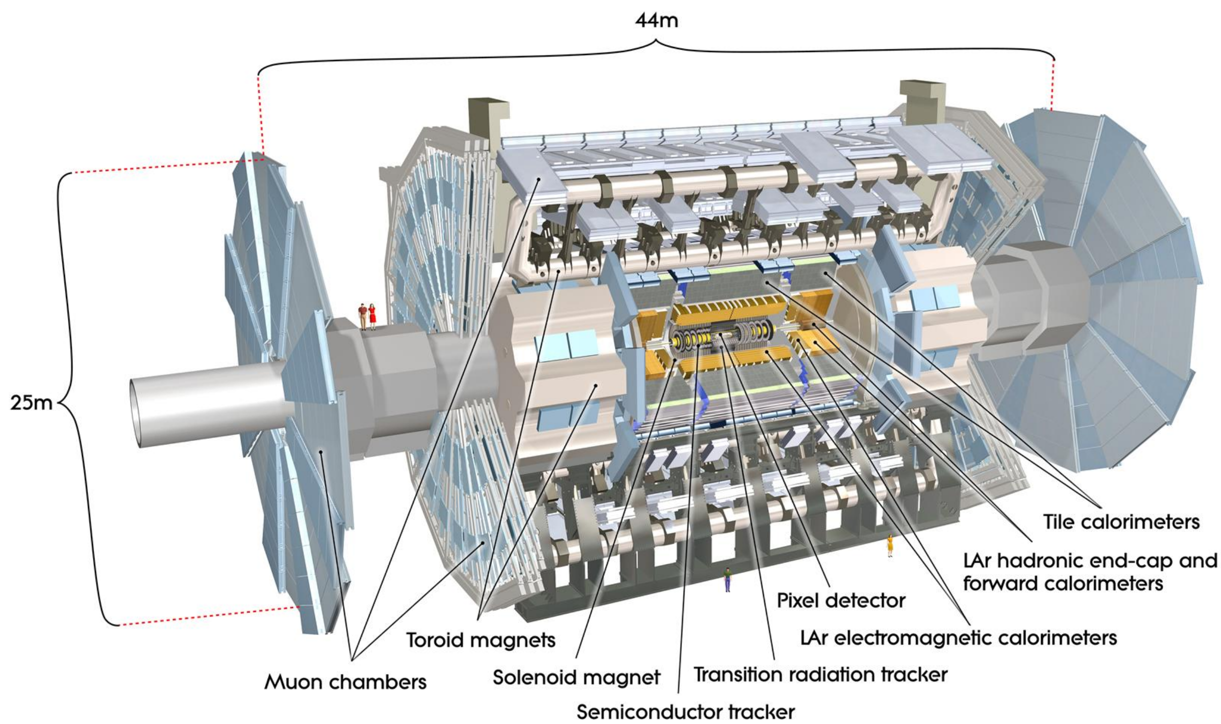


Figure 8: A detailed schematic of the ATLAS detector with all its sub-detectors [6].

angle measured around the beamline in xy-plane gives the azimuthal angle ϕ, whereas the angle measured with respect to the z-axis gives the polar angle θ. Transverse momentum (p_T) is particle's momentum in the xy-plane, defined as,

$$p_T = \sqrt{p_x^2 + p_y^2} = p \sin\theta \tag{6.1}$$

Rapidity (y) defined in terms of a particle's energy (E) and momentum (p) is a commonly used collider physics quantity that measures whether an outgoing particle is produced perpendicular or parallel to the z-axis. Rapidity is defined as,

$$y = \frac{1}{2} \ln\left(\frac{E + p_z}{E - p_z}\right) \tag{6.2}$$

Particles with larger momentum along the z-axis have larger values of rapidity, whereas particles with larger momentum values in the transverse plane have smaller values of rapidity.

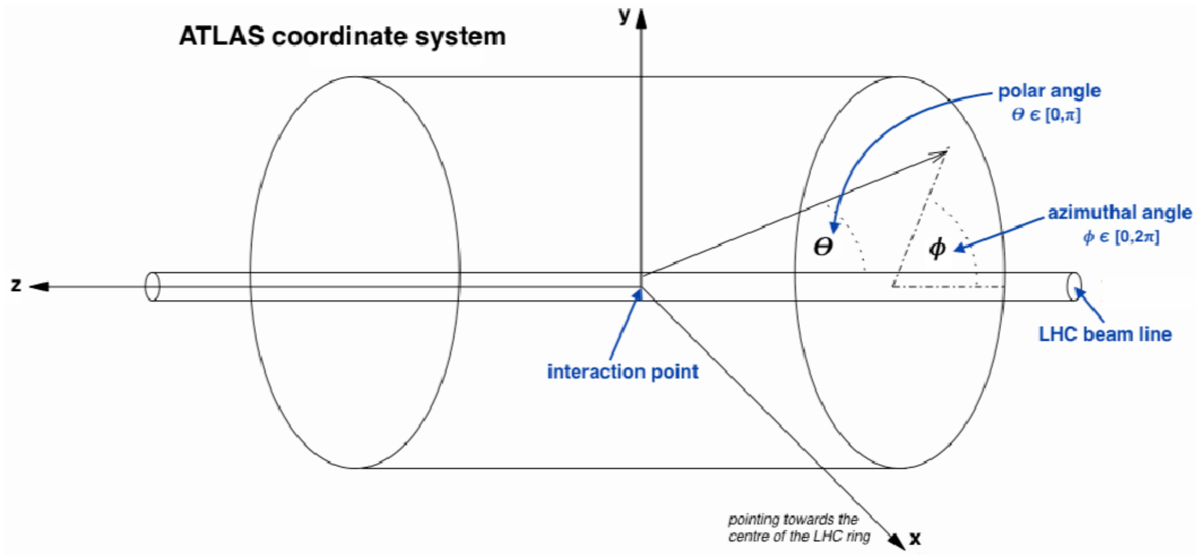


Figure 9: A schematic of the right-handed ATLAS coordinate system [7].

For particles with negligible mass, the rapidity approaches a purely angular variable called *pseudorapidity* (η) defined as,

$$\eta = \frac{1}{2}\ln\left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z}\right) = -ln\left[\tan\left(\frac{\theta}{2}\right)\right] \tag{6.3}$$

Higher values of rapidity and pseudorapidity refer to the forward region of the detector. ATLAS detector has full 2π coverage in ϕ and maximum coverage up to $|\eta| < 4.5$ corresponding to $1.3^\circ < \theta < 178.7^\circ$ [6].

6.2 Inner Detector

The inner detector is the innermost sub-detector of ATLAS and is responsible for tracking charged particles' trajectories and identifying the interaction point of the hard scatter. Closest to the interaction point is the Insertable B-Layer (IBL) [50], which was installed during the long-upgrade shutdown between Run-1 and Run-2 to meet the requirements for competent tracking at higher pile-up. The IBL is highly granular, consisting of roughly 12 million silicon pixel sensors with a size of $50 \times 250\ \mu m^2$ [50]. IBL is located 3.3 cm from the beamline and can reconstruct tracks within the pseudorapidity range of $|\eta| < 2.5$ [50].

Three layers of silicon-pixel detectors with $1,744$ pixel sensors, each comprising $47,232$ pixels of size 50×400 μm^2 surround the IBL [51]. The slightly larger pixel size is adequate for the pile-up at a distance larger than 5 cm from the interaction point. These pixel layers were also present during the Run-1 data-taking period and provided coverage up to $|\eta| < 2.5$ with a spatial resolution of tracks between 5 and 12 μm [51]. Surrounding the pixel layers is the Semiconductor Tracker (SCT) consisting of five layers of silicon microstrip detectors with a mean strip pitch of 80 μm in the barrel region and varying pitch of $57 - 94$ μm in the end-cap regions [52].

At a distance about 50 cm from the beamline lies the outermost layer of the ATLAS inner detector, the Transition Radiation Tracker (TRT), with $370,000$ straw tubes with a diameter of 4 mm [53]. Each TRT straw tube is filled with a Xenon-based gas mixture and consists of 31 μm diameter tungsten wires [53]. A charged particle passing through different layers of ID leaves a track via ionization.

Figure 10 schematically shows different parts of the inner detector and their distances from the interaction point.

6.3 Calorimeters

ATLAS has two calorimeters, electromagnetic and hadronic, designed to measure the energy of charged and neutral particles up to the range of $|\eta| \leq 4.9$ [6]. When interacting with a material, an electron loses its energy by photon emission, which could produce a pair of e^+e^- and vice versa, creating an electromagnetic shower in the detector. Similarly, the hadronic particles also result in a shower of particles through multiple scattering. The calorimeters measure the energy of the particles by reconstructing the electromagnetic and hadronic showers. The calorimeters are designed to capture all particles except muons and neutrinos. Therefore, motivated by the need to prevent *punch-through*[1] effect, materials

[1]particles' probabilities of passing through the calorimeters

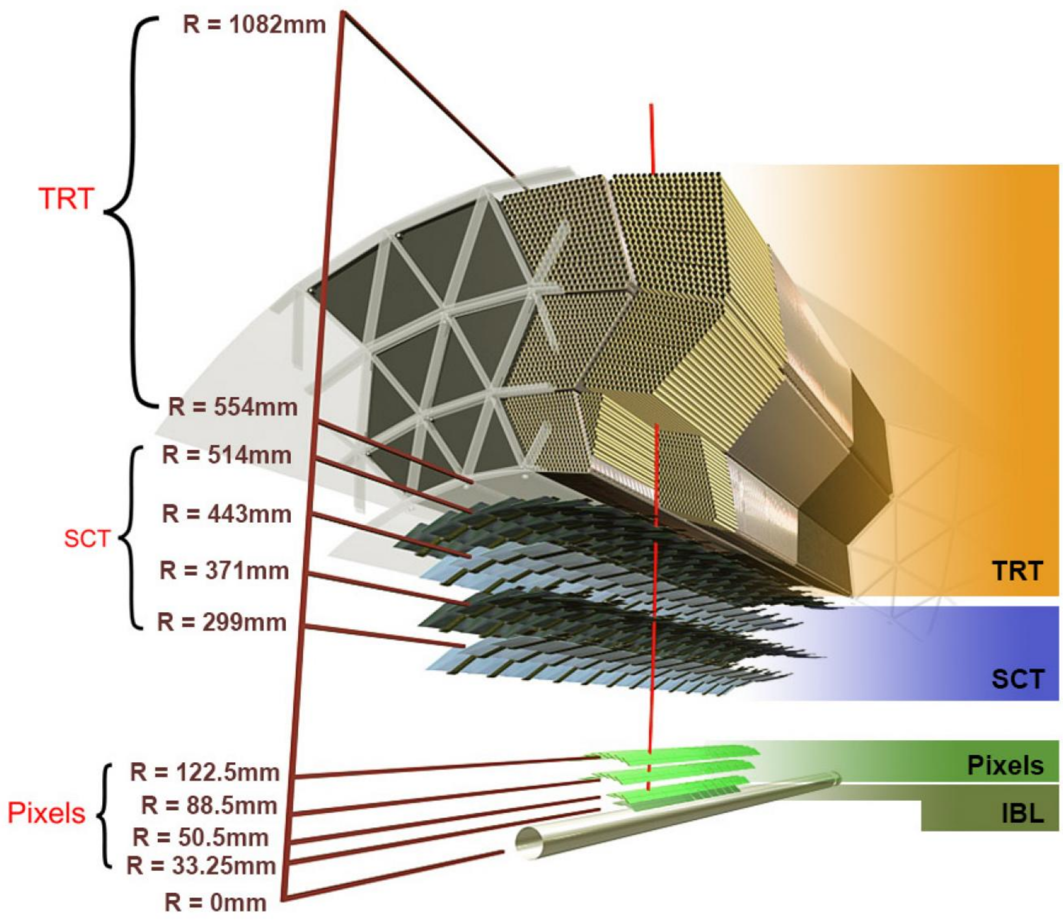


Figure 10: A schematic of the inner detector of ATLAS showing the IBL, pixel detectors, SCT, and TRT [8].

with high radiation length (X_0) and high interaction length (λ) are chosen to construct the calorimeters.

Outside the solenoid magnet surrounding the ID is the accordion-shaped electromagnetic calorimeter consisting of an alternate layer of lead absorber plates and highly granular liquid-argon (LAr) cells to precisely measure the energies of electrons and photons. It comprise of barrel section in $|\eta| < 1.475$ range and two end-caps in $1.375 < |\eta| < 3.2$ range [54]. The calorimeter's central region ($|\eta| < 2.5$) is designed to identify electrons and photons with high precision.

The hadronic calorimeter surrounds the ECAL and consists of a steel absorber and active scintillator tiles in the $|\eta| < 1.7$ range. In the end-caps range of $1.5 < |\eta| < 3.2$, it consists of a copper absorber and active LAr detectors. The forward region ranging from $3.2 < |\eta| < 4$ comprises the tungsten absorber followed by active LAr detectors [55].

Figure 11 schematically shows the layout of ATLAS calorimeters.

6.4 Muon Spectrometer

In ATLAS, muons are deeply-penetrating charged particles that leave minimum ionizing deposits in the calorimeter. The muon spectrometer, the outermost part of the ATLAS detector, tracks trajectories of muons deflected in 0.5 magnetic field provided by the superconducting toroidal magnets, giving an additional measure of muon's momentum [6]. The MS tracks muon with $p_T > 3$ GeV in $|\eta| < 2.7$ range [6]. As shown in figure 12, the muon spectrometer comprises four types of detectors; first, the three stations of Monitored Drift Tubes (MDT) in $|\eta| < 2.0$ region followed by the Cathode Strip Chambers (CSC) in $2.0 < |\eta| < 2.7$ region [6]. The other two detectors are the Resistive Plate Chambers (RPC) in $|\eta| < 1.05$ and the Thin-gap Chambers (TGC) beyond $|\eta| = 1.05$ comprising the trigger system in MS [6].

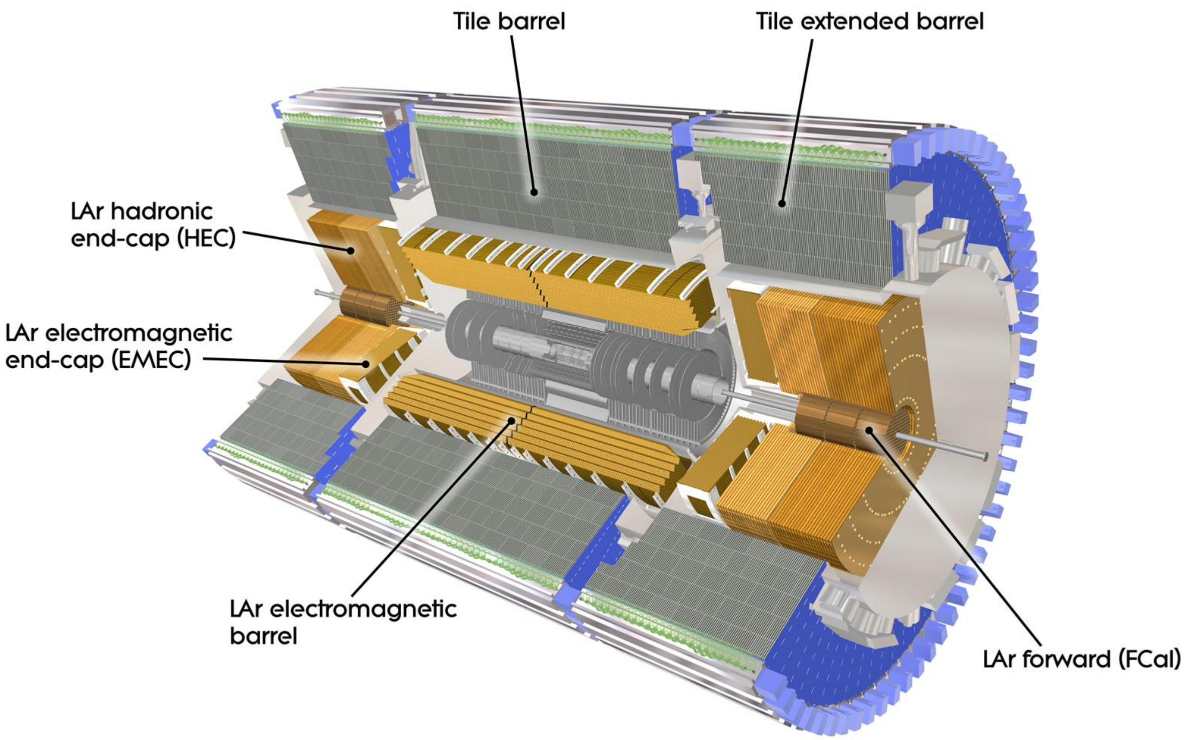


Figure 11: A schematic of electromagnetic and hadronic calorimeters In ATLAS [6].

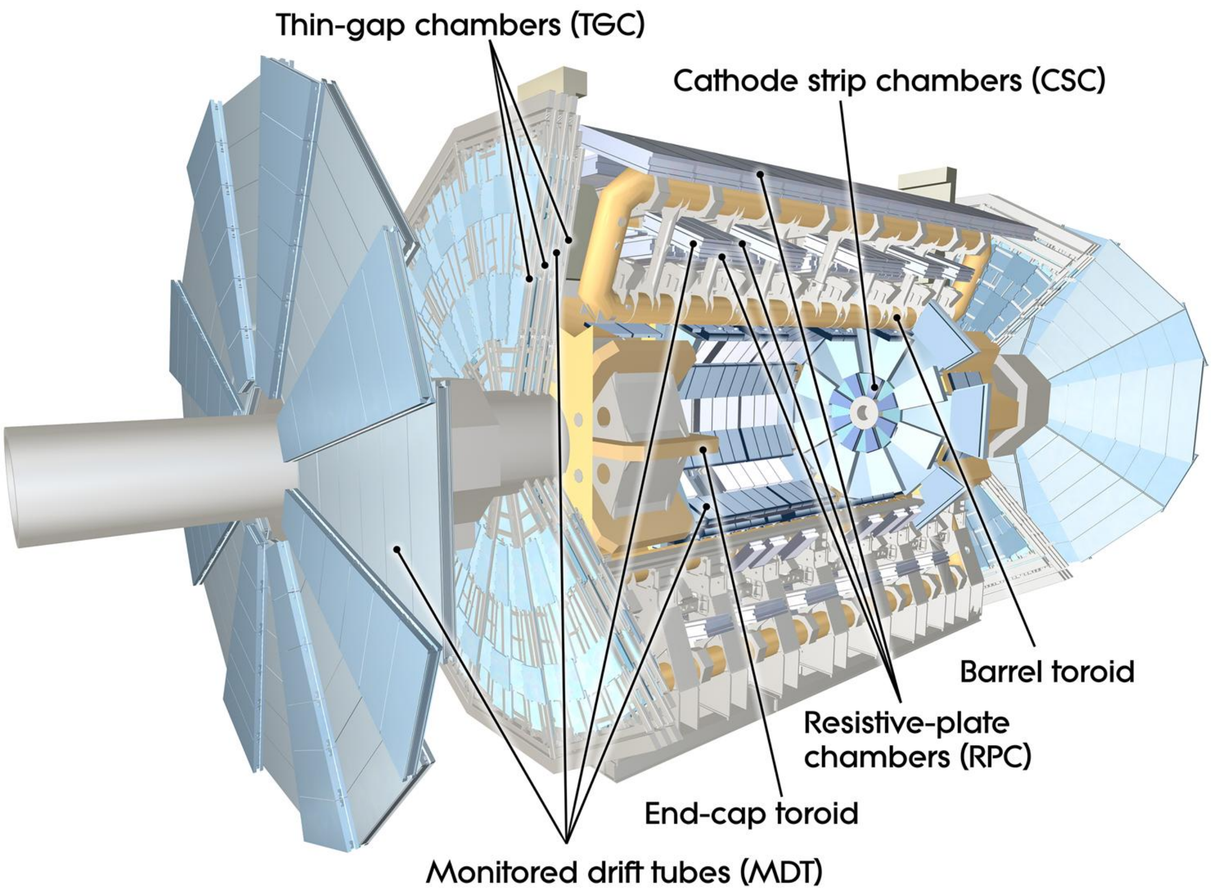


Figure 12: A schematic of different components of the muon spectrometer in ATLAS [6].

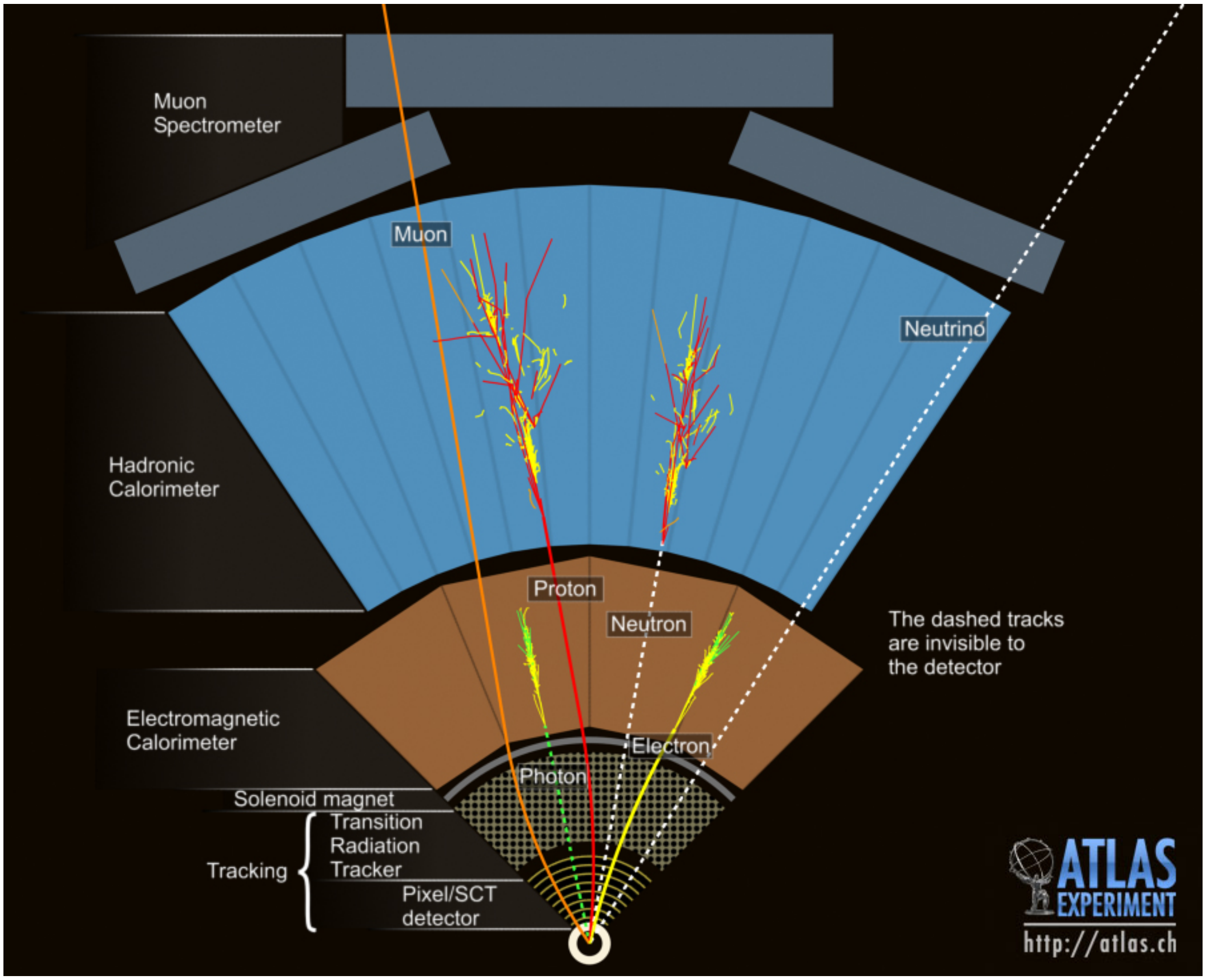


Figure 13: Simplified representation of various particles traversing through different layers of ATLAS sub-detectors and leaving unique signatures [9].

7 Physics Object Reconstruction

Different particles leave unique signatures in different sub-detectors of ATLAS. Figure 13 shows a schematic of simplified representation of various particles passing through different sub-detectors and leaving various signatures. Physics object reconstruction is the process of interpreting these signals to meaningful information about the outgoing particles. This section discusses the detail of reconstruction relevant to the thesis.

7.1 Trigger

The first step of particle and event reconstruction is selecting interesting high-energy events from a pool of lower-energy multiple-scattering signals. The high bunch crossing frequency of every 25 ns results in a large amount of data making it physically impossible to store all events. ATLAS trigger system filters the events interesting for physics measurements for permanent storage.

ATLAS trigger consists of two levels, Level 1 (L1) trigger integrated into the hardware and high-level software trigger (HLT) [56]. The L1 trigger is based on custom-built electronics, which uses signals from the calorimeters and muon trigger system (TGC and RPC) to identify event features such as electrons, photons, jets, taus, and missing energy. The L1 trigger reduces the 40 MHz incoming collision data-rate corresponding to 25 ns bunch crossing by a factor of 400 to 100 kHz output [56]. The events accepted by the L1 trigger defines regions of interest (ROI), and HLT algorithms are run on these events to select ones with candidate physics objects passing the kinematic requirements. The software-based HLT trigger further reduces the data rate by almost a factor of 100 to 1.5 kHz [6]. With the combination of L1 and HLT trigger system, the data rate is reduced by $400,000$, and the selected events corresponding to data readout of 1.5 GB per second are stored in the permanent storage. Figure 14 shows the schematic of ATLAS's trigger and data acquisition system.

Physics object reconstruction discussed below converts the raw data output stored in permanent storage to physics objects used in physics analyses.

7.2 Tracks and Vertices Reconstruction

Tracking a charged particle is a critical step in reconstruction. The tracks of the charged particles play an essential role in momentum measurement, particle identification, and primary vertex reconstruction through the extrapolation of tracks to the interaction point. As the inner detector is closest to the beamline and comprises minimally ionizing detector material

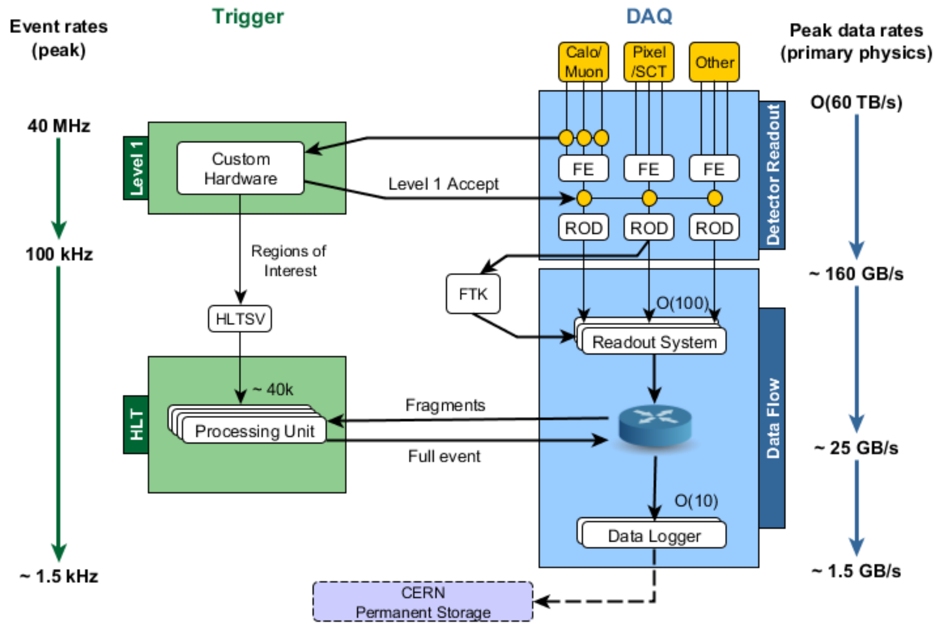


Figure 14: Trigger and data acquisition system in ATLAS [10].

with high granularity, it plays a crucial role in track reconstruction. The ID magnetic field is homogenous, resulting in circular tracks of the charged particles. Five parameters shown in Figure 15 define charged particle tracks; the ratio of charge and transverse momentum (q/p_T) defining the curvature; the distance of the closest approach to the primary vertex in xy-plane defining the transverse impact parameter (d_0), the longitudinal impact parameter (z_0) along the z-axis; the azimuthal angle (ϕ_0) and the polar angle (θ_0) of the particle direction at the closest approach point [57].

As shown by figure 16, used in Run-2 track reconstruction consists two different approaches; the primary *inside-out* approach and the secondary *outside-in* approach [57]. The first step in the inside-out track reconstruction is the space point and drift circle formation, formed respectively by the clusters of signals from the silicon detectors and drift-circles hits in the TRT. Second, track seeds are formed from a collection of three silicon-detector space points and extrapolated to the outer layers by including the compatible clusters in the track trajectory. Once the track is formed, an ambiguity resolution algorithm is applied to reassign

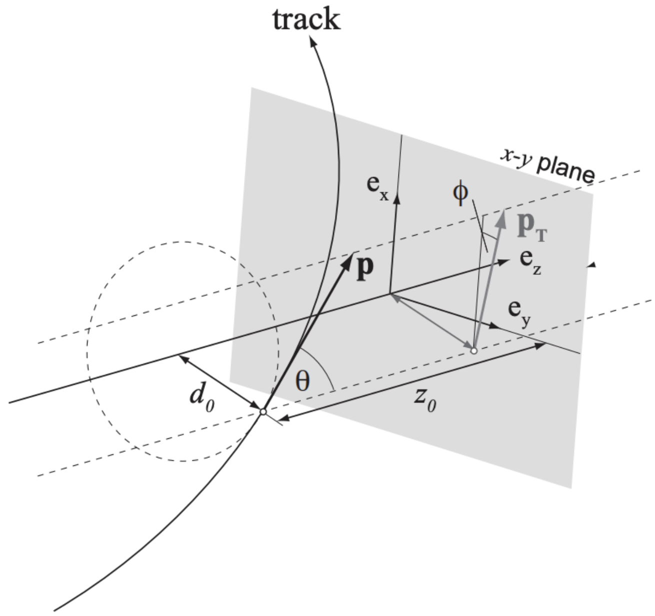


Figure 15: Schematic showing the five-track parameters [11].

shared clusters to the track with a better match, and the final track candidate is fitted using a global χ^2 method. The last step of inside-out track reconstruction is adding compatible TRT drift holes and refitting the tracks.

The inside-out method is optimal for particles that minimally interact with the inner-detector material. However, secondary backtracking with an outside-in approach is needed for particles interacting with the inner detector, such as reconverted photons. In this approach, the track pattern recognition starts at the TRT in the regions of interest flagged by the electromagnetic calorimeter and backtracks to the silicon detectors.

Tracks of the charged particle are extrapolated inward to the beamline and are assigned to vertices [58]. In most ATLAS analyses, including the one presented in this thesis, the space-point with the highest quadrature sum of track p_T ($\sum_{track} p_T^2$) is identified as the primary vertex of an event.

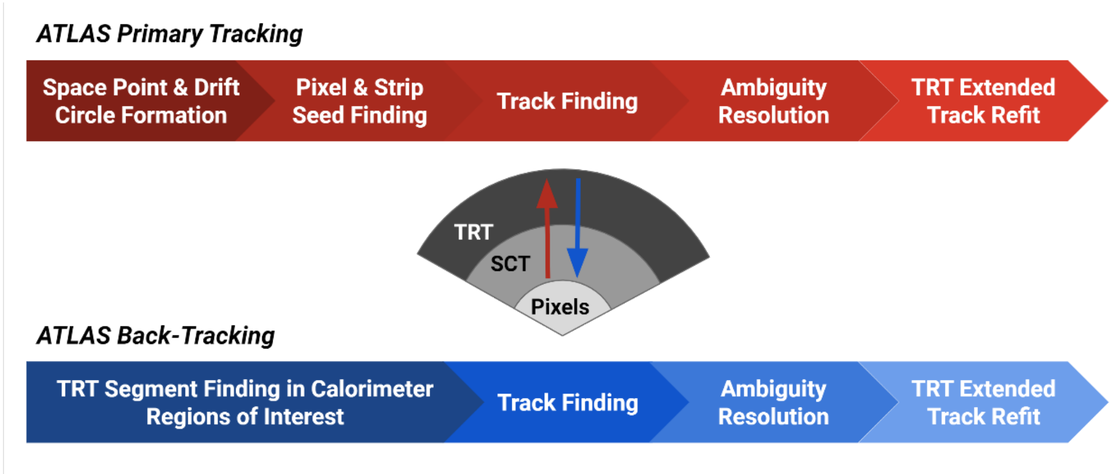


Figure 16: Schematic showing the two techniques of track reconstruction, primary inside-out and secondary outside-in. Figure taken from ATLAS Tracking CP group public tutorial https://atlassoftwaredocs.web.cern.ch/trackingTutorial/idoverview/.

7.3 Electron Reconstruction

Electrons, when interacting with a material, produce a photon by bremsstrahlung radiation, the process defining the interaction of a charged particle with the electric field of an atomic nucleus. Any energetic photon, either from a physics process or bremsstrahlung, can turn into a pair of e^+e^-, which can again radiate another set of photons, thus, giving rise to an electromagnetic shower. For given energy and material, electromagnetic showers have a characteristic penetration depth.

ATLAS electrons are reconstructed by combining the tracking information from the ID, and the energy deposits in nearby cells of the calorimeter, energy clusters [12]. The clusters are formed only if the energy deposit exceeds four times the expected deposits from the pile-up. The reconstruction efficiency for high-energy electrons with transverse energy ($E_T > 15$ GeV) is about $97 - 99\%$ [12]. *Prompt electrons* originate from the hard scattering and are the primary interest of physics analysis. However, the detector has electrons from *non-prompt sources* including the jets, misidentification, and pile-up. Therefore it is imperative to identify and isolate the prompt electrons in an event, and the efficiency for identification and isolation varies as a function of E_T. Limited by the coverage of the electromagnetic calorimeter, only electrons within the $|\eta| < 2.47$ range can be reconstructed and identified as prompt in ATLAS.

The electron identification is based on a multivariate-likelihood (LH) technique which takes information from tracking detectors and calorimeters as input. The tool is trained to separate signal and background probability density functions using simulated $Z \rightarrow e^+e^-$ and $J/\psi \rightarrow e^+e^-$ events. The LH tool provides four *working points*, VeryLoose, Loose, Medium, and Tight, at different values of the LH discriminant to cover various needs of several ATLAS analyses. The analysis presented in this thesis uses electrons satisfying the Loose identification working point with at least one hit in the IBL.

Prompt electrons originating from W, Z, or H decay are characterized by low activity around them in the $\eta - \phi$ plane. An isolation requirement is applied to the electron candidates to select ones from the hard scattering. Calorimeter and track-based requirements on isolation variables are defined to quantify the isolation. The variables are based on the amount of activity around an isolation cone of the candidate electron. Calorimeter-based isolation relies on the variable $E^{iso}_{T,cone}$, the sum of transverse energies inside a $\Delta R = 0.2$ cone of the electron candidate. Similarly, the track-based isolation variable is $p^{iso}_{T,cone}$, the sum of the transverse momentum of the electron candidate within a p_T-dependent ΔR, which is defined as,

$$\Delta R = min\left(\frac{10GeV}{p_T}, \Delta R_{max}\right) \tag{7.1}$$

where the maximum cone size is $\Delta R_{max} = 0.2$. Similar to the identification, several working points are available for electron isolation. The measurement in this thesis uses the *Loose_VarRad* isolation working point which requires $E^{iso}_{T,cone} < 0.3$ and $p^{iso}_{T,cone} < 0.15$. Figure 17 shows the electron identification and isolation efficiencies as a function of their E_T. The Loose working point has the highest identification and isolation efficiencies and is the optimal working point for the measurement because of the fully reconstructable final state.

