Measurement of the production cross section of jets in association with a Z boson in 8 TeV proton-proton collisions with the ATLAS detector

Rodger Mantifel

Department of Physics

McGill University

Montreal, Quebec

December, 2016

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

DEDICATION

This work is dedicated to my parents.

ACKNOWLEDGEMENTS

I would like to give special thanks to my supervisor Prof. Francois Corriveau, who not only helped immensely with this thesis, but also supported me throughout my (entirely too long) graduate studies. I am very grateful to the people of the McGill ATLAS group and the Department of Physics for the friendly and supportive work environment. Specifically, I would like to thank Mark Stocton, Camille Bélanger-Champagne, and Michael Stoebe for their valuable guidance and advice. I would also like to give a special thanks to Ulla Blumenschein of the ATLAS collaboration who helped develop this analysis. Finally, I wish to thank my family and friends, especially my parents, for their support throughout these years.

ABSTRACT

This thesis presents the measurement of the production cross sections of jets in association with a Z boson in proton-proton collisions at $\sqrt{s} = 8$ TeV with an integrated luminosity of 20.3 fb^{-1} recorded by the ATLAS experiment at the Large Hadron Collider at CERN. Inclusive and differential $Z/\gamma^*(\rightarrow e^+e^-)$+jets cross sections are measured for jets with a transverse momentum $p_T^{jet} > 30$ GeV and rapidity $|y^{jet}| < 4.4$. Z boson candidates are identified by their decay into electron pairs. In order to stringently probe the modelling of Z/γ^*+jets production by Monte Carlo generators in regions of high-p_T phase space typical for Higgs boson decay and searches for new physics, events with high-p_T final states are investigated. High-p_T final states are defined as events with either $p_T^Z > 300$ GeV or $p_T^{leading\ jet} > 300$ GeV. A set of four observables, sensitive to the topology of the event, are measured for the standard event selection as well as in the regions of high-p_T phase space. The results from data are unfolded to particle level and compared to calculations of the matrix element from Monte Carlo generators Alpgen+Pythia and Sherpa interfaced to a parton shower. The results show that Sherpa tends to model Z/γ^*+jets production in high-p_T regions of phase space much better than Alpgen+Pythia.

RÉSUMÉ

Cette thèse présente la mesure des sections efficaces de production de jets associés à un boson Z lors de collisions proton-proton à $\sqrt{s} = 8$ TeV avec une luminosité intégrée de 20.3 fb^{-1} enregist ée par l'expérience ATLAS au Grand Collisionneur Hadronique (LHC) du CERN. Les sections efficaces inclusives et différentielles $Z/\gamma^*(\rightarrow e^+e^-)$+jets sont mesurées pour des jets ayant une quantité de mouvement transversale $p_T^{jet} > 30$ GeV et une rapidité $|y^{jet}| < 4.4$. Les bosons Z sont identifiés par leur désintégration en une paire électron-positron. Afin d'explorer de façon stricte les modèles théoriques de production Z/γ^*+jets avec des générateurs Monte Carlo dans des régions d'espace de phase à grand p_T typiques pour la désintégration du boson de Higgs et pour les quêtes vers de la nouvelle physique, des événements avec des états finaux à grand p_T sont étudiés. Ces événements sont définis comme ayant soit $p_T^Z > 300$ GeV, soit $p_T^{1^{st}jet} > 300$ GeV. Un ensemble de 4 observables, sensibles à la topologie d'un événement, sont mesurés pour une sélection standard d'événements et pour la région spécifique d'espace de phase à grand p_T. Les distributions de ces données du détecteur sont déconvoluées au niveau des particules et comparés aux calculs d'éléments de matrice des générateurs Monte Carlo Alpgen+Pythia ainsi que Sherpa, tel qu'interfacé à la gerbe de parton. Les résultats démontrent que Sherpa tend à reproduire la production Z/γ^*+jets dans les régions à grand p_T de l'espace de phase beaucoup mieux qu'Alpgen+Pythia.

AUTHOR'S CONTRIBUTION

This analysis was performed on data collected by the ATLAS detector, which depends on the work of thousands of physicists and engineers for its successful operation. The reconstruction and calibration of all physics objects, including electrons and jets, is carried out by dedicated software developed by the ATLAS collaboration that relies on countless studies of detector performance conducted over the years. All simulated Monte Carlo event samples used in this analysis are generated internally by the ATLAS collaboration. Recommendations for implementing specific event selections are provided by various internal working groups. The author of this thesis is the sole analyzer of the measurements presented here. The original work by the author begins with the data-driven estimate of the multijet background presented in Section 6.3.3, and continues with the measurement of detector level distributions in Chapter 7, the unfolding of detector effects in Chapter 8, and a study of systematic uncertainties in Chapter 9. The analysis framework built by the author is original, and is not based on any software framework used by previous analyses of the ATLAS experiment.

TABLE OF CONTENTS

DEDICATION i

ACKNOWLEDGEMENTS ii

ABSTRACT iii

RÉSUMÉ iv

AUTHOR'S CONTRIBUTION v

1 Introduction 1

2 Theory 5

2.1 The Standard Model of Particle Physics 5
2.1.1 Quantum Chromodynamics 9
2.1.2 Electroweak Theory 10
2.2 Theory of Proton-Proton Collisions 11
2.2.1 The Quark Parton Model 14
2.2.2 Parton Distribution Functions 15
2.2.3 Cross Sections and the Factorization Theorem 15
2.2.4 Luminosity 17
2.2.5 Underlying Events 19
2.2.6 QCD Radation 19
2.2.7 Hadronization 20
2.3 Z/γ^*+jets production 21
2.3.1 Z Boson Production and Decay 22
2.3.2 Jet Production in Association with a Z Boson 23
2.3.3 Jet Scaling Patterns 25
2.3.4 Background Processes to Dilepton Final States 27
2.3.5 Experimental Measurements of Z/γ^*+jets Production 28

3 Experimental Setup 33

3.1 The Large Hadron Collider 33
3.2 The ATLAS Detector 38
3.2.1 Magnet System 40
3.2.2 Inner Detector and Tracking 41
3.2.3 Calorimetry 43
3.2.4 Muon System 48
3.2.5 Trigger and Data Acquisition System 49

4 Monte Carlo Simulation and Event Generators . . . 52

4.1 Pythia . . . 54
4.2 Herwig . . . 55
4.3 Alpgen . . . 55
4.4 Sherpa . . . 55
4.5 Powheg . . . 56

5 Electron and Jet Definitions . . . 57

5.1 Electrons . . . 57
5.1.1 Reconstruction . . . 58
5.1.2 Identification . . . 59
5.1.3 Calibration . . . 61
5.2 Jets . . . 67
5.2.1 Reconstruction . . . 69
5.2.2 Calibration . . . 71
5.2.3 Quality Criteria . . . 75
5.2.4 Performance . . . 77

6 Data Sets and Event Selection . . . 82

6.1 Data . . . 82
6.2 Event Selection . . . 83
6.3 Signal and Background Modelling . . . 84
6.3.1 Signal Monte Carlo . . . 84
6.3.2 Background Monte Carlo . . . 87
6.3.3 Data-Driven Estimate of Multijet Background . . . 87

7 Detector Level Measurements . . . 90

7.1 Detector Level Distributions . . . 91

8 Unfolding . . . 97

8.1 Bayes Unfolding Method . . . 98
8.2 Unfolding of Detector Effects . . . 100

9 Systematic Uncertainties . . . 108

9.1 Electron Selection . . . 109
9.2 Jet Energy Scale . . . 109
9.3 Jet Energy Resolution . . . 110
9.4 Pile-up Jet Removal . . . 112
9.5 Background Modelling and Normalization . . . 112
9.6 Unfolding . . . 114
9.7 Total Systematic Uncertainty . . . 117

10 Results . . . 120

10.1 Inclusive and Exclusive Jet Multiplicities . . . 121
10.2 Transverse Momentum Distributions . . . 123
10.3 Transverse Momentum Ratios . . . 125
10.4 Jet Rapidity . . . 128
10.5 Angular Separation Distributions . . . 130
10.6 Discussion . . . 133

11 Conclusions and Outlook . . . 139

Appendix . . .

A Detector Level Distributions . . . 142

B Unfolding Performance . . . 148

C Systematic Uncertainties . . . 169

D Glossary of Abbreviations . . . 190

References . . . 192

CHAPTER 1

Introduction

The study of the production of jets association with a Z boson is important for a variety of reasons. Measurements of the kinematics of $Z/\gamma^* + \text{jets}$ events provide a stringent test of the theory of the strong interaction, called Quantum Chromodynamics (QCD), at energy scales where the strong coupling is small enough for perturbation theory to be applied. Furthermore, these events form non-negligible backgrounds in studies of top quark pairs, diboson events, Higgs boson production, as well as searches for new physics, where the multiplicity and kinematics of jets in $Z/\gamma^* + \text{jets}$ events are used to achieve a significant separation of signal from background. This often results in final states containing highly energetic jets or boosted Z bosons, which introduce scales larger than the mass of the Z boson.

The kinematics of $Z/\gamma^* + \text{jets}$ can be predicted using fixed-order calculations at next-to-leading order (NLO) in perturbative QCD (pQCD) [1, 2] and from Monte Carlo (MC) generators based on calculations at leading order (LO) or NLO of the matrix element (ME) supplemented by parton showers (PS) [3–5]. Predictions from MC generators based on LO calculations are affected by large scale uncertainties and need to be tuned and validated using data, which has typically been done using data of low transverse momenta (p_T) and moderate jet multiplicity (N_{jet}). However, when moving to regions of high-p_T phase space, as is expected for Higgs boson decay and searches for new physics where we measure highly energetic jets and boosted Z bosons, the $Z/\gamma^* + \text{jets}$ topology is expected to change drastically. MC generators may fail to predict the kinematic in these regimes. Furthermore, in higher energy regimes, large logarithmic contributions to NLO predictions from pQCD are expected due to the appearance of kinematic topologies not present at LO in which a soft or

collinear Z boson is radiated from a quark line [6]. Finally, an increasing impact of higher-order electroweak (EW) corrections is expected with increasing transverse momentum of the Z boson or jets [7], which act to diminish the cross sections.

Measurements of the production cross section of jets in association with a Z boson have been performed by the CDF [8] and DØ [9] collaborations at the Tevatron for proton-antiproton collisions at a center-of-mass energy $\sqrt{s} = 1.96$ TeV. More recent results have been obtained by the ATLAS [10] and CMS [11, 12] collaborations at the LHC in 4.6 fb^{-1} of proton-proton collisions collected at $\sqrt{s} = 7$ TeV in 2011. The previous measurements from the LHC, as will be described later in Section 2.3.5, show good agreement in general with theoretical predictions from MC event generators based on LO calculations of the ME interfaced with PS. However, some discrepancies between measurements from data and predictions from MC of the p_T and rapidity spectra of Z bosons and jets has been observed in the analysis of the 7 TeV data set by the ATLAS collaboration.

The measurements made by this analysis are meant to determine if discrepancies observed in the previous analysis by the ATLAS experiment persist in the 8 TeV data set. Additionally, this analysis extends the measurements of the previous analysis of the 7 TeV data set by the ATLAS experiment into regions of high-p_T phase space using the much larger data set of 20.3 fb^{-1} of proton-proton collisions collected at $\sqrt{s} = 8$ TeV by the ATLAS detector in 2012. Selected events contain a Z boson decaying into a pair of electrons. Associated jets are identified in a rapidity range of $|y^{jet}| > 4.4$ and with transverse momentum $p_T^{jet} > 30$ GeV. Furthermore, certain phase space constraints are imposed that are designed to test the modelling of $Z/\gamma^* +$ jets production in regions of high-p_T phase space by the most important ME+PS generators used by the ATLAS experiment to model vector boson plus jets production: Alpgen [3] and Sherpa [4]. Alpgen calculates the matrix elements at LO in QCD for up to five jets in the final state, and is interfaced to Pythia [13] for the parton shower in this analysis. Sherpa calculates the inclusive cross section at NLO in QCD, while higher jet multiplicities with up to five jets in the final state are calculated at LO.

Cross sections are measured as a function of the inclusive and exclusive jet multiplicities (N_{jet}). Differential cross sections are measured as a function of the jet transverse momentum (p_T^{jet}), the jet rapidity (y^{jet}), and the transverse momentum of the Z boson (p_T^Z). Furthermore, differential cross sections are measured for a set of observables in events with at least two jets that are sensitive to the topology of the event. These are the ratio of the p_T of the sub-leading jet to the p_T of the leading jet (Rp_T^{jj}), the ratio of the p_T of the Z boson to the p_T of the leading jet (Rp_T^{Zj}), the absolute azimuthal separation between the two leading jets ($\Delta\phi^{jj}$), and the minimum $|\Delta\phi|$ between the Z boson and either of the two leading jets in the event ($min\Delta\phi^{Zjj}$). The set of observables which are sensitive to the topology of the event, and are meant to test the modelling of MC in high-p_T regions of phase space, are measured for events with high-p_T final states in addition to the standard event selection. High-p_T final states are defined as events with either $p_T^Z > 300$ GeV or $p_T^{leading\ jet} > 300$ GeV. The results of the measurements from data are unfolded for detector effects and quoted at the particle (hadron) level, where they are compared to predictions from the MC generators, Alpgen+Pythia and Sherpa.

This thesis begins with an introduction to the Standard Model of particle physics and a brief description of the phenomenology of hadronic collisions and Z boson production in Chapter 2. Chapter 3 provides a description of the LHC and the ATLAS detector. Chapter 4 explains the MC simulation chain and provides a general description of the different MC generators used in this analysis. An overview of the reconstruction algorithms and calibration methods for electrons and jets is given in Chapter 5. The original work of this thesis begins in Chapter 6 where the details of the 8 TeV data set and the individual MC samples are given, along with a description of the event selection requirements. Detector level measurements and their comparisons with predictions from MC generators are presented in Chapter 7. The unfolding procedure and evaluation of systematic uncertainties are detailed in Chapters 8 and 9 respectively. The results of the unfolded measurements and their comparisons with particle

level predictions from MC generators are presented in Chapter 10. An overall conclusion and outlook is given in Chapter 11.

CHAPTER 2

Theory

2.1 The Standard Model of Particle Physics

Particle physics is the science behind the fundamental building blocks of matter and the interactions between them. It is described by the theory of the Standard Model (SM) of particle physics [14]. The SM is a local Lorentz invariant quantum field theory in which the properties of the interactions between fundamental point-like particles result from the requirement of local gauge invariance. Mathematically, the SM can be described by group theory using the local symmetry group that is the direct product of

$$\mathrm{SU}(3) \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1), \tag{2.1}$$

where $\mathrm{U}(n)$ is the unitary group of degree n, and $\mathrm{SU}(n)$ is the special unitary group of degree n.

Within the SM, all known matter is comprised of a few elementary particles, called fermions, with intrinsic spin-$\frac{1}{2}$. The interactions between fermions are governed by forces mediated by spin-1 gauge bosons. Mass is acquired by particles in the SM via the Higgs mechanism, which requires the existence of an additional heavy particle of spin-0, called the Higgs boson. Figure 2.1 shows a representation of the known particles contained within the SM. Each particle is characterized by its mass, spin, and charge.

The SM describes the forces between particles as the exchange of spin-1 gauge bosons. The strong force is carried by gluons, the weak force is mediated by the charged $W^{\pm}$ and neutral Z^0 bosons, and the electromagnetic (EM) force is governed by the exchange of photons

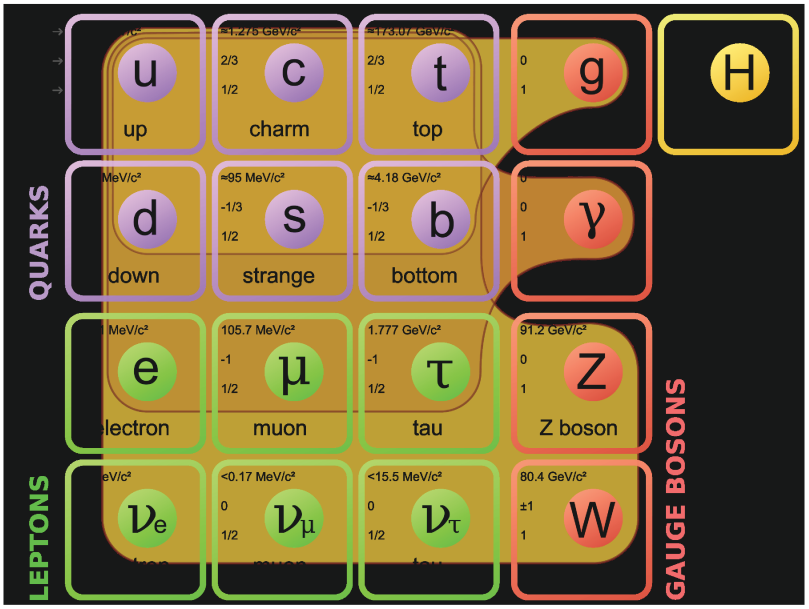


Figure 2.1: Fundamental particles of the Standard Model of particle physics [15].

(γ). The experimentally measured properties of the gauge bosons, as well as the forces they mediate, are detailed in Table 2.1. Gravity is the only known fundamental force that is not described by the SM.

Gauge Boson	Mass	Q	Spin	Force
gluon (g)	0	0	1	Strong
$W^{\pm}$	80.3850±0.0150 GeV	±1	1	Weak
Z^0	91.1876±0.0021 GeV	0	1	Weak
γ	$< 1 \cdot 10^{-18}$ GeV	0	1	Electromagnetic

Table 2.1: Standard Model gauge bosons and their properties [14].

Fermions are spin-$\frac{1}{2}$ particles that are further categorized into two sub-groups of particles, the leptons and the quarks. Each group has three generations, in which mass increases from generation to generation.

Leptons are classified according to their electric charge (Q), the lepton flavour number (electron number L_e, muon number L_μ, tau number L_τ), and the third component of the weak isospin (T_3). An additional quantity, known as the weak hypercharge (Y_W), is related to the

charge and the third component of the weak isospin by

$$Y_W = 2 \cdot (Q - T_3). \tag{2.2}$$

The properties of the leptons are summarized in Table 2.2. The charged leptons interact via the weak and EM forces, whereas the neutral leptons (the neutrinos) interact via the weak force only.

Lepton Flavour	Mass	Q	Y_W	L_e	L_μ	L_τ
ν_e	< 2 eV	0	-1	1	0	0
e^-	0.511 MeV	-1	-1	1	0	0
ν_μ	< 0.19 MeV	0	-1	0	1	0
μ^-	105.7 MeV	-1	-1	0	1	0
ν_τ	< 18.2 MeV	0	-1	0	0	1
τ^-	1776.8 MeV	-1	-1	0	0	1

Table 2.2: Standard Model leptons and their properties [14].

Quarks are classified according to their electric charge (Q), and their flavour quantum numbers (third component of the isospin (I_3), strangeness (S), charmness (C), bottomness (B), and topness (T)). Analogous to the electric charge of the electromagnetic force, quarks also carry a colour charge, which is the charge of the strong force. Colour charge can take on three different values: red, green, and blue. Consequently, quarks can interact via the strong force in addition to the weak and EM forces. The properties and quantum numbers of the quarks are summarized in Table 2.3.

Each fermion has a corresponding antiparticle with the same mass but with reversed signs of all other quantum numbers. Including the complete set of different colour charges for each quark, this leads to a total of 12 leptons and 36 quarks described by the SM.

Within the SM, the EM and weak forces are unified into a single electroweak (EW) interaction. The theories behind the strong and EW interactions of the SM are gauge theories

Quark Flavour	Mass	Q	I_3	C	S	T	B
u	$2.3^{+0.7}_{-0.5}$ MeV	2/3	1/2	0	0	0	0
d	$4.8^{+0.5}_{-0.3}$ MeV	−1/3	−1/2	0	0	0	0
c	1.275 ± 0.025 GeV	2/3	0	1	0	0	0
s	95 ± 5 MeV	−1/3	0	0	−1	0	0
t	173.21 ± 0.9 GeV	2/3	0	0	0	1	0
b	4.18 ± 0.03 GeV	−1/3	0	0	0	0	−1

Table 2.3: Standard Model quarks and their properties [14].

whose properties are the result of the application of local symmetry principles, i.e. the invariance of the Lagrange density function under gauge transformations. Predictions have been tested to a high level of accuracy, such that the SM is one of the most successful theories ever developed. However, within the electroweak theory, masses for the gauge bosons of the weak interaction violate invariance under local gauge transformations. Nonetheless, the weak gauge bosons, $W^{\pm}$ and Z, have large masses, as confirmed by experiment. The most popular and minimal solution to this problem is through the Higgs mechanism [16–21]. The Higgs mechanism leaves the fundamental symmetry of the electroweak theory unchanged and generates the masses of the weak gauge bosons, while leaving the photon massless, by spontaneous symmetry breaking of the quantum vacuum ground state. The masses of the W and Z bosons depend directly on the vacuum expectation value, allowing the vacuum expectation value to be determined by measuring both masses. However, the Higgs boson mass cannot be calculated from the vacuum expectation value since it also depends on a free parameter in the SM and must be determined by experiment. In 2012 a new boson with a mass around 125 GeV has been discovered [22, 23], which is consistent with the SM Higgs boson. Masses for the fermions are generated through Yukawa-like couplings of the fermions to the Higgs field.

A more mathematical description of the theories behind the strong and EW interactions is given in the following sections.

2.1.1 Quantum Chromodynamics

The theory of the strong interaction, called Quantum Chromodynamics (QCD) [14, 24], is a non-Abelian gauge theory based on a local SU(3) symmetry group of colour. The eight generators of the group correspond to the eight massless gluons, which mediate the interaction of coloured quarks. Each quark is a triplet of the QCD gauge group

$$q_f^T \equiv (q_f^1, q_f^2, q_f^3) \tag{2.3}$$

with 1, 2, 3 representing the 3 primitive colour charges, called red (r), green (g), and blue (b).

The Lagrangian density of QCD is given by

$$L_{QCD} = \sum_{j=1}^{n_f} \overline{q}_j (iD_\mu \gamma^\mu - m_j) q_j - \frac{1}{4} \sum_{A=1}^{8} F^{A\mu\nu} F^A_{\mu\nu} \tag{2.4}$$

with the quark-field spinors given by q_j , and the quark masses given by m_j. The γ^μ represent the Dirac matrices. The covariant derivative that preserves the gauge symmetry of QCD is given by $D_\mu = \partial_\mu - ig_s T_A \mathcal{A}^A_\mu$, where $\mathcal{A}^A_\mu$ corresponds to the eight gluon fields, and T_A to the eight generators of the SU(3) symmetry group. The field strength tensor $F^A_{\mu\nu}$ is based on the gluon field $\mathcal{A}^A_\mu$ by

$$F^A_{\mu\nu} = \partial_\mu \mathcal{A}^A_\nu - \partial_\nu \mathcal{A}^A_\mu + g_s f_{ABC} \mathcal{A}^B_\mu \mathcal{A}^C_\nu \tag{2.5}$$

where f_{ABC} are the structure constants of the SU(3) symmetry group, and the QCD coupling constant is given by $g_s = \sqrt{4\pi\alpha_s}$.

Since SU(3) is a non-Abelian group, the third term of Equation 2.5 does not vanish and consequently the gluon fields are able to self-interact. Due to this self-interaction, the effective coupling constant of the strong interaction decreases with increasing energy, leading to asymptotic freedom. For large energy scales, which probe short distances, the strong coupling constant converges asymptotically towards zero. For energy scales on the order of the mass of the Z boson, the value of the strong coupling is $\alpha_s(M_Z^2) = 0.1181 \pm 0.0013$ [14]. In

this regime, quarks and gluons can be treated as free and their interactions can be calculated within perturbation theory. Within perturbation theory, the matrix element of the hard scattering process can be expanded as an infinite power series in the strong coupling constant α_s. At small enough distance scales, where the value of the strong coupling is much less than one, contributions from higher-order terms in α_s become less and less significant. An approximate solution to the exact matrix element calculation in perturbative QCD (pQCD) can be obtained by keeping only a finite number of the dominant, lowest order terms.

With increasing distance between two quarks, the strength of the strong coupling grows too large for perturbation theory to be applied. At these distance scales, the quarks become bounded into colourless hadrons through a process called confinement. Confinement describes the fact that colour charged objects cannot be observed individually, but are only stable when confined in colour neutral combinations called hadrons. Colourless hadrons which consist of one quark and one anti-quark are called mesons, while colourless hadrons that consist of three quarks are called baryons.

2.1.2 Electroweak Theory

The electroweak (EW) theory [14, 25–27] is the gauge theory behind the EW interaction. It describes the unification of the weak interaction with the electromagnetic interaction under the SU(2) $\otimes$ U(1) symmetry group. The SU(2) group involves three gauge fields with a coupling strength g, and the U(1) symmetry group involves one gauge field with a coupling strength g'. The corresponding gauge bosons are W^i_μ, where $i = 1, 2, 3$ for SU(2), and B_μ for U(1). The Lagrangian of the EW theory is

$$L_{EW} = \sum_{j=1}^{3} i\overline{\psi}_j(x)\gamma^\mu D_\mu \psi_j(x) - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W^j_{\mu\nu}W_j^{\mu\nu} \tag{2.6}$$

where ψ_j represents the fermionic fields of the three generations of "up" and "down" type fermions, and D_μ describes the covariant derivative that preserves the gauge symmetry of

the EW theory,

$$D_\mu = \partial_\mu - ig\frac{\sigma_j}{2}W_\mu^j(x) - ig'\frac{Y}{2}B_\mu(x). \tag{2.7}$$

The fermionic part of the Lagrangian describes the kinetic energy of the fermions and their interactions, while the covariant derivative describes the interaction with the gauge field. The second part of the Lagrangian describes the gauge fields with a term for the kinetic energy and a term that describes the self interaction between the gauge fields.

The four gauge bosons of the SU(2) $\otimes$ U(1) symmetry group do not translate directly into the $W^\pm$, Z, and γ bosons. $W^\pm$ are linear combinations of W_μ^1 and W_μ^2

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \tag{2.8}$$

representing the charged part of the interaction. The two neutral fields are mixed in such a way that their mass eigenstates are

$$A_\mu = B_\mu \cos(\theta_W) + W_\mu^3 \sin(\theta_W), \tag{2.9}$$

$$Z_\mu = -B_\mu \sin(\theta_W) + W_\mu^3 \cos(\theta_W), \tag{2.10}$$

where θ_W is the so-called weak mixing angle. Hence, the photon field A_μ and the Z boson field Z_μ can be interpreted as an orthogonal combination of the two neutral gauge fields W_μ^3 and B_μ.

2.2 Theory of Proton-Proton Collisions

A description of hadronic collisions requires an understanding of the proton structure since the actual interaction takes place between the internal constituents of the proton, as well as an understanding of how the aftermath of the collision evolves into the observed final state. Figure 2.2 shows a schematic view of the sub-processes involved in the evolution of a hadronic collision.

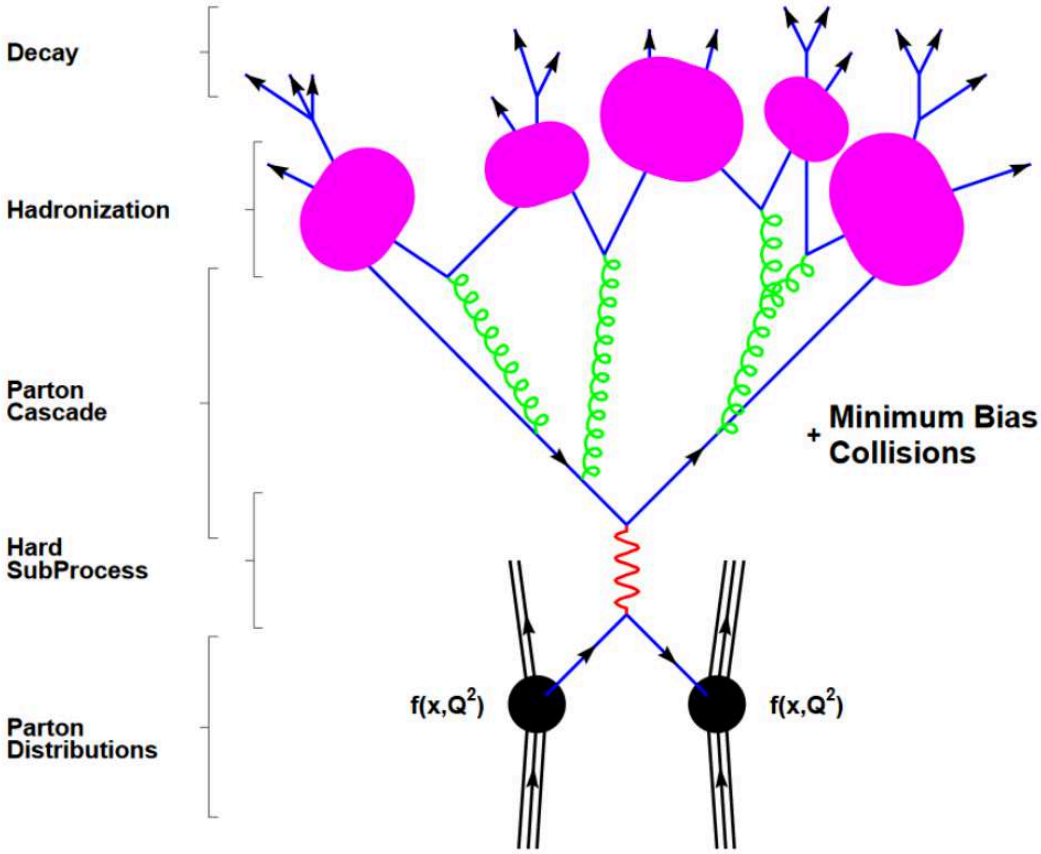


Figure 2.2: Schematic view of a hadronic collision event [28]. The collision begins with incoming partons, which are bound within hadrons and are represented as parton distributions. The hard sub-process refers to the hard interaction between partons of the colliding hadrons, which is calculable using pQCD. The incoming and outgoing partons can radiate soft quarks/gluons, which can initiate a cascading parton shower. Hadronization refers to the binding together of colour charged partons into hadrons due to colour confinement. Finally, massive particles will decay into lighter final state particles. Occurring in parallel, beneath the hard scattering event, is a myriad of soft minimum bias collisions that make up the underlying event.

Sub-processes in hadronic collisions can be classified into two types, hard and soft. For both types, the underlying theory is QCD. However, only the hard sub-process, representing the parton level interaction, can be described using fixed-order perturbation theory since the strength of the strong coupling is small at the sub-hadronic distance scales where the interaction between partons takes place . Fixed-order perturbation theory is able to provide parton level predictions at leading order (LO), next-to-leading order (NLO), and for a few processes, at next-to-next-to-leading order (NNLO) in α_S.

Soft processes occur at distance scales larger than the hadron, where the strength of the strong coupling is too large for quarks and gluons to interact freely. In order to describe soft processes, phenomenological models are needed since the properties of soft processes are dominated by non-perturbative QCD physics. The initial distributions of partons inside the colliding hadrons are governed by non-perturbative QCD physics. Additionally, the shower of soft QCD radiation from the incoming and outgoing partons, the binding of final state partons into hadrons due to colour confinement, and their subsequent decays into the lighter observed particles are all soft processes. For a complete description of the full final state, soft processes need to be combined with the calculation of the hard interaction.

Predictions of cross sections and event kinematics in hadronic collisions can be obtained from Monte Carlo (MC) event generators. MC generators calculate the matrix element of the hard partonic scattering cross section using pQCD. The parton level results are interfaced with various phenomenological models of soft physics processes in order to determine the hadron (particle) level results. The chosen methods implemented by the various models of soft QCD physics can affect the final result. The following sections describe different aspects of hadronic collisions, including the hard and soft physics sub-processes discussed above, which are necessary for a complete theoretical description of a hadronic collision.

2.2.1 The Quark Parton Model

The parton model in its original form was proposed by R. P. Feynman in 1969 [29] in order to describe the results of deep inelastic scattering (DIS) experiments in lepton-hadron collisions. Within the parton model, DIS is viewed as the sum of elastic scattering of weak probes off point-like spin-$\frac{1}{2}$ partons inside the nucleon through the exchange of a vector boson. At the heart of this model is the association of partons with the concept of quarks. The quark parton model (QPM), as it came to be known, proposed that protons are composed of three valence quarks that determine its flavour properties, surrounded by a flavour neutral sea of quark-antiquark pairs. However, it was discovered that the quarks within the proton seem to only carry about half of the total momentum of the proton. It was assumed that gluons, the mediators of the strong force, may carry this deficit in momentum. Thus, it was necessary to include gluon interactions using QCD into the framework of DIS, resulting in the QCD improved quark parton model [30]. Hence, the proton consists not only of three valence quarks, but also a sea of further quarks and gluons.

Adding QCD to the QPM had another benefit. Higher order QCD corrections to strong interactions led to the discovery of asymptotic freedom [31]. This states that as the energy of the interaction gets larger, which allows smaller distance scales to be probed, the strong coupling constant grows smaller. At the small distance scales within the proton, quarks interact almost as free particles. At lower energies, the coupling constant gets larger, and thus the strength of the interaction grows. At distance scales larger than the proton, the coupling is too large for quarks to exist on their own. This is called colour confinement.

In the context of proton-proton collisions, each proton with four-momentum P^μ is described by point-like partons (quarks or gluons), which are moving collinear to the proton, carrying a fraction of its momentum $p_i^\mu = x_i P^\mu$. The hard scattering of the hadronic collision takes place between the partons at the 4-momentum exchange Q^2, which, for example, could be the mass of the Z boson or the transverse momentum of a parton.

2.2.2 Parton Distribution Functions

In order to provide an exact description of the hard scattering process, the momentum distributions of the partons $f(x, Q^2)$, called Parton Distribution Functions (PDFs), have to be well known. PDFs give the probability density that a certain parton with momentum fraction x will be found inside the proton at the energy scale of the interaction Q^2.

Parton densities are governed by physics at long distances, where the relevant scale is that of the hadron (10^{-15} m), thus partons cannot be observed as free particles. Consequently, PDFs cannot be calculated from first principles using perturbation theory due to the inherent non-perturbative nature of partons at these energy scales. However, QCD does tell us how the PDFs evolve with the energy scale Q^2 using the DGLAP [32,33] evolution equations. The starting form of the PDF in terms of x at a starting scale Q_0^2 must be determined by fits to experimental data covering a wide range in x and Q^2. The data come from measurements of deep inelastic scattering, and from measurements of vector boson production, jet production, and photon production.

PDFs are determined at LO, NLO, and NNLO in QCD by various collaborations, such as CTEQ/CT14 [34, 35], MRST/MSTW [36–38], and NNPDF [39]. The LHAPDF [40] library provides a unified Fortran/C++ interface to all major PDF sets used in the LHC era. Figure 2.3 shows as an example the MSTW 2008 NLO PDFs for $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2.

2.2.3 Cross Sections and the Factorization Theorem

Cross sections for hadron-hadron interactions cannot be calculated directly using perturbative QCD due to the non-perturbative nature of the hadron structure at large distance scales. However, a basis for a theoretical description can be achieved using the QCD factorization theorem [41]. In a simplified interpretation, illustrated in Figure 2.4, the factorization theorem separates the short-distance parton cross section, calculable using pQCD, from the long-distance non-perturbative processes such as the parton distributions. Using the fac-

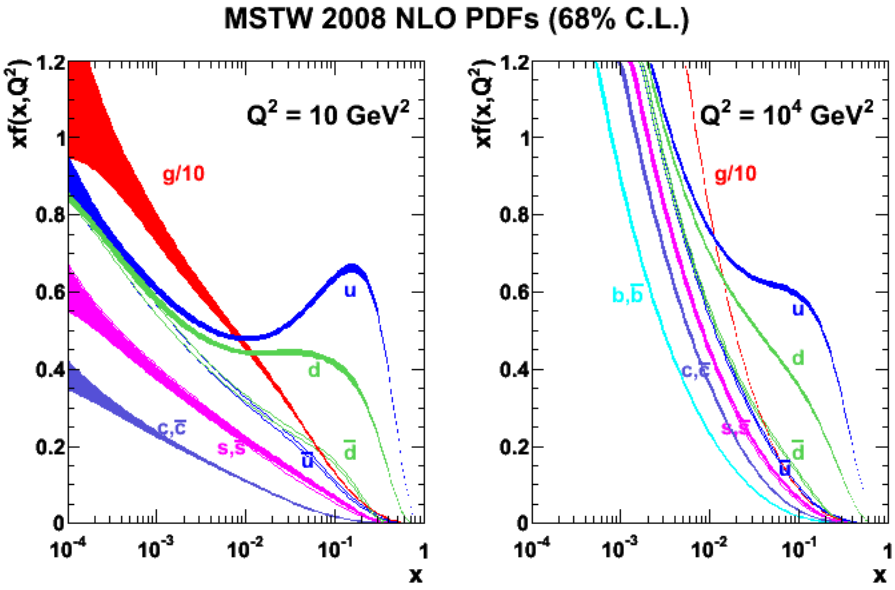


Figure 2.3: MSTW 2008 NLO PDFs for $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2 [36].

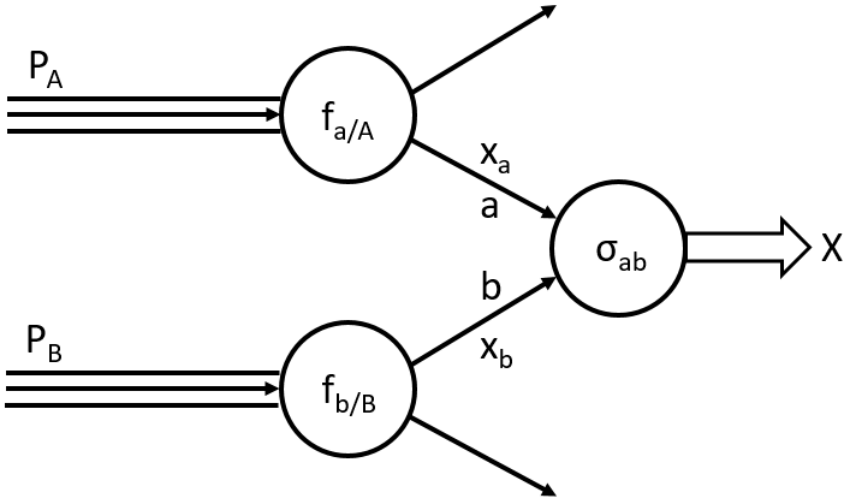


Figure 2.4: Schematic view of the factorization theorem applied to a hard scattering process in a hadronic collision. The colliding hadrons, P_A and P_B, are represented by parton distribution functions $f_{a/A}$ and $f_{b/B}$ respectively. The hard interaction takes place between partons a and b with momentum fractions x_a and x_b respectively. The factorization theorem separates the hard interaction σ_{ab} from the soft physics contained in the PDFs. Thus, the calculation of the partonic cross section does not depend on the incoming hadrons.

torization theorem, the hadronic cross section $\sigma_{AB\to X}$ of a hard-scattering process is given by

$$\sigma_{AB\to X} = \int \mathrm{d}x_a \mathrm{d}x_b f_{a/A}(x_a, Q^2) f_{b/A}(x_b, Q^2) \hat{\sigma}_{ab\to X} \tag{2.11}$$

where $\hat{\sigma}_{ab\to X}$ is the short-distance partonic cross section describing the hard-scattering of partons a and b. In order to calculate the hadronic cross section $\sigma_{AB\to X}$, the partonic cross section is convoluted with the long-distance non-perturbative physics contained in the PDFs, denoted by $f(x, Q_2)$. The partonic cross section does not depend on the incoming hadrons and can be calculated by a perturbative expansion in the strong coupling constant α_S

$$\sigma_{AB\to X} = \int \mathrm{d}x_a \mathrm{d}x_b f_{a/A}(x_a, \mu_F^2) f_{b/A}(x_b, \mu_F^2) \times [\hat{\sigma}_0 + \alpha_S(\mu_R^2)\hat{\sigma_1} + \dots]_{ab\to X} \tag{2.12}$$

where μ_F denotes the factorization scale, and μ_R denotes the renormalization scale at which the QCD coupling constant is evaluated. μ_R is introduced due to the renormalization of the QCD expansion. At higher order of the perturbative expansion, the dependence of the cross section on μ_F and μ_R decreases. Typically, the factorization scale and the renormalization scale are set equal and chosen to be at the order of the momentum scale Q of the hard scattering process.

2.2.4 Luminosity

For a scattering process with a given cross section σ, the reaction, or event, rate R is related to the instantaneous luminosity L by

$$R = \sigma \cdot L \tag{2.13}$$

In an experiment with two colliding beams of point-like particles, the instantaneous luminosity is given by

$$L = n_b \frac{n_1 \cdot n_2}{A_{eff}} f, \tag{2.14}$$

where n_b is the number of colliding bunches, n_1 and n_2 are the number of particles per bunch in each beam, and f is the collision frequency. A_{eff} is the effective interaction areas of the

colliding beams [42]. For round beams with a uniform transverse particle density, A_{eff} is given by

$$A_{eff} = 4\pi\sigma_r^2, \tag{2.15}$$

where σ_r is the Gaussian beam radius. The expected event rates and cross sections of various physics processes in proton-antiproton ($p\overline{p}$) collisions and proton-proton (pp) collisions are shown in Figure 2.5.

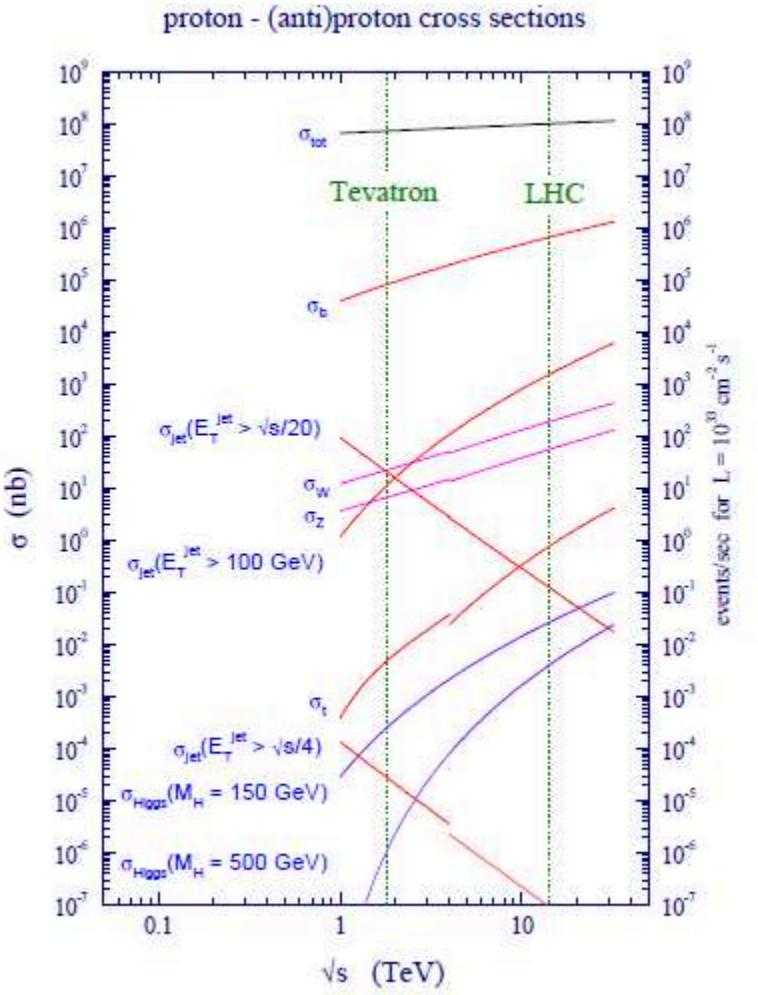


Figure 2.5: Expected cross section and event rates for various physics processes in $p\overline{p}$ collisions and pp collisions [24].

2.2.5 Underlying Events

The underlying event (UE) is everything seen in a collision which is not coming from the primary hard scattering process. It consists of particles that come from the breakup of the initial beam hardrons when a parton is knocked out during the collision ("beam remnants"), and particles arising from soft multiple particle interactions (MPI). The correct modelling of the UE is essential for precise physics measurements at high transverse momentum and large jet multiplicity, but it cannot be described by perturbation theory. Therefore, non-perturbative phenomenological models [43, 44] are needed to describe the UE, whose parameters can be modified, or "tuned", to match measurements from data from multiple experiments. MC generators can be interfaced with different UE and MPI models, where differences in the implementation of each model can have an affect on the final state kinematics. Multijet events at the LHC have been used to study the characteristics of the UE [45]. These types of events are ideal since the region in the direction perpendicular to the leading jet is very sensitive to the UE. The "Perugia Tunes" [46] are the most widely used models in the ATLAS experiment for simulating underlying event and multiple particle interactions.

2.2.6 QCD Radation

Particles that are electrically and/or colour charged can radiate photons or quarks/gluons before they enter the hard interaction, called initial state radiation (ISR), or after they exit the hard scattering process, called final state radiation (FSR). The emission of gluons dominates for hadronic interactions. The radiated quarks and gluons are again able to radiate, which leads to the formation of a partonic cascade where the quarks and gluons are emitted collinear and/or soft to the hard interacting partons.

The correct modelling of ISR and FSR is essential for a complete description of the physics process. Current perturbative calculations in QCD are only performed up to NLO, with a few cases at higher orders. Higher order calculations in α_S become increasingly difficult, thus LO and NLO calculations exist only for a limited number of multiple-parton final states, such

that collinear and/or soft parton emissions are not included. For some phase space regions, these perturbative corrections cannot be neglected. The parton shower (PS) approach [13] can be used to approximate all orders of perturbation theory not included in the calculation of the large-angle hard scattering process.

The PS method is based on a probabilistic approach to describe partonic cascades. Each radiation process is simulated by the branching of a parton into a parton of lower energy plus an emitted quark or gluon. The two new daughter partons might further undergo a branching process. The probability of each additional branching is governed by the DGLAP evolution equations [32, 33, 47] at the energy scale Q^2. The shower evolution is stopped when reaching a cut-off value Q_0 of approximately 1 GeV, well above the confinement regime. ISR is calculated using the backward-shower algorithm [48, 49], whereas FSR is calculated using the forward-shower algorithm [50, 51].

Since hard large-angle emissions are better covered by the full matrix element calculation from pQCD, and soft collinear emissions are better described by the parton shower approach, both methods need to be combined to provide an accurate model of physics processes with multiple parton final states at high energy. To avoid double counting partons in the overlapping regions of phase space described by the ME calculation and the parton shower, merging algorithms are employed. The primary merging algorithms used by ME+PS generators are the MLM [52, 53] and the CKKW [54] merging schemes.

2.2.7 Hadronization

The hard scattering process leaves coloured partons in the final state. The process by which colourless hadrons are formed from coloured objects is called fragmentation or hadronization. Due to the fact that α_S grows to values that are greater than one for energies on the order of $\Lambda_{QCD} \approx 200$ MeV, hadronization cannot be described by perturbation theory. Two different

phenomenological approaches are used to describe the hadronization, and thus the confinement of coloured objects: the string fragmentation model and the cluster hadronization model. The different fragmentation/hadronization models are tuned to match data.

The string fragmentation model [55, 56] is based on the string model [57] of the strong interaction, which predicts a linear rising potential with increasing distance between a quark $\boldsymbol{q}$ and antiquark $\overline{\boldsymbol{q}}$ caused by a coloured flux string due to the self interaction of gluons. If the energy stored in the string is large enough, a second pair of quarks $\boldsymbol{q'}$ and $\overline{\boldsymbol{q'}}$ is produced, which form two colour singlets $\overline{\boldsymbol{q}}\boldsymbol{q'}$ and $\boldsymbol{q}\overline{\boldsymbol{q'}}$. The breaking of strings stops when only on-mass-shell hadrons remain. Diquark-antidiquark pair production allows the production of baryons.

The cluster hadronization [58–61] model is based on the preconfinement property of QCD [62]. This states that at scales much less than that of the hard process, the partons in a shower are clustered into colourless groups with an invariant mass distribution that is independent of the nature of the hard process. These cluster are identified at the hadronization scale as proto-hadrons that decay into the observed final state hadrons. In the first step of this model, all gluons of the parton shower are split into light quark-antiquark or diquark-antidiquark pairs. Next, all quarks are combined with their nearest neighbours to form colour singlet clusters. The mass of these clusters is used to determine which hadrons they will decay into.

2.3 Z/γ^*+jets production

The measurement of the production of jets in association with a Z boson is the main focus of this analysis. Thus, it is important to know about the basic properties of the Z boson, as well as have a theoretical description for the production and decay of a Z boson with associated jets at hadron colliders.

In electroweak theory [14], the charged $W^{\pm}$ bosons and neutral Z boson mediate the weak interaction. Discovered in 1983 in $p\overline{p}$ collisions at the Super Proton Synchrotron (SPS) at

CERN [63], the Z boson is the heaviest of the gauge bosons and mediates the neutral current part of the weak interaction.

The properties of the Z boson are precisely known from e^+e^- collision from the Large Electron Positron Collider (LEP) at CERN and the Stanford Linear Collider (SLC) at SLAC. Due to its decay into two oppositely charged leptons, which leaves a very clean signature in the detector, and its rather large production cross section, on the order of 1 nb at the LHC (see Figure 2.5), measurements of the production of Z bosons in hadronic collisions serve as an important precision benchmark measurement of the SM. Furthermore, the selection of jets in association with a Z boson provides a valuable test of perturbative QCD predictions.

This section provides a summary of the properties of the Z boson, as well as a description of its production and decay with associated jets at the LHC. An overview of previous measurements of $Z/\gamma^* + \text{jets}$ production at hadron colliders is also given.

2.3.1 Z Boson Production and Decay

In hadronic collisions, the dominant production mechanism for Z bosons is the Drell-Yan (DY) process, depicted in Figure 2.6. A quark and an antiquark from the interacting hadrons annihilate to create a pair of oppositely charged leptons, via the exchange of a virtual photon γ^* or a Z boson.

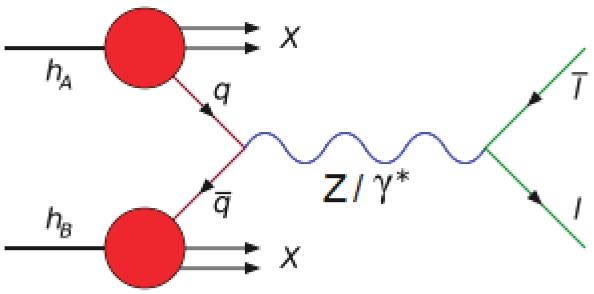


Figure 2.6: Z boson production via the Drell-Yan process.

Theoretically, the inclusive DY cross section is known at next-to-next-to-leading order (NNLO) in pQCD as determined by the FEWZ program [64–66] using the MSTW2008 NNLO PDFs [67]. For the calculation, the QCD factorization and renormalization scales are set to the mass of the Z boson. The NNLO pQCD prediction of the inclusive DY cross section reaches a precision of approximately 5% [68] when taking into account the uncertainties from the NNLO PDFs, the dependence on the factorization and renormalization scale, and the choice of α_S.

The decay width of the Z boson is small $\Gamma_Z = 2.4952 \pm 0.0023$ GeV [14] compared to its mass $m_Z = 91.1876 \pm 0.0021$ GeV [14]. Z bosons decay leptonically or hadronically into a fermion-antifermion pair, except for the higher mass top and antitop quark pair. The branching ratios for the different decay modes are summarized in Table 2.4.

Z Decay Mode	Branching Ratio
e^+e^-	3.363 ± 0.004
$\mu^+\mu^-$	3.366 ± 0.007
$\tau^+\tau^-$	3.370 ± 0.008
$\nu\ \overline{\nu}$	20.00 ± 0.06
hadrons	69.91 ± 0.06

Table 2.4: Branching ratios of the different Z boson decay modes [14].

2.3.2 Jet Production in Association with a Z Boson

The production of a single jet in association with a Z boson can be produced via the Compton process and through $q\overline{q}$ annihilation. The LO Feynman diagrams for both production mechanisms are shown in Figure 2.7. At pp colliders such as the LHC, the dominant production mechanism is the Compton process, whereas at $p\overline{p}$ colliders such as the Tevatron, $q\overline{q}$ annihilation is the dominant process.

Some LO diagrams that contribute to $Z/\gamma^* + 2$ jet final states are shown in Figure 2.8. These diagrams involve gluon radiation from the basic LO topology of a Z boson recoiling against a single jet. Using this approach, MC generators, such as Alpgen and Sherpa, calculate the

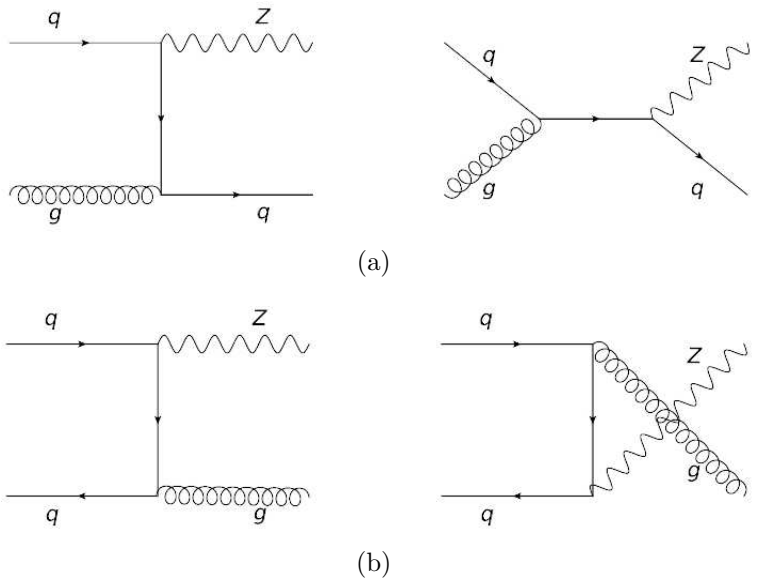


Figure 2.7: Feynman diagrams for the production of a single jet in association with a Z boson via (a) the Compton process and (b) $q\bar{q}$ annihilation.

explicit tree level matrix element at LO for exclusive multijet final states by successively adding real gluon emissions to the LO diagrams from Figure 2.7.

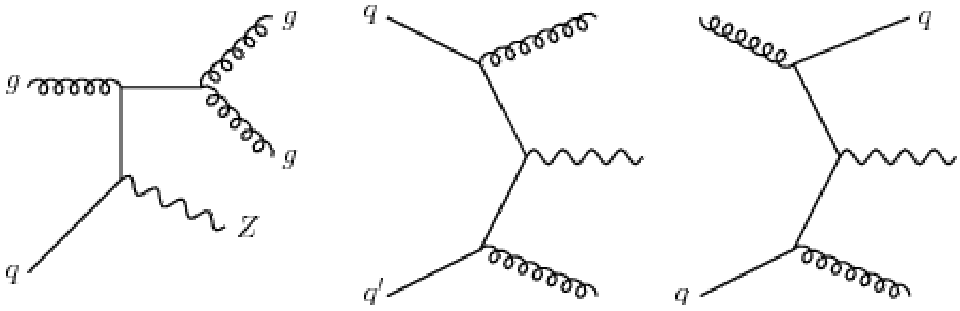


Figure 2.8: LO Feynman diagrams for the production of two jets in association with a Z boson.

However, there is a class of NLO diagrams, shown in Figure 2.9, that contribute to 2-jet final states whose topology is that of a dijet in which a soft or collinear Z boson is radiated from an incoming or outgoing quark line. These diagrams are considered to be higher-order QCD corrections to Z/γ^*+jet final states. In high energy regimes, where the transverse momentum of the jet exceeds the scale given by the Z boson mass, these types of topologies challenge pQCD calculations. At these scales, contributions from final states with a Z boson radiated

from a quark line result in large QCD corrections at the order of $\alpha_{EW}\alpha_s^2 \ln^2(p_T^{j1}/m_Z)$ [6], where the double logarithm comes from the integration over soft and collinear divergences for Z emission. The ratio of the NLO to the LO result (K-factor) is therefore $\mathcal{O}(\alpha_s \ln^2(p_T^{j1}/m_Z))$, which grows with p_T of the leading jet rather than just $\mathcal{O}(\alpha_S)$.

(a) (b)

Figure 2.9: Feynman diagrams for NLO corrections to Z/γ^*+jet observables whose topology is that of a dijet event with additional radiation of a soft or collinear Z boson either from a initial state quark (a) or a final state quark (b).

In addition, pQCD predictions are strongly affected by higher-order EW corrections. For p_T^Z, higher-order EW corrections are expected to reduce the differential cross section by 10%-20% in the range 100 GeV $< p_T^Z <$ 500 GeV [7], as shown in Figure 2.10.

2.3.3 Jet Scaling Patterns

Multijet final states are very common at hadron colliders. Thus, understanding the jet rate from QCD radiation is an important aspect of many physics analyses at the LHC.

Scaling patterns of the exclusive jet multiplicity can be displayed as the ratio of successive jet cross sections

$$R_{(n+1)/n} = \frac{\sigma_{n+1}}{\sigma_n} \tag{2.16}$$

where n describes the number of jets in addition to the hard process.

Two patterns provide limiting cases for the LHC, and are referred to as staircase scaling and Poisson scaling [69]. Staircase scaling is defined as a constant ratio between successive jet cross sections

$$R^{stair}_{(n+1)/n} = R \sim R_0 + n\frac{\mathrm{d}R}{\mathrm{d}n}. \tag{2.17}$$

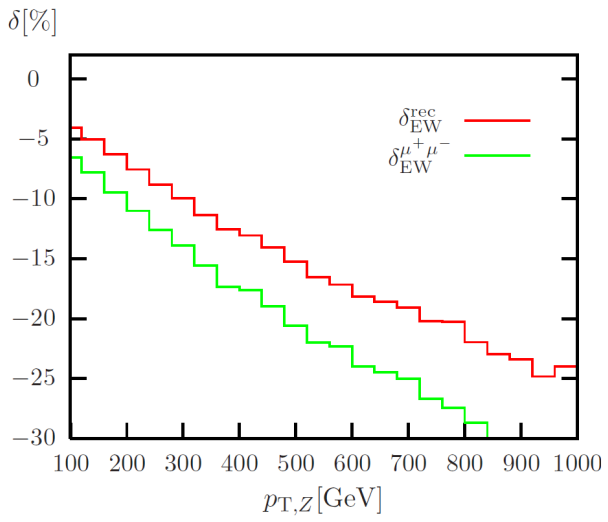


Figure 2.10: Relative electroweak corrections for p_T^Z at the LHC [7]. The EW correction $\delta_{EW}^{\mu^+\mu^-}$ for bare muons (green) and δ_{EW}^{rec} including lepton-photon recombination (red) are shown.

For Poisson scaling, the cross section ratio is given by

$$R_{(n+1)/n}^{Poisson} = \frac{\overline{n}}{n+1} \tag{2.18}$$

where $\overline{n}$ describes the expectation value of the Poisson distribution.

Scaling patterns derived from first principles in QCD for e^+e^- collisions indicate that cross section ratios at high jet multiplicities are described by staircase scaling. Conversely, ratios for low jet multiplicities are described by Poisson scaling. Poisson scaling is attributed to large scale differences between the energy of the first radiated jet and the subsequent soft parton emissions from FSR. For jet production in hadron collisions where the scale of the hard process is not much larger than the scale of additional jet emissions, such as in $Z/\gamma^* + \text{jets}$ production, current theoretical studies [70–72] predict the emission of the first parton to be more suppressed than the emission of additional partons. This restores the staircase scaling pattern down to the lowest jet multiplicities in Z/γ^*+jets events. For large scale differences between the leading jet and the additional radiated jets, the scaling behaviour is expected to transition to a Poisson scaling.

2.3.4 Background Processes to Dilepton Final States

Jet production in association with a Z boson for dielectron final states are characterized by two isolated oppositely charged electrons with additional well separated jets. Background contributions are expected from the production of $t\bar{t}$ pairs, single top quarks, W and Z boson pairs (WW, ZZ, WZ), $Z/\gamma^*(\rightarrow \tau\tau)$, $W(\rightarrow e\nu)$, and multijet events. Top quark pairs are mainly produced via gluon fusion. Single top quarks are produced in electroweak interactions via the s-channel, t-channel, and the associate production with a W boson (Wt-channel). Dibosons (WW, ZZ, WZ) are mainly produced in quark-antiquark interactions.

The background processes can be split into two groups, the irreducible backgrounds and the reducible backgrounds. Irreducible backgrounds are characterized by an identical final state with two isolated electrons. This is the case for $t\bar{t}$, diboson, and $Z/\gamma^*(\rightarrow \tau\tau)$ production. The reducible backgrounds can be split again into two groups, the backgrounds with one isolated electron and the backgrounds without isolated electrons. For both reducible backgrounds the additional electrons originate from misidentification of jets or additional activity in the detector. Events with only one isolated electron in the final state come from jet production in association with a W boson, where the W boson decays leptonically, or from single top-quark production. Events without isolated electrons originate from multijet processes. Even though the misidentification rate for jets faking an isolated electron is rather low, the background from multijet processes has a non-negligible contribution due to the large multijet production cross section.

The cross sections for the individual processes had been shown in Figure 2.5. At the energies of the LHC, the dominant background contributions are multijet processes, $t\bar{t}$, and diboson production. Single top production is a minor contribution, but is taken into consideration in this analysis. The remaining EW backgrounds, $Z/\gamma^*(\rightarrow\ \tau\tau)$ and $W(\rightarrow\ e\nu)$, are negligible when compared to those already listed. The analysis of the 7 TeV data set by the ATLAS experiment [10] found $Z/\gamma^*(\rightarrow \tau\tau)$ and $W(\rightarrow e\nu)$ processes contribute less than one tenth of

1% over the entire jet multiplicity range considered. Due to this fact, and in order to reduce the set of background MC samples to a manageable number, the $Z/\gamma^*(\rightarrow \tau\tau)$ and $W(\rightarrow e\nu)$ background contributions are not considered in this analysis.

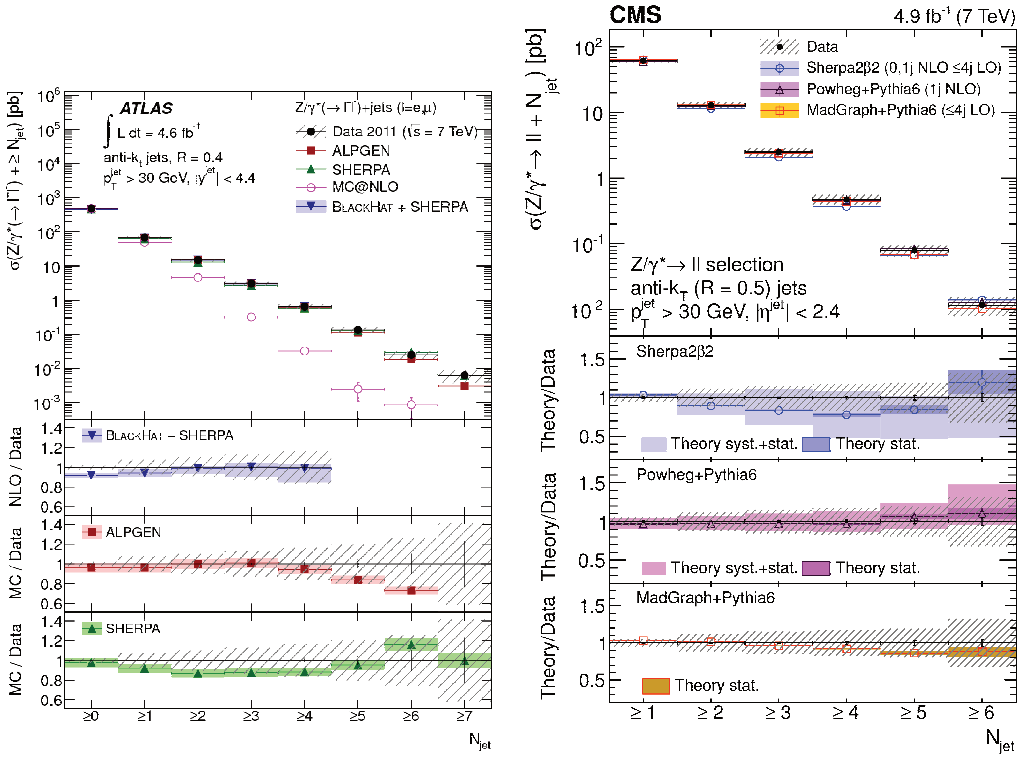
2.3.5 Experimental Measurements of Z/γ^*+jets Production

Measurements of the production cross section of jets in association with a Z boson have been performed by the CDF [8] and DØ [9] collaborations at the Tevatron for $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV. More recent results have been obtained in pp collisions at the LHC from the ATLAS [10] and CMS [11] collaborations using the $\sqrt{s} = 7$ TeV data set. CMS has also looked at event shapes and azimuthal correlations in Z/γ^*+jets events in 7 TeV data [12]. At both colliders, the analysis strategies are quite similar. The measured cross sections, determined by a cut and count method, are corrected for detector effects back to particle level and compared to pQCD predictions and predictions from various generators.