The total electron efficiency is defined as the product of the electron reconstruction,

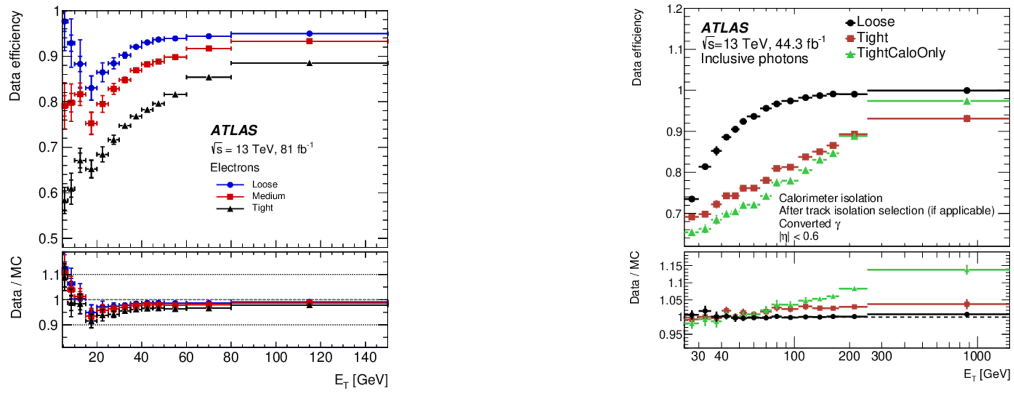


Figure 17: Distributions showing the identification (left) and isolation (right) efficiencies for electrons as a function of their E_T [12].

identification, isolation, and trigger efficiencies as,

$$\epsilon_{total} = \epsilon_{reco} \times \epsilon_{id} \times \epsilon_{iso} \times \epsilon_{trigger} \tag{7.2}$$

Each of the efficiency terms is evaluated on data and MC. *Scale Factors (SF)* defined as the ratio of the measured efficiency in data and the efficiency simulated in MC are derived and applied to the simulation to match the one observed in the data. Typically, SFs are close to one, and systematic uncertainties related to scale factors are considered in the measurement.

7.4 Muon Reconstruction

The rate of bremsstrahlung radiation is inversely proportional to the square of a particle's mass. Since muons are about 200 times heavier than electrons, they primarily interact with the detector material through ionization. Therefore, muons are minimally ionizing particles that do not create electromagnetic shower in the calorimeters and pass through all layers of the ATLAS detector. Hence, muon detection relies on track measurements from the inner detector and muon spectrometer. As shown in figure 18, four types of muons are defined based on the type of sub-detectors used during a muon reconstruction,

- **Combined muons:** muons reconstructed from a global refit of ID and MS tracks

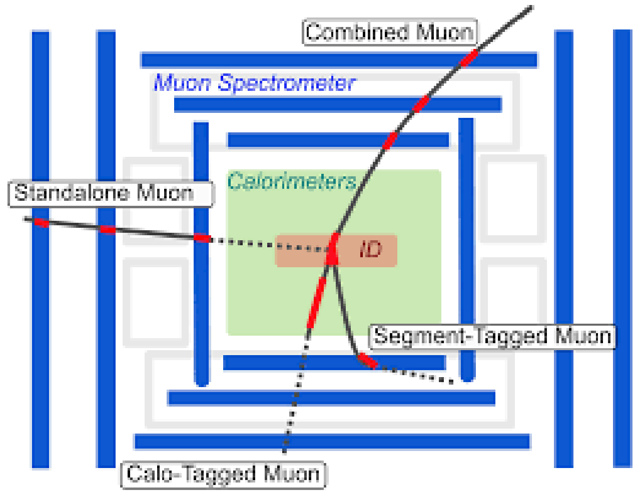


Figure 18: Schematic of four different types of muons reconstructed using several layers of sub-detectors [13].

- **Segment-tagged muons:** muons reconstructed from a fitted ID track and MS segment track
- **Calo-tagged muons:** muons reconstructed using ID track matched to the minimum ionizing energy deposits in the calorimeters
- **Standalone Muons:** muons reconstructed solely from the MS tracks

Similar to the electron reconstruction discussed in Section 7.3, reconstructed muons from the hard scatter are identified and isolated from the muons originating from secondary sources. Muon identification working points are developed by applying quality requirements in the simulated $t\bar{t}$ events where a W from top-quark decay decays to a muon and a neutrino. The quality cuts require at least one-pixel hit, five SCT hits, less than three pixel or SCT holes, and at least 10% of TRT hits to be included in the fit for the $0.1 < \eta < 0.9$ range with full TRT coverage [13]. Three variables are used in muon identification; q/p significance, which is defined as the ratio of muon's charge and momentum divided by the sum-quadrature of their uncertainties; ρ' defined as the absolute difference of transverse momentum measurements in the ID and MS divided by the combined track's p_T; and the normalized χ^2 of the combined track fit [13]. Four identification *working points*, Loose, Medium, Tight, and High-p_T are defined for muons. The measurement uses a Loose identification point, which comprises all four types of muons and is developed for processes with four leptons in the final state [13].

To evaluate the total reconstruction efficiency of muons, $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events are used. The reconstruction efficiency in the region with ID coverage $|\eta| < 2.5$ is obtained by using the tag-and-probe method, whereas, for $|\eta| > 2.5$ region, it is estimated by evaluating SFs based on a double ratio of data and MC in $Z \to \mu\mu$ events [59].

Analogous to electrons, muons are required to meet calo-based and ID-based isolation requirements. Muons in the measurement satisfy *PflowLoose_VarRad* isolation working point. Like in the case of electrons, systematic uncertainties on different scale factors for muon reconstruction, identification, isolation, and trigger efficiencies are propagated to the final cross-section measurement.

7.5 Jet Reconstruction

Due to color confinement in QCD, a quark or gluon cannot exist on its own and goes through hadronization to form a collimated color-neutral stream of particles, *jets*. Generally, jets are deeply penetrating, passing through the electromagnetic calorimeter, and gets fully absorbed by the material in the hadronic calorimeter. ATLAS jet reconstruction relies on the formation of *topo-clusters*, energy deposits in the calorimeter cells using a sequential recombination algorithm. Most commonly, the jets are reconstructed using a type of sequential algorithm, the anti-k_T algorithm. At first, pseudojets are formed by iteratively combining nearby particles based on a specific metric distance to other particles from the pseudojet. The iteration is continued over a set of nearby particles until the pseudojets' metric distance is smaller than that between the pseudojet and the nearby particle. For the anti-k_T algorithm, the separation distance metric is based on their transverse momenta [60]. A radius parameter for jets determines the angular size of resulting jets in the $y - \phi$ plane.

The analysis reported here uses jets reconstructed using the anti-k_T algorithm with a

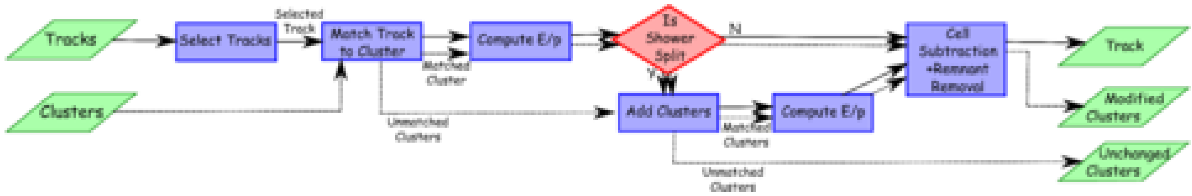


Figure 19: Schematic of particle flow jet reconstruction [14].

particle flow implementation [14] [60]. The particle flow reconstruction technique was first implemented by ATLAS reconstruction in Run-2. It combines information from the inner detector tracks and the calorimeter's topo-clusters. It has advantages over simple cluster-based algorithms because the tracking detector has a better momentum resolution for lower-energy charged particles, extending the reconstruction to include particles with topo-cluster energy deposits lower than the noise thresholds [14]. Moreover, the particle flow algorithm removes cluster contributions from pile-up using track-to-vertex association. Figure 19 shows a schematic of the particle flow jet reconstruction algorithm used in ATLAS.

First, ID tracks and topo clusters are reconstructed. If an ID track matches the topo-cluster, the tracks are used for jet reconstruction, and the cluster is modified by subtracting the energy of the used tracks. In the case of neutral jets or jets outside the geometrical acceptance of ID, only the topo-clusters are used as input for jet reconstruction. Figure 20 shows the improvement in jet transverse momenta resolution when using the particle flow algorithm over the cluster-based algorithm as a function of jet's p_T (left) and η (right). The distributions show the resolution is improved significantly for low$-p_T$ and central jets with low η with particle flow algorithm.

Jet energies are highly susceptible to pile-up density, and jets constitute various particles that could interact differently with various sub-detectors. Therefore, the reconstructed jets must be calibrated before being used in the measurement. Figure 21 schematically shows different steps of jet calibration, which is discussed in detail in Ref [15]. The first step in jet calibration is an event-by-event subtraction of pile-up-related average energy density (ρ) from jet candidate energy. The pile-up is first corrected as a function of p_T, then as a function of

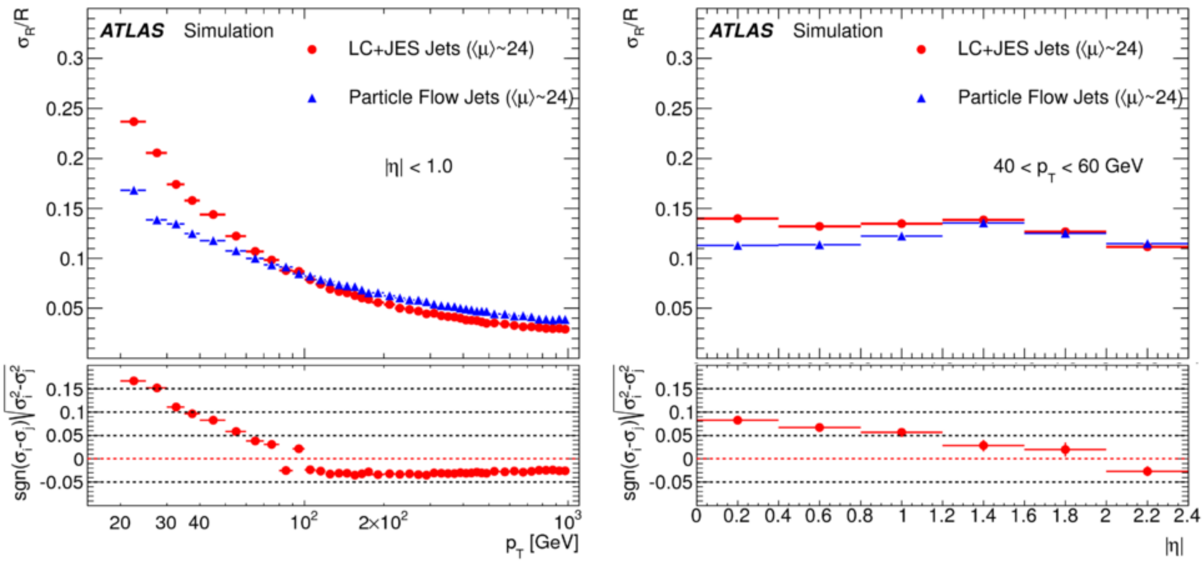


Figure 20: Resolution of jet transverse momentum for only cluster-based jets (LC+JES) and particle flow jets as a function of p_T (left) and η (right) [14].

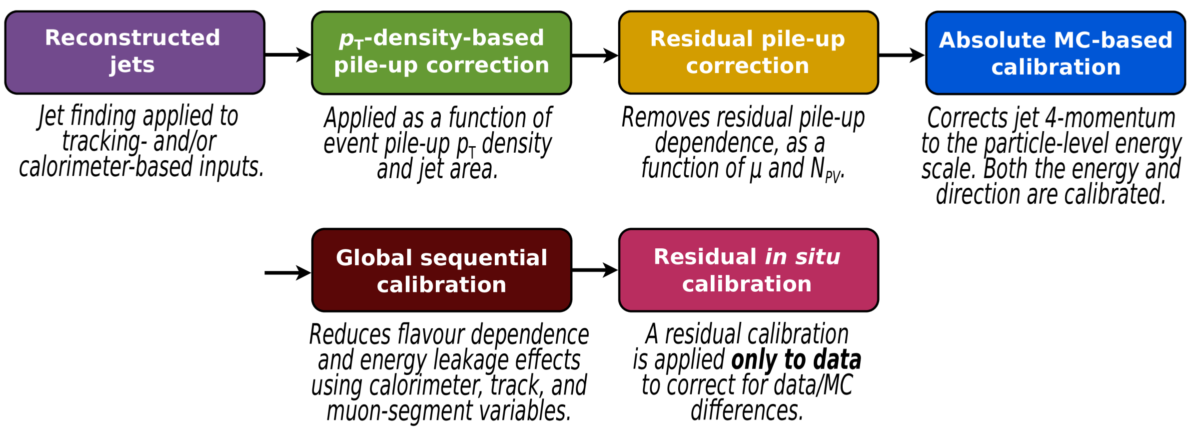


Figure 21: A schematic of different steps of jet calibration [15].

average pile-up density (μ) and the number of primary vertex per event (N_{PV}). The second set of jet calibration accounts for the detector response, correcting for the interaction with the detector's passive material and possible flavor dependence on the detector response. *In-situ* calibration is the final part of the jet calibration, which accounts for the differences in jet response measured in data and MC.

The analysis uses a *Jet Vertex Tagger* (JVT) variable to differentiate hard-scattering jets from pile-up [61]. The JVT variable is a multivariate combination of two variables; first, the *jet-veretex-fraction* defined as a fraction of the total momentum of tracks in jet associated with the primary vertex and second, the number of reconstructed primary vertices in an event (N_{PV}) [61]. To identify the jets from hard scattering outside the ID acceptance in $2.5 < |\eta| < 4.5$ range, a *forward Jet Vertex Tagger* (fJVT) discriminant is used, which is derived by exploiting the topological correlations among particles from a pile-up interaction [62].

The systematic uncertainties related to each step of jet reconstruction, calibration, and JVT/fJVT tagging are propagated to the final measurements of the unfolded differential cross-sections.

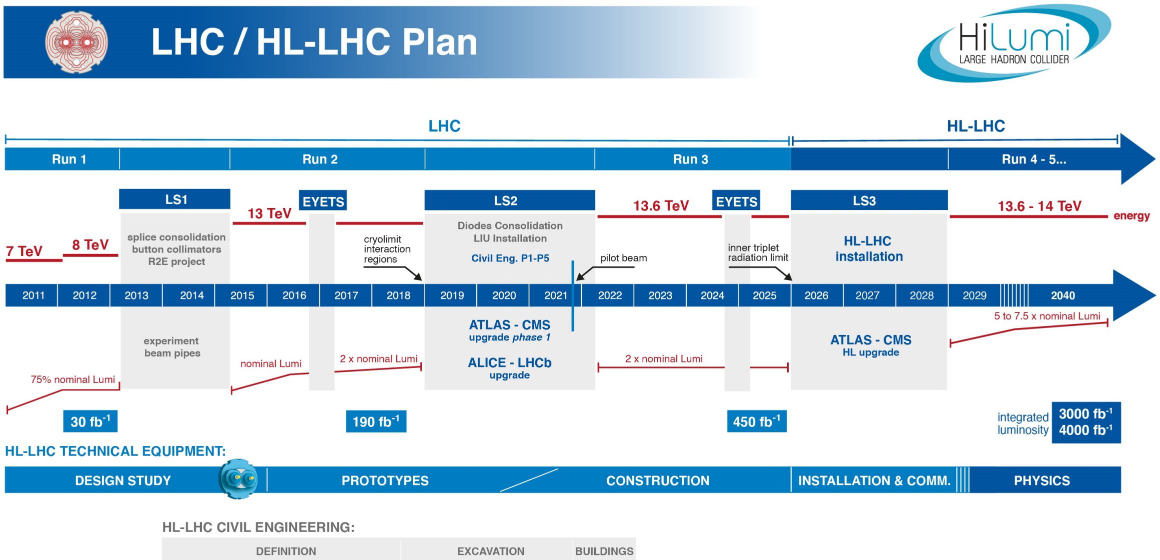


Figure 22: Timeline of LHC operation starting from 2011 to planned HL-LHC upgrade. Taken from https://hilumilhc.web.cern.ch/content/hl-lhc-project.

8 Future Upgrades

8.1 High Luminosity LHC

The planned High Luminosity Large Hadron Collider (HL-LHC) is expected to operate starting mid-2029. The primary goals of HL-LHC are to collect large high-quality data statistics needed to study rare SM processes such as Higgs self-interaction, Higgs couplings to lighter particles, the longitudinal component of vector boson scattering, and to extend the BSM searches beyond the current reach of LHC. The HL-LHC upgrade aims to increase the center-of-mass energy of proton-proton collisions to $\sqrt{s} = 14$ TeV and the instantaneous luminosity up to $\mathcal{L} = 7.5 \times 10^{34} cm^{-2} s^{-1}$ [63]. Figure 22 shows the complete operation of LHC starting in 2011 to the planned decade-long HL-LHC program.

8.2 ATLAS Upgrades

At HL-LHC, about 200 interactions per bunch crossing are expected, giving rise to several detector challenges, such as higher detector occupancy, harsher radiation conditions, and higher particle fluxes [63]. The atlas detector will undergo a Phase-II upgrade to upgrade several sub-systems of ATLAS to meet the challenges of HL-LHC. The main upgrades include upgrading the muon system by adding new chambers in the inner barrel region, upgrading the trigger & data acquisition system to meet challenges from higher detector occupancy, and upgrading the electronics of several sub-systems. Additionally, a new High Granularity Timing Detector (HGTD) will also be inserted in the end-caps to supplement the tracking system and, most importantly replacement of the current ID with all Silicon Inner Tracking Detector (ITk) [63].

The ITk consists of Silicon pixel and strip detectors to increase granularity and radiation hardness with less material in the detector. Figure 23 schematically shows the ITk layout with 5 inner layers of pixel detector and four outer layers of strips detector. The tracking for ITk is extended in the forward region up to $|\eta| < 4.0$ region [64].

More extensive statistics, extended tracking in the forward region, and the timing information from the HGTD HL-LHC program are highly beneficial to VBS $ZZjj$ measurements with extremely small cross-sections and two jets in the forward regions.

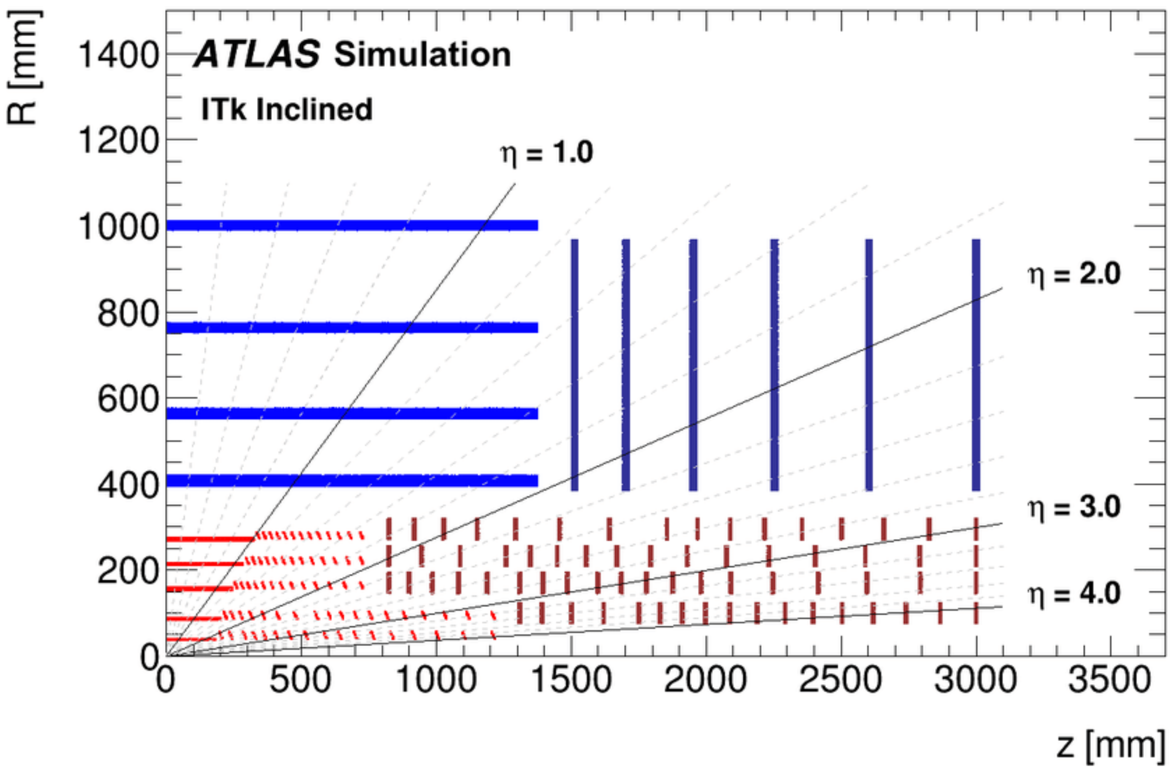


Figure 23: Schematic layout of ITK [16].

Chapter IV: Analysis Overview

9 Goals

The primary goal of the analysis is to measure the differential cross-sections of the kinematic observables sensitive to the EWK $ZZjj \rightarrow 4\ell jj$ production mode. The differential cross-sections measured in VBS-enhanced phase space are used in the precision study of the SM $4\ell jj$ production and constrain the effects of BSM physics. For simpler re-interpretation in the future without ATLAS detector simulations, the differential cross-sections are measured at a particle level using an unfolding technique, which corrects the detector effects. The details of the unfolding to extrapolate the particle-level yield from detector-level yield will be discussed in Section 17. The unfolded cross-sections shown in Section 20 are then used to constrain the effect of BSM in a model-independent framework using the Effective Field Theory (EFT) approach, which will be discussed in Section 21.

10 Phase Space Definition

The unfolded differential cross-sections are measured in a phase space within the acceptance of the detector. This section summarizes the selections defining the fiducial phase space of the analysis.

10.1 Fiducial Volume

The fiducial phase space consists of events with $pp \rightarrow ZZjj \rightarrow 4\ell jj$ ($\ell = e,\ \mu$) with four centrally produced prompt-leptons and two jets with large rapidity gap as motivated by section 4. The fiducial phase space does not contain any leptons from the decays of unstable taus. Both particle-level electrons and muons are required to be at a dressed level. Dressed leptons in MC generators are constructed by adding the four-momenta of nearby photons emitted by the lepton within a cone size of $\Delta R < 0.1$.

To ensure the selected events fall within detector acceptance, several kinematic cuts summarized in Table 4 are applied individually to the muons, electrons, and jets before defining the event quadruplet and dijet. Each electrons are required to have $p_T > 7$ GeV and $|\eta| < 2.47$, whereas the muons satisfy $p_T > 5$ GeV and $|\eta| < 2.7$. Lepton quadruplets are formed by requiring two same-flavor, SF-OC lepton pairs, with leading and sub-leading lepton $p_T > 20$ GeV and angular separation between any two leptons to satisfy $\Delta R > 0.05$. Additionally, the invariant mass of any SF-OC lepton pair is required to satisfy $m_{\ell\ell} > 5$ GeV to suppress the contamination from lower resonance backgrounds. Based on these requirements, the quadruplets can be of the following three types:

- $4e$: events with two e^+e^- pairs.
- 4μ: events with two $\mu^+\mu^-$ pairs.
- $2e2\mu$ or $2\mu2e$: events where one of the pair is e^+e^- and other is $\mu^+\mu^-$

In any event with more than two SF-OC lepton pairs, the quadruplet is formed by choosing the two pairs that minimize the distance to the Z resonance pole. Once the quadruplet is formed, the leading-lepton pair is defined as the one with a higher absolute rapidity value, i.e., $|y_{ij}|$. Finally, an additional criterion of $m_{4\ell} > 130$ GeV is imposed on the invariant mass of the quadruplet.

Similarly, the di-jet in the fiducial phase space are also constructed from the leading-dressed jets with opposite sign of pseudo-rapidity (η) to imitate the detector-level VBS di-jet production where jets are reconstructed on the opposite side of the detector. The jets are required to satisfy $|n| < 4.5$, $p_{T,\ leading\ jet} > 40$ GeV, and $p_{T,\ sub-leading\ jet} > 30$ GeV. The di-jet is required to have a large rapidity separation of $|\Delta y_{jj}| > 2$ and $m_{jj} > 300$ GeV to resemble dijet produced in electroweak $ZZjj$ production. Table 5 summarizes the requirements to select quadruplet and the di-jet in an event.

Table 4: Details of the kinematic pre-selection applied to the dressed baseline electrons, muons, and jets.

Selections	Electrons	Muons	Jets
p_{T}	> 7 GeV	> 5 GeV	> 30GeV
$\|\eta\|$	< 2.47	< 2.7	< 4.5

Table 5: Details of the selections applied to form a quadruplet and a dijet selection in the fiducial volume.

Selections	Cut
Lepton Kinematics	$P_{T,\ leading\ lepton} > 20$ GeV $P_{T,\ sub-leading\ lepton} > 20$ GeV
Pair Requirement	$\Delta R_{\ell i,\ell j} > 0.05$ SF-OC with $m_{\ell\ell} > 5$ GeV
Quadruplet Requirement	2 pair candidates with smallest $\|m_{12} - m_Z\| + \|m_{34} - m_Z\|$ Leading pair: pair with highest $\|y_{ij}\|$ Sub-leading pair: pair with lowest $\|y_{ij}\|$ $m_{4\ell} > 130$ GeV
Di-jet Requirement	$p_{T,\ leading\ jet} > 40$ GeV $\|\Delta y_{jj}\| > 2$ $m_{jj} > 300$ GeV

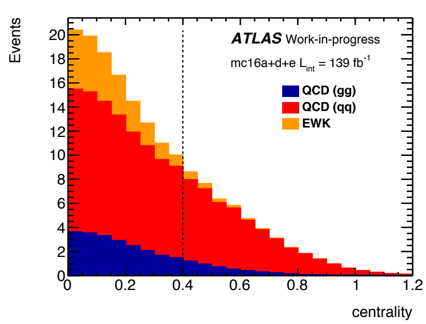
10.2 Signal Region

The signal region of the analysis is defined based on the centrality (ζ) of the di-Zboson production in an event. Centrality depends on the rapidity of the quadruplet and the rapidity of the dijet as:

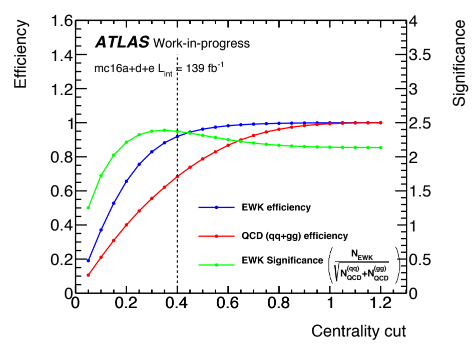
$$\zeta = \frac{|y_{quadruplet} - 0.5 * (y_{leading\ jet} + y_{sub-leading\ jet})|}{|y_{leading\ jet} - y_{sub-leading\ jet}|} \tag{10.1}$$

Figure 24a shows the distribution of centrality in MC for the three main production modes of $ZZjj$. The chosen cut value on the $ZZjj$ centrality maximizes the significance of the EWK component over the inclusive qq and gg-initiated QCD production (defined as $s = \frac{N_{EWK}}{\sqrt{N_{QCD}^{(qq)}+N_{QCD}^{(gg)}}}$) while maintaining a good selection efficiency of EWK events. The second distribution in 24b shows the efficiency and significance for various cut values.

A VBS-enhanced signal region is defined based on events with a quadruplet, a dijet, and $\zeta < 0.4$. The low value of the centrality and the requirements for a signal dijet ensures that the events in this signal region originate in a more significant fraction from the electroweak production of $ZZjj$. A VBS Suppressed control region is also defined based on events with a quadruplet, a dijet, and $\zeta > 0.4$. These events mainly originate from the QCD production of ZZjj and are used to optimize the analysis strategies.



(a) Yields of EWK(red) and QCD (parton initiated in blue, gg-loop initiated in green) ZZjj production as a function of centrality.



(b) Selection efficiency (EWK in blue, QCD in red) and EWK significance (green) for different centrality cut values. The dashed line highlights the selected cut values of 0.4.

11 Reconstruction Selection

This section summarizes the detector-level phase space selections applied to three physics objects, electrons, muons, and jets used in the measurement. Each physics object of the analysis has two categories: *baseline* and *signal* objects. Physics objects satisfying a set of kinematic selections or looser identification criteria are categorized as *baseline* whereas, the baseline leptons that pass either stricter kinematic selections or additional isolation and track-to-vertex association (TTVA) requirements are *signal*.

11.1 Electrons

As discussed in Section 7.3, electrons are reconstructed by matching the inner detector track (ID) to an energy cluster in the electromagnetic calorimeter. Baseline electron objects are required to satisfy the kinematic selections of $p_{\mathrm{T}} > 7$ GeV & $|\eta| < 2.47$ and a loose likelihood identification of working point *LHVeryLoose*. To avoid the electrons from pileup, a loose vertex association requirement of $|z_0 \sin\theta| < 0.5$ mm and an overlap removal discussed in section 11.4 is applied to the baseline electron candidates.

Signal electrons are required to pass a more stringent loose likelihood identification, *LHLooseBL*, which requires at least one hit in the innermost layer of the pixel detector. The signal electrons are distinguished by requiring the baseline electrons to have impact parameter requirements of $d0/\sigma_{d0} < 5$ and an isolation working point identification of *LooseVarRad*. Table 6 summarizes the several selections imposed to define the baseline and signal electrons.

11.2 Muons

As discussed in section 7.4, muons are reconstructed in multiple ways based on information from the inner detector (ID), the muon spectrometer (MS), and the calorimeters. All baseline muons are required to satisfy $|\eta| < 2.7$, $p_T > 5$ GeV, a loose impact parameter requirements of $|z_0 \sin\theta| < 0.5$ mm, lepton-favoring overlap removal and *Loose* identification working

Table 6: Definition of the baseline and signal electrons.

Selection Category	**Baseline**	**Signal**
Kinematic cuts	p_T > 7 GeV $\|\eta\| < 2.47$	p_T > 7 GeV $\|\eta\| < 2.47$
Identification	LHVeryLoose	LHLooseBL
Vertex Association	$\|z_0 \sin\theta\| < 0.5$ mm	$\|z_0 \sin\theta\| < 0.5$ mm
Overlap removal	Lepton-favored	Lepton-favored
Isolation Working Point	–	LooseVarRad
Impact Parameters	–	$d_0/\sigma_{d_0} < 5$

point. The signal muons are identified by requiring additional isolation identification of *PflowLooseVarRad* and TTVA requirements of $d_0/\sigma_{d_0} < 3$. Table 7 summarizes baseline and signal muons selection requirements.

Table 7: Definition of the baseline and signal muons.

Selection Category	**Baseline**	**Signal**
Kinematic cuts	p_T > 5 GeV Calo-tagged p_T > 15 GeV $\|\eta\| < 2.7$	p_T > 5 GeV Calo-tagged p_T > 15 GeV $\|\eta\| < 2.7$
Identification	Loose	Loose
Vertex Association	$\|z_0 \sin\theta\| < 0.5$ mm	$\|z_0 \sin\theta\| < 0.5$ mm
Overlap removal	Lepton-favored	Lepton-favored
Isolation Working Point	–	PflowLooseVarRad
Impact Parameters	–	$d_0/\sigma_{d_0} < 3$

11.3 Jets

Jets are reconstructed with the particle flow anti$-K_T$ clustering algorithm using a radius parameter of $R = 0.4$ as discussed in section 7.5. The jets reconstructed using the particle flow algorithm are required to satisfy p_T > 15 GeV, $|\eta| < 4.5$ kinematic cuts, and the lepton-favored overlap removal to be classified as baseline jets. Baseline jets satisfying the *Tight* working point of the jet to the vertex tagger tool are classified as signal jets. *Jet-vertex-tagger (JVT)* is applied to the baseline jets with $|\eta| < 2.4$ whereas the *forward-jet-vertex-tagger (fJVT)* tool is applied to the baseline jets with $|\eta| > 2.5$. Table 8 summarizes the details of baseline and signal jets selection.

Table 8: Definition of the baseline and signal jets.

Selection Category	**Baseline**	**Signal**
Kinematic cuts	$p_{\mathrm{T}} > 30$ GeV $\lvert\eta\rvert < 4.5$	$p_{\mathrm{T}} > 30$ GeV $\lvert\eta\rvert < 4.5$
Identification	AntiKt4EMPFlow	AntiKt4EMPFlow
Overlap removal	Lepton-favored	Lepton-favored
Jet-Vertex-Tagger	– –	$\lvert\eta\rvert < 2.4$ JVT ("Tight") $\lvert\eta\rvert > 2.5$ fJVT ("Tight")

11.4 Overlap Removal

An *overlap removal* procedure is applied to remove the physics objects reconstructed from the same detector signal. The measurement uses a lepton-favored overlap removal which selects leptons over jets. Overlap removal is an iterative process in which only objects surviving all previous steps are used in the subsequent steps. Table 9 summarizes the overlap removal steps, where the ΔR is the angular separation between objects calculated using rapidity.

Table 9: Overlap removal used in the analysis. An object removed in one step does not enter into the subsequent step.

Remove Object	Accept Object	Overlap Criteria
Electron	Electron	Share a track or have overlapping calorimeter cluster. Keep electron with higher p_T
Muon	Electron	Share ID track, and the muon is calo-tagged
Electron	Muon	Share ID track
Jet	Electron	$\Delta R_{e-jet} < 0.2$
Jet	Muon	$\Delta R_{\mu-jet} < 0.2$/ghost-associated and $N_{jet\ tracks} < 3$

12 Trigger

Due to the presence of four fully reconstructed leptons in the final state, the data events and detector-level MC events are preselected using a logical OR of different single and double-lepton triggers. The trigger menu varies according to the data-taking run periods to reflect the changes in the high-level trigger system, which are required to cope with increasing data rates. Additionally, trigger matching is required for the selected events. The trigger matching selects a subset of preselected events in which at least one lepton of the quadruplet is matched to one of the fired triggers. Table 10 shows the trigger menu used by the analysis per different data periods using either electrons, muons, or mixed electron-muon triggers.

The trigger efficiency of MC is defined as a ratio of events passing the logical OR selection of the triggers to the number of events passing reconstruction level pre-trigger selection. The trigger efficiency for MC events could differ from that for the data. Thus, trigger efficiency scale factors are applied to MC events to account for the differences. The scale factors are defined as a fraction of trigger efficiency for MC to that of data and retrieved from the ATLAS supported tool *TrigGlobalEffciencyCorrectionTool*[2].

Figure 25 shows the efficiency of trigger selection in events with a signal quadruplet and a dijet as a function of $m_{4\ell}$. The efficiency is almost 100% over the whole spectrum. The black distribution shows the total detector-level preselected events, the red distribution shows events passing at least one trigger requirement, and the blue distribution shows events passing both trigger and matching requirements.

[2]https://gitlab.cern.ch/atlas/athena/tree/21.2/Trigger/TrigAnalysis/TrigGlobalEfficiencyCorrection

Period	Leptons	Triggers
2015	Electron	HLT_e24_lhmedium_L1EM20VH
		HLT_e60_lhmedium
		HLT_e120_lhloose
		HLT_2e12_lhvloose_L12EM10VH
	Muon	HLT_mu20_iloose_L1MU15
		HLT_mu50
		HLT_2mu10
		HLT_mu18_mu8noL1
	Mixed	HLT_e7_lhmedium_mu24
		HLT_e17_lhloose_mu14
2016	Electron	HLT_e26_lhtight_nod0_ivarloose
		HLT_e60_lhmedium_nod0
		HLT_e140_lhloose_nod0
		HLT_2e17_lhvloose_nod0
	Muon	HLT_mu26_ivarmedium
		HLT_mu50
		HLT_2mu14
		HLT_mu22_mu8noL1
	Mixed	HLT_e7_lhmedium_nod0_mu24
		HLT_e17_lhloose_nod0_mu14
2017	Electron	HLT_e26_lhtight_nod0_ivarloose
		HLT_e60_lhmedium_nod0
		HLT_e140_lhloose_nod0
		HLT_2e24_lhvloose_nod0
	Muon	HLT_mu26_ivarmedium
		HLT_mu50
		HLT_2mu14
		HLT_mu22_mu8noL1
	Mixed	HLT_e17_lhloose_nod0_mu14
		HLT_e26_lhmedium_nod0_mu8noL1
2018	Electron	HLT_e26_lhtight_nod0_ivarloose
		HLT_e60_lhmedium_nod0
		HLT_e140_lhloose_nod0
		HLT_2e24_lhvloose_nod0
	Muon	HLT_mu26_ivarmedium
		HLT_mu50
		HLT_2mu14
		HLT_mu22_mu8noL1
	Mixed	HLT_e17_lhloose_nod0_mu14
		HLT_e26_lhmedium_nod0_mu8noL1

Table 10: Trigger menu used in the analysis for event preselection

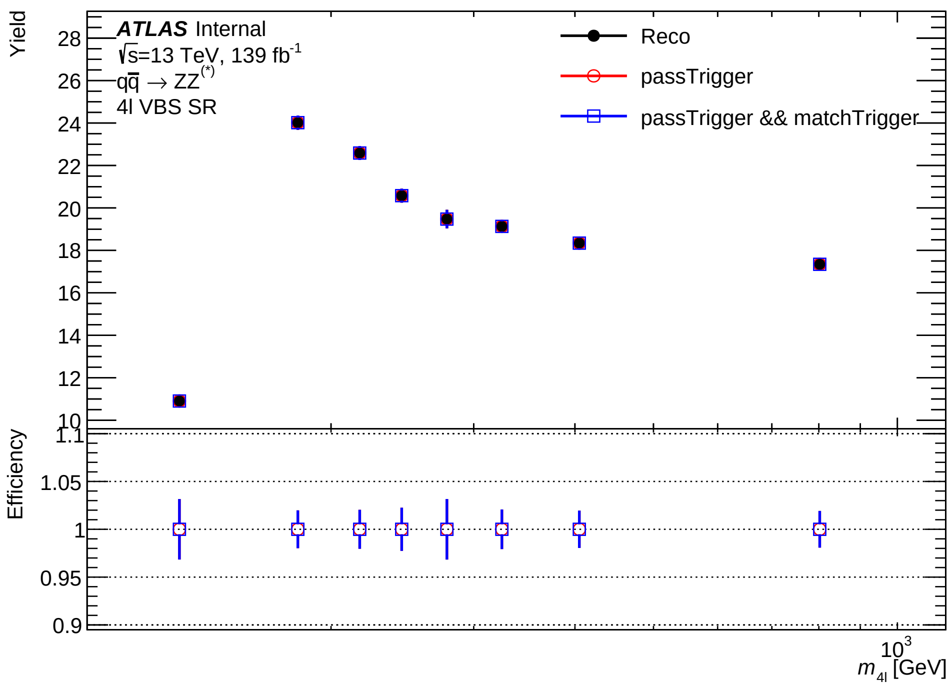


Figure 25: Trigger efficiency as a function of $m_{4\ell}$.

13 Event Selection

A $ZZjj$ event at the detector level consists of a lepton quadruplet formed from SF-OC baseline-lepton pairs and a dijet, passing similar selections as the fiducial level defined in section 10. The leading and sub-leading leptons are required to satisfy $p_T > 20$ GeV to ensure a high trigger efficiency. From the leptons passing these requirements, at least two SF-OC lepton pairs with $\Delta R > 0.05$ and $m_{\ell\ell} > 5$ GeV are formed. A quadruplet is formed from the two SF-OC lepton pairs whose invariant masses are closest and next closest to the mass of the Z-boson (m_Z). Similar to the fiducial level selection, the lepton pair with the highest value of absolute rapidity is identified as the leading pair. The quadruplets with all four leptons passing the signal lepton criteria of the TTVA and isolation are the *signal quadruplet* defining the signal region. While on the contrary, the quadruplets where one lepton fails either isolation or TTVA requirement used in the fake background estimation are the *not-signal quadruplets.*

A dijet in an event is selected by requiring two signal jets defined in section 11.3 from the opposite side of the detector i.e., $\eta_{lead\ jet} \times \eta_{sub-leading\ jet} < 0$). To maximize the probability of selecting an event from EWK $ZZjj$ production, a requirement of significant rapidity difference between the jets of $\Delta Y_{jj} > 2$ and a large invariant mass of $m_{jj} > 300$ GeV are imposed on the dijet selection. Table 11 summarizes all selections applied to select $ZZjj$ detector-level events.

Figure 26 illustrates a signature of two Z-bosons production in an association of two jets. The event display corresponds to an event recorded during Run Number 340368 of the 2017 data-taking period. The two light-yellow cones on two opposite sides of the detector with a large rapidity gap represent the reconstructed dijet of the event with $m_{jj} = 2228$ GeV. In this event, one of the SF-OC lepton pairs decays to e^+e^- ($Z \to e^+e^-$), and the other decays into $\mu^+\mu^-$ ($Z \to \mu^+\mu^-$).

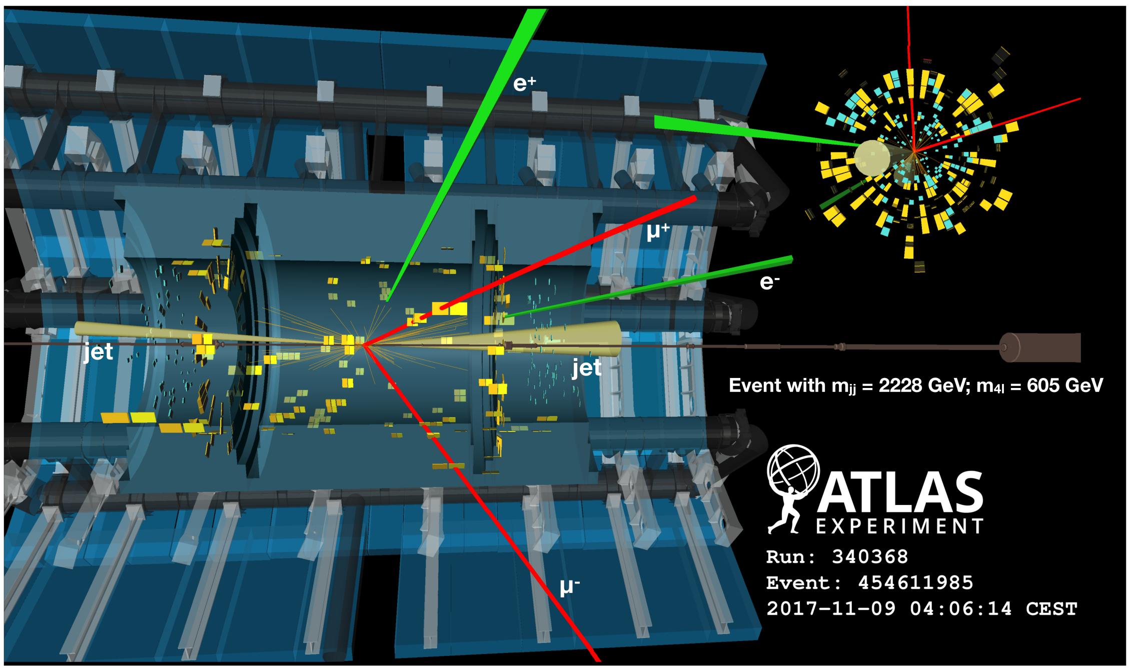


Figure 26: Event display of a candidate $pp \to ZZjj \to e^+e^-\mu^+\mu^-jj$ recorded by the ATLAS experiment in Run-2 2017 data-taking period. The invariant mass of the reconstructed four leptons is $m_{4\ell} = 605$ GeV, and that of the reconstructed di-jet is $m_{jj} = 2228$ GeV. The large rapidity separation between the two jet cones (light yellow) on the opposite sides of the ATLAS detector and centrally produced two Z bosons defines the characteristic feature of the EWK production of ZZjj [17].

Table 11: Details of event selection.

Event Selection	Cut	Requirement
Event Preselection	Trigger	Fire at least one lepton trigger
	Vertex	At least one vertex with 2 or more tracks
Quadruplet Selection	Lepton Kinematics	p_T > 20 GeV for two leading leptons
	Lepton Separation	$\Delta R_{ij} > 0.05$ between leptons in quadruplet
	Pair Requirement	Two SF-OC lepton pairs
		$m_{\ell\ell} > 5$ GeV
	Minimal Δm_Z	quadruplet with smallest $\lvert m_{12} - m_Z\rvert + \lvert m_{34} - m_Z\rvert$
		Leading Pair: pair with highest $\lvert y_{ij}\rvert$
	ZZ Mass	$m_{4\ell} > 130$ GeV
Quadruplet Categorisation	Signal Quadruplet	Quadruplet with all **signal leptons**
	Not-Signal Quadruplet	Quadruplet with ≥ 1 **baseline-not-signal lepton**
Dijet Selection	Different Detector Sides	$\eta_{lead\ jet} \times \eta_{sub-leading\ jet} < 0$
	Rapidity Separation	$\Delta Y_{jj} > 2$
	Leading Jet p_T	$p_{T,\ leading\ jet} > 40$ GeV
	Dijet Mass	$m_{jj} > 300$ GeV
	Dijet	Both jets required to pass either JVT or FJVT
Event Categorisation	VBS-Enhanced Region	signal quadruplet & dijet and centrality $(\zeta) < 0.4$
	VBS-Suppressed Region	signal quadruplet & dijet and centrality $(\zeta) > 0.4$

14 Datasets and Monte Carlo Simulation

14.1 LHC Dataset

The measurement uses the LHC collision data, named the ATLAS Run-2 dataset collected by the ATLAS experiment during its operation in 2015, 2016, 2017, and 2018. This dataset corresponds to proton-proton collisions at the center-of-mass energy of $\sqrt(s) = 13$ TeV and total integrated luminosity of 139 ± 2.4 fb^{-1} measured by the LUCID-2 detector [65] [66]. The uncertainty on the integrated luminosity is obtained by combining the measurements of LHC runs each year. Each data-taking run period is further divided into sub-periods of one to three weeks that vary in beam and detector conditions. The dataset used in physics analyses is required to satisfy a series of data quality checks discussed in detail in Ref [18]. The data passing these requirements collectively form a Good Run List (GRL) consisting of several luminosity blocks (LB). Figure 27 shows the total integrated luminosity delivered by LHC in the green distribution, recorded by the ATLAS experiment in the yellow distribution and part of the GRL in the blue distribution. The plateaus correspond to the end-of-year shutdowns of LHC, and the slopes correspond to the increasing instantaneous luminosity in different data-taking periods.

The measurement uses the following data samples from the GRL,

- GoodRunsLists/data15_13TeV/20170619/PHYS_StandardGRL_All_Good_25ns_276262-284484_OflLumi-13TeV-008.root
- GoodRunsLists/data16_13TeV/20180129/PHYS_StandardGRL_All_Good_25ns_297730-311481_OflLumi-13TeV-009.root
- GoodRunsLists/data17_13TeV/20180619/physics_25ns_Triggerno17e33prim.lumicalc.OflLumi-13TeV-010.root
- GoodRunsLists/data18_13TeV/20180924/physics_25ns_Triggerno17e33prim.lumicalc.OflLumi-13TeV-001.root

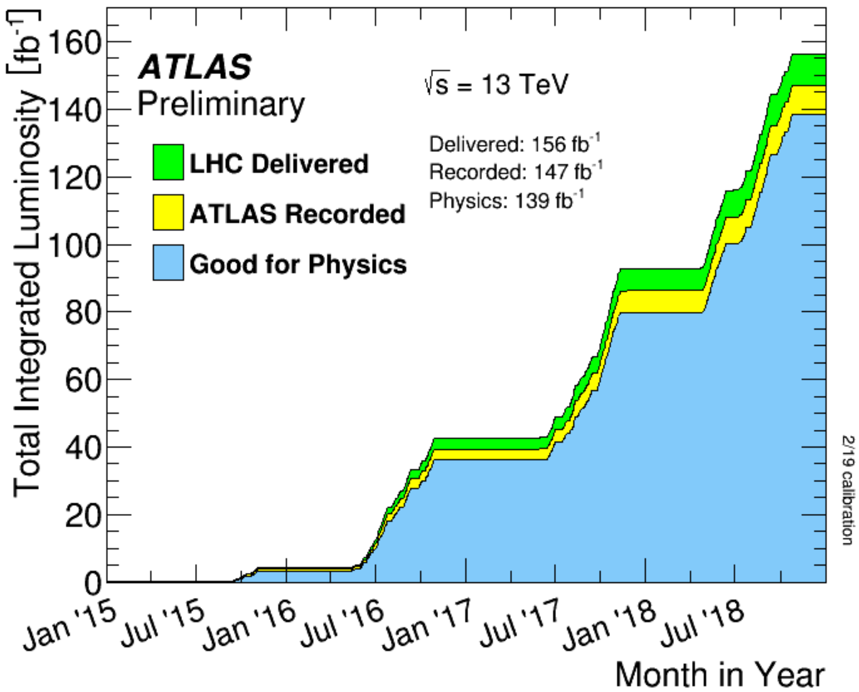


Figure 27: Total integrated luminosity collected during data taking period in Run-2 [18].

14.2 Monte Carlo Samples

As briefly mentioned in Section 3, MC generates the $pp \to ZZjj \to 4\ell jj$ events incorporating the matrix element calculations for the hard-scatter $ZZjj \to 4\ell jj$ production, the parton showering, hadronization, the effect of the underlying events, and pile-up. The generated events are then simulated to interact with the ATLAS material using the Geant4 simulation toolkit following the description in Ref [67]. The energy deposits of the simulated events in the detectors are then digitized and reconstructed using a detector geometry corresponding to the data-taking period. Figure 28 shows a schematic overview of the MC generation.

14.2.1 Signal Samples

As discussed in section 4, two types of interaction, QCD and EWK, give us $pp \to ZZjj \to 4\ell jj$ final state. The two types of QCD process, quark induced $qqZZ$ ($qq \to ZZ^* \to 4\ell jj$) and gluon induced $ggZZ$ ($gg \to ZZ^* \to 4\ell jj$) are simulated using the Sherpa 2.2.2 MC

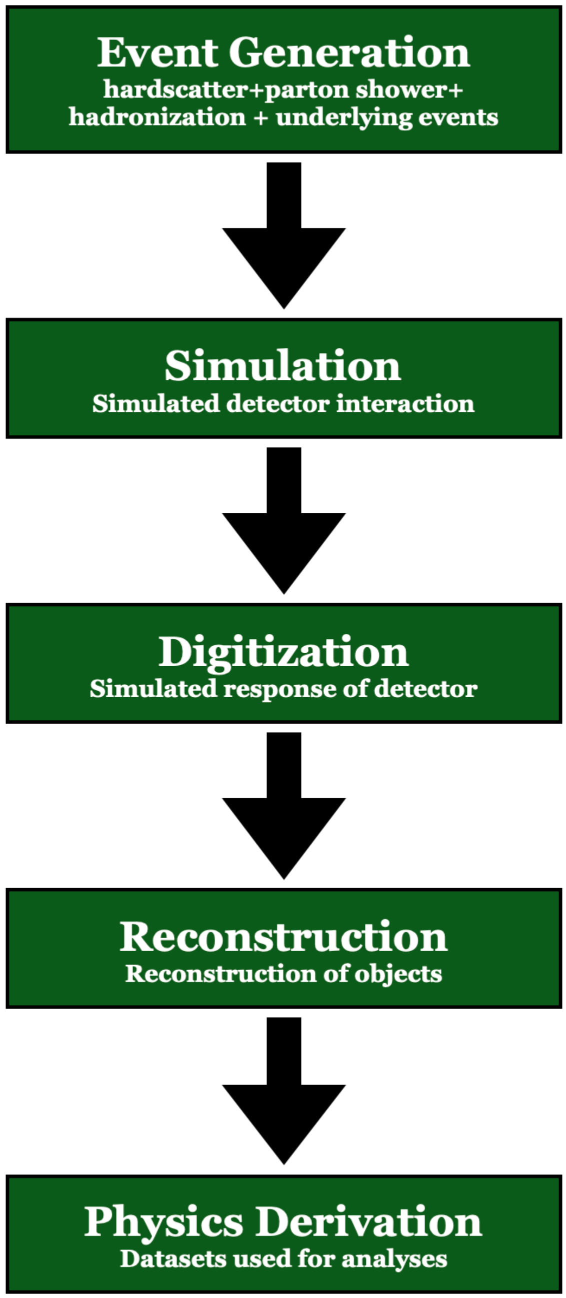


Figure 28: Various steps in MC sample generation.

generator. The $qqZZ$ and $ggZZ$ samples corresponding to figure 4 are generated with NLO accuracy in QCD up to one additional parton emission and LO accuracy for up to three additional partons emission. The loop-induced $ggZZ$ samples emerging at NNLO in α_S corresponding to figure 5 are generated using LO-accurate matrix elements for up to one additional parton emission [68]. Results from The generator uses an NNPDF3.0NNLO PDF set evaluated using different measurements from several experiments, such as deep-inelastic inclusive cross-sections measurement from HERA-II, the combined charm data from HERA, jet production, vector boson rapidity and transverse momentum measurements from ATLAS, CMS and LHCb, total cross sections of top quark pair production from ATLAS and CMS and W+c data from CMS [69]. Parton showering is done by SHERPA's internal algorithm based on Catani–Seymour dipole factorization matrix element [70]. The matrix element calculations are matched and merged using the $ME + PS@NLO$ prescription [71].

An alternative MADGRAPH5 samples produced at NLO accuracy for up to one additional parton emission and LO accuracy for up to three additional parton emission [72] are also used in the measurement for the parton induced $qqZZ$ samples. The generator uses A14NNPDF23LO PDF set, and the ME is interfaced with PYTHIA8 for parton showering, merging, and matching [73].

The EWK production $qqZZjj$ ($qq \to ZZ^{(*)}jj \to 4\ell jj$) is simulated using a POWHEGV2 generator using an MSTW2008 PDF set with NLO accuracy in QCD correction and interfaced with PYTHIA8 for parton showering and hadronization [74]. An alternative sample at LO accuracy is also used in the measurement from MADGRAPH5 with A14NNPDF23LO PDF set and PYTHIA8 showering [72]. The POWHEGV2 NLO prediction of electroweak $qqZZjj$ does not contain the contribution from electroweak triboson VZZ processes where two vector bosons decay leptonically and one decays hadronically. The contribution from these electroweak triboson processes is predicted using the Sherpa 2.2.2 MC generator at LO accuracy for up to two additional parton emissions and added to the POWHEGV2 predictions. Table 12 summarizes the signal MC used in the measurement.

Process	Description	Generator	PDF	Accuracy
QCD				
$q\bar{q} \to ZZ^{(*)} \to 4\ell$	inclusive	SHERPA2.2.2 MADGRAPH	NNPDF3.0NNLO A14NNPDF23LO	$0, 1j@NLO + 2, 3j@LO$
QCD gg loop $gg \to ZZ^{(*)} \to 4\ell$	$m_{4\ell} > 130$ GeV	SHERPA2.2.2	NNPDF3.0NNLO	$0, 1j@LO$
EWK				
$q\bar{q} \to ZZ^{(*)}jj \to 4\ell jj$	$m_{4\ell} > 130$ GeV	POWHEGV2 MADGRAPH	MSTW2008 A14NNPDF23LO	$\geq 2j$ (EWK) @ NLO QCD $\geq 2j$ (EWK) @LO
EWK $q\bar{q} \to ZZ^{(*)}jj \to 4\ell jj$		SHERPA2.2.2	NNPDF3.0NNLO	$1, 2j@LO$

Table 12: List of signal MC samples used in the analysis. Each process consists of three different generation campaigns corresponding to the data-taking conditions of the ATLAS Run2 data-taking periods.

14.2.2 Background Samples

In addition to the QCD and EWK production discussed above, two other processes, triboson (WWZ, WZZ, ZZZ) and Z-bosons production in association with top quark pair ($t\bar{t}Z$), also contributes to the $ZZjj \to 4\ell jj$ final state. The triboson processes are modeled with SHERPA2.2.2 generator at NLO accuracy in QCD for zero or one additional parton emissions and LO accuracy for up to two additional parton emissions. The triboson samples only include the fully leptonic decays of the vector bosons. Therefore, there is no overlap between the background triboson and the signal EWK $qqZZjj$ samples. The $t\bar{t}Z$ processes are modeled by SHERPA2.2.0 generator at LO accuracy with up to one additional parton emission using the MEPS@LO set-up [75]. The same algorithms as in the QCD $qqZZ$ sample generation are used for parton showering, matching, and merging. The MC simulation of the triboson and $t\bar{t}Z$ samples are subtracted directly from the data. Table 13 summarizes the details of these samples.

Process	Description	Generator	PDF	Accuracy
$pp \to W^{(*)}W^{(*)}Z^{(*)} \to 4\ell 2\nu$		SHERPA2.2.2		
$pp \to W^{(*)}Z^{(*)}Z^{(*)} \to 5\ell 1\nu$	inclusive	SHERPA2.2.2	NNPDF3.0NNLO	$0, 1j@NLO + 2j@LO$
$pp \to Z^{(*)}Z^{(*)}Z^{(*)} \to 6\ell$		SHERPA2.2.2		
$pp \to t\bar{t} + Z(\to 2\ell)$	$m_{ll} > 5$ GeV	SHERPA2.2.0	NNPDF3.0NNLO	LO

Table 13: List of background MC samples used in the analysis. Each process consists of three different generation campaigns corresponding to the data-taking conditions of the ATLAS Run2 data-taking periods.

14.2.3 Samples for Fake Background

In addition to the triboson and $t\bar{t}$Z samples, the analysis has additional backgrounds coming from events with one or more non-prompt or fake leptons. These fake backgrounds are estimated using a data-driven method discussed in detail in Section 16.1. MC samples are used to develop and validate the data-driven fake background estimation procedure. There are three sources of events that could contribute as a source for fake background events. The first type of events is from a Z-boson production in association with jets $pp \to Z^{(*)} \to 2\ell + jets$, which is simulated for both three or more leptons using SHERPA2.2.1. The subdominant process is events from $t\bar{t} \to 2\ell$ production in which both top quarks decay semileptonically, which is simulated with POWHEG+PYTHIA8 and uses the A14NNPDF23LO PDF set [76]. The third type of fake backgrounds arises from the WZ production in which both bosons decay leptonically $pp \to WZ \to 2\ell 1\nu$ and is simulated using SHERPA2.2.2. Table 14 summarizes the different processes and MC generators used to estimate the fake background.

14.3 Event Weights

The raw predictions from the MC generators are completely unscaled and cannot be compared to the data from the detector directly. Each event generated by the MC needs to be scaled based on the cross-section of a given process normalized to the total sum of all the weights from events generated and multiplied by the integrated luminosity of the data-

Process	Description	Generator	PDF	Accuracy
$pp \to Z^{(*)} \to 2e + jets$ $pp \to Z^{(*)} \to 2\mu + jets$ $pp \to Z^{(*)} \to 2\tau + jets$	inclusive	SHERPA2.2.2	NNPDF3.0NNLO	$NLO + 2j, LO + 4j$
$pp \to t\bar{t} \to 2\ell$	inclusive	POWHEG+PYTHIA8	A14NNPDF23LO	LO
$pp \to WZ \to 2\ell 1\nu$	inclusive	SHERPA2.2.2	NNPDF3.0NNLO	$NLO + 1j, LO + 3j$

Table 14: List of MC samples used in the estimation and validation of the data-driven fake background estimation.

taking period. As shown by figure 29, the pile-up distribution is different for the different data-taking periods. The MC-generated events are modified to correctly simulate the effect of pile-up distribution imitating that of the data. Additionally, a set of measurement-related corrections are included in the event weight. These corrections, named *scaled factors (SF)*, correct the reconstruction, identification, isolation, and trigger efficiencies in the MC to match that of measured data. The total event weight for MC generated event is a product of the normalized generator weight scaled to match the pile-up profile and all scale factors.

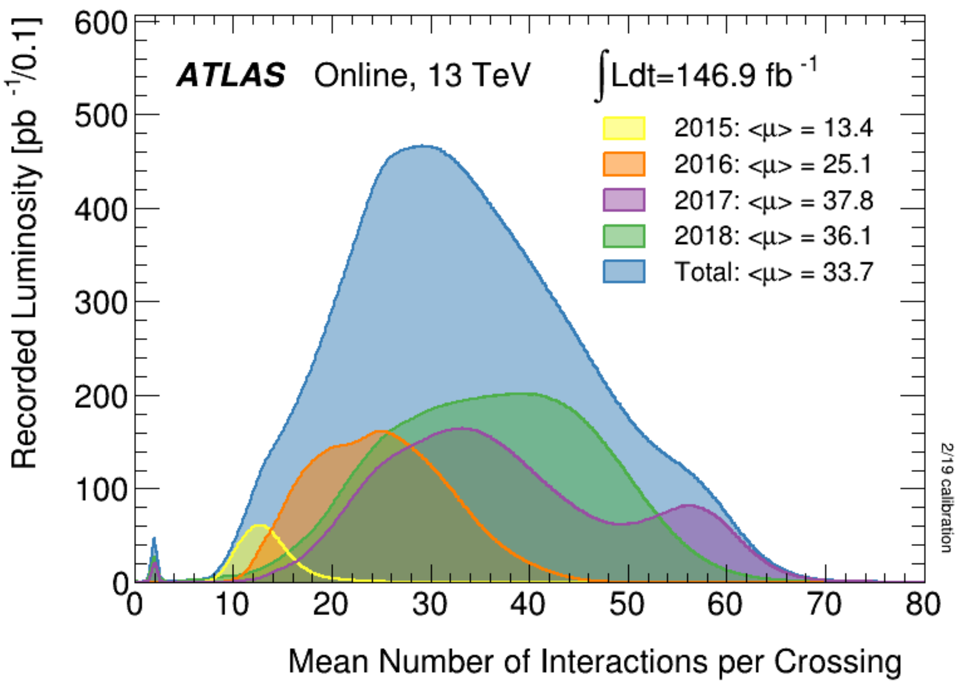


Figure 29: Pile-up distributions in different Run-2 data-taking period. [18]

15 Definition of Measured Observables

The primary results of the thesis are differential cross-sections of the following 11 different kinematic observables:

- $m_{4\ell}$: invariant mass of the four-leptons (or 2 Z-bosons)
- m_{jj}: invariant mass of the dijet
- $p_{T,4\ell}$: transverse momentum of the four-leptons
- $p_{T,jj}$: transverse momentum of the dijet
- $p_{T,4\ell jj}$: transverse momentum of the four-leptons and the dijet
- $s_{T,4\ell jj}$: scalar transverse momentum of the four-leptons and the dijet
- $\Delta\phi_{jj}^{signed}$: difference in the azimuthal angle between the two jets in the dijet, ordered according to their rapidity,i.e.

$$\Delta\phi_{jj}^{signed} = \begin{cases} \phi(j_1) - \phi(j_2) & \text{if } y_{j_1} > y_{j_2} \\ \phi(j_2) - \phi(j_1) & \text{otherwise} \end{cases}$$

- Δy_{jj}: the absolute value of rapidity difference between the leading and the sub-leading jets in the dijet
- ζ: centrality of the system
- $\cos\theta^*_{\ell 1\ell 2}$: cosine of the decay angle of the negative lepton of the leading pair in the pair's rest frame as shown by figure 30
- $\cos\theta^*_{\ell 3\ell 4}$: cosine of the decay angle of the negative lepton of the sub-leading pair in the pair's rest frame as shown by figure 30

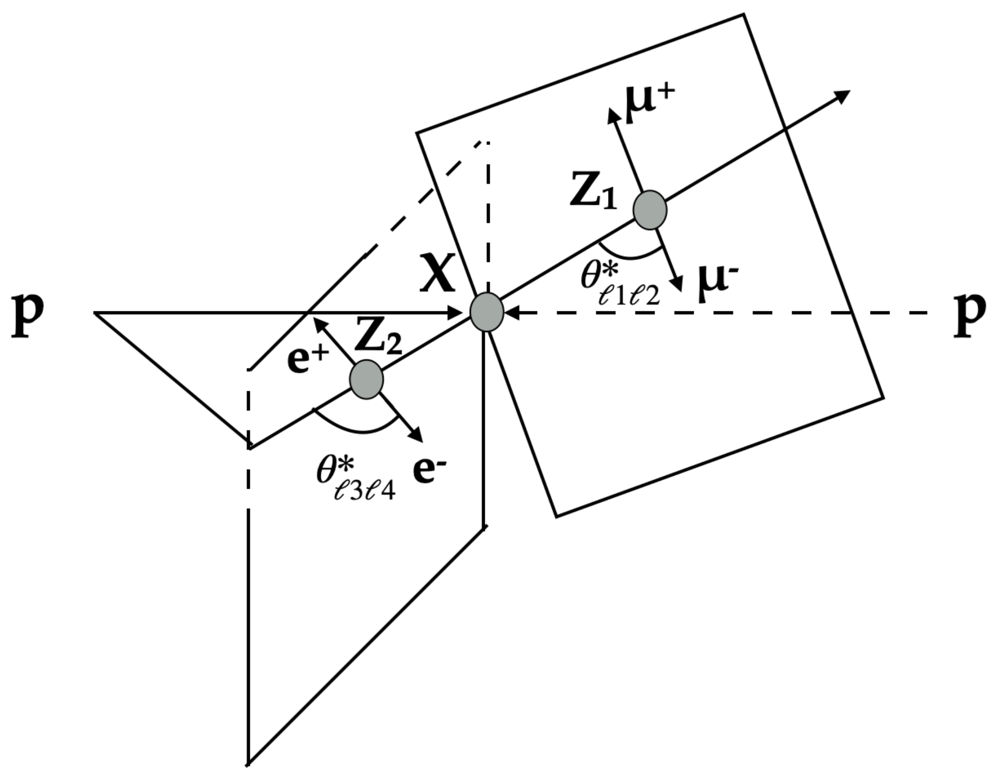


Figure 30: Figure showing the decay angle $\theta^*_{\ell1\ell2}$ ($\theta^*_{\ell3\ell4}$) of the negative lepton in the primary (secondary) pair's rest frame. [19].

Chapter V: **Analysis Strategy**

16 Background Estimation

16.1 Data Driven Estimate of Fake Background

Non-prompt leptons originate from a non-hard scatter source, either from a secondary interaction such as jet decay or from charged tracks misidentification. Figure 31 shows an example of non-prompt lepton production. The hard scatter process produces a b-jet which in secondary interaction produces a muon whose track does not point towards the interaction vertex and is surrounded by jet activities. The signal lepton criteria of isolation and TTVA discussed in Section 11 discards most of the non-prompt leptons. However, some non-prompt leptons pass the signal criteria and, in association with other prompt leptons, form a quadruplet in the signal region. Thus, giving rise to the *fake background* events for the analysis. The origins of non-prompt leptons are discussed in detail in Section 16.1.1.

The fake backgrounds could be predicted using the MC for $Z(\rightarrow \ell\ell) + jets$, $t\bar{t}$ and WZ processes where one or more non-prompt leptons in association with the prompt leptons form a signal quadruplet. However, the MC predictions of the fake background events are statistically limited. It is challenging to precisely model the non-prompt leptons originating from the reconstruction effects. Therefore, the fake backgrounds are estimated using an entirely data-driven technique discussed in this Section. Figure 32 shows the schematic of the whole background estimation process. The fake factors are evaluated from a combined control region (CR), formed by combining two independent control regions $Z + jets$ and $t\bar{t}$. Both regions are enriched in non-prompt leptons, and the combination is discussed in Section 16.1.2. Section 16.1.3 discusses the technical aspects of the fake factor method, and Section 16.1.4 discusses the fake efficiencies. The fake background is estimated by applying

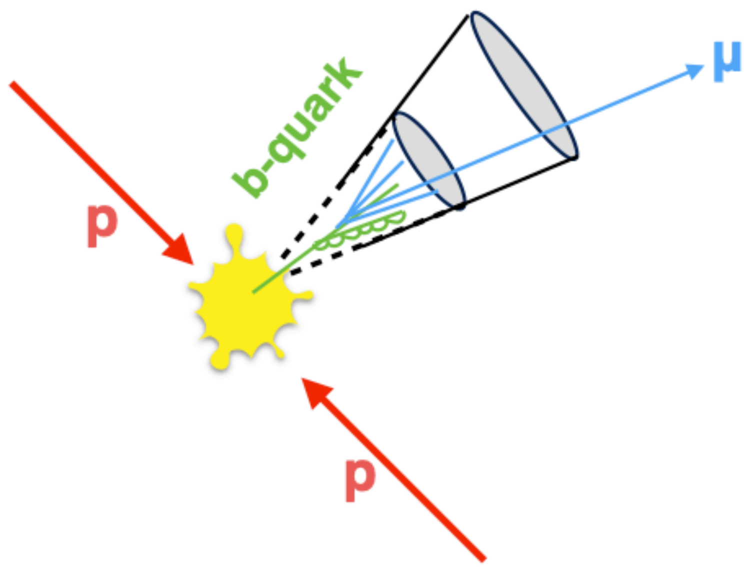


Figure 31: A schematic of the non-prompt lepton production from secondary interaction. Jet activities surround the non-prompt muon, and the muon track does not point to the hard scatter interaction point.

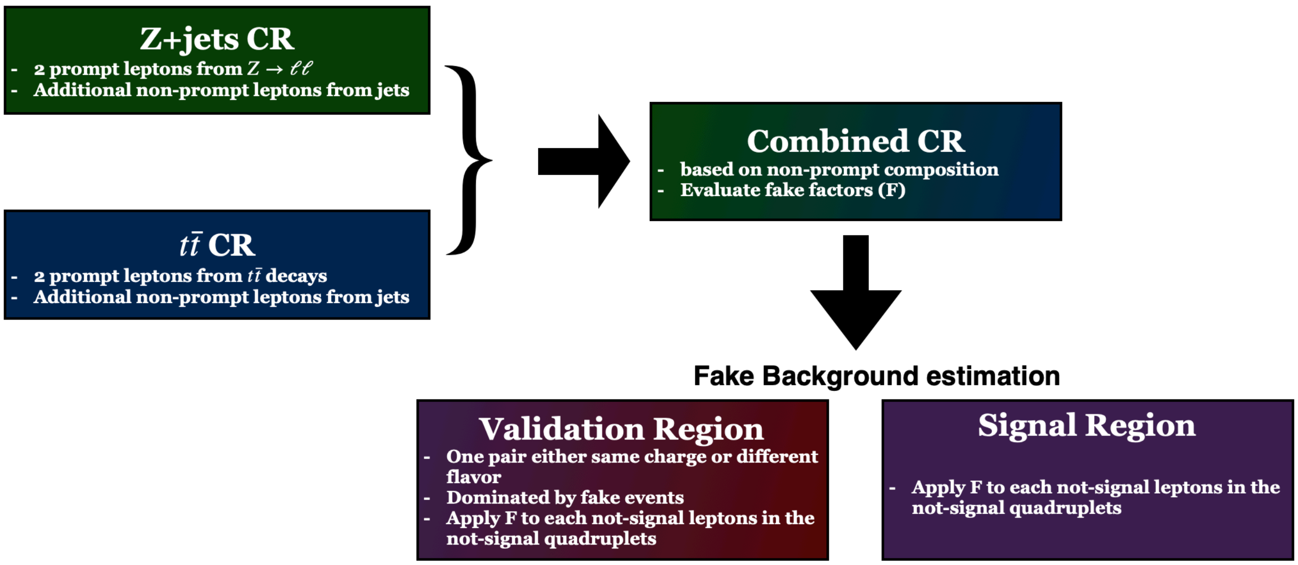


Figure 32: An overview of the fake background estimation.

the fake factors to each anti-signal lepton in not-signal quadruplets. First, the background estimation technique is validated in fake-enriched validation regions discussed in Section 16.1.5 and applied to the signal region, which is discussed in Section 16.1.6.

16.1.1 Lepton Composition

The fake background MC predictions provide essential insight into the origin of the non-prompt leptons. A classification tool developed by the Isolation and Fake Forum (IFF) identifies the true origin of the leptons, which is studied to understand the composition of non-prompt leptons in various phase-space regions of the analysis [3]. The tool has the

[3]https://gitlab.cern.ch/atlas/athena/-/tree/21.2/PhysicsAnalysis/AnalysisCommon/TruthClassification

following classification of truth origin for a non-prompt lepton

- *Unknown or KnownUnknown*: leptons with insufficient truth-level information to be classified by the tool.
- *IsoElectron*: electrons originate either from the hard scatter or a boson decay. These electrons are treated as prompts in signal and background control regions.
- *ChargeFlipIsoElectron*: electrons whose charge is mismeasured at detector level and is classified as a non-prompt.
- *PromptMuon*: muons originate from either the hard scatter or a boson decay. These muons are treated as prompts for signal and background control regions.
- *PromptPhotonConversion*: non-prompt electrons originating from photon conversion.
- *TauDecay*: leptons originating from tau decays are treated as prompt leptons.
- *BHadronDecay*: leptons originating from hadrons containing a b-quark. These types of leptons are one of the primary sources of non-prompt leptons.
- *CHadronDecay*: leptons originating from hadrons containing a c-quark.
- *LightFlavourDecay*: leptons originating from mesons and lighter hadrons.