From ATLAS, comparisons are made with predictions from MC generators Alpgen [3] interfaced to Herwig [73] for PS, which calculate the explicit matrix element at LO in QCD with up to five additional partons in the final state, and Sherpa [4], which calculates the matrix element for the inclusive cross section at NLO in QCD while events with 1 – 5 jets in the final state are calculated at LO. Additional comparisons are made with fixed-order pQCD calculations for the production of $Z/\gamma^* + \geq 1$ jet up to $Z/\gamma^* + \geq 4$ jets using the BlackHat+Sherpa [1, 2] program.

Results from CMS are compared to LO predictions from MadGraph [74] with up to four additional partons in the final state interfaced to Pythia for PS simulation. Additional comparisons are made with NLO predictions from BlackHat for $Z/\gamma^* + 0$ jet and $Z/\gamma^* + 1$ jet production which are merged with LO matrix elements from Sherpa with up to four partons in the final state, and from an NLO calculation for the $Z/\gamma^* + 1$ jet matrix element provided by the PowHeg-Box package [75–77] interfaced to Pythia for the PS evolution.

The measured inclusive jet multiplicity from both ATLAS and CMS, shown in Figure 2.11, is consistent with theoretical predictions within uncertainties, although a trend is seen with Alpgen+Herwig to underestimate high multiplicity final states. The measurement has been performed for jet multiplicities up to seven jets in the final state by ATLAS, and up to six jets by CMS.

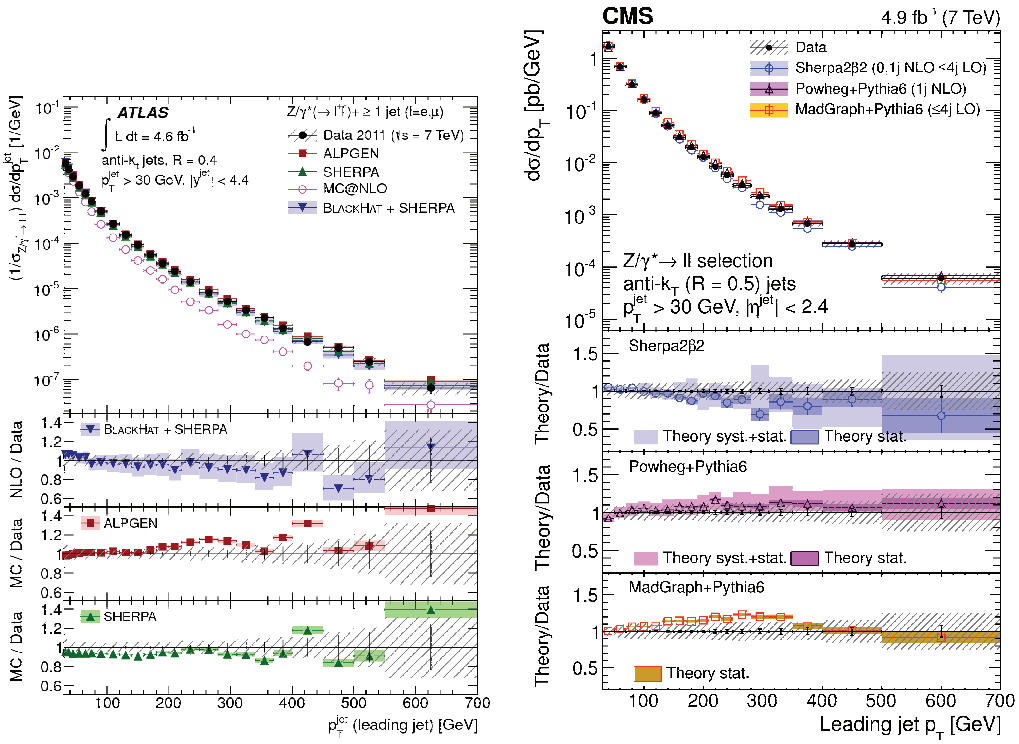


(a) (b)

Figure 2.11: Inclusive jet multiplicity as measured in pp collisions at $\sqrt{s} = 7$ TeV and compared to various theoretical predictions from the ATLAS collaboration (a) and the CMS collaboration (b) [10, 11].

Both ATLAS and CMS have measured the transverse momentum and rapidity of the 1$^{\text{st}}$ – 4$^{\text{th}}$ leading jet in Z/γ^*+jets events. Figure 2.12 shows the transverse momentum distribution of the leading jet as measured by the ATLAS and CMS experiments. In ATLAS, the MC generators Alpgen+Herwig and Sherpa are used heavily to model the kinematics of vector boson plus jets events, whereas CMS relies primarily on MadGraph. Alpgen+Herwig predicts a too hard spectrum of the transverse momentum of the leading jet in a regime where large corrections from higher-order electroweak and QCD processes are expected. The pQCD

predictions from MadGraph provide a satisfactory description of data, but shows an excess in the p_T spectra of the leading jet at $p_T > 100$ GeV.



(a) (b)

Figure 2.12: Tranverse momentum of the leading jet as measured in pp collisions at $\sqrt{s} =$ 7 TeV and compared to various theoretical predictions from the ATLAS collaboration (a) and the CMS collaboration (b) [10, 11].

Additional measurements from the ATLAS experiment, which are relevant to this analysis, are shown in Figures 2.13(a) – (d).

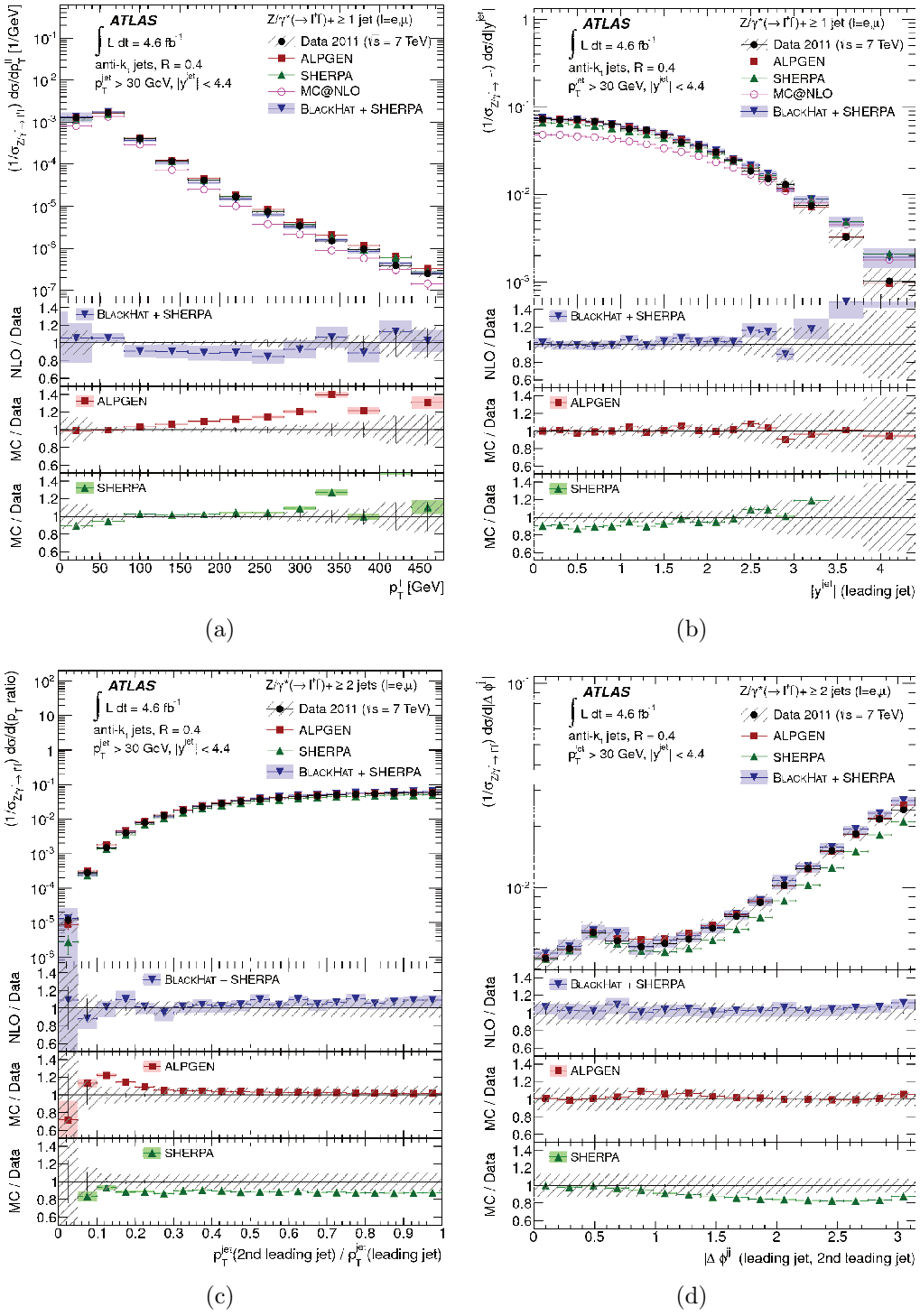
Figure 2.13(a) shows the p_T measurement of the recoiling Z boson with at least one jet in the event. Alpgen+Herwig predict a too hard p_T spectrum. The discrepancy with data is comparable to the expected higher-order electroweak corrections, although higher-order QCD corrections could equally account for this.

Figure 2.13(b) shows the rapidity spectrum of the leading jet. Both BlackHat+Sherpa and Sherpa predict a rapidity spectrum for the leading jet that is wider than observed in data. Alpgen+Herwig predictions are compatible with the measurements.

The measurement of the p_T ratio of the 2$^{\text{nd}}$ leading jet to the 1$^{\text{st}}$ leading jet, which is shown in Figure 2.13(c), provides a more precise test of pQCD since part of the systematic uncertainties cancel. Alpgen+Herwig overestimates the cross section in the range 0.1–0.2, which corresponds to a topology with a much harder leading jet compared to the 2$^{\text{nd}}$ leading jet.

The final measurement of interest is the azimuthal angular separation ($\Delta\phi^{jj}$) between the two leading jets in the event, shown in Figure 2.13(d), which explores the topology of dijet final states. The azimuthal separation is well modelled by Alpgen+Herwig and Black+Sherpa, while Sherpa models a too flat spectrum.

Since the MC generators, Alpgen and Sherpa, are so heavily relied upon by the ATLAS experiment for the modelling of vector boson plus jet final states, it is of vital interest to determine if the trends and discrepancies seen in the previous measurement using the 7 TeV data set persist in the 8 TeV data set. In addition, the significantly larger data set will allow the exploration of extreme regions of phase space important to many analyses that are part of the ATLAS research program where Z/γ^*+jets events form significant backgrounds.



(a) (b)

(c) (d)

Figure 2.13: (a) tranverse momentum of the Z boson in events with at least one jet in the final state (p_T^Z), (b) rapidity of the leading jet (y^{jet}), (c) transverse momentum ratio of the 2^{nd} leading jet to the 1^{st} leading jet (Rp_T^{jj}), and (d) the absolute azimuthal separation between the two leading jets in the event ($|\Delta\phi^{jj}|$) as measured in pp collisions at $\sqrt{s} =$ 7 TeV and compared to various theoretical predictions from the ATLAS collaboration [10].

CHAPTER 3

Experimental Setup

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [78] is currently the world's largest particle accelerator, located at CERN (European Organization for Nuclear Research) on the Swiss-French border near the city of Geneva in Switzerland. It is a circular machine, 27 km in circumference, built in the former tunnels of the Large Electron-Positron collider [79]. It is designed to collide counter-rotating beams of protons up to a center-of-mass energy of 14 TeV with an instantaneous luminosity of 10^{34} $s^{-1}cm^{-2}$, and counter-rotating beams of Pb^{+} ions at a center-of-mass energy of 5.5 TeV with an instantaneous luminosity of 10^{27} $s^{-1}cm^{-2}$.

Protons are produced by the ionization of hydrogen atoms. Before being injected into the main LHC ring, the protons are brought up to an energy of 450 GeV by a series of smaller linear and circular particle accelerators. The LINAC 2 accelerates protons up to an energy of 50 MeV before injecting them into the Booster (PSB) which brings the energy up to 1.4 GeV. The Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) accelerate the protons up to energies of 25 GeV and 450 GeV respectively. Finally, the protons are injected into the main LHC ring in opposite directions where they are accelerated up to the nominal beam energy before being brought into collision. By design, each proton beam contains 2808 proton bunches, with approximately 10^{11} protons per bunch. These conditions result in an average of roughly 20 interactions per bunch crossing, referred to as pile-up (μ). Super conducting magnets, which are able to create magnetic fields up to a maximum strength of 8.33 T, are used to bend the the trajectory of the protons around the LHC ring.

The complete LHC accelerator complex is shown in Figure 3.1. The LHC has four main collider experiments: ATLAS, CMS, ALICE, and LHCb. ATLAS (described in Section 3.2) and CMS [80] are two complementary particle detectors whose main purposes are the discovery of the Higgs Boson, measurement of its properties, and searches for new physics. The ALICE [81] detector was designed to study heavy ion collisions and the properties of matter at high energy densities. The LHCb [82] detector was built to investigate proton-proton collisions with a focus on physics with b-quarks.

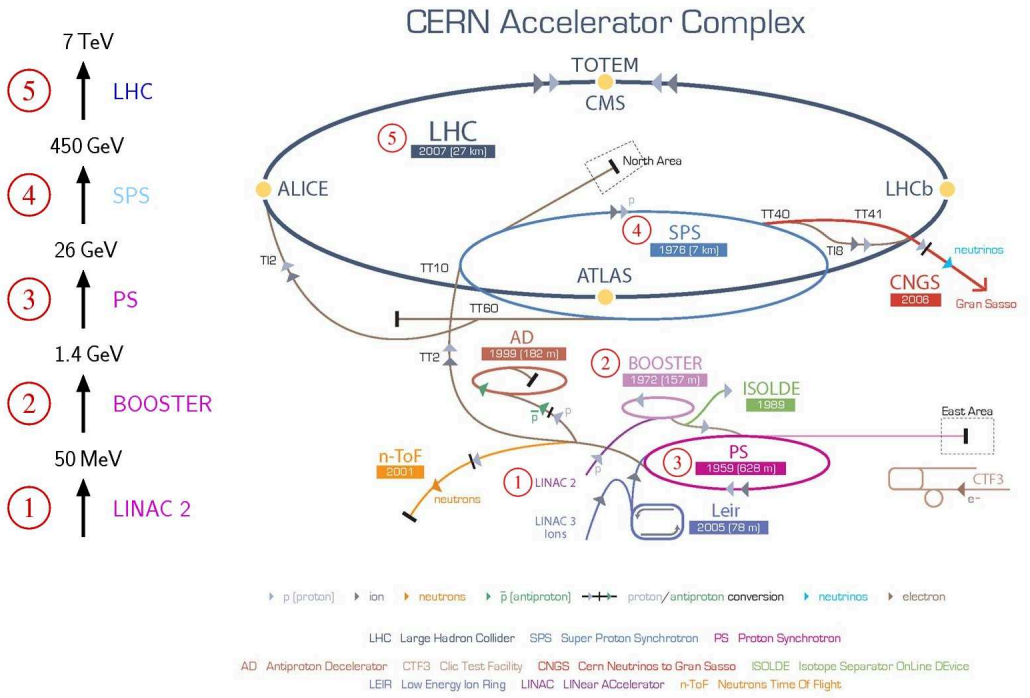


Figure 3.1: The Large Hadron Collider accelerator complex showing the main LHC ring, preaccelerators, and the 4 main experiments.

The LHC began operation in 2009 after an unsuccessful first attempt in 2008 was stopped by a catastrophic magnet quench caused by an electrical fault. First proton-proton collisions were recorded at a center-of-mass energy of 7 TeV, which was a new world energy record at the time. The operational conditions of the LHC for the data taking periods of 2010, 2011, and 2012 are summarized in Table 3.1.

Parameter	2010	2011	2012
Center-of-mass energy (TeV)	7	7	8
Maximum number colliding bunch pairs	348	1331	1380
Minimum bunch spacing (ns)	150	50	50
Peak instantaneous luminosity (10^{33} $s^{-1}cm^{-2}$)	0.2	3.6	7.73
Maximum interactions per bunch crossing	∼5	∼20	∼70
Average interactions per bunch crossing	∼ 2	∼9	∼20
Total integrated luminosity delivered	48.1 pb^{-1}	5.61 fb^{-1}	22.8 fb^{-1}

Table 3.1: Operational conditions of the LHC for the 2010, 2011 [83], and 2012 [84] data taking periods of proton-proton collisions.

The instantaneous luminosity during the first year of running was relatively small. This allowed measurements to be made in low pile-up conditions, however only 48.1 pb^{-1} of integrated luminosity was delivered. In 2011, the center-of-mass energy remained at 7 TeV. However, the bunch spacing was decreased, and the number of protons per bunch and the number of colliding bunch pairs were increased. Consequently, the total delivered integrated luminosity was increased to 5.61 fb^{-1}, but with a corresponding increase in the amount of pile-up. In 2012, the center-of-mass energy was increased to 8 TeV and the maximum number of colliding bunch pairs was increased to 1380. For the 8 TeV run period, 22.8 fb^{-1} of total integrated luminosity was delivered by the LHC, and 20.3 fb^{-1} was recorded by the ATLAS detector resulting in a data taking efficiency of 93.4%. The total integrated luminosity delivered by the LHC over time and recorded by the ATLAS detector for the 2011 and 2012 data taking period is shown in Figure 3.2. The distribution of the mean number of interactions per crossing for the 2011 and 2012 data taking period is shown in Figure 3.3.

In 2013, the LHC shut down for a planned technical stop in order to upgrade the machine for operation closer to its design parameters. In 2015, the LHC began operation again at a center-of-mass energy of 13 TeV.

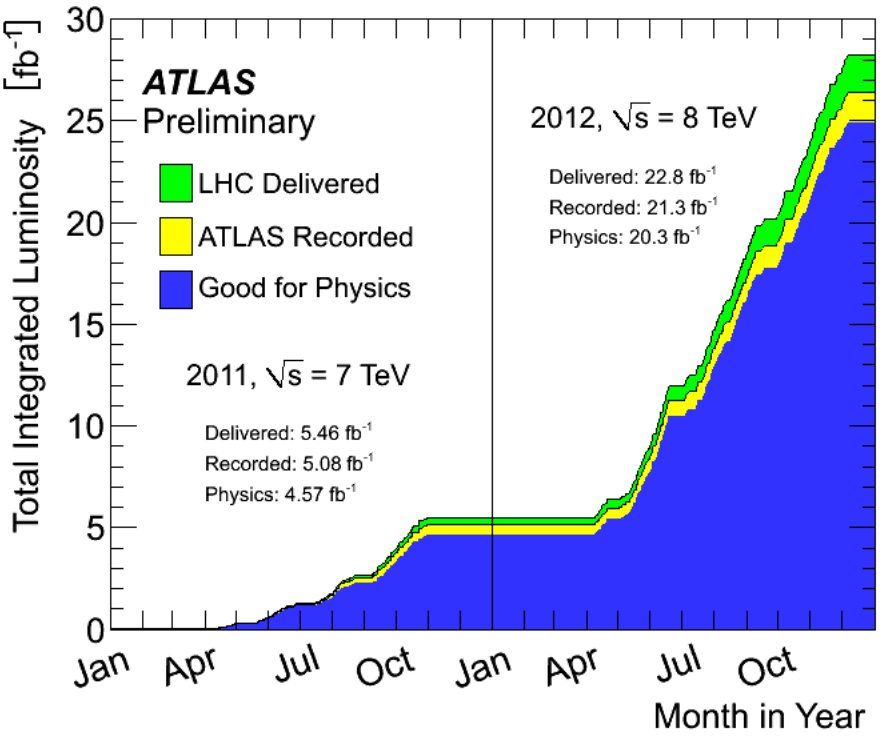


Figure 3.2: Cumulative luminosity versus time delivered by the LHC (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams and for proton-proton collisions at 7 and 8 TeV center-of-mass energy in 2011 and 2012 [85] respectively.

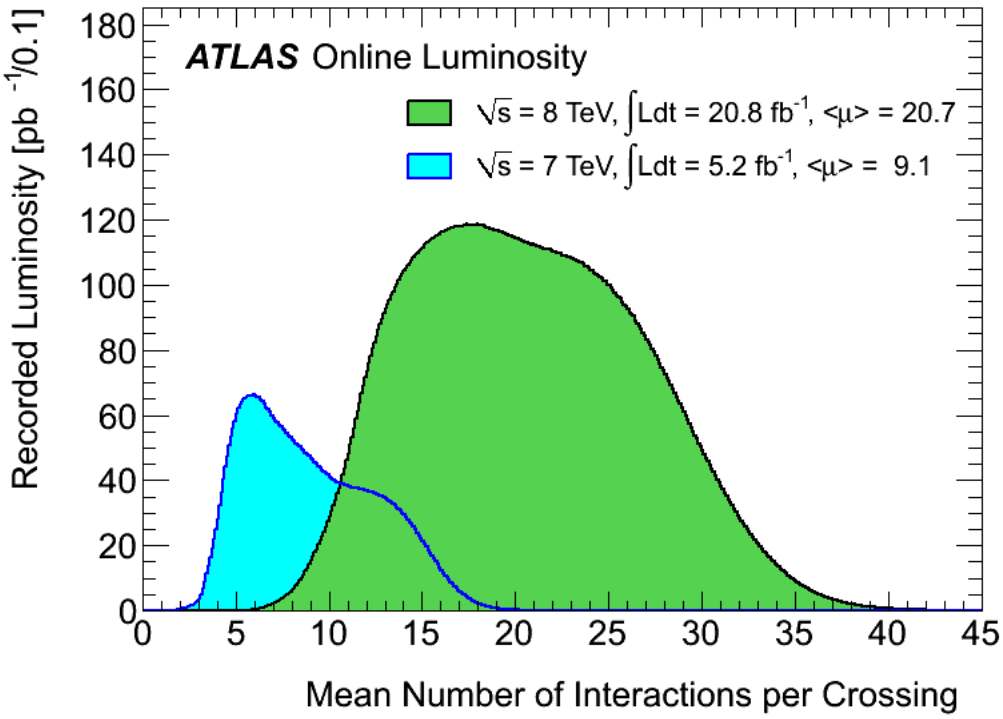


Figure 3.3: Luminosity-weighted distribution of the mean number of interactions per crossing for the 2011 and 2012 proton-proton runs. The integrated luminosity and the mean μ (pile-up) values are given in the figure. The mean number of interactions per crossing corresponds to the mean of the Poisson distribution of the number of interactions per crossing calculated for each bunch [85].

3.2 The ATLAS Detector

The ATLAS detector [86] is the largest of the experiments at the LHC, with a length of 45 m and a height of 22 m. A schematic of the detector is shown in Figure 3.4.

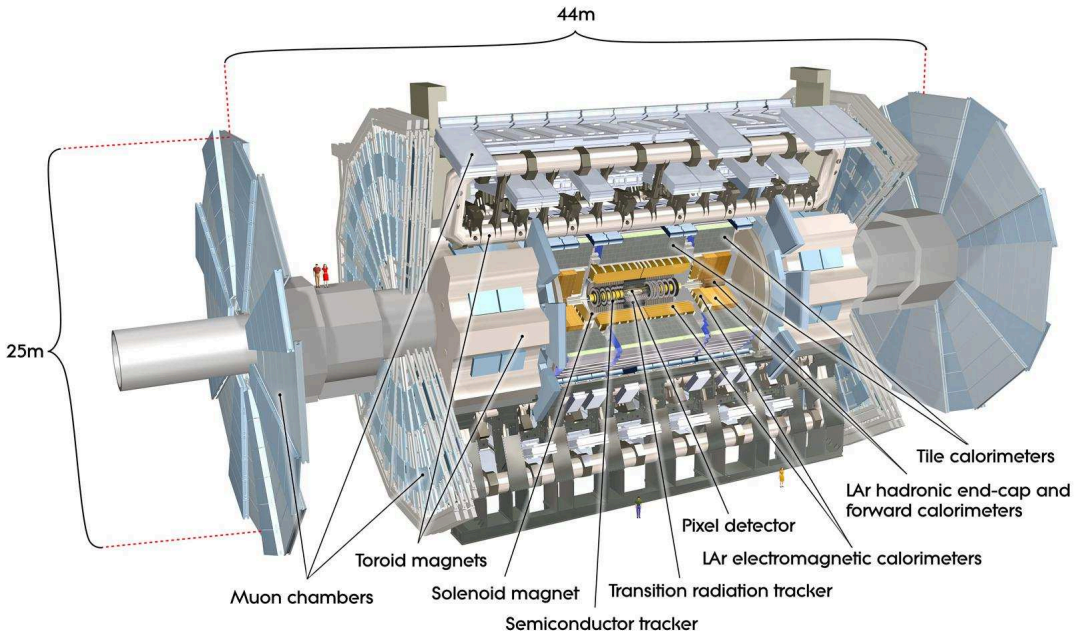


Figure 3.4: The ATLAS detector.

The design requirements of the detector were driven by the physics program of the LHC. The broad physics spectrum covers topics such as precise measurements of both established and rare Standard Model processes, physics beyond the Standard Model, the discovery of the Higgs boson and measurement of its properties. Many of these interesting physics processes have, or are expected to have, very low cross sections. Therefore, a large integrated luminosity is required to reach the significance of a discovery or observation. The ATLAS detector has been optimized for such measurements. In addition, the detector must deal with extremely high collision rates. Roughly 10^9 proton-proton interactions per second are expected during the high luminosity phase of the LHC. This introduces the problematic issues of in-time and out-of-time pile-up. Out-of-time pile-up refers to the effect where there are physical objects, e.g. pions, from previous bunch crossings still in the active part of the detector. In-time

pile-up is related to the number of additional proton-proton interactions occurring in each bunch crossing. During the 2012 running of the LHC, there were on average a further 20 soft hadronic interactions accompanying each interesting physics event.

Most of the interesting physics of the ATLAS scientific programme can be characterized by the presence of high-p_T leptons, highly energetic jets of particles, and a large amount of missing transverse energy (MET) in the final state. This implies basic design requirements which satisfy a broad field of specifications. The detector and its electronics must work reliably in a high radiation environment and provide precise measurements of various physical quantities, such as charge, transverse momentum and energy of leptons, photons, and jets, as well as MET. It must cover as much as possible a hermetically closed volume around the interaction point to ensure that no stable particles can escape without being detected. A very good electromagnetic calorimeter is needed for electron and photon identification and energy measurements, together with a hadronic calorimeter for jet and MET measurements. In addition, high performance muon detectors are needed to provide good muon identification and energy resolution.

As with most colliding beam experiments, the ATLAS detector has a cylindrical shape surrounding the beam pipe and is centered around the interaction point. It consists of five main components: the magnet system, the inner detector system, the calorimeter system, the muon system, and the data acquisition system. The innermost part of the detector contains a silicon-based tracker inside a transition radiation tracker, which is surrounded by the electromagnetic calorimeter and the hadronic calorimeter. The outermost part of the ATLAS detector is the muon spectrometer. The different detector components are described in more detail in the following subsections.

The ATLAS detector uses a right-handed coordinate system. The nominal interaction point is defined as the origin of the coordinate system, while the beam direction defines the z-axis and the xy plane is transverse to the beam direction. The positive x-axis is defined as

pointing from the interaction point to the centre of the LHC ring and the positive y-axis is defined as pointing upwards. The azimuthal angle ϕ is measured around the beam axis and the polar angle θ is the angle from the beam axis. The transverse momentum p_T, the transverse energy E_T, and the missing transverse energy E_T^{miss} are defined in the xy plane. The direction of a particle's track can be described by the pseudorapidity, defined as

$$\eta = -\frac{1}{2} \ln \left(\frac{|\overrightarrow{p}| + p_z}{|\overrightarrow{p}| - p_z} \right) = -\ln \left(\tan \frac{\theta}{2} \right). \tag{3.1}$$

For massive objects like jets, the rapidity is used

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right). \tag{3.2}$$

3.2.1 Magnet System

Strong magnetic fields are used to bend the trajectories of charged particles in order to measure their momentum. The ATLAS detector incorporates two large superconducting magnet systems [87] to produce these magnetic fields: the solenoid system and the toroid system. A schematic view of the magnet systems are shown in Figure 3.5.

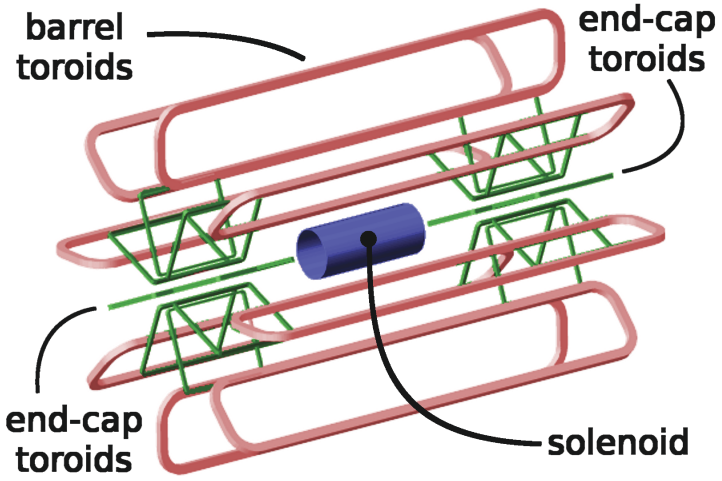


Figure 3.5: Magnet system of the ATLAS detector.

Central Solenoid Magnet System

The central solenoid (CS) magnet system provides a magnetic field for for the momentum

measurement of charged particles by the Inner Detector. It surrounds the tracking system and produces an axial magnetic field of about 2 T at the interaction point.

Air-core Toroid Magnet System

The toroid magnet system provides the magnetic field for the Muon Spectrometer. It consists of one barrel toroid (BT) and two endcap toroids (ECT). The barrel and the endcap toroid magnets together cover a pseudorapidity range up to $|\eta| = 2.7$. The bending power provided by the BT ranges between 2 to 6 Tm, while the ECT contributes 4 to 8 Tm, in the pseudorapidity ranges 0–1.3 and 1.6–2.7 respectively. The bending power is lower in the transition regions where the two magnets overlap, $1.3 < |\eta| < 1.6$.

3.2.2 Inner Detector and Tracking

The Inner Detector (ID) [88] is the subsystem closest to the interaction point. It has full coverage in ϕ and covers a pseudorapidity range up to $|\eta| = 2.5$. Its primary task is the precise reconstruction of the tracks of charged particles. It is immersed in the axial magnetic field of 2 T generated by the central solenoid magnetic system. By reconstructing the trajectory of the charged particles, and knowing the magnetic field inside the ID, the charge, initial momentum, direction of flight, and impact parameter can be calculated. The impact parameter describes the point of closest approach of the trajectory to the beam line. A schematic view of the ID is shown in Figure 3.6.

The ID consists of three subsystems: the silicon pixel detector, the silicon micro-strip detector (SCT), and the transition radiation tracker (TRT). The ID is used for vertex and track reconstruction of charged particles. The transverse momentum resolution is meant to be better than 30% for charged particles with a transverse momentum of 500 GeV. Moreover, the ID must provide precise primary and secondary vertex reconstruction, which, for example, is important for the identification of B-mesons and converted photons.

Silicon Pixel Detector

The pixel detector is the innermost part of the ID. It provides high precision measurements

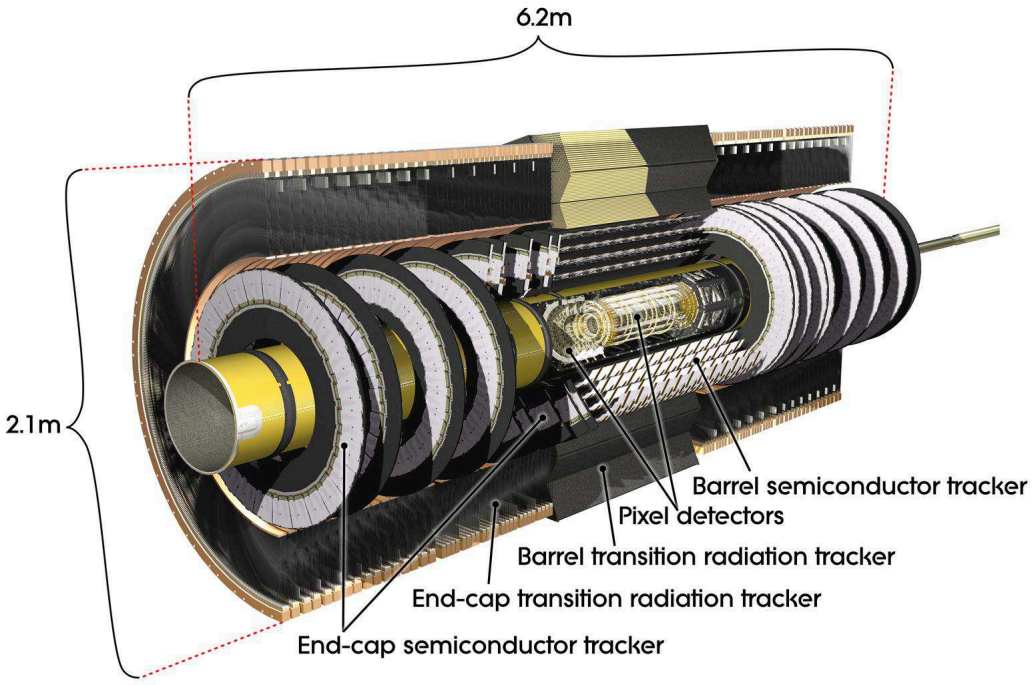


Figure 3.6: ATLAS inner detector.

of position with the best resolution of the impact parameter due to its high granularity and its close proximity to the beam pipe. The active material of the pixel detector is silicon, which is structured in rectangular cells called pixels with a nominal size of 50 μm in the $R\phi$-plane by 400 μm in z-direction. Charged particles that pass through the silicon produce electron-hole pairs. A bias voltage applied to each cell cause the electrons and holes to drift to the readout side of the cell. The cells are placed in three cylindrical layers in the barrel region, with distances to the beam-line of 5.05 cm, 8.85 cm, and 12.25 cm. The endcap regions are covered by four disks of cells on each side. The pixel detector typically provides three space points per track over the full acceptance region of the ID and reaches a resolution of up to 9 μm – 10 μm in the $R\phi$-plane and a resolution of 110 μm in the z-direction.

Silicon Micro-Strip Detector (SCT)

The SCT encloses the pixel detector and is responsible for tracking at radii from 30 cm to 60 cm. As with the pixel detector, silicon is used as the active material. Each silicon detector is 6.36 cm $\times$ 6.40 cm with 768 readout strips, each with 80 μm pitch. The barrel SCT uses

8 layers of silicon micro-strip detectors, which are grouped into pairs and glued back-to-back with a small stereo angle of 40 mrad to provide precision measurements in the z and $R\phi$ coordinates. The barrel region covers a pseudorapidity range up to $|\eta| = 1.4$. The endcaps consist of 9 discs of similar construction and cover a pseudorapidity region $|\eta| = 1.4 - 2.5$. The SCT typically provides four space points per track and reaches a resolution of 16 μm in the $R\phi$-plane and 580 μm in the z-direction.

Transition Radiation Tracker (TRT)

The TRT, which is a gaseous detector, was chosen for radii between 60 cm and 107 cm due to the high cost per unit area of semiconductor layers and their relative high radiation length. It covers a pseudorapidity range up to $|\eta| = 2.0$, and consists of a barrel region and two endcaps. The TRT is composed of 4 mm diameter drift straw tubes, each filled with a 70:27:3 gas mixture of Xe:CO_2:O_2. Each tube has a gold-plated tungsten wire in the middle. Charged particles passing through the straw tubes lead to ionization of the gas mixture. In addition, the walls of the straw tubes contain a polyethylene radiator material which enhances the production of transition radiation photons. These photons can be detected in the Xe gas. The production of transition radiation photons, which depends on the mass, charge, and momentum of the particle can be used to distinguish between electrons and pions. The TRT typically provides 36 space points per track and reaches a resolution of up to 130 μm per straw tube in the $R\phi$-plane.

3.2.3 Calorimetry

The ATLAS calorimeter system [89] encloses the solenoid system and the inner detector as shown in Figure 3.7. The purpose of the calorimeter system is to measure the energy and position of charged and neutral particles. It is composed of an inner electromagnetic calorimeter (EMCAL) for the measurement of electrons and photons, and an outer hadronic calorimeter (HCAL) for the measurement of hadrons. Both calorimeter systems have full coverage in ϕ. Furthermore, the calorimeters are used for the determination of missing

energy. This requires a large coverage in η. The main calorimeter system consists of one barrel and two endcaps which cover a pseudorapidity range up to $|\eta| = 3.2$. A special radiation resistant forward calorimeter is placed at either end of the detector to provide η coverage in the range $|\eta| = 3.1 - 4.9$. Measurement of the position of the incident particle is achieved by segmenting the calorimeter.

The ATLAS calorimeter employs the sampling technique for energy measurement. In general, a calorimeter measures the energy of an incident particle via absorption within the calorimeter material. A sampling calorimeter separates the absorption and energy measurement using alternating layers of passive and active materials. The advantage of this technique is that each material can be chosen to best suit the task. For example, a very dense material can be used to produce a particle shower that evolves quickly in a limited space, even if the material itself is not meant to measure the energy deposited by the shower. Both hadronic and electromagnetic particles produce showers of particles in a cone around the original direction of the incident particle as it interacts with the dense absorber material of the calorimeter. The active layer then measures the energy deposited by this shower. The disadvantage of this method is that some of the particles in the shower are stopped completely within the absorber material and their energy is not measured.

Electromagnetic Calorimeter

The ATLAS Electromagnetic Calorimeter (EMCAL) takes advantage of the interaction of electrons and photons with matter. For electrons at high energies ($E \gg mc^2$), the most important interaction is bremsstrahlung, which leads to a production of additional photons. The probability of this interaction is proportional to the square of the charge, or number of protons, (Z) of the nuclei of the absorber material. The photons themselves can produce further electrons via pair production, which is the dominant process for high-energy photons. The probability of this interaction is also dependent on Z^2. The leads to a cascade of electrons and photons in the general direction of travel of the original particle.

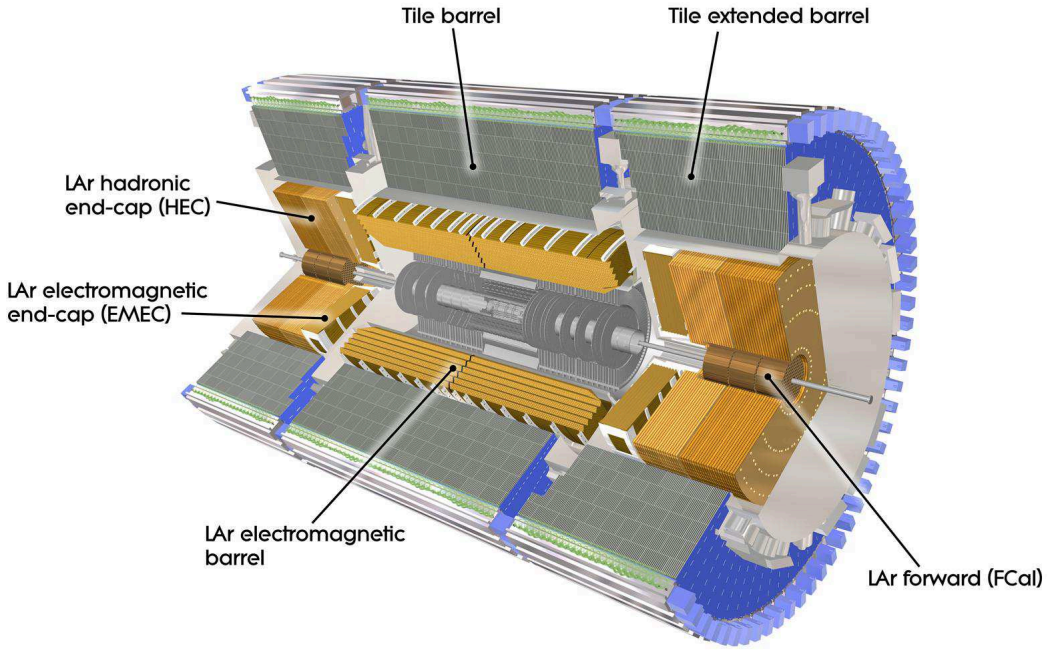


Figure 3.7: ATLAS calorimeter system.

The EMCAL consists of a barrel and two endcap calorimeters, covering a pseudorapidity range up to $|\eta| = 3.2$. It uses lead absorber plates as a passive medium for the shower production of photons and electrons due to the high Z-number of lead. Liquid argon acts as a simple ionization chamber. The lead absorber plate are arranged in an accordion-style structure as shown in Figure 3.8, allowing for full coverage in ϕ without azimuthal cracks. This design ensures that approximately all tracks transverse the same amount of material.

The EMCAL has three longitudinal layers (called strip, middle, and back) with fine segmentation in the lateral direction of the showers in the ID coverage. At high energy, most of the EM shower energy is collected in the middle layer. The energy resolution of the EMCAL has been determined in electron test beam measurements [90] to be:

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus b, \text{ with } a = (10.1 \pm 0.1)\% \text{ and } b = (0.17 \pm 0.04)\% \tag{3.3}$$

An important aspect for the performance is the material budget in front of the EMCAL because a significant fraction of a particles energy is lost in the inactive material in front of the EMCAL. Thus, the EMCAL is preceded by a thin LAr based pre-sampling calorimeter in the barrel covering a pseudorapidity region up to $|\eta| = 1.52$, designed to correct for energy losses prior to the EMCAL.

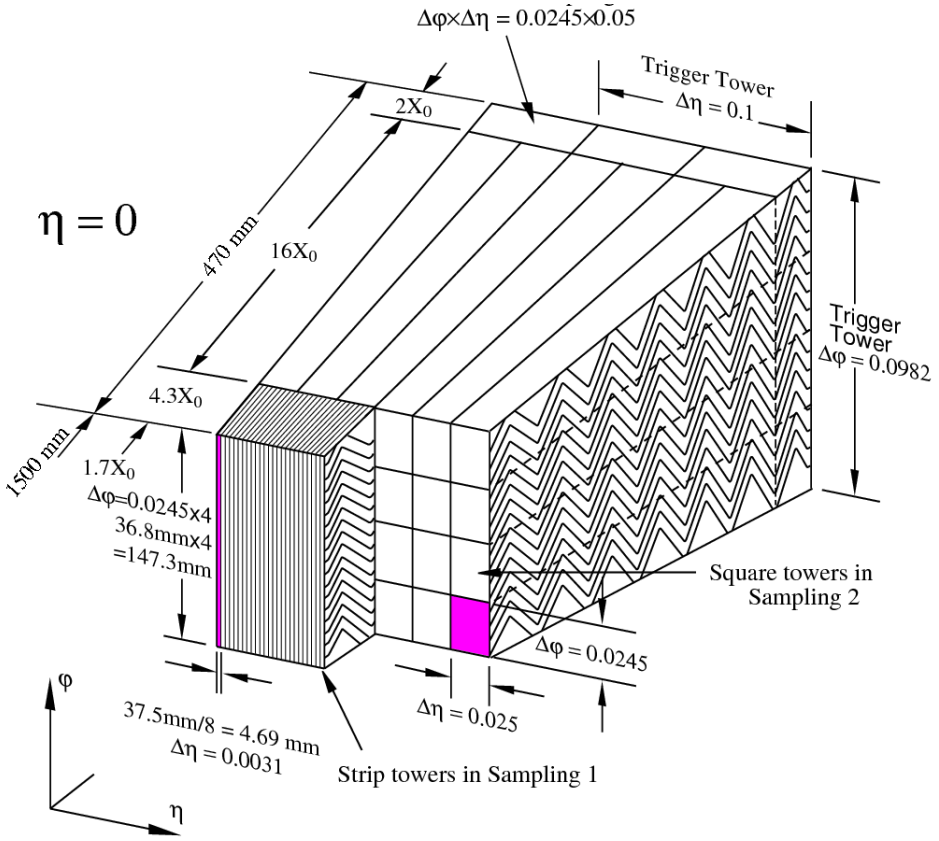


Figure 3.8: The layout of an Electromagnetic Calorimeter module with the accordion shape of the lead plates.

Hadronic Calorimeter

The purpose of the Hadronic Calorimeter (HCAL) is the identification, reconstruction, and energy measurement of jets of particles which result from the hadronization of quarks and gluons, and the measurement of the missing transverse energy in an event. The major design differences between the HCAL and the EMCAL is a result of the fact that hadronic showers are larger since they are produced via strong interactions. A hadronic shower penetrates much

further into the calorimeter than an electromagnetic shower since the interaction length for hadrons is larger than that for electromagnetic particles. The incoming hadrons interact with the atomic nuclei and produce further neutrons, protons, and primarily pions, which then interact with further nuclei. A significant fraction of the incoming energy of the hadrons is used for breaking nuclear binding, which is one of the reasons for the poorer energy resolution of the hadronic calorimeter. The decay of π^0 into two photons also induces electromagnetic showers, which accompanies the hadronic counterpart. The hadronic interactions leave behind highly excited nuclei, which may undergo fission or radiate to lower energy states, that lead to a part of a hadronic shower. The signal response of the calorimeter to hadronic interactions is generally less than the response to electromagnetic interactions, thus the ratio of the calorimeter response of electrons to hadrons becomes greater than 1, i.e. $e/h > 1$. Since the ATLAS HCAL does not correct for this effect, it is called a non-compensating calorimeter. An additional software-based offline calibration algorithm is required in order to compensate for this effect.

The energy of the hadronic shower is measured again by the sampling technique. Due to the larger interaction length of hadrons, resulting in larger shower sizes, more material is needed in the HCAL. The ATLAS HCAL is positioned around the electromagnetic calorimeter from a radius of 2.28 m to 4.23 m. In the central barrel region, the Tile Calorimeter covers a pseudorapidity range up to $|\eta| = 1$. An extended barrel calorimeter is responsible for a pseudorapidity range $|\eta| = 0.8 - 1.7$. Steel plates are used as the absorber material and also as the return yoke for the solenoid magnetic field. Scintillator plastic tiles are used as the active medium. The readout of the tiles is achieved with optical fibers. Readout cells are formed by a cluster of tiles that projects to the interaction point. They provide a granularity of $\delta\eta \times \delta\phi = 0.1 \times 0.1$. A single module of the Tile Calorimeter is shown in Figure 3.9. The region $|\eta| = 1.5 - 3.2$ is covered by the Hadronic Endcap Calorimeter (HEC). The HEC uses copper plates as the absorber material and liquid argon as the active medium. The tile and endcap calorimeters are segmented into three independent layers. The readout cells

provide a three dimensional measurement of the deposited energy, which is needed for the jet reconstruction and the jet triggering. The energy resolution of hadrons in the tile calorimeter has been measured by test beam studies [91] to be

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus b, \text{ with } a = (52.9 \pm 0.9)\% \text{ and } b = (5.7 \pm 0.2)\%. \tag{3.4}$$

A special forward calorimeter is placed at either end of the detector to provide η coverage in the range $|\eta| = 3.1 - 4.9$ in order to improve the measurement of missing transverse energy by providing more complete coverage. They are located 4.7 m along the beam direction from the interaction point, and because of its close proximity to the beam line, it must be resistant against hard radiation coming directly from the proton beam.

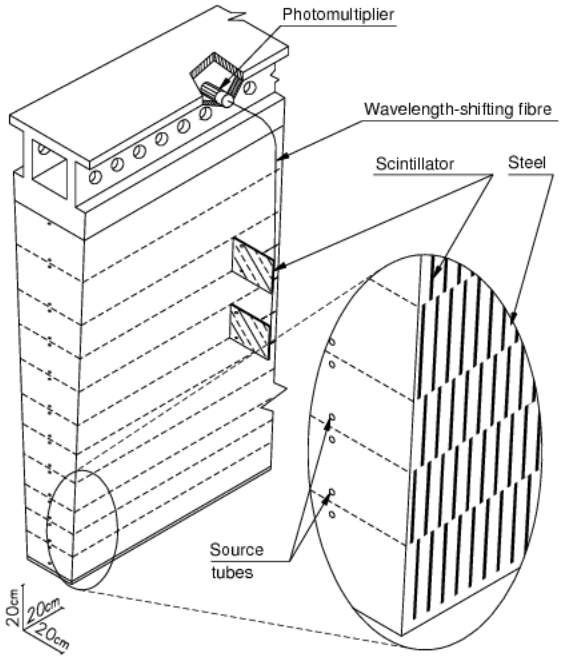


Figure 3.9: A single module of the Tile Calorimeter.

3.2.4 Muon System

The ATLAS muon spectrometer [92], the outermost part of the ATLAS detector, is designed to measure the momentum of muons from 3 GeV to 3 TeV with high resolution covering a pseudorapidity range up to $|\eta| = 2.7$. The accuracy of the momentum measurement is

required to be within 10% for 1 TeV muons. The muon system is composed of four components: the monitored drift tubes (MDT), the cathode strip chambers (CSC), the resistive plate chambers (RPC), and the thin gap chambers (TGC). The MDTs are used as high precision tracking chambers in the full pseudorapidity range, except for the innermost endcap layer in the pseudorapidity region $|\eta| = 2.0 - 2.7$ where the CSCs are used due to their higher granularity. The MDTs and CSCs achieve a resolution of 80 μm and 60 μm, respectively. The RPCs and TGCs are used for triggering with a response time of 15 ns – 25 ns. The RPCs cover the barrel region ($|\eta| < 1.05$) and the TGCs cover the endcap regions ($|\eta| = 1.05 - 2.4$). A schematic view of the the ALTAS muon system is shown in Figure 3.10.

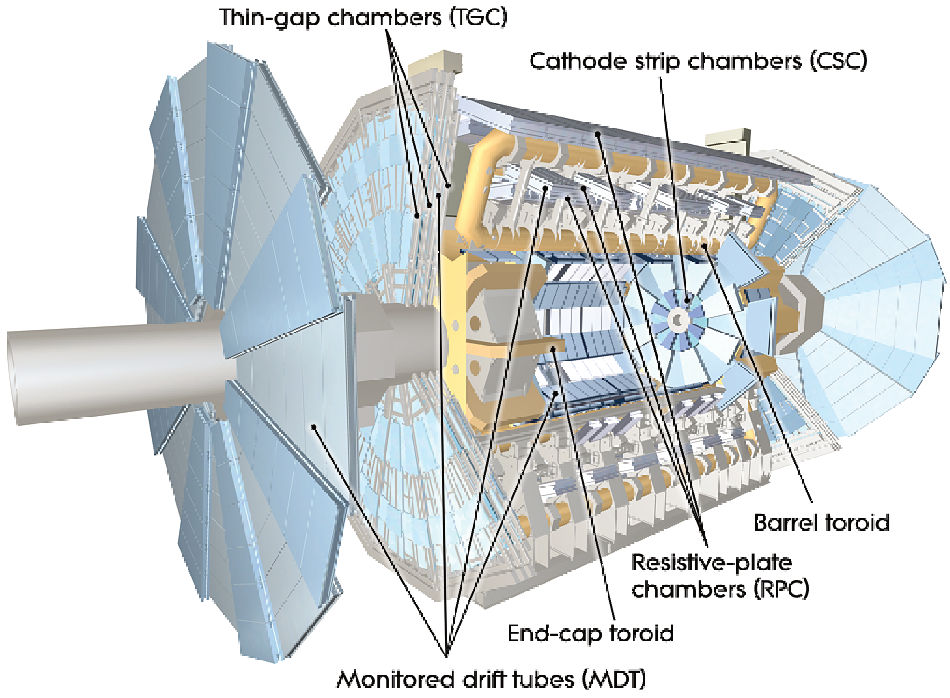


Figure 3.10: ATLAS muon system.

3.2.5 Trigger and Data Acquisition System

The size of one fully reconstructed event from the ATLAS detector is on the order of 1.5 MB. Given a bunch crossing rate of roughly 40 MHz, combined with an average of approximately 20 interactions per bunch crossing, a very high event rate up to an order of 1 GHz is expected. Current technology is not capable of recording this magnitude of data volume. To be handled

by the ATLAS computing system, a reduction to a more manageable rate is required. The goal of the ATLAS trigger and data acquisition system (TDAQ) is to reduce the event rate from 40 MHz to less than 1 kHz without a loss of interesting physics events, e.g. events which contain a Higgs boson or a Z boson.

The ATLAS trigger system [93, 94] consists of three levels, the level-1 trigger (LVL1), the level-2 trigger (LVL2), and the event filter (EF). The combination of the LVL2 trigger and EF is represented by the high-level trigger (HLT). The event rate is reduced consecutively at each level to give sufficient time for data processing. The amount of time needed by the trigger to process an event is given by the latency. A trigger signature is a condition placed on any combination of reconstructed objects, e.g. at least one trigger-level reconstructed electron with $p_T > 12$ GeV. The content of each active trigger menu is strongly dependent on the run and data taking conditions. Certain high-rate triggers need an additional pre-scale. A pre-scale randomly selects a fraction of events which pass the trigger decision in order to further reduce the trigger rate. A functional overview of the ATLAS trigger system, along with a plot of the nominal trigger rates and processing times for various physics processes at each trigger level, is shown in Figure 3.11. A brief description of each trigger level follows.

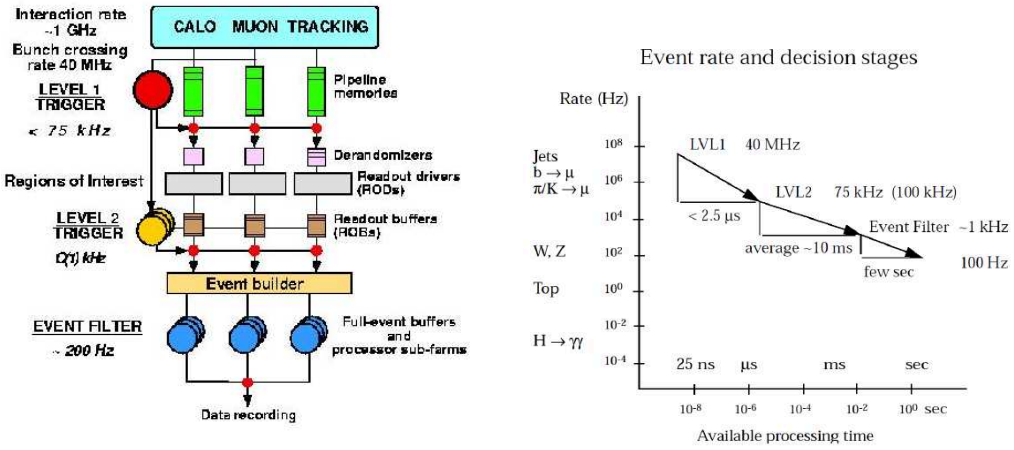


Figure 3.11: (a) Functional overview of the ATLAS TDAQ system and (b) the event rates and processing times for various physics processes at each trigger level.

Level-1 Trigger

The LVL1-trigger is a hardware-based trigger. It uses information from the calorimeters with reduced granularity and from the RPC and TGC muon trigger chambers. The latency of the LVL1-trigger is 2 ms, which leads to a target rate of 75 kHz. This level also defines regions of interests (RoI) for each event. The RoIs are regions in the detector where possibly interesting objects might be present, e.g. a high energetic electron. The LVL1-trigger passes the event information within the RoIs from the readout buffers (ROBs) to the LVL2-trigger.

Level-2 Trigger

The LVL2-trigger is software based and uses the full granularity in the RoIs of the detector, including the inner detector. The latency is from 1 ms to 10 ms depending on the complexity of the event, leading to a target rate of 1 kHz. Full access to the detector information by the LVL2-trigger for the event would exceed the required maximum latency, thus the concept of RoIs had to be introduced. The disadvantage of this approach is that interesting objects which have failed the LVL1-trigger cannot be found at LVL2. If an event passes the LVL2-trigger requirements, all information of the event is collected from the ROBs by the Event Builder and passed on to the Event Filter.

Event Filter

The Event Filter (EF) makes the final decision whether an event is recorded for further analysis. Its output target rate is on the order of a few hundred Hz. The EF is software based and runs on a computer farm near the ATLAS cavern. This gives the EF on the order of one second to make a decision. Given this relatively long decision time, the EF is able to access to the full event information with maximum granularity. More sophisticated reconstruction algorithms, similar to those used offline, can be applied. Events which are accepted by the EF are written to mass-storage devices for offline analysis.

CHAPTER 4

Monte Carlo Simulation and Event Generators

Monte Carlo (MC) simulation plays an important role in particle physics analyses by modelling the particle interactions in the experiment. Predictions from MC generators are supplemented with detector simulation in order to compare theoretical predictions to experimental results. Since many different MC generators are used in this analysis to provide predictions of signal and background events, it is important to have an understanding of MC simulation. This chapter gives a description of MC generators and details the production chain for simulating events at ATLAS.

There are three general types of MC generators usually used in particle physics studies: multipurpose leading order (LO) hadronization and parton showering generators, "multi-leg" LO generators for multi-parton final states, and next-to-leading order (NLO) generators. For a complete description of the event, MC generators need to be interfaced to models of soft physics processes such as underlying event (UE) and multiple particle interactions (MPI), parton shower (PS), hadronization and decays, and initial and final state radiation.

Multipurpose PS generators, such as Herwig [73] and Pythia [13], are based on calculations of the LO hard process. Higher order corrections, which are enhanced by large logarithmic terms, are modelled by parton showers. Parton showers result in the production of extra jets in the event, but mainly account for soft and collinear emissions, and not for hard large angle emissions.

Multi-leg generators, such as Alpgen [3] and Sherpa [4], are parton-level generators which use fixed-order matrix element (ME) calculations in lowest order in perturbation theory, not

including virtual corrections, for several exclusive multi-parton final states, e.g. $Z/\gamma^* + N_{jet}$ where $N_{jet} = 0, 1, 2, 3, 4, 5$. Some multi-leg generators, such as Alpgen, do not incorporate any kind of hadronization, resulting in final states containing bare quarks and gluons. Results from these generators need to be supplemented by hadronization and parton showering programs to obtain a complete model of the physics process.

The phase space for jets at low p_T and small angle emission is dominated by parton shower jets, whereas the phase space at high p_T and large angle emission is dominated by jets from the hard matrix element. A small overlap exists between these two regions of phase space. In order to avoid double counting jets during the merging of ME predictions with PS, ME+PS matching algorithms, like CKKW [54] or MLM [52, 53], are needed. These algorithms allow the combination of tree-level matrix elements for hard large angle emissions with a parton shower for soft and collinear emissions.

NLO generators, such as MC@NLO [95] and Powheg [75, 77], are parton-level generators which are based on hard matrix element calculations including the full NLO QCD corrections. They are expected to give a better description of the hard process, and a give a better prediction of the overall rate compared to LO ME+PS generators and standalone parton shower generators such as Pythia. In order to provide a more complete model for the physics process, NLO generators have to be interfaced to hadronization and showering programs as well.

The full production chain for simulating MC events at the ATLAS experiment consists of 5 steps, beginning with the generation process. The intermediate steps are the event simulation, digitization, and reconstruction. The final step ends with the creation of Analysis Object Data (AOD). All steps in the production process are performed within the Athena framework [96], which is a flexible component-based architecture designed for a wide range of distributed data-processing applications.

In the first step, particle four-vectors for desired physics processes are generated and saved as HepMC [28], which is an object oriented event record written in C++ for high energy physics MC generators. The MC generators used in this analysis for signal and background modelling are described in more detail in the following sections. The four-vectors are then passed through the GEANT4 [97, 98] simulation or the ATLFAST II [99] fast detector simulation. The latter replaces the more CPU intensive steps (simulation, digitization, and reconstruction) by faster, approximate simulations. This allows for the production of more statistics in the same amount of time, but with the drawback of a less precise overall detector simulation. The digitization step is used to transform the output from the GEANT4 simulation into the actual output format of the detector, the Raw Data Objects (RDOs). In the final step, which is identical for both MC simulated events and real data, the digitized output is reconstructed as tracks and energy deposits, and stored as Event Summary Data (ESD). For analysis purposes, a reduced version of ESDs containing a summary of reconstructed events is saved as AODs that can be analyzed within the Athena framework. Another data format, D3PDs (Derived Physics Data), can be created from AODs and easily accessed within the ROOT analysis framework [100].

4.1 Pythia

Pythia [13] is a LO multi-purpose parton showering and hadronization generator, which can be used to simulate hard and soft interactions in hadronic and leptonic collisions. It contains a large library of about 240 different subprocesses. Pythia describes initial and final state radiation, UE and MPI, hadronization as well as decays. Initial and final state radiation is simulated using either Q^2-ordered or p_T-ordered parton showers. For hadronization, the (Lund) string fragmentation model [55, 56] is used (see Section 2.2.7). MPI and UE simulations are modelled using a p_T-ordered parton shower and string fragmentation model, whose governing parameters can be "tuned" with input from measurements using the Perugia Tunes [46].

4.2 Herwig

Herwig [73], just as Pythia, is a leading order multipurpose hadronization and showering generator, which can be used to simulate hadronic and leptonic collisions. It also contains a large library of hard scattering processes. Herwig includes initial and final state radiation, UE and MPI simulation, hadronization as well as decays. In contrast to Pythia, Herwig uses an angular-ordered parton shower model for the description of the initial and final state radiation. For hadronization, Herwig uses the cluster hadronization model [58–61] (see Section 2.2.7). The UE simulation in Herwig is based on the high energy proton-proton simulation program of the UA5 collaboration [101], incorporating the cluster hadronization model. In addition, Herwig can be interfaced to Jimmy [102,103] for the UE simulation.

4.3 Alpgen

Alpgen [3] is a multi-leg generator and has been designed to generate hard multi-parton processes in hadronic collisions. It provides fixed order matrix element calculations at tree-level in QCD and EW interactions for a large set of processes, such as the production of W/Z bosons with up to 5 associated partons in the final state. Alpgen can be interfaced to Herwig or Pythia to evolve the event by parton shower and finally form hadrons from the partonic cascades. In order to avoid double counting when merging the tree-level matrix elements with the parton shower jets, the MLM [52,53] matching scheme is used which relies on a ΔR match between final state partons from the matrix element calculation and parton shower jets. For QED radiative corrections, Alpgen is interfaced to PHOTOS [104].

4.4 Sherpa

Sherpa [4], just as Alpgen, is a multi-leg generator that can calculate tree-level matrix elements for a wide range of hard scattering processes in hadronic collisions. Sherpa provides its own algorithms for parton shower [105, 106], initial and final state radiation, UE and

MPI [107], and hadronization and decays. For the formation of hadrons, Sherpa uses a modified cluster hadronization model [108], which is based on the model used in Herwig. In order to describe the effects of QED corrections, Sherpa incorporates the YFS [109] approach.

Sherpa has the option to generate the hard process using the MENLOPS approach [5, 110, 111]. This approach combines NLO accuracy for the inclusive process with leading-log accuracy from parton shower resummation. In the MENLOPS approach, the hardest parton emission is generated at NLO. To consistently combine the NLO matrix element with the QCD parton cascades, the CKKW [106] merging scheme is employed. Multi-parton final states beyond the first hard emission are generated at LO in QCD and merged with the parton shower.

4.5 Powheg

Powheg stands for POsitive Weight Hard Emission Generator [75, 77]. It generates the hard matrix element at NLO for inclusive quantities which results in the the hardest emission generated at NLO. The Sherpa framework employs Powheg in the MENLOPS approach to generate the hard process at NLO. The modelling of the parton shower and non-perturbative effects is obtained by interfacing the output from Powheg with any parton shower MC generator. In order to not double count the first hard emission when merging with the parton shower, the parton shower MC generator must use a p_T-ordered algorithm.

CHAPTER 5

Electron and Jet Definitions

The identification, reconstruction, and calibration of electrons and jets in ATLAS is a very complex procedure involving a multitude of steps. Since the analysis presented in this thesis focuses on final states containing electrons and jets, an in-depth understanding of these processes is necessary.

The identification and reconstruction of physical objects are based on the information obtained from the relevant subdetector systems. The increase in the average number of bunch crossings from 2011 to 2012, as shown Figure 3.3, has necessitated the development of more refined methods that can perform efficiently in high-activity environments. This chapter summarizes the techniques and algorithms used by the ATLAS detector to identify and reconstruct electrons and jets in proton-proton collisions.

5.1 Electrons

The electron reconstruction and identification is performed separately for electrons in the central region of $|\eta| < 2.47$, which is within the tracking volume of the inner detector, and for electrons in the forward region, $|\eta| = 2.5 - 4.9$. For central electrons, the reconstruction and identification is based on information from the electromagnetic (EM) calorimeter and the inner detector tracking system, whereas for forward electrons, only information from the EM calorimeter can be used since the tracking detectors only cover a range up to $|\eta| < 2.5$. The following subsections detail the reconstruction and identification algorithms for central electrons, as well as the performance of the electron calibration. The algorithms for forward

electrons will not be described since the analysis presented in this thesis is performed with central electrons only.

5.1.1 Reconstruction

The reconstruction of electrons [112] in the region $|\eta| < 2.47$ starts with clusters reconstructed from energy deposits within the EM calorimeter. The EM calorimeter is divided into a grid of $N_\eta \times N_\phi = 200 \times 256$ towers of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$, corresponding to the granularity of the middle layer of the EM calorimeter. Inside each tower, the energy of all cells in all longitudinal layers (strip, middle, back, and presampler for $|\eta| < 1.8$) is summed into the tower energy. The total EM energy deposited in an EM calorimeter cell has been calibrated from test-beam studies [113]. This is referred to as the EM scale, and is the fundamental starting energy calibration of the Liquid Argon (LAr) calorimeter.

Next, clusters are seeded by towers with total transverse energy above 2.5 GeV and searched for by a sliding-window algorithm [114], with a window size of 3×5 towers. The size of the window and the threshold are optimized to obtain the best efficiency for finding clusters, and to limit the rate of fake clusters due to noise. The cluster is large enough so that it contains most of the energy deposited by the particle in the calorimeter, thus limiting the effect of lateral shower fluctuations on the energy resolution. However, including more cells also means including more noise

Clusters that are matched to a well-reconstructed ID track in a window of size $\Delta\eta \times \Delta\phi = 0.05 \times 0.10$ and originating from a vertex found in the beam interaction region are classified as electrons. For the case in which more than one track matches the seed cluster, the track closest in $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ to the seed cluster is used. Tracks containing silicon hits from the pixel or SCT detectors are preferred rather than tracks which only contain hits from the TRT.

The electron cluster is then rebuilt using a window of calorimeter cells in the second layer corresponding to 3×7 and 5×5 cells in the barrel and endcaps respectively. Since the magnetic

field curves the electron trajectory in the ϕ direction, the ϕ size of the cluster is increased in order to contain most of the energy. Similarly, converted photons lead to electron-positron pairs that spread in the ϕ direction due to the magnetic field. In the endcaps, because the effect of the magnetic field is smaller, the cluster size is the same for all particle types. The number of cells in η is larger in the endcaps than in the barrel because of the smaller physical cell size in the endcaps.

The cluster energy is then determined by applying correction factors computed by a calibration scheme based on the full detector simulation as described in Section 5.1.3. The energy of the electron is obtained from the calorimeter cluster, while the ϕ and η directions are taken from the tracking information from the inner detector of the track associated to the electron cluster. In the cases where the track does not contain any silicon hits, the ϕ direction is determined from the associated track, and the η direction is given by the η position of the EM cluster.

5.1.2 Identification

Not all objects built by the electron reconstruction algorithms are signal electrons. Background objects include hadronic jets as well as electrons from photon conversions, electrons from leptonic meson decays, and electrons from semileptonic heavy flavour hadron decays. In order to reject as much of these backgrounds as possible while keeping the efficiency for signal electrons high, electron identification in ATLAS is based on discriminating variables, which are combined into a menu of selections of varying background rejection. Both cut-based and multivariate analysis techniques are used.

Due to its simplicity, the cut-based electron identification [115–118], which is based on sequential cuts on selected variables, has been used in ATLAS for identifying electrons since the beginning of data-taking. In the central region of $|\eta| < 2.47$, variables describing the longitudinal and transverse shapes of the electromagnetic showers in the calorimeters, the number of tracking points provided by the inner detector, as well as the matching between

tracks and energy clusters are used to discriminate against the different background sources. In 2011, the cut-based performance (efficiency / background rejection) was improved by loosening cuts on existing variables and introducing cuts on additional variables [118]. In 2012, due to the higher instantaneous luminosity, the amount of pile-up increased, and therefore the number of particles in an event increased. Furthermore, due to the higher energy per event, the electron showers tend to look more often like jets. In order to cope with this, cuts were loosened on the variables most sensitive to pile-up and tightened on others to keep the identification efficiency roughly constant as a function of the number of reconstructed primary vertices.

The different cut-based selections available for physics analyses, labelled as *loose*, *medium*, *tight*, and *multilepton*, are optimized in 10 bins in $|\eta|$ and 11 bins in E_T. This binning allows one to take into account the variation of the electrons' characteristics due to dependencies such as shower shapes and the amount of passive material traversed before entering the electromagnetic calorimeter. Shower shapes and track properties also change with the energy of the particle. The *tight* selection is a subset of *medium*, while *medium* is a subset of *loose*. With increasing tightness, more variables are added and cuts are tightened on the variables already used in the looser selections. The rejection power of jets faking electrons has been determined from MC simulation to be on the order of 500, 5000, and 50000 with respect to the number of truth jets in the MC sample for the *loose*, *medium*, and *tight* operating points respectively [116]. In 2012, an additional cut-based operating point was added, called *multilepton*, which was optimized for the $H \to ZZ^* \to 4\ell$ analysis involving multiple low energy electrons in the final state. For these electrons, *multilepton* has a similar efficiency to the *loose* operating point, but a better background rejection. In comparison to *loose* , cuts on the shower shapes are loosened and more variables added, including those measuring bremsstrahlung effects.

In this thesis, the *loose* and *medium* cut-based selections are used. The *loose* selection is used in a data-driven method to measure the multijet background. The *medium* selection is used

to select signal electrons with high efficiency and high jet rejection. The *tight* selection would increase the jet rejection but with a significant reduction in the number of signal statistics. The selection criteria of these operating points are described in Table 5.1.

5.1.3 Calibration

The energy scale and energy resolution of electrons have been determined using in-situ techniques on $Z \to e^+e^-$ decays for central electrons in $\sqrt{s} = 8$ TeV data collected in 2012 [119], and then tested on $\sqrt{s} = 7$ TeV data from 2011 as a consistency check. The procedure shifts its focus from the calibration of the energy deposited in the individual cells of the EM calorimeter, as mentioned in Section 5.1.1, to the absolute energy scale calibration on fully reconstructed electrons in the detector. The calibration procedure allows for η and E_T dependent energy corrections, accounts for energy losses in the passive material in front of the calorimeter, for energy deposited in cells of neighbouring clusters, and implements a set of corrections to account for response variations not included in the detector simulation.

The energy of an electron candidate is built from the energy of a cluster of cells in the EM calorimeter calibrated to the EM energy scale. The calibration procedure can then be broken down into four standard steps. An additional fifth calibration step is used to correct electron trigger, reconstruction, and identification efficiencies to account for differences between data and MC.

1) Intercalibration of the LAr calorimeter layers

Since the EM calorimeter is longitudinally segmented, the energy scales of the different longitudinal layers must be equalized in data with respect to simulation prior to the determination of the overall energy scale. This is necessary to ensure the correct extrapolation of the response in the full p_T range used in the various physics analyses of the following MC-based calibration step. The intercalibration of the first and second EM calorimeter layers uses muons from $Z \to \mu\mu$ decays as probes since most muons in this energy range are minimum ionizing particles and are insensitive to the passive material upstream of the EM calorimeter.

Type	Description	Name
Loose Selection		
Hadronic leakage	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\| < 0.8$ or $\|\eta\| > 1.37$)	R_{Had1}
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \|\eta\| < 1.37$)	R_{Had}
Middle layer of EM calorimeter	Lateral shower width, $\sqrt{(\sum E_i\eta_i^2)/(\sum E_i) - ((\sum E_i\eta_i)/(\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells	$W_{\eta 2}$
	Ratio of the energy in 3×7 cells over the energy in 7×7 cells centred at the electron cluster position	R_η
Strip layer of EM calorimeter	Shower width, $\sqrt{\sum E_i(i - i_{max})^2/(\sum E_i)}$, where i runs over all strips in a window of $\Delta\eta \times \Delta\phi \approx 0.0625 \times 0.2$, corresponding typically to 20 strips in η, and i_{max} is the index of the highest-energy strip	ω_{stot}
	Ratio of the energy difference between the largest and second largest energy deposits in the cluster over the sum of these energies	E_{ratio}
Track Quality	Number of hits in the pixel detector	n_{pixel}
	Number of total hits in the pixel and SCT detectors	n_{Si}
Track-cluster matching	$\Delta\eta$ between the cluster position in the strip layer and the extrapolated track	$\Delta\eta_1$
Medium Selection (in addition to all *Loose* selection critieria)		
Back layer of EM calorimeter	Ratio of the energy in the back layer to the total energy in the EM accordion calorimeter	f_3
Track quality	Number of hits in the B-layer (discriminates against photon conversions)	ω_{stot}
	Transverse impact parameter	d_0
TRT	Total number of hits in the TRT	n_{TRT}
	Ratio of the numberof high-threshold hits to the total number of hits in the TRT	F_{HT}

Table 5.1: Definitions of the variables used for the *loose* and *medium* electron identification cuts for the central region of the detector, $|\eta| < 2.47$ [118].

The ratio between the energies deposited by the muons in the cells of the first and second layer are compared between data and MC in order to determine calibration constants. The pre-sampler energy scale is determined from the ratio of pre-sampler energies in data and MC using electrons from W and Z decays. No dedicated intercalibration of the third EM longitudinal layer is carried out as its contribution is negligible in the energy range covered by the calibration techniques.

2) MC-based calibration

Reconstructed electron clusters in data and MC simulation are calibrated to correct for the energy lost in the material upstream of the calorimeter, the energy deposited in the cells neighbouring the cluster in η and ϕ, and energy leakage beyond the LAr calorimeter. Further corrections are applied to correct for the response dependence as a function of the particle impact point within the central cluster cell. The cluster-level calibration constants are extracted from simulation using a multivariate algorithm.

3) Uniformity and stability corrections

A set of corrections are implemented to account for response variations not included in the simulation in specific detector regions. Four classes of effects are considered: high-voltage (HV) inhomogeneities due to short circuits in specific LAr gaps resulting in non-nominal HV values, variations in the pre-sampler response due to changes in the nominal HV operating point between run periods, energy loss between the barrel calorimeter modules due to a gravity-induced widening of the intermodule spacing, and a percent-level dependence of the energy response on the signal gain of the calorimeter electronics in two specific $|\eta|$ regions.

4) Energy scale and resolution determination with electrons from $Z \to e^+e^-$ decays

The overall electron response in data is calibrated so that it agrees with the expectation from simulation using a large sample of $Z \to e^+e^-$ decays. Scale factors are extracted for each electron and applied to electron candidates in data. The absolute scale determination is

carried out as a function of η only, and independently of azimuthal angle and run period. The energy miscalibration is defined as the difference in response between data and simulation, and is parameterized as $E^{data} = E^{MC}(1 + \alpha_i)$, where E^{data} and E^{MC} are the electron energy in data and MC, and α_i represents the departure from the optimal calibration in a given pseudorapidity bin i. The resulting energy scale corrections are illustrated in Figure 5.1 as a function of η. The energy scale corrections are large in the transition regions between the barrel and endcap calorimeters $|\eta| = 1.37 - 1.52$.

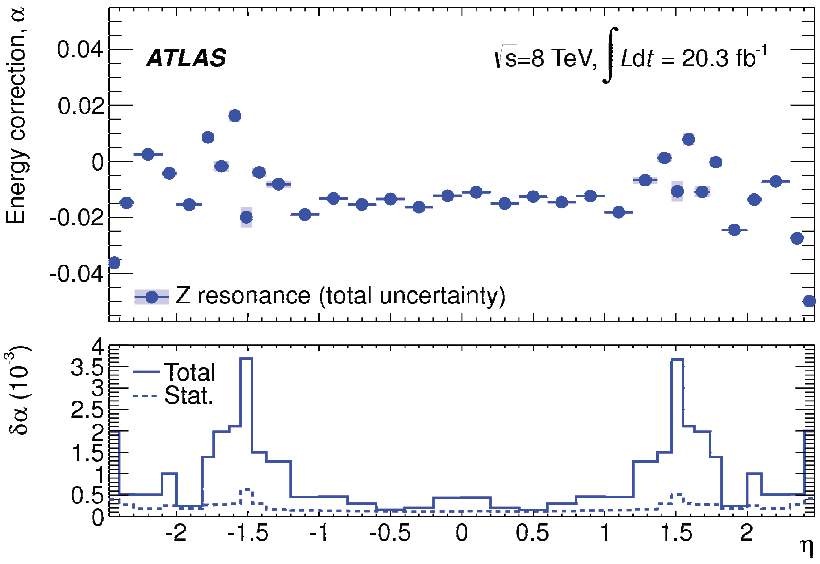


Figure 5.1: The energy scale corrections α as a function of η, defined as $E^{data} = E^{MC}(1+\alpha)$, and derived from $Z \to ee$ events. The error bands include statistical and systematic uncertainties [119].

Using the same event sample, it was found that the resolution in data is slightly worse than that in simulation, which results in the mass peak being slightly broader in data than in the MC. This effect is illustrated in Figure 5.2 which shows the invariant mass distribution for $Z \to ee$ decays in data and compared to MC before and after resolution correction factors are are applied to MC.

The electron resolution curve is assumed to be well modelled by simulation up to a constant correction term,

$$\left(\frac{\sigma_E}{E}\right)^{data} = \left(\frac{\sigma_E}{E}\right)^{MC} \otimes c \tag{5.1}$$

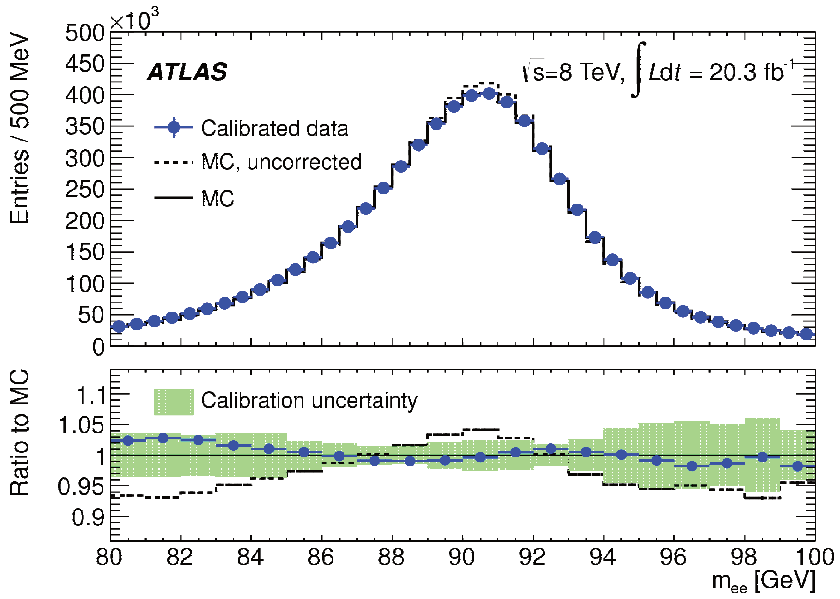


Figure 5.2: Electron pair invariant mass distribution for $Z \to ee$ decays in data and improved simulation. Energy scale corrections are applied to the data. The distribution is shown before and after energy resolution corrections, and is normalized to the number of events in data [119].

The resolution correction terms, c, are derived and applied to simulation in order to match the mass distribution observed in data. The corrections to the expected resolution as a function of η are illustrated in Figure 5.3.

The width of the invariant dielectron mass distribution gives an estimate of the electron energy resolution. The resolution curve and its uncertainty as a function of E_T for electrons with $|\eta| = 0.2$ is shown in Figure 5.4. For electrons with E_T around 40 GeV, which is the average transverse energy for electrons from Z decays, the total resolution at $|\eta| = 0.2$ is about 2%.

5) Electron trigger, reconstruction, and identification efficiency corrections using tag-and-probe methods

The accuracy with which the MC-based detector simulation models the electron efficiency plays an important role in cross section measurements and various searches for new physics. In order to achieve reliable physics results, the MC samples need to be corrected to reproduce the measured data efficiencies as closely as possible. This is achieved by a multiplicative

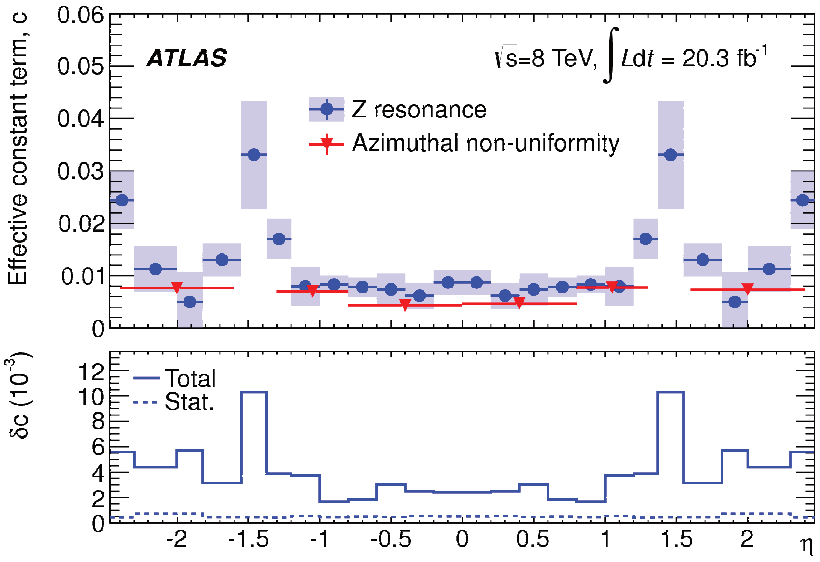


Figure 5.3: The effective constant term corrections c to the expected resolution derived from $Z \to ee$ as a function of η. The error bands include statistical and systematic uncertainties [119].

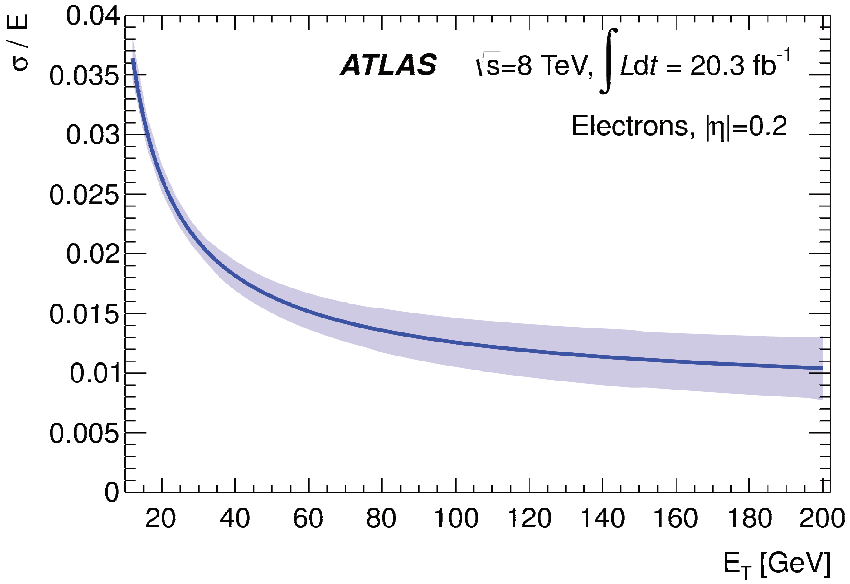


Figure 5.4: The resolution curve and its uncertainty as a function of E_T for electrons with $|\eta| = 0.2$ [119].

correction factor defined as the ratio of the efficiency measured in data to that in simulation. These data-to-MC correction factors are usually close to unity. Deviations stem from the mismodeling of tracking properties or shower shapes in the calorimeters. The efficiency to detect an electron in the ATLAS detector is not measured as a single quantity but is divided into different components, namely trigger, reconstruction, and identification efficiency. The full efficiency ϵ_{total} for a single electron can be written as:

$$\epsilon_{total} = \epsilon_{reconstruction} \times \epsilon_{identification} \times \epsilon_{trigger} \tag{5.2}$$

Residual effects due to differences in the kinematics of the physics processes used in the measurements of each efficiency are expected to cancel out in the data-to-MC efficiency ratio. Therefore, the combination of the different efficiency measurements is carried out using the data-to-MC ratios instead of the efficiencies themselves. Since the electron efficiencies are a function of the transverse energy and pseudorapidity, the measurements are performed in two-dimensional bins in (E_T, η) using the tag-and-probe method on $Z \to e^+e^-$ and $J/\Psi \to e^+e^-$ [118]. In both cases, strict selection criteria are applied on one of the two decay electrons (called tag), and the second electron candidate (probe) is used for the efficiency measurements. Figure 5.5 shows as an example the electron identification efficiency as measured in data as a function of η and compared to MC for the various cut-based identification selections in the E_T bin of 45 - 50 GeV.

5.2 Jets

Quarks and gluons, produced as final state partons in the hard interaction of proton-proton collisions, hadronize and manifest themselves in the detector as a spray of particles in a cone around the original direction of the quark or gluon. This spray of particles is called a jet, and is the physical object which is measured. Therefore, the kinematic properties of the jet represent those of the final state parton which initiated the particle shower. However, since the jet is not a true physical object, but a collection of hadrons and other particles, its attributes, such as pseudorapidity and transverse energy, as well as the jet multiplicity

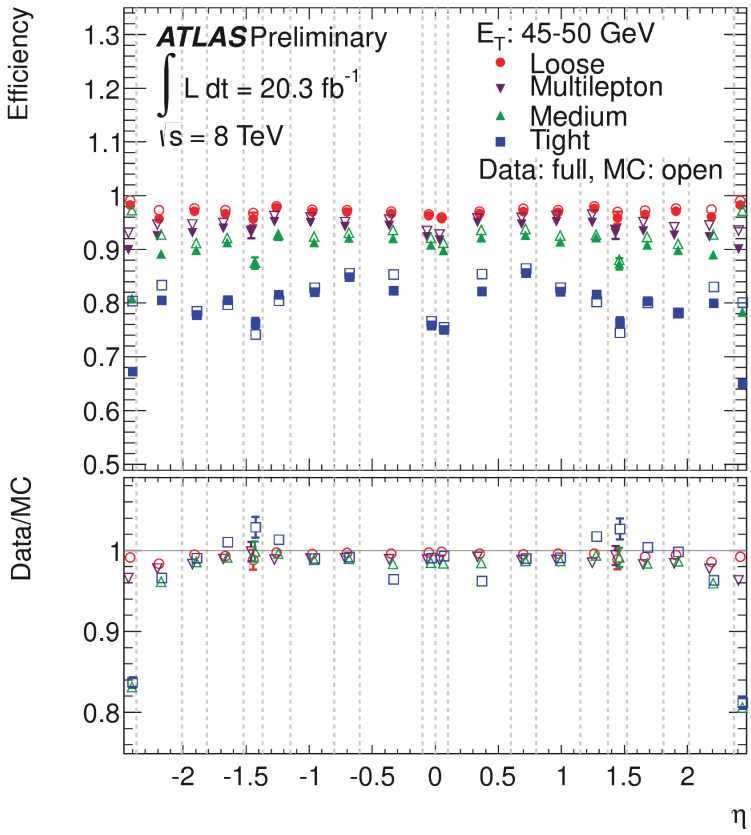


Figure 5.5: Identification efficiency in data as a function of η for E_T = 45 - 50 GeV for the cut-based loose, multilepton, medium and tight selections, compared to the MC expectation for electrons from $Z \to ee$ decay [119].

of the event, depend on the reconstruction algorithms used to build the jet. The following subsections detail the reconstruction and calibration of jets, as well as the jet quality criteria and the performance of the energy and resolution measurements.

5.2.1 Reconstruction

Jets are reconstructed using the anti-k_t algorithm [120] with the radius parameter set to $R = 0.4$ using the FastJet software package [121]. The inputs to the jet algorithm can be simulated "truth" particles (truth jets), reconstructed tracks in the inner detector (track jets) or energy deposits in the calorimeter (calorimeter jets) from both data and simulation. Truth jets are simulated particles with a lifetime τ, defined by $c\tau > 10$ mm, that are neither final-state muons nor neutrinos. Calorimeter jets can be reconstructed from various different inputs, such as calorimeter towers or topologically connected calorimeter clusters (topo-clusters) [122]. In this analysis, locally calibrated topo-clusters are used as inputs to the jet finding algorithm. A schematic overview of the ATLAS jet reconstruction is shown in Figure 5.6.

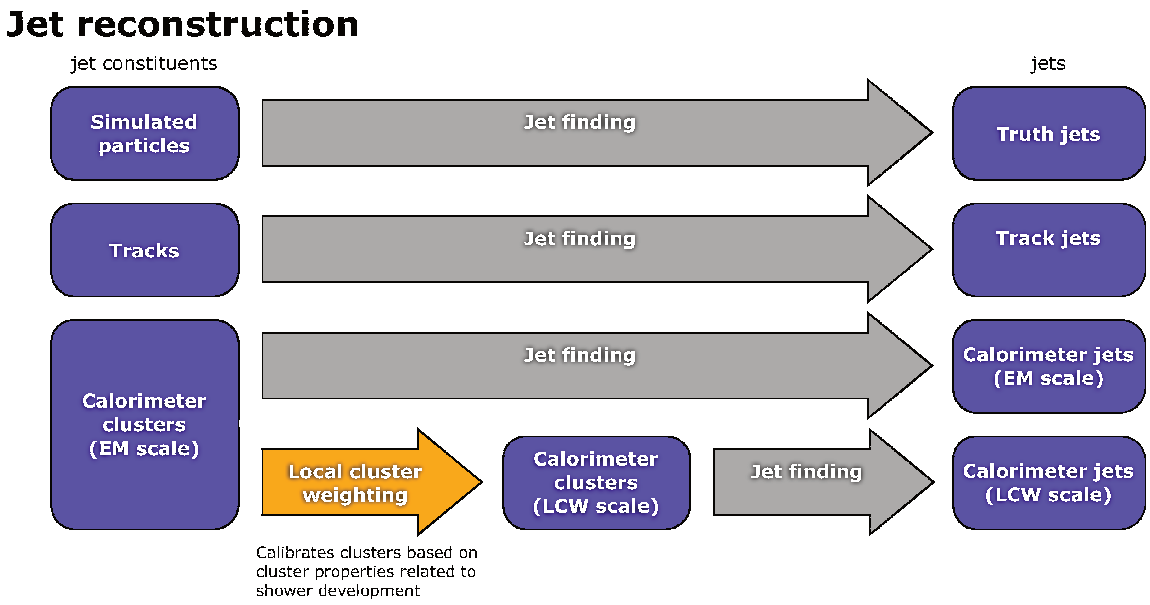


Figure 5.6: Schematic overview of the ATLAS jet reconstruction. After the jet finding, the jet four momentum is defined as the four momentum sum of its constituents.

Topo-clusters are initially reconstructed at the EM scale [113, 118, 119], correctly measuring the energy deposited in the calorimeter by particles produced in electromagnetic showers. The formation of topo-clusters follows a 4-2-0 clustering algorithm [123]. Starting with cells which have an energy greater than four times the electronic plus pile-up noise as seeds, the algorithm then proceeds with the recursive aggregation of all neighbouring cells which have a signal-to-noise ratio above two. A second topo-cluster collection is built by calibrating the calorimeter cells such that the response of the calorimeter to hadrons is correctly reconstructed [124, 125]. This calibration uses the local cell signal weighting (LCW) method [126] that improves the resolution compared to the EM scale by correcting the signals from hadronic energy deposits for fluctuations due to the non-compensating nature of the ATLAS calorimeter. The LCW method classifies topo-clusters as either electromagnetic or hadronic, primarily based on the measured energy density and the longitudinal shower depth. Energy corrections, derived from MC simulations of single charged and neutral pion interactions with the detector, are applied to topo-clusters classified as hadronic. The jet energy scale (JES) of calorimeter jets, built by running the jet finding algorithm on either the EM-scale or LCW-scale calibrated topo-clusters, are referred to as EM+JES or LCW+JES respectively.

Due to its accuracy for resolving jets, the anti-k_t jet finding algorithm is used to combine topo-clusters into jets. Accurate jet reconstruction is necessary for determining the kinematics of the unobserved originating partons. Some aspects of jet finding algorithms that need to be considered are the jet size and whether the algorithm is infra-red and collinear (IRC) safe. The jet size and area determine the susceptibility of a jet to soft radiation. A larger jet radius is important as it allows enough of the jet's particles to be captured by the algorithm for the accurate calculation of the jet's mass and energy. However, a smaller jet radius is useful in reducing the amount of the underlying event and pile-up captured by the jet, preventing the overestimation of the jet's mass and energy. An IRC safe algorithm will prevent the splitting of a jet if a soft gluon is emitted or if a hard particle splits into two softer particles.

There are two main classes of jet algorithms in use: iterative cone algorithms, and sequential clustering algorithms. Cone algorithms assume that particles in jets will show up in conical regions, and thus cluster in $\eta - \phi$ space, resulting in jets with rigid circular boundaries. Cone algorithms are generally IRC unsafe as well. On the other hand, sequential clustering algorithms, such as the anti-k_t algorithm, are IRC safe. Sequential clustering algorithms assume that particles within jets will have small differences in transverse momenta, and thus cluster particles based on momentum space, resulting in jets that have fluctuating areas in $\eta - \phi$ space. However, the anti-k_t algorithm prefers to cluster hard particles first, thus the area only fluctuates slightly and the algorithm is only mildly susceptible to the underlying event and pile-up.

5.2.2 Calibration

Calorimeter jets reconstructed by the jet finding algorithm using the EM or LCW jet energy scale are calibrated further using the procedure as outlined in Figure 5.7. The calibration consists of five sequential steps: a jet origin correction, a pile-up correction, a MC-based energy and η correction, a global sequential calibration, and a final in-situ correction applied only to data. This procedure is based on the jet energy scale calibration employed for the 7 TeV data set from 2011 [124], which considered only the energy and pseudorapidity of the jets. The global sequential (GS) calibration [127] was developed as an extension to this baseline calibration for the 8 TeV data set. The GS calibration makes use of additional jet variables, such as jet flavour, jet charge, jet sub-structure, and tracking information from the inner detector in order to improve the resolution of the jet response and decrease the sensitivity to jet flavour. The different steps of the jet calibration are outlined below.

1) Jet origin correction

The direction of the jet is adjusted such that it points back to the primary event vertex instead of the nominal center of the ATLAS detector. The primary vertex is defined as the one with the highest $\sum (p_T^{track})^2$, where the scalar sum is taken over all tracks associated to

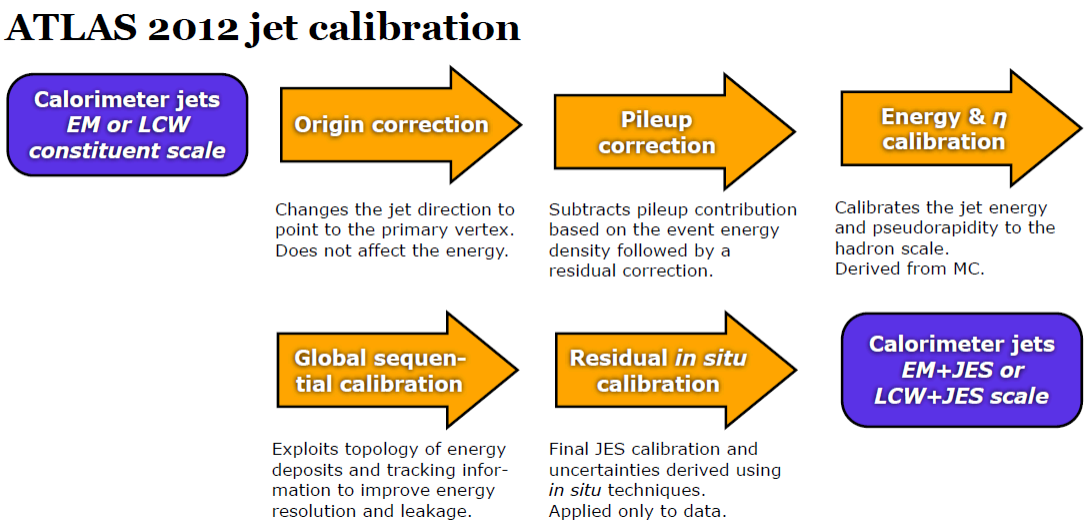


Figure 5.7: Schematic overview of the ALTAS jet calibration procedure. All steps are derived and applied separately for EM+JES and LCW+JES calorimeter jets.

the vertex. This vertex selection criteria was found to be correct in at least 98% of the events used in the calibration studies [128].

2) Pile-up correction

The pile-up dependence of the calorimeter response to jets is reduced by subtracting the energy offset caused by both in-time and out-of-time pile-up interactions. The contribution from pile-up is estimated on an event-by-event and jet-by-jet basis as the product of the p_T-density of pile-up in the event ρ and the active area of the jet A^{jet}. The active area is determined using the active area implementation [129], where the reconstructed jet collection is rebuilt after adding a uniform density of randomly distributed particles in η and ϕ with infinitesimal p_T, called ghost particles. The addition of ghost particles does not affect the four momentum of the jet. The number of ghost particles clustered into each jet by the jet-finding algorithm is a measure of the active area of the jet. The active area of the jet is a representation of the average susceptibility of the jet to a high density of soft radiation (e.g. pile-up), since the many soft particles will cluster between each other and into jets much in the same way as do the ghost particles. The pile-up density ρ is estimated on an event-by-event basis by taking the median of p_T/A^{jet} for all such jets.

Any remaining residual dependence of the jet p_T on pile-up is removed by corrections that are parametrized as a function of the number of primary vertices N_{PV}, and the expected average number of interactions $\langle\mu\rangle$. The parametrization is derived using MC simulation.

3) MC Based Jet Energy and η calibration

After the origin and pile-up corrections have been applied, a baseline jet energy scale correction is applied that corrects the reconstructed jet energy to the true particle level jet energy. The energy correction factors are determined from MC simulations in narrow bins of pseudorapidity in the detector frame of reference η_{det} by comparing the energy of the reconstructed jet to the energy of its associated truth jet at particle level. Truth jets are jets in an event from MC simulation that have not passed through a simulation of the detector. Reconstructed jets are the objects which are reconstructed and measured by the detector. Thus, the properties of truth jets represent the true properties of the associated reconstructed jets. An angular matching of $\Delta R < 0.3$ is used to associate a particle level jet to a reconstructed jet. The jet response $\langle R_{jet}\rangle$ is defined as the peak position of a Gaussian fit to the E_{reco}/E_{truth} distribution. Figure 5.8 shows the jet energy response $\langle R_{jet}\rangle$, which is the inverse of the calibration function, after the application of the jet energy scale calibration as determined from MC simulation for the 2011 JES calibration. This calibration restores the measured jet energy to the correct hadronic energy scale. Good closure, which quantifies the level of agreement in MC between the corrected hadronic jet energy and the true jet energy, is demonstrated across the full pseudorapidity range. However there is some small disagreement between the corrected hadronic jet energy and the true jet energy, called non-closure, for low p_T jets due to non-perfect fits arising from non-Gaussian energy response.

4) Global sequential (GS) calibration

The GS calibration [127] is an extension of the baseline jet calibration outlined above. Whereas the baseline calibration considers only the jet energy and pseudorapidity in the determination of the jet energy scale, the GS calibration makes use of global jet observables

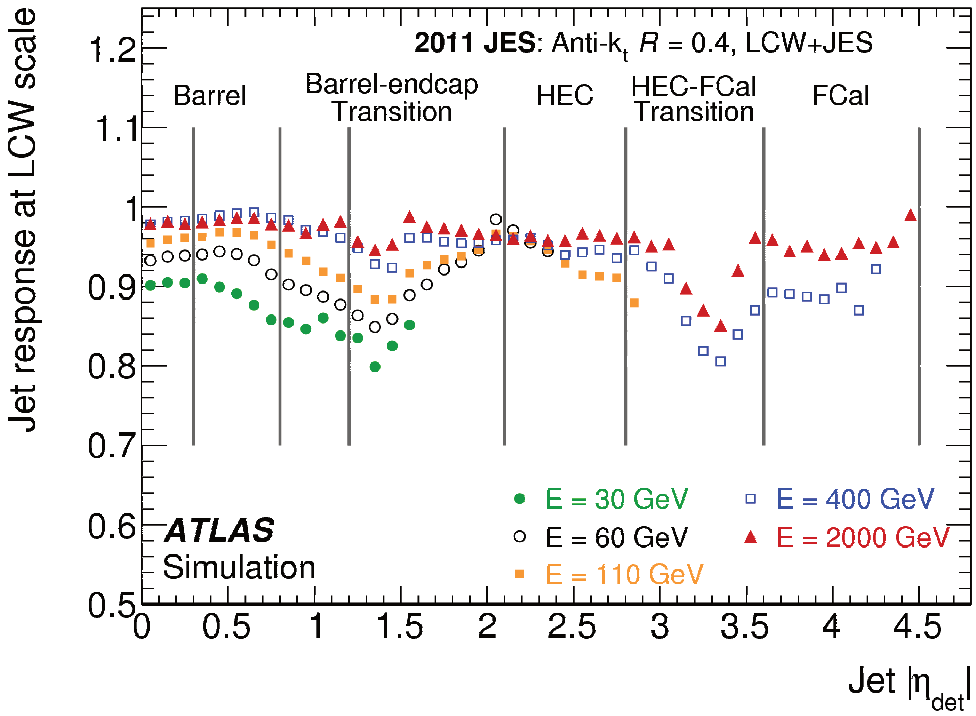


Figure 5.8: Average response of simulated jets formed from topo-clusters, calculated for the LCW scale for the 2011 JES calibration. The response is shown separately for various truth-jet energies as function of the uncorrected (detector) jet pseudorapidity $|\eta_{det}|$. The inverse of the response corresponds to the average jet energy scale correction in each $|\eta_{det}|$ bin [124].

such as the the longitudinal structure of the energy depositions within the calorimeters, tracking information associated to the jet, and information related to the activity in the muon chambers behind a jet, in order to improve the energy resolution and reduce the sensitivity of the response to jet flavour. In principle, any variable that carries information about the response of the jet can be used. The correction factors are binned in $|\eta_{det}|$, and constructed as a function of p_T^{jet} and the global jet property being exploited. The corrections are derived from MC simulations of inclusive QCD jet production, and are designed such that the mean jet energy is left unchanged in the sample used to derive the corrections. This reduces the spread of the reconstructed jet energy, thus improving the resolution.

5) In-situ calibration (data only)

Following the MC-based jet calibration, in-situ techniques employing the energy balance of physics objects in the transverse plane, are used in the final step of the JES calibration. The energy balance technique is applied to dijet, multijet, photon+jet and Z+jet events. The p_T of well measured reference objects (photons, Z bosons, or other jets) recoiling against the jet being calibrated are compared in data and MC simulations to measure the ratio

$$\langle p_T^{jet}/p_T^{ref}\rangle_{data}/\langle p_T^{jet}/p_T^{ref}\rangle_{MC}. \tag{5.3}$$

This quantity defines a residual correction binned in p_T and $|\eta|$ which is applied to jets reconstructed in data.

5.2.3 Quality Criteria

Jets with high transverse momenta produced in proton-proton collisions must be distinguished from background jet candidates (fake jets) not originating from hard-scattering events. The main sources of these fake jets are:

1. Beam-gas events, where a proton of the beam collides with residual gas within the beam pipe.

2. Beam-halo events, where outlying particles in the beam line collide with the tertiary collimators.

3. Cosmic ray muons overlapping in-time with collision events.

4. Calorimeter noise.

Much like the electron identification in ATLAS, jet quality is based on a menu of selections of varying fake-jet rejection and jet selection efficiencies [124, 130]. Four sets of cut-based selections are available for physics analysis, labelled as *Looser*, *Loose*, $Medium$, and $Tight$. They correspond to different levels of fake-jet rejection and jet selection efficiencies, with the *Looser* criterion being the one with the highest jet selection efficiency while the $Tight$ criterion is the one with the best rejection. The discrimination between jets produced in the hard collision and fake jet candidates is based on several variables, including the quality of the energy reconstruction at the cell level, jet energy deposits along the direction of the shower development, and reconstructed tracks that are matched to jets. The efficiency of the different quality selections have been measured in data using tag-and-probe methods in dijet events [124]. Figure 5.9 shows the jet quality selection efficiency in the 7 TeV data set from 2011 for anti-k_t jets with R = 0.4 in the region $|\eta| < 0.3$ as an example.

The jet selection efficiency of the *Looser* selection is greater than 99.8% over all calibrated p_T^{jet} and $|\eta|$ bins. A slightly lower efficiency is measured for the *Loose* selection. A very good agreement between data and simulation is observed for the *Looser* and *Loose* selections. The jet quality selection criteria remained unchanged for the 8 TeV data set. For this analysis, the *Looser* selection is applied as it has the highest jet selection efficiency.

Furthermore, jets can originate from additional proton-proton interactions per bunch crossing. These jets typically have a low transverse momenta. In order to reject these jets, the jet vertex fraction (JVF) is defined as the ratio of the sum of the p_T of the tracks associated to the jet and belonging to the primary vertex of the hard interaction, to the sum of the p_T of all the tracks associated to the jet. Since the JVF requires tracking information, it is only

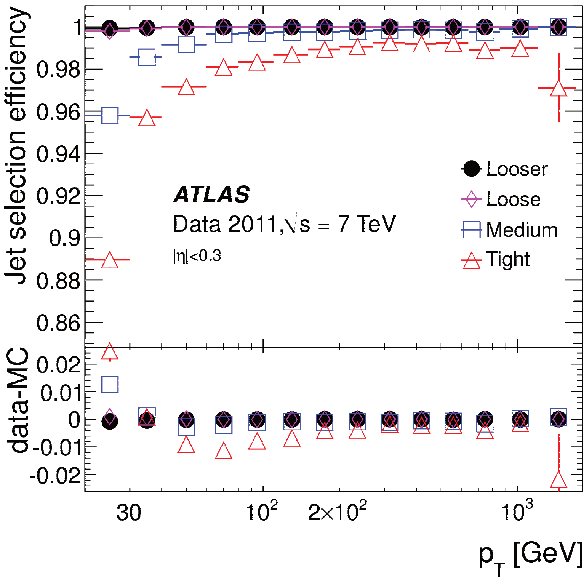


Figure 5.9: Jet quality selection efficiency for anti-k_t jets with R = 0.4 measured with a tag-and-probe technique as a function of p_T^{jet} for $|\eta| < 0.3$, for the four sets of selection criteria. Only statistical uncertainties are shown. Differences between data and MC simulations are also shown [130].

defined within the tracker acceptance region. Jets originating from the primary vertex are expected to have a value close to 1 for the JVF, whereas the value for jets from additional pp-interactions are expected to be close to 0. For this analysis, the absolute value of the JVF is required to be greater than 0.25 for jets with $p_T < 50$ GeV and and to be within $|\eta| < 2.4$ so that their charged particles are within the coverage of the inner tracking detector.

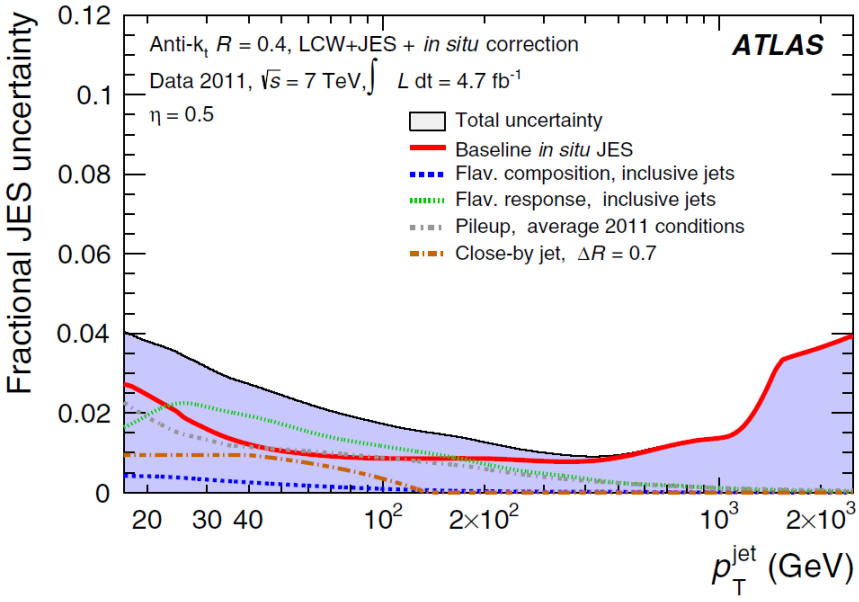
5.2.4 Performance

As described in the previous section, the jet energy scale is determined using MC simulation (baseline + GS calibration) with an additional correction derived from in-situ measurements to correct for the residual differences between data and MC simulation. Systematic uncertainties of the jet energy scale account for uncertainties on the different in-situ measurements entering in the jet calibration, uncertainties on the pile-up correction, and uncertainties due the dependence of the jet energy scale on whether a jet originates from a light quark, heavy quark, or gluon. For the 2012 data set, the different systematic uncertainties have been evaluated as a function of η and p_T of the jet. In addition, the uncertainty due to the pile-up

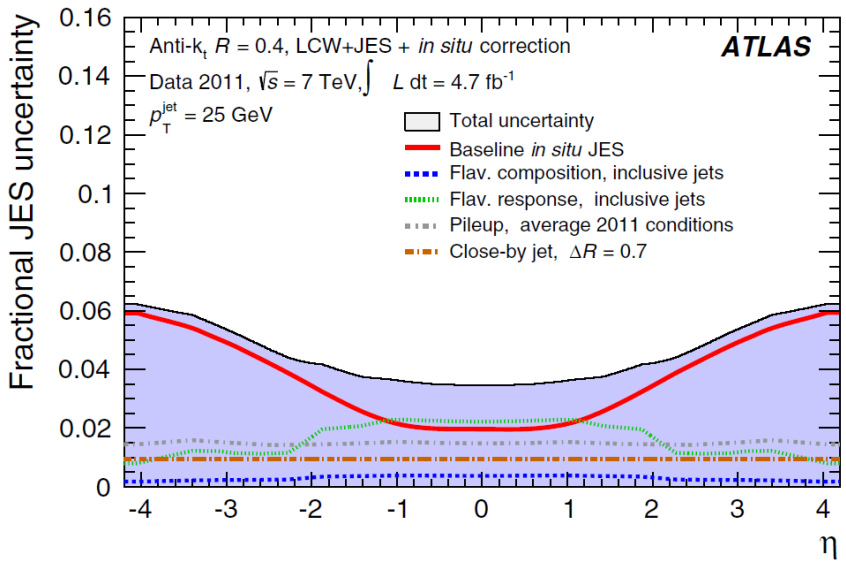
correction is determined as a function of the number of primary vertices N_{PV} and the number of additional interactions due to pile-up μ. In total, the jet energy scale uncertainty in the 2012 data set is decomposed into 65 uncertainty components (so-called nuisance parameters) which are added in quadrature to obtain the total JES uncertainty.

The performance of the jet energy scale has been evaluated for the 8 TeV data set and is very similar to that of the 7 TeV data set. As an example, Figure 5.10 shows the total fractional jet energy scale uncertainty binned in p_T^{jet} at constant $\eta = 0.5$, and binned in η at constant $p_T^{jet} = 25$ GeV in the 2011 data set for $\sqrt{s} = 7$ TeV [124]. In these figures, "Baseline in-situ JES" refers to the combined uncertainty arising from the in-situ measurements used in the calibration. The JES uncertainty as a function of p_T^{jet} for central jets is dominated by the uncertainties on the in-situ calibration, especially for high-p_T jets, and the uncertainty arising from the flavour composition of the MC samples used in the calibration. The uncertainty from the pile-up correction becomes more dominant for low p_T jets. As a function of η of the detector for 25 GeV jets, the JES uncertainty is dominated by the in-situ calibration, especially in the forward regions of the detector. The uncertainty on the jet response due to the flavour composition of the jets becomes important in the central region.

To simplify some physics analyses, the set of 65 uncertainty components have been reduced to a set of 14 components plus 2 additional uncertainties due to jet flavour. This was done by diagonalizing the total covariance matrix of the JES correction factors related to the different in-situ measurements, identifying a new set of uncertainty sources derived from the eigenvectors and eigenvalues of such a matrix, and then selecting a subset of 5 uncertainty sources with the largest eigenvalues and combining the remaining nuisance parameters into a sixth residual term. Only uncertainties dependent on a single parameter (p_T^{jet}) are combined in this way, and any η, μ, N_{PV}, or flavour dependent uncertainties are not included in the combination. The final reduced set of 16 uncertainty components used in this analysis are:



(a)



(b)

Figure 5.10: Fractional jet energy scale systematic uncertainty as a function of (a) p_T^{jet} and (b) jet pseudorapidity for anti-k$_t$ jets with distance parameter of R = 0.4 calibrated using the LCW+JES calibration scheme for $\sqrt{s} = 7$ TeV. The uncertainty shown applies to inclusive QCD jets with average 2011 pile-up conditions [124].

- 6 from the reduction of the in-situ analyses nuisance parameters
- 2 from η-intercalibration
- 1 from the behaviour of high-p_T jets derived from single hadron response
- 1 nuisance parameter from MC non-closure (imperfect closure when the MC-based calibration factors are applied to an alternate MC sample)
- 4 nuisance parameters from the pile-up offset corrections (3 of which are μ and/or N_{PV} dependent)
- 2 from the flavour composition and topology of the MC samples

The precision of the measurement of the jet energy is also an important quantity for physics measurements as well as the central value of the jet energy scale. For the majority of the jet p_T spectrum, the width of the distribution of the balance between jets and well measured photons or reconstructed Z bosons is used to measure the detector resolution [131, 132]. Additionally the balance between jets in dijet events is used to extend these measurements to higher $|\eta|$ and p_T. For very low p_T jets there is a significant contribution to the jet energy resolution from pile-up particles and electronic noise. Figure 5.11 shows the individual in-situ measurements of the resolution in the central region and the result of the combination and the associated uncertainty. The combination of these measurements results in a measured jet energy resolution with an uncertainty below 3% at 20 GeV and less than 1% above 100 GeV for anti-k_t R = 0.4 jets.

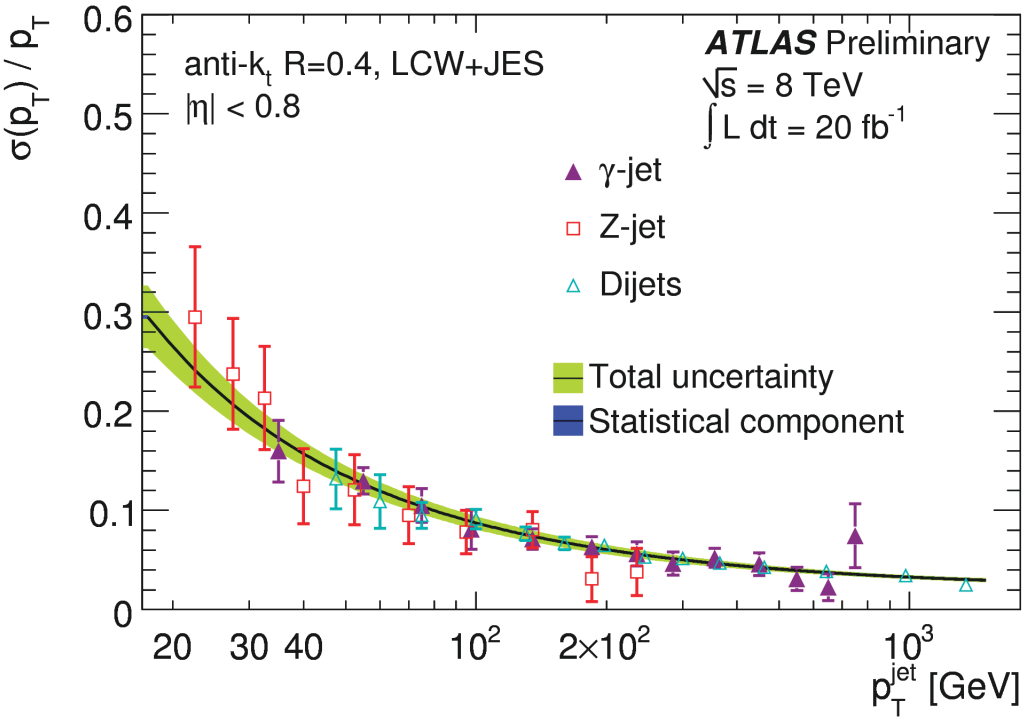


Figure 5.11: The jet resolution as a function of p_T for LCW+JES anti-k_t R = 0.4 jets in the central region $|\eta| < 0.8$. The three in-situ measurements of the resolution are shown to illustrate the compatibility between the different methods. The final fit is included with its associated statistical and total uncertainty [131].

CHAPTER 6

Data Sets and Event Selection

6.1 Data

The data used in this analysis was collected from proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 8$ TeV using the ATLAS detector in 2012. Events were selected online using an unprescaled dielectron trigger (`EF_2e12Tvh_loose1`) with a minimum transverse energy threshold of $E_T = 12$ GeV for each of the two required electrons in the region $|\eta| < 2.5$. The high-level trigger E_T threshold lies in the region affected by the first-level E_T turn-on, as denoted by "`T`" in the trigger name. The "`h`" in the trigger name indicates that an additional cut on the energy deposited in the hadronic calorimeter directly behind the central core of the EM cluster of ≤ 1 GeV has been applied at LVL1. The "`v`" component in the name stands for varied threshold, and is effectively a coarse dead material correction applied on the trigger threshold at LVL1. The full trigger chain is `L1_2EM10VH` $\rightarrow$ `L2_2e12Tvh_loose1` $\rightarrow$ `EF_2e12Tvh_loose1`. All of these triggers are unprescaled. At LVL2 and EF, electrons are required to pass the *loose* identification requirement as described in Section 5.1.2. For the dielectron trigger used in this analysis, the trigger efficiency is between 95% and 99% for electrons with $p_T > 20$ GeV. Only events which have been collected during periods of stable beam operation and with all ATLAS subdetectors fully operational are considered by applying the Good Run List (GRL). The selected data correspond to an integrated luminosity of $\int Ldt = 20.3$ fb$^{-1} \pm 2.8\%$.

6.2 Event Selection

Each selected event is required to have at least one primary vertex with a minimum of three associated tracks. Z bosons are selected by requiring the event to have exactly two oppositely charged, reconstructed electrons which fulfill the medium identification quality requirement as described in section 5.1.2. The electron energy in data is calibrated as described in Section 5.1.3. Further corrections are applied to MC predictions in order to compensate for differences between data and MC for the electron trigger, reconstruction, and identification efficiencies. Each of the selected reconstructed electrons must have a minimum transverse momentum of $p_T^e > 20$ GeV and a pseudorapidity in the region $|\eta| < 2.47$, excluding the transition regions between the barrel and endcap calorimeter sections $|\eta| = 1.37-1.52$. A veto on events with more than two electrons is used to further suppress background contributions from diboson production. The dielectron invariant mass, $m_{e^+e^-}$, is required to be within a ± 20 GeV mass window centered on the mass of the Z boson, 71 GeV $< m_{e^+e^-} <$ 111 GeV. This mass window is smaller than that used in the 2011 analysis on 7 TeV data in order to reduce the multijet background component and increase the signal purity.

Associated jets are build from locally calibrated calorimeter topo-clusters using the anti-k_t algorithm with a distance parameter of $\Delta\mathrm{R} = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$, as described in Section 5.2.1. A baseline calibration with an additional global sequential calibration is applied offline as a further correction to the energy of reconstructed jets in data, as detailed in section 5.2.2. Selected jets are required to be in the rapidity region $|y^{jet}| < 4.4$, and have a transverse momentum $p_T^{jet} > 30$ GeV. In order to suppress jets due to pile-up, jets whose $p_T^{jet} < 50$ GeV and which are within the pseudorapidity region $|\eta| < 2.4$ are required to have their value of the jet vertex fraction (JVF) be greater than 0.25, as described in Section 5.2.3. In order to account for beam-gas events, beam-halo events, cosmic ray muons and calorimeter noise, further jet cleaning procedures are applied, as detailed in section 5.2.3. Studies in reference [133] have shown that the reconstruction efficiency for jets decreases in the proximity

of electrons, which directly translates into a bias on the measured jet multiplicity. Therefore, jets closer than 0.4 in ΔR to any of the two selected electrons are removed.

6.3 Signal and Background Modelling

Fully simulated samples from MC event generators are used to model signal and background processes, which are summarized in Table 6.1. Signal samples are used for comparisons to data at detector level and particle level, as well as for determining corrections for removal of detector effects in the later unfolding procedure and studying systematic uncertainties on the measurement. Therefore a good modelling of the signal process is essential. A detailed description of the signal MC used in this analysis is provided in Section 6.3.1.

Major background contributions to $Z/\gamma^*(\rightarrow ee)$+jets production are composed of electroweak diboson production, single top, $t\bar{t}$, and multijet processes as described in Section 2.3.4. MC predictions are used to estimate the background contributions from electroweak, single top, and $t\bar{t}$ processes, which are detailed in Section 6.3.2. MC predictions for multijet processes suffer from large systematic uncertainties in the signal region due to the large multijet cross section and limited MC statistics, thus an estimate of this background contribution is made using a data-driven method, which is described in Section 6.3.3.

In order to account for the presence of pile-up, the predictions from the different MC generators are overlaid with additional soft interactions simulated by Pythia. Any remaining residual pile-up differences between data and the MC samples are corrected by reweighting the MC predictions to the average number of interactions per bunch crossing measured in data. Finally, all MC event samples are passed through the GEANT4 simulation of the ATLAS detector and trigger system, as described in Chapter 4.

6.3.1 Signal Monte Carlo

The nominal $Z/\gamma^*(\rightarrow ee)$+jets MC event sample is modelled with Alpgen [3] v2.14 using CTEQ6L1 PDFs and $\sqrt{m_Z^2 + p_{T,Z}^2}$ as both the QCD factorization and renormalization scale.

Signal Process	Data Set	Generator	PDF Set	Cross Section (nb)	ϵ_{filter}
$Z/\gamma^*(\rightarrow ee)$	147105-147110	Alpgen (+Pythia)	CTEQ6L1	1.15 ± 0.06	1.0
$Z/\gamma^*(\rightarrow ee) + cc$	110805-110808	Alpgen (+Pythia)	CTEQ6L1	0.027 ± 0.001	1.0
$Z/\gamma^*(\rightarrow ee) + bb$	110817-110820	Alpgen (+Pythia)	CTEQ6L1	0.013 ± 0.0007	1.0
$Z/\gamma^*(\rightarrow ee)$	147770	Sherpa v1.4.1	CT10	1.15 ± 0.06	1.0

Background Process	Data Set	Generator	PDF Set	Cross Section (nb)	ϵ_{filter}
WW	105985	Herwig	CTEQ6L1	$(5.49 \pm 0.55) \cdot 10^{-2}$	0.38
ZZ	105986	Herwig	CTEQ6L1	$(7.33 \pm 0.31) \cdot 10^{-3}$	0.21
WZ	105987	Herwig	CTEQ6L1	$(2.23 \pm 0.09) \cdot 10^{-2}$	0.31
$t\bar{t}$	117050	Powheg (+Pythia)	CTEQ6L1	0.21 ± 0.01	0.54
single top, t-channel	110090	Powheg (+Pythia)	CTEQ6L1	$(1.75 \pm 0.12) \cdot 10^{-2}$	1.0
single anti-top, t-channel	110091	Powheg (+Pythia)	CTEQ6L1	$(9.39 \pm 0.64) \cdot 10^{-3}$	1.0
single top, s-channel	110119	Powheg (+Pythia)	CTEQ6L1	$(1.64 \pm 0.11) \cdot 10^{-3}$	1.0
single top, Wt-channel	110140	Powheg (+Pythia)	CTEQ6L1	$(2.05 \pm 0.14) \cdot 10^{-2}$	1.0

Table 6.1: Details and settings of MC event samples used for the analysis. The cross sections quoted for the $Z/\gamma^*(\rightarrow ee)$ is the inclusive NNLO cross section. The cross sections are used to normalize the estimates of the expected number of events, and ϵ_{filter} is the event filter efficiency. The PDF set is the name of the set of parton distributions used by the generator.

Samples are generated at LO in QCD for up to five partons in the final state, and with an invariant dilepton mass above 60 GeV. Alpgen is interfaced to Pythia [13] v6.426.2 for parton shower evolution, hadronization, and underlying event simulation using the Perugia2011c tune [46]. For the parton shower interface between Alpgen and Pythia, the MLM matching cut is set to 20 GeV. These MC event samples only include heavy flavour jets due to gluon splittings in the parton shower. Therefore the production of Z bosons in association with jets originating from b-quarks and c-quarks is modelled separately using Alpgen including LO matrix elements with up to three additional partons. The overlap between the light and heavy flavour samples from Alpgen is removed. PHOTOS [104] is used by Alpgen to describe QED radiation.

A second $Z/\gamma^*(\rightarrow ee)$+jets sample is generated with Sherpa [4] v1.4.1 using CT10 PDFs. Events from Sherpa are generated using the MENLOPS approach [5, 110, 111], which calculates the inclusive $Z/\gamma^*+ \geq 0$ jet matrix element at NLO, and higher jet multiplicities with up to five partons in the final state are calculated at LO in QCD. The predictions from Sherpa include matrix elements containing b-quarks, which are assumed to be massless. Samples are generated with a dilepton invariant mass above 40 GeV. Parton shower evolution is handled internally and is matched to the matrix element using the CKKW merging scheme which preserves the NLO accuracy of the inclusive process. For the formation of hadrons, Sherpa uses a modified cluster hadronization model. The YFS [109] approach is used to describe virtual and real QED corrections.

The Alpgen+Pythia and Sherpa samples are normalized globally to the NNLO pQCD inclusive Drell-Yan cross section of 1.15 nb $\pm$ 5% [68] as determined by the FEWZ program [64–66] for $\sqrt{s} = 8$ TeV pp-collisions. The uncertainty on the NNLO cross section accounts for PDF uncertainties, uncertainties due to the choice of α_S, and renormalization and factorization scale uncertainties on the pQCD predictions. The signal MC event samples do not include contributions from vector boson fusion production of $Z/\gamma^*(\rightarrow ee)$+jets.

6.3.2 Background Monte Carlo

Background MC event samples for diboson processes (WW, WZ and ZZ) are modelled with Herwig [73] v6.520.2 using CTEQ6L1 PDFs interfaced to Jimmy [102, 103] using the AUET2 [134] underlying event tune. Diboson processes are normalized globally to NLO pQCD predictions with uncertainties of 10%, 4%, and 4.2% for WW, WZ, and ZZ processes respectively [68]. Background samples for single top (s-channel, t-channel, Wt inclusive production) and $t\bar{t}$ production are modelled with Powheg [75–77] interfaced to Pythia using CTEQ6L1 PDFs and the Perugia2011C tune [46]. Cross sections for top quark production have uncertainties of 6.8% and 5.7% for single top [135] and $t\bar{t}$ [136] processes respectively. Uncertainties on the various cross sections account for PDF uncertainties, uncertainties due to the choice of α_S, and renormalization and factorization scale uncertainties on the pQCD predictions. For the background samples described here, PHOTOS is used to describe QED radiation, and TAUOLA [137] is used to simulate tau decays. Background contributions from multijet events are estimated using a data-driven technique.

6.3.3 Data-Driven Estimate of Multijet Background

The multijet background constitutes a major background to $Z/\gamma^*(\rightarrow ee)$+jets for low jet multiplicities. This is a background where jets in the event get misidentified as electrons, which then go on to fulfill the electron and Z boson selection criteria. MC predictions of the multijet background have large uncertainties due to the large multijet cross section and limited available MC statistics. Thus, this analysis uses a data-driven estimate to provide a more reliable shape for each of the measured distributions, which results in a significantly lower uncertainty.

The idea is to select a sample of data events which model the multijet background in the signal region so that its shape can be extracted for any observable. Multijet events are selected similarly to the standard $Z/\gamma^*(\rightarrow ee)$+jets selection, with modifications to the electron identification and charge requirements that are meant to enhance the multijet component.

Events are selected online using the same dielectron trigger as the standard signal selection. Offline, the electron candidates are required to pass the *loose* identification criterion, but fail the *medium* one (see Section 5.1.2 for description of electron identification). Furthermore, selected multijet events must have exactly two electron candidates with the same electric charge. No bias is expected between events with two electrons of same-sign electric charge and events with two electrons of opposite-sign electric charge stemming from multijet production. Therefore, it is expected that the shape of multijet events passing this selection is similar to the shape of multijet events passing the standard selection. Additionally, requiring two electrons of same-sign electric charge greatly reduces contributions from signal events and other background processes. The small contamination from non-multijet processes is removed by subtracting predictions from signal and background MC with the same multijet selection applied.