Figure 33 shows the origin of all leptons that are part of the quadruplet in the events with a signal quadruplet and a dijet. Most of the leptons in these regions are prompt and predominantly originate from $gg \to ZZjj$, $qq \to ZZjj$, and $EWKqq \to ZZjj$ processes. The leptons are classified *Unknown/KnownUnknown* due to insufficient truth information and mainly originate from $ttZ(\to \ell\ell)$ and VVV processes. The event record lacks information on the intermediary bosons for these samples, thus failing to identify the lepton origin. The *Unknown/KnownUnknown* leptons are treated as prompt leptons in the signal region. This treatment relies on the fact that ΔR between the *Unknown/KnownUnknown* classified truth leptons and reconstruction level lepton is observed to be close to 0. The *Unknown/KnownUnknown* classified leptons are treated as non-prompt leptons in the background control regions.

Figure 34 shows the fraction of non-prompt electrons (left) and non-prompt muons (bottom) in the events with a signal quadruplet and a dijet. The non-prompt leptons originating from b-hadrons or c-hadrons are collectively called *heavy flavor (HF)* non-prompt leptons. All other non-prompt leptons are categorized as *light flavor (LF)* non-prompt leptons. About 50% of non-prompt electrons in the signal region originate from heavy flavor sources, whereas more than 90% of non-prompt muons originate from the heavy flavor decays.

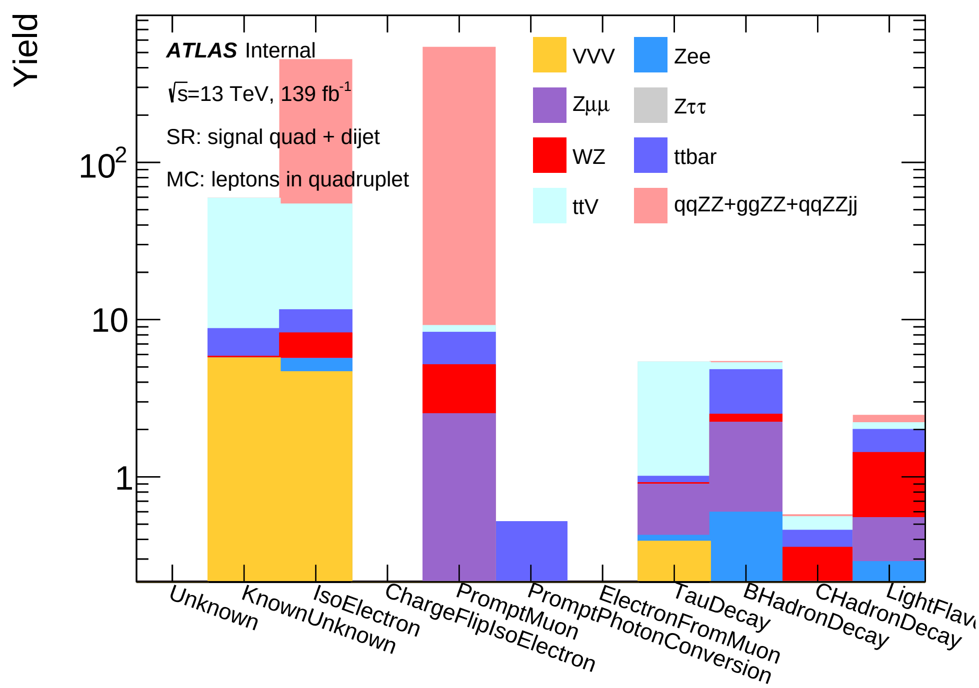


Figure 33: Origins of leptons in the signal region in events with a quadruplet and a dijet. The lepton origin is classified by the IFF classifier tool. Only leptons that are part of the signal quadruplet are shown. remake plots with label and larger y-axis

16.1.2 Control Regions

The fake factors are measured from data in a fake enriched background control region formed by combining two independent control regions, the $Z + jets$ control region and the $t\bar{t}$ control

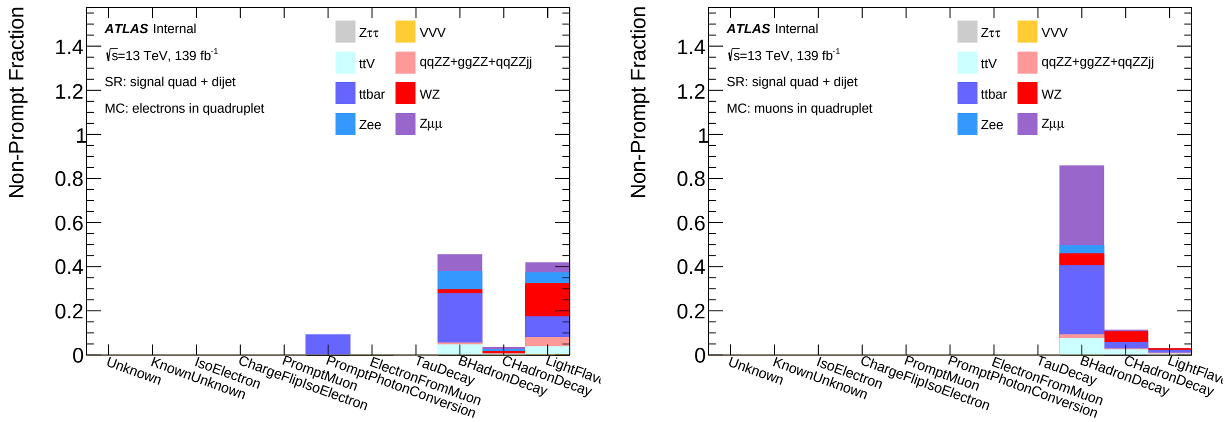
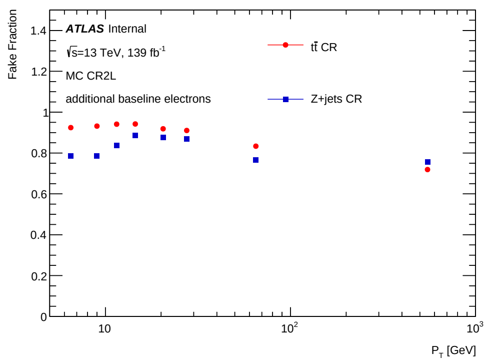


Figure 34: Origins of non-prompt leptons in the signal region in events with a signal quadruplet and a dijet. The events are normalized to the number of non-prompt electrons (left) and non-prompt muons (right). remake plots w/wo ATLAS label

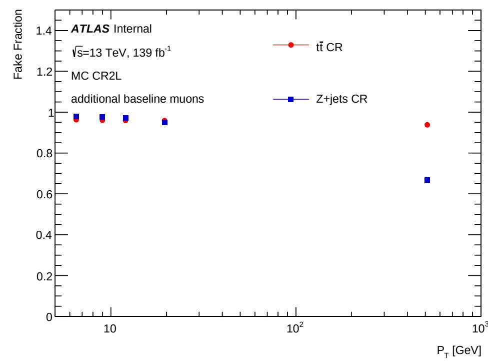
region. Events in the control regions consist of a prompt lepton pair from a physics process and additional leptons from non-prompt sources. Both control regions use a single or dilepton trigger similar to the signal region and require the leading and sub-leading leptons in an event to satisfy $p_{T,\ leading\ lepton} > 20$ GeV and $p_{T,\ sub-leading\ lepton} > 15$ GeV. An event in the $Z + jets$ CR consists of an SF-OC prompt-lepton pair from the Z boson decay with an invariant mass of 76 $GeV < m_{\ell\ell} < 106$ GeV, and additional leptons. Additionally, no events can have missing transverse energy higher than 50 GeV to suppress the contamination from the WZ process.

Similarly, the $t\bar{t}$ CR consists of events with different flavor prompt-lepton pairs and additional leptons. An event in the $t\bar{t}$ CR requires at least one b-tagged jet to reduce the WZ contamination. The b-tagging in the $t\bar{t}$ CR is performed by a flavor tagging tool described in Ref [77].

Figure 35 shows the fractions of the additional baseline electrons (left) and muons (right) that originate from a non-prompt source as a function of their p_T in the $Z + jets$ CR (blue) and the $t\bar{t}$ CR (red). A high fraction ($\geq 80\%$) of baseline electrons originate from non-prompt sources in both $Z + jets$ CR and $t\bar{t}$ CR. More than 95% of the low-p_T baseline muons are from non-prompt sources in both control regions. These distributions show that most of the additional leptons in either control region are expected to be from non-prompt sources, thus, motivating the control regions to evaluate the fake factors.



(a) Fake-fraction of baseline electrons.



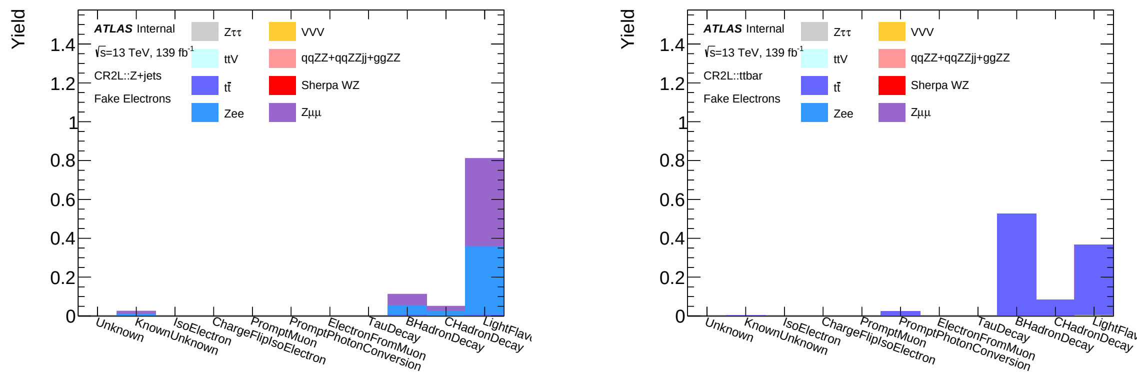
(b) Fake-fraction of baseline muons.

Figure 35: Fraction of non-prompt electrons and muons in the $Z + jets$ and $t\bar{t}$ control regions.remake plots w/wo ATLAS label

The control regions have a unique non-prompt lepton composition as shown by figures 36 and 37. More than 80% of the non-prompt electrons in the $Z + jets$ CR originate from the light flavor decays, but about 60% are from the light flavor decays in the $t\bar{t}$ CR. Similarly, about 80% of the non-prompt muons in the $Z + jets$ CR originate from the heavy flavor, whereas more than 90% are from the heavy-flavor decays in the $t\bar{t}$ CR. The non-prompt compositions of the signal region shown in figure 34 are different from either control region. The two independent control regions are combined to form a single control region with a similar non-prompt lepton composition as the signal region.

The b-jet requirement applied to suppress the prompt-lepton contamination from the WZ process in $t\bar{t}$ CR ensures the presence of at least one jet in all events. Therefore, events without jets in the combined control region only contain the $Z + jets$ $n_{jet} = 0$ events. The two control regions are first weighted and combined for the events with the jets to match the heavy flavor composition of the $n_{jet} > 0$ events in the signal region. The combination weights are evaluated by solving the following equation:

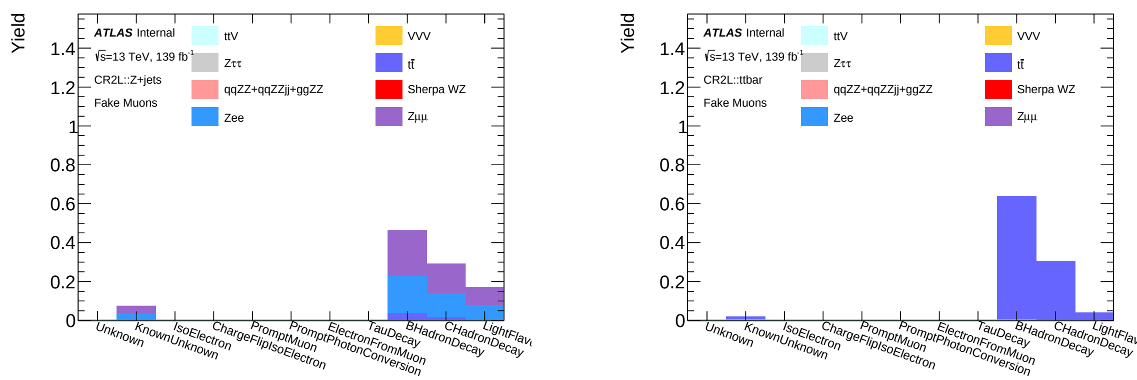
$$\frac{\{w \times N_{Z+jets} \times f_{HF,Z+jets}\} + \{(1-w) \times N_{t\bar{t}} \times f_{HF,t\bar{t}}\}}{\{w \times N_{Z+jets} + (1-w) \times N_{t\bar{t}}\}} = f_{HF,SR} \tag{16.1}$$



(a) Non-prompt electrons in $Z + jets$ CR.

(b) Non-prompt electrons in $t\bar{t}$ CR.

Figure 36: Sources of non-prompt electrons in background control regions. Fake composition is unique in these control regions. remake plots w/wo ATLAS label



(a) Non-prompt muons in $Z + jets$ CR.

(b) Non-prompt muons in $t\bar{t}$ CR.

Figure 37: Sources of non-prompt muons in $Z + jets$(left) and $t\bar{t}$(right) control regions. remake plots w/wo ATLAS label

where N is the total yield in the control region, f_{HF} is the ratio of the non-prompt leptons from heavy-flavor decays to total non-prompt leptons, and w is the combination weight to be determined.

As the composition of non-prompt electrons and muons are different in different regions, the weights are evaluated separately for electrons and muons and evaluated as $w_\mu = 0.26$ and $w_e = 0.06$. Figure 38 shows the composition of the non-prompt electrons and muons in the combined control region, which is formed by a weighted combination of the $Z + jets$ CR and the $t\bar{t}$ CR.

Figure 39 shows the distributions of additional baseline electrons as a function of their

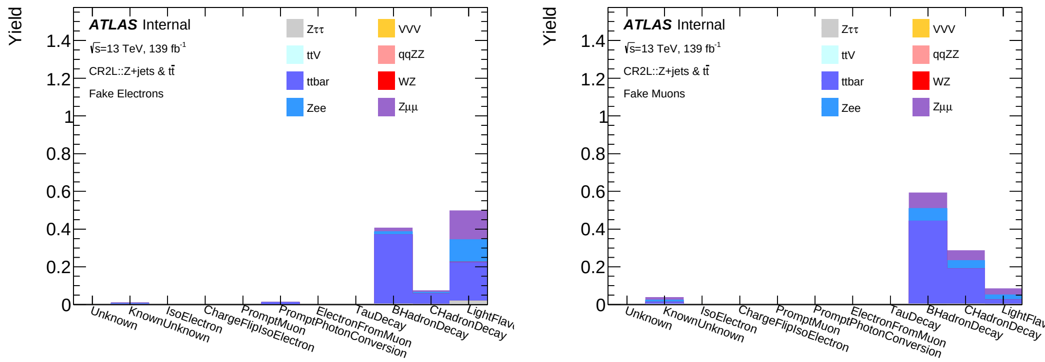


Figure 38: Origins of non-prompt electrons (left) and muons (right) in the combined control region. remake plots w/wo ATLAS label

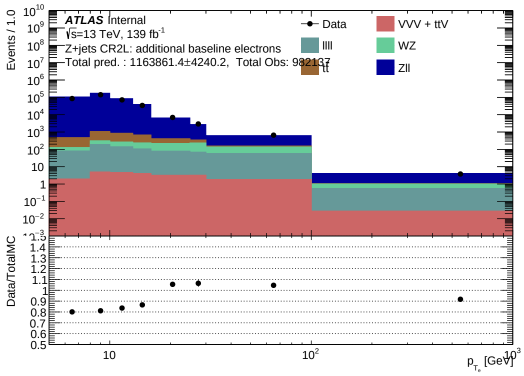
p_T in the $Z + jets$ CR (left) and $t\bar{t}$ CR (right). The bottom distribution shows the same for the combined control region. For $Z + jets$ CR at low p_T region, additional baseline electrons are overestimated in MC by about 20% showing the limited precision of the MC to estimate the non-prompt leptons. Similarly, figure 40 shows the distributions of additional baseline muons as a function of their p_T in the three control regions. In $Z + jets$ CR, additional muons mainly originate from $Z \to \ell\ell$ process in low p_T region, whereas at high p_T contribution from $t\bar{t}$ and WZ is more significant.

16.1.3 Fake Factor Strategy

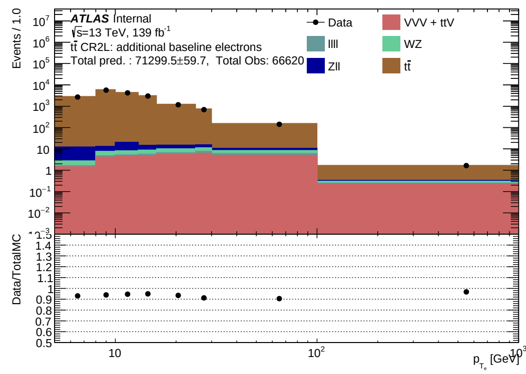
The centrally developed *fake factor tool* by ATLAS IFF is used to estimate the fake background [78]. The tool takes the ratio of signal to baseline leptons, i.e., *fake efficiency (f)*, calculated in the combined control region as an input where,

$$f = \frac{N_{non-prompt\ signal\ leptons}}{N_{non-prompt\ baseline\ leptons}} \tag{16.2}$$

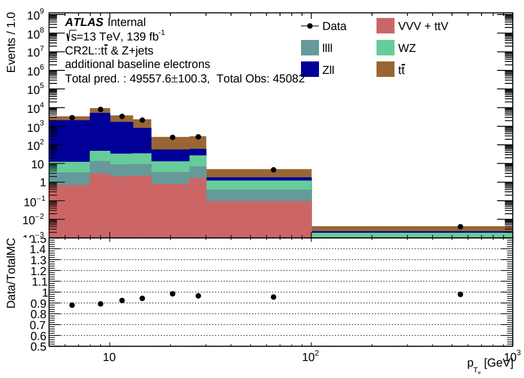
For a simple case of a signal region with one signal lepton, the observed signal lepton (N^T) and baseline-anti-signal lepton (N^L) can be estimated in terms of the number of prompt or real baseline leptons (N^B_R) and the number of non-prompt or fake baseline leptons (N^B_F) as



(a) $Z + jets$ CR



(b) $t\bar{t}$ CR



(c) combined control region

Figure 39: Additional baseline electrons as a function of p_T in control regions. remake plots w/wo ATLAS label

$$N^T = rN_R^B + f_F^B \tag{16.3}$$

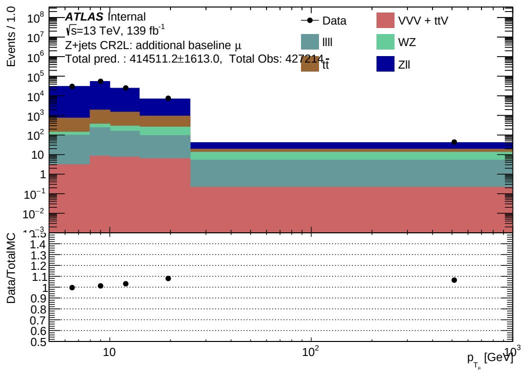
and

$$N^L = (1-r)N_R^B + (1-f)N_F^B \tag{16.4}$$

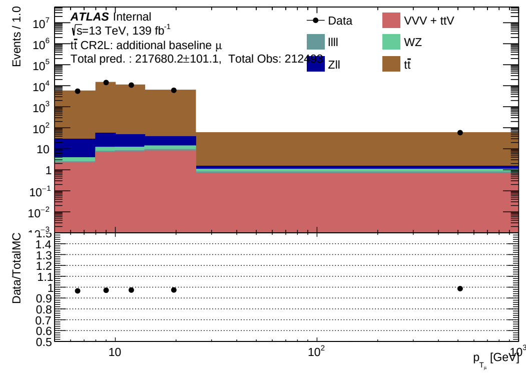
where, r is the *real efficiency* such that,

$$r = \frac{N_{prompt\ signal\ leptons}}{N_{prompt\ baseline\ leptons}} \tag{16.5}$$

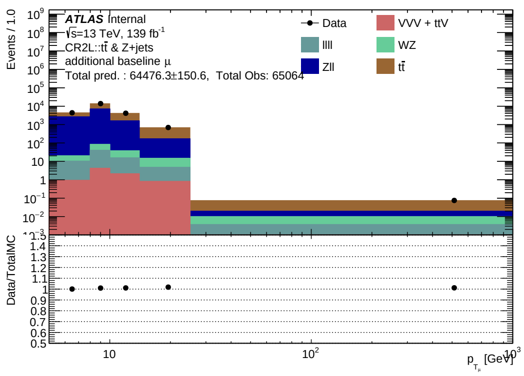
Equations 16.3 and 16.4 can be written as a 2×2 matrix equation as



(a) $Z + jets$ CR



(b) $t\bar{t}$ CR



(c) combined control region

Figure 40: Additional baseline muons as a function of p_T in control regions. remake plots w/wo ATLAS label

$$\begin{pmatrix} N^T \\ N^L \end{pmatrix} = \begin{pmatrix} r & f \\ 1-r & 1-f \end{pmatrix} \begin{pmatrix} N_R^B \\ N_F^B \end{pmatrix} \tag{16.6}$$

The number of non-prompt baseline contributions is estimated by ignoring the higher-order term of the fake efficiency as

$$N_F^B = \frac{1}{r-f}[r(N^T + N^L) - N^T] \tag{16.7}$$

Therefore, the predicted number of non-prompt signal leptons is

$$N_F^T = \frac{f}{r-f}[r(N^T + N^L) - N^T] \tag{16.8}$$

The fake factor method assumes the $r \to 1$ limit, which simplifies equation 16.8. However, since the real efficiency of any measurement is less than one, this approximation overestimates the fake background. To account for this overestimation, the number of genuine baseline-anti-signal prompt leptons (N_R^L) are measured in MC and subtracted to get the final background yield as,

$$N_F^T = \frac{f}{1-f}[N^L - N_R^L] \tag{16.9}$$

The method makes a typically safe assumption that the real anti-signal prompt leptons are modeled precisely in MC. The coefficient F is the fake factor where,

$$F = \frac{f}{1-f} \tag{16.10}$$

As the fake efficiency f is estimated from data in the combined control region, the fake factor background estimation method does not rely on any efficiencies or yield in the tight signal region.

This method can be extended to the four-lepton signal region where there are four baseline leptons, of which one or more could be non-prompt. Corresponding to the permutation of individual leptons to be either signal or baseline-anti-signal, there are $2^4 = 16$ $\{N^{TTTT}$, N^{TTTL}, N^{TTLL},...N^{LLLL} $\}$ observations to consider. The analysis considers N^{TTTT} the signal region; therefore, the background is estimated from the quadruplets with at least one baseline-anti-signal lepton.

16.1.4 Fake Efficiency & Systematics

Fake efficiency (f), defined in previous Section 16.1.3, is evaluated from the combined control region using the total number of additional leptons from data as

$$f = \frac{N_{Data}^{Signal} - N_{MC}^{Prompt\ Signal}}{N_{Data}^{Baseline} - N_{MC}^{Prompt\ Baseline}} \tag{16.11}$$

Since some additional leptons could originate from prompt sources, such contributions are estimated from MC and subtracted as shown in equation 16.11.

Figures 35, 39, and 40 show that the fake-fraction and the total yield of the additional leptons are dependent on their transverse momentum p_T . Therefore, the fake efficiency evaluated using equation 16.11 depends on the lepton p_T . Because of the low resolution of the detector in forward regions, a higher number of non-prompt leptons are expected; thus, the fake efficiency depends on the leptons' pseudorapidity η. Additionally, since the non-prompt leptons predominantly originate from jets, the fake efficiency also depends on the number of jets n_{jets} in an event.

Figures 41 and 42 show the fake efficiencies for electrons and muons respectively as a function of p_T (top-left), η (top-right) and n_{jets} (bottom-center). Since high-p_T leptons are most likely to originate from a prompt source, fake efficiency typically decreases as a function of p_T for leptons. The dependency on η is most likely from lower detector resolution causing a higher number of misidentifications and lower efficiency for TTVA.

As discussed in Section 11.4, the lepton-favored overlap removal used in the analysis rejects jets if they overlap with leptons. Due to the $b - jet$ requirement in $t\bar{t}$ CR, the $njet = 0$ events only consist of contributions from the $Z + jets$ CR, which does not have an explicit event requirement on the number of jets. The probability of non-prompt leptons passing the isolation requirement is higher in events with no jets or surrounding hadronic activity. Therefore, as observed, a higher fake efficiency is expected in events without jets.

The fake efficiency is parametrized in three-dimensional distributions of p_T , η, and n_{jets}. Only two bins ($n_{jet} = 0$ & $n_{jet} > 0$) are used for number of jets. The distributions in figure 43 show the fake efficiency of an electron as a function of p_T & η for $n_{jet} = 0$ bin (left) and for $n_{jet} > 0$ bin (right). Similar distributions are shown in figure 44 for muons.

The fake efficiency distributions' binomial errors are propagated as the statistical uncertainties on the fake estimate. The subtracted prompt component of equation 16.11 is estimated using MC predictions. As discussed in Section 3, the prediction relies on the

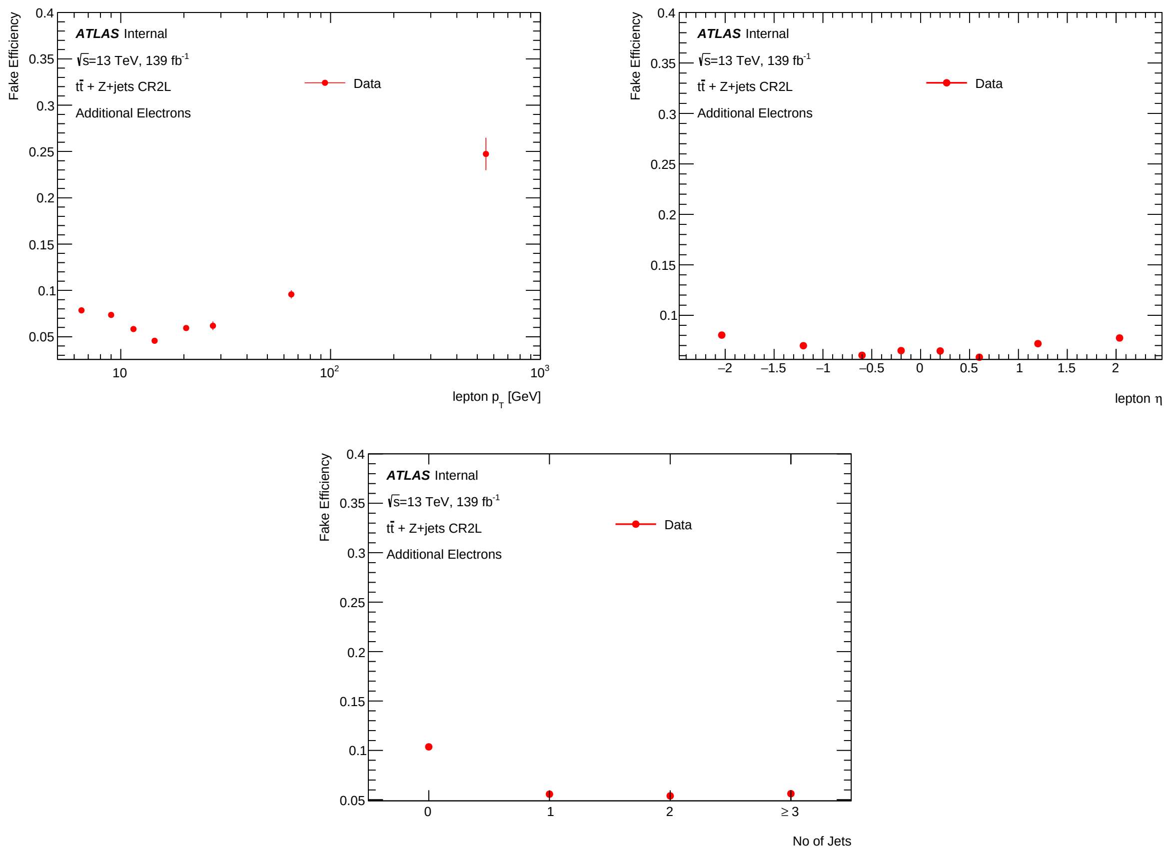


Figure 41: Fake efficiency of fake electrons measured in the combined control region from data as a function of its p_T , η, and n_{jets}. remake plots w/wo ATLAS label and change color

PDF, the energy-dependent QCD factorization and renormalization scale, and the strong coupling constant (α_S). Therefore, the theory uncertainties on these three parameters are propagated as systematic uncertainties of the fake efficiency.

For each theory uncertainty, a variation-applied fake efficiency is evaluated by separately varying the numerator and denominator of the fake efficiency equation 16.11. The difference between the variation-applied fake efficiency and the nominal fake efficiency is considered systematic uncertainty. Figures 45 and 46 show the statistical and systematic uncertainties on the fake efficiency calculated in the combined control region. For both electrons and muons, the statistical uncertainty is dominant.

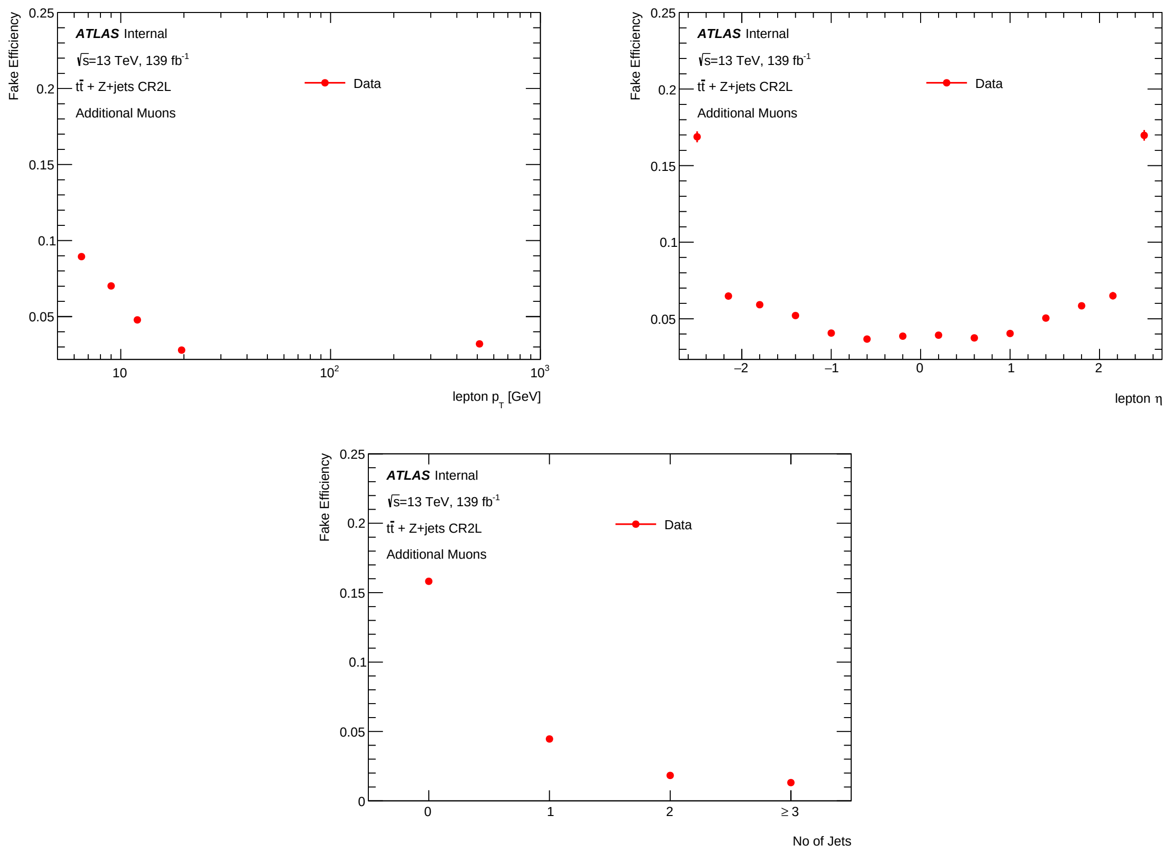


Figure 42: Fake efficiency of fake muons measured in the combined control region from data as a function of its p_T , η, and n_{jets}. remake plots w/wo ATLAS label and change color

16.1.5 Method Validation

Before implementing the fake-factor method to estimate the fake background in the signal region, the method is validated in two separate validation regions

1. Different-flavor validation region (VRDF): one pair in the quadruplet must have two different-flavor leptons.

2. Same-charge validation region (VRSC): one pair in the quadruplet must have two same-charge leptons.

The low statistics in both regions result from requiring one of the pairs to be either same-charge or different-flavor. Therefore, events in the validation regions only have a signal

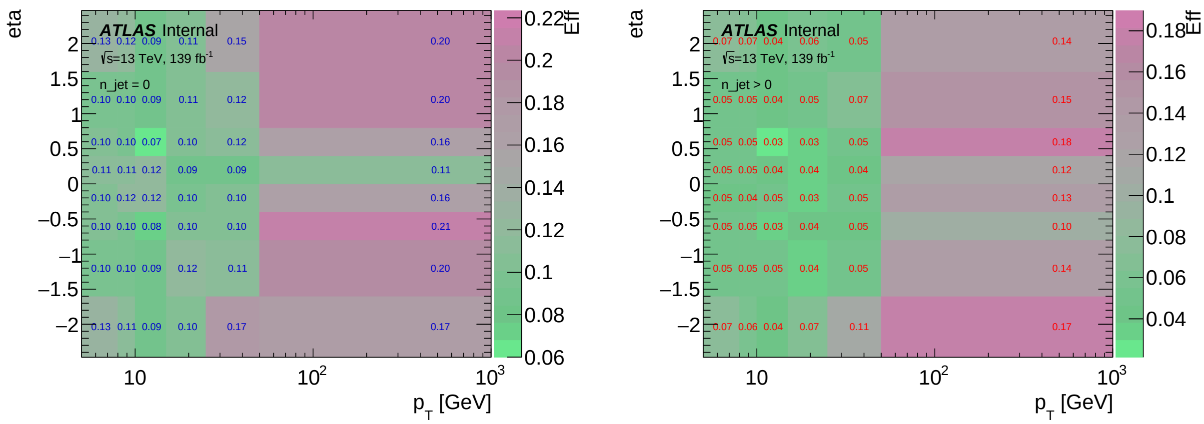


Figure 43: Fake efficiency of fake electrons measured in the combined control region from data as a function of its p_T , and η in two slices of n_{jets} ($n_{jet} = 0$ left) and ($n_{jet} > 0$ right). remake plots w/wo ATLAS label, change the color of text for second plot and y-label

quadruplet without any dijet requirement. The validation regions' quadruplets are formed by requiring the same kinematic criteria as that of the signal region discussed in Section 13. The trigger requirement, object selection, and overlap removal are identical to the signal region. Additionally, events in the VRDF are required not to have any b-tagged jet to reduce the contribution from ttZ processes. Reducing the ttZ component further reduces the significant modeling uncertainties related to the ttZ process.

By constructing either a same-charge or a different-flavor pair, the event yield in validation regions is dominated by events where at least one lepton in the quadruplet is a non-prompt-signal lepton known as the fake background in the signal region. The events also originate from other physics processes, such as $qqZZ$, $qqZZjj$, $ggZZ$, ttZ, and VVV whose contribution is predicted by the same MC generators as that of the signal region.

Figures 47 show the non-prompt composition in the different flavor validation region (left) and same-charge validation region (right). The non-prompt compositions in the two validation regions are different from that of the signal region or the background control regions composition as shown in figures 34, 36 and 37. Therefore, to validate the fake background estimation strategy, it is imperative to observe a good correspondence between data and a combination of the MC prediction with the fake background yield in both validation regions.