The normalization of the multijet contribution is extracted from a maximum likelihood fit of two template distributions to the measured dielectron mass distribution from data for events passing the standard $Z/\gamma^*(\rightarrow ee)$+jets selection. The template which is fitted (T_{fit}) to the invariant mass distribution from data is constructed as shown in Eqation 6.1.

$$T_{fit} = \alpha \cdot T_{multijet} + (1 - \alpha) \cdot T_{EW} \tag{6.1}$$

$T_{multijet}$ is the multijet background template that comes from the selected multijet events in data as described above. T_{EW} is created by combining the MC signal sample with the non-multijet background samples with the standard signal selection applied to all. The normalization constant α of the multijet template is a free parameters of the fit to the Z boson mass peak. The fit is performed in the invariant mass range of 60 GeV $< m_{e^+e^-} <$ 150 GeV in order to take advantage of the side-bands of the Z boson mass peak. The result of the maximum likelihood fit for the inclusive jet selection is shown in figure 6.1. The number of predicted multijet events for different inclusive jet multiplicities is shown in Table 7.1 from Chapter 7

The different sources of uncertainties that were considered are: the choice of the invariant mass fit range and the binning of the mass distribution, the choice of the signal MC, and the impact of electron energy scale and electron energy resolution. In order to obtain the total systematic uncertainty on the normalization of the multijet template, the maximum difference in each category from the nominal normalization are taken and added in quadrature. The results of the systematic uncertainty study on the multijet background estimate are described in Section 9.5.

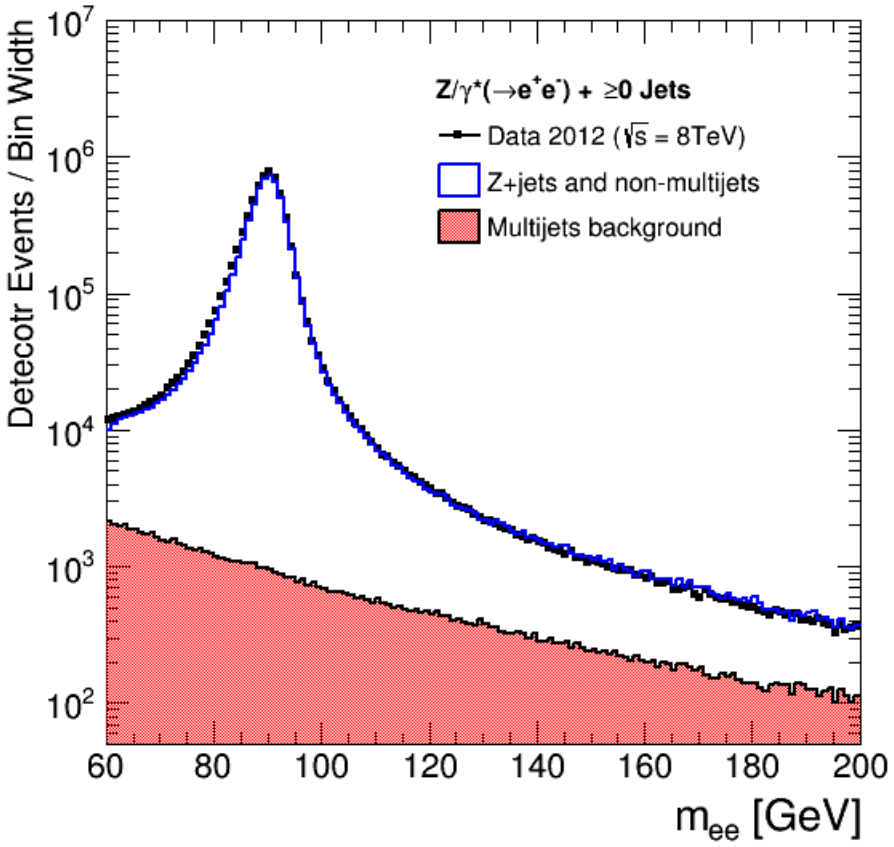


Figure 6.1: Result of the maximum likelihood fit to the dielectron mass distribution in an extended mass range using two templates: the selected multijet events in data and the combined MC signal events and non-multijet background events. The normalization of both templates are free parameters of the fit.

CHAPTER 7

Detector Level Measurements

Total inclusive and exclusive cross sections are measured as a function of the inclusive and exclusive jet multiplicities (N_{jet}). Differential cross sections are measured as a function of the jet transverse momentum (p_T^{jet}), the jet rapidity (y^{jet}), and the transverse momentum of the Z boson (p_T^Z). Furthermore, differential cross sections are measured for a set of observables in events with at least two jets that are sensitive to the topology of the event. These are the ratio of the p_T of the sub-leading jet to the p_T of the leading jet (Rp_T^{jj}), the ratio of the p_T of the Z boson to the p_T of the leading jet (Rp_T^{Zj}), the absolute azimuthal separation between the two leading jets ($\Delta\phi^{jj}$), and the minimum $|\Delta\phi|$ between the Z boson and either of the two leading jets in the event ($min\Delta\phi^{Zjj}$).

The set of observables which are sensitive to the topology of the event, and are meant to test the modelling of MC in high-p_T regions of phase space, are measured for events with high-p_T final states in addition to the standard event selection. High-p_T final states are defined as events with either $p_T^Z > 300$ GeV or $p_T^{leading\ jet} > 300$ GeV. The p_T threshold was chosen to be as high as possible, while at the same time the bin widths were selected such that the statistical uncertainty in each bin was no more than 10% after applying the high-p_T event selection.

In this chapter, measurements are first presented as detector level distributions and compared to predictions from MC event generators that have been passed through the detector simulation. Detector level distributions correspond to electrons and jets reconstructed in the detector. However, in order to provide results which are independent of the detector setup,

measured distributions are corrected for detector effects in Chapter 8 using the iterative Bayesian method. The different sources of systematic uncertainties on the measurement are discussed in Chapter 9. The unfolded results are presented as cross sections at particle level in Chapter 10.

To illustrate the different steps of the measurements, results are presented in this chapter, as well as in Chapters 8 – 10, for a few key observables: the exclusive jet multiplicity (N_{jet}), the transverse momentum (p_T) and rapidity (y^{jet}) of the leading jet, the transverse momentum of the Z boson in events with at least one jet (p_T^Z), and the p_T ratio between the Z boson and the leading jet in events with at least two jets (Rp_T^{Zj}).

7.1 Detector Level Distributions

The uncorrected distributions measured in data for $Z/\gamma^*(\rightarrow ee)$+jets production are compared to MC predictions of signal and background processes. For signal predictions, two different ME+PS generators have been used as described in Section 6.3.1: Alpgen interfaced to Pythia for parton shower, and Sherpa. Electron energy scale, resolution, and efficiency corrections are applied to MC predictions as described in Section 5.1. Jet energy calibrations are applied on both data and MC, as described in Section 5.2. Table 7.1 summarizes the events in data and the expectations from MC for both Alpgen+Pythia and Sherpa as signal MC. The uncertainties shown for data and MC predictions are the statistical uncertainties. The uncertainties on the multijet background include both the statistical and the systematic uncertainty on the normalization of the multijet template fit described in Section 9.5.

The expected background fraction for $Z/\gamma^*(\rightarrow ee)$+jets increases with increasing jet multiplicity from 0.8% for $N_{jet} \geq 0$, to 23% for $N_{jet} \geq 5$, and up to 45% for $N_{jet} \geq 7$. For low jet multiplicities, $N_{jet} \leq 1$, the main contributions come from QCD multijet, $t\bar{t}$ and diboson events. For higher jet multiplicities, contributions from $t\bar{t}$ become dominant. The multijet background is estimated using a data-driven method as described in Section 6.3.3. The fraction of multijet events in data is estimated to vary between $(0.46 \pm 0.15)\%$ for $N_{jet} \geq 0$

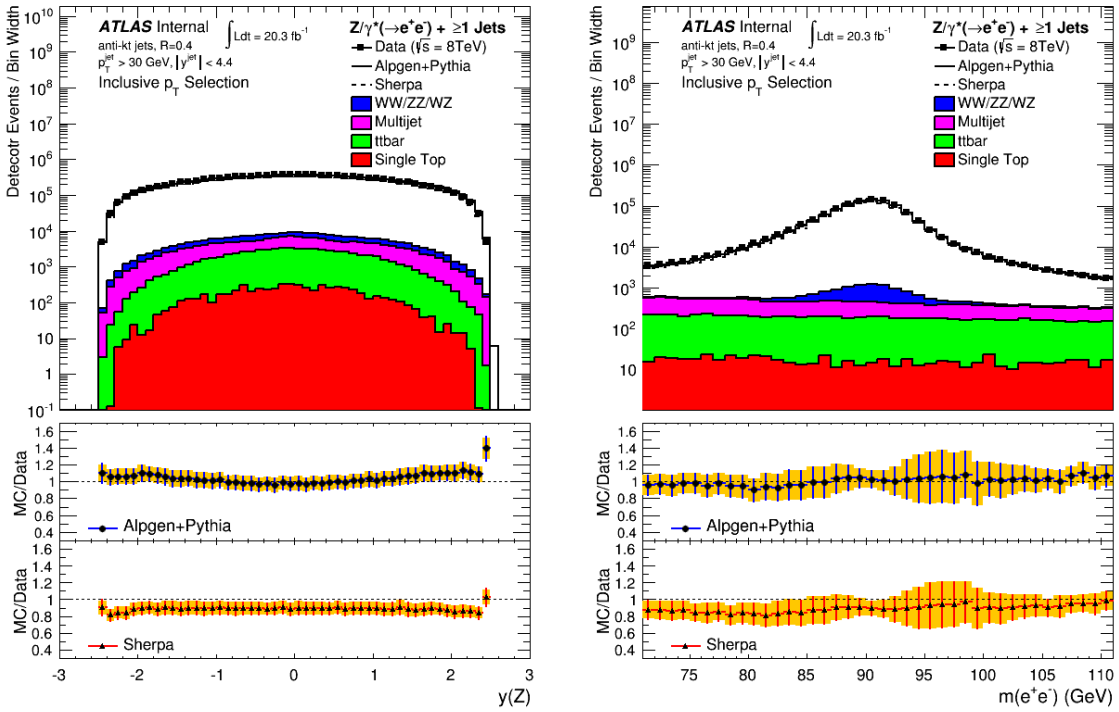
	≥ 0 jets	≥ 1 jet	≥ 2 jets	≥ 3 jets	≥ 4 jets
$Z/\gamma^*(\rightarrow ee)$ (Alpgen)	6043879 ± 4239	1159777 ± 1052	278548 ± 454	57829 ± 181	12259 ± 71
$Z/\gamma^*(\rightarrow ee)$ (Sherpa)	5743884 ± 2860	1013743 ± 899	226940 ± 364	50063 ± 162	11308 ± 76
WW, WZ, ZZ	9497 ± 43	6485 ± 35	2953 ± 24	709 ± 12	137 ± 5
Single top	723 ± 20	655 ± 19	309 ± 13	112 ± 8	31 ± 4
$t\bar{t}$	6966 ± 39	6784 ± 39	5404 ± 35	2664 ± 24	1002 ± 15
Multijets	29525 ± 9455	8486 ± 2713	2391 ± 766	605 ± 192	142 ± 45
Total predicted (Alpgen)	6090590 ± 10362	1182187 ± 2910	289605 ± 892	61919 ± 265	13571 ± 86
Total predicted (Sherpa)	5790595 ± 9878	1036153 ± 2858	237997 ± 849	54143 ± 253	12620 ± 90
Data observed	6454794 ± 2541	1164977 ± 1079	279272 ± 528	62892 ± 251	13994 ± 118

	≥ 5 jets	≥ 6 jets	≥ 7 jets	≥ 8 jets	
$Z/\gamma^*(\rightarrow ee)$ (Alpgen)	2588 ± 29	552 ± 12	116 ± 5	23 ± 2	
$Z/\gamma^*(\rightarrow ee)$ (Sherpa)	2607 ± 36	604 ± 17	131 ± 8	30 ± 4	
WW, WZ, ZZ	22 ± 2	3 ± 1	0.4 ± 0.3	< 0.4	
Single top	9 ± 2	1 ± 0.8	< 1	< 1	
$t\bar{t}$	328 ± 9	91 ± 4	24 ± 2	7 ± 1	
Multijets	31 ± 10	7 ± 2	2 ± 0.5	0.4 ± 0.1	
Total predicted (Alpgen)	2978 ± 32	654 ± 13	147 ± 5	31.8 ± 2.2	
Total predicted (Sherpa)	2997 ± 39	706 ± 18	162 ± 8	38.8 ± 4.1	
Data observed	3150 ± 56	618 ± 25	109 ± 10	18 ± 4	

Table 7.1: Number of predicted events from MC simulation and observed in data for the different jet multiplicities. Only the statistical uncertainty is shown. The predictions for the multijet background have been estimated from data, whose uncertainty includes both the estimated statistical and systematic uncertainties.

and $(1.5 \pm 0.45)\%$ for $N_{jet} \geq 7$. The other backgrounds, such as single top, $t\bar{t}$, and diboson are estimated using MC and normalized to the integrated luminosity using their respective cross sections. Contributions from single top processes remain below the percent level for all jet multiplicities. The background contributions from $t\bar{t}$ processes increase steadily with increasing jet multiplicity, from 0.1% for $N_{jet} \geq 0$ up to 21.9% for $N_{jet} \geq 7$.

Figure 7.1 shows the detector level distributions of the rapidity of the Z boson (y^Z) and the dielectron invariant mass (m_{ee}) of Z boson candidates in events with at least one associated jet in the final state. The invariant dielectron mass range shown corresponds to the invariant mass window used in the standard signal selection, 71 GeV $< m_{ee} <$ 111 GeV. The shape of these distributions are well modelled by both Alpgen+Pythia and Sherpa signal samples.



(a) Z rapidity

(b) Dielectron mass

Figure 7.1: (a) rapidity and (b) dielectron mass for Z boson candidates with at least one jet in the event in data and simulation. The multijet background has been estimated from data. The yellow band in the ratio plots represent the total systematic uncertainty. The error bars represent the total systematic and statistical uncertainty.

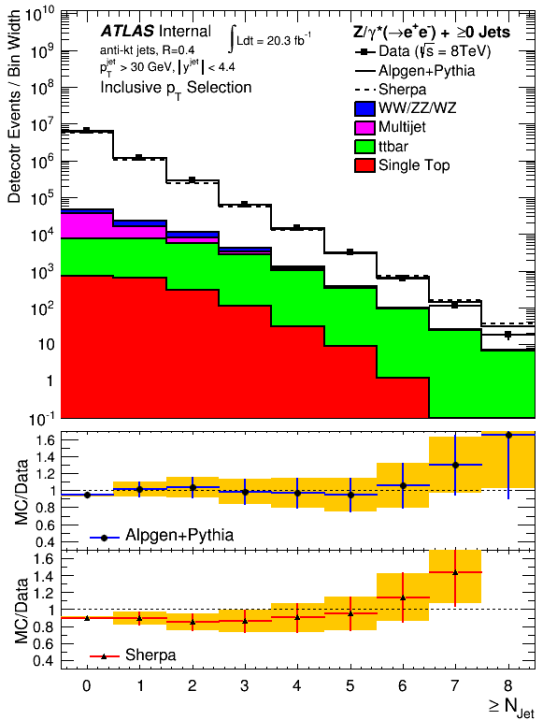


Figure 7.2: Inclusive jet multiplicity. The multijet background has been estimated from data. The yellow band in the ratio plots represent the total systematic uncertainty. The error bars represent the total systematic and statistical uncertainty.

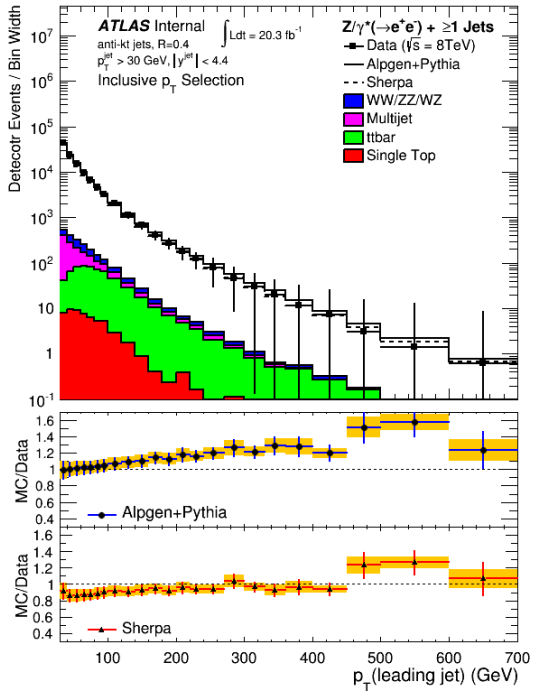
The detector level distribution of the inclusive jet multiplicity is shown in Figure 7.2. The available statistics in data are of sufficient accuracy for access up to the 8^{th} jet bin. For both Alpgen+Pythia and Sherpa, data and MC are consistent up to $N_{jet} \geq 6$. For larger jet multiplicities, $N_{jet} \geq 7$, MC predictions seem to overestimate the number of associated jets. However, care must be taken in this region since both Alpgen+Pythia and Sherpa calculate matrix elements only up to five partons in the final state. Higher jet multiplicities are modelled with parton showers, which are already known to not fully describe the jet kinematics in this region [10].

Figure 7.3 shows the detector level distributions p_T^{jet} and y^{jet} for the leading jet in events with at least one jet in the final state, as well as the p_T of the Z boson and the ratio of the p_T of the Z boson to the p_T of the leading jet Rp_T^{Zj} in events with at least one jet.

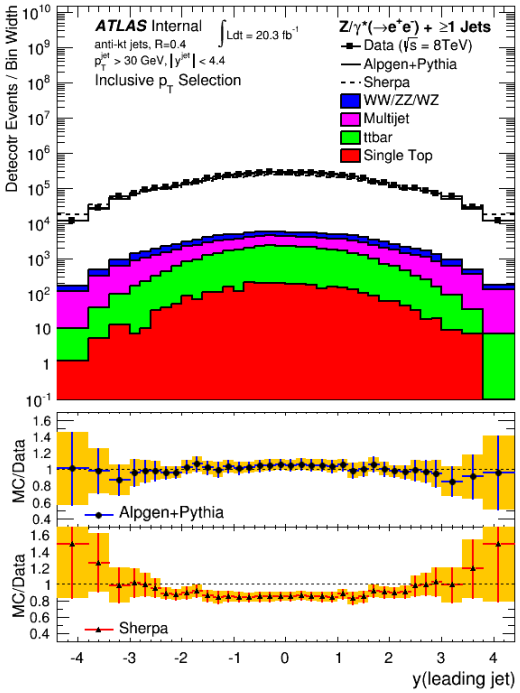
The predictions from Alpgen+Pythia model a too hard p_T distribution for the leading jet and for the Z boson, while they are consistent with the data for the other distributions. Sherpa models a too broad rapidity distribution for jets, which is seen as an overestimation

of the number of forward jets. For the Rp_T^{Zj} distribution, results from both Alpgen+Pythia and Sherpa are consistent with data when using the standard signal selection.

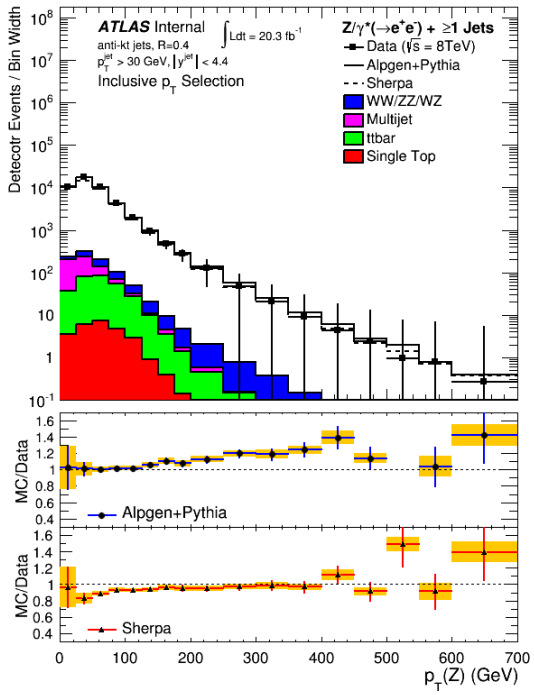
The full set of detector level distributions can be found in Appendix A. All plots illustrated in this section and in Appendix A are shown with the combined statistical and systematic uncertainty as described later in Chapter 9. Since these distributions have not been corrected for detector effects, there are no systematic uncertainties introduced by the unfolding process yet. They provide a valuable reference for a comparison between generator predictions and data. In general, the predictions from Alpgen+Pythia and Sherpa are consistent with the data. This gives confidence that the signal MC event samples provide a reasonable description of the event kinematics and the detector response, which is necessary since they are used as inputs in the unfolding procedure. A discussion of all distributions is given in Chapter 10 together with comparisons at particle level.



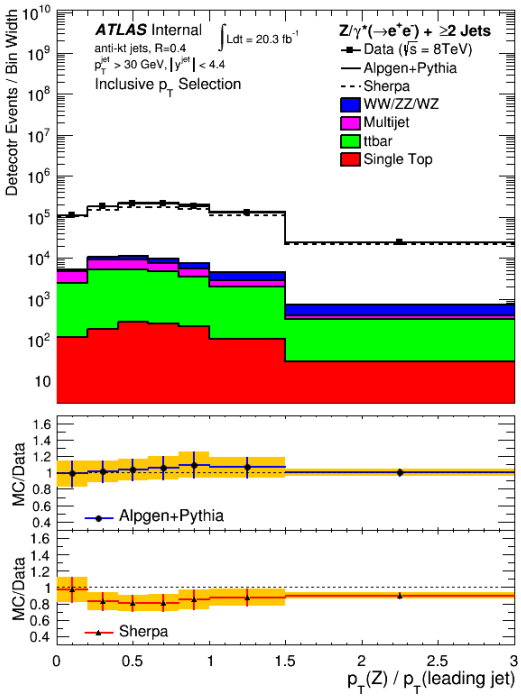
(a) Leading jet p_T



(b) Leading jet rapidity



(c) Z boson p_T



(d) Rp_T^{Zj}

Figure 7.3: Leading jet p_T (a) and leading jet rapidity (b) for events with at least one jet in the final state, (c) Z boson p_T for events with at least one jet in the final state, and (d) ratio of the p_T of the Z boson to the p_T of the leading jet in events with at least two jets in the final state in data and simulation. The multijet background has been estimated from data. The yellow band in the ratio plots represent the total systematic uncertainty. The error bars represent the total systematic and statistical uncertainty.

CHAPTER 8

Unfolding

Measured distributions in particle physics differ from those of the corresponding "true" physical quantities due to smearing effects, such as those caused by the tracking resolution, energy resolution, acceptance, non-linearities of the detector, and imperfect reconstruction of the measured objects. In order to facilitate comparisons with perturbative QCD predictions that are not passed through a simulation of the ATLAS detector, as well as results from other experiments, measured distributions are unfolded from these distortions to extract the true, or particle level, distributions.

When dealing with only one variable, the simplest method for unfolding the true distribution from the measured distribution is the bin-by-bin correction. A generalized efficiency is calculated using MC simulation by taking the ratio between the number of events falling in a certain bin of the true variable and the number of events in the same bin of the reconstructed variable. This efficiency is then used to estimate the number of true events from the number of observed events in that bin of the measured distribution. Unfortunately, this method cannot take into account large migrations of events from one bin to others, and neglects correlation between adjacent bins. An alternative option sometimes used to solve the problem of migration is to build a migration matrix using a MC simulation which connects the number of true events generated in one bin with the number of reconstructed events observed in the other bins. This matrix is then inverted and applied to the measured distribution. However, from a probabilistic point of view, there is no reason to expect the inverse matrix to exist. Additionally, this method cannot handle large statistically fluctuations, as this would result in negative terms of the inverted matrix.

In this analysis, the unfolding of detector effects is performed using the iterative Bayes method [138]. This method is a multidimensional unfolding procedure based on Bayes theorem of statistics, which takes into account both bin migration and correlation, and avoids the problem of matrix inversion. Section 8.1 describes the Bayes unfolding method in more detail, while Section 8.2 covers the application of this unfolding method to the measured distributions. All plots related to the performance of the unfolding procedure can be found in Appendix B. The uncertainties related to the unfolding procedure are covered in Chapter 9. The unfolded distributions, which correspond to the true or particle level distributions, are presented in Chapter 10 together with comparisons of MC predictions at particle level.

8.1 Bayes Unfolding Method

Measured distributions are unfolded to the particle level using the iterative Bayes method [138] implemented in the RooUnfold software package [139]. It can be applied to multidimensional unfolding problems, does not require matrix inversion, and can take into account any kind of smearing and bin migration.

The method can be explained in terms of several independent causes ($C_i,\ i = 1, 2, \ldots, n_C$), where each cause can produce one of several effects ($E_j,\ j = 1, 2, \ldots, n_E$). These individual causes and effects can be thought of as the bins of a histogram of a variable such as the jet p_T. The causes represent the true distribution before smearing, and the effects represent the measured distribution after smearing, migration, reconstruction efficiency, and fakes. Each cause can produce any one of the different effects. However, for each given effect, as in the case of a measurement, the exact cause is unknown.

The goal of the unfolding process is to determine the true number of events in each cause bin, $\hat{n}(C_i)$. In order to do this, one needs to estimate the conditional probabilities $P(C_i|E_j)$ that an observation of a single event E_j was due to the i^{th} cause C_i. Using Bayes theorem, this can be expressed as

$$P(C_i|E_j) = \frac{P(E_j|C_i)P_0(C_i)}{\sum_{l=1}^{n_C} P(E_j|C_l)P_0(C_l)}. \tag{8.1}$$

The conditional probabilities expressed by Equation 8.1 depend on the probability for an effect to be produced from a defined cause $P(E_j|C_i)$, and the initial probability of the causes $P_0(C_i)$. The conditional probabilities $P(E_j|C_i)$ carry information about the migration and smearing, and must be calculated or estimated using MC simulation. Then, if one observes $n(E_j)$ events with effect E_j, the expected number of events assignable to each of the causes is

$$\hat{n}(C_i) = \frac{1}{\epsilon_i} \sum_{j=1}^{n_E} n(E_j) P(C_i|E_j), \quad \epsilon_i \neq 0. \tag{8.2}$$

Equation 8.2 takes into account the efficiency ϵ_i of detecting the cause C_i in any one of the possible effect bins, where $0 \leq \epsilon_i \equiv \sum_{j=1}^{n_E} P(E_j|C_i) \leq 1$. This is required as there is no need for each cause to produce at least one of the effects taken under consideration. The expected number of events in the different cause bins can then be used to calculate the total true number of events $\hat{N}_{true}$ and the individual probabilities of the causes $\hat{P}(C_i)$:

$$\hat{N}_{true} = \sum_{i=1}^{n_C} \hat{n}(C_i) \tag{8.3}$$

$$\hat{P}(C_i) \equiv P(C_i|n(E)) = \frac{\hat{n}(C_i)}{\hat{N}_{true}}. \tag{8.4}$$

As mentioned already, the conditional probabilities $P(C_i|E_j)$ expressed in Equation 8.1 also depend on the initial probability of the causes $P_0(C_i)$. This suggests that the unfolding method based on Bayes theorem should be performed iteratively using the result as a prior for each consecutive iteration. Starting with the first iteration, the initial distribution $P_0(C)$, and hence the initial expected number of events $n_0(C)$ is chosen from the best knowledge of the process under study. The starting point is typically taken as the truth distribution from MC simulation, but in the case of complete ignorance, satisfactory results are obtained starting from a uniform distribution [138]. For each subsequent iteration, $\hat{n}(C)$ and $\hat{P}(C)$ are calculated. A χ^2 comparison can be made between $n_0(C)$ and $\hat{n}(C)$, and if the χ^2 value is "small enough" the procedure can be stopped. If not, then $P_0(C_i)$ is replaced by $\hat{P}(C)$, and $n_0(C)$ is replaced by $\hat{n}(C)$, and the the procedure is repeated.

8.2 Unfolding of Detector Effects

The correction for detector effects is done using the iterative Bayes method using the Alpgen+Pythia signal MC event sample. The correction is performed in one dimension for each distribution separately. Background events, estimated from various MC simulations or data-driven methods as described in Sections 6.3.2 and 6.3.3, are subtracted from the measured detector level distribution prior to the unfolding procedure. A multidimensional unfolding method performed with respect to the leading and sub-leading jet p_T would take into account the complete event kinematics more fully, but is not performed here. Multidimensional unfolding requires a very large number of simulated MC events. The number of bins in the migration matrix increases exponentially with the number of unfolding dimensions, resulting in fewer events per bin. In order to keep the statistical uncertainty under control, a much larger sample of simulated signal events is required than what is available.

The migration matrix is created from MC simulation using only matched events, i.e. those events which fulfill the selection criteria at both the particle level and detector level. At the detector level, the selection criteria are described in Section 6.2. At the particle level, the event must contain exactly two electrons with additional jets that fulfill the phase space requirements as detailed in Table 8.1. Electrons are "dressed" with QED radiation by summing their 4-momenta with the 4-momenta of all photons found within a cone radius of 0.1 around the electron. This defines the fiducial phase space of the unfolded measurements. In addition, every particle level jet needed for the variable under study must be matched to its corresponding jet at detector level within a cone of radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \leq 0.4$. Events which do not fulfill the requirements at particle level are added to the distribution of not-matched reconstructed events. Events which do not fulfill all the requirements at detector level are counted as an inefficiency.

In matched events, the jet hierarchy at particle level, defined as the ordering of jets in terms of decreasing p_T^{jet}, is preserved at detector level, i.e. jets in matched events in MC are ordered

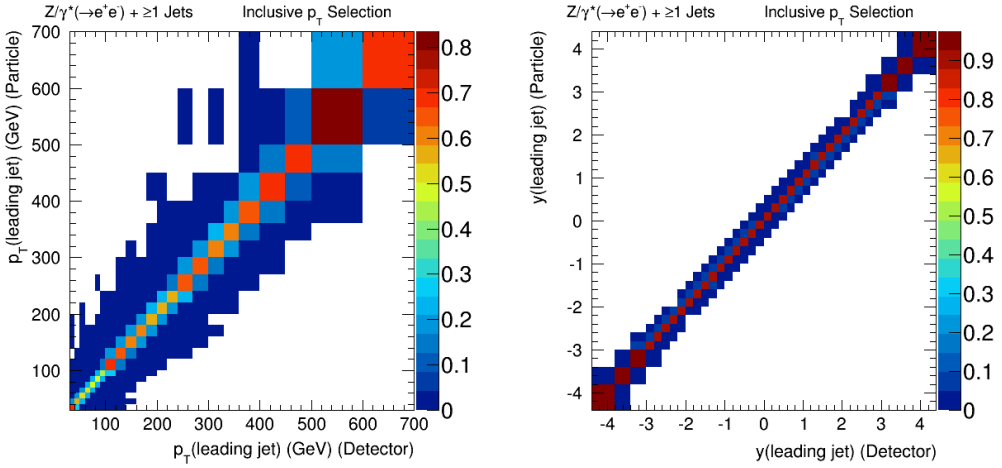
Variable	Criteria
electron p_T	$p_T > 20$ GeV
electron $\|\eta\|$	$\|\eta\| < 1.37$ or $1.52 < \|\eta\| < 2.47$
electron charge	opposite
dielectron invariant mass m^{ee}	71 GeV $\leq m^{ee} \leq$ 111 GeV
jet p_T	$p_T^{jet} > 30$ GeV
jet rapidity y^{jet}	$y^{jet} < 4.4$
electron-jet separation ΔR^{ej}	$\Delta R^{ej} > 0.4$

Table 8.1: Summary of the complete particle level phase space selection requirements.

according to their particle level p_T. Detector resolution effects or energy mismeasurement can alter the p_T ordering of jets when going from the particle level to the detector level. This can result in a different jet hierarchy for the matched distribution compared to the reconstructed distribution. Therefore, prior to the unfolding of detector effects, a "jet order" correction must be applied to the measured background-subtracted distribution from data in order to correct for the misordering of jets at the detector level. This correction is based on bin-by-bin factors calculated using MC simulations.

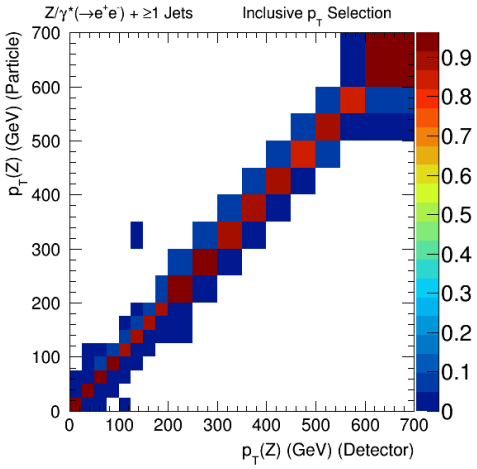
The detector level distributions in matched events do not include events which are not matched to a event at particle level. Thus, a second bin-by-bin correction for not-matched events is also applied to the measured background-subtracted distribution from data prior to the unfolding. This correction is determined from the detector level distribution of not-matched events from the same MC event sample used to build the migration matrix.

Figure 8.1 shows the migration matrices determined from matched events for the exclusive jet multiplicity (N_{jet}), the transverse momentum p_T^{jet} and rapidity y^{jet} of the leading jet, the transverse momentum of the Z boson in events with at least one jet (p_T^Z), and the p_T ratio between the Z boson and the leading jet in events with at least two jets (Rp_T^{Zj}). The migration matrices have been normalized for presentation such that each bin represents the probability that an event in a certain bin at particle level will end up in a particular bin at detector level.

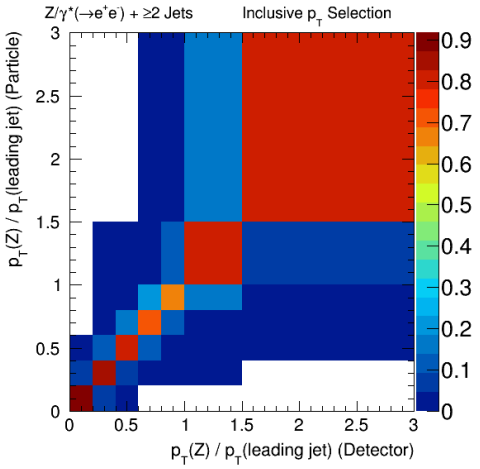


(a) Leading jet p_T

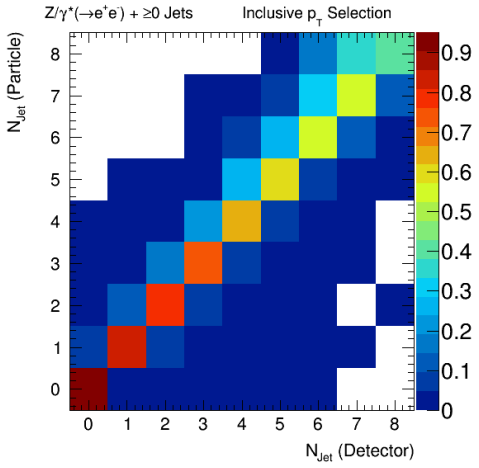
(b) Leading jet rapidity y^{jet}



(c) Z boson p_T



(d) Rp_T^{Zj}



(e) Exclusive N_{jet}

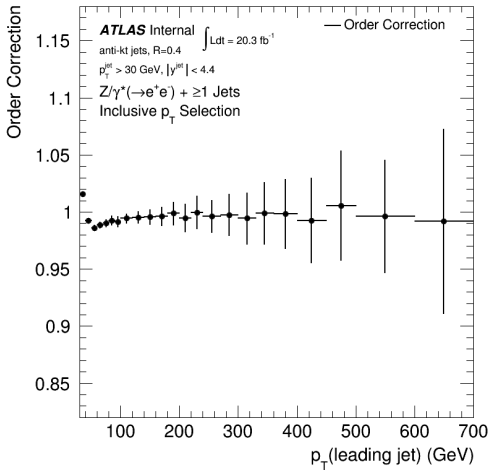
Figure 8.1: Migration matrices for (a) the p_T of the leading jet, (b) the rapidity of the leading jet, (c) the p_T of the Z boson in events with at least one jet (p_T^Z), (d) the p_T ratio between the Z boson and the leading jet in events with at least two jets (Rp_T^{Zj}), (e) and the exclusive jet multiplicity (N_{jet}).

Some significant migration is seen for the transverse momentum of the leading jet, more so in the low p_T region. This can be caused by threshold effects due to the 30 GeV cut on the jet p_T or the presence of pile-up jets in this region, which can lead to an incorrect match between particle and detector level jets. Conversely, the migration matrix for the jet rapidity of the leading jet is approximately diagonal, as would be expected from jets that are spatially matched in $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.4$. Very little migration is seen for the p_T of the Z boson, as well as for the p_T ratio between the Z boson and the leading jet.

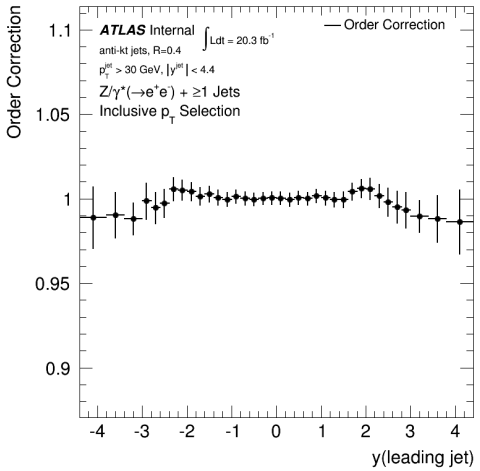
For the exclusive jet multiplicity, an increasing amount of migration between bins is seen for larger jet multiplicities. In order to preserve the overall momentum balance of the event, the average jet p_T in events with a large number of jets is typically smaller than that in events with only a few jets. Since low p_T jets are most sensitive to the p_T threshold cut, it is not surprising that in high multiplicity events, the number of jets at particle level does not match the number of jets at detector level.

Figures 8.2 and 8.3 show the jet order correction and the correction for not-matched events for the same example variables as used above. The order correction for the p_T^Z and N_{Jet} distributions act as a small test of the validity of the algorithm. These variables do not depend on the jet hierarchy, therefore, the order correction should be zero, as is shown by these distributions. The order correction for the remaining variables is rather small, on the order of 1% over the whole phase space region. Only for very low leading jet p_T does the order correction increase. This is due to the steeply falling p_T distribution, which makes it more likely for low p_T jets to change their ordering due to resolution effects. This effect manifests itself in the largest values of Rp_T^{Zj}, where the order correction reaches a few percent, and in the p_T ratio between the second leading jet and the leading jet, Rp_T^{jj}, as the ratio approaches unity (shown in Appendix B).

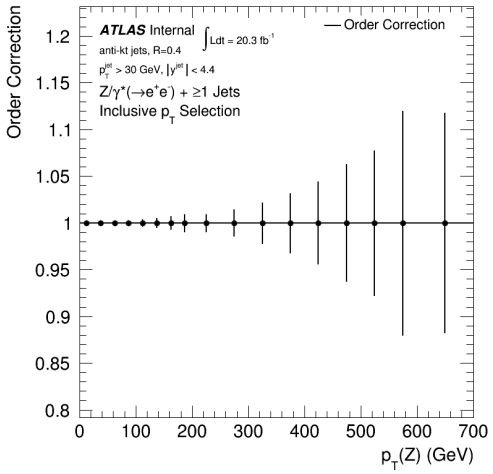
For some phase space regions, the correction for not-matched events can be quite large. This is especially true for jets in the very forward regions of the detector. These regions



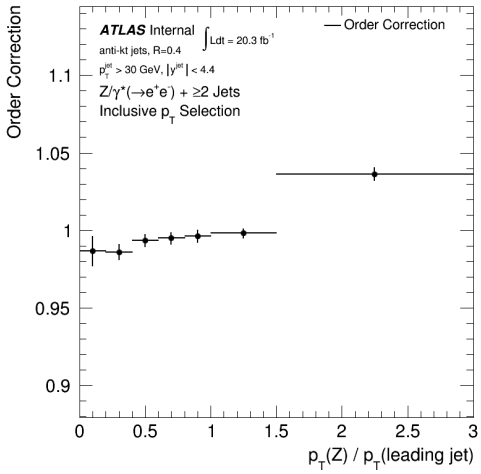
(a) Leading jet p_T



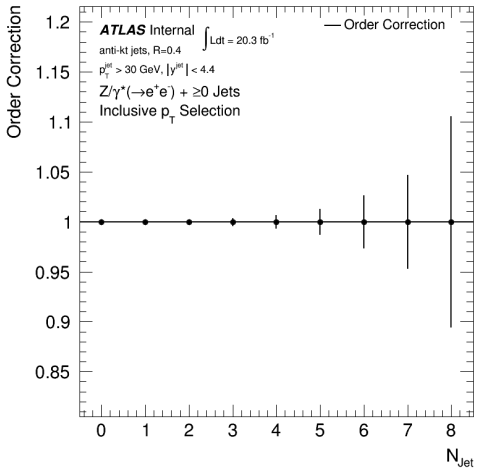
(b) Leading jet rapidity y^{jet}



(c) Z boson p_T

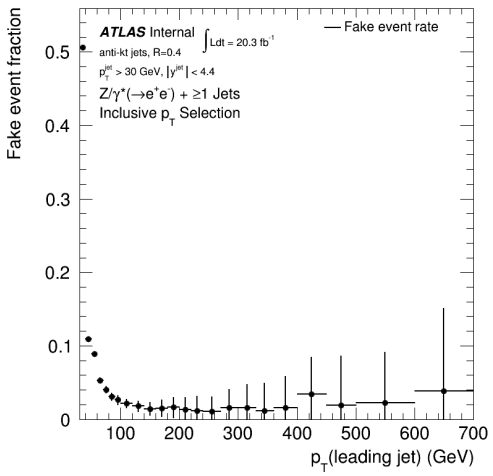


(d) Rp_T^{Zj}

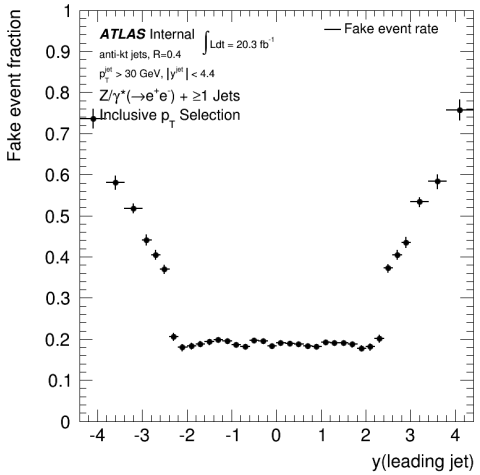


(e) Exclusive N_{jet}

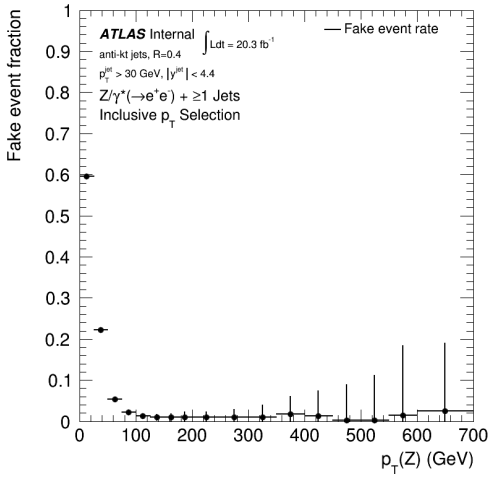
Figure 8.2: Order correction for (a) the p_T of the leading jet, (b) the rapidity of the leading jet, (c) the p_T of the Z boson in events with at least one jet (p_T^Z), (d) the p_T ratio between the Z boson and the leading jet in events with at least two jets (Rp_T^{Zj}), (e) and the exclusive jet multiplicity (N_{jet}).



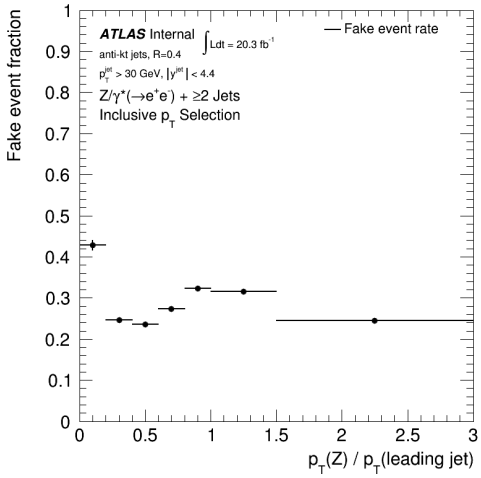
(a) Leading jet p_T



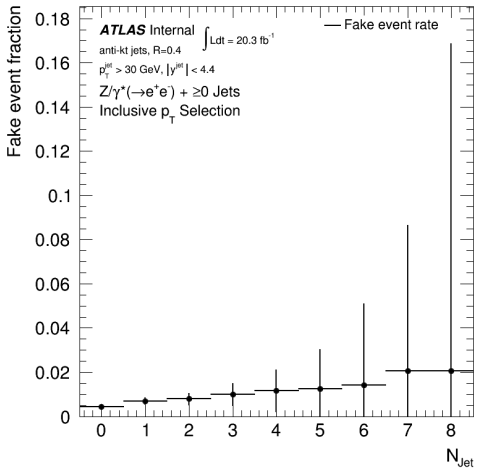
(b) Leading jet rapidity y^{jet}



(c) Z boson p_T



(d) Rp_T^{Zj}



(e) Exclusive N_{jet}

Figure 8.3: Correction factors for non-matched, or "fake", events for (a) the p_T of the leading jet, (b) the rapidity of the leading jet, (c) the p_T of the Z boson in events with at least one jet (p_T^Z), (d) the p_T ratio between the Z boson and the leading jet in events with at least two jets (Rp_T^{Zj}), (e) and the exclusive jet multiplicity (N_{jet}).

are not covered by the tracking detector, thus no tracking information is available for the identification and removal of pile-up jets. Additionally, the modelling of jets in the very forward region of the detector by MC simulation is difficult. As was seen in the detector level measurement of the leading jet rapidity in Figure 7.3(b), the predictions from Alpgen+Pythia and Sherpa diverge in the forward region, which leads to large systematic uncertainties on the final measurement. In the future, the analysis of jets in the forward region needs to be scrutinized in more detail.

As a consistency check and to ensure that the unfolding algorithms are working properly, a closure test can be performed. The reconstruction level distribution from the very same MC event sample that was used to build the migration matrix can be unfolded and compared to its truth distribution. Any discrepancy between the unfolded and truth distributions would indicate a problem with the unfolding procedure. For every distribution considered, the resulting ratio is unity.

Since the migration matrix and corrections are estimated using MC simulation, uncertainties have to account for the different modelling of jet kinematics and composition, such as underlying event tunes, hadronization modelling, and the heavy flavour fraction within the jets, by different MC simulations. Thus, the migration and corrections are estimated twice, once with the Alpgen+Pythia signal sample as the nominal correction, and once with the Sherpa signal MC event sample as a systematic variation. Furthermore, due to the limited size of the MC event sample used to build the migration matrix, an additional statistical uncertainty is introduced by the unfolding method. This uncertainty is estimated using MC toys, where the bin contents of the migration matrix are varied within their statistical uncertainty according to a Gaussian distribution before unfolding the measurements from data. This procedure is repeated 1000 times and the RMS of the modified results is taken as an estimate of the statistical uncertainty. Finally, since the unfolding method is iterative, the final distribution will depend on the number of unfolding iterations. Statistical fluctuations are magnified by the unfolding procedure. After an optimal number of iterations, the results

will begin to diverge. The optimal number of iterations for each variable was not determined in this analysis. The analysis of the 7 TeV data set [10] did not state the exact number of unfolding iterations for each specific variable, but did find that the optimal number of iterations was typically two or three. Therefore, in this analysis, two iterations was chosen as the nominal case for unfolding, with the results obtained using one and three iterations acting as systematic variations. Details of the uncertainty analysis related to the unfolding are covered in Section 9.6.

CHAPTER 9

Systematic Uncertainties

Several different sources of systematic uncertainty have been studied in detail for each observable presented in this analysis. Systematic shifts, corresponding to $\pm 1\sigma$ variations of quantities affecting the electron and jet selections, are applied to the predictions from MC event samples. The migration matrices and correction factors needed in the unfolding procedure are built from the modified MC predictions, which are then used to unfold the background-subtracted data distributions, as described in Section 8.2. The differences between the nominal unfolded result and the modified unfolded results from the up and down systematic variations are taken as the systematic uncertainties. Uncertainties related to the normalization of a given background prediction are only applied to that specific background contribution. The modified background distributions are subtracted from the data distributions before unfolding. Uncertainties directly related to the unfolding method are also considered. The following sections describe the different sources of systematic uncertainties that were taken into account in this analysis.

To mitigate artificial fluctuations due to a lack of statistics in the MC event samples, systematic uncertainties are symmetrized. The symmetrization is performed by taking the average of the final result of the up and down variations of the different uncertainty components on the unfolded distribution. Finally, the different sources of systematic uncertainties are added in quadrature.

9.1 Electron Selection

Electron trigger, identification and reconstruction efficiencies have been measured in data and MC event samples using tag and probe methods. Differences between the data and the MC predictions are corrected by applying scale factors on the MC event samples as described in Section 5.1.3. The individual systematic uncertainties on these scale factors are propagated into the uncertainty on the cross section by increasing or decreasing the scale factors by their associated uncertainty.

Additional uncertainties come from the electron energy scale and resolution calibration as detailed in Section 5.1.3. The systematic uncertainty on the electron energy scale receives contributions from seven components. The uncertainty due to each component is propagated into the uncertainty on the cross section in a correlated manner by increasing or decreasing the corresponding electron energy correction simultaneously for both electrons within $\pm 1\sigma$.

The electron energy resolution in MC is not reproduced by data. In order to account for this difference, the electron energies in MC predictions are smeared according to a Gaussian distribution so that the dielectron mass distribution in $Z \to ee$ events matches data. The corresponding uncertainty is taken into account by smearing the energy of each electron within the uncertainty of the resolution.

The different sources of uncertainty are symmetrized and added in quadrature, resulting in an uncertainty of about 2%, roughly independent of jet p_T, jet rapidity, and jet multiplicity. The dominant contributions come from the electron trigger and identification uncertainties. The uncertainty on the electron energy scale is minor, but grows larger with increasing values of the p_T of the Z boson as well as with N_{jet}.

9.2 Jet Energy Scale

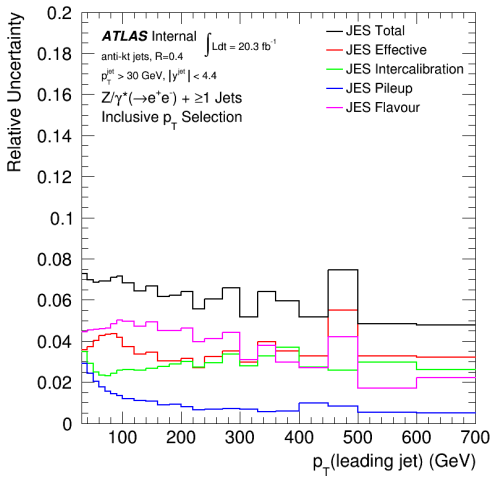
In this analysis, the set of 65 uncertainty components have been reduced to a set of 14 components plus 2 additional uncertainties due to jet flavour and topology, as described in

Section 5.2.4. The different components are propagated separately for each observable to the final cross section. The total JES uncertainty is then determined by symmetrizing each component and adding them in quadrature.

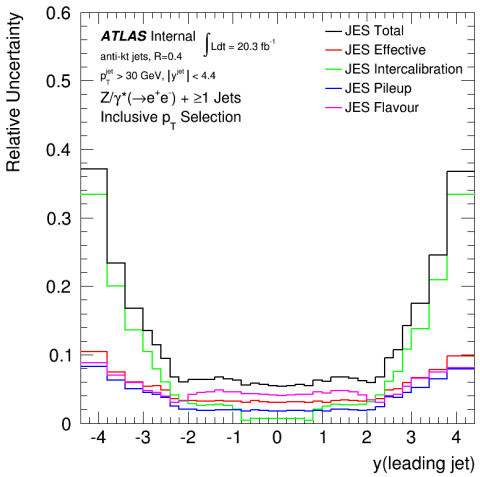
Figure 9.1 shows the different contributions to the total JES uncertainty for the transverse momentum of the leading jet, the rapidity of the leading jet, the transverse momentum of the Z boson in events with at least one jet (p_T^Z), the transverse momentum ratio between the Z boson and the leading jet in events with at least two jets (Rp_T^{Zj}), and the exclusive jet multiplicity (N_{jet}). The uncertainty components related to the different in-situ measurement used in the JES calibration (see Section 5.2.2) are combined and displayed as "JES Effective". The uncertainty on the JES translates into an uncertainty on the final cross section measurement of 6.6% for $N_{jet} \geq 1$, to 15.3% for $N_{jet} \geq 4$, and up to 25.9% for $N_{jet} \geq 7$. For large values of the jet rapidity, the total JES uncertainty reaches up to 35%. The dominating contribution to the JES uncertainty for forward jets comes from the uncertainty on the η-intercalibration, while for central jets the uncertainty on the flavour composition is the largest. The total JES uncertainty as a function of the leading jet p_T decreases from approximately 7.5% for low-p_T jets to 5.5% for high-p_T jets. The dominating contributions to the JES uncertainty for low-p_T jets come from the uncertainty from the different in-situ measurements and the jet flavour, while for high-p_T jets, the uncertainties on the flavour, in-situ measurements, and η-intercalibration are roughly equal, as seen in Figure 9.1(a). The uncertainty due to the pile-up energy correction is the least dominant. The total JES uncertainty constitutes the dominant contribution to the total systematic uncertainty.

9.3 Jet Energy Resolution

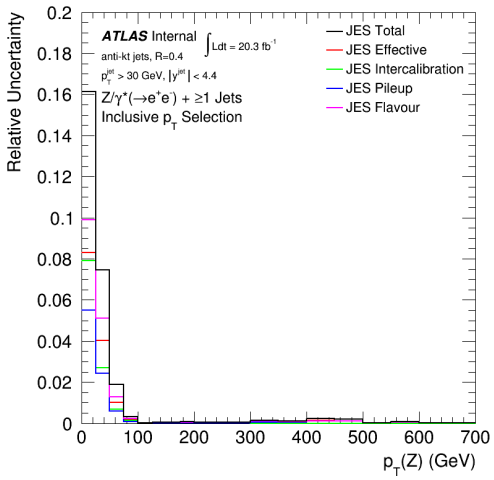
The uncertainty on the jet energy resolution (JER) described in Section 5.2.2 is propagated to the final cross section by over-smearing the jet energy in the MC event samples. The energy is smeared using a random number sampled from a Gaussian distribution centered at one, with a width given by $\sigma = \sqrt{(\sigma_{data} + \Delta\sigma_{data})^2 - \sigma_{data}^2}$, where σ_{data} and $\Delta\sigma_{data}$ are



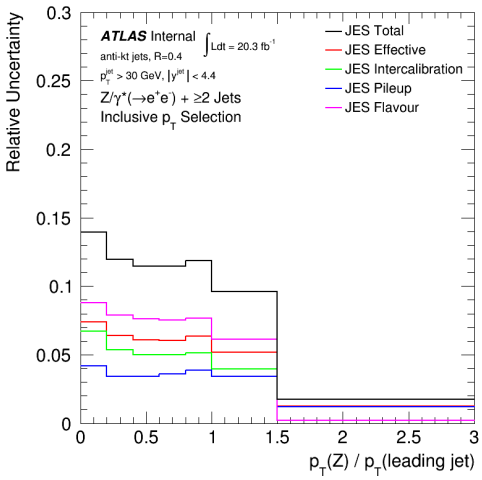
(a) Leading jet p_T



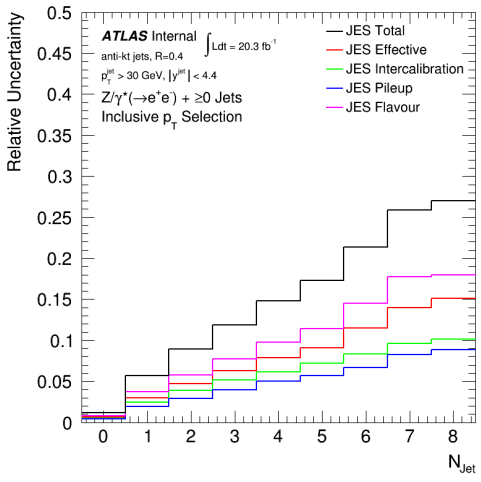
(b) Leading jet rapidity y^{jet}



(c) Z boson p_T



(d) Rp_T^{Zj}



(e) Exclusive N_{jet}

Figure 9.1: Jet energy scale uncertainty for (a) the p_T of the leading jet, (b) the rapidity of the leading jet, (c) the p_T of the Z boson in events with at least one jet, (d) the p_T ratio between the Z boson and the leading jet in events with at least two jets (Rp_T^{Zj}), (e) and the exclusive jet multiplicity (N_{jet}).

the measured jet energy resolution and its uncertainty, respectively. Given the nature of the stochastic process, the procedure is repeated 1000 times by modifying the random seed. For each trial, a modified migration matrix and correction factors are created for unfolding the background-subtracted data distributions. The difference between the nominal unfolded result and the modified unfolded results are averaged among the individual trials. The final result is symmetrized and taken as the systematic uncertainty.

This translates into an uncertainty of 4.1% for $N_{jet} \geq 1$ to 10.2% for $N_{jet} \geq 8$. For large values of the leading jet rapidity, the uncertainty on the energy resolution reaches 20%, and in the central region averages about 4%. The JER uncertainty for jets with $p_T < 100$ GeV reaches a few percent, but falls to roughly 1% for jets with $p_T > 100$ GeV.

9.4 Pile-up Jet Removal

The jet vertex fraction (JVF), as described in Section 5.2.2, is not well modelled in MC. As a consequence, its use for the identification and removal of pile-up jets constitutes another source of systematic uncertainty on the measurement. The systematic uncertainty associated with the cut on the JVF is estimated using the MC by varying the cut up and down by an amount such that the relative fraction of jets which pass the varied cut differs from the nominal fraction by the same difference which is seen between data and MC. The JVF cut, as well as the up/down systematic variations, are applied to jets with $p_T < 50$ GeV and $|\eta| < 2.4$. The difference between the modified unfolded result and the nominal result is symmetrized and taken as the systematic uncertainty. The resulting uncertainty is less than 1% over the entire phase space considered, indicating that pile-up has a negligible influence on the final results.

9.5 Background Modelling and Normalization

The uncertainties on the background predictions are related to normalization and statistical uncertainties. These two types of uncertainties are treated differently. The normalization

uncertainty is assumed to be correlated between the bins, whereas the statistical uncertainty is assumed to be uncorrelated between bins.

The normalization uncertainty is propagated to the final cross section by shifting the different background contributions up and down by an amount equal to the uncertainty on their respective normalization. The dominant contribution comes from the uncertainty on the normalization of the data-driven estimate of the multijet background for $N_{jet} \leq 4$, and on the uncertainty on the $t\bar{t}$ background for $N_{jet} \geq 5$. The diboson and single top backgrounds are negligible.

MC event samples of background processes are normalized to the given luminosity using their respective cross sections listed in Table 6.1. The normalization uncertainty on the various background contributions are propagated to the final cross section by shifting the generator cross section up and down by an amount equal to the cross section uncertainty given in Section 6.1. These uncertainties account for PDF uncertainties, uncertainties due to the choice of α_s, and scale uncertainties on the pQCD predictions [68,135,136]. The uncertainties due to the diboson and single top backgrounds remain below 0.1%. The uncertainty due to the $t\bar{t}$ background contribution is less than 0.5% for $N_{jet} \leq 4$, and rises to about 2% only for the highest jet multiplicity.

For the overall uncertainty on the normalization of the multijet background described in Section 6.3.3, three categories of systematic uncertainty were considered: the choice of the parameters used in the fit to the dielectron mass distribution, the modelling of the $Z/\gamma^*(\rightarrow ee)$+jets signal, and the impact of the electron energy scale and energy resolution. The uncertainty due to the choice of the fit parameters was estimated by varying the bin width and the fit range of the dielectron mass distribution. The uncertainty due to the signal modelling was estimated by repeating the multijet background estimate using Sherpa to model the signal instead of Alpgen+Pythia. Finally, the normalization uncertainty due to the electron energy scale and energy resolution uncertainty was estimated in the same manner

as the overall electron energy scale and resolution uncertainty described in Section 9.1. In each category, the maximum up and down deviations are taken and symmetrized. In order to obtain the total systematic uncertainty on the multijet background, the uncertainties from each category are added in quadrature. This translates into an uncertainty of about 0.5% roughly independent of the jet multiplicity.

Since the fractions of these background contributions are rather small, the resulting uncertainty on the final cross section is less than 1%. Only for the highest multiplicity events does the uncertainty due to the $t\bar{t}$ background rise to 2%.

The statistical uncertainties from the background predictions are added in quadrature to the statistical uncertainty in data. They are propagated simultaneously to the final cross section.

9.6 Unfolding

For the systematic uncertainty on the unfolding procedure described in Chapter 8, three uncorrelated sources of uncertainty were considered.

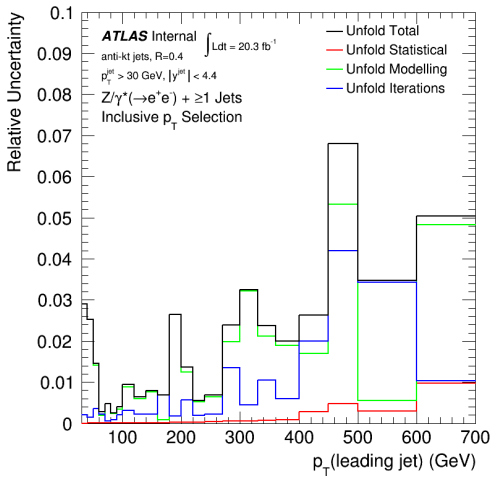
First, to account for the potential mismodelling of the jet kinematics and composition, the unfolding procedure is repeated using the Sherpa MC signal sample. The unfolded results are compared to the nominal unfolded results using the Alpgen+Pythia MC signal sample and the difference is taken as a systematic uncertainty.

An additional comparison is made between the nominal result with two unfolding iterations and the results obtained with one and three iterations. This attempts to account for the fact that the optimal number of unfolding iterations was not determined. Even though each additional iteration is meant to improve the estimation of the unfolding matrix, it also amplifies statistical fluctuations. An optimum number of iterations likely exists where the difference between the results of the previous iteration is a minimum. However, as is typical with this method, good enough results are usually obtained after only one iteration. Previous

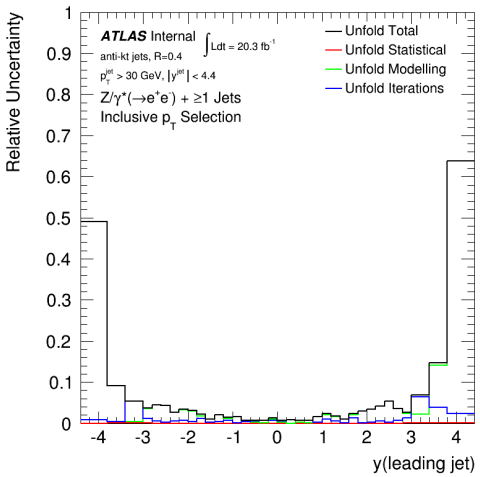
experience with the application of this unfolding method to the measurement of differential cross sections in Z/γ^*+jets events in the 2011 data set [10] found that the optimal number of iterations was two or three. Since the optimal number of iterations was not determined in this analysis, and considering that one iteration is usually sufficient, two iterations was chosen as the nominal since it was the minimum number of iterations which was found to be optimal in the 2011 analysis. A conservative estimate of the uncertainty introduced by this choice is made by taking the maximum difference between the nominal result with two iterations and the results obtained with one and three unfolding iterations.

The final component of the systematic uncertainty accounts for the limited statistics of the MC event samples. This uncertainty is propagated to the particle level distribution using a toy MC. Each bin of the migration matrix is fluctuated independently according to a Gaussian distribution with σ equal to the statistical uncertainty of the bin content. The total number of events is left unchanged, while the correction factors are recalculated. This is repeated 1000 times, and the RMS value of the difference between the nominal unfolded result and the modified results is taken as a systematic uncertainty.