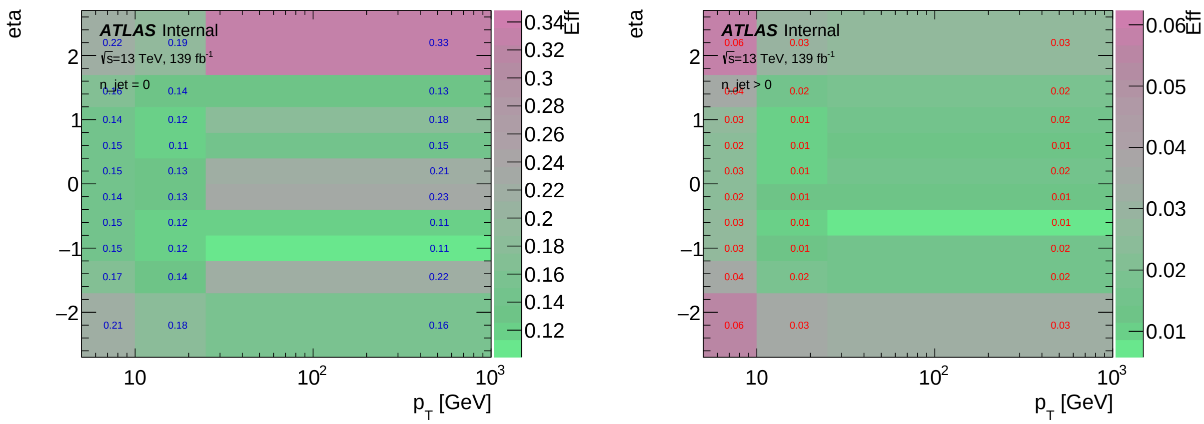


Figure 44: Fake efficiency of fake muons measured in the combined control region from data as a function of its p_T , and η in two slices of n_{jets} ($n_{jets} = 0$ left) and ($n_{jet} > 0$ right). remake plots w/wo ATLAS label, change the color of text for second plot and y-label

The fake backgrounds for the validation regions are estimated by applying the fake factor to each baseline-anti-signal leptons in the not-signal quadruplet, as discussed in Section 16.1.3. Figure 48a shows the data and the predicted MC yield in VRDF as a function of $m_{4\ell}$ where the fake backgrounds are estimated from $Z+jets$, $t\bar{t}$, and WZ MC predictions. Figure 48b shows the same but the reducible estimated using the fake factor method. Similarly, figures 48c and 48d show the yields as a function of $m_{4\ell}$ in the same charge validation region. Both regions have similar characteristics, and the fake background dominates the event yield with some contribution from other physics processes.

The systematic gray bands in figures 48b and 48d include the systematic and statistical uncertainties from the fake factor method, as well as the uncertainties on PDF, QCD scale, and strong coupling (α_S) on the $qqZZ$, $qqZZjj$ & $ggZZ$ MC prediction. The bands also include the uncertainties in the cross-section measurements of the ttZ and VVV processes. The treatment of the systematic theoretical uncertainties is the same as that of the signal region and will be discussed in Section 18.1. Other experimental uncertainties related to the lepton reconstruction and identification, trigger, and luminosity discussed in Section 18.2 are assumed to be negligible for the validation regions. For most bins, the data and MC yield are compatible with both regions' systematic and statistical uncertainties. Moreover,

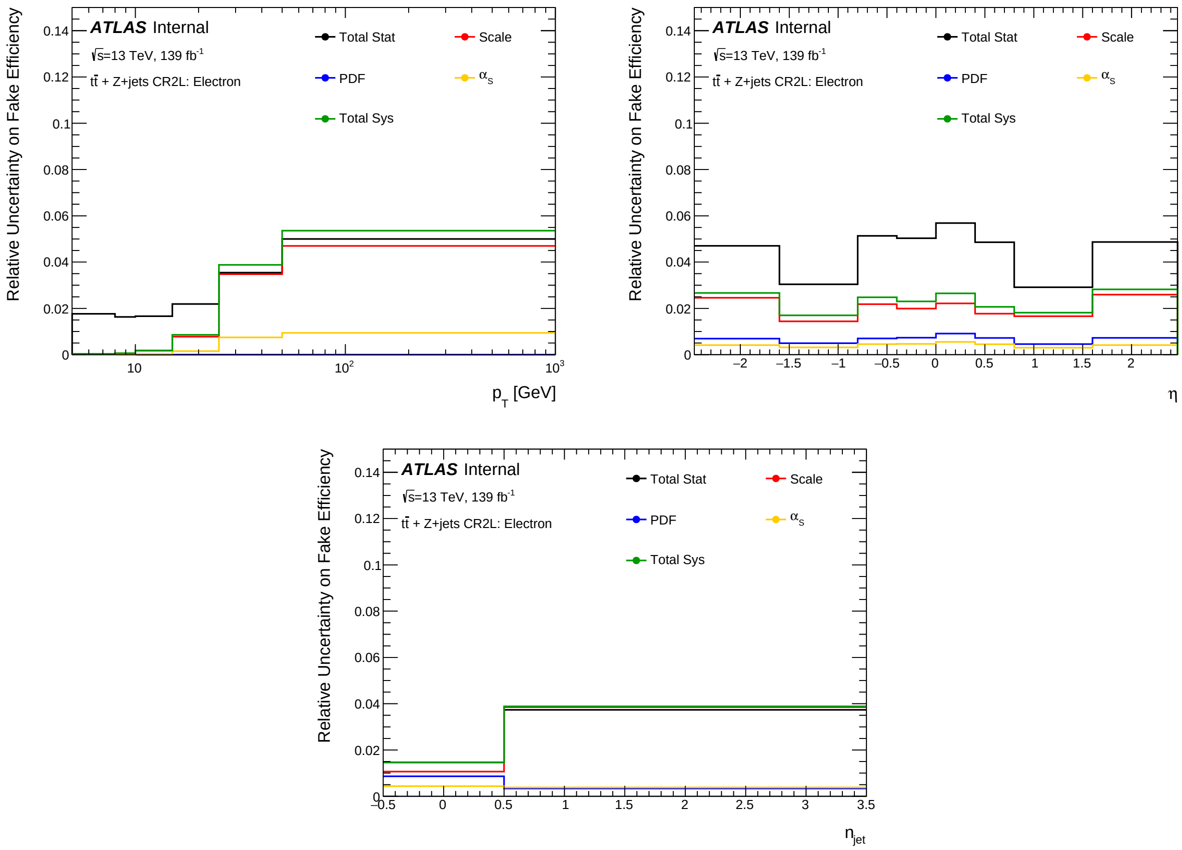


Figure 45: Uncertainties on the fake efficiency of the fake electrons measured in the combined control region from data as a function of its p_T , η, and n_{jets}. remake plots w/wo ATLAS label

the agreement between data and MC simulation is better when the reducible events are estimated using the fake factor method, thus, fully validating the method.

The data and MC yield comparisons for several kinematic observables in VRDF (left) and VRSC (right) are shown by distributions in figure 49. The data and MC prediction are compatible in most bins within the systematic uncertainties for all the observables.

16.1.6 Signal Region Estimation

Similar to the validation regions, the background in the signal region is estimated by applying the fake factor to the not-signal quadruplets, as discussed in section 16.1.3. Distributions in figure 50 compare the fake background predicted from MC and estimated from the fake-

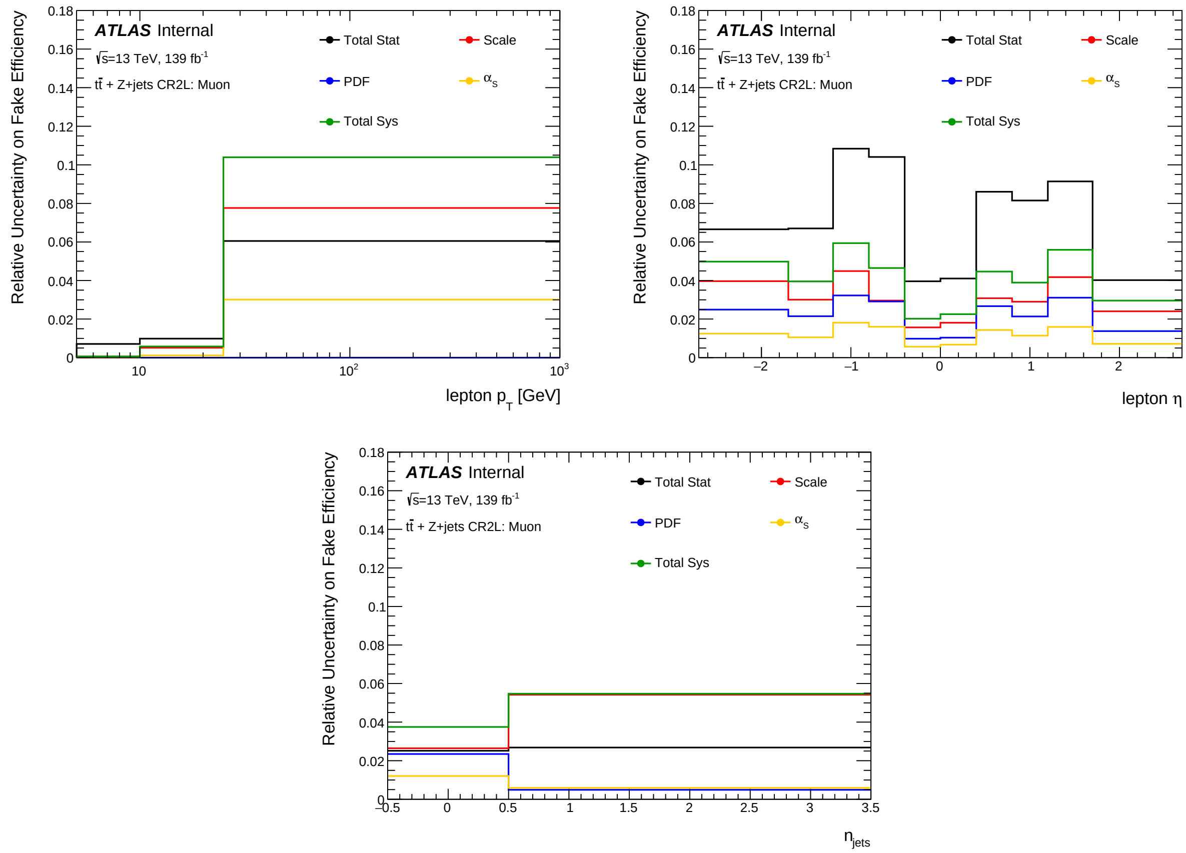
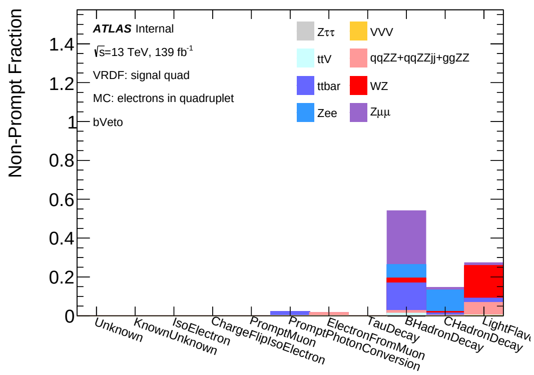
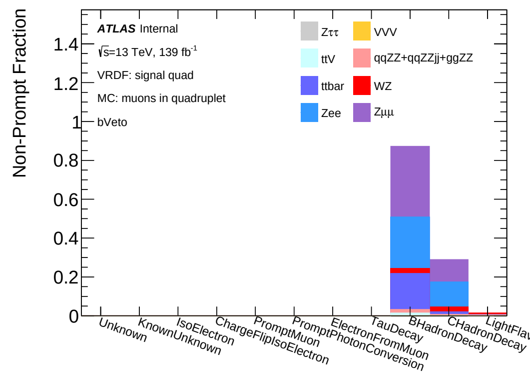


Figure 46: Uncertainties on the fake efficiency of the fake muons measured in the combined control region from data as a function of its p_T , η, and n_{jets}. remake plots w/wo ATLAS label

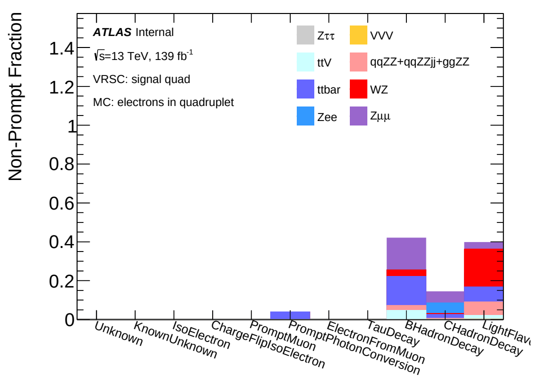
factor method in the VBS-Enhanced signal events as a function of $m_{4\ell}$ (left) and $p_{\mathrm{T},4\ell}$ (right). Figure 51 shows identical distributions but also includes the total SM prediction in the same region. The lower panel of the plot shows the fake background to the predicted signal ratio, which is small. The gray bands in the plots are from systematic uncertainties of the fake factor method, whose effect is negligible on the overall yield of the signal region.



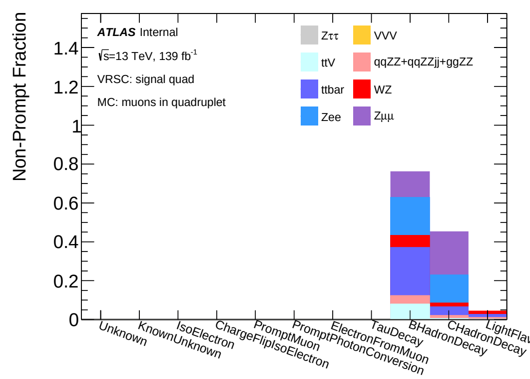
(a) Non-prompt electrons in VRDF.



(b) Non-prompt muons in VRDF.

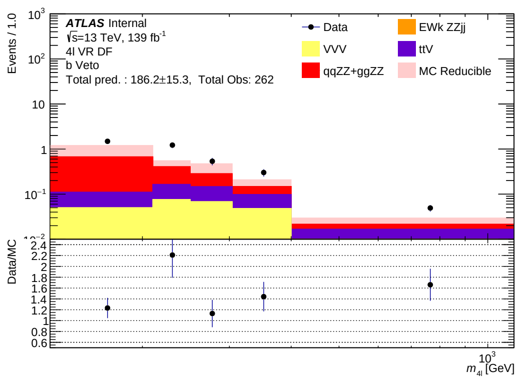


(c) Non-prompt electrons in VRSC.

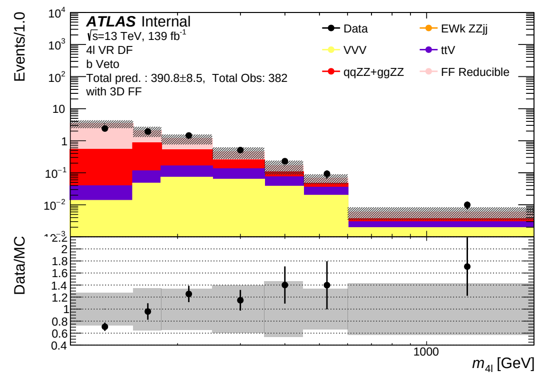


(d) Non-prompt muons in VRSC.

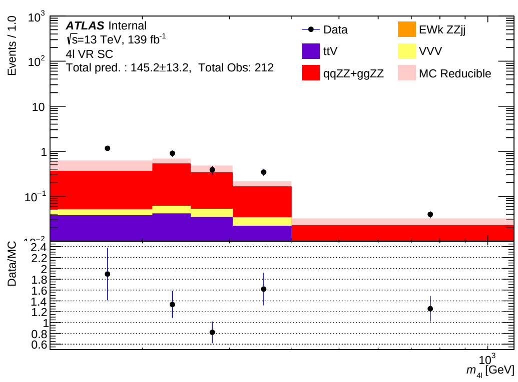
Figure 47: Sources of non-prompt electrons and muons in the different flavors and the same charge validation regions.remake plots with ATLAS Label



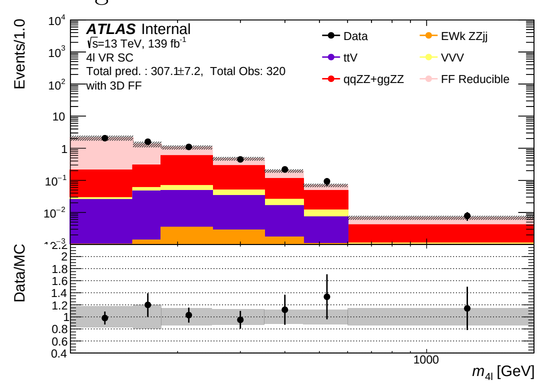
(a) VRDF: MC predicted fake background.



(b) VRDF: data-driven fake factor estimate of fake background.

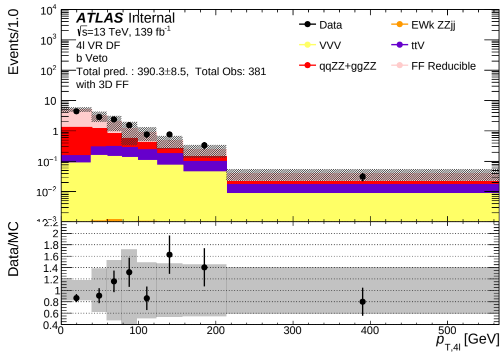


(c) VRSC: MC predicted fake background.

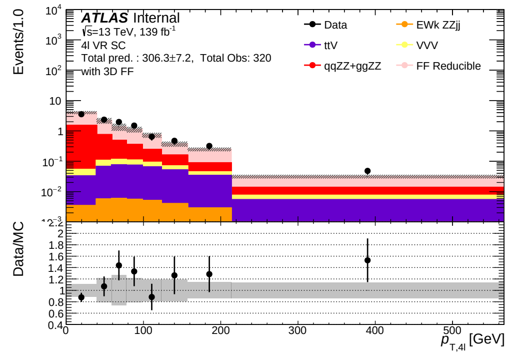


(d) VRSC: data-driven fake factor estimate of fake background.

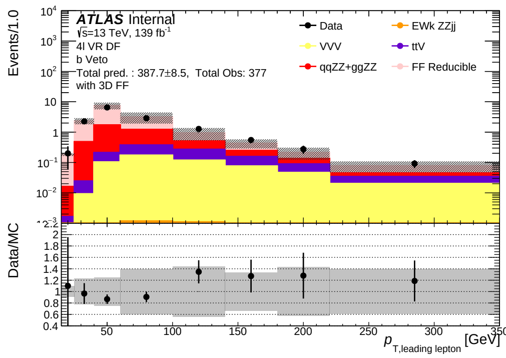
Figure 48: Yield as a function of $m_{4\ell}$ in the different flavor (top) and same charge (bottom) validation regions. In both regions, the MC prediction matches more closely with data when the fake background events are estimated using the data-driven fake-factor method.remake plots with ATLAS Label and cleaning other labels



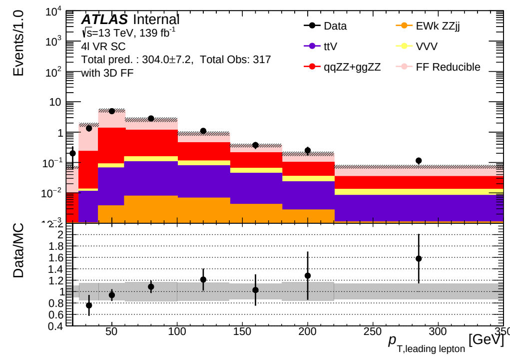
(a) VRDF: $p_{T,4\ell}$



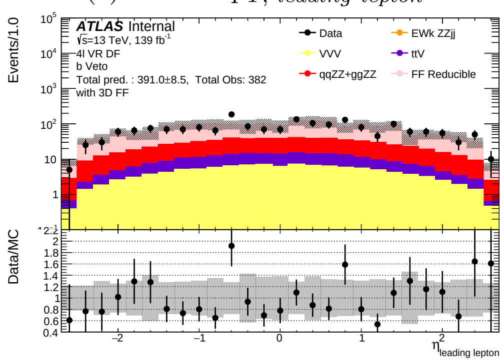
(b) VRSC: $p_{T,4\ell}$



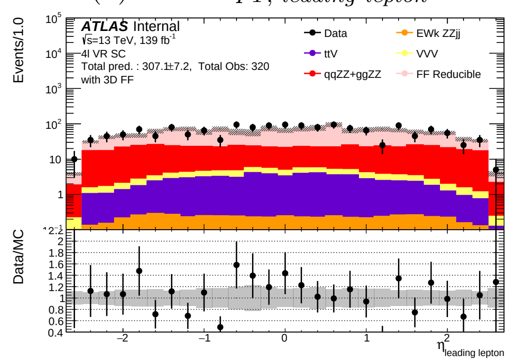
(c) VRDF: $p_{T,\ leading\ lepton}$



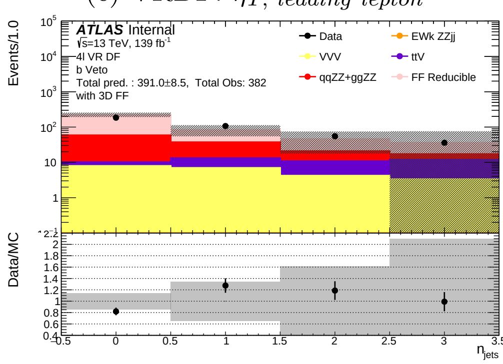
(d) VRSC: $p_{T,\ leading\ lepton}$



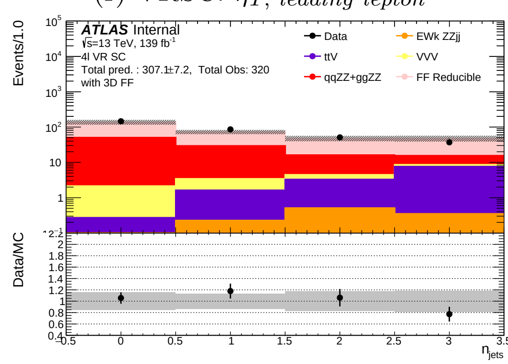
(e) VRDF: $\eta_{T,\ leading\ lepton}$



(f) VRSC: $\eta_{T,\ leading\ lepton}$



(g) VRDF: n_{jets}



(h) VRSC: n_{jets}

Figure 49: Data and MC yield comparison for different flavor validation regions (left) and same charge validation region (right) as a function of several kinematic observables. remake plots with ATLAS Label and cleaning other labels

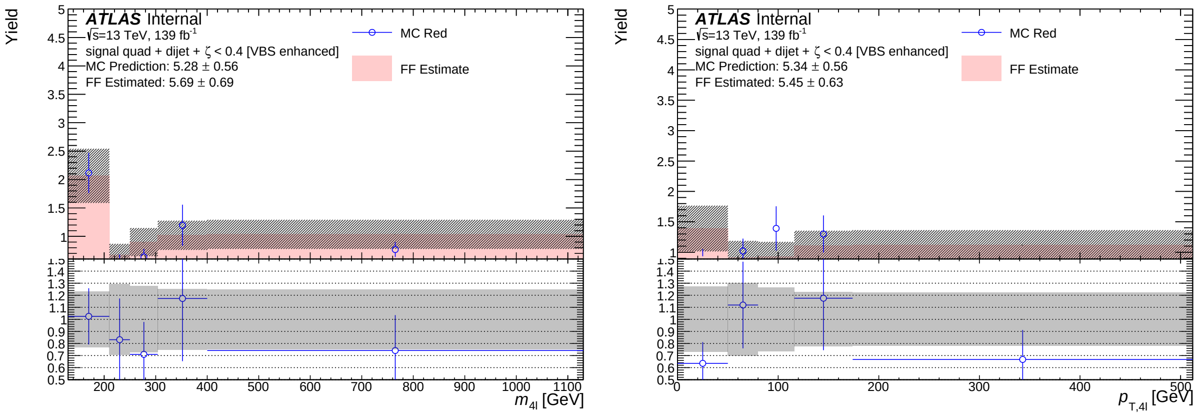


Figure 50: MC prediction and fake-factor method estimate of the fake background as a function of $m_{4\ell}$(left) and $p_{\mathrm{T},4\ell}$ (right) in the VBS-Enhanced region. Black bands represent the systematic uncertainties from the fake factor method. remake plots with ATLAS Label and cleaning other labels

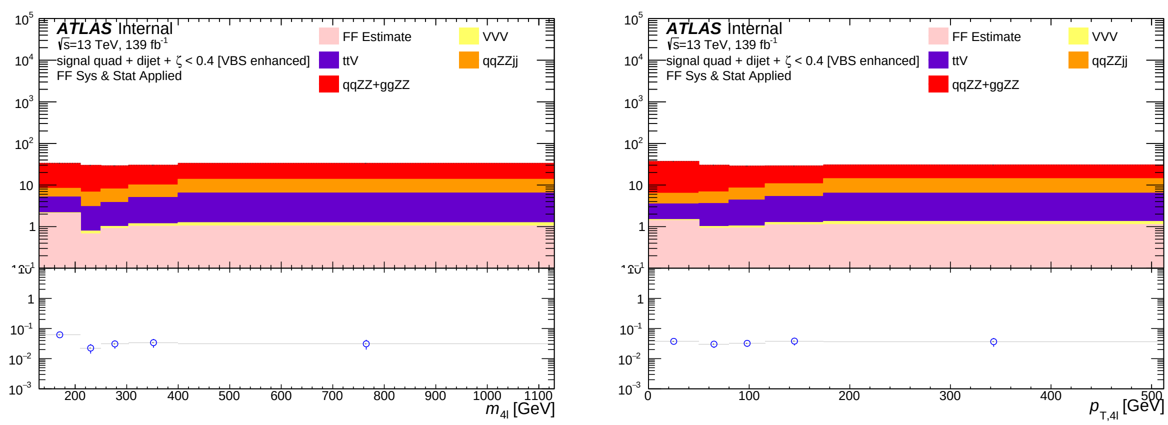


Figure 51: SM prediction and fake background estimated from the fake-factor method as a function of $m_{4\ell}$(left) and $p_{\mathrm{T},4\ell}$ (right) in the VBS-Enhanced region. Black bands represent the systematic uncertainties from the fake factor method, which are negligible on the full signal region distribution. remake plots with ATLAS Label, cleaning other label and y-xis/ratio-axis title

17 Unfolding

The main results of the thesis are differential cross-section measurements at the particle level. The inclusive detector level cross-section for a given physics process $p_1p_2 \to X$ is,

$$\sigma_{p_1p_2\to X}^{detector\ level} = A \times \epsilon \times \sigma_{p_1p_2\to X}^{particle-level} \tag{17.1}$$

where $\sigma_{p_1p_2\to X}^{particle-level}$ is the *true* cross-section of the physics process predicted by the theory. The physical layout of the ATLAS detector does not cover all areas of the phase space. A accounts for the limited acceptance of the ATLAS detector. Several parts of the detector have several reconstruction efficiencies, which are accounted for by the factor ϵ. The detector level cross-section is measured experimentally in terms of the number of particles in a given final state (N) and integrated Luminosity L as $\sigma_{p_1p_2\to X}^{detector\ level} = \frac{N}{L}$. The *true* particle level inclusive cross-section can be estimated by correcting for detector acceptance and detector efficiency for the measured cross-section $\sigma_{p_1p_2\to X}^{detector\ level}$.

For differential cross-sections where the cross-section is measured in different bins of the kinematic observables, additional correction is needed to correct the resolution-induced migration between nearby bins.

This Chapter discusses the unfolding technique in detail. Section 17.1 gives an Overview on the unfolding algorithm, whereas Section 17.3 validates the unfolding method. Section 17.4 discusses the bias from unfolding and the attempts to optimize the bias.

17.1 Method Overview

The analysis uses an *iterative Bayesian unfolding* algorithm based on Baye's theorem [79] [80] using ATLAS-supported *RooUnfold* package [81]. Bayes' theorem formulates a mathematical relation to obtain a probability of an effect E caused by several independent causes C_i, given the initial probability of the causes $P(C_i)$ and the conditional probability of the $i-th$ cause to produce the effect $P(E|C_i)$ as,

$$P(C_i|E) = \frac{P(E|C_i).P(C_i)}{\sum_j P(E|C_j).P(C_j)} \tag{17.2}$$

The obtained probability depends on the prior probability of the cause and the conditional probability of cause and effect. The prior dependency is reduced by using an iterative technique, where the outcome of the previous iteration is used as a prior for the subsequent iteration.

For a single iteration, the algorithm can be summarized as,

$$U_i = \frac{1}{\epsilon_i} \times \sum_{j}^{reco\ bins} (R_j - F_j).f_i.\frac{M_{ji}T_i}{\sum_k^{truth\ bins} M_{jk}T_k} \tag{17.3}$$

where U_i is the unfolded yield in the target bin i, T_i is the predicted truth level yield in particle bin i, R_j is the observed detector level yield in reco bin j and F_j is the subtracted detector level reducible background yield. M_{ij} is the migration matrix element from particle level bin j to detector level bin i.

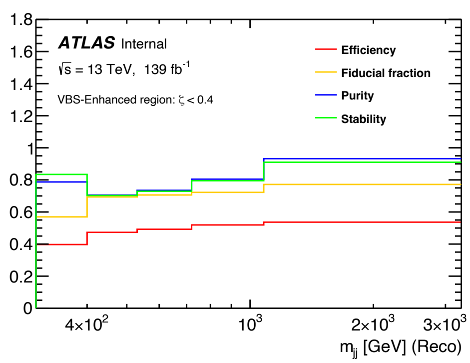
Based on the discussion, conceptually, three corrections from the SM MC prediction need to be applied to estimate the unfolded yield. The three unfolding inputs are

- ***Reconstruction efficiency (ϵ):*** The reconstruction efficiency accounts for the limited acceptance and efficiency of the detector. Technically, it is defined as a fraction of events that pass both detector and fiducial level selection to the events passing only the fiducial level selection.

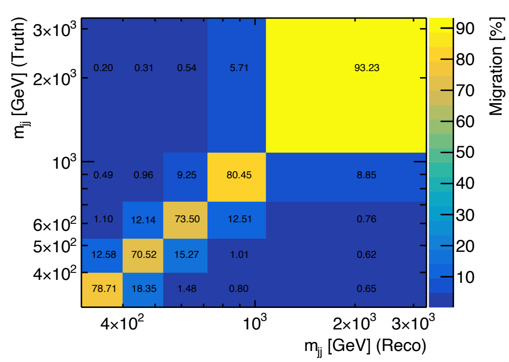
- ***Fiducial fraction (f):*** The fiducial fraction accounts for events that are outside the fiducial region at the particle level, which due to limited detector resolution entered in the measured distribution. An example of such an event would be a signal $4\ell + jj$ event where one of the jets originates from pile-up instead of hard-scatter. Technically, it is defined as a fraction of events that pass both detector and fiducial level selection to the events passing only the detector level selection.

- ***Migration matrix*** (M_{ij})***:*** The migration matrix is a two-dimensional matrix that accounts for events migrated from particle level bin j to detector level bin i. The migration matrix corrects the probability of bin migration. It is measured in MC by comparing particle and detector levels distributions for events that pass both fiducial and detector-level selections. Bin migrations result from resolution effects and smearing of the reconstructed distributions. The diagonal component of the migration matrix is related to the *fiducial purity*, which corresponds to the fraction of detector-level events that originate from the same bin at the particle level.

Figure 52 show all three unfolding inputs along with the purity as a function of m_{jj}. The reconstruction efficiency is less than 50% caused by the poor jet reconstruction efficiency. The fiducial fraction and purity is smaller in lower bins of m_{jj}, which mainly corresponds to contribution from pileup jets faking the event selection. The normalized migration matrix shown in the second plot with the particle level prediction in $y-axis$ and the detector level prediction in $x-axis$ is diagonal.



(a) reconstruction efficiency (red), fiducial fraction (yellow) and purity (blue).



(b) migration matrix

Figure 52: Unfolding inputs from SM MC as a function of m_{jj}.remake first plot with ATLAS Label and stability

17.2 Binning for Unfolding

Choosing optimal binning to perform the unfolding procedure for all kinematic observables effectively is imperative. Two factors drive the choice of binning; first, the necessity to have large enough bin statistics to maintain the Gaussian approximation while preserving the shape of the differential distributions, and second, the necessity to minimize large bin migrations and statistical uncertainties from unfolding. Therefore, each bin must have at least 15 events in the VBS-Suppressed region and at least 20 events in the VBS-Enhanced signal region.

To maintain a good performance of the unfolding, each bin for the kinematic observable has at least 70% purity except for $p_{T,4\ell jj}$ where at least 50% purity is required. Moreover, for each observable, every bin width must be equal to or greater than the resolution of the same bin. The resolution in each particle-level bin is evaluated from MC by comparing the difference of particle and detector level yield for events that pass both fiducial- and detector-level event selection. The difference is fitted using Gaussian approximation, and twice the resulting standard deviation is taken as the resolution. Table 15 shows the final bin choices for all the kinematic observables used in differential cross-section measurement. .

17.3 Method Validation

The unfolding method is validated using three different tests.

17.3.1 MC Closure Test

The first validation of the unfolding technique is with the SM MC. An SM-predicted detector level distribution for a kinematic observable is unfolded using the unfolding inputs from the same MC. Figure 53 shows an example of the MC-based closure test for m_{jj} in the VBS-Enhanced region. The blue detector-level MC prediction is unfolded using the inputs from the same MC, and the resulting black unfolded distribution is compared with the red particle-

Table 15: Binning for all unfolded observables in VBS-Enhanced and suppressed regions.

Observable	Region	Binning
m_{jj} [GeV]	VBS-Enhanced VBS-Suppressed	[300, 400, 530, 720, 1080, 3280] [300, 410, 600, 178]
$m_{4\ell}$ [GeV]	VBS-Enhanced VBS-Suppressed	[130, 210, 250, 304, 400, 1130] [130, 226, 304, 752]
$p_{T,4\ell}$ [GeV]	VBS-Enhanced VBS-Suppressed	[0, 50, 80, 116, 174, 512] [0, 76, 140, 424]
$p_{T,jj}$ [GeV]	VBS-Enhanced VBS-Suppressed	[0, 52, 82, 116, 172, 524] [0, 80, 146, 448]
$p_{T,4\ell jj}$ [GeV]	VBS-Enhanced VBS-Suppressed	[0, 20, 42, 64, 298] [0, 36, 70, 254]
$s_{T,4\ell jj}$ [GeV]	VBS-Enhanced VBS-Suppressed	[70, 240, 320, 420, 580, 1410] [70, 330, 500, 1210]
$\|\Delta y_{jj}\|$	VBS-Enhanced VBS-Suppressed	[2, 3.08, 3.74, 4.32, 5.06, 7.4] [2, 2.94, 3.78, 5.4]
$\Delta\phi_{jj}^{signed}$	VBS-Enhanced VBS-Suppressed	[$-\pi$, -2.1, 0, 2.1, π] [$-\pi$,0,π]
$cos\theta^*_{\ell i\ell j}$	VBS-Enhanced VBS-Suppressed	[-1, -0.5, 0, 0.5, 1] [-1, 0, 1]
ζ	VBS-Enhanced VBS-Suppressed	[0, 0.06, 0.12, 0.18, 0.26, 0.4] [0.4, 0.5, 0.64, 1.02]

level prediction. Since both detector-level prediction and unfolding inputs are from the same MC, a perfect closure between the unfolded and particle-level distribution is observed.