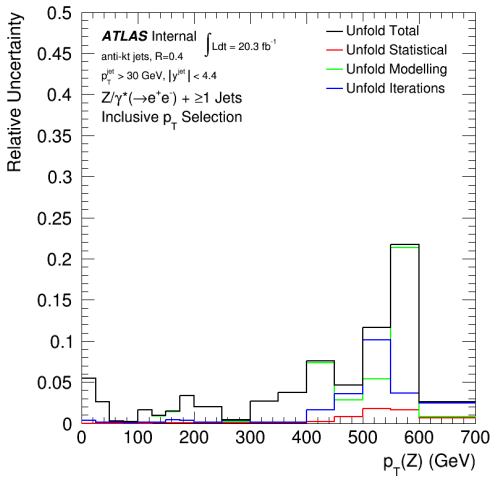
The three different sources of systematic uncertainty are symmetrized and added in quadrature, resulting in an uncertainty of 4.1% to 9.3% on the inclusive jet multiplicity for $N_{jet} \geq 1$ to $N_{jet} \geq 6$. The different contributions to the total unfolding systematic uncertainty are shown in Figure 9.2 for a few example variables. The dominant contribution comes from the choice of the MC event sample. This is especially evident for very large values of the jet rapidity. The predictions between Sherpa and Alpgen+Pythia diverge significantly in the very forward regions of the detector, which results in a very large unfolding uncertainty for high jet rapidity. The unfolding uncertainty on the leading jet p_T reaches a few percent. For jets with $p_T < 100$ GeV, the choice of the MC signal sample is again the most dominant uncertainty as Sherpa and Alpgen+Pythia disagree slightly in this region. The next dominant uncertainty comes from the number of unfolding iterations, which grows larger for jets and Z bosons with transverse momentum in the 400 - 600 GeV range, as well as for events with



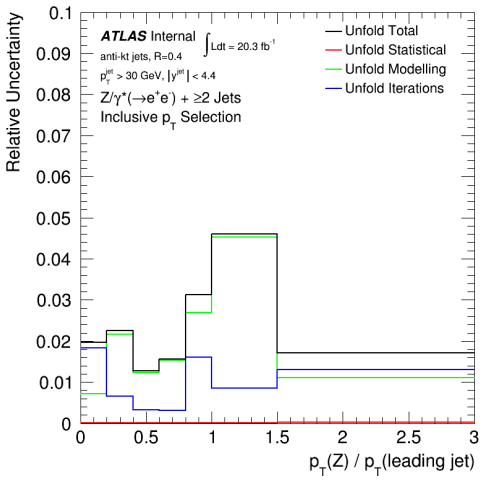
(a) Leading jet p_T



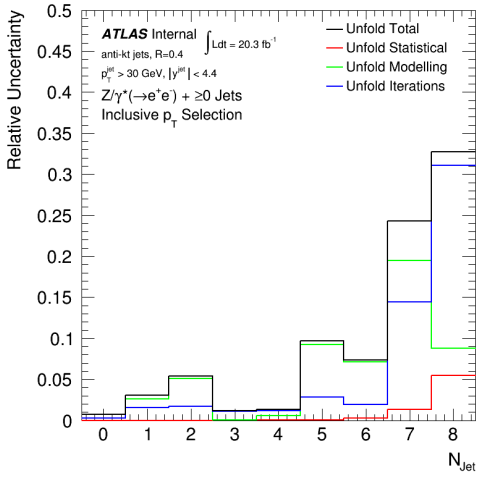
(b) Leading jet rapidity y^{jet}



(c) Z boson p_T



(d) Rp_T^{Zj}



(e) Exclusive N_{jet}

Figure 9.2: Unfolding uncertainty for (a) the p_T of the leading jet, (b) the rapidity of the leading jet, (c) the p_T of the Z boson in events with at least one jet, (d) the p_T ratio between the Z boson and the leading jet in events with at least two jets (Rp_T^{Zj}), (e) and the exclusive jet multiplicity (N_{jet}).

$N_{jet} \geq 7$. The statistical uncertainty introduced by the unfolding procedure is less than 1% for most of the phase space under study.

The first two sources of systematic uncertainty are highly influenced by statistical fluctuations in some phase space regions, which results in artificial peaks in the uncertainty distributions. For future analyses, a smoothing algorithm using a Gaussian kernel could be applied to the difference between the modified results and the nominal unfolded results in order to reduce these fluctuations.

9.7 Total Systematic Uncertainty

Table 9.1 summarizes the different components of the total systematic uncertainty on the measured cross section for several inclusive jet multiplicities. The dominant systematic uncertainty on the inclusive $Z/\gamma^*(\rightarrow ee)$ selection comes from the electron identification and trigger efficiency, while the measurement of the production cross section of $Z/\gamma^*(\rightarrow ee)$ with associated jets is dominated by the JES and JER uncertainties.

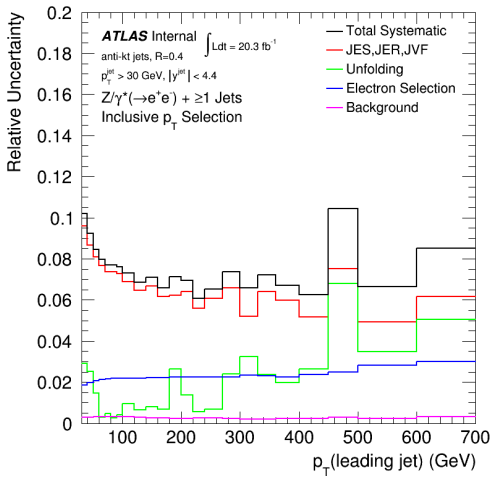
Figure 9.3 shows the components of the total systematic uncertainty for the particle level cross section as a function of the transverse momentum of the leading jet, the rapidity of the leading jet, the transverse momentum of the Z boson in events with at least one jet (p_T^Z), the ratio between the Z boson and the leading jet in events with at least two jets (Rp_T^{Zj}), and the exclusive jet multiplicity (N_{jet}). The dominant contribution to the total systematic uncertainty comes from the JES, except for a few regions, such as large values of the jet rapidity, high-p_T jets, and high jet multiplicity events, where the uncertainty introduced by the unfolding procedure is comparable. The uncertainty on the distribution of the p_T of the Z boson with at least one jet is dominated by the electron selection for $p_T^Z > 100$ GeV, with the unfolding uncertainty becoming more important at high-p_T. Similarly, the total systematic uncertainty on the p_T ratio between the Z boson and the leading jet is completely dominated by the JES in the region below 1.5, while the uncertainty on the unfolding and

Section	Source	≥ 0 jets	≥ 1 jet	≥ 2 jets	≥ 3 jets	≥ 4 jets
9.1	Electron Trigger	0.90%	1.05%	1.07%	1.09%	1.11%
	Elec. Energy Scale	0.26%	0.37%	0.39%	0.39%	0.44%
	Elec. Reconstruction	0.41%	0.53%	0.57%	0.61%	0.65%
	Elec. Energy Resolution	0.01%	0.02%	0.02%	0.04%	0.06%
	Electron ID	1.12%	1.51%	1.57%	1.62%	1.66%
9.2	Jet Energy Scale	-	6.6%	9.7%	12.7%	15.5%
9.3	Jet Energy Resolution	-	4.1%	4.7%	6.4%	7.5%
9.4	Jet Vertex Fraction	-	0.26%	0.33%	0.43%	0.52%
9.5	Multijet Background	0.21%	0.34%	0.41%	0.47%	0.51%
	$t\bar{t}$ Background	0.006%	0.037%	0.11%	0.25%	0.42%
	Single Top Background	0.0008%	0.004%	0.008%	0.012%	0.016%
	Diboson Background	0.012%	0.045%	0.083%	0.094%	0.087%
9.6	Unfolding MC	0.16%	2.9%	3.9%	0.58%	2.4%
	Unfolding Iterations	0.0001%	1.4%	1.1%	1.2%	1.3%
	Unfolding Statistical	0.001%	0.003%	0.004%	0.01%	0.02%
Total Systematic Uncertainty		1.6%	8.7%	11.7%	14.4%	17.6%
Total $\otimes$ 5%(theo.) $\otimes$ 2.8%(lumi.)		5.9%	10.4%	13.0%	15.5%	18.5%

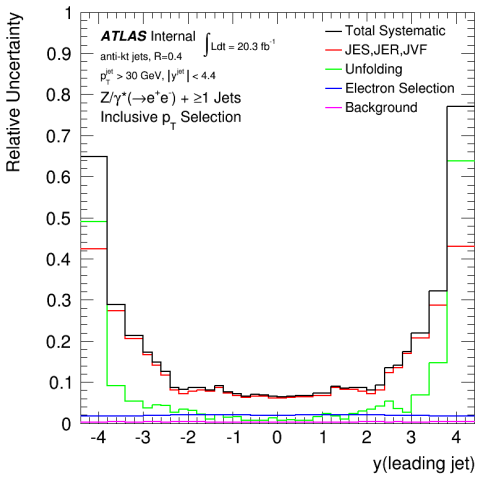
Table 9.1: Summary of the different components of the total systematic uncertainty on the measured cross section for several inclusive jet multiplicities. Also shown is the total systematic uncertainty added in quadrature with a flat 5% theoretical uncertainty on the inclusive NNLO Z/γ^*+jets cross section and a flat 2.8% uncertainty on the luminosity.

electron selection is dominant in the region above 1.5. Further distributions for the study of systematic uncertainties can be found in Appendix C.

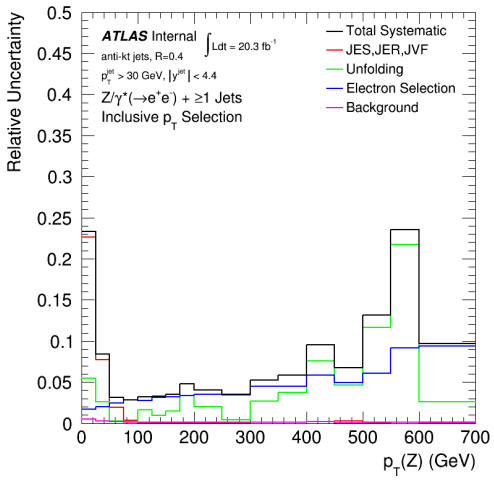
The $Z/\gamma^*(\rightarrow ee)$+jets signal MC is normalized to the inclusive NNLO cross section as determined by the FEWZ program [64–66], which has an uncertainty of 5% that accounts for PDF uncertainties, uncertainties due to the choice of α_s, and scale uncertainties on the pQCD predictions [68]. Therefore, a flat $\pm$5% theoretical uncertainty is added in quadrature to the total systematic uncertainty. Finally, the uncertainty on the total integrated luminosity of the 8 TeV data set is added in quadrature. This uncertainty was derived following the same methodology as for the 7 TeV data set, which used beam separation scans [83]. For the 8 TeV data set, this results in a flat uncertainty of $\pm$2.8%. The total systematic uncertainty is then added in quadrature to the statistical uncertainty of the measurement.



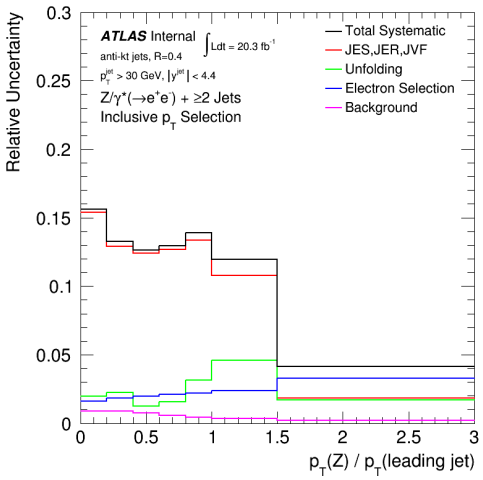
(a) Leading jet p_T



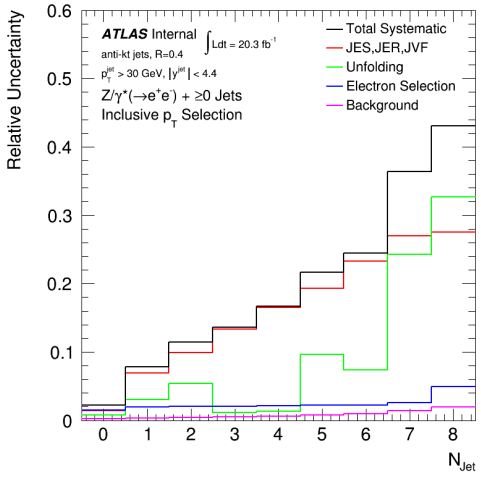
(b) Leading jet rapidity y^{jet}



(c) Z boson p_T



(d) Rp_T^{Zj}



(e) Exclusive N_{jet}

Figure 9.3: Total systematic uncertainty for (a) the p_T of the leading jet, (b) the rapidity of the leading jet, (c) the p_T of the Z boson in events with at least one jet, (d) the p_T ratio between the Z boson and the leading jet in events with at least two jets (Rp_T^{Zj}), (e) and the exclusive jet multiplicity (N_{jet}).

CHAPTER 10

Results

This section presents the final results of the measurement using 20.3 fb^{-1} of data from pp collisions at $\sqrt{s} = 8$ TeV. Measured distributions, including those presented in Chapter 7 and Appendix A, are unfolded to particle level and compared to predictions from Alpgen+Pythia and Sherpa as described in Section 6.3. Alpgen+Pythia calculates matrix elements for up to five partons in the final state at lowest level in QCD. Sherpa calculates the matrix element for the inclusive process at NLO in QCD, while higher jet multiplicities with up to five partons in the final state are calculated at LO. For both Alpgen and Sherpa, jet multiplicities greater than five are modelled by parton showers. The systematic uncertainties on the measurements are detailed in Chapter 9.

Total inclusive and exclusive cross sections are measured as a function of the inclusive and exclusive jet multiplicities (N_{jet}). Differential cross sections are measured as a function of the jet transverse momentum (p_T^{jet}), the jet rapidity (y^{jet}), and the transverse momentum of the Z boson (p_T^Z). Furthermore, differential cross sections are measured for a set of observables in events with at least two jets that are sensitive to the topology of the event. These are the ratio of the p_T of the sub-leading jet to the p_T of the leading jet (Rp_T^{jj}), the ratio of the p_T of the Z boson to the p_T of the leading jet (Rp_T^{Zj}), the absolute azimuthal separation between the two leading jets ($\Delta\phi^{jj}$), and the minimum $\Delta\phi$ between the Z boson and either of the two leading jets in the event ($min\Delta\phi^{Zjj}$).

The unfolded measurements refer to distributions with exactly two dressed electrons and additional jets which fulfill the phase space requirements as detailed in Table 8.1 from Chapter 8. The differential cross sections are defined in the fiducial region as a function of a variable ξ

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\xi} = \frac{1}{\mathcal{L}}\frac{1}{\Delta\xi}(N_{data} - N_{bkg}) \times U(\xi), \tag{10.1}$$

where N_{data} and N_{bkg} are the number of observed events in data and the number of expected background events in bin ξ respectively, $\Delta\xi$ is the bin width, and $U(\xi)$ represents the corresponding unfolding factor for bin ξ. Finally, $\mathcal{L}$ is the total integrated luminosity. In order to reduce the uncertainties on the differential cross sections related to the flat uncertainties of the luminosity estimate and the normalization of the MC signal samples, the measured and predicted cross sections are divided by the total inclusive $Z/\gamma^*(\rightarrow ee)$ + jets cross section as measured in data and predicted by the MC generators respectively.

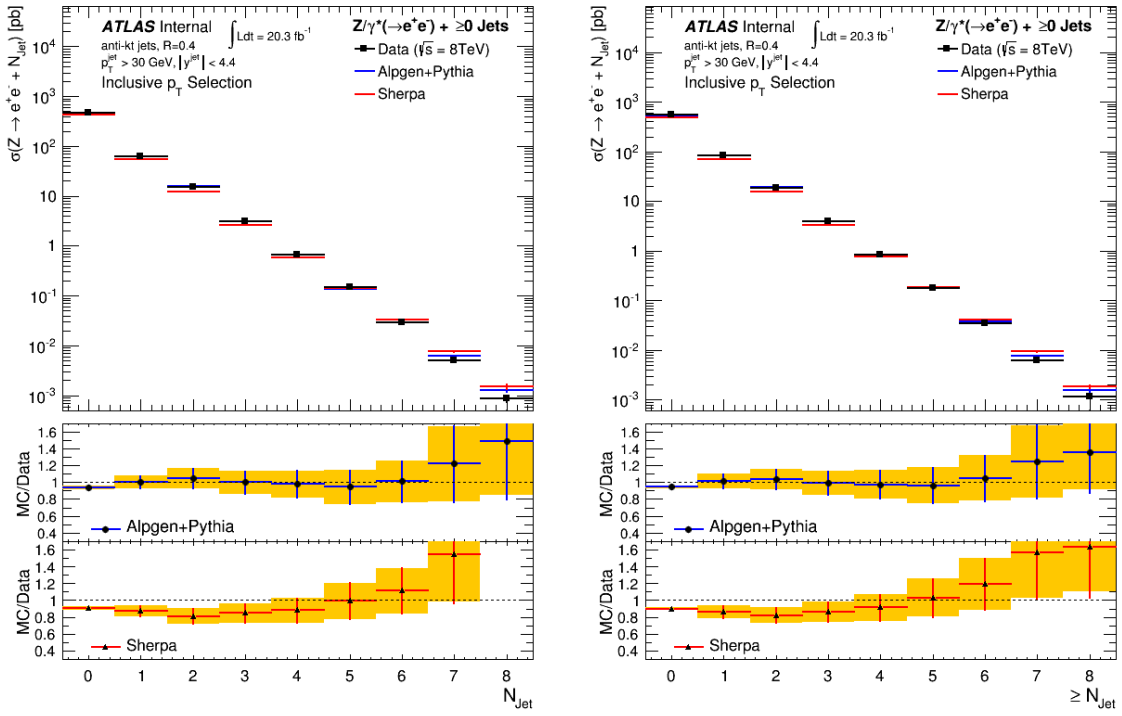
The figures in this section are organized such that they show the absolute or normalized cross sections in the upper part of the figure, with the ratios of the predictions from the MC generators (Alpgen+Pythia and Sherpa) divided by the measurement from data in the lower part. The total systematic uncertainty is shown as a yellow uncertainty band on the ratio plots, while the total combined statistical and systematic uncertainty is shown as a black error bar.

10.1 Inclusive and Exclusive Jet Multiplicities

Final states with leptons and jets are a typical signature of Higgs production and new physics processes. Since jet counting and jet vetoes serve as important discriminators against background events, the correct modelling of jet multiplicity in typical SM processes like $Z/\gamma^*(\rightarrow ee)$ + jets is crucial. ME+PS generators are usually used for this purpose, however the predicted multiplicity depends on the implementation of higher-multiplicity matrix elements, the modelling of soft gluon radiation, and the PS matching procedure employed to prevent double-counting during merging with the ME.

Figure 10.1 shows the absolute cross sections for the inclusive and exclusive jet multiplicity in $Z/\gamma^*(\rightarrow ee)$ + jet events. The predictions include a 5% uncertainty from the normalization to the NNLO Drell-Yan cross section and a 2.8% uncertainty on the luminosity. The uncertainty is dominated by the JES+JER at the order of 5-20% for $Z/\gamma^*(\rightarrow ee)$ + (1 – 4) jets (see Chapter 9). Final states with up to 8 jets are now accessible. The expected staircase pattern of the inclusive and exclusive jet multiplicity (see Section 2.3.3) is observed in data and is well modelled by predictions from both Alpgen+Pythia and Sherpa. Predictions of the multiplicity from both Alpgen and Sherpa agree well with data for up to six jets in the final state. For both Alpgen+Pythia and Sherpa, events with up to five jets in the final state are produced by explicit LO matrix elements. Larger multiplicities are created by parton shower. MC predictions overestimate the number of observed events with more than six jets in the final state. Slightly better agreement is seen with Alpgen+Pythia for the highest jet multiplicity states.

The staircase pattern of the exclusive jet multiplicity changes after introducing a hard scale via the requirement of p_T(leading jet)$>$ 300 GeV or $p_T^Z > 300$ GeV as shown in Figure 10.2. The exclusive jet multiplicity is dominated by events with exactly two jets in the final state. With a hard leading jet ($p_T^{jet} > 300$ GeV), 50% of the Z events are now accompanied by a second jet with $p_T > 30$ GeV. These events are characterized by a large angular separation between the two leading jets, typically two jets back-to-back in the transverse plane with a Z boson radiated from one of the two leading jets. This can be seen later in a comparison between the Figures 10.7(a)–(b) which show the $\Delta\phi^{jj}$ distribution for the inclusive p_T selection and the selection with a hard cut on the p_T of the leading jet respectively. Conversely, events with a hard cut of the Z boson ($p_T^Z > 300$ GeV) are characterized by a small $\Delta\phi^{jj}$ and small Rp_T^{jj} where a high-p_T Z boson recoils against a dijet.



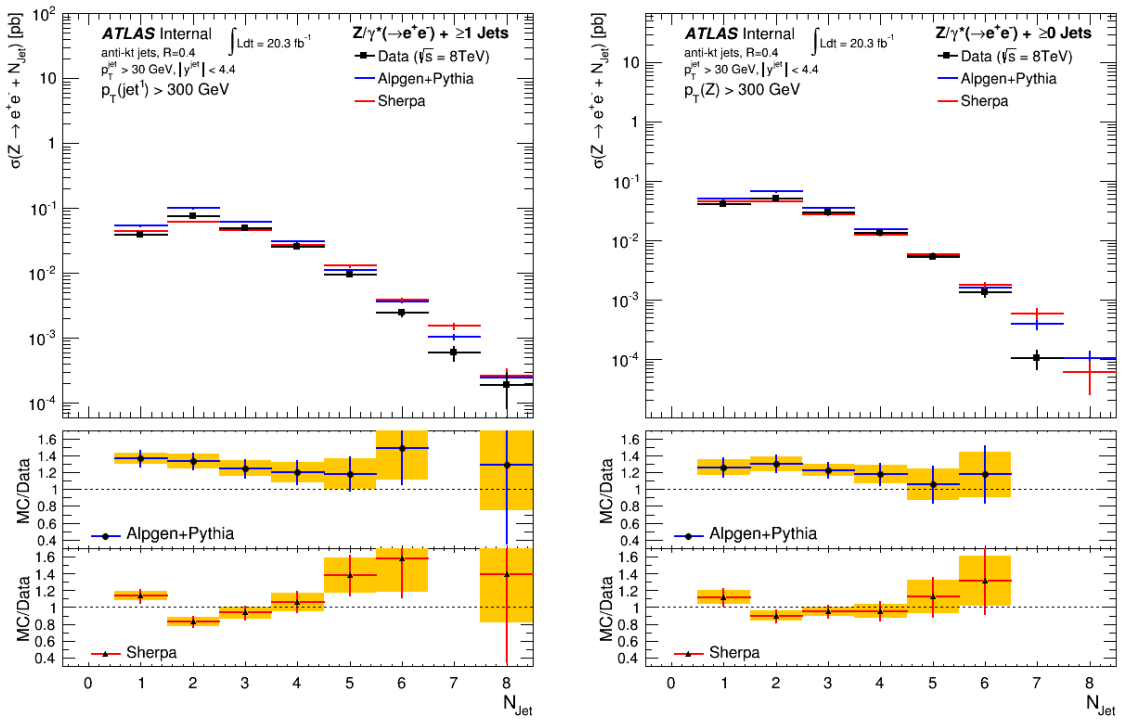
(a) Exclusive jet multiplicity

(b) Inclusive jet multiplicity

Figure 10.1: Exclusive jet multiplicity (a) and inclusive jet multiplicity (b) in $Z/\gamma^*(\rightarrow ee)$+jets events for data unfolded to particle level compared to predictions from Alpgen+Pythia and Sherpa. The yellow band in the ratio plots represent the total systematic uncertainty. The error bars represent the total systematic and statistical uncertainty.

10.2 Transverse Momentum Distributions

A correct modelling of the radiation of multiple low-energy partons does not necessarily imply a correct prediction for the production of higher energy jets. In particular, if the jet transverse momentum surpasses the scale given by the gauge boson mass, the resulting final state is expected to not be dominated by a gauge boson recoiling against the leading jet, but instead acquire significant contributions from diagrams which correspond to a gauge boson radiated off a quark line, leading to large NLO corrections for p_T^{jet} of the order $\alpha_s \ln^2(p_T^{jet}/m_Z)$ [6]. This leads to large scale uncertainties on LO ME+PS predictions as well as on fixed order NLO calculations. An increasing impact of higher-order electroweak corrections is also expected with increasing transverse momentum of the Z boson or jets, which are expected to diminish the cross sections by 10-20% for 100 GeV $< p_T^Z <$ 500 GeV [7]. Neither Alpgen+Pythia



(a) $p_T^{leadingjet} > 300$ GeV (b) $p_T^Z > 300$ GeV

Figure 10.2: Exclusive jet multiplicity for (a) $p_T^{leading\ jet} > 300$ GeV and (b) $p_T^Z > 300$ GeV in $Z/\gamma^*(\rightarrow ee)$+jets events for data unfolded to particle level compared to predictions from Alpgen+Pythia and Sherpa. The yellow band in the ratio plots represent the total systematic uncertainty. The error bars represent the total systematic and statistical uncertainty.

or Sherpa currently include EW corrections. Only the predictions from Sherpa partly include higher-order electroweak corrections, due to its implementation of additional photon radiation

Figures 10.3(a)–(b) show the normalized transverse momentum distributions of the leading jet and sub-leading jets in data unfolded to particle level compared to normalized predictions from Alpgen+Pythia and Sherpa. Sherpa predictions of the leading jet and sub-leading jet p_T are consistent with data. Alpgen+Pythia predictions of the sub-leading jet p_T are roughly consistent with the spectrum observed in data, however there are significant discrepancies between Alpgen+Pythia and data for the leading jet p_T. For a small transverse momentum of the leading jet, the Alpgen+Pythia predictions are consistent with the observed spectrum,

but for larger values of p_T^{jet}, Alpgen+Pythia predicts a much too hard spectrum that is not compatible with the spectrum observed in data.

A similar trend is observed in the complementary measurement of the p_T of the Z boson in events with at least one jet, as shown in Figure 10.3(c), which can be performed with a greater precision since it does not rely on the JES. Sherpa predictions are consistent with the spectrum measured in data. Alpgen+Pythia predicts a too hard p_T^Z distribution to an extent compatible with higher-order EW corrections.

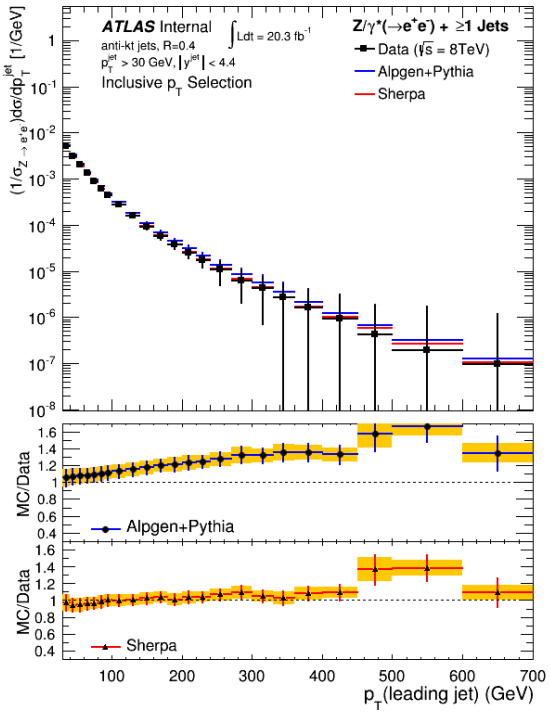
10.3 Transverse Momentum Ratios

Ratios between the transverse momenta of objects in the final state provide a more precise test of the SM than the individual transverse momentum distributions due to the fact that part of the systematics related to the JES and JER cancel, therefore predictions can be tested to a higher level of accuracy in high-p_T regimes where statistical uncertainties are large.

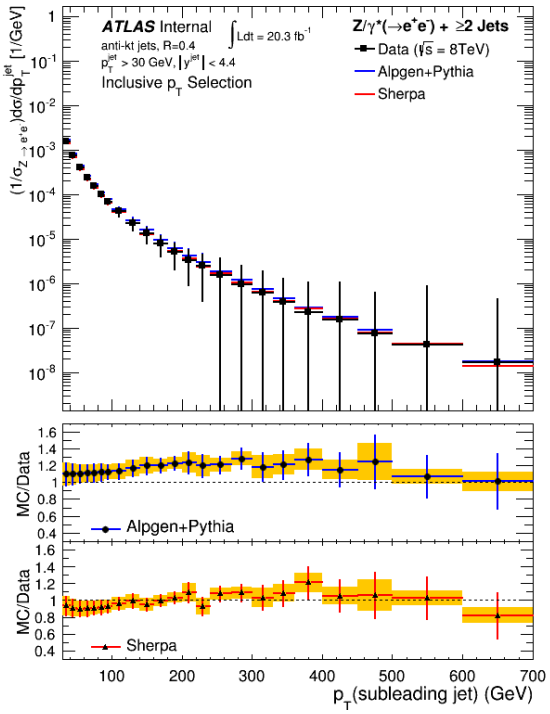
Figure 10.4 shows the transverse momentum ratio of the sub-leading jet to the leading jet (Rp_T^{jj}) for events with at least two jets in the final state. For the inclusive p_T selection shown in Figure 10.4(a), predictions from Sherpa model the transverse momentum ratio quite well. Alpgen+Pythia predictions are consistent with data for high values of the ratio only. Low values of the the ratio $Rp_T^{jj} < 0.3$ correspond to a hard leading jet ($p_T^{jet} > 300$ GeV) where 2-jet multiplities dominate. In this region, the predictions from Alpgen+Pythia overestimate the measurement substantially. This is consistent with the results of the measurement of the jet p_T spectrum shown in Figures 10.3(a)–(b).

The region where the leading jet $p_T > 300$ GeV is explored in more detail in Figure 10.4(b). Sherpa predictions remain consistent with the measured ratio in data, whereas the deviation with Alpgen+Pythia becomes more significant.

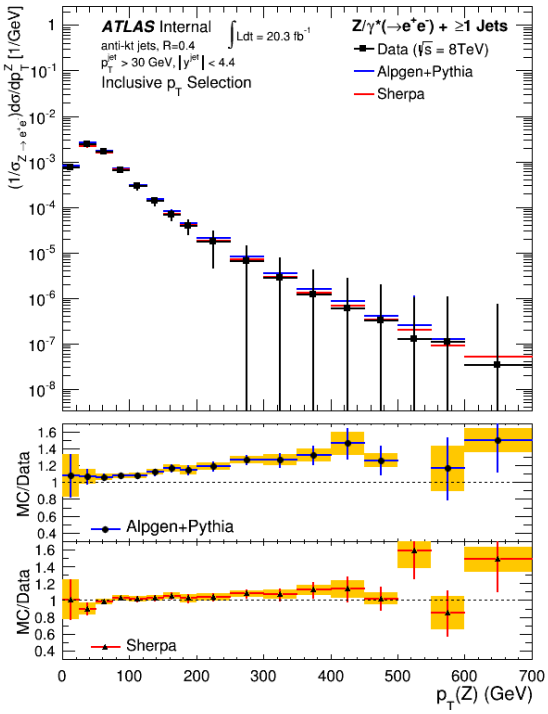
Events with a hard leading jet typically have two jets of similar p_T^{jet} back-to-back in the transverse plane, whereas events with large p_T^Z are characterized by a small $\Delta\phi$ between the



(a) Leading jet p_T for $N_{jet} \geq 1$

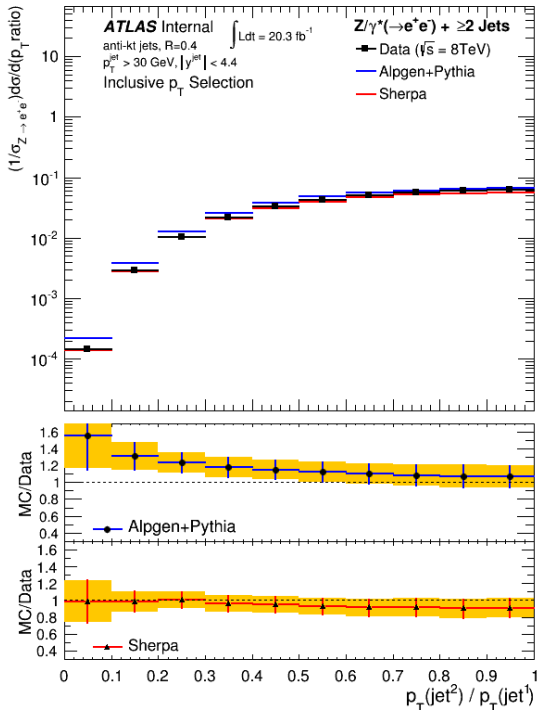


(b) Sub-leading jet p_T for $N_{jet} \geq 2$

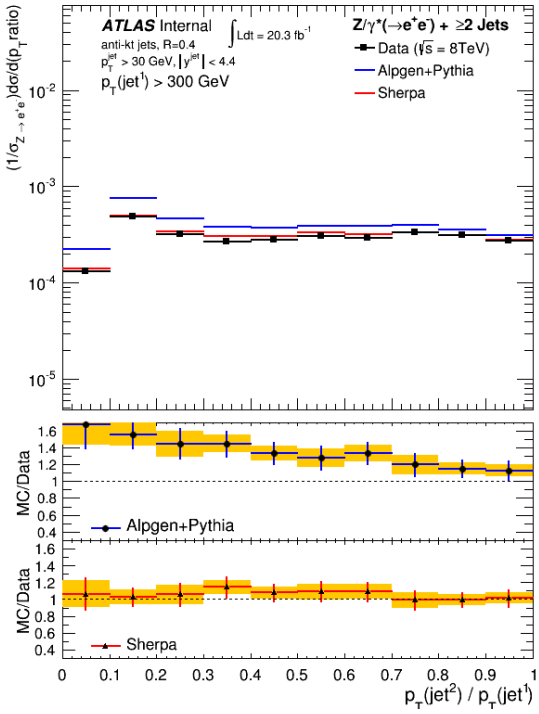


(c) Z boson p_T for $N_{jet} \geq 1$

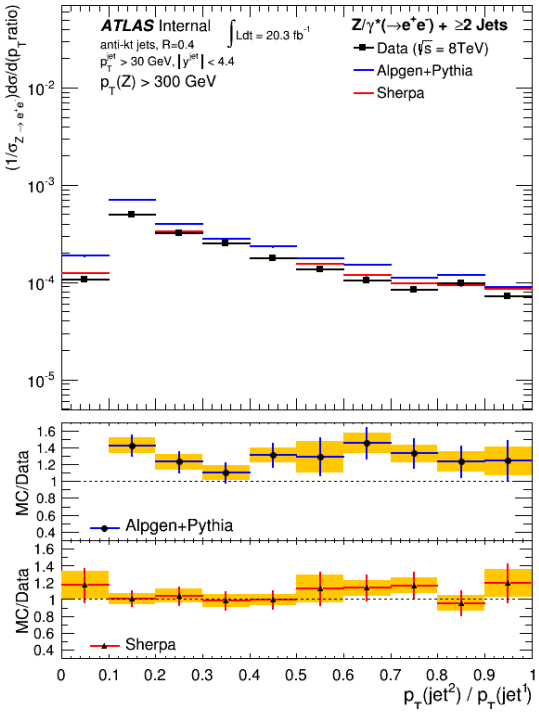
Figure 10.3: (a) p_T of the leading jet in events with at least one jet, (b) p_T of the sub-leading jet in events with at least two jets, and (c) p_T of the Z boson in events with at least one jet for data unfolded to particle level compared to predictions from Alpgen+Pythia and Sherpa. The yellow band in the ratio plots represent the total systematic uncertainty. The error bars represent the total systematic and statistical uncertainty.



(a) Inclusive p_T Selection



(b) p_T(leading jet) > 300 GeV



(c) p_T(Z) > 300 GeV

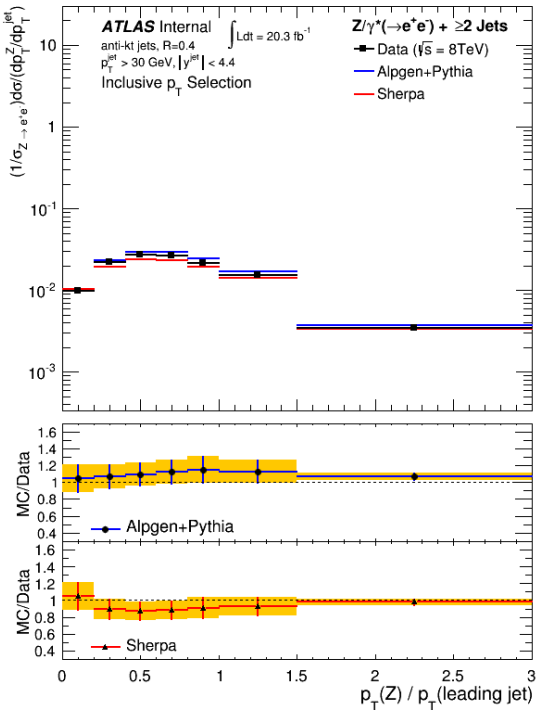
Figure 10.4: Tranverse momentum ratio between the sub-leading jet and the leading jet (Rp_T^{jj}) in events with at least two jets in the final state for the (a) inclusive p_T selection, (b) p_T(leading jet) > 300 GeV, and (c) p_T(Z) > 300 GeV.

two leading jets and small Rp_T^{jj}, thus introducing a large scale difference between the two leading jets. Figure 10.4(c) shows the transverse momentum ratio Rp_T^{jj} for events where $p_T^Z > 300$ GeV. Sherpa seems able to describe 2-jet final states with large scale differences. Alpgen+Pythia fails to model the shape of the distribution, especially for low values of the ratio.

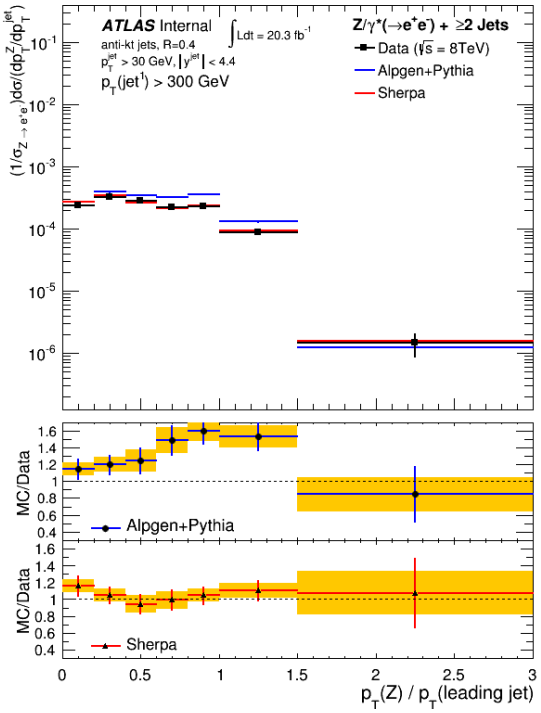
A complementary measurement to Rp_T^{jj} can be made of the ratio between the transverse momentum of the Z boson and the transverse momentum of the leading jet Rp_T^{Zj}. Figure 10.5 shows the transverse momentum ratio Rp_T^{Zj} for events with at least two jets in the final state. Predictions from both Sherpa and Alpgen+Pythia are consistent with data for the inclusive p_T selection. In events with a hard leading jet ($p_T^{jet} > 300$ GeV) or high-p_T Z boson ($p_T^Z > 300$ GeV), Sherpa predictions remain consistent with data. However, predictions from Alpgen are not able to reproduce the distribution from data for values of $Rp_T^{Zj} < 1.5$.

10.4 Jet Rapidity

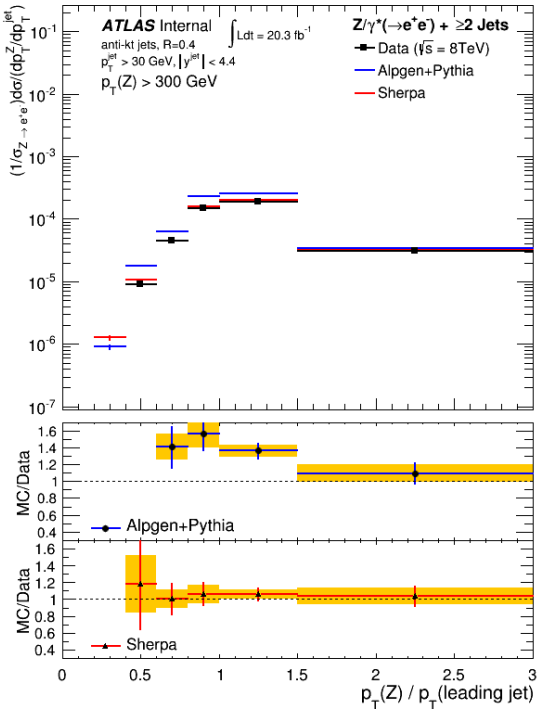
The correct modelling of the rapidity distribution of the jets, especially in the forward region, is an essential quantity since it plays an important role in searches for new physics and studies of the Higgs boson. Such processes are characterized by highly energetic, well separated forward jets. Therefore, an excellent knowledge of the rapidity distribution can be used to distinguish signal from background. Figures 10.6 shows the rapidity spectra of the leading jet and of the sub-leading jet. Alpgen+Pythia predictions for the leading jet and sub leading jet are consistent with data. However, Sherpa predicts a too wide rapidity spectrum. In the forward regions, the modelling of hadronization effects, pile-up, and of the transverse momentum and resulting migration leads to substantial systematic uncertainties from the JES and unfolding procedure. This limits the precision of the measurement and the conclusion that can be drawn from comparisons with MC predictions in this extreme region. The unfolding uncertainty may be optimized using a multidimensional unfolding technique. Multidimensional unfolding could reduce the model dependence since the full event kinematics



(a) Inclusive p_T Selection



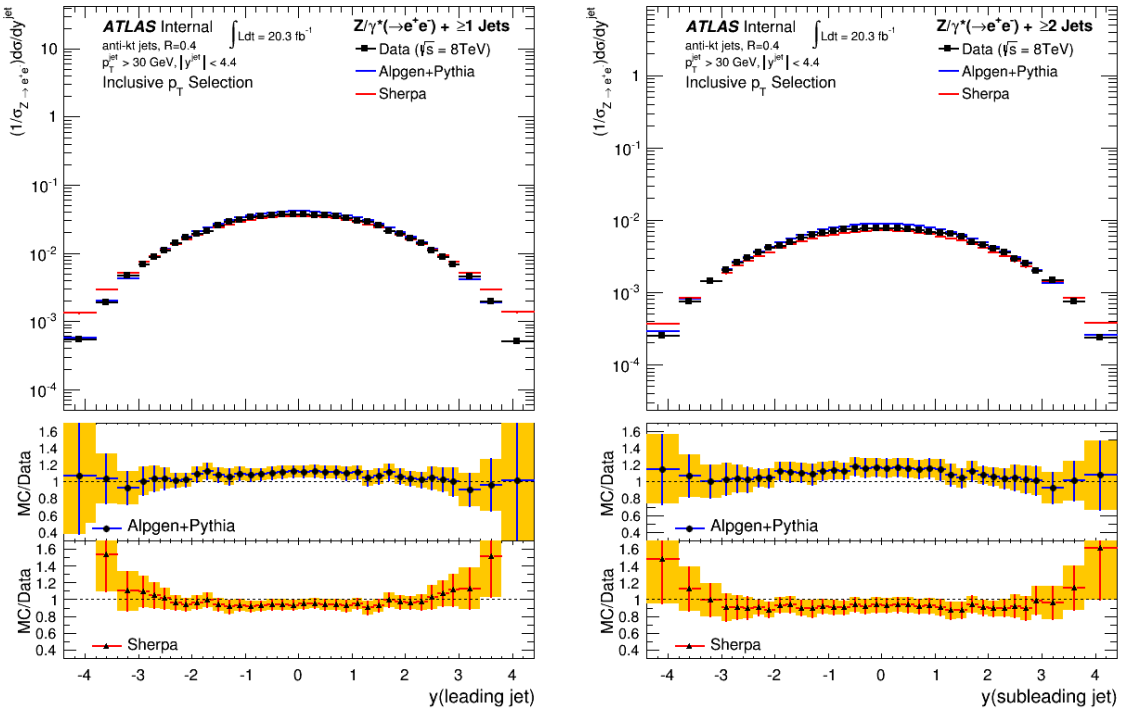
(b) p_T(leading jet) > 300 GeV



(c) p_T(Z) > 300 GeV

Figure 10.5: Tranverse momentum ratio between the Z boson and leading jet (Rp_T^{Zj}) in events with at least two jets in the final state for the (a) inclusive p_T selection, (b) p_T(leading jet) > 300 GeV, and (c) p_T(Z) > 300 GeV.

are taken into account. However, as additional unfolding dimensions are added (i.e. p_T^{j1}, p_T^{j2}), this technique becomes very computationally intensive. Furthermore, multiple dimensions result in fewer statistics in each bin of the migration matrix. In order to keep the statistical uncertainty under control, a much larger sample of simulated signal events is required than what is available.



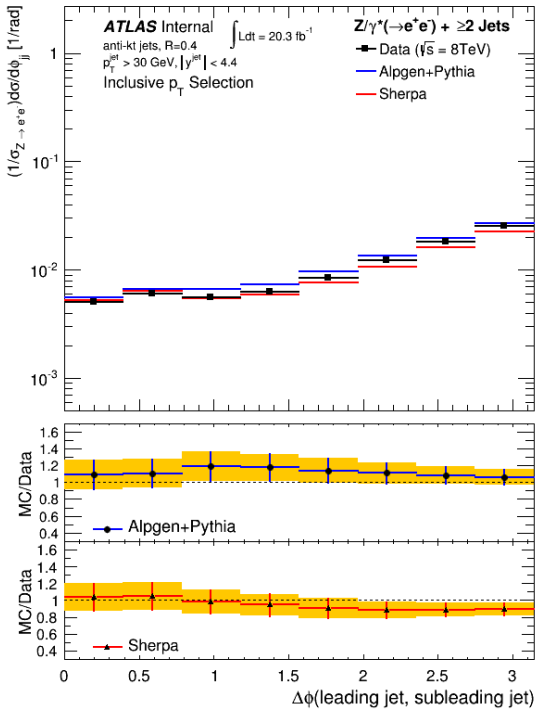
(a) Leading jet rapidity

(b) Sub-leading jet rapidity

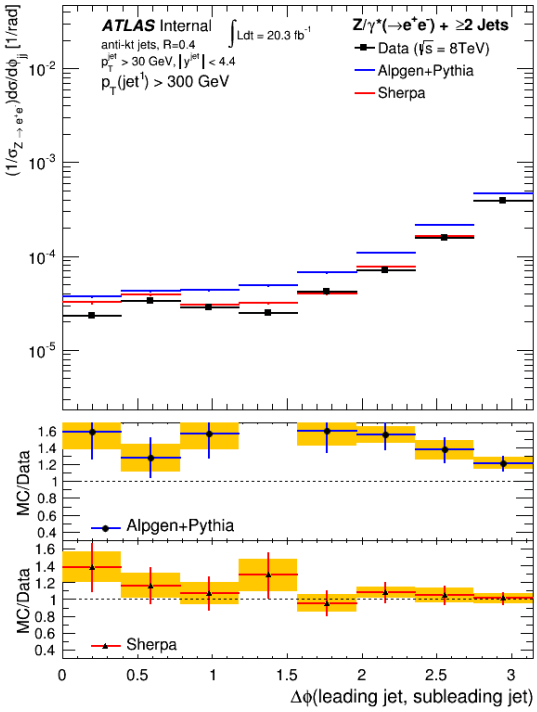
Figure 10.6: Jet rapidity for (a) the leading jet and (b) the sub-leading jet in $Z/\gamma^*(\rightarrow ee)$ + jet events for data unfolded to particle level compared to predictions from Alpgen+Pythia and Sherpa. The yellow band in the ratio plots represent the total systematic uncertainty. The error bars represent the total systematic and statistical uncertainty.

10.5 Angular Separation Distributions

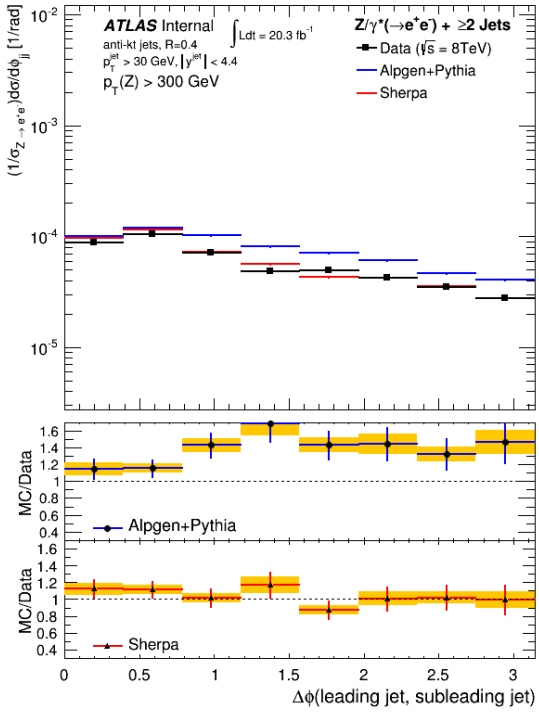
The topology of dijet final states in Z/γ^* + jets events can be explored with the measurement of angular separation variables. Figures 10.7 shows the absolute azimuthal separation ($\Delta\phi^{jj}$) between the two leading jets, and Figures 10.8 shows the minimum azimuthal separation between the Z boson and either one of the two leading jets ($min\Delta\phi^{Zjj}$) in events with at least two jets in the final state. Predictions from Alpgen+Pythia and Sherpa are consistent



(a) Inclusive p_T Selection

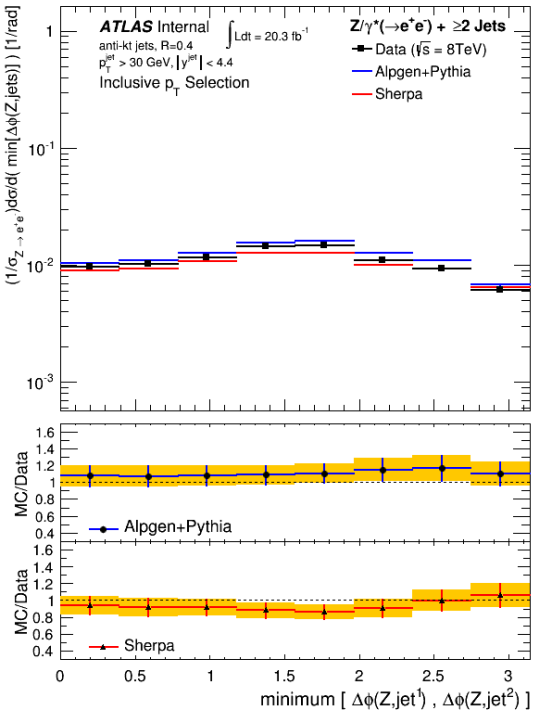


(b) p_T(leading jet) > 300 GeV

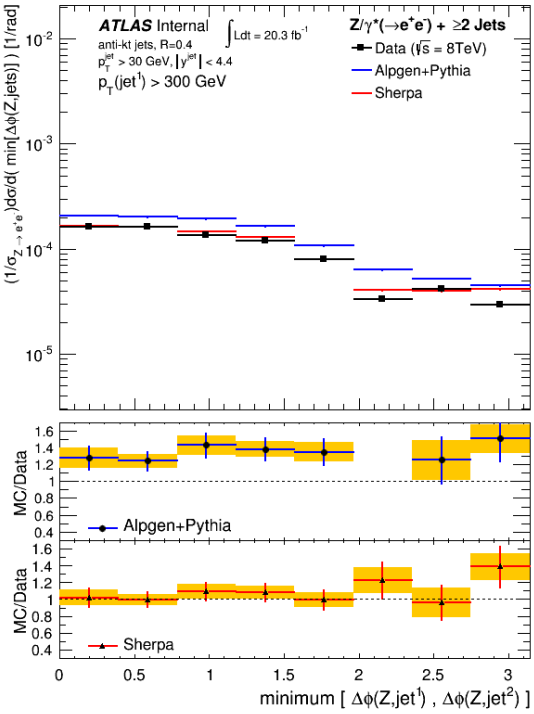


(c) p_T(Z) > 300 GeV

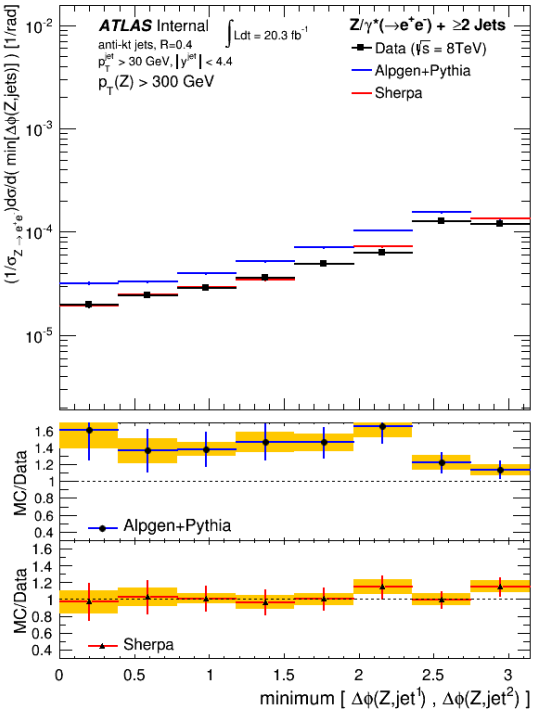
Figure 10.7: Absolute azimuthal separation between the two leading jets ($\Delta\phi^{jj}$) in events with at least two jets in the final state for the (a) inclusive p_T selection, (b) p_T(leading jet) > 300 GeV, and (c) p_T(Z) > 300 GeV.



(a) Inclusive p_T Selection



(b) p_T(leading jet) > 300 GeV



(c) p_T(Z) > 300 GeV

Figure 10.8: Minimum $\Delta\phi$ between the Z boson and either of the two leading jets ($min\Delta\phi^{Zjj}$) in events with at least two jets in the final state for the (a) inclusive p_T selection, (b) p_T(leading jet) > 300 GeV, and (c) p_T(Z) > 300 GeV.

with data for both $\Delta\phi^{jj}$ and $min\Delta\phi^{Zjj}$ with the inclusive p_T selection applied as shown in Figure 10.7(a) and Figure 10.8(a) respectively.

A hard cut on the p_T of the leading jet ($p_T^{jet} > 300$ GeV) enhances the contribution of events where the two leading jets are widely separated in the transverse plane with the Z boson being emitted relatively close to one of the two leading jets. This can be seen in the shape of the distributions of $\Delta\phi^{jj}$ in Figure 10.7(b) and $min\Delta\phi^{Zjj}$ in Figure 10.8(b) where events with large $\Delta\phi^{jj}$ and small $min\Delta\phi^{Zjj}$ become more prominent when compared to the inclusive p_T selection. Sherpa predicts a slightly too flat distribution for the $\Delta\phi^{jj}$ with a hard p_T cut on the leading jet, but is mostly consistent with the distribution measured in data. Sherpa predictions for $min\Delta\phi^{Zjj}$ agree well with data. Alpgen+Pythia predicts a too flat distribution for both $\Delta\phi^{jj}$ and $min\Delta\phi^{Zjj}$ that is not compatible with the distribution measured in data when a hard cut on the p_T the leading jet is applied.

Conversely, dijet events containing a high-p_T Z boson ($p_T^Z > 300$ GeV) are typically characterized by a small $\Delta\phi^{jj}$ between the two leading jets which recoil against the Z boson resulting in a large angular separation between the Z boson and the dijet system. This is reflected in the measurements of $\Delta\phi^{jj}$ and $min\Delta\phi^{Zjj}$ as shown in Figures 10.7(c) and 10.8(c). Predictions from Sherpa are consistent with the observations from data. Alpgen+Pythia again predicts a much too flat distribution for both $\Delta\phi^{jj}$ and $min\Delta\phi^{Zjj}$, where the cross section is overestimated by as much as 30% to 60%.

10.6 Discussion

Results from Sherpa for the N_{jet} distributions shown in Figure 10.1 are similar to the results from Sherpa from the 7 TeV analyses by the ATLAS and CMS experiments [10,11], as shown in Figure 2.11. The results show a slight underestimation by Sherpa for jet multiplicities N_{jet} between 1 and 5. The results from Alpgen+Pythia for the N_{jet} distributions in this analysis seem to indicate a tendency to overestimate the highest jet multiplicities, although results are consistent within experimental uncertainties. However, the results from the 7 TeV analysis

by ATLAS show a tendency by Alpgen+Herwig to underestimate the high jet multiplicity states. In events generated by Alpgen+Pythia, jets beyond the 5^{th} leading jet are modelled by parton shower. In this analysis of the 8 TeV data set, Alpgen was interfaced with Pythia for parton shower, which uses a p_T-ordered parton shower model. In the previous analysis, Alpgen was interfaced with Herwig for parton shower, which uses an angular-ordered parton shower model. Therefore, some differences between the results of the 8 TeV and 7 TeV analyses of high jet multiplicity states might be caused by differences in the PS model.

The expected staircase scaling pattern of the jet multiplicity is observed in data and is well modelled by both Alpgen+Pythia and Sherpa, as was also seen in the 7 TeV results from ATLAS and CMS. A hard cut on the p_T of the leading jet or the p_T of the Z boson induces changes in the scaling pattern of the exclusive jet multiplicity, as seen in Figure 10.2, which are not well reproduced by Alpgen+Pythia. Sherpa seems to reproduce the scaling pattern within uncertainties when a hard cut on the p_T of the Z boson is applied, but does not reproduce the scaling pattern when a hard cut on the p_T of the leading jet is applied. Theory predicts a transition to Poisson scaling (see Section 2.3.3) for large scale differences between the leading jet and the additional radiated jets. For further studies of $Z/\gamma^*(\rightarrow ee)$ + jets production beyond the scope of this analysis, it would be of interest to compare fits of a Poisson distribution to the jet cross section ratio measured in data and MC when a hard scale is applied in order to test this prediction.

Trends observed in the distributions of the p_T and rapidity of the leading jet, and the p_T of the Z boson by the 7 TeV analyses persist in this analysis of the 8 TeV data set. This analysis found that Sherpa predicts a wider rapidity distribution of the leading jet than observed in data, while predicitions from Alpgen+Pythia are in agreement with data, as shown in Figure 10.6(a). This result is consistent with what was observed in the 7 TeV analysis by the ATLAS experiment, as shown in Figure 2.13(b).

Sherpa predictions of the p_T distribution of the leading jet, shown in Figure 10.3(a), and of the p_T distribution of the Z boson, shown in Figure 10.3(c), are consistent with data. The level of agreement between Sherpa and data is very similar to that which was seen in the results from the 7 TeV analyses of the p_T of the leading jet, shown in Figure 2.12, and the p_T of the Z boson, shown in Figure 2.13(a). This analysis also found that Alpgen+Pythia predicts a much too hard p_T spectra of the leading jet and the Z boson, which was the same observation made by the analysis of the 7 TeV data set by the ATLAS experiment. The level of agreement between Alpgen+Pythia and data in the distribution of the p_T of the Z boson is very similar with what was observed in the 7 TeV analysis from ATLAS. However, worse disagreement is seen in the comparison between Alpgen+Pythia and data in the distribution of the p_T of the leading jet than in the previous analysis by ATLAS.

The analysis of the 7 TeV data set by the ATLAS experiment also measured the distributions of Rp_T^{jj} and $\Delta\phi^{jj}$, as shown in Figures 2.13(c) and 2.13(d) respectively, for an inclusive p_T selection. Results from this analysis of the measurements of Rp_T^{jj} and $\Delta\phi^{jj}$, as shown in Figures 10.4(a) and 10.7(a) respectively, are similar to the results from the 7 TeV analysis. Predictions from Alpgen+Pythia overestimate low values of the Rp_T^{jj} distribution, which corresponds to a region with a large scale difference between the two leading jets, while predictions from Sherpa are consistent with data. Predictions from Alpgen+Pythia and Sherpa for the $\Delta\phi^{jj}$ distribution are consistent with data, however Sherpa predictions in the 7 TeV analysis were found to model a too flat spectrum.

In addition to the Rp_T^{jj} and $\Delta\phi^{jj}$ distribtions, this analysis measured the Rp_T^{Zj} and $min\Delta\phi^{Zjj}$ distributions in events with at least 2 jets in the final state. These four distributions have been measured for the standard inclusive p_T selection, as well as in two regions of high-p_T phase space which are defined by a hard cut on either the p_T of the leading jet ($p_T^{leading\ jet} > 300$ GeV) or on the p_T of the Z boson ($p_T^Z > 300$ GeV). In general, without a high-p_T selection applied, the observed spectra for these distributions are well modelled by

Alpgen+Pythia and Sherpa, with the exception of Rp_T^{jj} in which Alpgen+Pythia overestimates the low values of this distribution. In regions of high-p_T phase space, Sherpa continues to model these distributions well. Alpgen+Pythia, however, fails to reproduce the shapes of the distributions observed in data, where typical deviations of 40–60% are seen. The angular separation variables ($\Delta\phi^{jj}$ and $min\Delta\phi^{Zjj}$) modelled by Alpgen+Pythia in the high-p_T region are too flat. The shapes of the transverse momentum ratios (Rp_T^{jj} and Rp_T^{Zj}) in the high-p_T region are poorly modelled by Alpgen+Pythia, especially for large scale differences between the two leading jets. A summary of the comparisons between measurements from data and predictions from Alpgen+Pythia and Sherpa for the observables measured in this analysis is shown in Table 10.1.

These observations show that Sherpa provides an improved description of the transverse momentum of the Z boson and leading jet at high-p_T compared to that provided by Alpgen+Pythia. As a result, Sherpa seems to model the distributions of the tranverse momentum ratios and angular separation variables much better than Alpgen+Pythia in high-p_T regions of phase space for $Z/\gamma^* + 2$ jet final states. A few differences exist between Alpgen+Pythia and Sherpa in their implementation that could influence the modelling of $Z/\gamma^* +$ jets final states. Most importantly, Sherpa calculates the inclusive $Z/\gamma^* +$ jets cross section at NLO, which by design, produces the hardest parton emission (leading jet) at NLO, with higher jet multiplicities produced at LO. Alpgen+Pythia, on the other hand, provides LO calculations for both the inclusive process and higher jet multiplicity final states. Additional differences between Alpgen+Pythia and Sherpa exist regarding parton shower merging, underlying event modelling, hadronization, and the implementation of photon radiation. For parton shower, underlying event, and hadronization, Alpgen is interfaced with Pythia, while Sherpa provides its own models. Both Pythia and Sherpa employ a p_T-ordered parton shower based on a dipole showering model. Alpgen uses the MLM [52, 53] matching scheme to merge the ME calculation with the parton shower from Pythia, while Sherpa uses the CKKW [54] matching scheme for its merging of the NLO ME calculation with its own parton shower.

(a) N_{jet}, p_T, rapidity

	Alpgen+Pythia	**Sherpa**
N_{jet}	good for $N_{jet} \leq 6$	good for $N_{jet} \leq 6$
p_T^Z	too hard for $p_T > 100$ GeV	good
$p_T^{leading\ jet}$	too hard for $p_T > 100$ GeV	good
$p_T^{sub-leading\ jet}$	fair	good
$y^{leading\ jet}$	good	too wide
$y^{sub-leading\ jet}$	good	too wide

(b) Alpgen+Pythia

	Alpgen+Pythia		
	Standard	$p_T^{j1} > 300$ GeV	$p_T^Z > 300$ GeV
Rp_T^{jj}	fail at low values	fail	fail
Rp_T^{Zj}	good	fail	fail
$\Delta\phi^{jj}$	good	fail	fail
$min\Delta\phi^{Zjj}$	good	fail	fail

(c) Sherpa

	Sherpa		
	Standard	$p_T^{j1} > 300$ GeV	$p_T^Z > 300$ GeV
Rp_T^{jj}	good	good	good
Rp_T^{Zj}	good	good	good
$\Delta\phi^{jj}$	good	fair	good
$min\Delta\phi^{Zjj}$	good	good	good

Table 10.1: Summary of the comparisons between measurements from data and predictions from Alpgen+Pythia and Sherpa for the N_{jet}, p_T, and rapidity distributions, as well as the observables sensitive to the topology of events with at least 2 jets in the final state for the standard inclusive selection and for the two high-p_T event selections.

Sherpa simulates underlying events using a multiple-interaction model based on a modified version used by Pythia. Hadronization is handled in Pythia using the Lund string model, while Sherpa uses the cluster model. Radiative photon corrections are described in Sherpa as soft wide angle emission using the YFS [109] implementation. Alpgen describes radiative photon corrections as hard collinear radiation. These differences could affect the kinematics of the final state objects in $Z/\gamma^*(\rightarrow ee) + \text{jets}$ events.

Although beyond the scope of this analysis, a detailed systematic study of the effects of the different models implemented by the various MC generators would be useful for identifying the exact origins of the discrepancies. Information gleaned from these studies could be used to tune the MC generators in order to improve the modelling of Z/γ^*+jets final states in high-p_T regions of phase space that are important to many other physics analyses. Meanwhile, the results of this validation study indicated that Sherpa is the better choice for the modelling $Z/\gamma^* + \text{jets}$ processes.