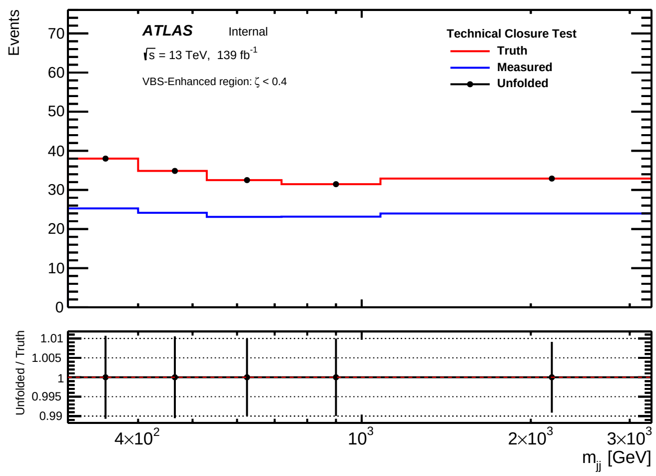
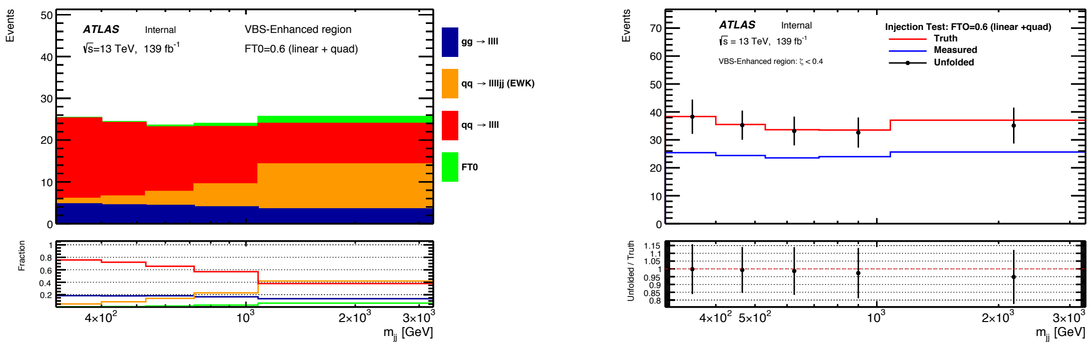


Figure 53: MC technical closure test of the unfolding procedure for m_{jj}. The detector-level MC distribution (in blue) is unfolded with the nominal SM unfolding inputs and compared to the particle-level distribution (in red) from the same MC. A perfect closure between unfolded and particle level distribution is observed

17.3.2 Injection Test

The analysis uses a model-independent EFT approach discussed in Section 21 to constrain the effect of BSM physics. Therefore, it is essential to test the ability of the unfolding algorithm to uncover the accurate particle-level prediction from data containing BSM physics via injection test. In an injection test, a BSM physics contribution is added to the SM detector-level prediction, unfolded with the nominal SM unfolding inputs, and compared with the BSM-added particle-level distribution. Figure 54a shows an injection test for m_{jj} in the VBS-Enhanced region where a BSM contribution (green distribution) is added to the SM MC. The BSM contribution is from linear and quadratic contributions of an $FT0$ EFT operator. Figure 54b shows the result of the injection test. The BSM-added detector-level MC prediction (blue) is unfolded (black) using nominal SM MC unfolding inputs and compared against the BSM-added particle-level distribution (red). A small non-closure of about 5% in the last bin of m_{jj} is observed, which is well within the uncertainties of the unfolded distribution.

Note to self: perhaps it makes sense to discuss EFT theory motivation in theory section?



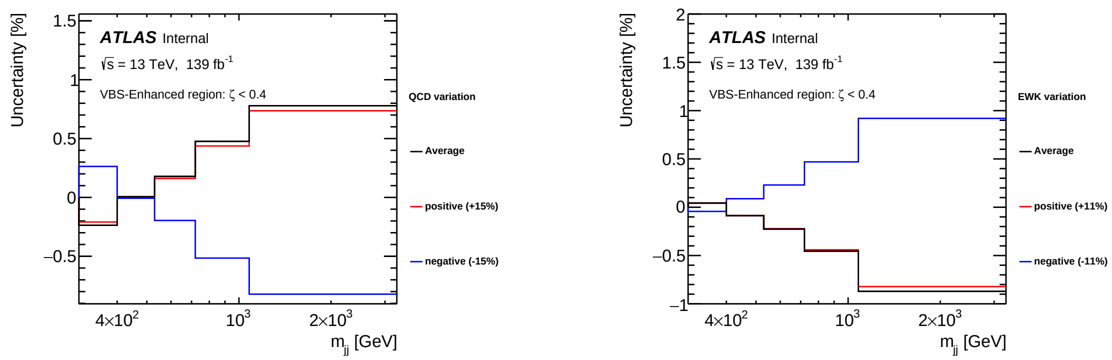
(a) Detector level MC prediction with contribution from dimension−8 $FT0$ EFT operator.

(b) Unfolded SM+EFT MC detector-level distribution with response matrix from SM MC.

Figure 54: Injection test with dimension−8 $FT0$ EFT operator. remake plots with ATLAS Label

17.3.3 Physics Variation

FFrom the previous ATLAS electroweak $ZZjj$ analysis, a slight enhancement on the central value of the EWk $ZZjj$ cross-section was measured [17]. The final unfolding validation tested the ability of the algorithm to recover the actual shape of particle-level distribution if a physics process cross-section was different from the SM prediction. First, as shown by figure 55a, the cross-section for parton-initiated QCD $qqZZjj$ is varied by a factor equal to the total statistical uncertainty on data in the VBS-Suppressed region ±15%. The varied detector-level distribution is then unfolded using the nominal SM MC unfolding inputs and compared with the varied fiducial level prediction. Figure 55b shows the same test where the $EWKqqZZjj$ cross-section is varied by ±11% based on the enhanced cross-section observed in the previous measurement. In both cases, a non-closure of about 1% is observed, well below the uncertainties from unfolding.



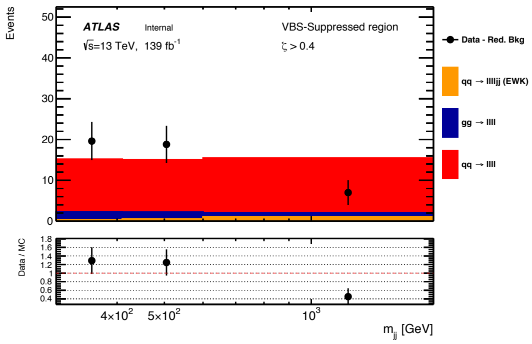
(a) QCD cross-section is varied by $\pm 15\%$ (b) EWK cross-section is varied by $\pm 11\%$

Figure 55: Unfolding validation using physics variation where parton-initiated QCD (left) or the EWK process cross-sections are varied.

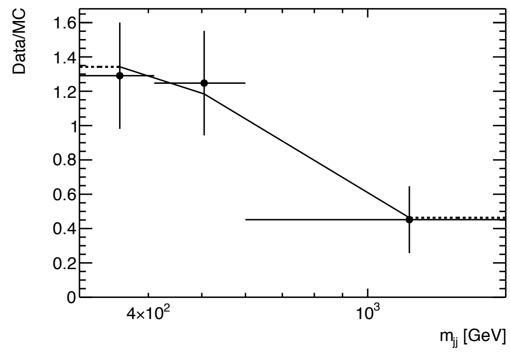
17.4 Bias and Optimization

The unfolded procedure relies on a prior value depending on the SM MC which naturally biases the unfolded cross-sections. With each iteration of unfolding, the algorithm improves the knowledge of the prior, thus, reducing the unfolding bias. However, with increasing number of iterations, the repeated bin migrations amplifies the statistical fluctuations in data, resulting in larger values of statistical uncertainties. Therefore, a finite number of iteration is chosen and the resulting unfolding bias is taken as the systematic uncertainty for the measurement.

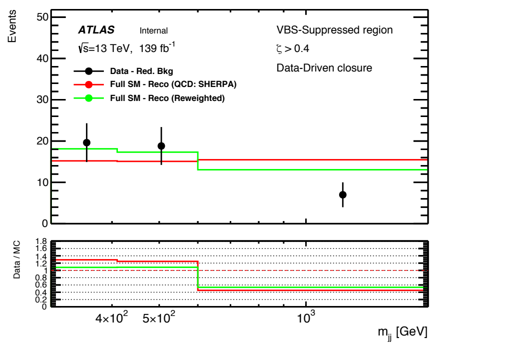
The unfolding bias is evaluated by the *data-driven closure test*, where a pseudo dataset is developed utilizing the ratio of observed data and SM-predicted detector-level yield. First, for each observable the data and MC ratio is smoothed using Friedman's Super Smoother technique [82], fixing the end points to the value of ratio in the first and last bins. A reweighing function for each observable is developed to reweigh the SM fiducial- and detector-level yields. The reweighed detector-level signal-yield is then unfolded with the nominal unfolding inputs from SM and compared with the reweighed fiducial-level yield to get the final unfolding bias. Figure 56 shows step-by-step procedure for the data-driven closure test. As shown by the ratio panel of figure 56d, unfolding bias of order 10% is observed.



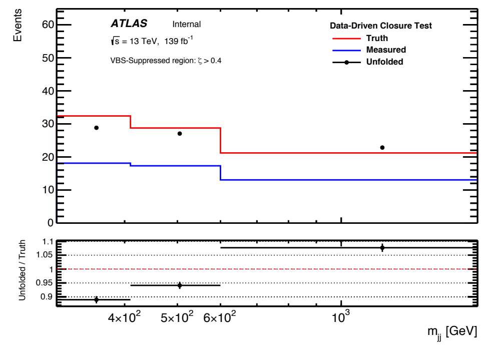
(a) Data and MC for m_{jj}



(b) Smoothed ratio of Data and MC.



(c) Nominal SM (red) detector level yield and reweighted-detector level yield(green).



(d) Unfolding bias.

Figure 56: A step-by-step overview of the data driven closure test to get the unfolding bias. remake plots with ATLAS Label

The bias observed in figure 56d is obtained by using one number of iteration for unfolding. With a goal to reduce the unfolding bias, the data-driven closure test was repeated for several number of iterations. The resulting unfolding bias and systematic uncertainties up to 4 iterations are shown in figure 57. As expected the unfolding bias decreases whereas the statistical uncertainty increases with the higher number of iteration. To balance between the statistical uncertainty and bias uncertainty, one number of iteration is chosen as optimal choice for the measurement.

Unfolding bias is the largest source of the systematic uncertainty of the analysis and is studied in detail using a MC-driven toy studies to understand the source. The observed large bias is from detector-level pileup jets at lower p_T or higher η that are not part of the fiducial phase space. The jet-vertex-tagger and forward-jet-vertex-tagger has lower efficiency to select

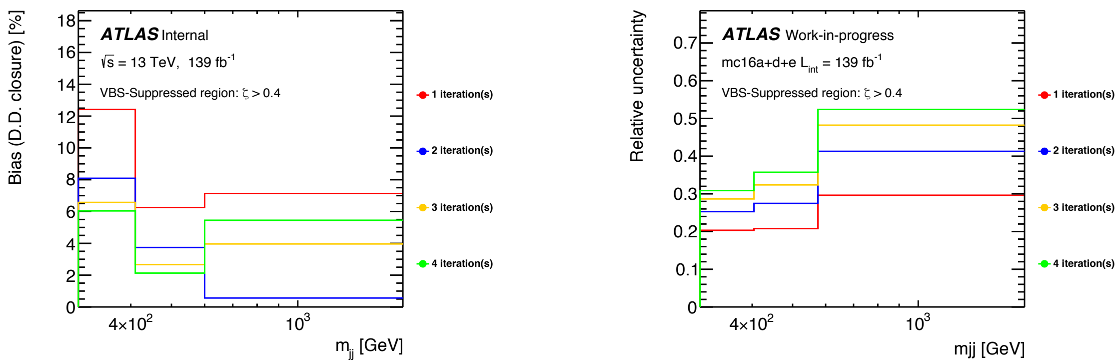


Figure 57: Unfolding bias (left) and statistical uncertainty (right) with up to 4 unfolding iterations as a function of m_{jj} in VBS-Suppressed region.

the hard scattering jets at lower p_T or higher η, thus resulting in more *fiducial-fake-event* contamination. The additional MC-based studies on the unfolding bias are summarized in Appendix B.

18 Uncertainties on the Measurement

The differential cross-section measurements discussed in this thesis are affected by three sources of systematic uncertainties, experimental sources, theoretical sources, and intrinsic systematics related to the unfolding process. The statistical uncertainty of the measurements is dominant as data statistics limit the cross-section measurements. This section discusses the source of theoretical, experimental, unfolding uncertainties and propagation of the statistical uncertainties to the unfolded cross-sections.

18.1 Theoretical Uncertainties

The following sources of theoretical uncertainties are considered in the measurement.

- **Uncertainties on QCD Scale:** As discussed in Section 3, the theoretical predictions of cross-sections depend on the factorization scale (μ_F) and renormalization scale (μ_R) [83]. To account for this dependence, a QCD scale uncertainty is evaluated by scaling μ_F and μ_R independently using on-the-fly variations provided by the MC generators. The variations constitute of six-point variations of μ_F and μ_R from -50% or $+100\%$ around their nominal values of 1, such that $\{\mu_R = 0.5, \mu_F = 0.5\}$, $\{\mu_R = 0.5, \mu_F = 1.0\}$, $\{\mu_R = 1.0, \mu_F = 0.5\}$, $\{\mu_R = 1.0, \mu_F = 2.0\}$, $\{\mu_R = 2.0, \mu_F = 1.0\}$, and $\{\mu_R = 2.0, \mu_F = 2.0\}$. The final uncertainty is evaluated as the absolute envelope of the six variations. The QCD scale uncertainties are evaluated for $qqZZ$, $ggZZ$, and $EWKqqZZjj$ samples.

- **Uncertainties on PDF & α_S:** The cross-sections also depend on the choice of the PDF used by the MC generators. Thus, the PDF uncertainty for Sherpa and MadGraph5 samples that use NNPDF3.0NNLO is evaluated using the prescription described in Ref [69] using on-the-fly variation weights. The PDF variations include a set of 100 internal variations, two additional variations from the nominal PDF reweighted to the alternative MMHT2014nnlo [84] & CT14nnlo [85] PDF sets, and variations of the strong coupling constant by ± 0.0001 where the nominal value of α_S is 0.118. The total uncertainty is taken as the absolute envelope of all standard deviations of 100 internal variations, the two alternate PDF variations, added in quadrature with the envelope of the α_S variations,

$$\sigma_{PDF}^{NNPDF3.0NNLO} = \sqrt{[max(\sigma_{std.\ dev.\ int.}, |\sigma_{MMHT2014nnlo}|, |\sigma_{CT14nnlo}|\ )]^2 + \sigma_{\alpha_S}^2} \quad (18.1)$$

The PDF uncertainty is evaluated for $qqZZ$, $ggZZ$ and MadGraph5 $EWKqqZZjj$ samples.

The electroweak EWK $qqZZjj$ samples generated by POWHEG-V2 do not have on-the-fly variations to evaluate the PDF uncertainty. Therefore, PDF uncertainty from the MadGraph5 sample is taken as the PDF uncertainty for POWHEG-V2 EWK $qqZZjj$ samples.

- **Uncertainties on $gg \to ZZ^{(*)}$ NLO Corrections:** The uncertainty is related to the NLO QCD k-factor applied to the $ggZZ$ sample [86]. The NLO QCD k-factors applied are evaluated differentially as a function of the $m_{4\ell}$.

- $t\bar{t}V$ & VVV **cross-sections:** The experimental uncertainties on recently published cross-section measurements of the ttV [87] and WZZ [88] processes by ATLAS are propagated for the analysis. On the entire ttV process, a flat conservative variation of 15% is applied, taken from the cross-section measurement of ttZ. Similarly, for VVV conservative 10% variation taken from the cross-section measurement of WWZ is applied to the whole VVV samples.

As shown above, the theoretical uncertainties are process specific and are evaluated separately for each MC sample. The theory uncertainties need to be propagated to the unfolded cross-section measurements. For each theory uncertainty, variation-applied particle and detector level yields are built by substituting the varied distribution for the selected process instead of the nominal one. The variation-applied detector level yield is unfolded using the unfolding inputs from nominal SM predictions. The difference between the unfolded result to the variation-applied truth MC yields gives systematic uncertainty for each variation. In general, the theoretical variations significantly affect the predicted particle-level and detector-level yields; however, they have a negligible impact on the shape of the distributions. Since the variation is applied to both detector and particle level yields, the resulting uncertainties from theory systematics on the unfolded cross-sections are small.

18.2 Experimental Uncertainties

The experimental uncertainties arise from the measurement of the energy and momentum scales of the reconstructed objects and the uncertainties on object reconstruction, identification, and selection efficiencies. The following category summarizes the sources of experimental uncertainties,

Jet Related Uncertainties: The analysis requires two jets in the final state. Therefore, jet reconstruction and selection uncertainties are the measurement's most significant source of systematic experimental uncertainties.

- **Jet Reconstruction Uncertainty:** The jet-related uncertainties associated with reconstruction and different steps of calibration discussed in Section 7.5 are provided by ATLAS-supported tool *JetUncertainties*[4]. The tool provides several configurations for jet-related uncertainties adjusted to the various needs of several analyses. The measurement in this thesis uses *GlobalReduction_FullJER* configuration with a total of 36 uncertainties, each with upward and downward components, corresponding to 36×2 variations, 20×2 variations are related to JES, and 13×2 to JER. 6×2 of the 36×2 variations are related to the η inter-calibration procedure, 4×2 to the pile-up energy subtraction procedure, and 8×2 to the in-situ calibration of jets. Additional

[4]https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetUncertainties

1×2 variations arise separately from the flavor composition, flavor response, a single particle response at high p_T, and possible punch-through effects.

- **JVT & fJVT Uncertainties:** Additional sets of jet uncertainties (1×2) arising from the efficiencies of jet vertex selections, JVT, and fJVT cut requirements are also considered in the analysis.

An envelope of the 13 JER uncertainty added in quadrature to an envelope of each of the other sources gives the final impact of jet-related uncertainties.

Lepton Related Uncertainties: The following categories define the lepton-related uncertainties in the analysis

- **Electron Efficiencies:** The electron efficiency uncertainty consists of uncertainties on the trigger, identification, reconstruction, and isolation efficiencies of electrons. These uncertainties are provided by an ATLAS-supported tool *ElectronEffciencyCorrection*[5]. There are a total of 61 nuisance parameters related to electron efficiencies, each with upward and downward components corresponding to 61×2 variations. 34×2 out of 61 is related to uncertainties in identification efficiency, 25×2 related to the reconstruction efficiencies, and a single nuisance parameter (1×2) each from the isolation efficiency and trigger efficiency scale factors.

- **Muon Efficiencies:** Similar to the electrons, muon efficiency uncertainty consists of variations on the trigger, identification, reconstruction, and isolation efficiencies of muons, which are provided by another ATLAS-supported tool *MuonEfficiencyCorrections*[6]. In total, there are 10×2 nuisance parameters, sets of two (2×2) variations corresponding to trigger efficiency scale factors, sets of four (4×2) related to the identification and reconstructed efficiency, two sets of two (2×2) each corresponding to the isolation efficiency and track-to-vertex association efficiency.

[5]https://gitlab.cern.ch/atlas/athena/-/tree/21.2/PhysicsAnalysis/ElectronPhotonID/ElectronEfficiencyCorrection
[6]21https://gitlab.cern.ch/atlas/athena/-/tree/21.2/PhysicsAnalysis/MuonID/MuonIDAnalysis/MuonEfficiencyCorrectio

- **Electron Scale & Resolution:** The electron scale and resolution uncertainty is accounted for by two sets of nuisance parameters corresponding 2×2 variations.

- **Muon Scale & Resolution:** For muons resolution and scale uncertainties, there are 5 sets of nuisance parameters, 2×2 corresponding to the muon momentum resolution as measured separately by the Inner Detector and the Muon Spectrometer. One set of nuisance parameters (1×2) corresponds to the uncertainties on the muon momentum scale, and two sets of 2×2 are associated with the uncertainties in Sagitta correction.

Other Experimental Uncertainties:

- **Pileup Reweighting:** As discussed in Section 14.3, the MC predictions are reweighted to match the pile-up profile of data. A single 1×2 nuisance parameter accounts for upward and downward variations in the factors used for pile-up reweighting.

- **Luminosity:** As discussed in Section 14.1, the uncertainty in the collected integrated luminosity of $139 fb^{-1}$ is $\pm 1.7\%$, which is applied as a flat variation to both particle and detector level yields.

The experimental uncertainties affect all detector-level MC predictions and the estimate of the fake backgrounds. The experimental uncertainties need to be propagated to the unfolded cross-sections. For each systematic variation, a detector-level signal ($qqZZ + ggZZ + EWK\ qqZZjj$) and background ($ttV + VVV$) distribution, a variation applied prediction is built. The variation is also applied to the fake background estimate. The variation-applied background MC and fake backgrounds are subtracted from the variation-applied total MC prediction and then unfolded using the unfolding inputs from the nominal SM prediction. The individual systematic uncertainty corresponds to the difference between the variation-applied and nominal unfolded distributions for each variation.

18.3 Unfolding Uncertainties

The following two uncertainties are intrinsic to the unfolding process itself and are included in the uncertainties for the unfolded differential cross-sections.

- **Unfolding Bias:** The inherent unfolding bias can be estimated either using the data-driven method discussed in Section 17.4 or using an MC-toy-based method. It is the most significant source of systematic uncertainty for the measurement. Both methods yield similar results. However, the MC-toy-based method is free from the influence of statistical fluctuation in data and thus is chosen to assign the systematic uncertainty on the intrinsic unfolding bias.

  First, ten thousand toys are constructed from the truth level MC distribution, independently fluctuating each bin of the truth distribution with a random number extracted from a Gaussian with mean zero and variance of N^{bin}_{truth}. A pseudo detector level dataset is created by *folding* using the nominal response from MC. Folding is a technique to create a detector level distribution from the MC by applying the detector's limited acceptance and inefficiencies, adding detector resolution-induced bin migration between the truth and detector level distributions, and including the contribution of the fiducial fakes. The pseudo dataset is then unfolded using the nominal unfolding inputs and plotted as the average bias of the toys as a function of the toy's truth-level yield. The final unfolding bias is the value where this distribution intersects with the 68.3% confidence interval of the toys' true value. Figure 58 shows the unfolding bias evaluated in each bin of m_{jj} in the VBS Enhanced region, estimated using the MC-toy-based approach. The unfolding bias ranges from $8 - 15\%$ in different bins, similar to what was observed using the data-driven method. 699

- **QCD $qqZZ$ Modeling Uncertainty:** There are known differences between different generators driven by differences in parton shower and hadronization. Therefore, the second source of unfolding systematics is required to account for the differences in the unfolding input modeling for the dominant $qqZZ$ process. To avoid double-counting of the unfolding method covered by the data-driven uncertainties, an alternative $qqZZ$ sample predicted by MadGraph5 is first reweighted to match the nominal-Sherpa lineshape. The relative difference in reweighted detector-level distribution is unfolded using the inputs from nominal-Sherpa and compared with the reweighted-MadGraph5 particle level distribution. The relative difference between these two distributions is taken as modeling systematic uncertainty. Figure 59 shows the estimation of the modeling uncertainty for m_{jj} in the VBS-Enhanced region. The ratio panel of the right plot shows the QCD modeling uncertainties, which range from $2 - 4\%$ varying in different bins.

18.4 Background Uncertainties

There are additional sources of uncertainties from the data-driven estimate of the fake background. The statistical and systematic uncertainties on the fake efficiency discussed in Section 16.1.4, estimated in the combined control region, are propagated to the final unfolded cross-section yield. First, the variation-applied fake background is calculated and subtracted from the nominal detector-level prediction for each variation. The subtracted altered distribution is then unfolded with nominal unfolding inputs. The difference between the altered-unfolded distribution and the nominal-unfolded distribution gives the impact of the background uncertainties on the unfolded cross-section measurements.

18.5 Statistical Uncertainties

The statistical uncertainty from the reconstructed data needs to be propagated to the estimated unfolded yield. Equation 17.3 gives the unfolded yield for a target bin i with a single iteration. As the background subtracted detector yields are filled event by event, the reconstruction distribution is uncorrelated. However, as shown by the equation, an unfolded yield in one single bin depends on all detector-level bins due to the resolution effects via

Bin m_{jj} [GeV]	[300, 400)	[400, 530)	[530, 720)	[720, 1080)	[1080, 3280)
QCD MC modelling	2.91	0.05	0.77	0.56	4.21
Jet	**7.5**	**7.6**	**9.1**	**9.1**	**9.0**
Trigger	0.13	0.14	0.13	0.45	0.53
Leptons	1.7	1.7	1.6	2.4	2.6
PRW	0.38	0.58	0.83	0.88	0.62
Theory ($qqZZ$)	2.7	2.3	2.6	2.0	0.74
Theory (EWK $qqZZjj$)	0.08	0.05	0.07	0.14	0.96
Theory ($ggZZ$)	0.19	0.10	0.13	0.13	0.51
MC Bkg. (ttV+VVV)	2.7	2.6	2.4	1.8	1.2
Fake Bkg. (stat + syst)	3.3	3.1	2.1	2.7	2.8
Luminosity	1.3	1.3	1.3	1.4	1.4
Unfolding Bias	**10.1**	**10.7**	**11.0**	**9.6**	**5.9**
Total	14.0	14.2	15.1	14.1	12.4

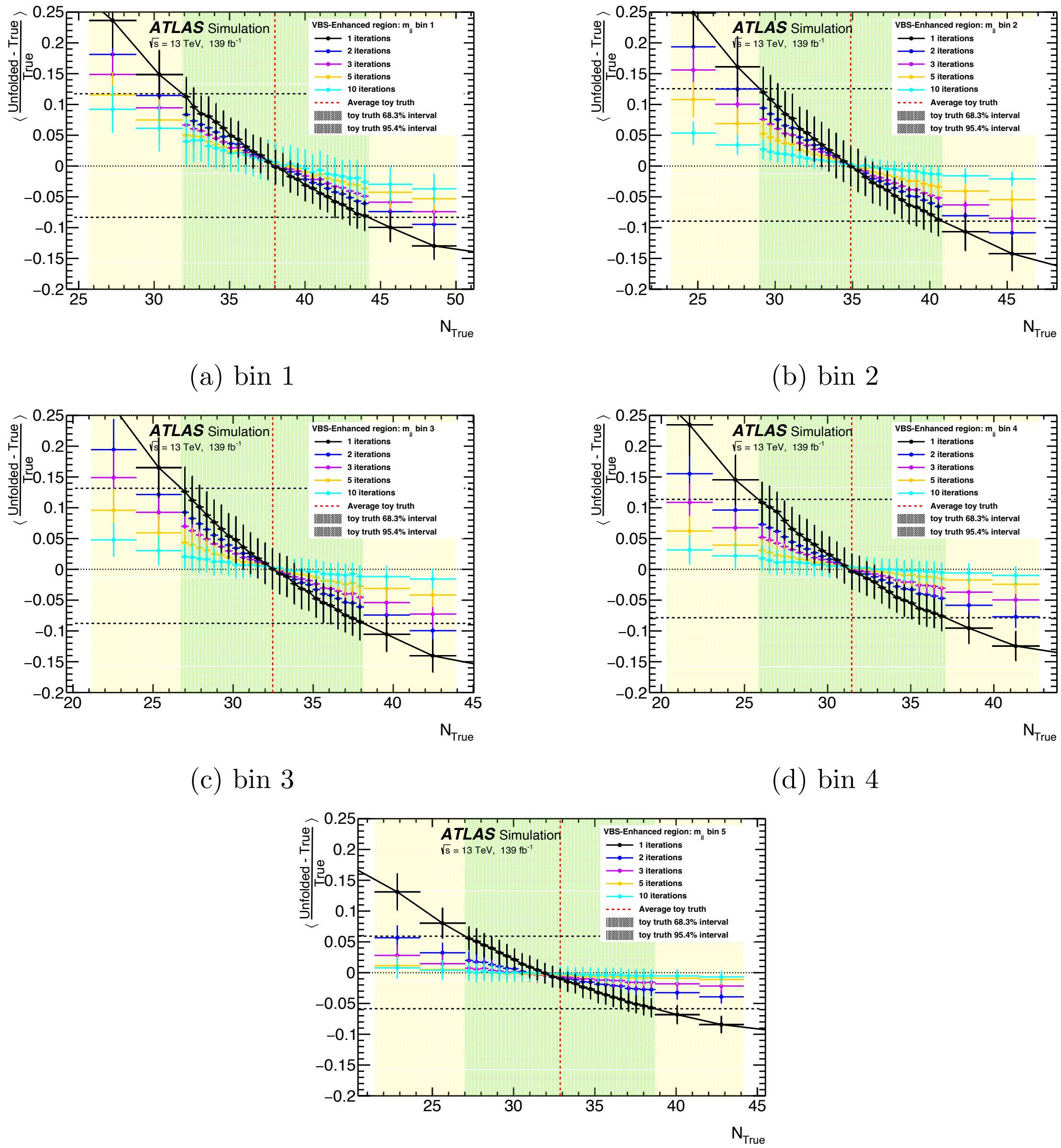
Table 16: Breakdown of the relative systematic uncertainties (%) for each bin of m_{jj} in the VBS-Enhanced region.

bin migration. Therefore, the statistical uncertainty on the unfolded yield is a combination of the uncertainties in detector-level bins and uncertainties on the migration probabilities, which takes the covariance between the detector-level bins into account. The statistical uncertainties at the unfolded level are evaluated by the *RooUnfold* package, which propagates both of these uncertainties.

18.6 Breakdown of Uncertainties

Table 16 shows the impact of systematic uncertainties in the VBS-Enhanced region for each bin of m_{jj}. In most bins, the unfolding bias is the dominant source of systematic uncertainty, followed by the jet systematics.

Figure 60a schematically shows different types of uncertainties affecting the measurements in the VBS-Enhanced region as a function of m_{jj}. The solid black line represents the total statistical uncertainty and has the most significant impact in all bins. Total systematic uncertainty, represented by the dashed black line, combines all systematic uncertainties. Depending on the bin, it is dominated either by unfolding bias or jet-related uncertainties. Figure 60b shows the impact of different categories of the jet systematic uncertainties. In most bins of m_{jj}, the dominant jet uncertainties are from the pile-up energy correction step in the jet calibration. The uncertainties from jet eta-dependent calibration and jet energy resolution are also significant. Overall, the jet reconstruction uncertainties have about $8-9\%$ effect on each bin of the unfolded cross-sections.



(a) bin 1

(b) bin 2

(c) bin 3

(d) bin 4

(e) bin 5

Figure 58: MC-based unfolding bias in each bin of m_{jj} in the VBS-Enhanced region using Gaussian toys. The distribution shows the relative difference between unfolded and true values for toys as a function of the true value in each bin. The number of unfolding iterations is varied, and the final bias is chosen as the one-sigma intersection with the average value of the toy truth distribution (light-green region intersecting with the black curve).

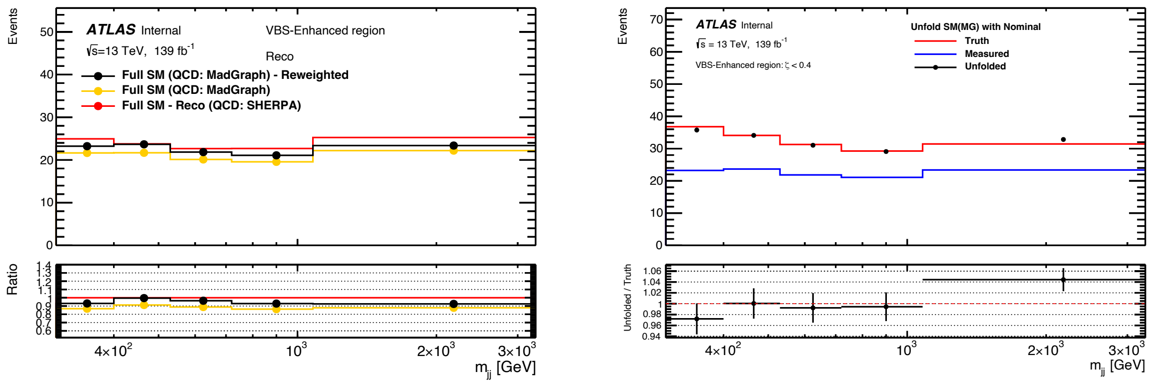
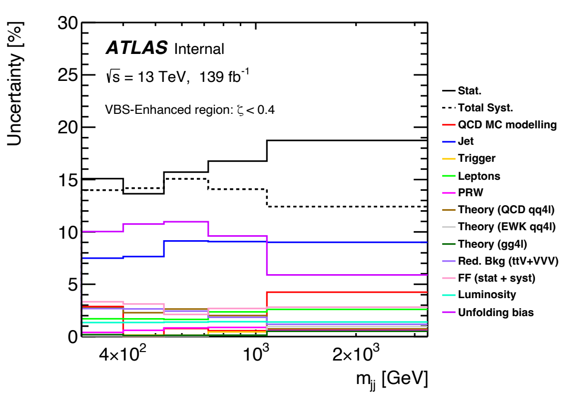
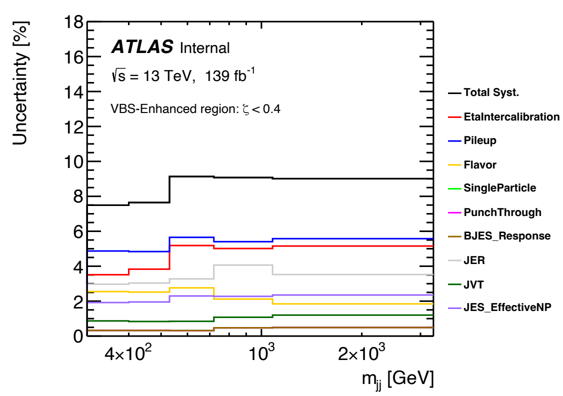


Figure 59: The left plot shows three distributions of m_{jj} in VBS-Enhanced region at detector-level, red corresponding to SM predictions where $qqZZ$ is taken from Sherpa, yellow shows the same but $qqZZ$ is taken from MadGraph5 and black shows the reweighted-MadGraph5 distribution to match the Sherpa lineshape. The right plot shows reweighted-MadGraph5 detector-level (blue) distribution, which is unfolded (black) using unfolding inputs from nominal-Sherpa and compared to the particle-level reweighted-MadGraph5 distribution (red). The ratio panel of the right plot showing the ratio between reweighted truth-level MadGraph5 and reweighted unfolded-level MadGraph5 gives the QCD modeling uncertainties.



(a) Statistical and systematic uncertainties affecting the differential cross-section measurements.



(b) Breakdown of the total jet-related uncertainties.

Figure 60: Uncertainties as a function of m_{jj} in the VBS-Enhanced region.

Chapter VI: Results

This chapter presents the main results of this thesis. Section 19 presents the detector level measurements of the eleven kinematic observables introduced in Section 15, whereas Section 20 presents the unfolded differential cross-sections. Section 21 reinterprets the unfolded cross-sections shown in Section 20 to constrain parameters of physics beyond the SM affecting the quartic gauge vertices of electroweak $ZZjj$ production.

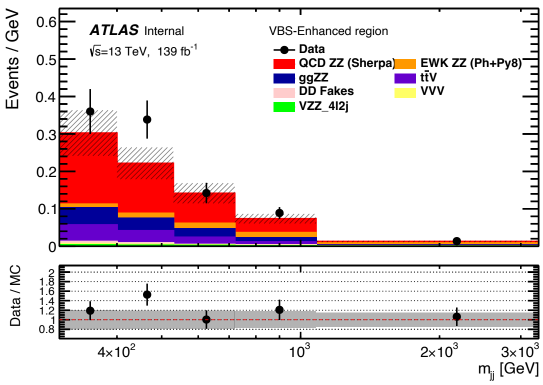
19 Detector Level Measurements

Figure 61 shows the measured detector level data and predicted detector level yield for six kinematic observables; the invariant mass of the dijet [m_{jj}], invariant mass of the two Z bosons [$m_{4\ell}$], transverse momentum of the two tagging jets [$p_{T,jj}$], transverse momentum of the two Z bosons [$p_{T,4\ell}$], transverse momentum of the two Z bosons and two tagging jets [$p_{T,4\ell jj}$], and scalar transverse momentum of the two Z bosons and dijet [$s_{T,4\ell jj}$]. Similarly, figure 62 shows the same for remaining kinematic observables, cosine of the decay angle of the negative lepton of the leading (sub-leading) pair in the pair's rest frame [$\cos\theta^*_{\ell 1(3)\ell 2(4)}$], signed difference between the azimuthal angle of two jets [$\Delta\phi_{jj}^{signed}$], rapidity difference between two jets [Δy_{jj}], and centrality of the system [ζ].