CHAPTER 11

Conclusions and Outlook

This thesis covers a large range of measurements of the production of jets in association with a Z boson, which provides valuable input for the tuning of Monte Carlo (MC) event generators. The clean decay signature of the associated Z boson in the detector allows for the easy identification of jets produced in the hard collision. Inclusive and differential $Z/\gamma^*(\rightarrow ee) + \text{jets}$ cross sections have been measured in 20.3 fb^{-1} of proton-proton collisions at a center-of-mass energy $\sqrt{s} = 8$ TeV collected by the ATLAS detector at the LHC in 2012. Total inclusive and exclusive cross sections have been measured as a function of the inclusive and exclusive jet multiplicities (N_{jet}). Differential cross sections have been measured as a function of the jet transverse momentum (p_T^{jet}), the jet rapidity (y^{jet}), and the transverse momentum of the Z boson (p_T^Z). Furthermore, differential cross sections have been measured for a set of observables in events with at least two jets that are sensitive to the topology of the event. These are the ratio of the p_T of the sub-leading jet to the p_T of the leading jet (Rp_T^{jj}), the ratio of the p_T of the Z boson to the p_T of the leading jet (Rp_T^{Zj}), the absolute azimuthal separation between the two leading jets ($\Delta\phi^{jj}$), and the minimum $|\Delta\phi|$ between the Z boson and either of the two leading jets in the event ($min\Delta\phi^{Zjj}$). Several sources of systematic uncertainty on the measurements have been evaluated.

Comparisons are made between measurements from data and predictions from matrix element (ME) calculations interfaced to a parton shower (PS) provided by MC generators Alpgen+Pythia and Sherpa. Alpgen calculates the matrix elements at leading order (LO) in perturbative Quantum Chromodynamics (pQCD) for up to five jets in the final state, and is interfaced to Pythia for the parton shower. Sherpa calculates the inclusive cross section at

next-to-leading order (NLO) in pQCD, while higher jet multiplicities up to five jets in the final state are calculated at LO.

Compared to the previous measurements with the 7 TeV data set from the ATLAS experiment, this analysis has been extended into high-p_T regions of phase space. This was done in order to test the modelling by ME+PS MC generators in regions not fully accessible before but are typical of searches for new physics and Higgs boson production. Therefore, the set of observables which are sensitive to the topology of the event are measured for events with high-p_T final states in addition to the standard event selection. High-p_T final states are defined as events with either $p_T^Z > 300$ GeV or $p_T^{leading\ jet} > 300$ GeV.

For some phase space regions, discrepancies between predictions from the ME+PS generators and measurements from data have been observed. The jet multiplicity is well described by Alpgen+Pythia and Sherpa up to six jets, but additional jet emissions are overestimated. A hard cut on the p_T of the leading jet or the Z boson induces a change in the scaling pattern of the jet multiplicity, which is not well reproduced by Alpgen+Pythia, while Sherpa is consistent with data for a hard cut on p_T^Z only.

Furthermore, Alpgen+Pythia overestimates the cross section for high energy jets or highly boosted Z bosons, while Sherpa predicts a rapidity distribution for jets that is too wide. Additionally, without a high-p_T selection applied, the observed spectra of the transverse momentum ratios (Rp_T^{jj} and Rp_T^{Zjj}) and angular separation variables ($\Delta\phi^{jj}$ and $min\Delta\phi^{Zjj}$) are generally well modelled by Alpgen+Pythia and Sherpa. In regions of high-p_T phase space, Sherpa continues to model these distributions well. Alpgen+Pythia, however, fails to reproduce the shapes of the distributions observed in data, where typical deviations of 40–60% are seen.

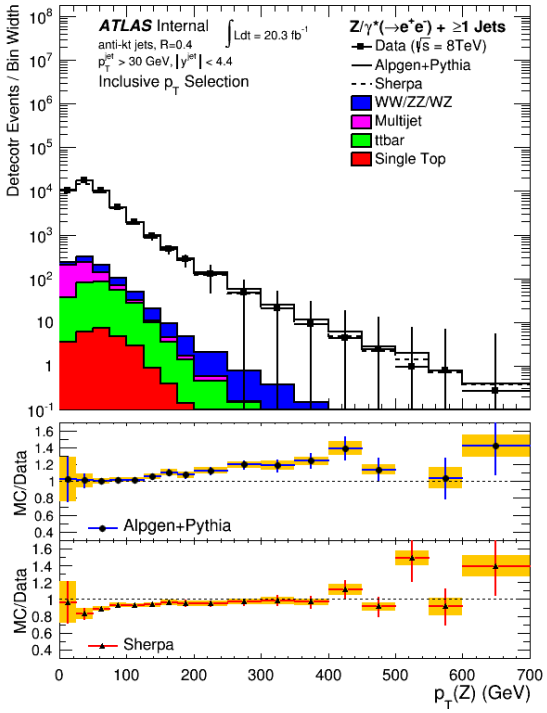
These observations indicate that Sherpa provides a better description of $Z/\gamma^* + \text{jets}$ production than Alpgen+Pythia, especially for high-p_T final states. Alpgen+Pythia and Sherpa

differ in multiple aspects regarding their respective implementation of the hard ME calculation and models of soft physics. This leaves the possibility for future analyses to identify the exact origins of the discrepancies. Finally, for high energy regimes, large logarithmic contributions appear in higher order pQCD predictions of 2-jet final states. Thus, an important extension of this analysis could provide tests of future fixed-order NLO pQCD calculations of multi-parton final states.

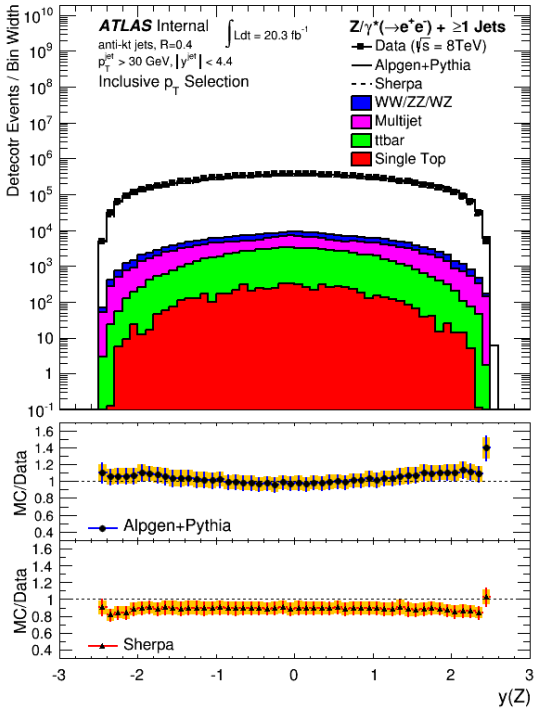
And finally, this analysis would benefit from the much larger center-of-mass energy of the 13 TeV data set, which would allow access to regions of phase space of even higher energy. These regions are currently dominated by statistical uncertainties. Thus, an increase in the amount of statistics in high-energy regions of phase space would increase the sensitivity and precision of this measurement.

APPENDIX A

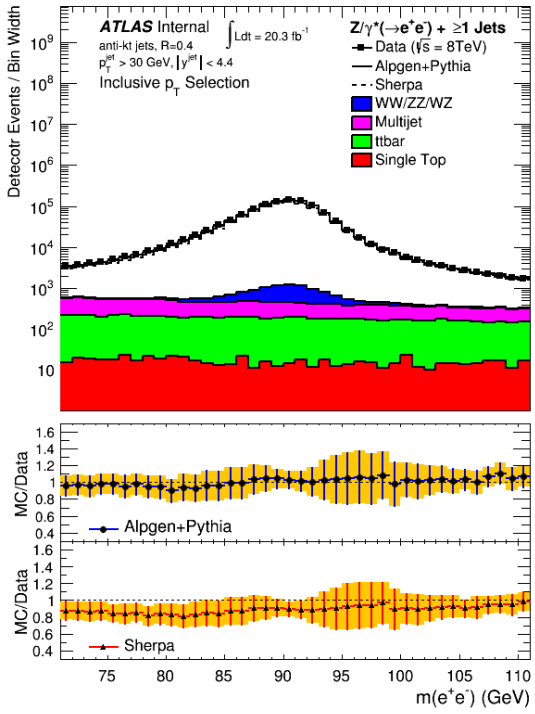
Detector Level Distributions



(a) Z boson p_T

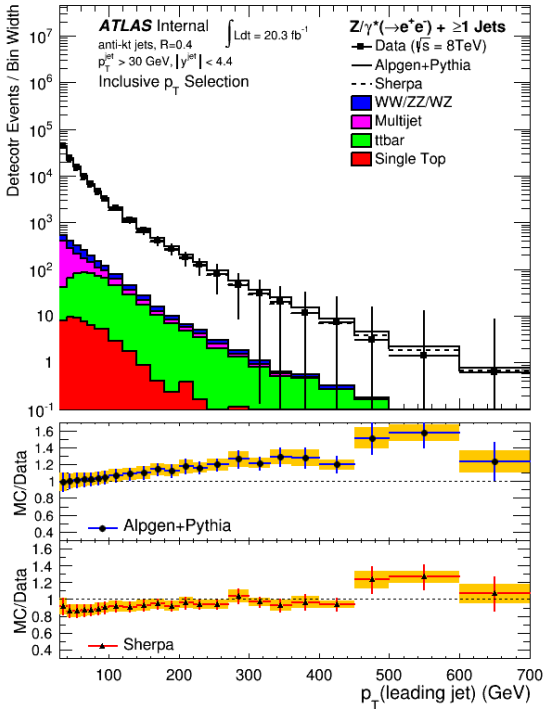


(b) Z boson rapidity

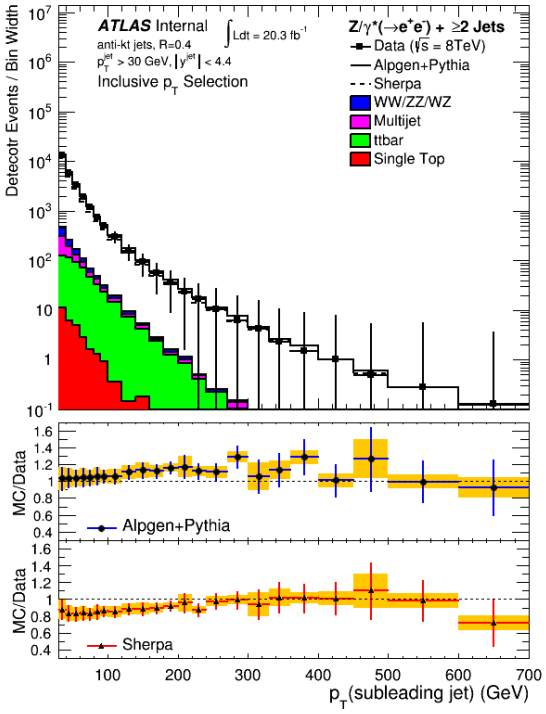


(c) Z boson mass

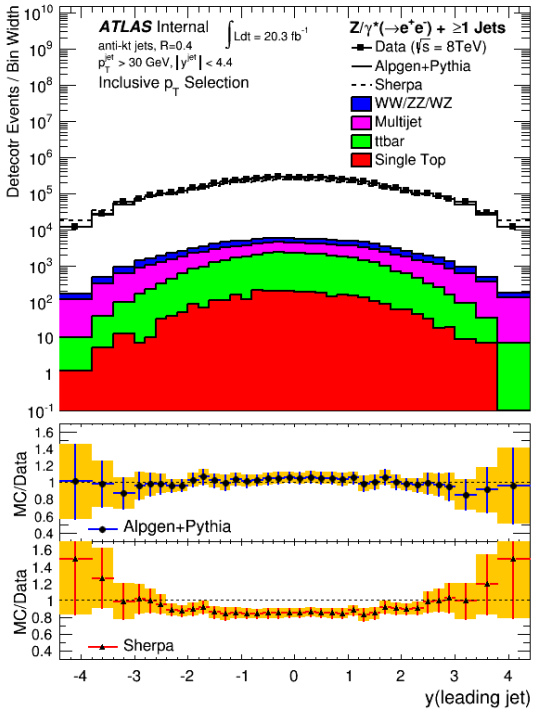
Figure A.1: (a) p_T of the Z boson, (b) rapidity of the Z boson, and (c) mass of the Z boson in events with at least one jet in $Z/\gamma^*(\rightarrow ee)$+jets events in data and simulation.



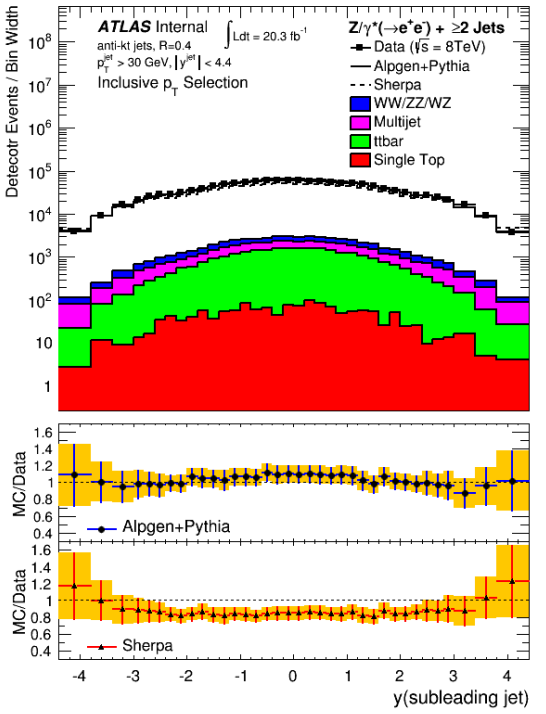
(a) Leading jet p_T



(b) Sub-leading jet p_T

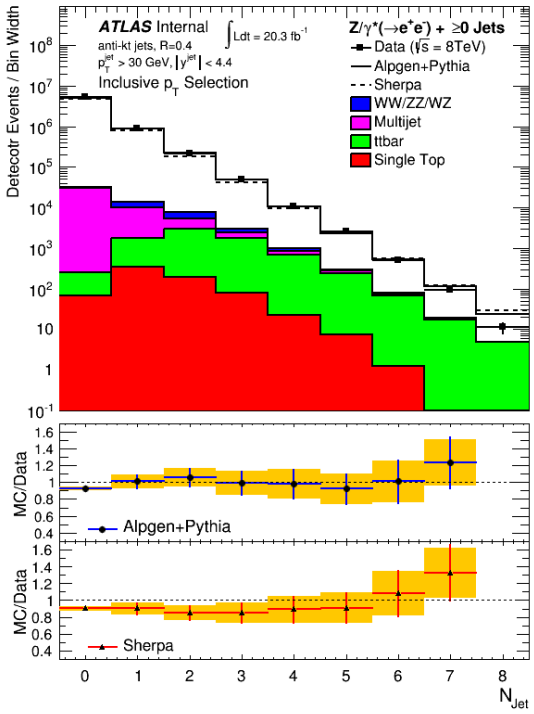


(c) Leading jet rapidity

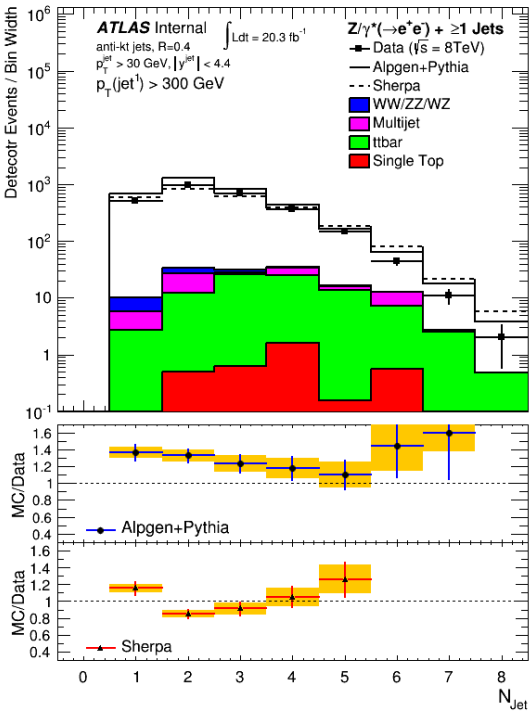


(d) Sub-leading jet rapidity

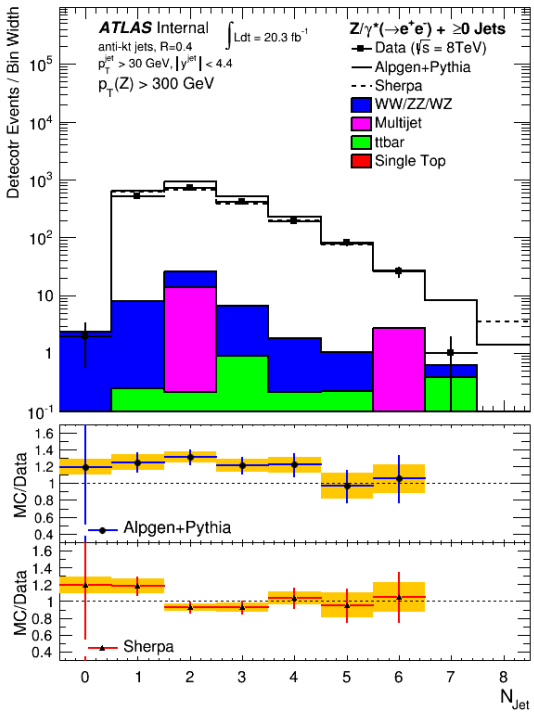
Figure A.2: (a) p_T of the leading jet in events with at least one jet, (b) p_T of the sub-leading jet in events with at least two jets, (c) rapditiy of the leading jet in events with at least one jet, and (d) rapidity of the sub-leading jet in events with at least two jets in $Z/\gamma^*(\rightarrow ee)$+jets events in data and simulation.



(a) Inclusive p_T Selection

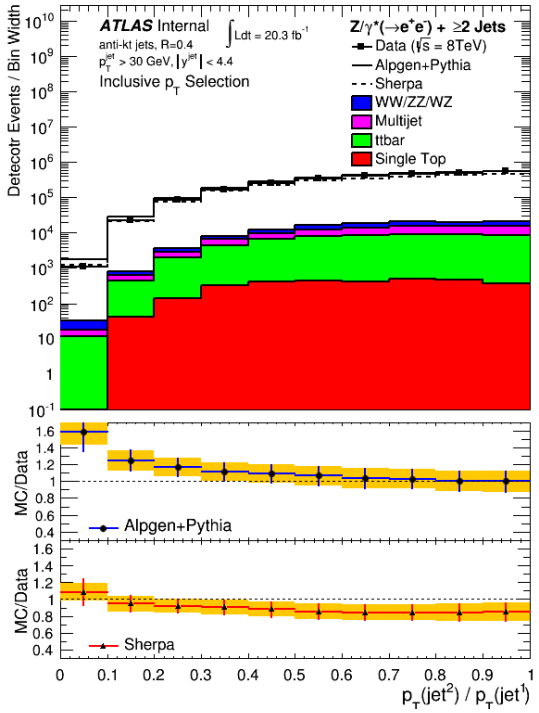


(b) p_T(leading jet) > 300 GeV

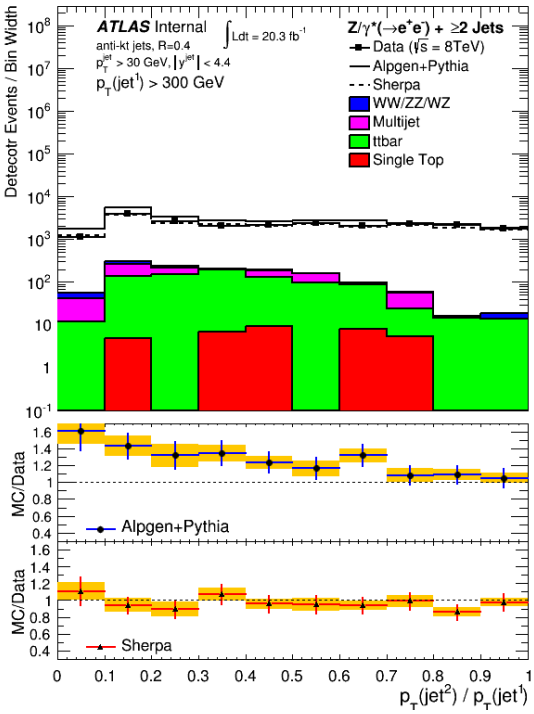


(c) p_T(Z) > 300 GeV

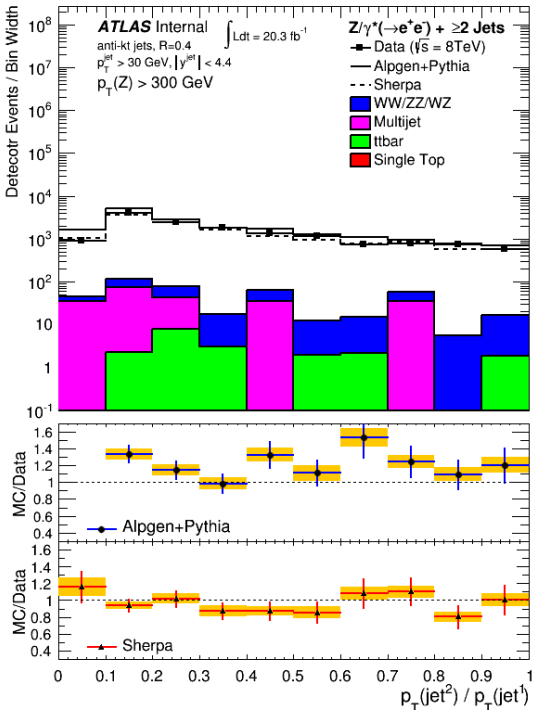
Figure A.3: Exclusive jet multiplicity (N_{jet}) in $Z/\gamma^*(\rightarrow ee)$+jets events in data and simulation for the (a) inclusive p_T selection, (b) p_T(leading jet) > 300 GeV, and (c) p_T(Z) > 300 GeV



(a) Inclusive p_T Selection

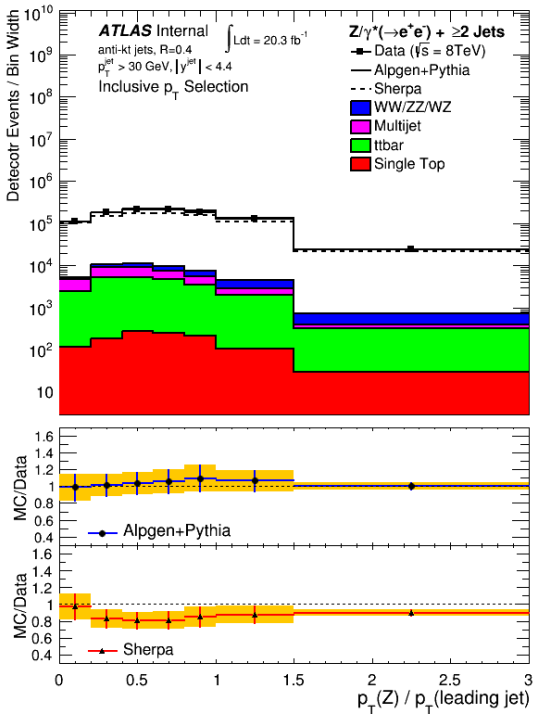


(b) p_T(leading jet) > 300 GeV

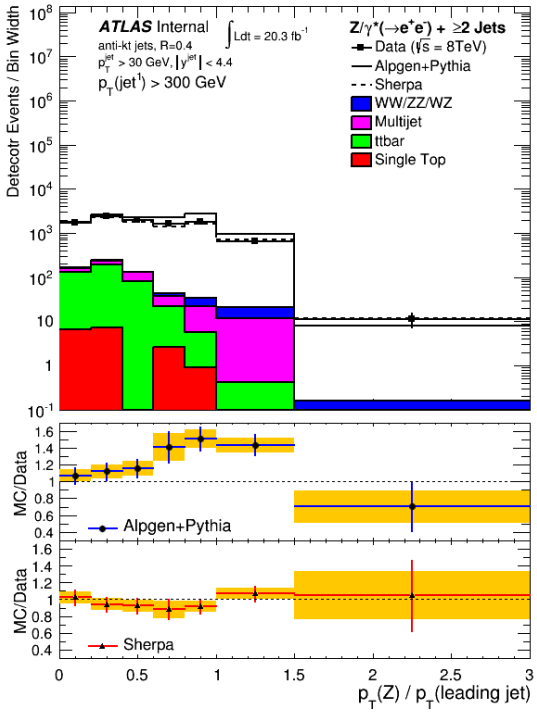


(c) p_T(Z) > 300 GeV

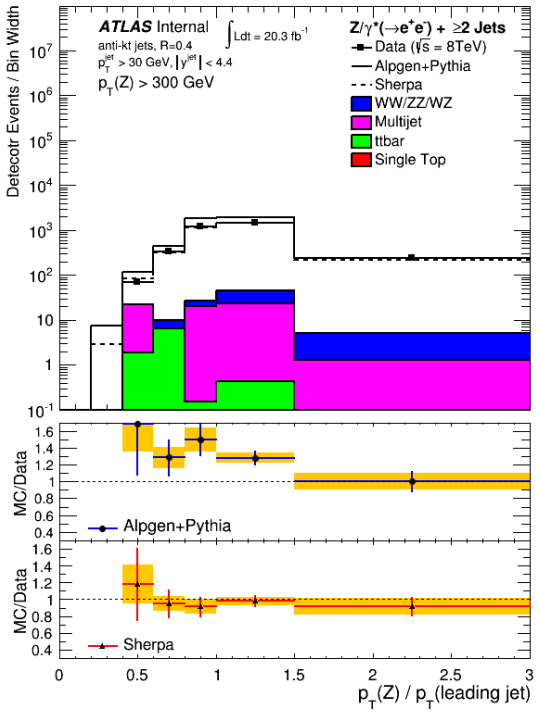
Figure A.4: Transverse momentum ratio between the sub-leading jet and the leading jet (Rp_T^{jj}) in events with at least two jets in the final state in $Z/\gamma^*(\rightarrow ee)$+jets events in data and simulation for the (a) inclusive p_T selection, (b) p_T(leading jet) > 300 GeV, and (c) p_T(Z) > 300 GeV



(a) Inclusive p_T Selection



(b) p_T(leading jet) > 300 GeV

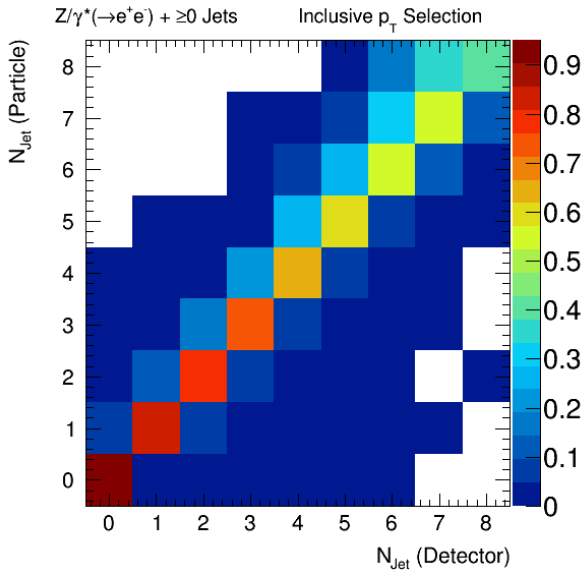


(c) p_T(Z) > 300 GeV

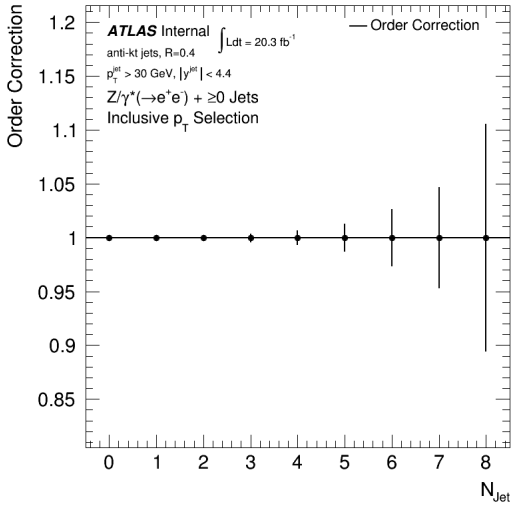
Figure A.5: Transverse momentum ratio between the Z boson and leading jet (Rp_T^{Zj}) in events with at least two jets in the final state in $Z/\gamma^*(\rightarrow ee)$+jets events in data and simulation for the (a) inclusive p_T selection, (b) p_T(leading jet) > 300 GeV, and (c) p_T(Z) > 300 GeV.

APPENDIX B

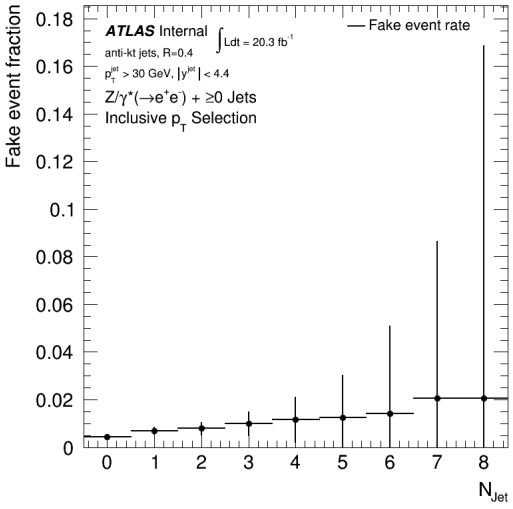
Unfolding Performance



(a) Migration Matrix

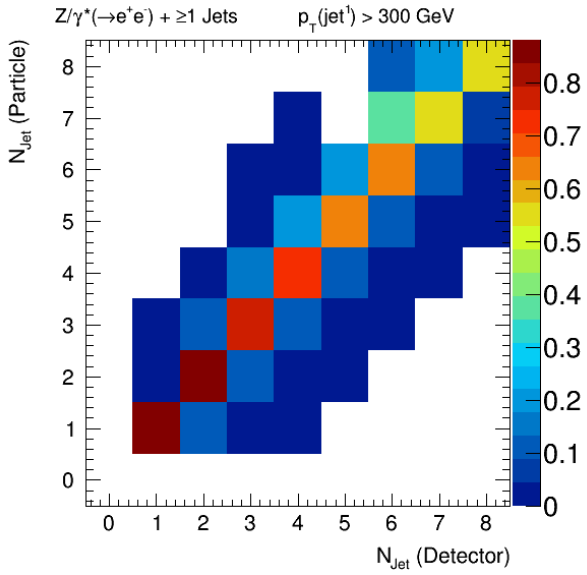


(b) Order Correction

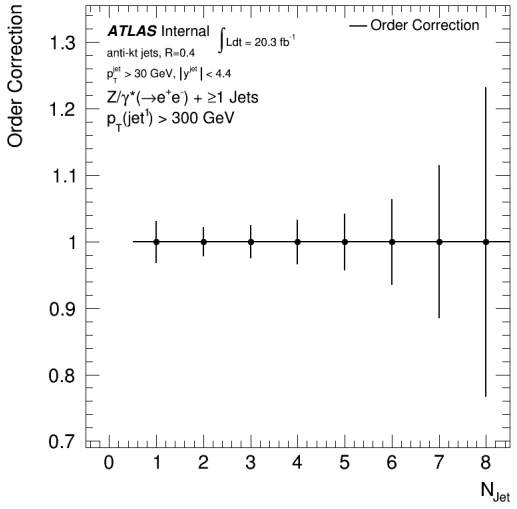


(c) Fake Correction

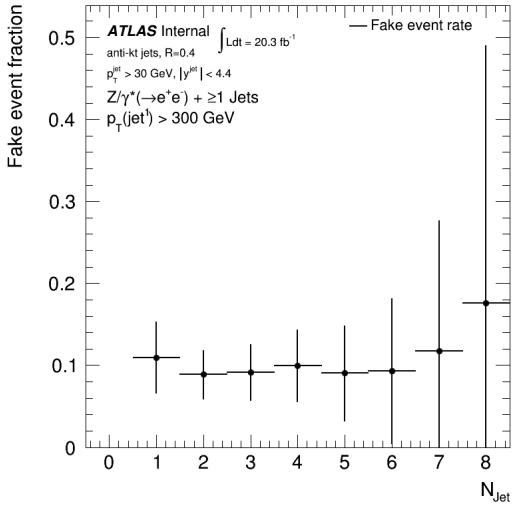
Figure B.1: Exclusive jet multiplicity (N_{jet}).



(a) Migration Matrix

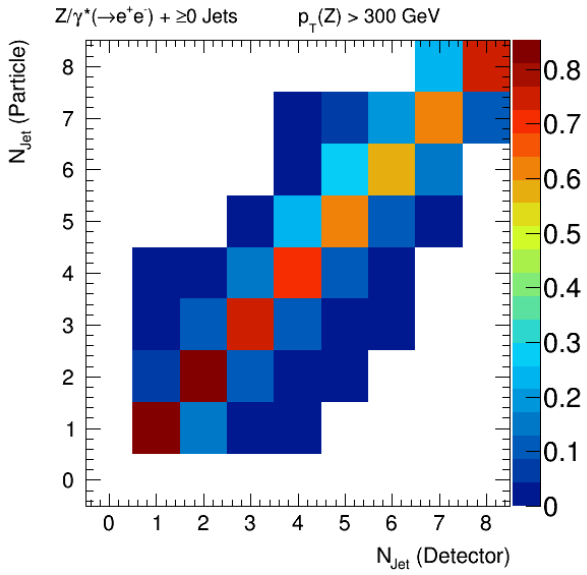


(b) Order Correction

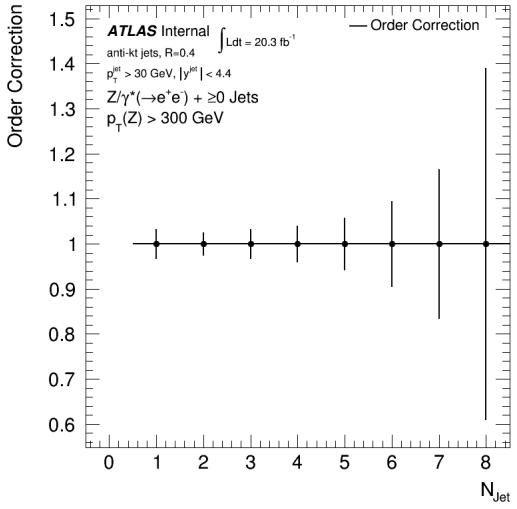


(c) Fake Correction

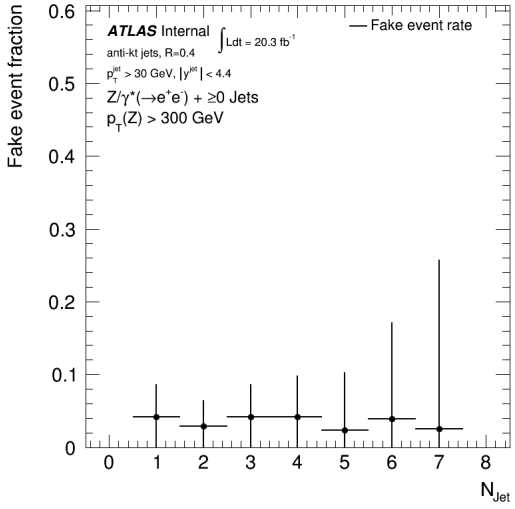
Figure B.2: Exclusive jet multiplicity (N_{jet}) for $p_T^{\text{leading jet}} > 300$ GeV.



(a) Migration Matrix

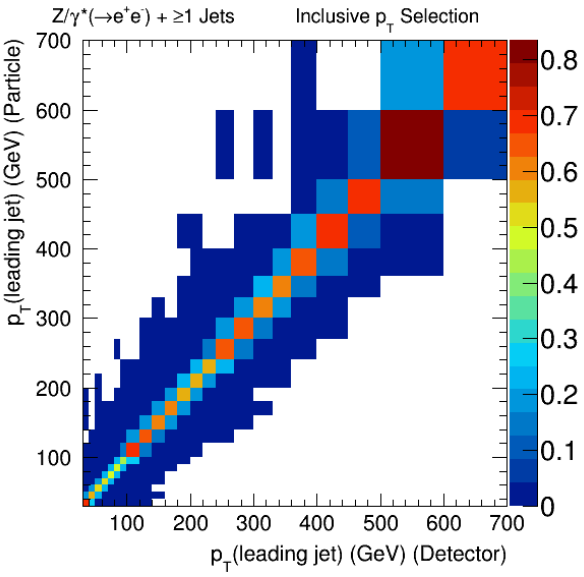


(b) Order Correction

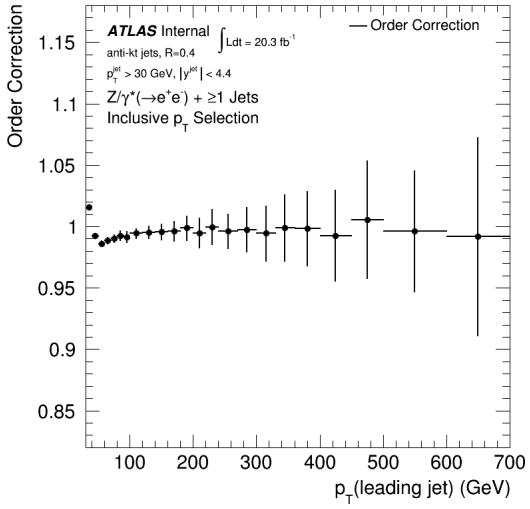


(c) Fake Correction

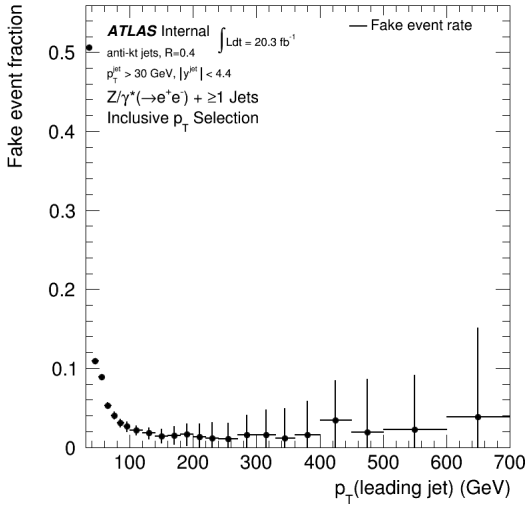
Figure B.3: Exclusive jet multiplicity (N_{jet}) for $p_T^Z > 300$ GeV.



(a) Migration Matrix

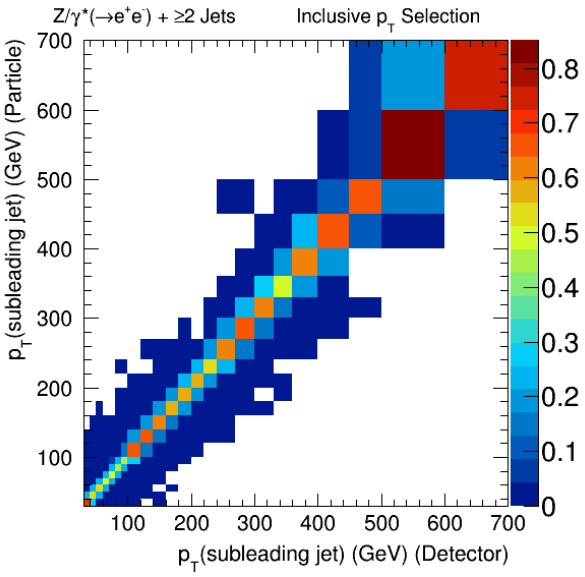


(b) Order Correction

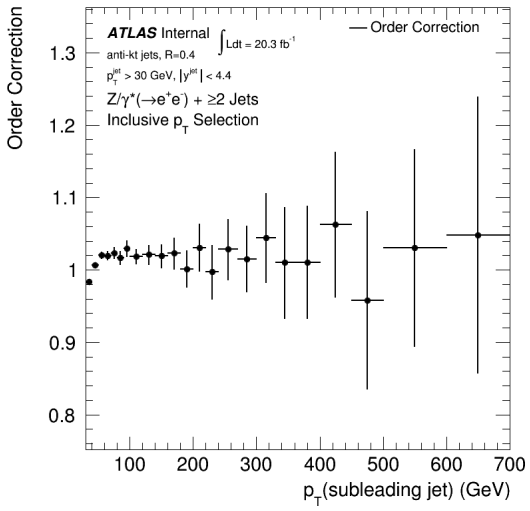


(c) Fake Correction

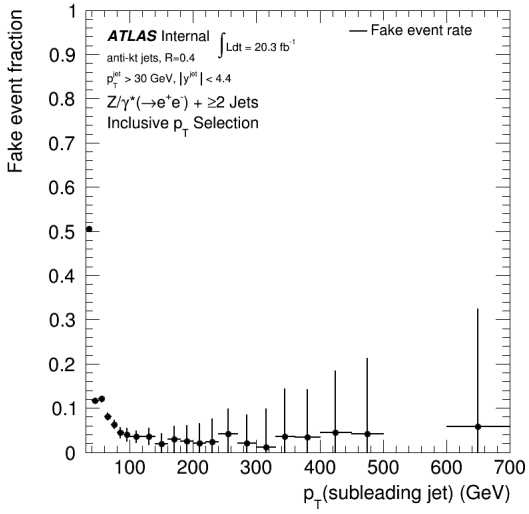
Figure B.4: Leading jet p_T in events with at least one jet.



(a) Migration Matrix

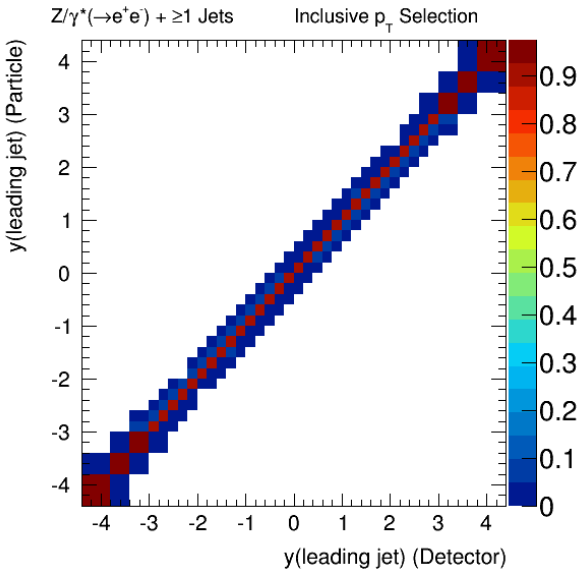


(b) Order Correction

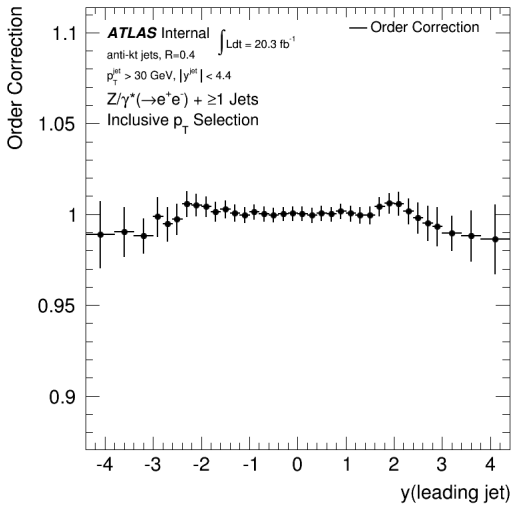


(c) Fake Correction

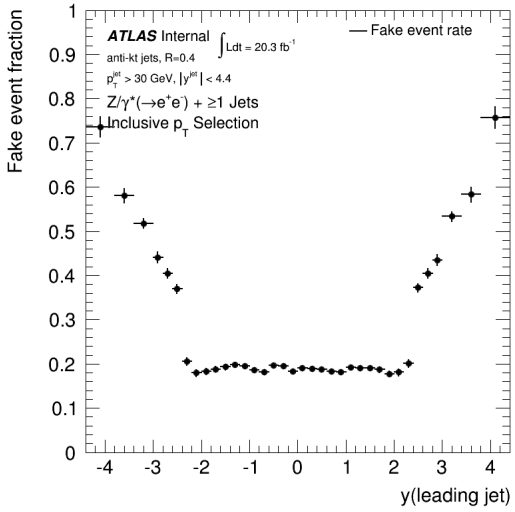
Figure B.5: Sub-leading jet p_T in events with at least two jets.



(a) Migration Matrix

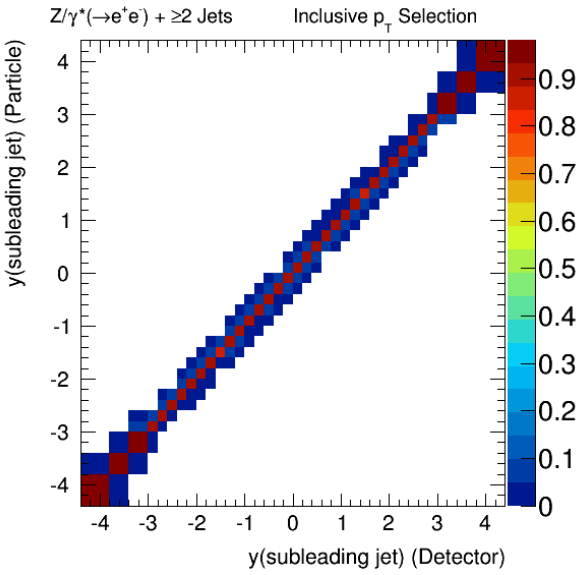


(b) Order Correction

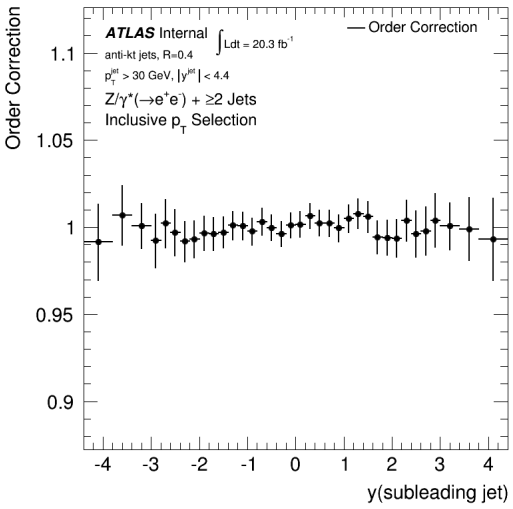


(c) Fake Correction

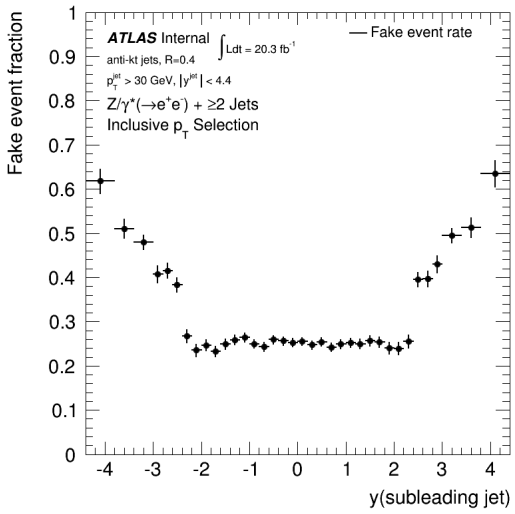
Figure B.6: Leading jet rapidity in events with at least one jet.



(a) Migration Matrix

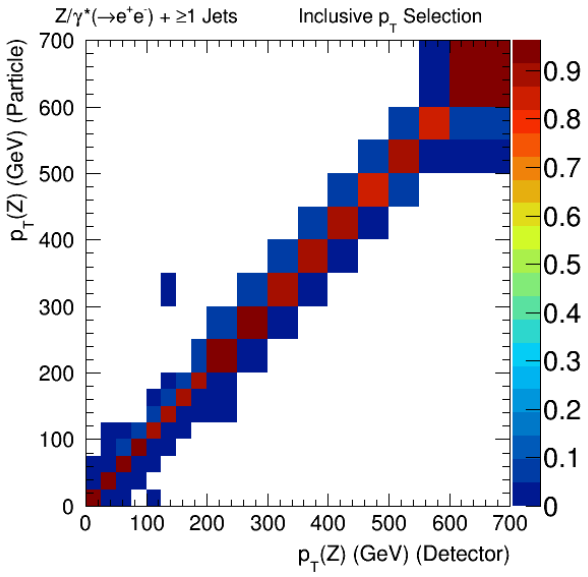


(b) Order Correction

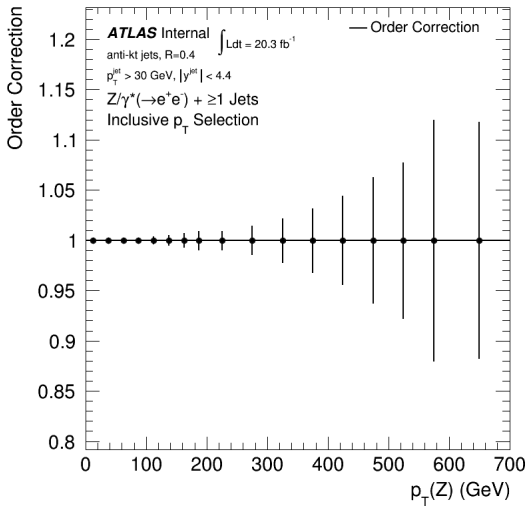


(c) Fake Correction

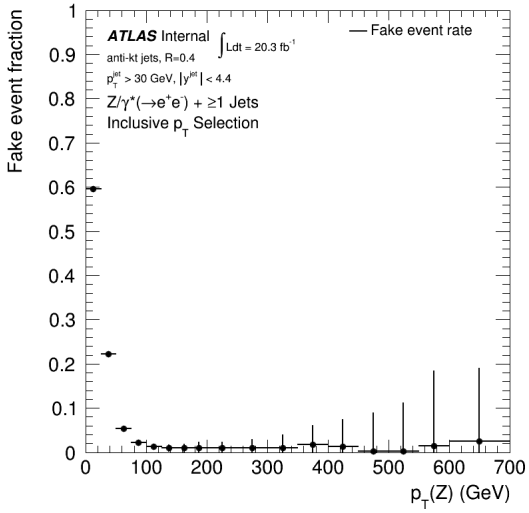
Figure B.7: Sub-leading jet rapdity in events with at least two jets.



(a) Migration Matrix

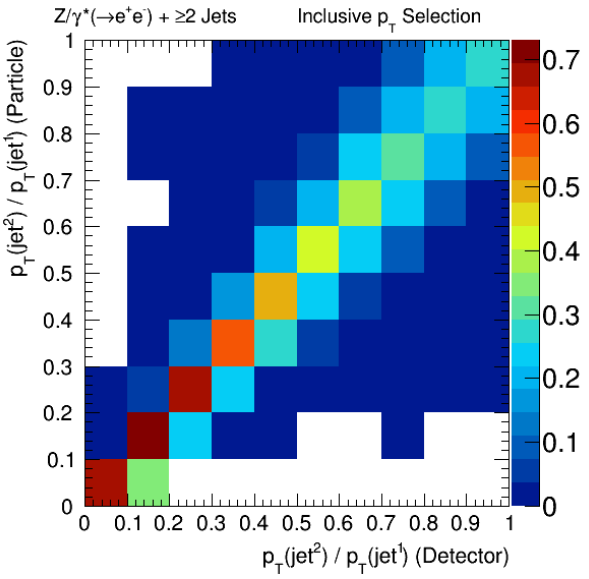


(b) Order Correction

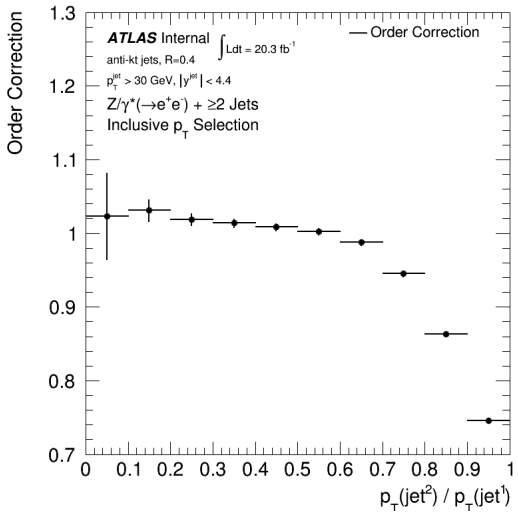


(c) Fake Correction

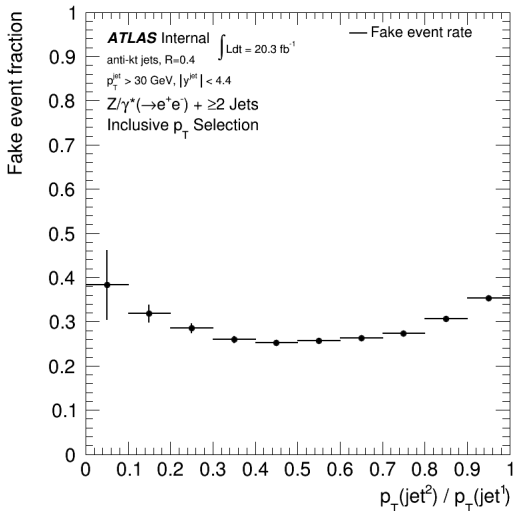
Figure B.8: Z boson p_T in events with at least one jet.



(a) Migration Matrix

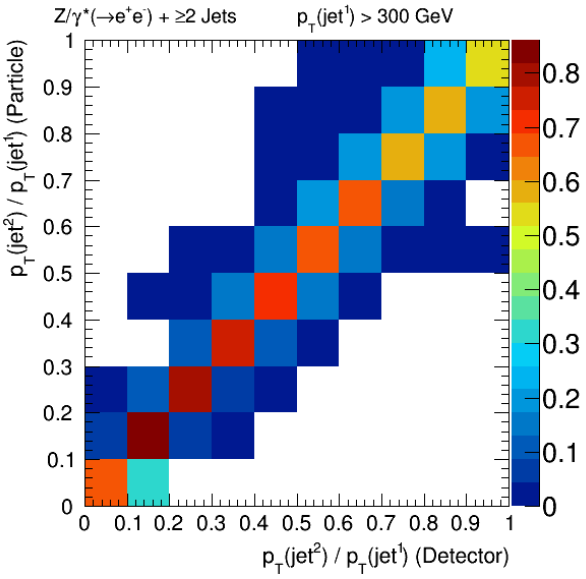


(b) Order Correction

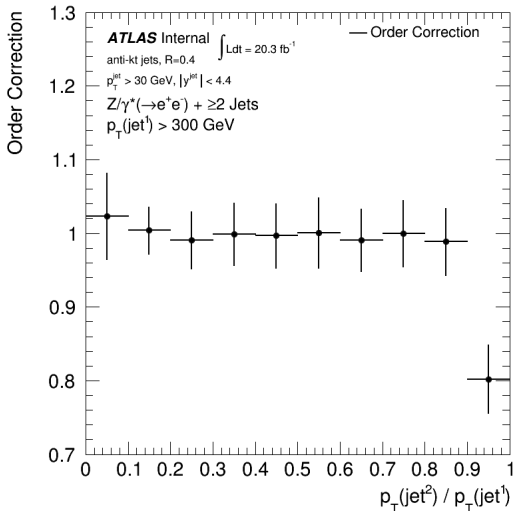


(c) Fake Correction

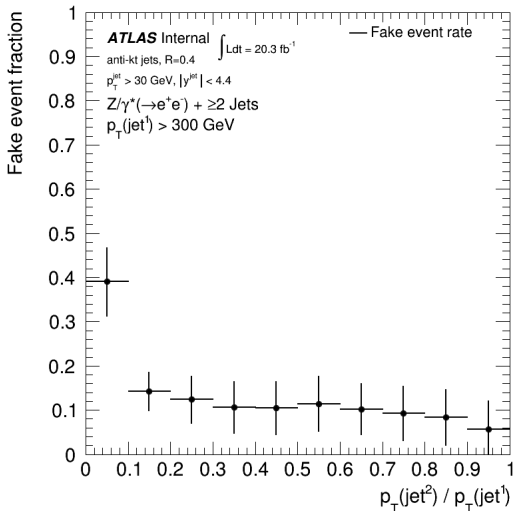
Figure B.9: p_T ratio between sub-leading jet and leading jet (Rp_T^{jj}) in events with at least two jets.



(a) Migration Matrix

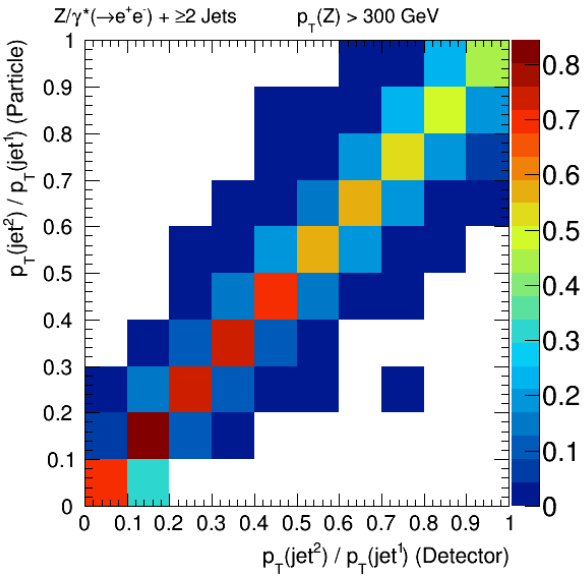


(b) Order Correction

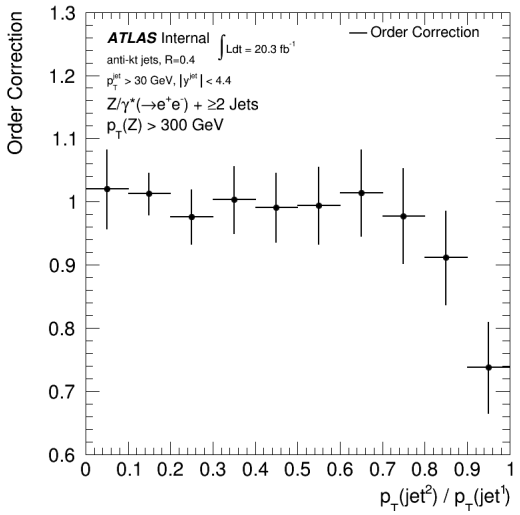


(c) Fake Correction

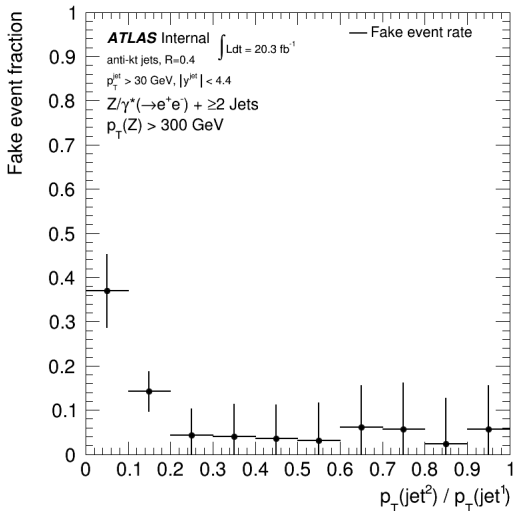
Figure B.10: p_T ratio between sub-leading jet and leading jet (Rp_T^{jj}) in events with at least two jets for $p_T^{\text{leading jet}} > 300$ GeV.



(a) Migration Matrix

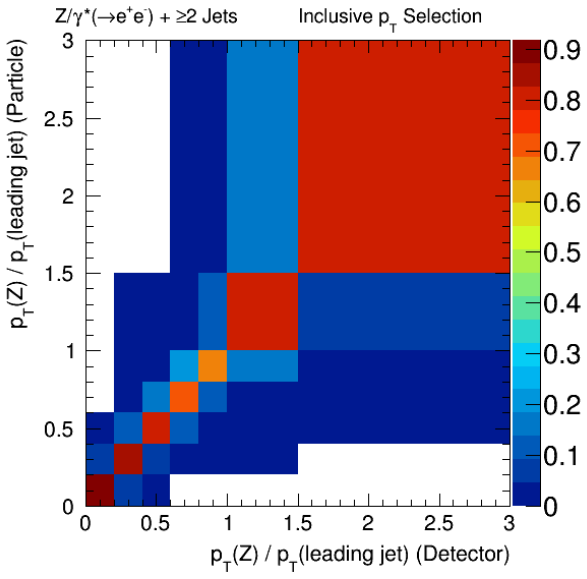


(b) Order Correction

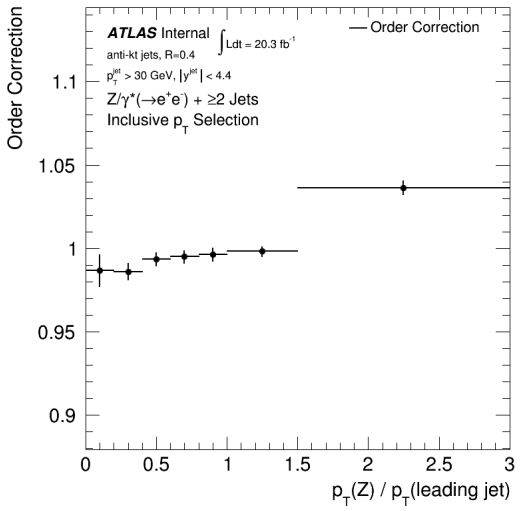


(c) Fake Correction

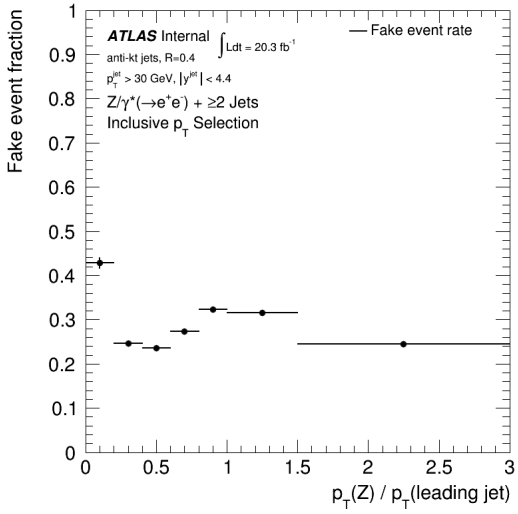
Figure B.11: p_T ratio between sub-leading jet and leading jet (Rp_T^{jj}) in events with at least two jets for $p_T^Z > 300$ GeV.



(a) Migration Matrix

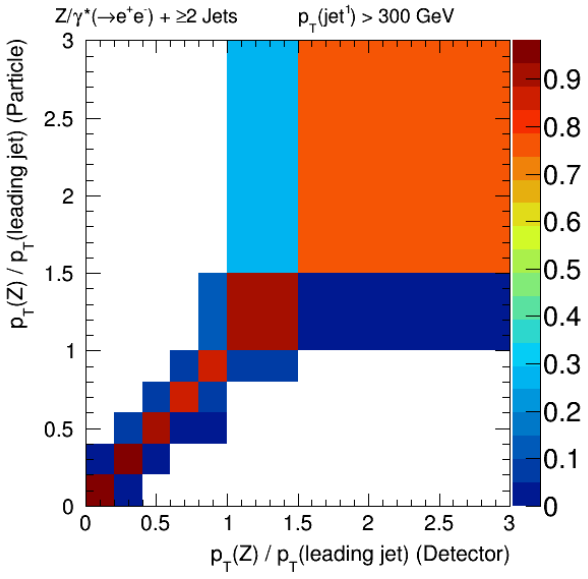


(b) Order Correction

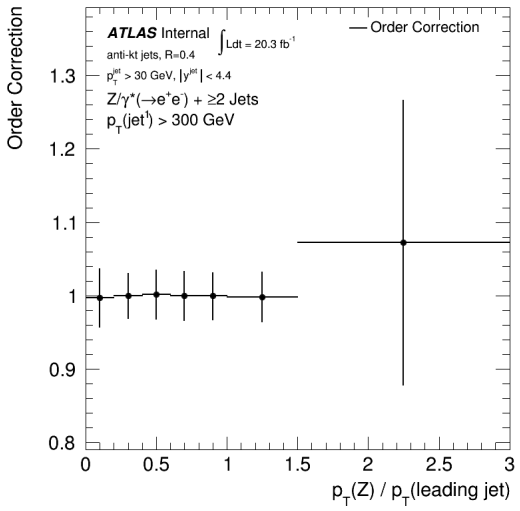


(c) Fake Correction

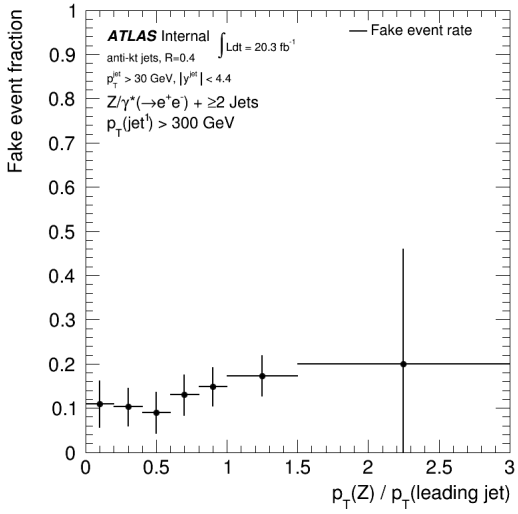
Figure B.12: p_T ratio between Z boson and leading jet (Rp_T^{jj}) in events with at least two jets.