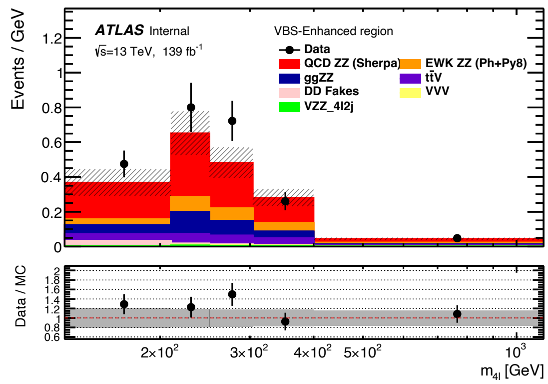
For each of these distributions, the measured data (black dot) is compared with state-of-the-art SM predictions, where the two QCD signal $qqZZjj$ (red) & $ggZZjj$ (blue), and the two MC predicted background, VVV (yellow) and ttV (purple) are taken from Sherpa predictions. The contribution from fake backgrounds in light pink is estimated using the data-driven method. The $qqZZjj$ electroweak signal is taken from the PowhegV2 predictions, and the electroweak production of triboson and two jets ($VZZjj$) is taken from the Sherpa predictions. The vertical-solid black line on the data points represents the statistical uncertainty in the data, and the dashed black box in each bin represents the impact of the total theoretical and experimental uncertainties on the predicted detector-level yields. The impact of the total systematic uncertainties and the statistical precision each ranges from 15 to 20% depending on the bins and the distributions.

The ratio panel in these distributions shows the data yield ratio to the total predicted yield, which agrees with the SM predictions with the uncertainties. Some discrepancies are observed in some distributions; for instance, the MC yield is underpredicted in the second bin of m_{jj} and in the third bin of m_{4l}. However, this difference is statistically insignificant. Moreover, a slight but statistically insignificant asymmetry is also observed in the measured distribution of $\Delta\phi_{jj}^{signed}$, which in SM is expected to be symmetric.

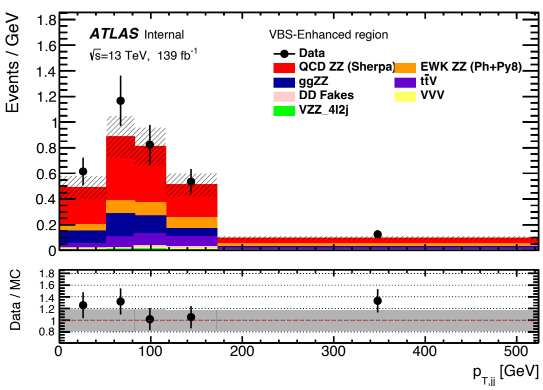
A simple chi-squared per degree of freedom (χ^2/NDF) is estimated to quantify whether the measured data agree with the SM prediction. The χ^2/NDF takes the residual difference in each bin between the unweighted data yield and the weighted MC prediction yield, along with the statistical and systematic uncertainties for both distributions. The reported values of χ^2/NDF for each distribution in figures 61 and 62 show statistically good agreement between the measured data and MC predictions.



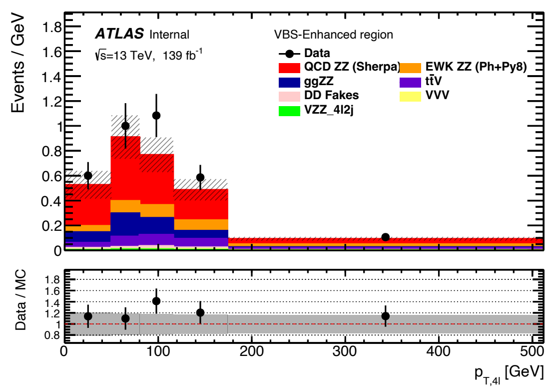
(a) m_{jj}: $\chi^2/NDF = 0.43$



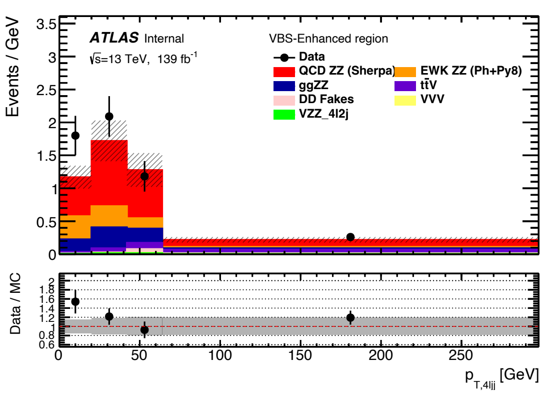
(b) $m_{4\ell}$: $\chi^2/NDF = 0.49$



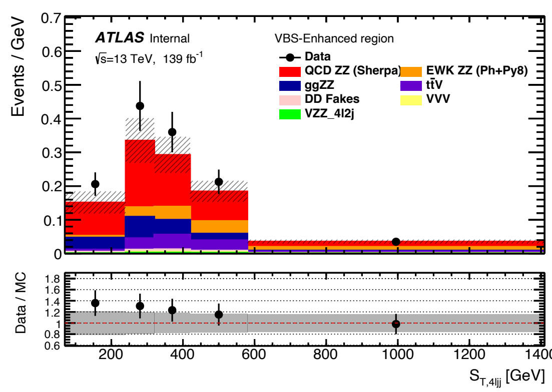
(c) $p_{T,jj}$: $\chi^2/NDF = 0.30$



(d) $p_{T,4\ell}$: $\chi^2/NDF = 0.15$

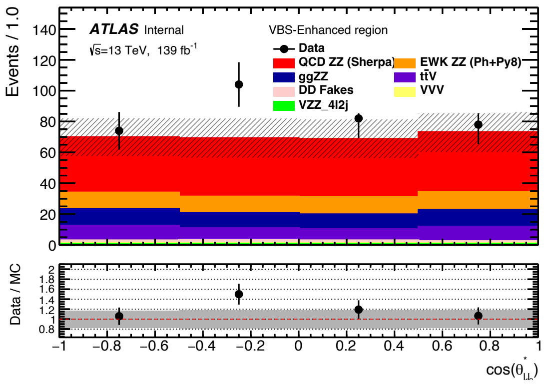


(e) $p_{T,4\ell jj}$: $\chi^2/NDF = 0.70$

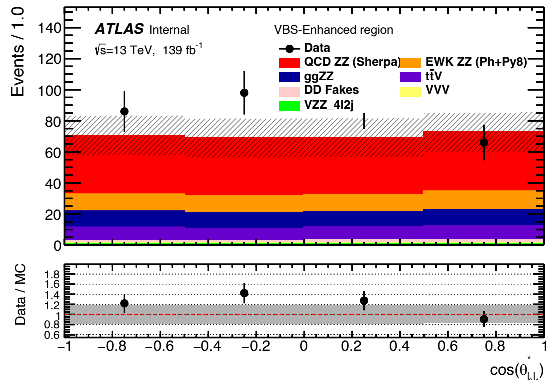


(f) $s_{T,4\ell jj}$: $\chi^2/NDF = 0.24$

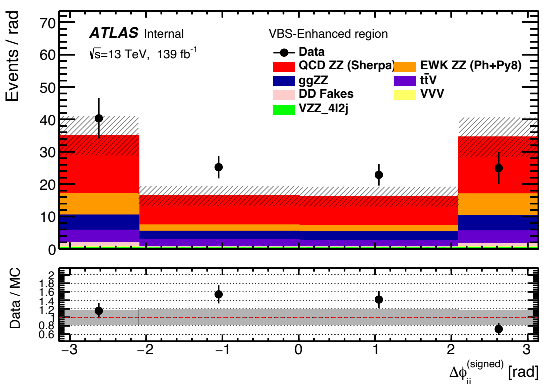
Figure 61: Detector level distributions in the VBS-Enhanced region.



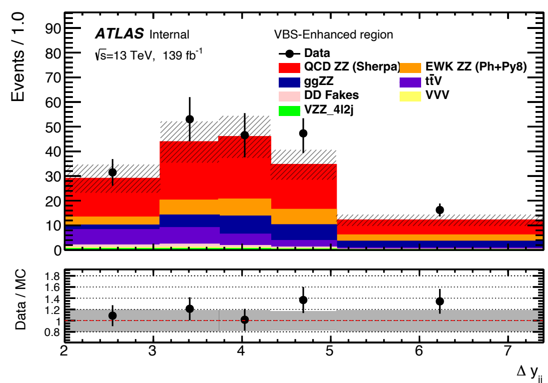
(a) $\cos\theta^*_{\ell 1\ell 2}$: $\chi^2/NDF = 0.51$



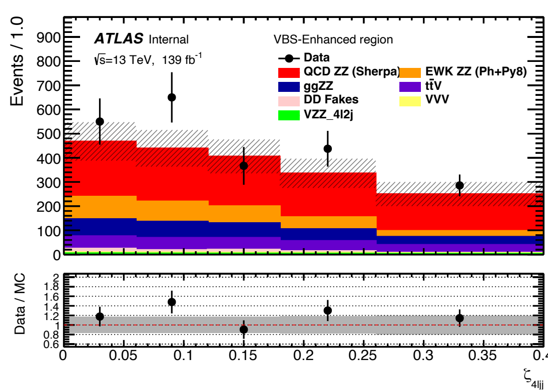
(b) $\cos\theta^*_{\ell 3\ell 4}$: $\chi^2/NDF = 0.61$



(c) $\Delta\phi_{jj}^{signed}$: $\chi^2/NDF = 1.69$



(d) Δy_{jj} : $\chi^2/NDF = 0.26$



(e) ζ : $\chi^2/NDF = 0.50$

Figure 62: Detector level distributions in the VBS-Enhanced region.

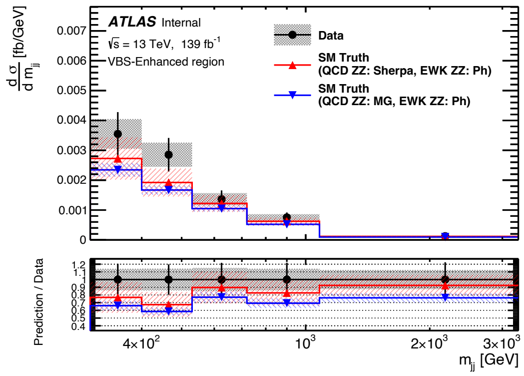
20 Unfolded Differential Cross-sections

The background-subtracted data yield is unfolded using the iterative Bayesian Unfolding discussed in Section 17. The unfolded differential cross-sections, the main results of this thesis, are obtained by multiplying the inverse of integrated luminosity to the background-subtracted unfolded yield in each bin. The unfolded differential cross-sections for eleven kinematic observables are shown in figures 63 and 64.

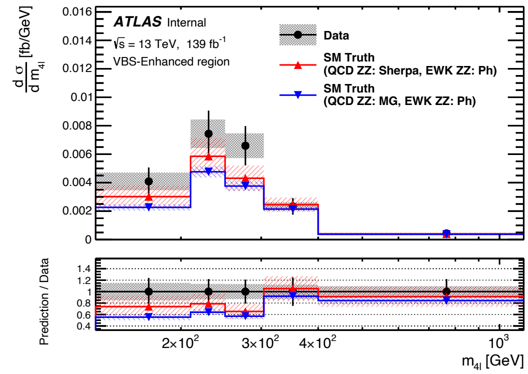
Each distribution of unfolded differential cross-sections in black is compared to two different state-of-the-art particle-level SM predictions in red and blue. The red distribution represents the SM particle level differential cross-sections where the QCD $qqZZjj$ contribution is predicted by SHERPA generator. Similarly, in the blue distribution, the QCD $qqZZjj$ contribution is predicted by MADGRAPH generator. The light-red and light-blue bands are the fiducial level systematics on the particle-level yield, respectively, by SHERPA and MADGRAPH. The vertical-solid line on the unfolded data points is the statistical uncertainty, whereas the black band represents the total systematic uncertainties from theoretical, experimental, and unfolding sources. The unfolded cross-sections are limited by statistical precision in all distributions and all bins. For some distributions, one or two events are found in the overflow bin resulting from bin migrations. These events are added to the content of the last bin of the unfolded distribution.

Generally, for all distributions, the data is well modeled by the MC simulations within 2σ of the uncertainty band. Two different p-values are determined by comparing unfolded cross-sections to the two predicted cross-sections to quantify the agreement between the experimentally measured unfolded and the SM-predicted cross-sections. The p-value is calculated based on a technique discussed in Ref [89] by taking the residual and uncertainties of two weighted histograms. For all kinematic observables, the reported p-values obtained by comparing to either generator are more significant than 0.05. Therefore, in the analyzed LHC Run-2 dataset, for the $ZZ(\rightarrow 4)\ell jj$ process, all differential cross-sections in the

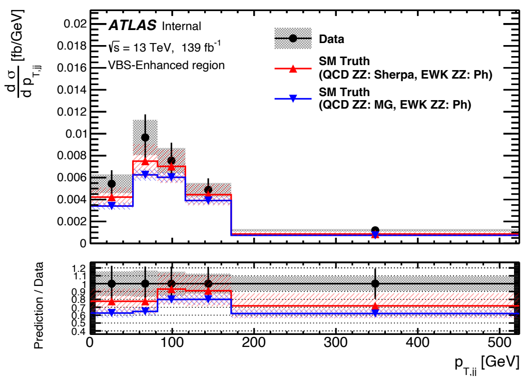
VBS-Enhanced region are concluded to agree with the SM predictions.



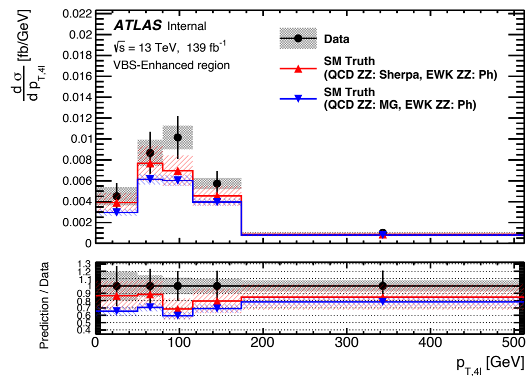
(a) m_{jj}: p-value(Sherpa=0.91 & MG=0.79)



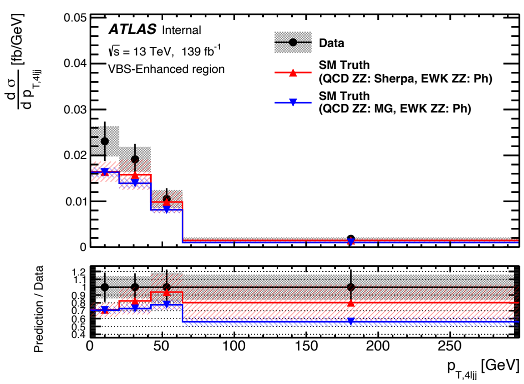
(b) $m_{4\ell}$: p-value(Sherpa=0.74 & MG=0.16)



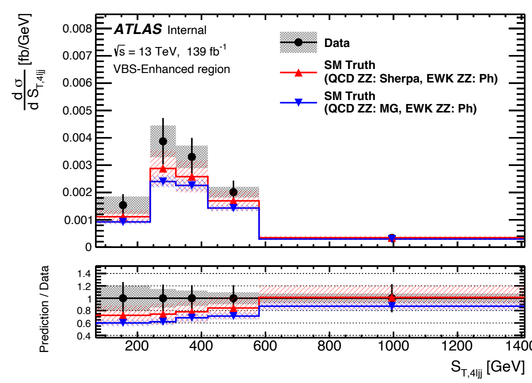
(c) $p_{T,jj}$: p-value(Sherpa=0.95 & MG=0.71)



(d) $p_{T,4\ell}$: p-value(Sherpa=0.96 & MG=0.83)

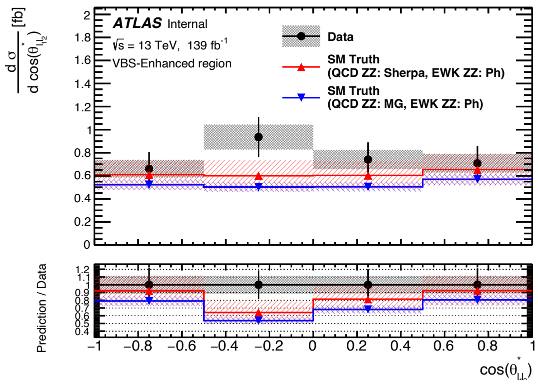


(e) $p_{T,4\ell jj}$: p-value(Sherpa=0.88 & MG=0.57)

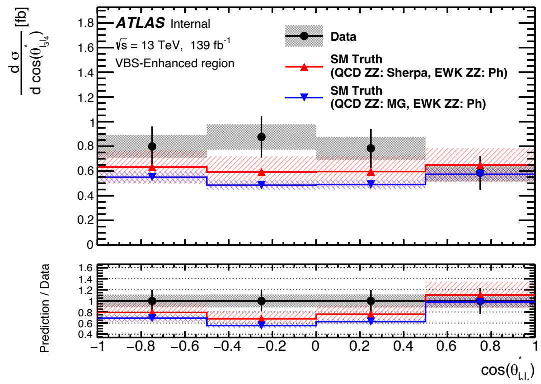


(f) $s_{T,4\ell jj}$: p-value(Sherpa=0.90 & MG=0.63)

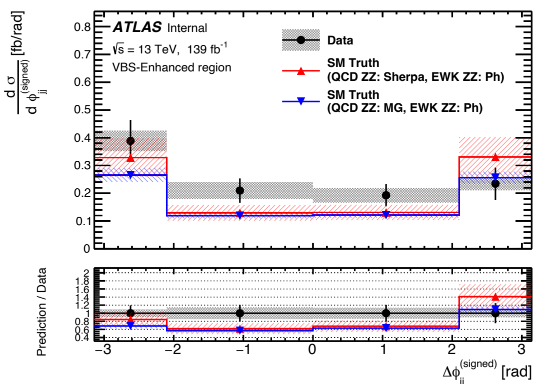
Figure 63: Unfolded differential cross-sections in the VBS-Enhanced region.



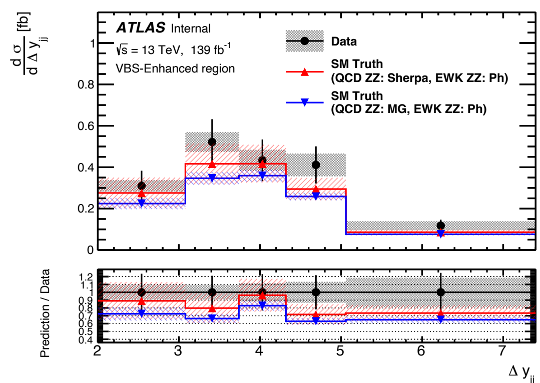
(a) $\cos\theta^*_{\ell 1\ell 2}$: p-value(Sherpa=0.70 & MG=0.23)



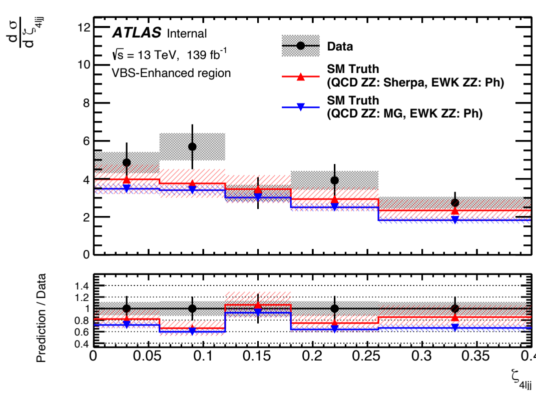
(b) $\cos\theta^*_{\ell 3\ell 4}$: p-value(Sherpa=0.61 & MG=0.11)



(c) $\Delta\phi_{jj}^{signed}$: p-value(Sherpa=0.16 & MG=0.05)



(d) Δy_{jj}: p-value(Sherpa=0.91 & MG=0.83)



(e) ζ: p-value(Sherpa=0.79 & MG=0.55)

Figure 64: Unfolded differential cross-sections in the VBS-Enhanced region.

21 Effective Field Theory ReInterpretation

Chapter VII: Conclusion and Outlook

22 Conclusion

Vector boson scattering is a critical phenomenon in the electroweak sector of the Standard Model of particle physics. Vector boson scattering processes include rare triple and quartic self-couplings of the electroweak gauge bosons whose production rate at high energies is sensitive to possible modifications from physics beyond the Standard Model. The quartic gauge-self-couplings are experimentally accessible for the first time with the ATLAS datasets of 139 fb^{-1} recorded during the $2015-2018$ data-taking period. The presence of clean signature of two same-flavor, opposite-sign lepton pairs in the $ZZ(\rightarrow 4\ell)jj$ final state with minor contributions from background processes offers an excellent approach to studying the high energy behavior of the vector boson scattering. However, given the low production cross-section of electroweak $ZZjj$ and small branching ratio of $Z \rightarrow e^{+}e^{-}(\mu^{+}\mu^{-})$, these processes are statistically limited with the current dataset. Therefore, unfolded differential cross-section measurements of $ZZ(\rightarrow 4\ell)jj$ in an electroweak enhanced phase-space are measured as a function of eleven kinematic observables and compared to the state-of-the-art Standard Model predictions. The measured cross-sections agree with the theoretical predictions within the experimental and statistical uncertainties. The unfolded cross-sections are then used to put competitive constraints on beyond the Standard Model effects using a model-independent, effective field theory approach.

23 Outlook

With Run-2 datasets, the electroweak production of several multiboson processes such as VBS same-sign WW [90], VBS WWW [91], VBS WZ [92], and VBS ZZ [17] were experimentally observed for the first time with the ATLAS experiment. These VBS measurements are still statistically dominated and could gain greater precision from more extensive statistics. The third physics operation of LHC after three years of upgrade started in July of 2022 and is expected to continue until 2025 at the highest to date center-of-mass-energy of proton-proton collisions, at $\sqrt{s} = 13.6$ TeV [93]. In Run-3, the ATLAS experiment is expected to record more than twice the dataset of Run-2 corresponding to the integrated luminosity of 300 fb^{-1}. Run-3 statistics is crucial to study the cross-sections of any VBS processes differentially. The Run-3 datasets are expected to make the differential cross-sections measurement of the fully electroweak $ZZ(\rightarrow 4\ell)jj$ process statistically feasible and put stringer constraints on the BSM parameters causing anomalous self-interactions of the gauge bosons.

However, high-luminosity LHC discussed briefly in Section 8.1 is expected to be a golden era for the vector boson scattering measurements. The ATLAS experiment is expected to record about ten times more data and more precise reconstruction of the forward jets, essential physics objects defining the VBS processes, which is driven by the extended η coverage of the inner tracker and the additional timing information from the high granularity timing detectors. With the extensive statistics and the unprecedented proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 14$ TeV at HL-LHC, the scattering of the longitudinally polarized vector bosons is expected to be within the experimental reach [94]. As discussed in section 4, the self-interactions of the longitudinally polarized vector bosons are regularized by the Higgs-mediated processes to restore the unitarity at high energies. Therefore, the ultimate goal of the Standard Model electroweak multiboson processes is experimentally accessing the electroweak cross-sections of the longitudinally polarized vector bosons.

References

[1] P.A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020.

[2] Francis Halzen and Alan Martin. *Quarks & Leptons: An introductory course in modern particle physics*. John Wiley & Sons, New York, USA, 1984.

[3] Wikipedia contributors. Standard Model of Elementary Particles. https://en.wikipedia.org/wiki/Standard_Model, 2019. [Online; accessed 05-November-2022].

[4] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J. C*, 63:189–285, 2009.

[5] Fabienne Marcastel. CERN's Accelerator Complex. La chaîne des accélérateurs du CERN. 2013. General Photo.

[6] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003, aug 2008.

[7] Sebastian Murk. Trigger Algorithms and Monte Carlo Event Generation for Dijet Searches in ATLAS and Beyond. 2017.

[8] ATLAS Collaboration. Alignment of the ATLAS Inner Detector in Run-2. *Eur. Phys. J. C*, 80(12), 2020.

[9] Joao Pequenao. Event Cross Section in a computer generated image of the ATLAS detector. 2008.

[10] ATLAS Collaboration. DAQ approved plots, 2020.

[11] Cornelissen, T G and Van Eldik, et al. . Updates of the ATLAS Tracking Event Data Model (Release 13). Technical report, CERN, Geneva, 2007.

[12] 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. *The European Physical Journal C*, 79(8), aug 2019.

[13] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton-proton collision data $\sqrt{s} = 13$ TeV . *The European Physical Journal C*, 76(5), may 2016.

[14] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS Detector . *The European Physical Journal C*, 77(7), jul 2017.

[15] ATLAS Collaboration. Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 81(8):689, 2021.

[16] Technical Design Report for the ATLAS Inner Tracker Pixel Detector. Technical report, CERN, Geneva, 2017.

[17] ATLAS Collaboration. Observation of electroweak production of two jets and a Z-boson pair, 2020.

[18] The ATLAS Collaboration. ATLAS data quality operations and performance for 2015-2018 data-taking. *Journal of Instrumentation*, 15(04):P04003–P04003, apr 2020.

[19] Yanyan Gao and Andrei V. Gritsan and Zijin Guo and Kirill Melnikov and Markus Schulze and Nhan V. Tran. Spin determination of single-produced resonances at hadron colliders. *Physical Review D*, 81(7), apr 2010.

[20] Mark Thomson. *Modern particle physics.* Cambridge University Press, New York, 2013.

[21] E. Noether. Invariante variationsprobleme. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1918:235–257, 1918.

[22] Bernabeu, Jose", and Murillo, Jesús Carnicer and Michelini, Marisa and Perea, Carmen. *Fundamental Physics and Physics Education Research*, chapter Symmetries in the Standard Model, pages 3–16. Springer International Publishing, Cham, 2021.

[23] The OPAL Collaboration. Measurement of the running of the QED couplingin small-angle Bhabha scattering at LEP. *The European Physical Journal C*, 45(1):1–21, jan 2006.

[24] Alexandre Deur, Stanley J. Brodsky, and Guy F. de Téramond. The QCD running coupling. *Progress in Particle and Nuclear Physics*, 90:1–74, sep 2016.

[25] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588, 1961.

[26] Weinberg, Steven. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[27] A. Salam and J.C. Ward. Electromagnetic and weak interactions. *Physics Letters*, 13(2):168–171, 1964.

[28] Nguyen, K. The Higgs Mechanism, 2009.

[29] M. E. Peskin , D. V. Schroeder. *An Introduction to Quantum Field Theory.* Westview PRESS, Colorado, USA, 1995.

[30] Buchmüller, W. and Lüdeling, C. Field Theory and Standard Model, 2006.

[31] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, sep 2012.

[32] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, 2012.

[33] The CDF II Collaboration. High-precision measurement of the W boson mass with the CDF II detector. *Science*, 376(6589):170–176, 2022.

[34] Richard Massey, Thomas Kitching, and Johan Richard. The dark matter of gravitational lensing. *Reports on Progress in Physics*, 73(8):086901, jul 2010.

[35] Y. Ashie, J. Hosaka, and Ishihara et al. Evidence for an oscillatory signature in atmospheric neutrino oscillations. *Phys. Rev. Lett.*, 93:101801, Sep 2004.

[36] Peter Skands. QCD for Collider Physics, 2011.

[37] Sergey Alekhin, Kirill Melnikov, and Frank Petriello. Fixed target Drell-Yan data and NNLO QCD fits of parton distribution functions. *Phys. Rev. D*, 74:054033, Sep 2006.

[38] Sergey et al. Alekhin. The PDF4LHC Working Group Interim Report, 2011.

[39] Andy Buckley et al. General-purpose event generators for LHC physics. *Physics Reports*, 504(5):145–233, 2011.

[40] The CMS Collaboration. Evidence for electroweak production of four charged leptons and two jets in proton-proton collisions at s=13TeV. *Physics Letters B*, 812:135992, 2021.

[41] Jager, B., Karlberg, A. and Zanderighi, G. . Electroweak ZZjj production in the Standard Model and beyond in the POWHEG-BOX V2. *J. High Energ. Phys.*, 141, 2014.

[42] A. Denner and T. Hahn. Radiative corrections to W+W- $\rightarrow$ W+W- in the electroweak standard model. *Nuclear Physics B*, 525(1-2):27–50, 1998.

[43] Tao Han and David Krohn and Lian-Tao Wang and Wenhan Zhu. New physics signals in longitudinal gauge boson scattering at the LHC. *J. High Energ. Phys.*, 2010(3), mar 2010.

[44] Eduardo da Silva Almeida and O. J. P. Éboli and M. C. Gonzalez–Garcia. Unitarity constraints on anomalous quartic couplings. *Physical Review D*, 101(11), jun 2020.

[45] Jung Chang, Kingman Cheung, Chih-Ting Lu, and Tzu-Chiang Yuan. WW Scattering in the Era of Post Higgs Discovery. *Physical Review D*, 87(9), may 2013.

[46] Khoze, V., Ryskin, M., Stirling, W. et al. A Z-monitor to calibrate Higgs production via vector boson fusion with rapidity gaps at the LHC. *Eur. Phys. J.*, C 26:429–440, 2003.

[47] LHC Machine. *JINST*, 3:S08001, 2008.

[48] *LEP design report.* CERN, Geneva, 1984.

[49] LHC Guide. 2017.

[50] Capeans, M and Darbo, G and Einsweiller, et al. ATLAS Insertable B-Layer Technical Design Report. Technical report, 2010.

[51] Technical Design Report for the ATLAS Inner Tracker Pixel Detector. Technical report, CERN, Geneva, 2017.

[52] Technical Design Report for the ATLAS Inner Tracker Strip Detector. Technical report, CERN, Geneva, 2017.

[53] Particle Identification Performance of the ATLAS Transition Radiation Tracker. Technical report, CERN, Geneva, 2011.

[54] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1996.

[55] *ATLAS tile calorimeter: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1996.

[56] The ATLAS collaboration. Operation of the ATLAS trigger system in Run 2. *Journal of Instrumentation*, 15(10):P10004, oct 2020.

[57] The ATLAS Collaboration. Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2. *The European Physical Journal C*, 77(10), 2017.

[58] Federico Meloni. Primary vertex reconstruction with the ATLAS detector. Technical Report 12, CERN, Geneva, 2016.

[59] ATLAS Collaboration. Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data. *The European Physical Journal C*, 74(11), nov 2014.

[60] Matteo Cacciari and Gavin P. Salam and Gregory Soyez. The anti-kt jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, apr 2008.

[61] ATLAS Collaboration. Tagging and suppression of pileup jets with the ATLAS detector. Technical report, CERN, Geneva, 2014.

[62] ATLAS Collaboration. Forward jet vertex tagging using the particle flow algorithm. Technical report, CERN, Geneva, 2019.

[63] ATLAS Collaboration. *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2020.

[64] Technical Design Report for the ATLAS Inner Tracker Strip Detector. 4 2017.

[65] G. Avoni et al. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS. *Journal of Instrumentation*, 13(07):P07017, jul 2018.

[66] The ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. Technical report, CERN, 6 2019.

[67] S. Agostinelli et al. GEANT4–a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250–303, 2003.

[68] Enrico Bothmann, Gurpreet Singh Chahal, Stefan Höche, Johannes Krause, Frank Krauss, Silvan Kuttimalai, Sebastian Liebschner, Davide Napoletano, Marek Schönherr, Holger Schulz, Steffen Schumann, and Frank Siegert. Event generation with Sherpa 2.2. *SciPost Physics*, 7(3), sep 2019.

[69] Richard D. Ball, , Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, Juan Rojo, and Maria Ubiali. Parton distributions for the LHC run II. *Journal of High Energy Physics*, 2015(4), apr 2015.

[70] S. Schumann and F. Krauss. A parton shower algorithm based on catani-seymour dipole factorisation. *Journal of High Energy Physics*, 2008(03):038, mar 2008.

[71] Höche, S., Krauss, F., Schönherr, M. et al. QCD matrix elements + parton showers. The NLO case. *J. High Energ. Phys.*, 2013(27), 2013.

[72] J. Alwall and R. Frederix and S. Frixione and V. Hirschi and F. Maltoni and O. Mattelaer and H.-S. Shao and T. Stelzer and P. Torrielli and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), jul 2014.

[73] Torbjörn Sjöstrand and Stephen Mrenna and Peter Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852–867, jun 2008.

[74] Barbara Jäger and Alexander Karlberg and Giulia Zanderighi. Electroweak ZZjj production in the Standard Model and beyond in the POWHEG-BOX V2. *Journal of High Energy Physics*, 2014(3), mar 2014.

[75] Stefano Catani and Frank Krauss and Bryan R Webber and Ralf Kuhn. Qcd matrix elements + parton showers. *Journal of High Energy Physics*, 2001(11):063–063, nov 2001.

[76] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.

[77] ATLAS Collaboration. ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13\,\text{TeV}$. 2019.

[78] J. Maurer, J. Reichert, E. Varnes. Technical discussion of FakeBkgTools. 2018.

[79] Giulio D'Agostini. A multidimensional unfolding method based on Bayes' Theorem. Technical report, DESY, Hamburg, 1994.

[80] D'Agostini, G. Improved iterative Bayesian unfolding, 2010.

[81] Brenner, Lydia and Balasubramanian, Rahul and Burgard, Carsten and Verkerke, Wouter and Cowan, Glen and Verschuuren, Pim and Croft, Vincent. Comparison of unfolding methods using RooFitUnfold. *International Journal of Modern Physics A*, 35(24):2050145, 2020.

[82] Friedman, Jerome H. A VARIABLE SPAN SMOOTHER. 10 1984.

[83] Enrico Bothmann and Marek Schönherr and Steffen Schumann. Reweighting QCD matrix-element and parton-shower calculations. *The European Physical Journal C*, 76(11), oct 2016.

[84] L. A. Harland-Lang and A. D. Martin and P. Motylinski and R. S. Thorne. Parton distributions in the LHC era: MMHT 2014 PDFs. *The European Physical Journal C*, 75(5), may 2015.

[85] Sayipjamal Dulat and Tie-Jiun Hou and Jun Gao and Marco Guzzi and Joey Huston and Pavel Nadolsky and Jon Pumplin and Carl Schmidt and Daniel Stump and C.-P. Yuan. New parton distribution functions from a global analysis of quantum chromodynamics. *Physical Review D*, 93(3), feb 2016.

[86] Fabrizio Caola and Matthew Dowling and Kirill Melnikov and Raoul Röntsch and Lorenzo Tancredi. QCD corrections to vector boson pair production in gluon fusion including interference effects with off-shell Higgs at the LHC. *Journal of High Energy Physics*, 2016(7), jul 2016.

[87] ATLAS Collaboration. Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 99:072009, Apr 2019.

[88] ATLAS Collaboration. Evidence for the production of three massive vector bosons with the ATLAS detector. *Physics Letters B*, 798:134913, 2019.

[89] Gagunashvili, N. D. Comparison of weighted and unweighted histograms. 2006.

[90] ATLAS Collaboration. Observation of electroweak production of a same-sign W boson pair in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Physical Review Letters*, 123(16), oct 2019.

[91] ATLAS Collaboration. Observation of WWW Production in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector. *Physical Review Letters*, 129(6), aug 2022.

[92] ATLAS Collaboration. Observation of electroweak $W^{\pm}Z$ boson pair production in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Physics Letters B*, 793:469–492, jun 2019.

[93] ATLAS Collaboration. ATLAS Experiment records "first physics" at new high-energy frontier. CERN, Jul. 5, 2022 [Online].

[94] Prospects for the measurement of the $W^{\pm}W^{\pm}$ scattering cross section and extraction of the longitudinal scattering component in pp collisions at the High-Luminosity LHC with the ATLAS experiment. Technical report, CERN, Geneva, 2018.

Appendices

Supplementary materials to the main body of this thesis are presented in the following appendices. Appendix A summarizes my contributions to the measurement presented in this thesis and to the ATLAS experiment. The largest source of systematic uncertainty on the results is the unfolding bias; thus, Appendix B summarizes the studies conducted to understand the source of this bias. Finally, Appendix C presents the differential unfolded cross-sections measured in the VBS-suppressed region of the phase space.