(a) Migration Matrix

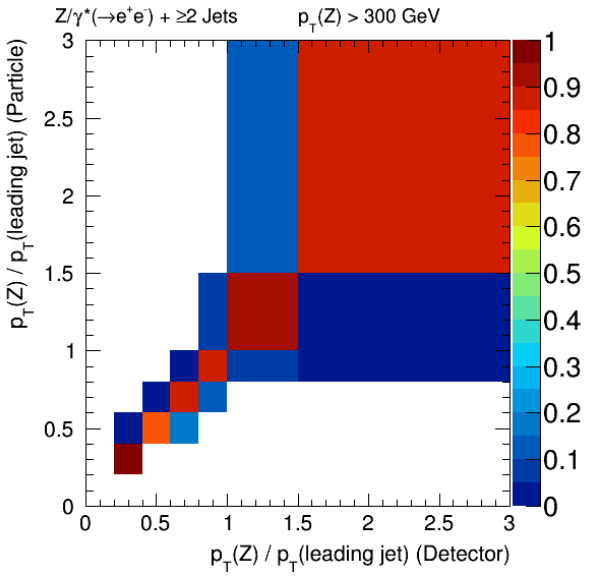


(b) Order Correction

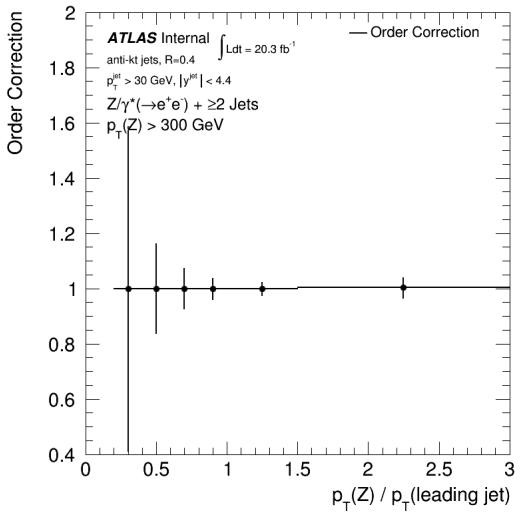


(c) Fake Correction

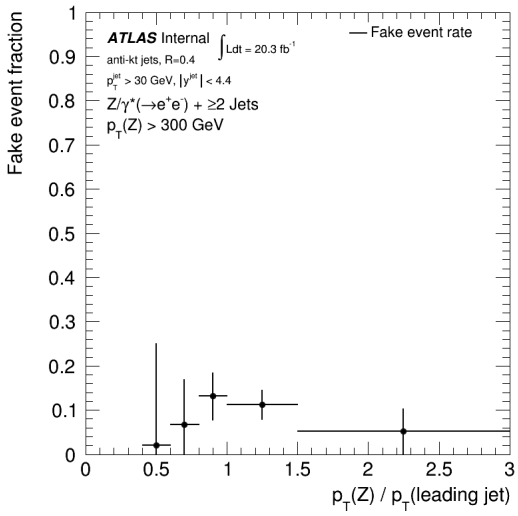
Figure B.13: p_T ratio between Z boson and leading jet (Rp_T^{jj}) in events with at least two jets for $p_T^{\text{leading jet}} > 300$ GeV.



(a) Migration Matrix

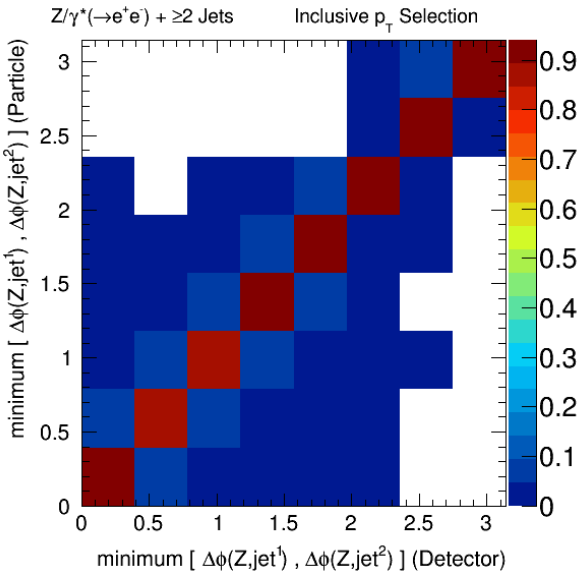


(b) Order Correction

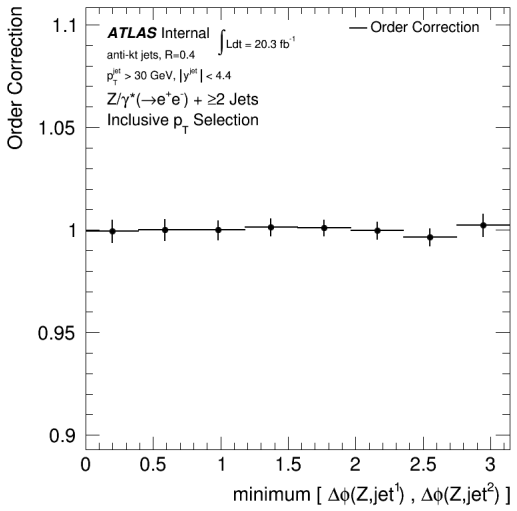


(c) Fake Correction

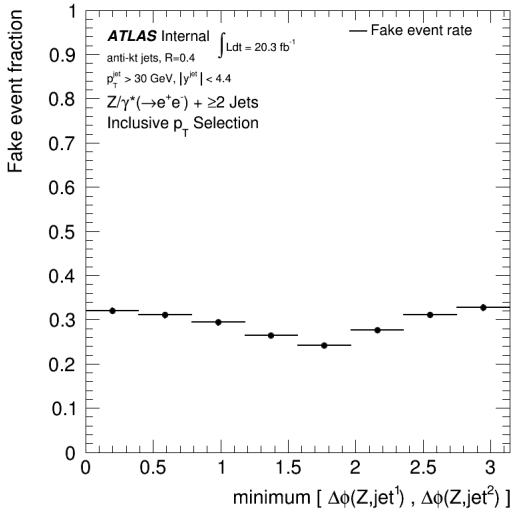
Figure B.14: p_T ratio between Z boson and leading jet (Rp_T^{jj}) in events with at least two jets for $p_T^Z > 300$ GeV.



(a) Migration Matrix

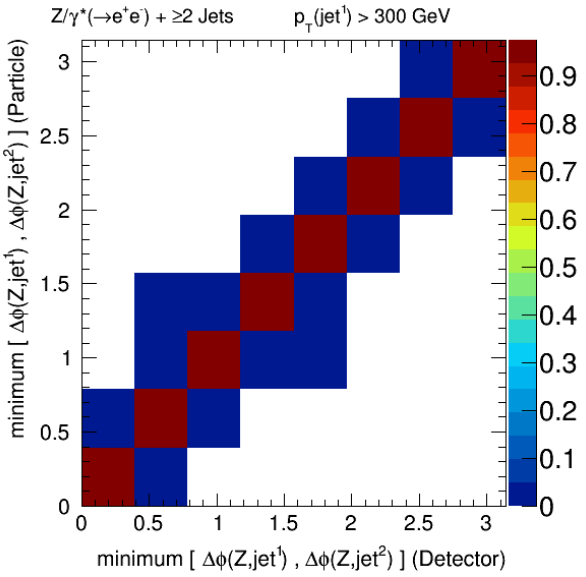


(b) Order Correction

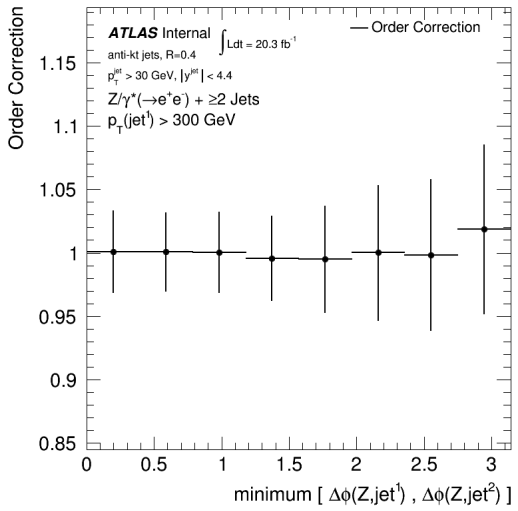


(c) Fake Correction

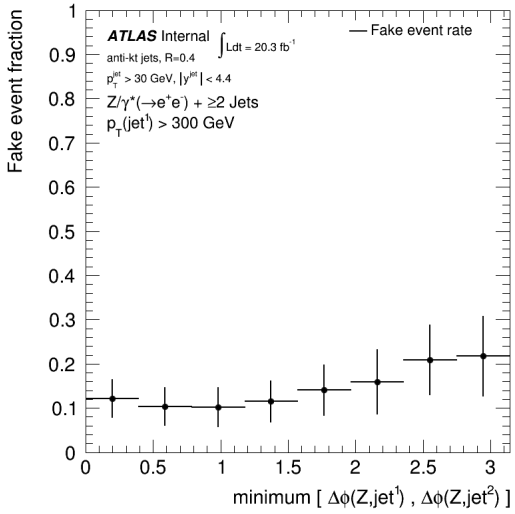
Figure B.15: Minimum $\Delta\phi$ between the Z boson and either of the two leading jets in the event ($min\Delta\phi^{Zjj}$).



(a) Migration Matrix

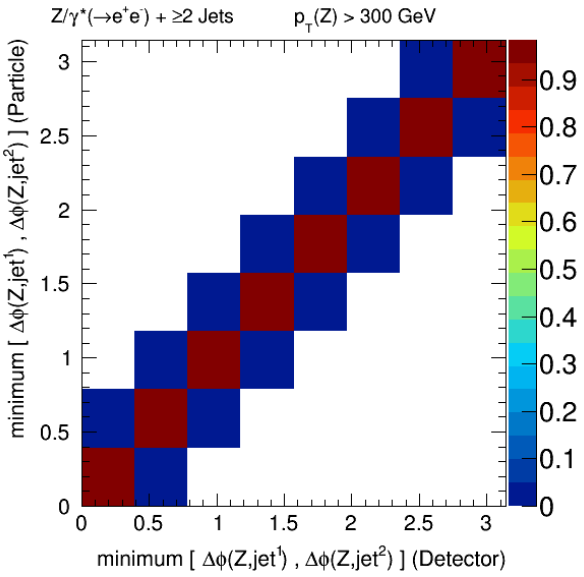


(b) Order Correction

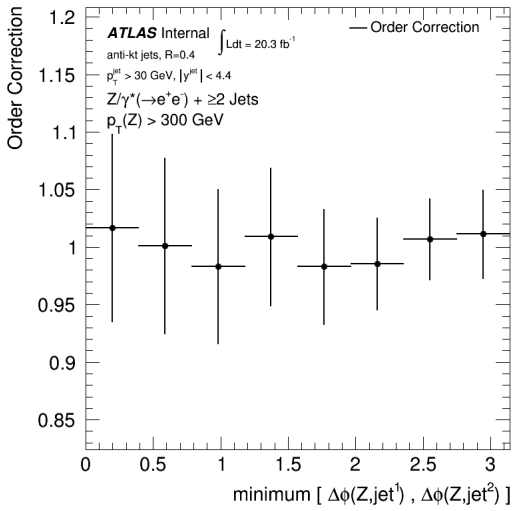


(c) Fake Correction

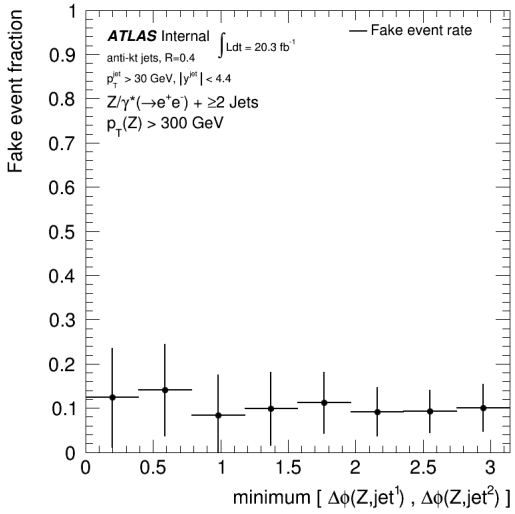
Figure B.16: Minimum $\Delta\phi$ between the Z boson and either of the two leading jets in the event ($min\Delta\phi^{Zjj}$) for $p_T^{\text{leading jet}} > 300$ GeV.



(a) Migration Matrix

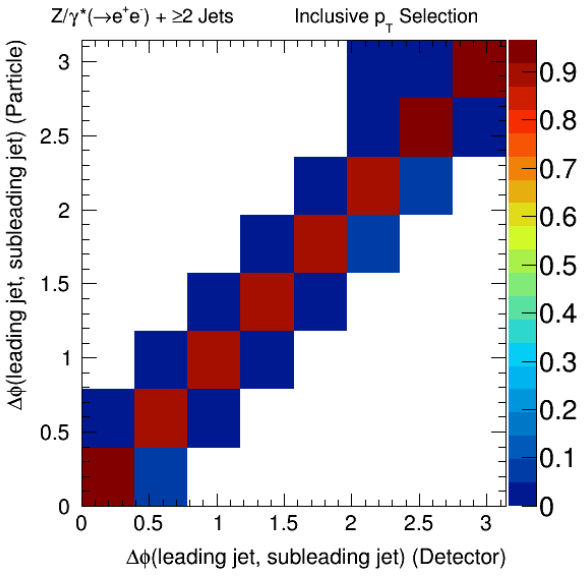


(b) Order Correction

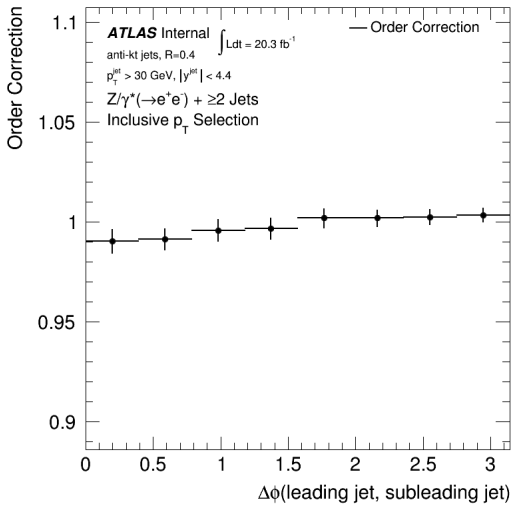


(c) Fake Correction

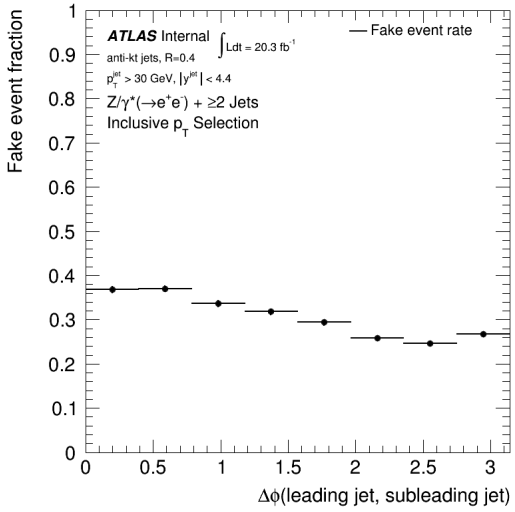
Figure B.17: Minimum $\Delta\phi$ between the Z boson and either of the two leading jets in the event ($min\Delta\phi^{Zjj}$) for $p_T^Z > 300$ GeV.



(a) Migration Matrix

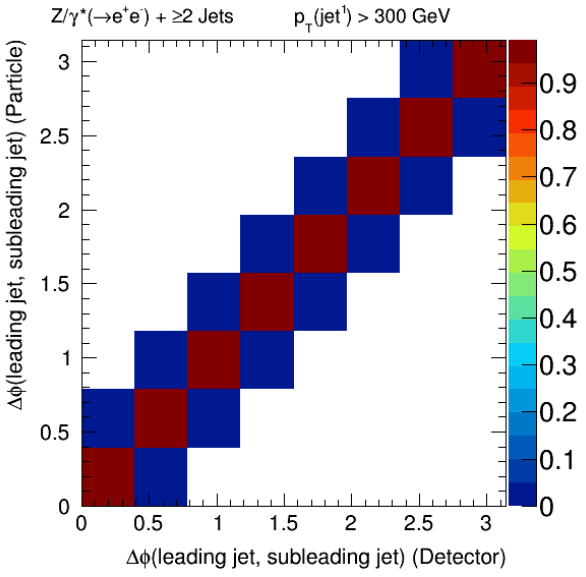


(b) Order Correction

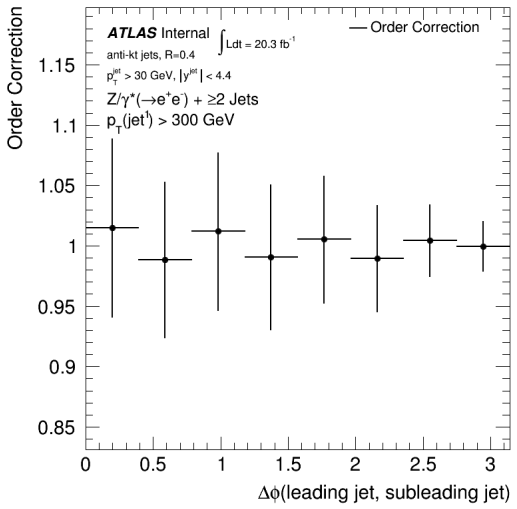


(c) Fake Correction

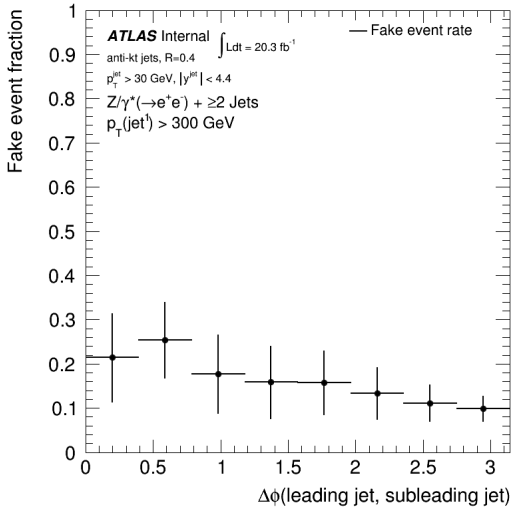
Figure B.18: Absolute azimuthal separation between the two leading jets ($\Delta\phi^{jj}$) in events with at least two jets.



(a) Migration Matrix

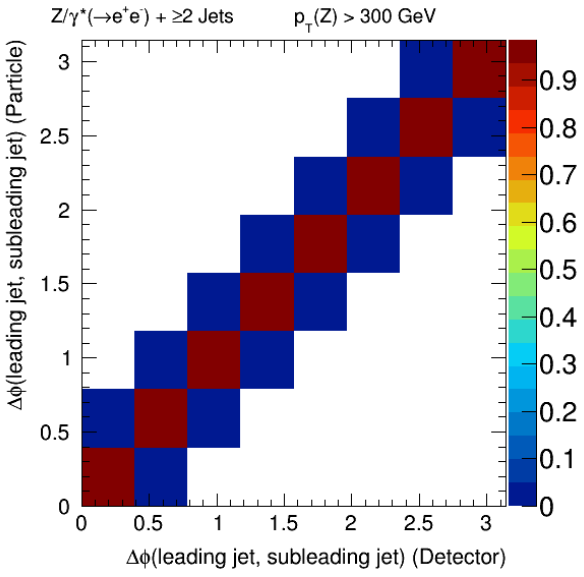


(b) Order Correction

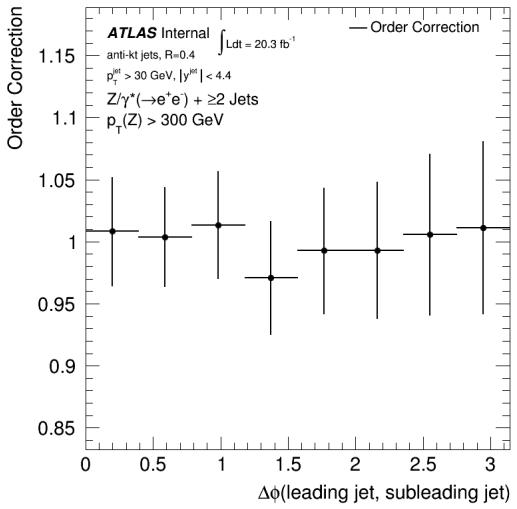


(c) Fake Correction

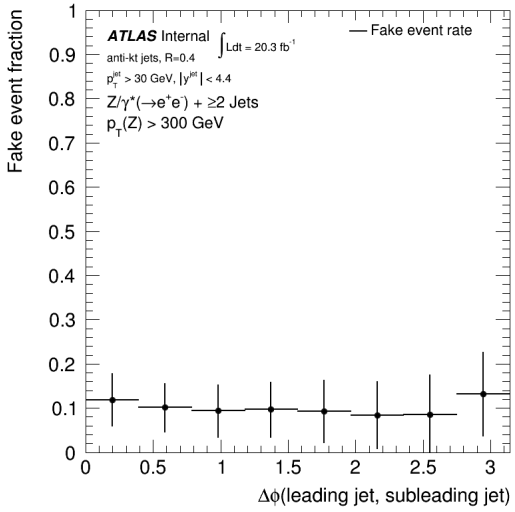
Figure B.19: Absolute azimuthal separation between the two leading jets ($\Delta\phi^{jj}$) in events with at least two jets for $p_T^{\text{leading jet}} > 300$ GeV.



(a) Migration Matrix



(b) Order Correction

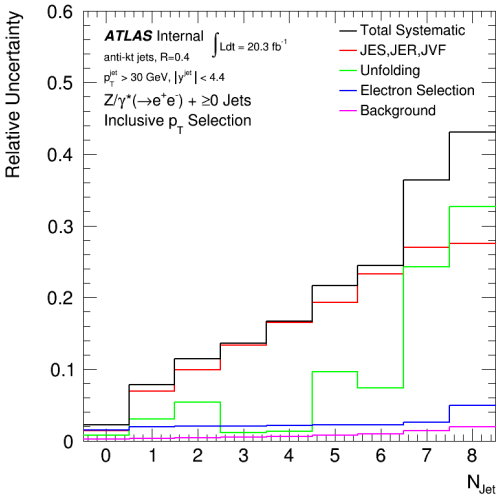


(c) Fake Correction

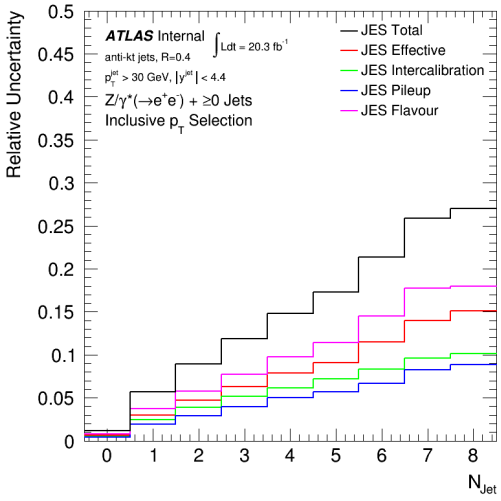
Figure B.20: Absolute azimuthal separation between the two leading jets ($\Delta\phi^{jj}$) in events with at least two jets for $p_T^Z > 300$ GeV.

APPENDIX C

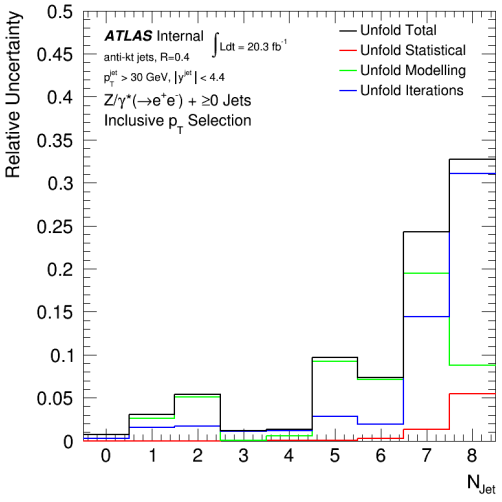
Systematic Uncertainties



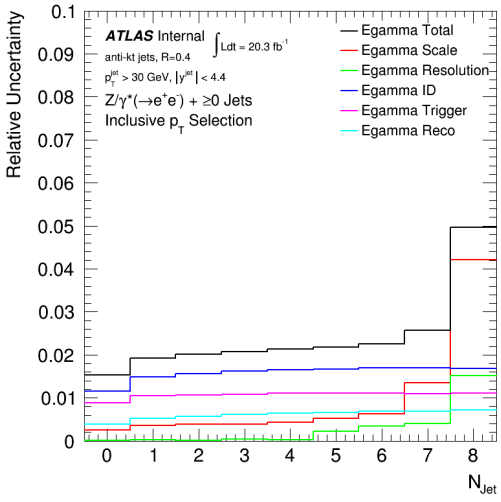
(a) Total Systematic Uncertainty



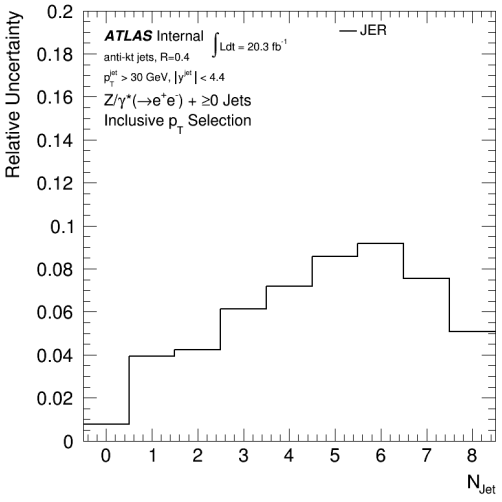
(b) JES Uncertainty



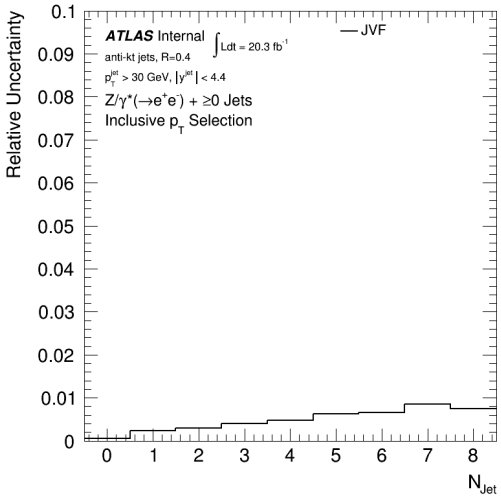
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

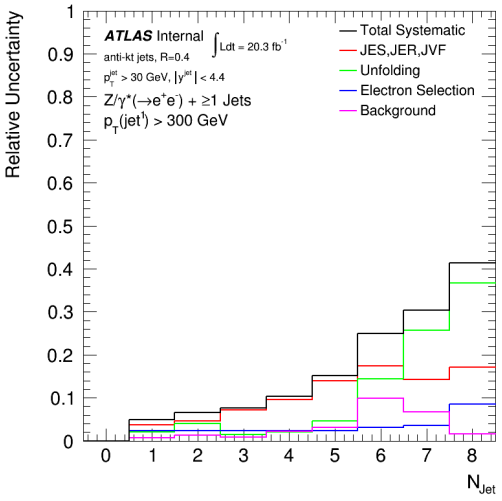


(e) JER Uncertainty

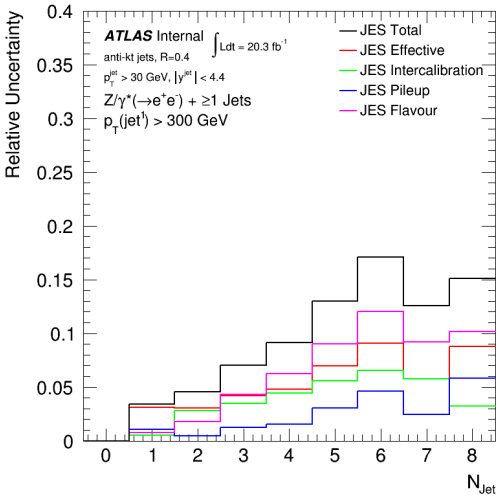


(f) Pile-up Jet Removal Uncertainty

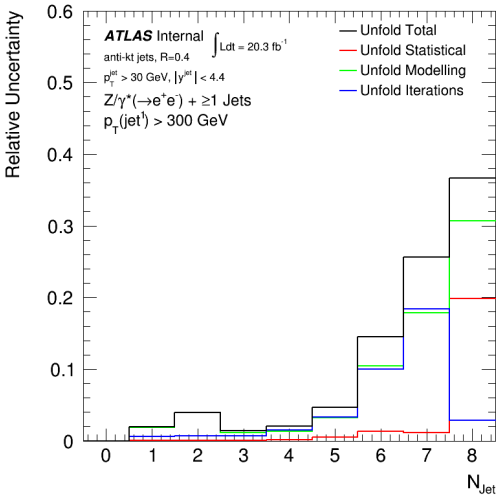
Figure C.1: Exclusive jet multiplicity (N_{jet}).



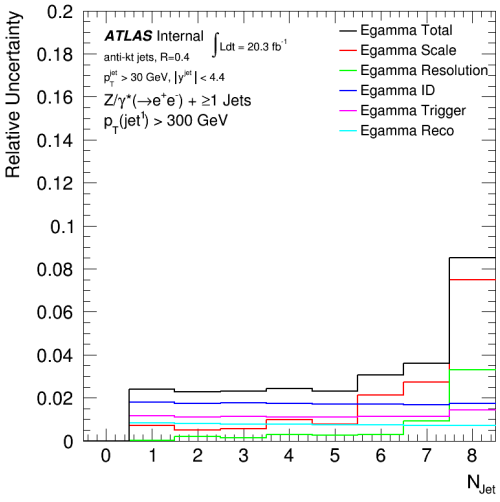
(a) Total Systematic Uncertainty



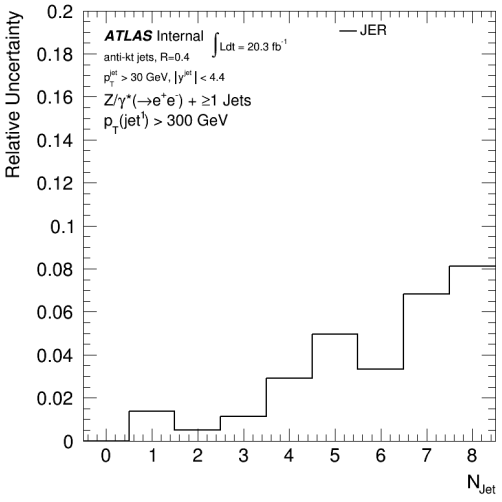
(b) JES Uncertainty



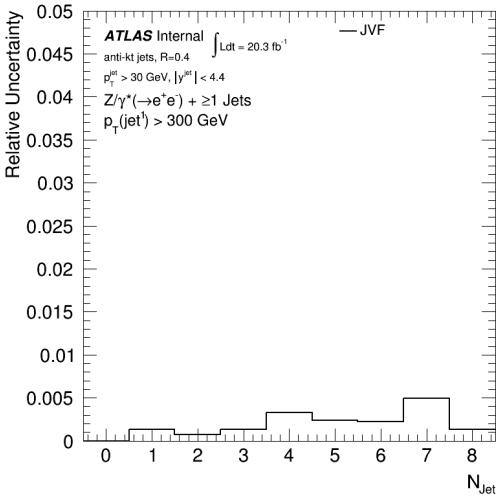
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

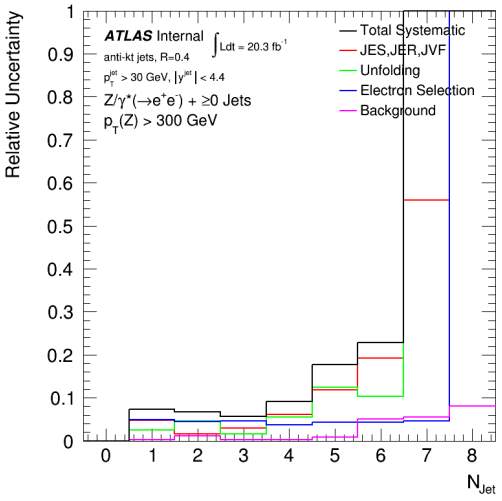


(e) JER Uncertainty

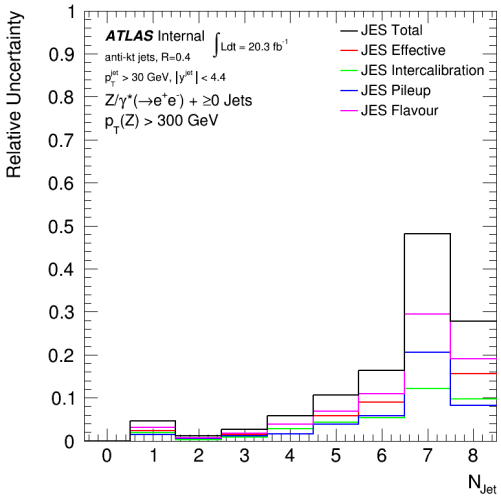


(f) Pile-up Jet Removal Uncertainty

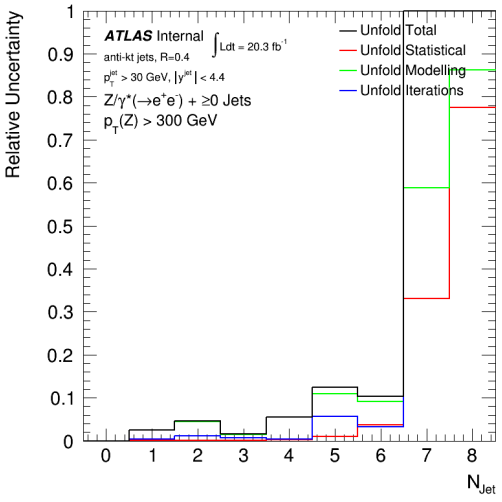
Figure C.2: Exclusive jet multiplicity (N_{jet}) for $p_T^{\text{leading jet}} > 300$ GeV.



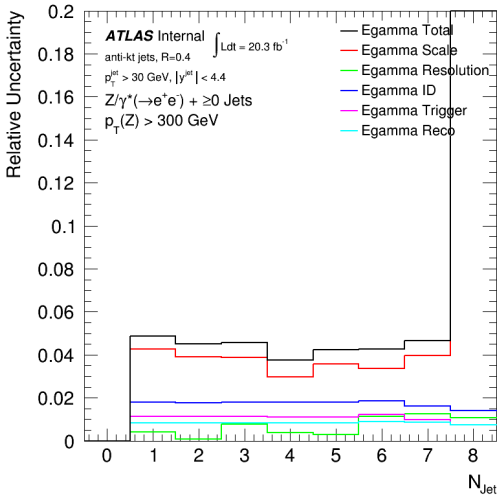
(a) Total Systematic Uncertainty



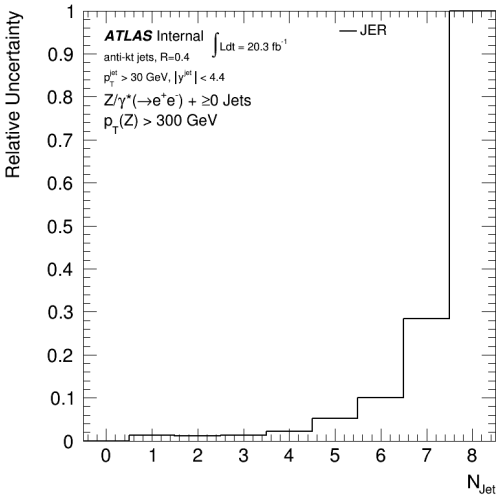
(b) JES Uncertainty



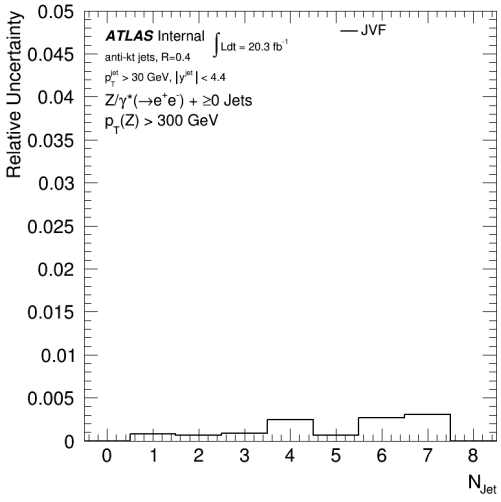
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

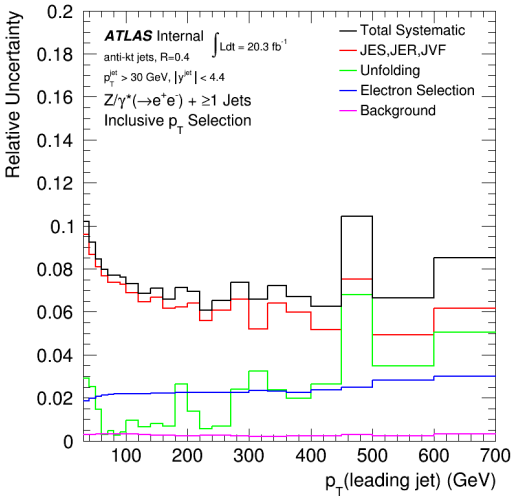


(e) JER Uncertainty

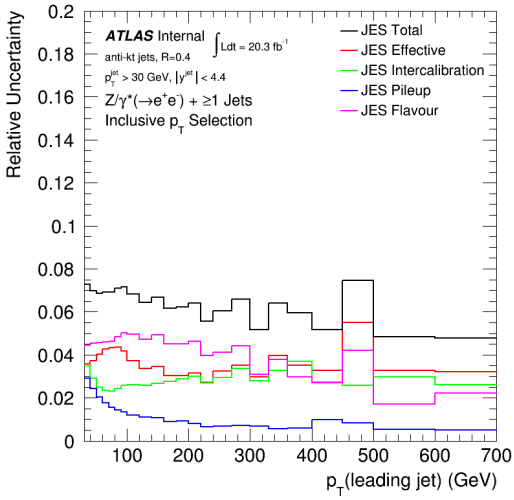


(f) Pile-up Jet Removal Uncertainty

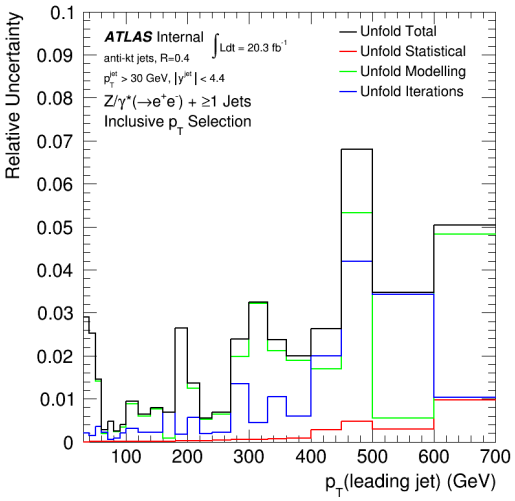
Figure C.3: Exclusive jet multiplicity (N_{jet}) for $p_T^Z > 300$ GeV.



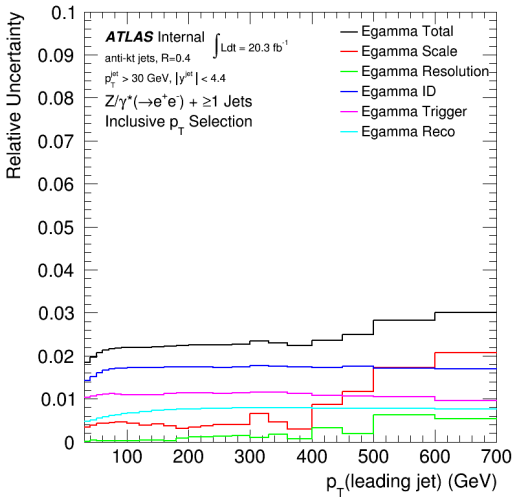
(a) Total Systematic Uncertainty



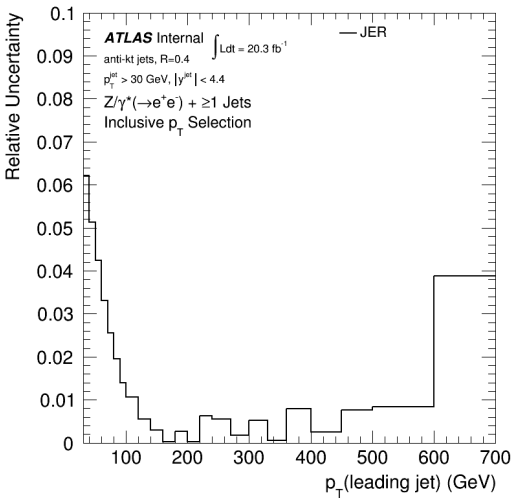
(b) JES Uncertainty



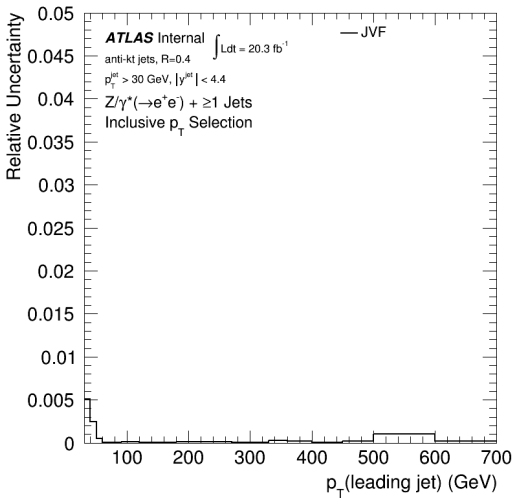
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

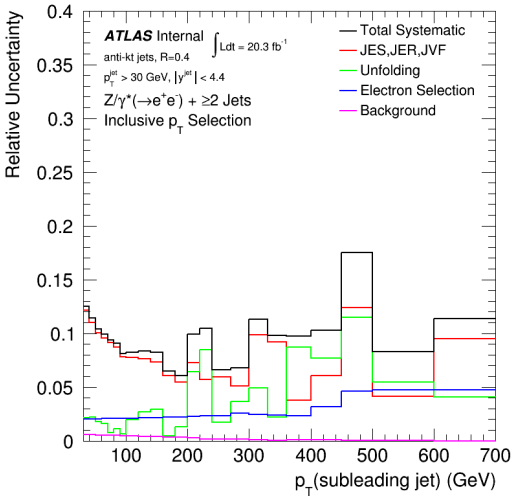


(e) JER Uncertainty

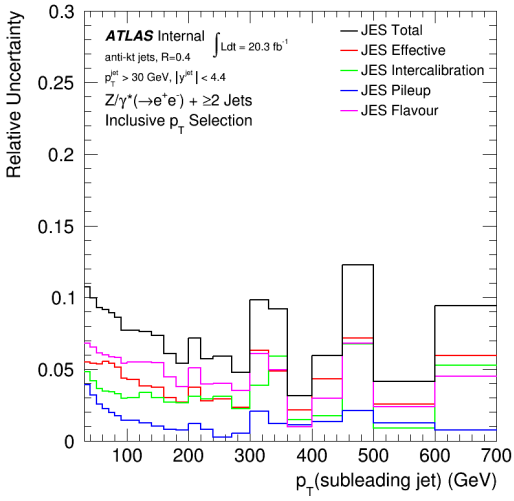


(f) Pile-up Jet Removal Uncertainty

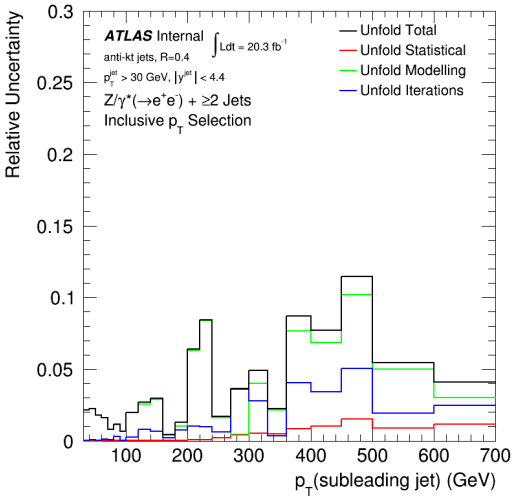
Figure C.4: Leading jet p_T.



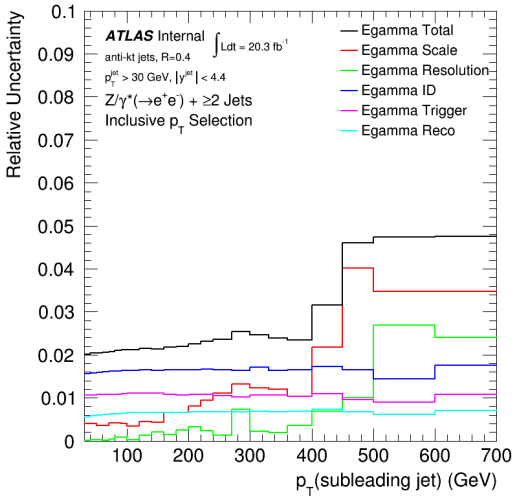
(a) Total Systematic Uncertainty



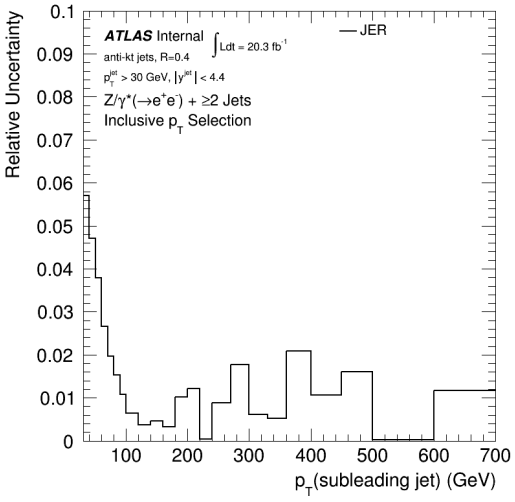
(b) JES Uncertainty



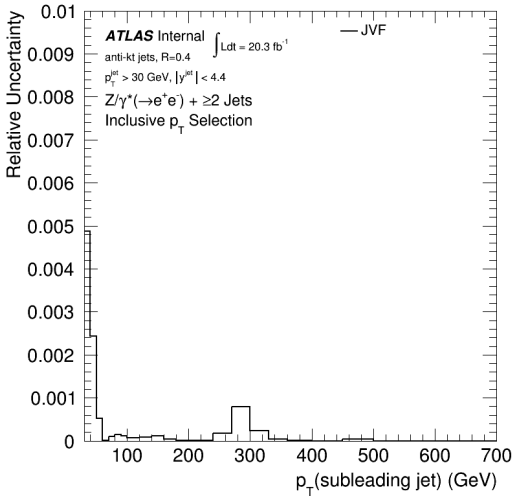
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

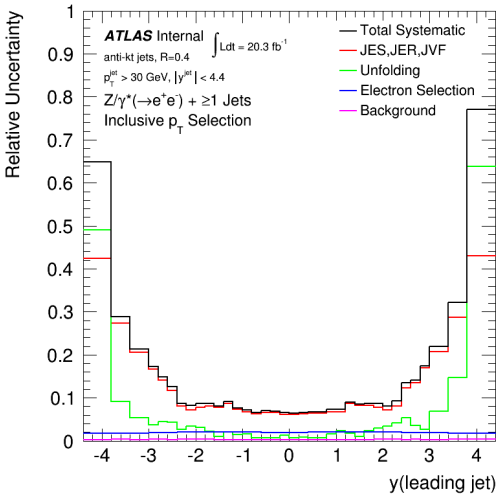


(e) JER Uncertainty

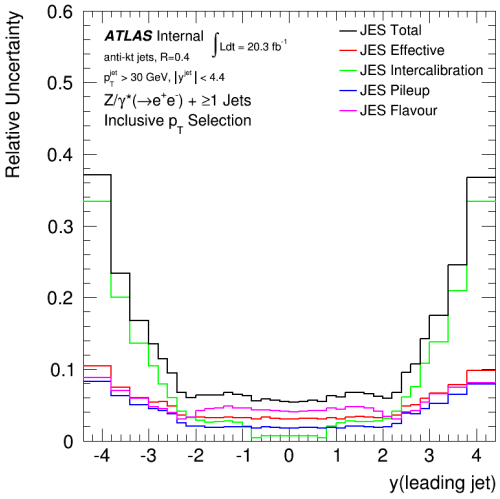


(f) Pile-up Jet Removal Uncertainty

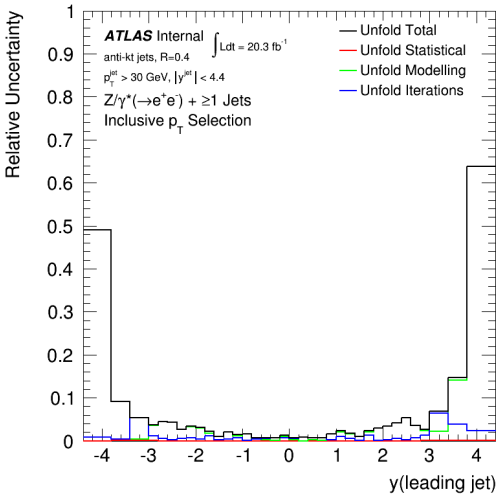
Figure C.5: Sub-leading jet p_T.



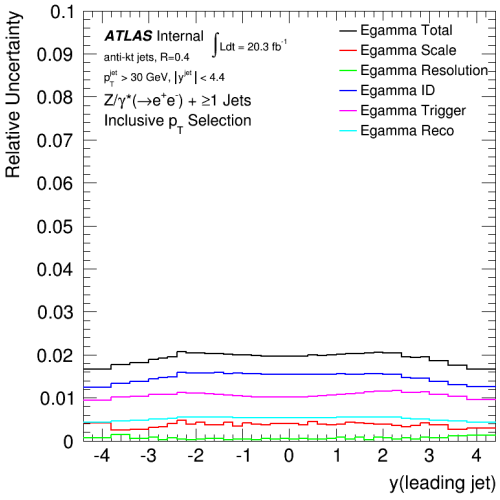
(a) Total Systematic Uncertainty



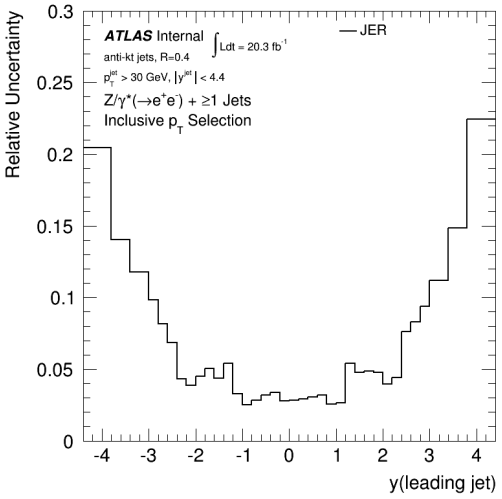
(b) JES Uncertainty



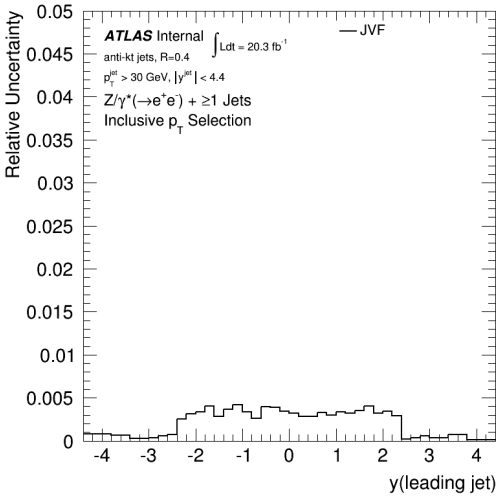
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

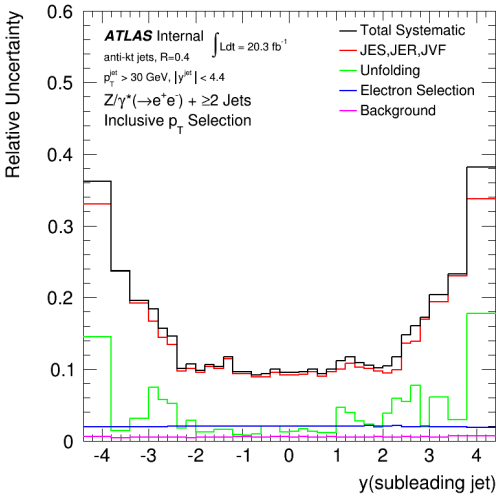


(e) JER Uncertainty

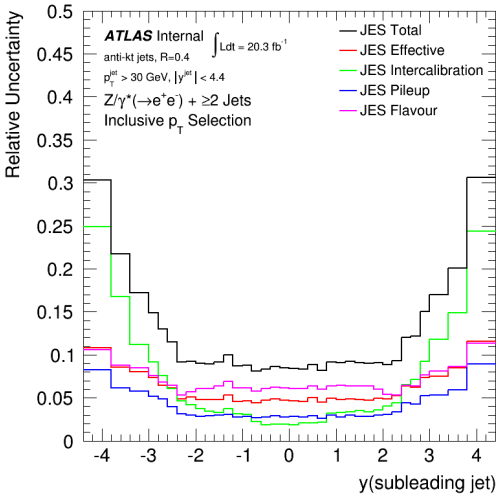


(f) Pile-up Jet Removal Uncertainty

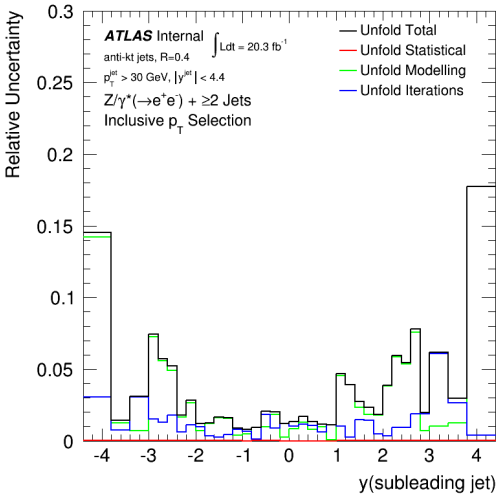
Figure C.6: Leading jet rapidity.



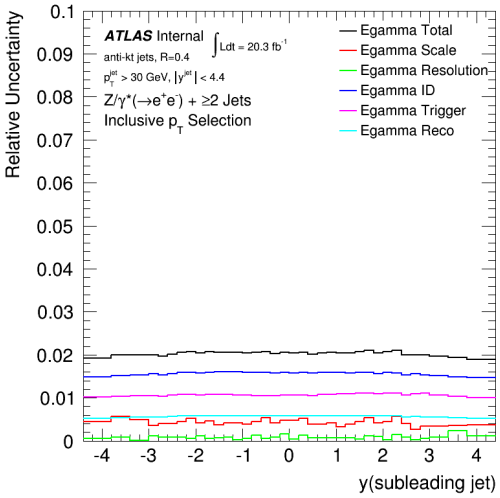
(a) Total Systematic Uncertainty



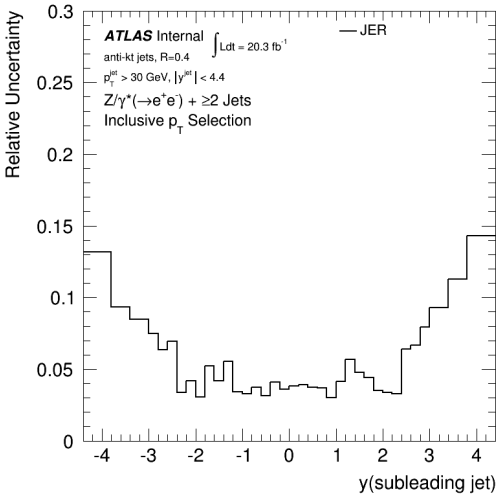
(b) JES Uncertainty



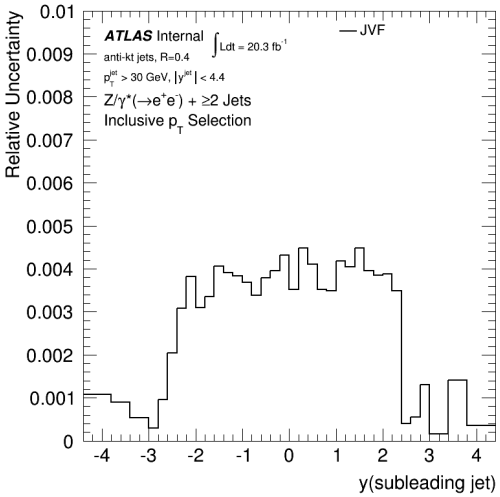
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

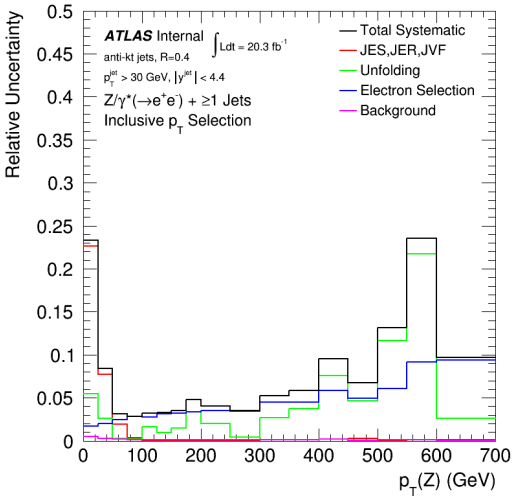


(e) JER Uncertainty

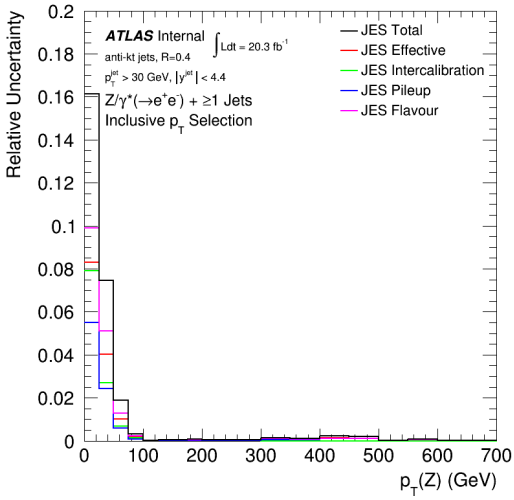


(f) Pile-up Jet Removal Uncertainty

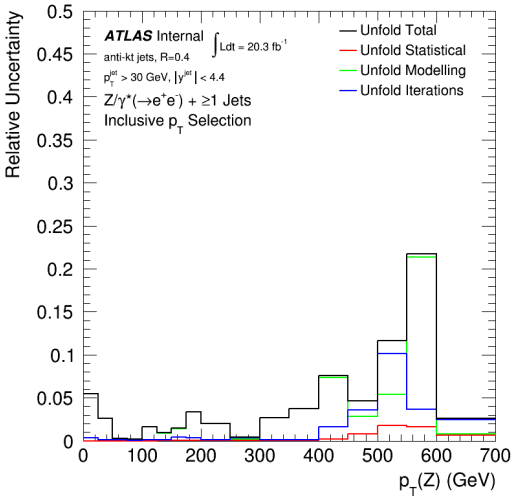
Figure C.7: Sub-leading jet rapdity.



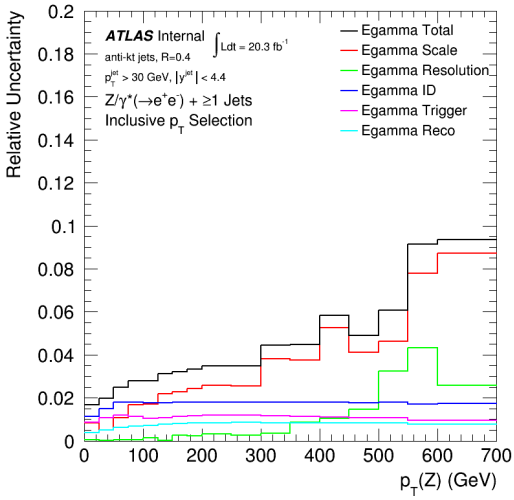
(a) Total Systematic Uncertainty



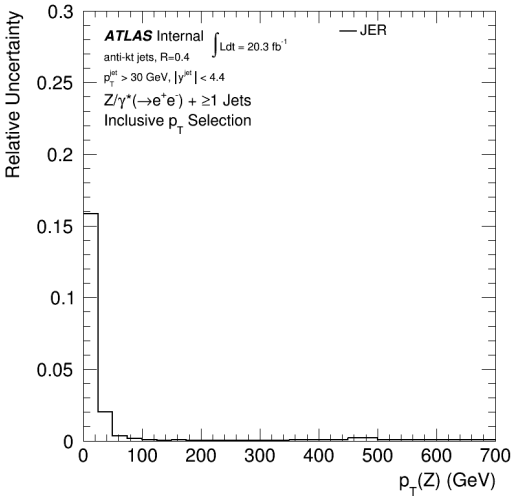
(b) JES Uncertainty



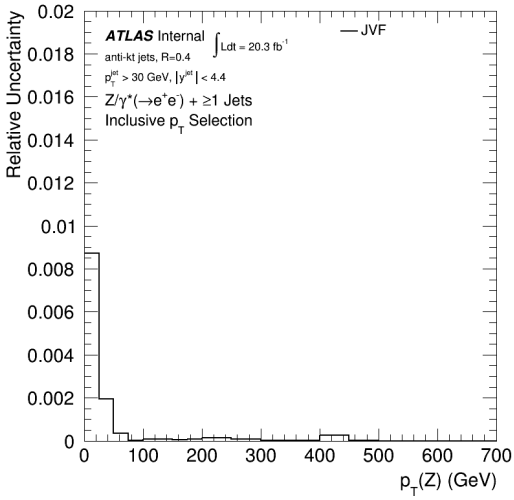
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

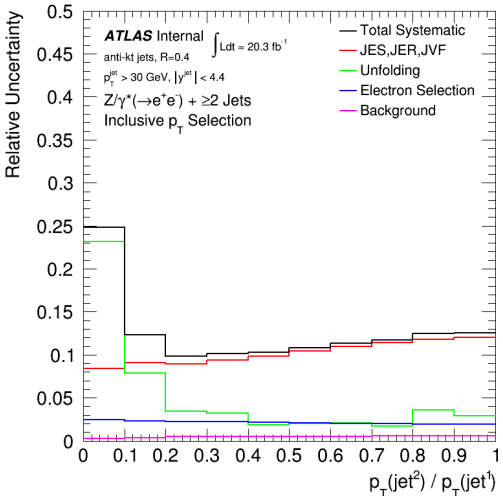


(e) JER Uncertainty

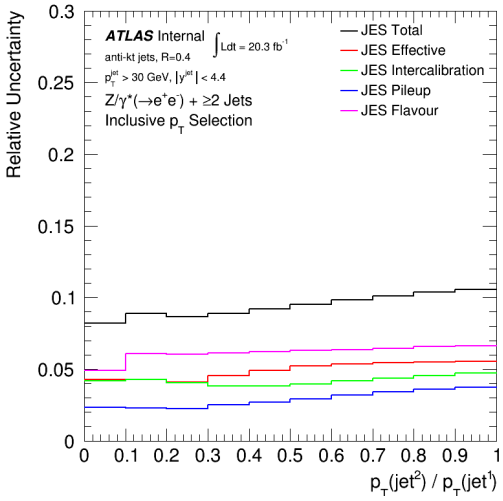


(f) Pile-up Jet Removal Uncertainty

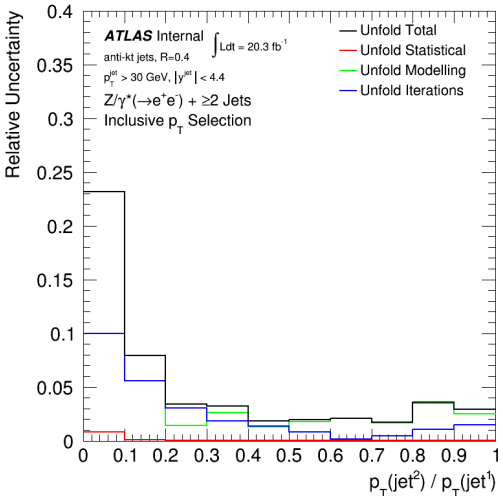
Figure C.8: Z boson p_T in events with at least one jet.