A Personal Contribution

A.1 Contribution to $ZZ(\to 4\ell)jj$ Measurement

The measurement presented in this thesis is only possible from the effort of the entire analysis team. I am one of the two primary analyzers who have contributed to several measurement areas from its initial formation to the current stage of finalization and publication. I contributed significantly to defining the analysis phase space discussed in Chapter IV. The phase space defined in the preceding ATLAS analysis [17], which observed the electroweak $ZZjj$ process was not optimal for differential measurement. Therefore, I worked on loosening the kinematic selections to increase the acceptance, defining the isolation working point to maintain optimal signal-selection and background-rejection probabilities, and establishing novel pair sorting strategy to reduce the bin migration in unfolding. Moreover, I was responsible for maintaining the main analysis framework, which implements the latest recommendations from combined performance groups for physics object reconstruction discussed in Chapter III, and applies most of the kinematic cuts for object and event selection discussed in Chapter IV. This framework applies all the scale factors and event weights in MC events, and selects a few hundred relevant physics events that passes all kinematic selections for both data and MC. Additionally, I estimated the fake backgrounds of the analysis using the background estimation technique discussed in Section 16 and assisted in selecting relevant systematics discussed in Section 18 of Chapter V. I also validated the novel next-to-leading-order Powheg $qqZZjj$ MC sample used as the primary sample for the electroweak production of $ZZ(\to 4\ell)jj$.

A.2 Contribution to the ATLAS Experiment

I have been a member of the ATLAS collaboration since 2017 and have contributed to the three critical areas of the experiment; detector development, detector performance, and physics analyses.

Detector Developement: I spent my first year as a graduate student working at Brookhaven National Lab, where I contributed to the prototype development of the all-silicon inner tracker for the high luminosity LHC. During this period, I assembled the first three prototype staves, the detector units of the ITk strip barrel region detector, using a semi-automated loading setup consisting of an Aeoretch robotic arm, cameras, and an alignment system. I developed a $3D$-printed tooling pins used in the alignment of loading the sensors on staves. I led comprehensive thermal and mechanical tests of the first prototype, the thermo-mechanical stave using IR imaging and laser metrology, respectively. These tests validated the cooling system designs of the ITk stave core structure and the stability of the mechanical design, giving the green light to the production of 200 out of 400 strip-staves needed to build the barrel region of the inner tracker. During this year, I learned the fundamentals of particle physics detectors, their development, and their operations.

Physics Analysis: Apart from the $ZZ(\to 4\ell)jj$ analysis presented in this thesis, I have worked on two other ATLAS analyses involving four leptons in the final state, analyzing the complete Run-2 datasets. After my Ph.D. candidacy and ATLAS qualification, I joined the inclusive four-lepton measurement analysis team, whose goal was to inclusively measure the unfolded cross-sections of the Standard Model four-lepton process. In a year as a part of the team, I worked on finding the suitable lepton isolation working points, studying the impact of including electrons and muons originating from tau leptons in the unfolded differential cross-sections, and the most precise measurement to date of the branching ratio of $Z \to 4\ell$ with full Run-2 dataset.

Since early 2021, I have been a crucial part of the on-shell ZZ CP and polarization analysis team, where the main goal is to extract the fraction of two Z bosons simultaneously longitudinally polarized and search for additional CP violation. Like the $ZZ(\to 4\ell)jj$ analysis, I have contributed to background estimation, phase space optimization, and event selection. Additionally, I have contributed to deriving the unfolded cross-sections with all relevant systematic and statistical uncertainties used in CP violation searches using an effective field theory approach.

Detector Performance: The training of a particle physicist is incomplete without understanding the detector's performance. Therefore, in 2021 I joined the Tracking Combined Performance group of ATLAS and have contributed to the validation of the performance of the Run-3 tracking reconstruction software in both early Run-3 data and different types of MC simulation.

Similar to $ZZ(\rightarrow 4\ell)jj$ measurement, most ATLAS physics analyses use the vertex with the highest value of the sum-squared of track's transverse momenta ($\sum_{tracks} p_T^2$) as the hard scattering vertex of the measurement. However, in processes with softer leptons and invisible tracks (including photons), $\sum_{tracks} p_T^2$ is inadequate to identify the hard scatter vertex. Therefore, I am currently working on developing an alternative algorithm that is suitable for a variety of different physics processes.

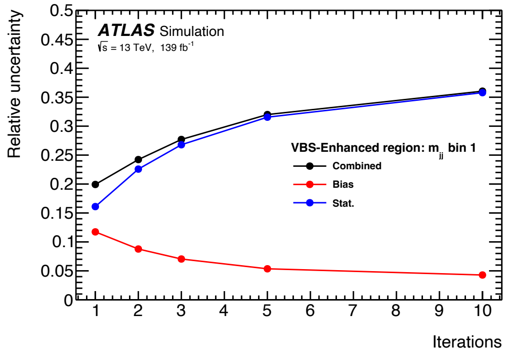
The experience I have gained in different areas of particle physics has significantly shaped my discussion of the measurement presented in this document.

B Additional Study on Unfolding Bias

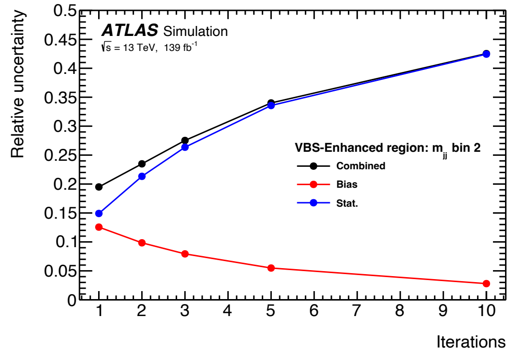
As discussed in Section 17.4, the bias from the unfolding process is the largest source of the systematic uncertainty for the measurement. Additional studies were conducted to understand the underlying source of this bias. Figure 65 shows the unfolding bias and statistical uncertainty on the unfolded yield as a function of the increasing number of unfolding iterations for each bin of m_{jj} in the VBS-Enhanced region. The bias is evaluated using the MC-toy-based method introduced in Section 17.4. The total uncertainty shown in the black distribution is always the smallest for a single iteration, further assuring the choice of iteration is optimal. As expected, the bias decreases and the statistical uncertainty increases with the increasing number of unfolding iterations.

The unfolding bias is expected to converge to a value of zero with a higher number of iterations. However, as observed in figure 65, the rate of convergence of unfolding bias is lower, suggesting that the fiducial fakes present in the detector level distributions are not fully corrected by the unfolding method. The fiducial fraction, as shown by figure 52, is usually between $60-80\%$ in this measurement. To confirm that a high fraction of fiducial fakes causes the unfolding bias, these are subtracted manually from the MC predictions of nominal and toy distributions. The MC-toy-based unfolding bias estimate is repeated and figure 66 shows the resulting unfolding bias in each bin of m_{jj} in the VBS-Enhanced region. Compared to the nominal unfolding bias shown in figure 58, figure 66 has a smaller bias in each bin. Moreover, the bias converges to zero at a higher rate.

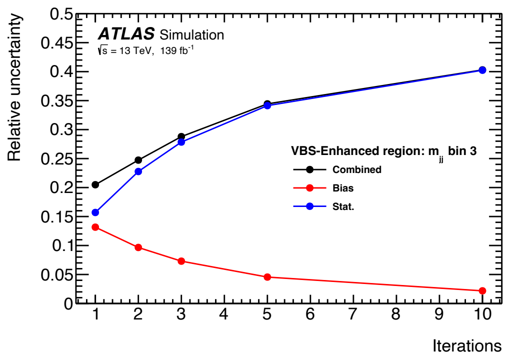
The differential measurements of the $ZZ(\rightarrow 4\ell)jj$ process are statistically limited, so it is impossible to directly subtract the fiducial fakes from data without increasing both statistical and systematic uncertainties from the fiducial fake estimate. However, it is imperative to understand the origin and topology of the fiducial fake events to reduce their impact without degrading the unfolding procedure's performance. Figure 67 shows the fake fraction, the fraction of detector-level events passing detector-level selection but failing the particle-level selection as a function of p_T (left) and η (right) of the leading and the sub-leading jets. More significant fractions of fakes are observed in low-p_T and high η region, which is likely related to the worse resolution of jet reconstruction in low-p_T and smaller efficiency of fJVT tagging in forward regions. The large fraction of fiducial fakes is understood to originate either from migrations outside the fiducial volume due to jet resolution effects or from wrongfully selecting events with pile-up jets. More stringer kinematic selections were applied to the leading and sub-leading jets in an attempt to reduce the bias, but this resulted in the degradation of unfolding performance due to low statistics. Therefore, the nominal bias shown in Section 18.3 was deemed optimal.



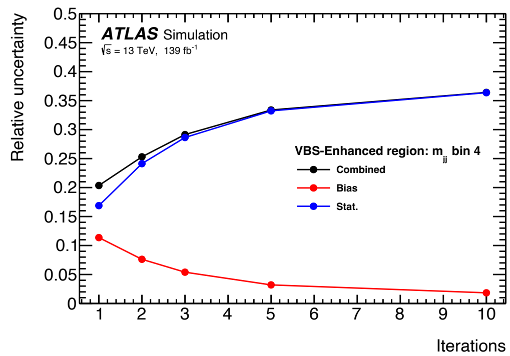
(a) bin 1



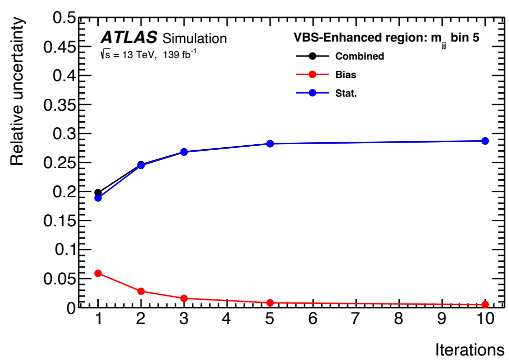
(b) bin 2



(c) bin 3

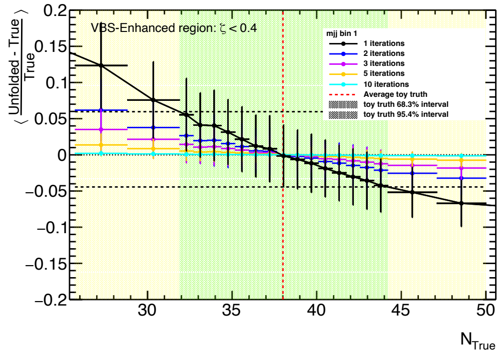


(d) bin 4

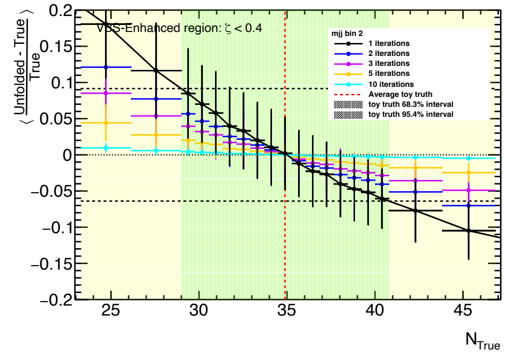


(e) bin 5

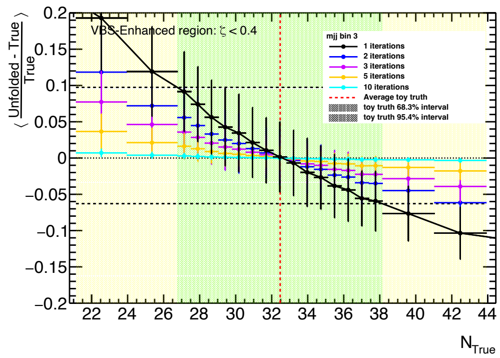
Figure 65: MC-toy-based unfolding bias and statistical uncertainty as a function of several numbers of iterations in each bin of m_{jj} distribution in the VBS-Enhanced region.



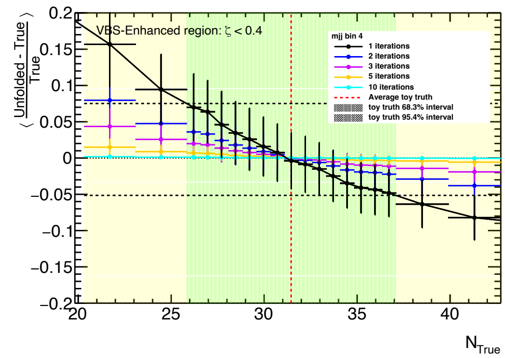
(a) bin 1



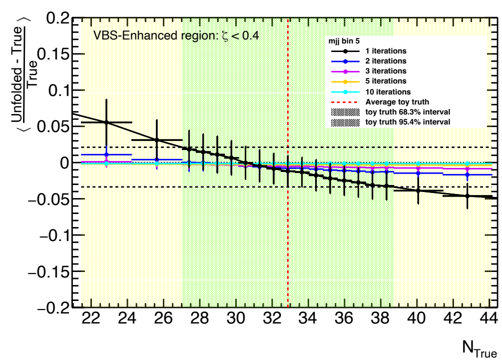
(b) bin 2



(c) bin 3

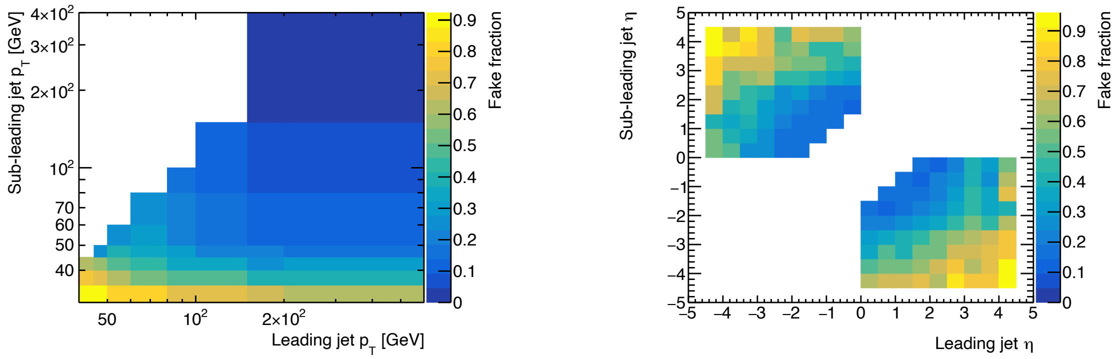


(d) bin 4



(e) bin 5

Figure 66: MC-toy-based unfolding bias in each bin of m_{jj} in the VBS-Enhanced region using Gaussian toys after subtracting the contribution of the fiducial fake events from both nominal and toy MC predictions.



(a) p_T (b) η

Figure 67: Fraction of fiducial fake events as a function of p_T & η of the leading and the sub-leading jets in the VBS-Enhanced region.

Table 17: Breakdown of the relative systematic uncertainties (%) for each bin of m_{jj} in the VBS-Suppressed region.

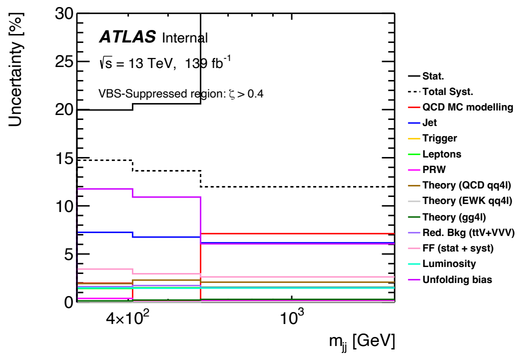
Bin m_{jj} [GeV]	[300, 410)	[410, 600)	[600, 1780)
QCD MC modelling	2	0.16	**7.1**
Jet	**7.3**	**6.8**	**6.2**
Trigger	0.03	0.06	0.08
Leptons	1.4	1.5	1.6
PRW	0.39	0.06	0.16
Theory ($qqZZ$)	1.9	2.3	2.1
Theory (EWK $qqZZjj$)	0.02	0.02	0.04
Theory ($ggZZ$)	0.13	0.22	0.28
MC Bkg. (ttV+VVV)	1.6	1.7	1.6
Fake Bkg. (stat + syst)	3.4	2.9	2.6
Luminosity	1.5	1.4	1.4
Unfolding Bias	**12.1**	**11.0**	**6.1**
Total	14.7	13.6	12.0

C VBS Suppressed Region

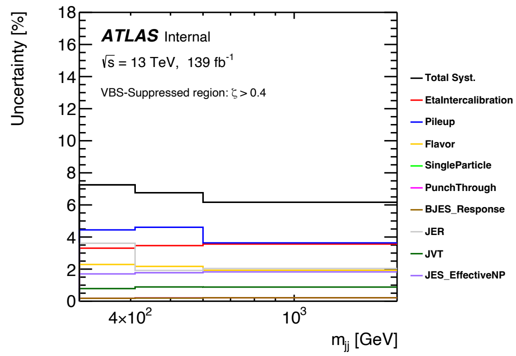
This section summarizes the results of the detector level yield and the unfolded differential cross-sections measured in the VBS-Suppressed region. The systematics affecting these results are also discussed briefly.

C.1 Systematics

The same systematic uncertainties discussed in Chapter V also impact the measurements in the VBS-Suppressed region. Table 17 shows the impact of several systematic uncertainties on the unfolded differential cross-sections in each bin of m_{jj} for the VBS-Suppressed region. Like the VBS-Enhanced region, unfolding bias followed by the jet-related uncertainties and QCD MC modeling uncertainty are the most significant sources of systematic uncertainties. Figure 68a shows the impact of the statistical and different systematic uncertainties on the unfolded cross-sections as a function of m_{jj}. The unfolded cross-sections are statistically limited. Figure 68b shows the breakdown of the jet-related uncertainties, where uncertainties related to the punch through calibration are the dominant source.



(a) Statistical and systematic uncertainties are affecting the differential cross-section measurements.



(b) Breakdown of the total jet-related uncertainties.

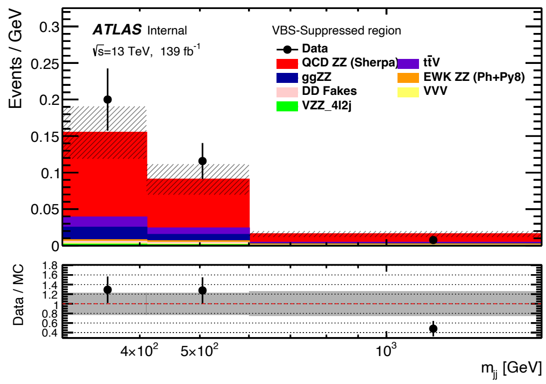
Figure 68: Uncertainties as a function of m_{jj} in the VBS-Suppressed region.

C.2 Detector Level Measurements

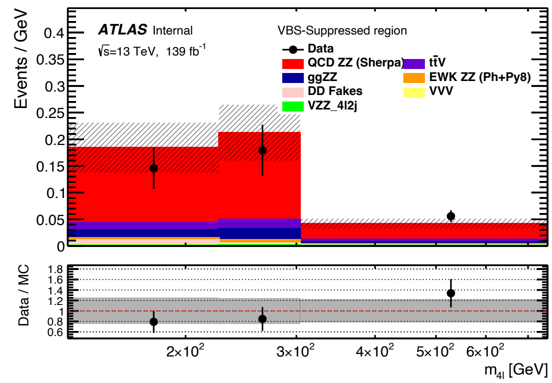
Figures 69 and 70 show the measured data yield compared to the SM detector-level predictions as a function of the eleven kinematic observables in the VBS-Suppressed region. Distributions are statistically limited, and the impact of theoretical and experimental systematic uncertainties ranges from about 20 to 30% depending on bins and distributions. Similar to the VBS-Enhanced region, a χ^2/NDF for each distribution is also reported. These values suggest a good agreement between the measured data and the SM prediction.

C.3 Unfolded Cross-sections

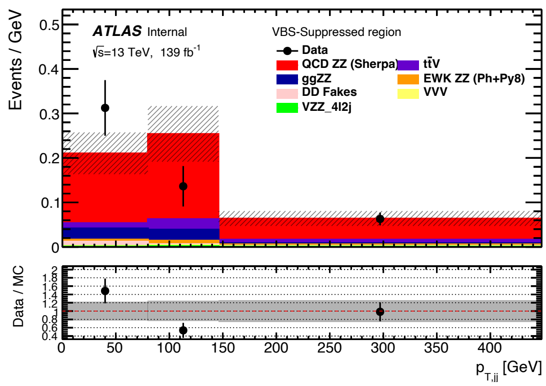
Figures 71 and 72 show the measured unfolded differential cross-sections as a function of the eleven kinematic observables in the VBS-Suppressed region. The unfolded differential cross-sections (black) are compared with two different state-of-the-art SM particle-level predictions, first, where the QCD $qqZZjj$ is generated by SHERPA (red) and second, where the QCD $qqZZjj$ is generated by MADGRAPH (blue). The vertical-solid line in the black distribution shows the statistical uncertainty of the unfolded cross-sections, and the black band represents the total systematic uncertainties. Similarly, the dashed-red and dashed-blue bands represent the fiducial uncertainties on the SHERPA and MADGRAPH particle



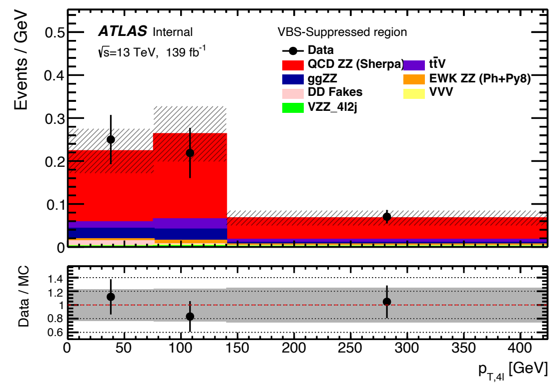
(a) m_{jj}: $\chi^2/NDF = 1.78$



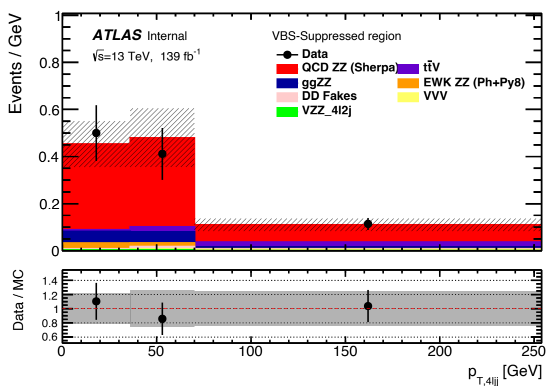
(b) $m_{4\ell}$: $\chi^2/NDF = 0.88$



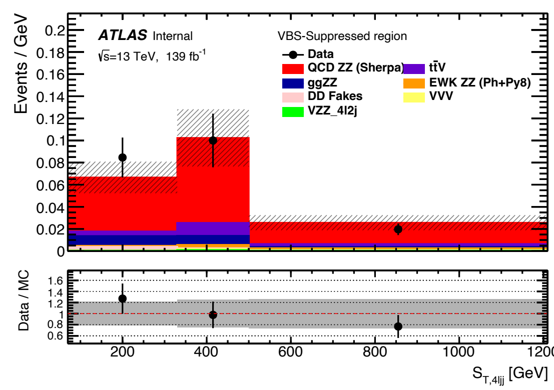
(c) $p_{T,jj}$: $\chi^2/NDF = 2.0$



(d) $p_{T,4\ell}$: $\chi^2/NDF = 0.20$



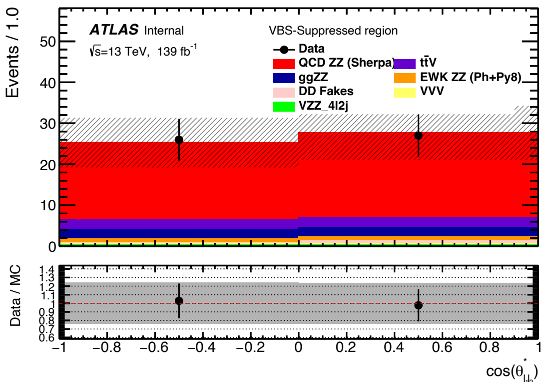
(e) $p_{T,4\ell jj}$: $\chi^2/NDF = 0.15$



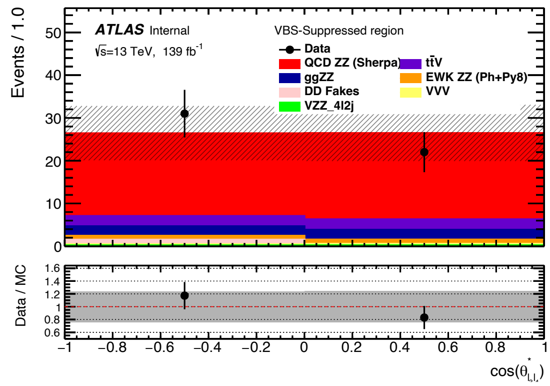
(f) $s_{T,4\ell jj}$: $\chi^2/NDF = 0.58$

Figure 69: Detector level distributions in the VBS-Suppressed region.

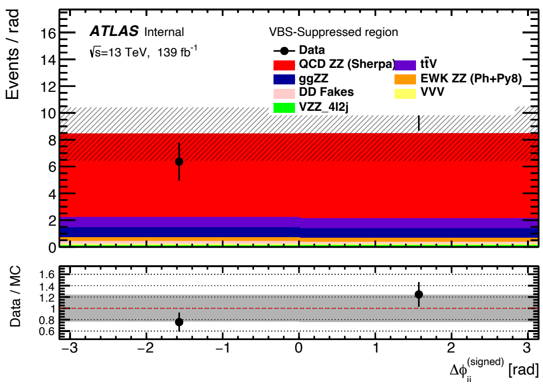
level yield, respectively. Similar to the VBS-Enhanced region, the p-values comparing the unfolded cross-sections with the two truth level yields are also reported. The p-values larger than 0.05 and the ratios of particle level yields to the unfolded-data yields suggest no statistical deviations of the differential cross-section from the SM predicted values. Therefore, in



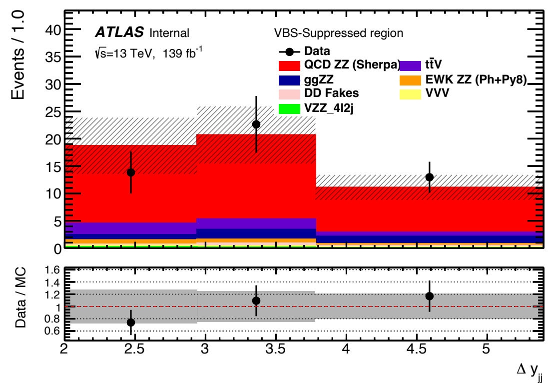
(a) $\cos\theta^{*}_{\ell 1\ell 2}$: $\chi^2/NDF = 0.01$



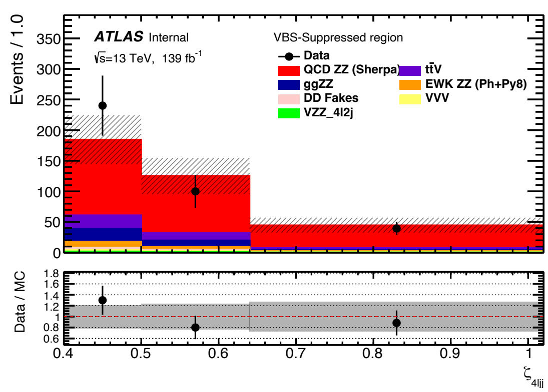
(b) $\cos\theta^{*}_{\ell 3\ell 4}$: $\chi^2/NDF = 0.60$



(c) $\Delta\phi_{jj}^{signed}$: $\chi^2/NDF = 1.26$



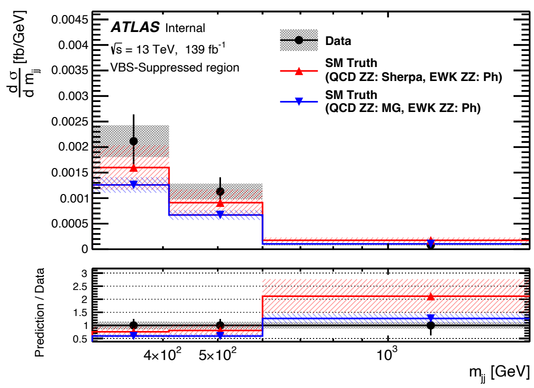
(d) Δy_{jj} : $\chi^2/NDF = 0.55$



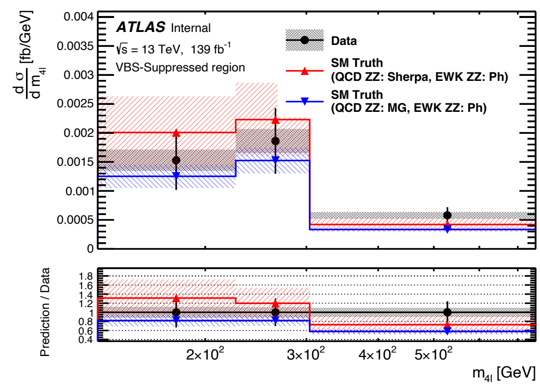
(e) ζ : $\chi^2/NDF = 0.60$

Figure 70: Detector level distributions in the VBS-Suppressed region.

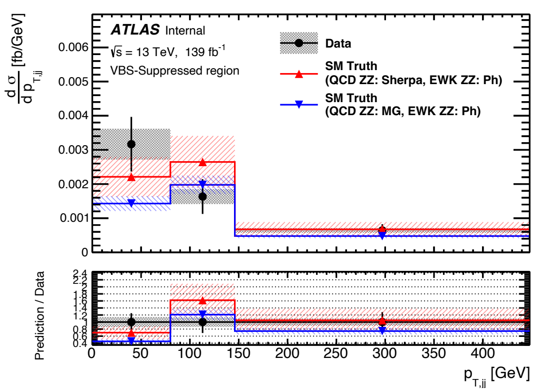
the analyzed LHC Run-2 dataset, for the $ZZ(\rightarrow 4)\ell jj$ process, all differential cross-sections in the VBS-Suppressed region are concluded to agree with the SM predictions.



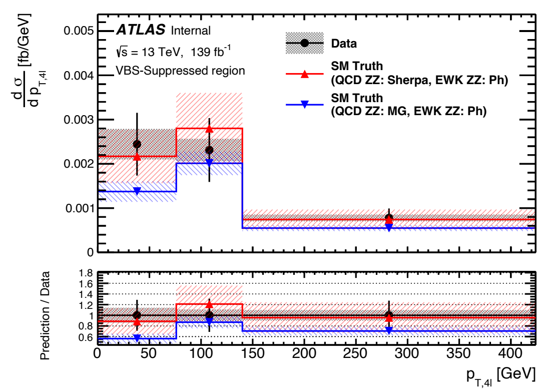
(a) m_{jj}: p-value(Sherpa=0.15 & MG=0.08)



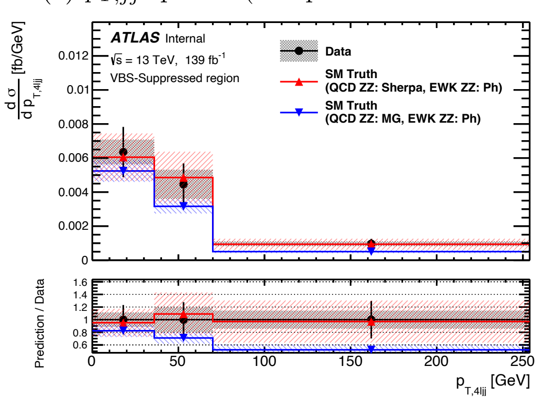
(b) $m_{4\ell}$: p-value(Sherpa=0.40 & MG=0.15)



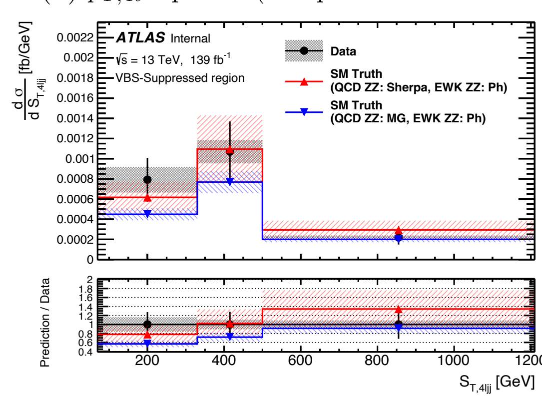
(c) $p_{T,jj}$: p-value(Sherpa=0.24 & MG=0.02)



(d) $p_{T,4\ell}$: p-value(Sherpa=0.82 & MG=0.40)

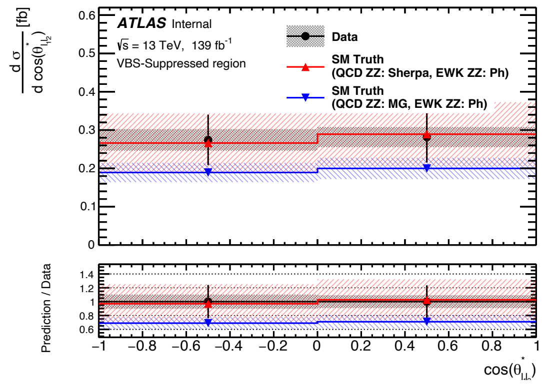


(e) $p_{T,4\ell jj}$: p-value(Sherpa=0.96 & MG=0.08)

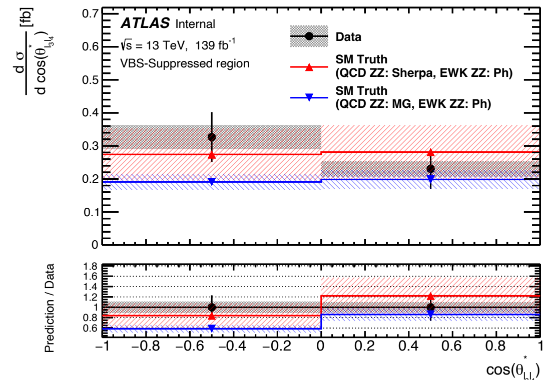


(f) $s_{T,4\ell jj}$: p-value(Sherpa=0.52 & MG=0.11)

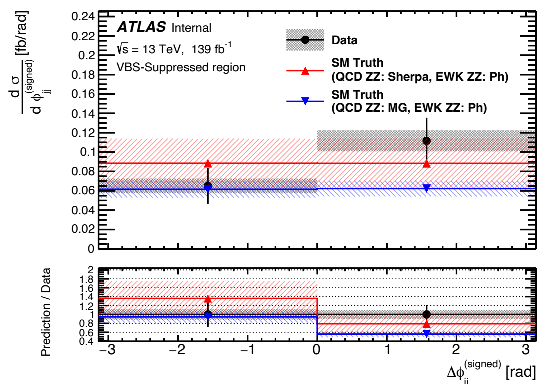
Figure 71: Unfolded differential cross-sections in the VBS-Suppressed region.



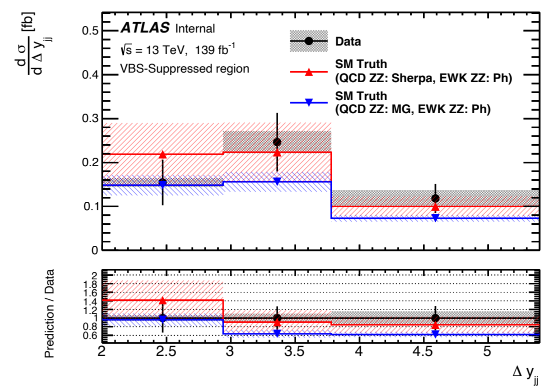
(a) $\cos\theta^*_{\ell 1\ell 2}$: p-value(Sherpa=0.90 & MG=0.81)



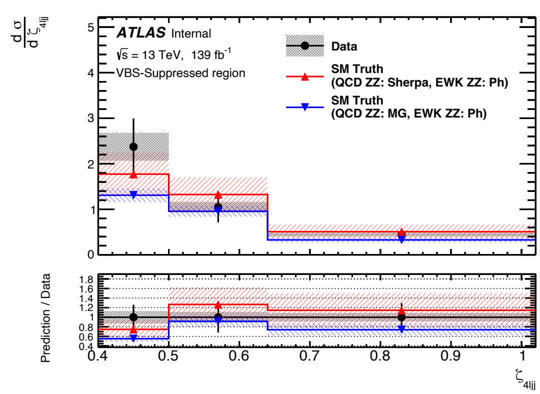
(b) $\cos\theta^*_{\ell 3\ell 4}$: p-value(Sherpa=0.43 & MG=0.41)



(c) $\Delta\phi_{jj}^{signed}$: p-value(Sherpa=0.26 & MG=0.02)



(d) Δy_{jj} : p-value(Sherpa=0.57 & MG=0.32)



(e) ζ : p-value(Sherpa=0.48 & MG=0.31)

Figure 72: Unfolded differential cross-sections in the VBS-Suppressed region.