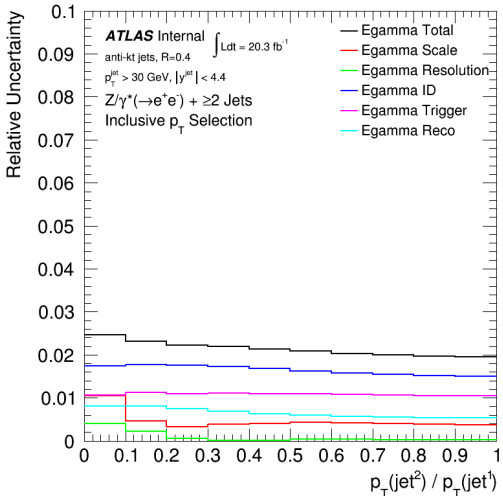
(a) Total Systematic Uncertainty



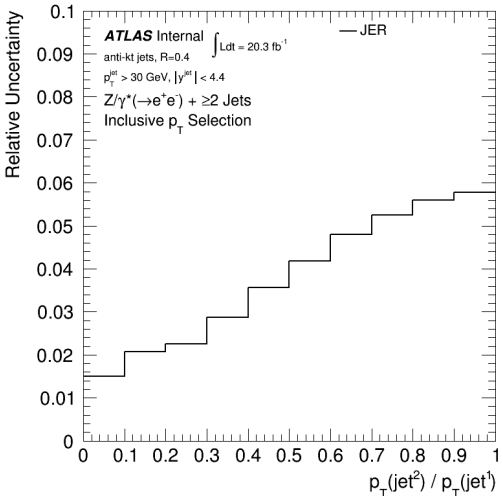
(b) JES Uncertainty



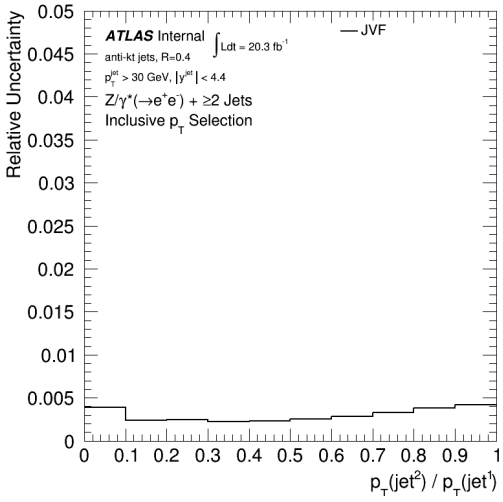
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

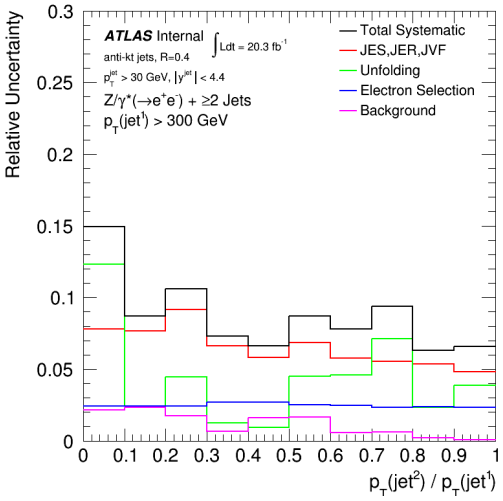


(e) JER Uncertainty

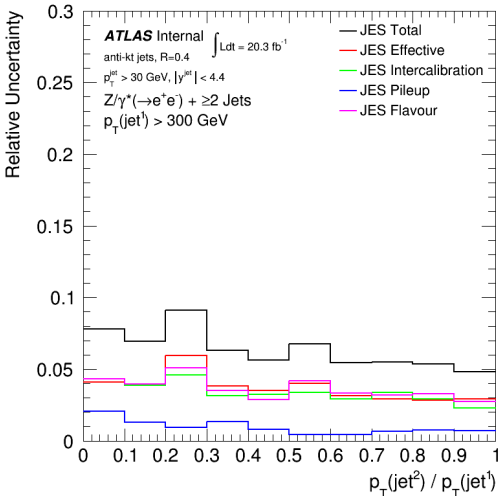


(f) Pile-up Jet Removal Uncertainty

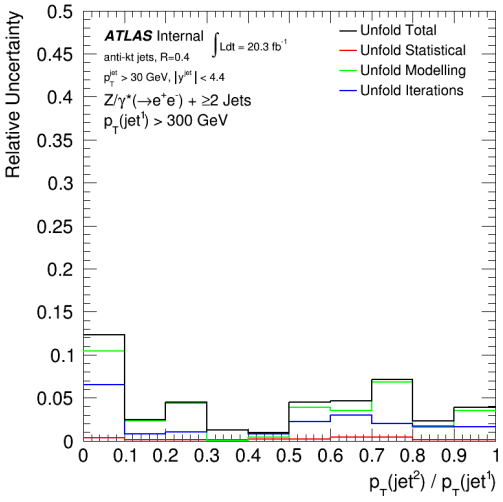
Figure C.9: p_T ratio between sub-leading jet and leading jet (Rp_T^{jj}) in events with at least two jets.



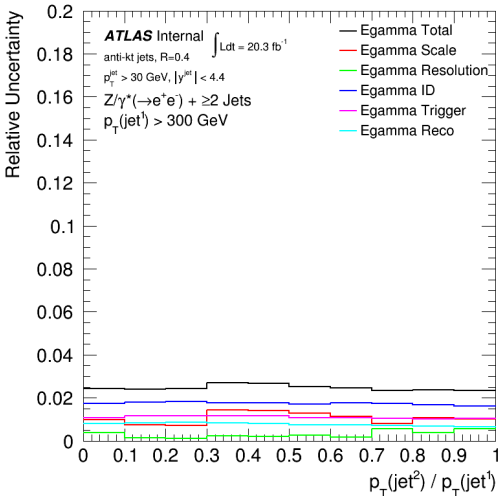
(a) Total Systematic Uncertainty



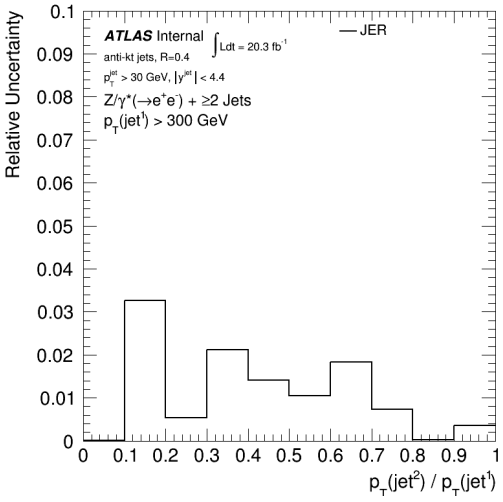
(b) JES Uncertainty



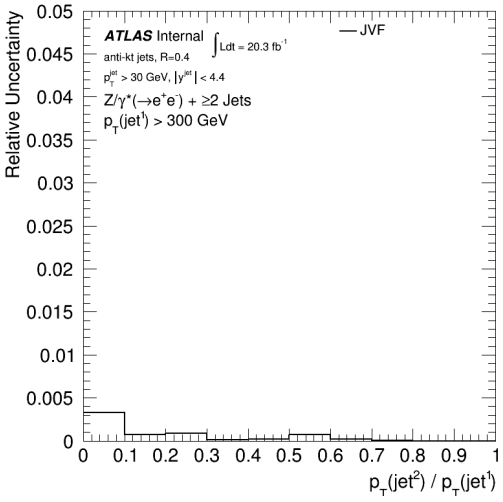
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

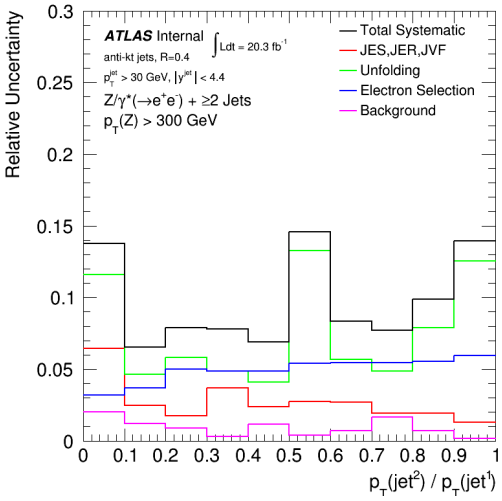


(e) JER Uncertainty

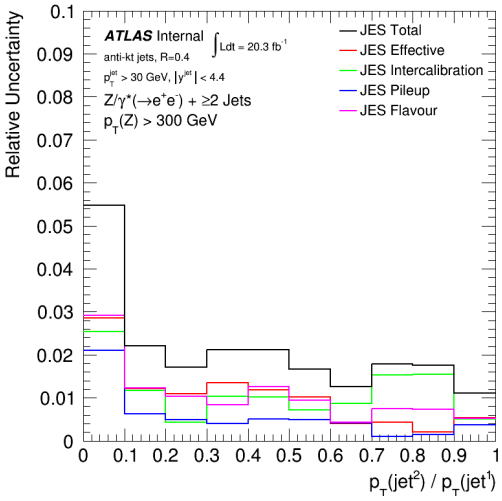


(f) Pile-up Jet Removal Uncertainty

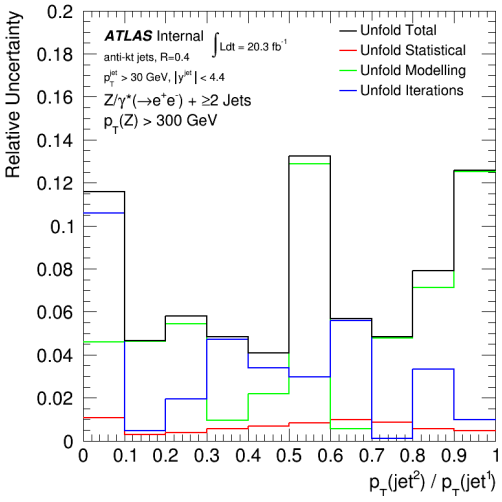
Figure C.10: p_T ratio between sub-leading jet and leading jet (Rp_T^{jj}) in events with at least two jets for $p_T^{\text{leading jet}} > 300$ GeV.



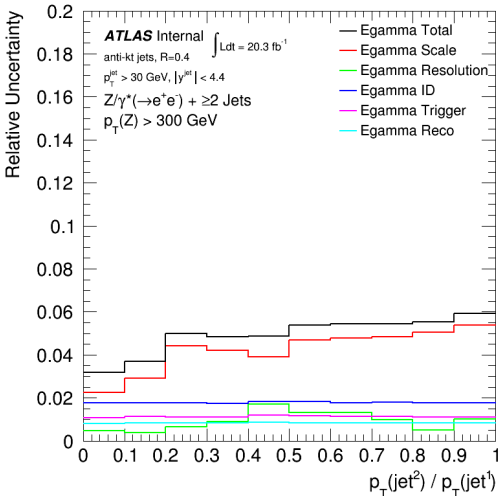
(a) Total Systematic Uncertainty



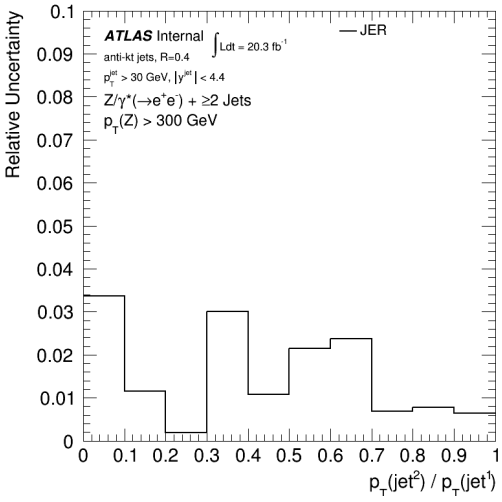
(b) JES Uncertainty



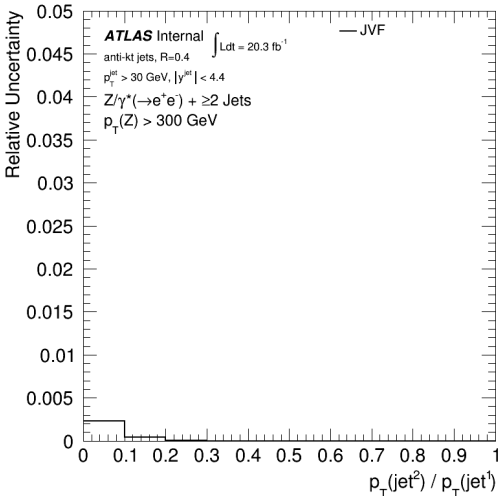
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

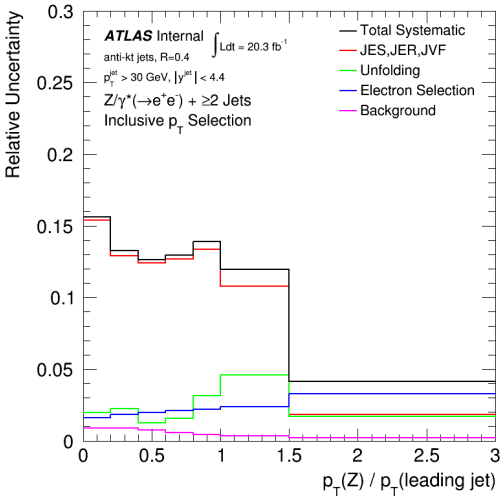


(e) JER Uncertainty

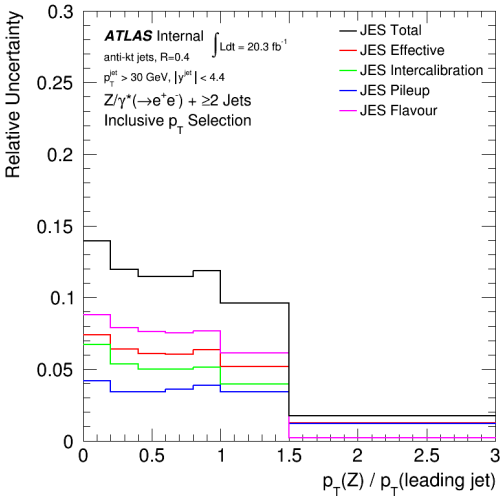


(f) Pile-up Jet Removal Uncertainty

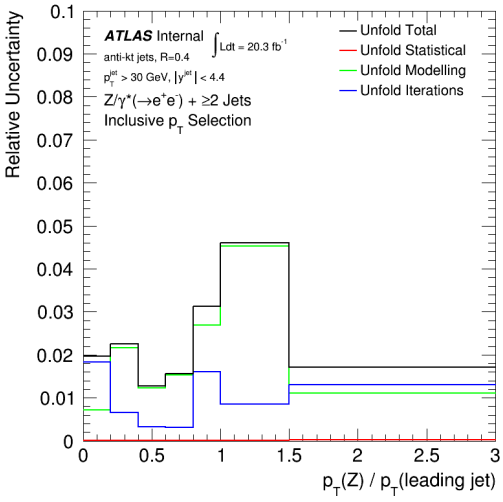
Figure C.11: p_T ratio between sub-leading jet and leading jet (Rp_T^{jj}) in events with at least two jets for $p_T^Z > 300$ GeV.



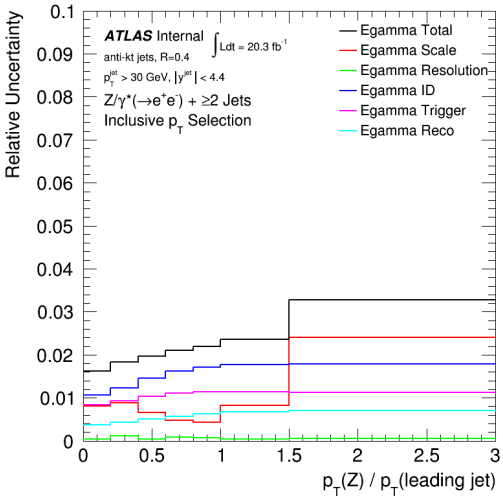
(a) Total Systematic Uncertainty



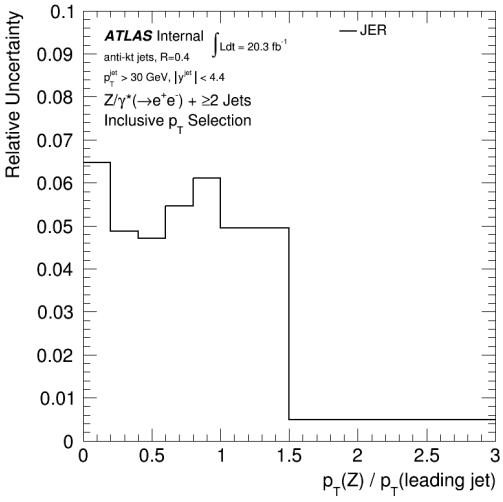
(b) JES Uncertainty



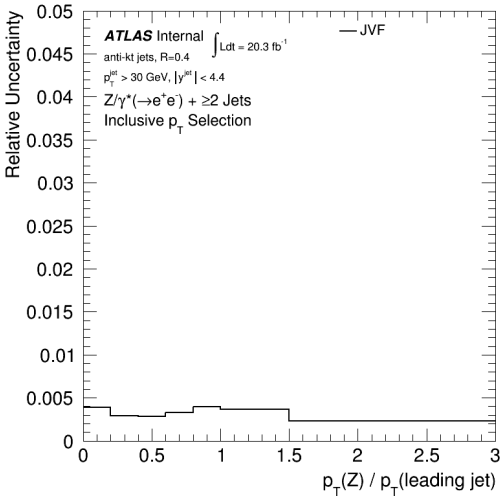
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

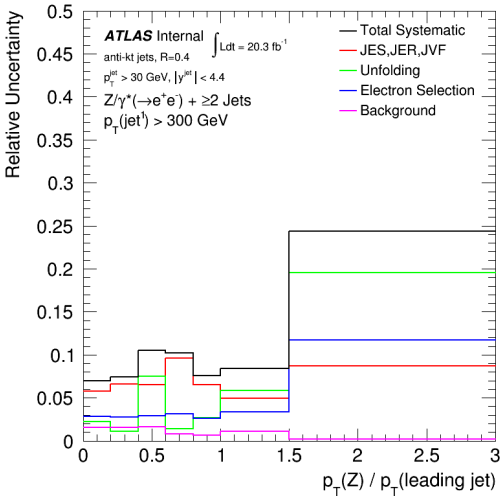


(e) JER Uncertainty

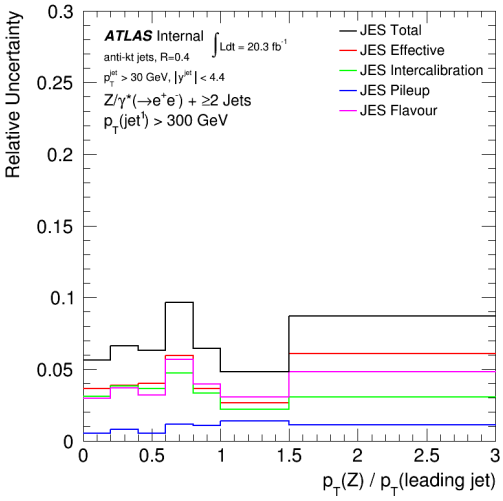


(f) Pile-up Jet Removal Uncertainty

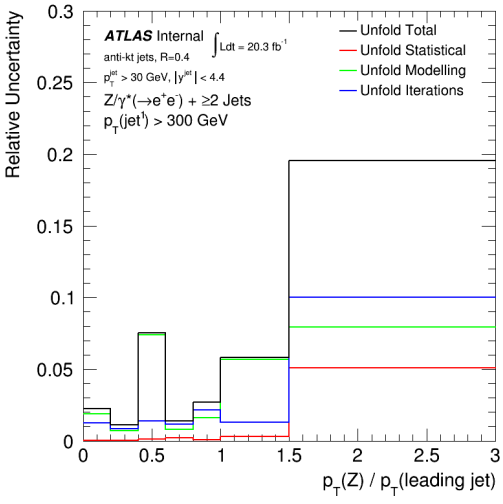
Figure C.12: p_T ratio between Z boson and leading jet (Rp_T^{jj}) in events with at least two jets.



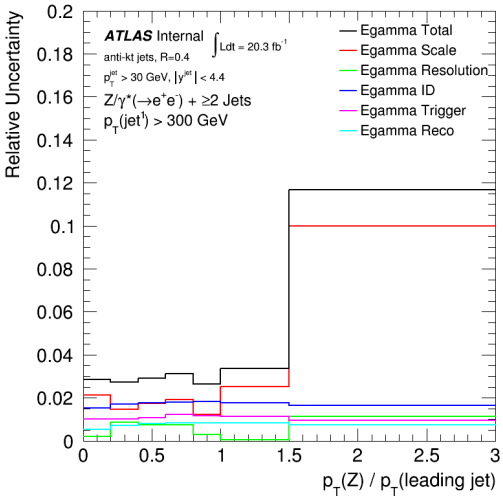
(a) Total Systematic Uncertainty



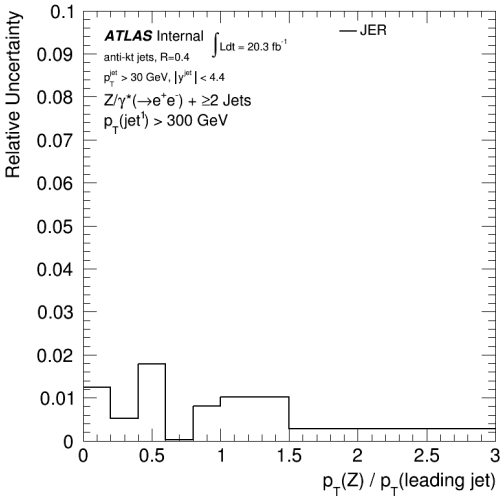
(b) JES Uncertainty



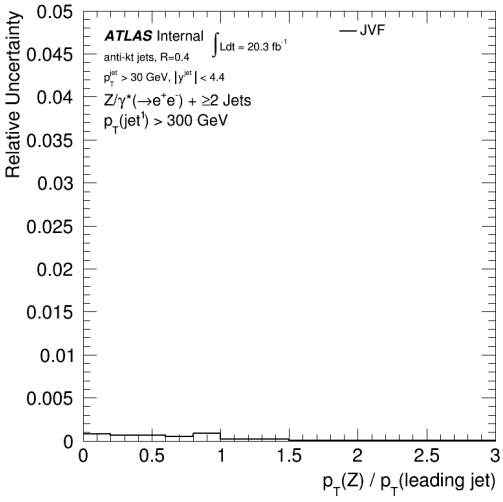
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

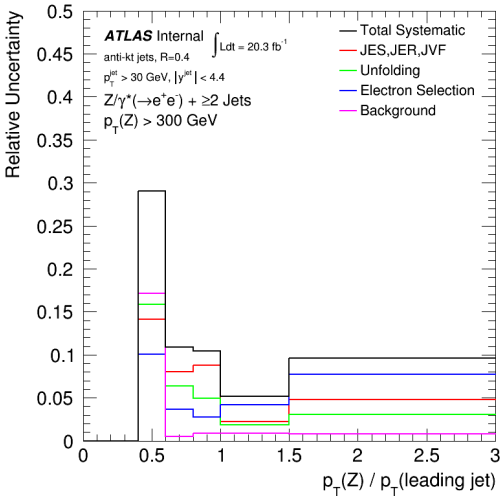


(e) JER Uncertainty

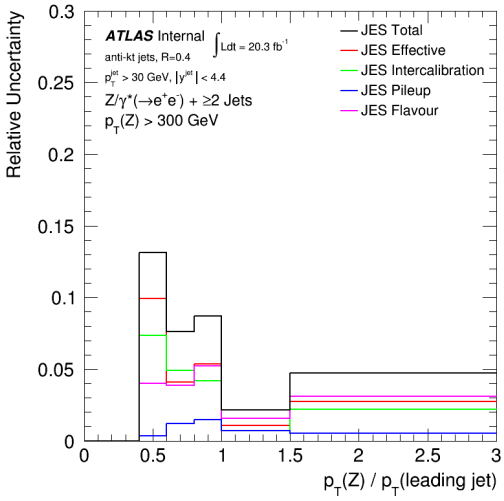


(f) Pile-up Jet Removal Uncertainty

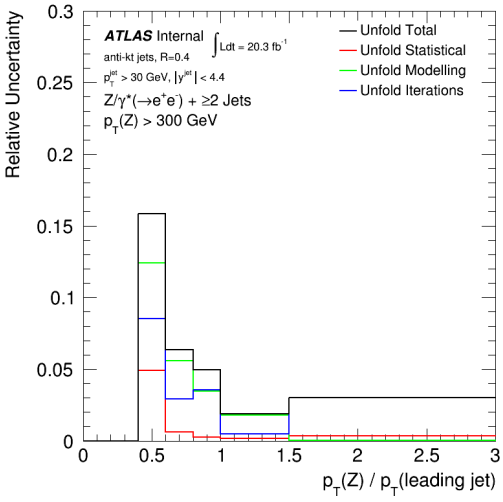
Figure C.13: p_T ratio between Z boson and leading jet (Rp_T^{jj}) in events with at least two jets for $p_T^{\text{leading jet}} > 300$ GeV.



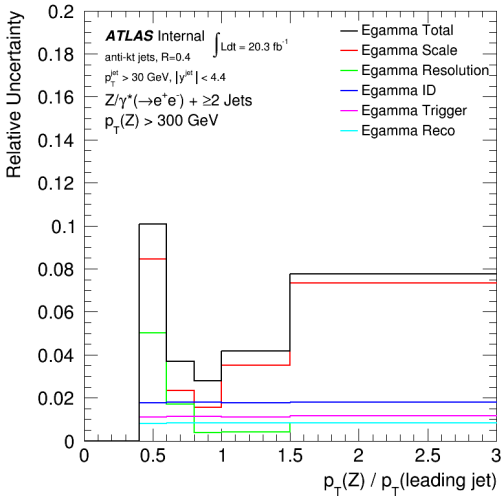
(a) Total Systematic Uncertainty



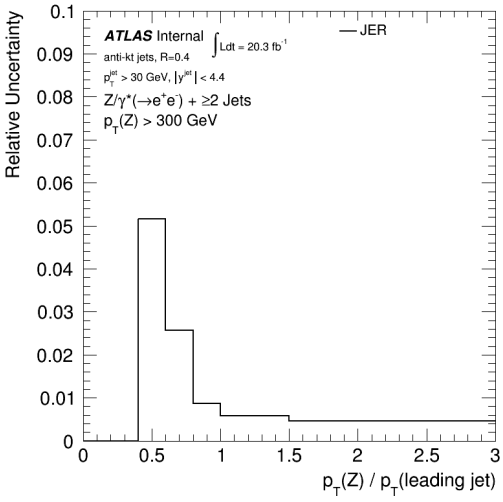
(b) JES Uncertainty



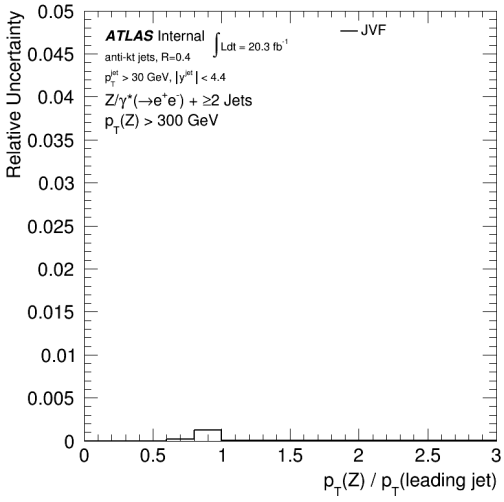
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

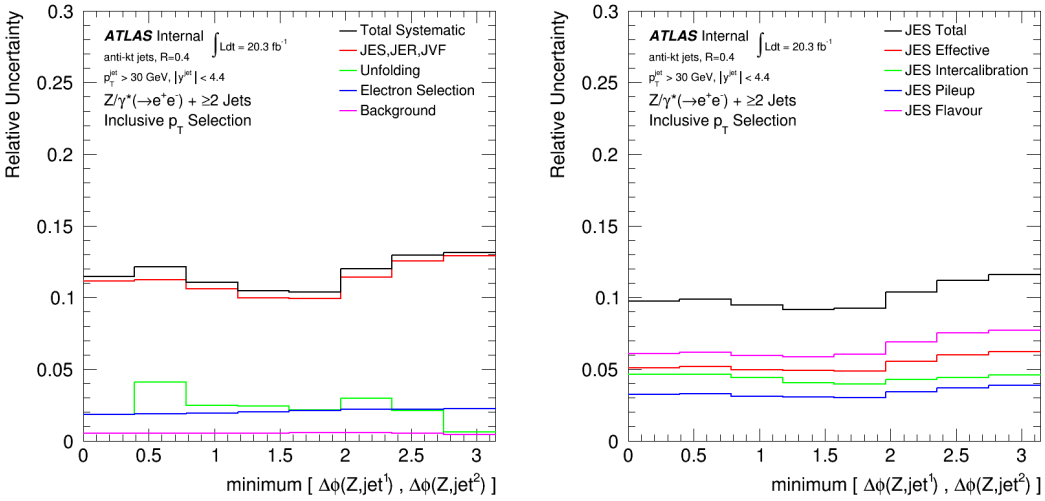


(e) JER Uncertainty



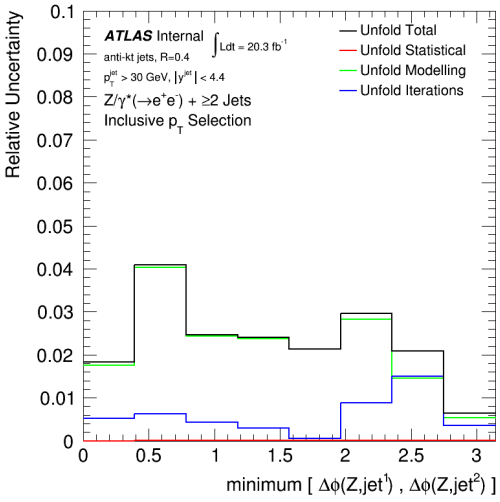
(f) Pile-up Jet Removal Uncertainty

Figure C.14: p_T ratio between Z boson and leading jet (Rp_T^{jj}) in events with at least two jets for $p_T^Z > 300$ GeV.

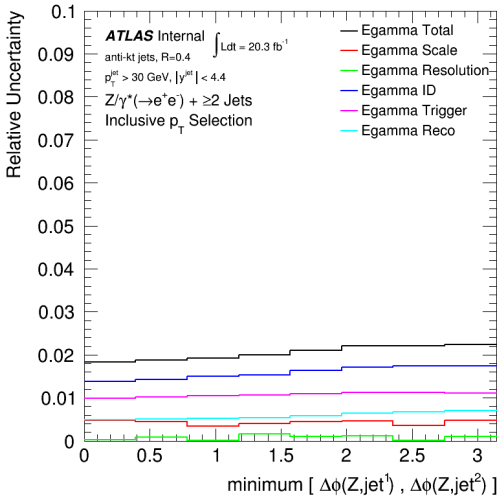


(a) Total Systematic Uncertainty

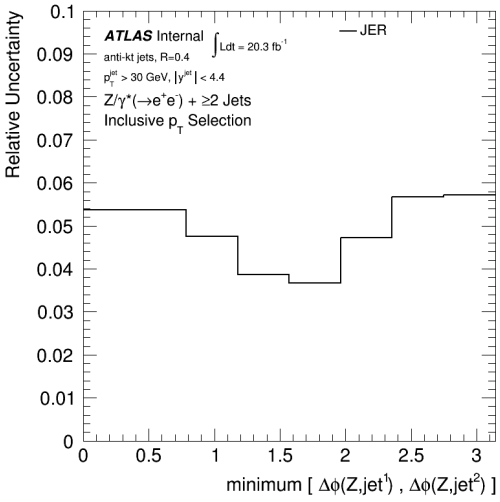
(b) JES Uncertainty



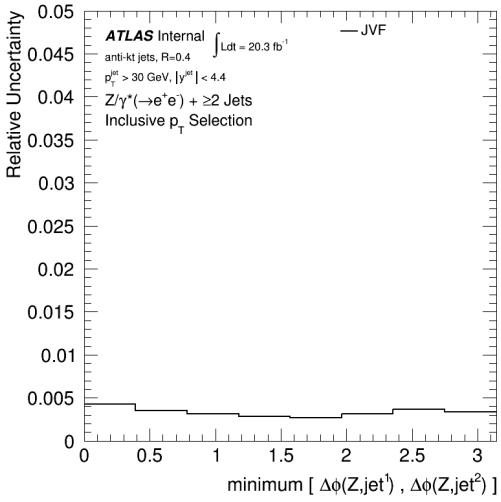
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

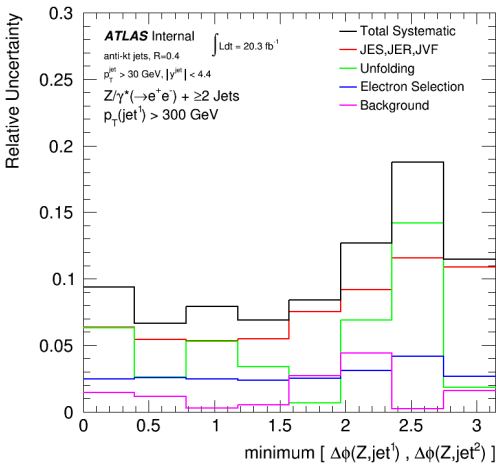


(e) JER Uncertainty

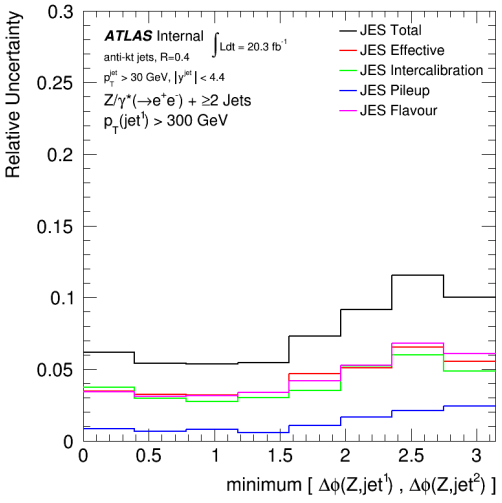


(f) Pile-up Jet Removal Uncertainty

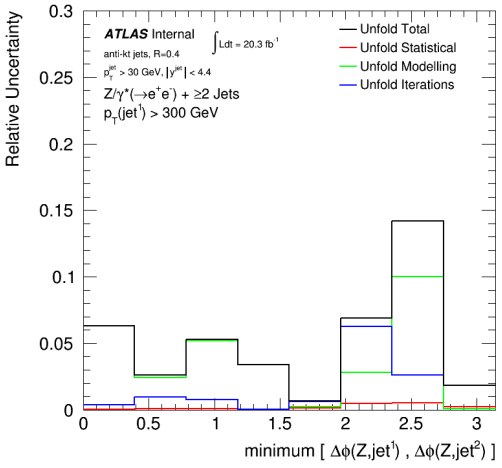
Figure C.15: Minimum $\Delta\phi$ between the Z boson and either of the two leading jets in the event ($min\Delta\phi^{Zjj}$) in events with at least two jets.



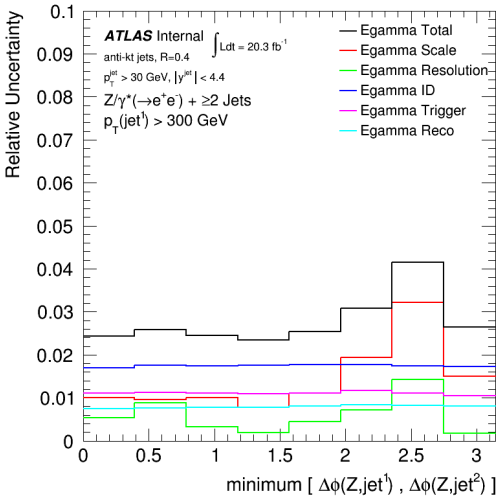
(a) Total Systematic Uncertainty



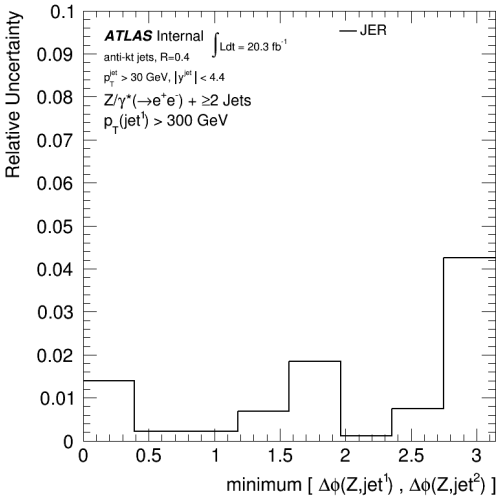
(b) JES Uncertainty



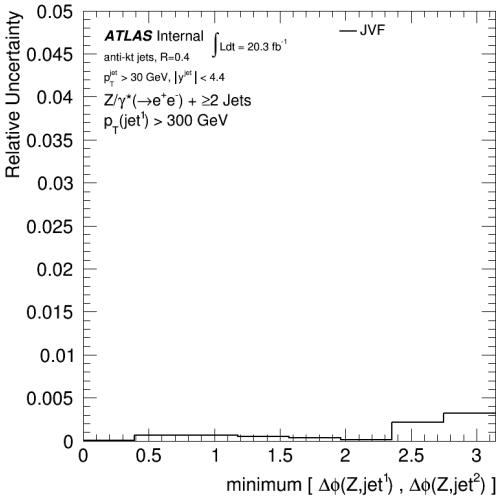
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

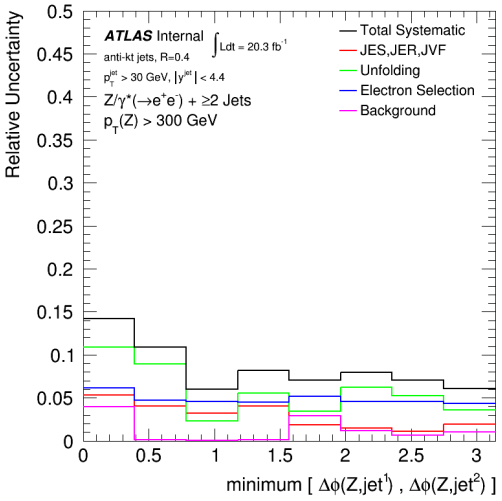


(e) JER Uncertainty

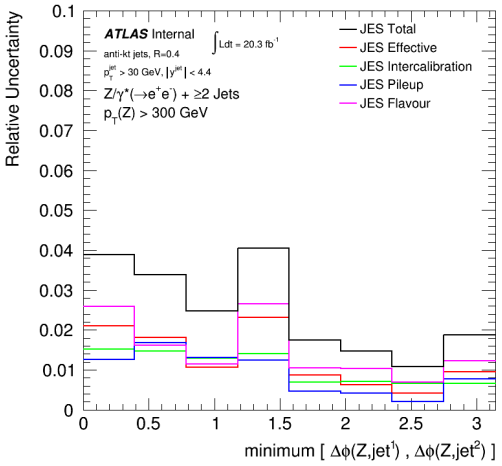


(f) Pile-up Jet Removal Uncertainty

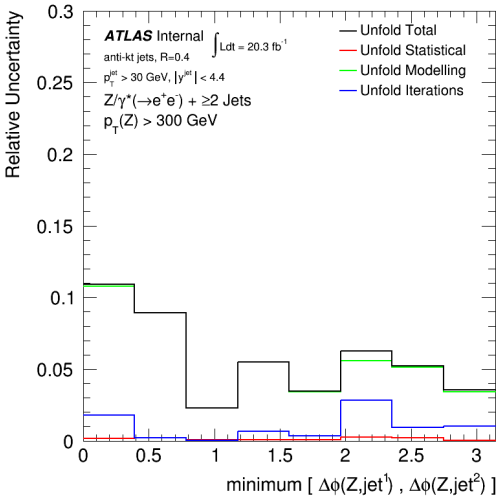
Figure C.16: Minimum $\Delta\phi$ between the Z boson and either of the two leading jets in the event ($min\Delta\phi^{Zjj}$) in events with at least two jets for $p_T^{\text{leading jet}} > 300$ GeV.



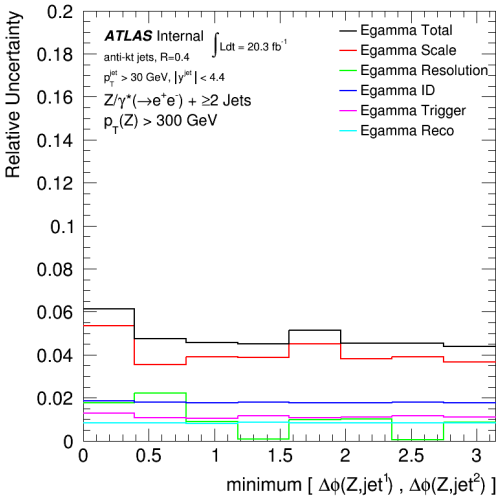
(a) Total Systematic Uncertainty



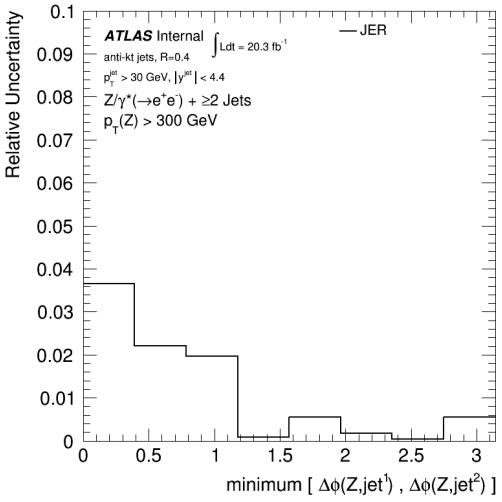
(b) JES Uncertainty



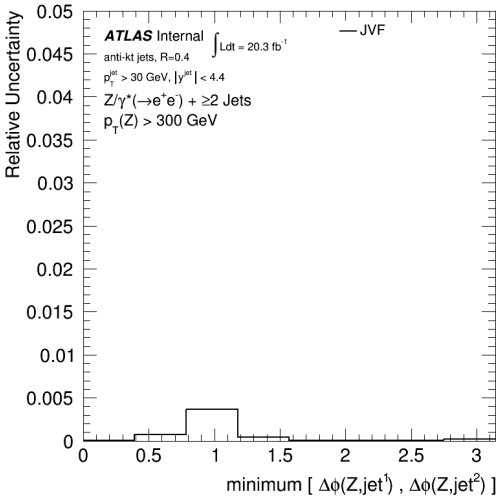
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

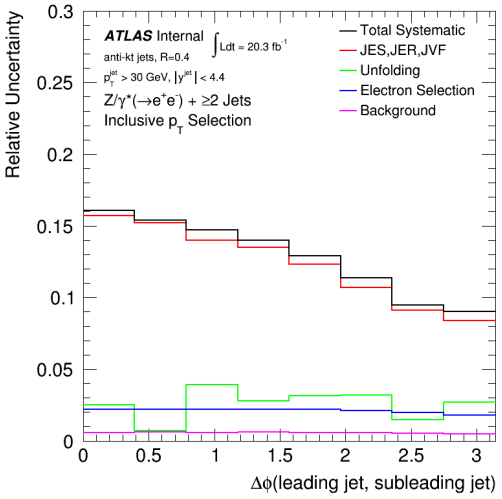


(e) JER Uncertainty

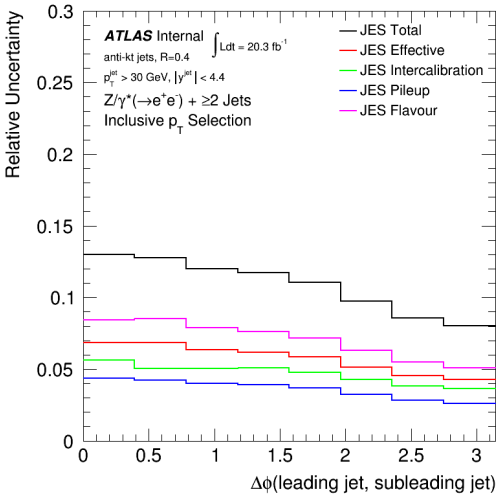


(f) Pile-up Jet Removal Uncertainty

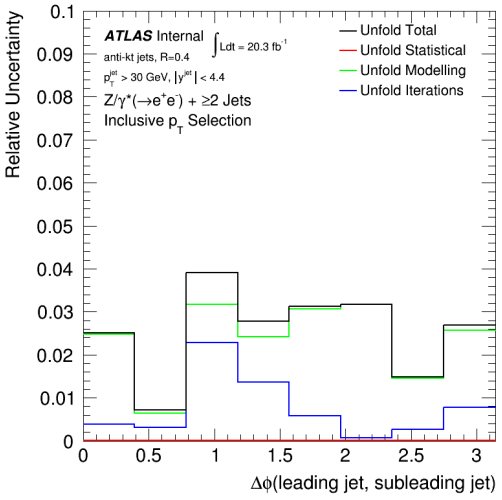
Figure C.17: Minimum $\Delta\phi$ between the Z boson and either of the two leading jets in the event ($min\Delta\phi^{Zjj}$) in events with at least two jets for $p_T^Z > 300$ GeV.



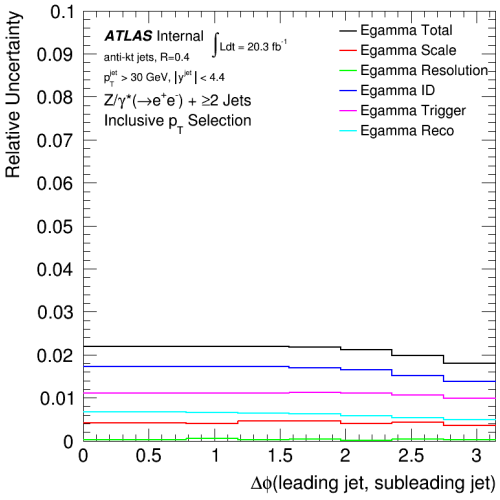
(a) Total Systematic Uncertainty



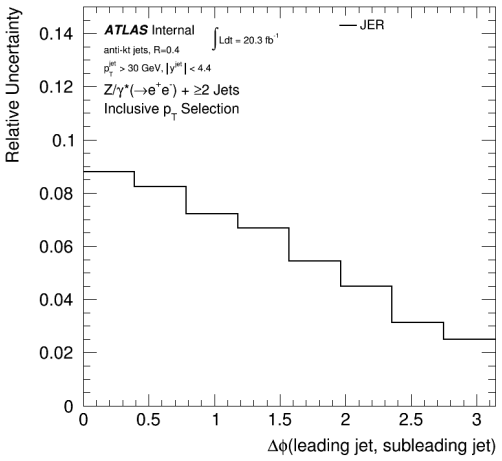
(b) JES Uncertainty



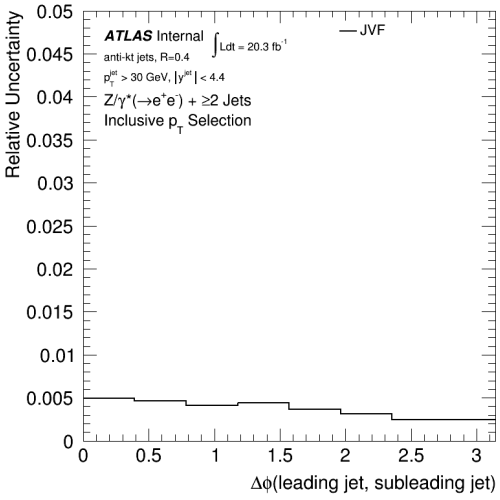
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

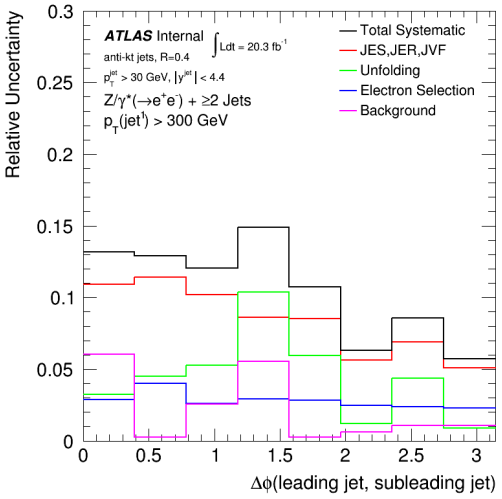


(e) JER Uncertainty

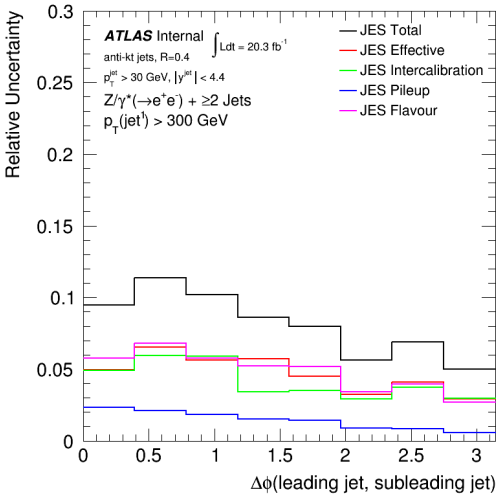


(f) Pile-up Jet Removal Uncertainty

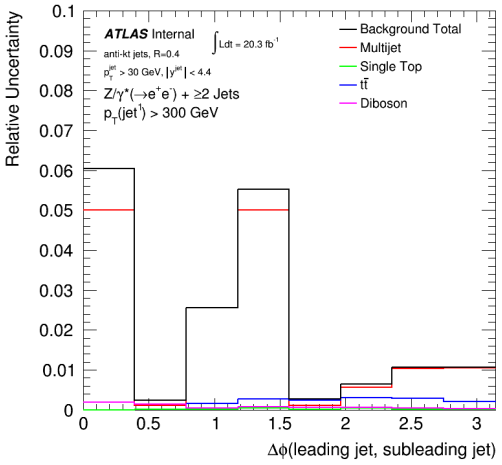
Figure C.18: Absolute azimuthal separation between the two leading jets ($\Delta\phi^{jj}$) in events with at least two jets.



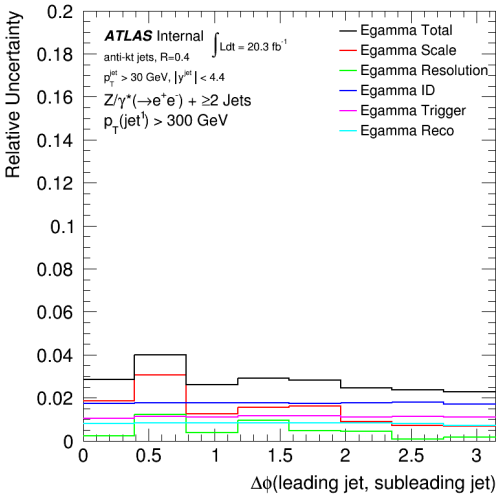
(a) Total Systematic Uncertainty



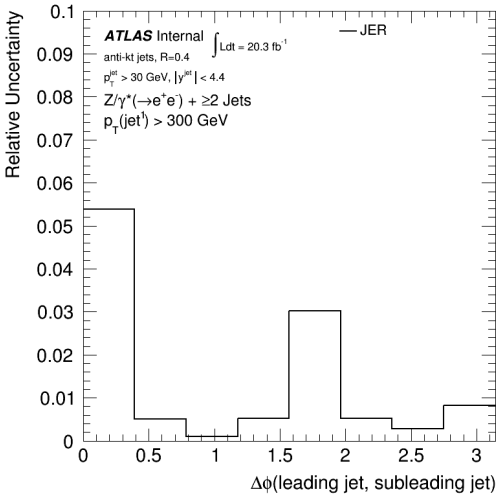
(b) JES Uncertainty



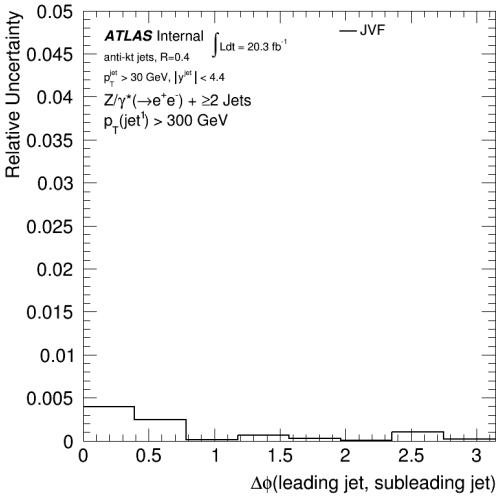
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty

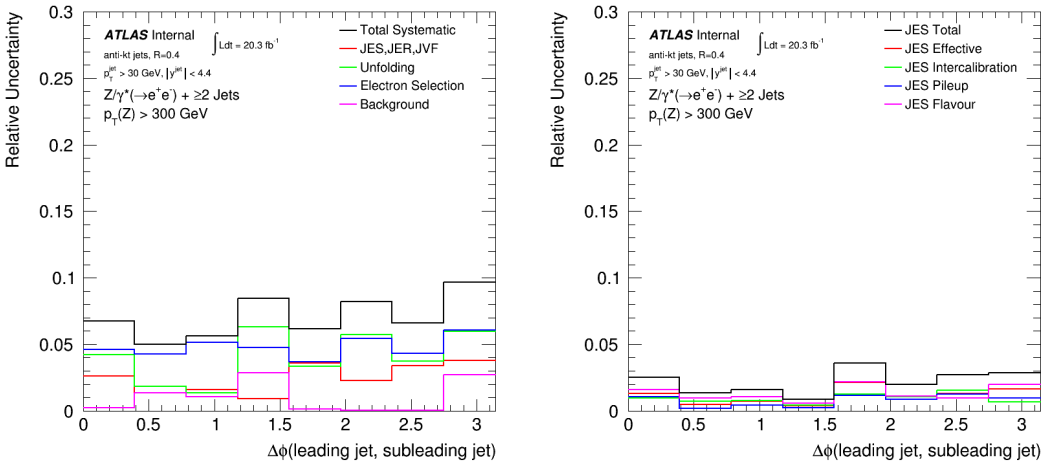


(e) JER Uncertainty



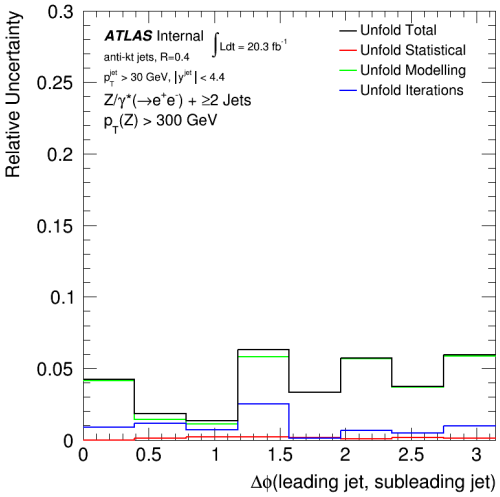
(f) Pile-up Jet Removal Uncertainty

Figure C.19: Absolute azimuthal separation between the two leading jets ($\Delta\phi^{jj}$) in events with at least two jets for $p_T^{\text{leading jet}} > 300$ GeV.

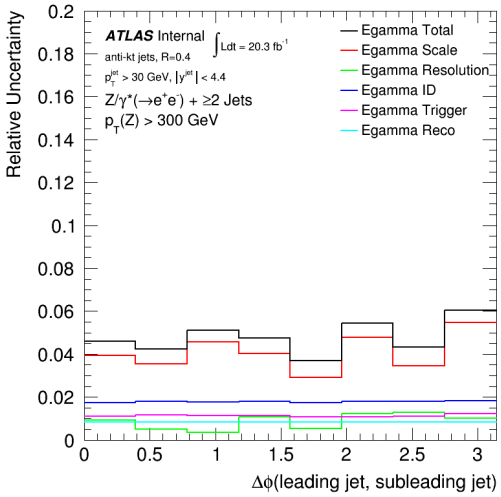


(a) Total Systematic Uncertainty

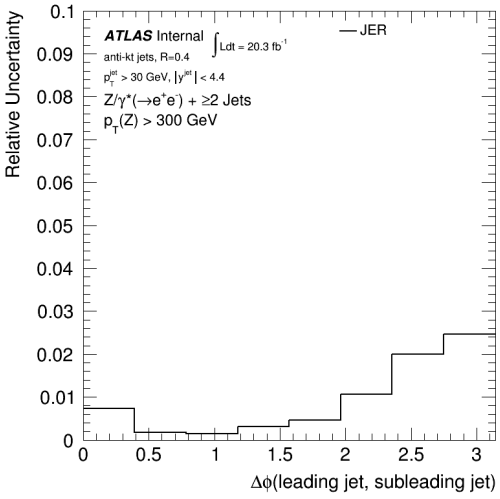
(b) JES Uncertainty



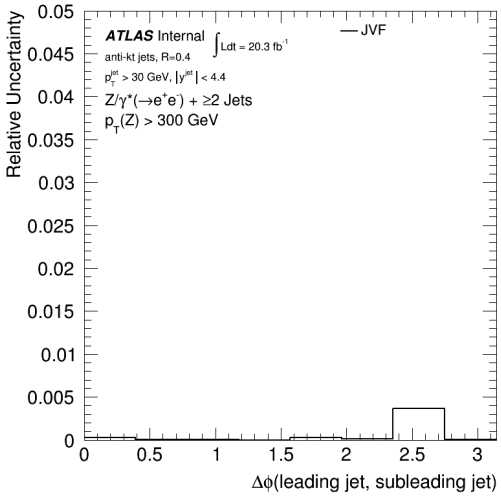
(c) Unfolding Uncertainty



(d) Electron Calibration Uncertainty



(e) JER Uncertainty



(f) Pile-up Jet Removal Uncertainty

Figure C.20: Absolute azimuthal separation between the two leading jets ($\Delta\phi^{jj}$) in events with at least two jets for $p_T^Z > 300$ GeV.

APPENDIX D

Glossary of Abbreviations

AOD - Analysis Object Data
ATLAS - A Toroidal LHC ApparatuS
CKKW - Catani-Krauss-Kuhn-Webber
CMS - Compact Muon Solenoid
CSC - Cathode Strip Chamber
CTEQ - Coordinated Theoretical-Experimental Project on QCD
D3PD - Derived Physics Data
DGLAP - Dokshitzer-Gribov-Lipatov-Altarelli-Parisi
DIS - Deep Inelastic Scattering
DY - Drell-Yan
EF - Event Filter
EM - Electromagnetic
EMCAL - Electromagnetic Calorimeter
ESD - Event Summary Data
EW - Electroweak
FCAL - Forward Calorimeter
FSR - Final State Radiation
HCAL - Hadronic Calorimeter
HEC - Hadronic Endcap Calorimeter
HLT - High Level Trigger
ID - Inner Detector
ISR - Initial State Radiation
JER - Jet Energy Resolution
JES - Jet Energy Scale
JVF - Jet Vertex Fraction
LEP - Large Electron-Positron Collider
LAr - Liquid Argon
LHAPDF - Les Houches Accord PDF Interface
LHC - Large Hadron Collider
LO - Leading Order
LVL1 - Level 1 Trigger
LVL2 - Level 2 Trigger

MC - Monte Carlo
MDT - Monitored Drift Tube
ME - Matrix Element
MENLOPS - Matrix Element Next-to-Leading Order Parton Shower
MET - Missing Transverse Energy
MLM - Michelangelo-L-Mangano
MPI - Multiple Particle Interactions
MRST - Martin-Roberts-Stirling-Thorne
MSTW - Martin-Stirling-Thorne-Watt
NLO - Next-to-Leading Order
NNLO - Next-to-Next-to-Leading Order
NNPDF - Neural Network Parton Distribution Functions
PDF - Parton Distribution Function
pQCD - Perturbative Quantum Chromodynamics
PS - Parton Shower
PXL - Pixel
QCD - Quantum Chromodynamics
QED - Quantum Electrodynamics
QPM - Quark Parton Model
RDO - Local Cell signal Weighting
RDO - Raw Data Object
ROB - Readout Buffer
RoI - Region of Interest
RPC - Resistive Plate Chamber
SCT - Silicon Micro-Strip
SM - Standard Model
SPS - Super Proton Synchrotron
TDAQ - Trigger and Data Acqusition
TGC - Thin Gap Chamber
TRT - Transition Radiation Tracker
UE - Underlying Event
YFS - Yennie-Frautschi-Suura

REFERENCES

[1] C. F. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, H. Ita, D. A. Kosower, and D. Maitre. An Automated Implementation of On-Shell Methods for One-Loop Amplitudes. *Phys. Rev.*, D78:036003, 2008.

[2] H. Ita, Z. Bern, L. J. Dixon, Fernando Febres Cordero, D. A. Kosower, and D. Maitre. Precise Predictions for Z + 4 Jets at Hadron Colliders. *Phys. Rev.*, D85:031501, 2012.

[3] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003.

[4] T. Gleisberg, S. Hche, F. Krauss, M. Schnherr, S. Schumann, F. Siegert, and J. Winter. Event generation with sherpa 1.1. *JHEP*, 2009(02):007, 2009.

[5] Stefan Hoeche, Frank Krauss, Marek Schonherr, and Frank Siegert. QCD matrix elements + parton showers: The NLO case. *JHEP*, 04:027, 2013.

[6] Mathieu Rubin, Gavin P. Salam, and Sebastian Sapeta. Giant QCD K-factors beyond NLO. *JHEP*, 09:084, 2010.

[7] Ansgar Denner, Stefan Dittmaier, Tobias Kasprzik, and Alexander Muck. Electroweak corrections to dilepton + jet production at hadron colliders. *JHEP*, 06:069, 2011.

[8] T. Aaltonen et al. Measurement of inclusive jet cross-sections in $Z/\gamma^* \to e^+e^-$ + jets production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$-TeV. *Phys. Rev. Lett.*, 100:102001, 2008.

[9] Victor Mukhamedovich Abazov et al. Measurement of Z/γ^*+jet+X angular distributions in p anti-p collisions at $\sqrt{s} = 1.96$.TeV. *Phys. Lett.*, B682:370–380, 2010.

[10] Georges Aad et al. Measurement of the production cross section of jets in association with a Z boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 07:032, 2013.

[11] Vardan Khachatryan et al. Measurements of jet multiplicity and differential production cross sections of Z+ jets events in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev.*, D91(5):052008, 2015.

[12] Serguei Chatrchyan et al. Event shapes and azimuthal correlations in Z + jets events in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Lett.*, B722:238–261, 2013.

[13] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.

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

[15] Wikipedia. Standard model — Wikipedia, the free encyclopedia, 2016. [Online; accessed 10-July-2016].

[16] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[17] P.W. Higgs. Broken symmetries, massless particles and gauge fields. *Physics Letters*, 12(2):132 – 133, 1964.

[18] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[19] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.

[20] Peter W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.*, 145:1156–1163, May 1966.

[21] T. W. B. Kibble. Symmetry breaking in non-abelian gauge theories. *Phys. Rev.*, 155:1554–1561, Mar 1967.

[22] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012.

[23] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[24] John M. Campbell, J. W. Huston, and W. J. Stirling. Hard Interactions of Quarks and Gluons: A Primer for LHC Physics. *Rept. Prog. Phys.*, 70:89, 2007.

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

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

[27] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc.*, C680519:367–377, 1968.

[28] Matt Dobbs and Jrgen Beck Hansen. The hepmc c++ monte carlo event record for high energy physics. *Computer Physics Communications*, 134(1):41 – 46, 2001.

[29] R. P. Feynman. *Photon-hadron interactions*. Addison-Wesley, 1972. Reading 1972, 282p.

[30] R.Keith Ellis, Howard Georgi, Marie Machacek, H.David Politzer, and Graham G. Ross. Perturbation theory and the parton model in qcd. *Nuclear Physics B*, 152(2):285 – 329, 1979.

[31] R. Devenish and A. Cooper-Sarkar. *Deep inelastic scattering*. 2004.

[32] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys.*, B126:298–318, 1977.

[33] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977. [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[34] Pavel M. Nadolsky, Hung-Liang Lai, Qing-Hong Cao, Joey Huston, Jon Pumplin, Daniel Stump, Wu-Ki Tung, and C. P. Yuan. Implications of CTEQ global analysis for collider observables. *Phys. Rev.*, D78:013004, 2008.

[35] Sayipjamal Dulat, Tie-Jiun Hou, Jun Gao, Marco Guzzi, Joey Huston, Pavel Nadolsky, Jon Pumplin, Carl Schmidt, Daniel Stump, and C. P. Yuan. New parton distribution functions from a global analysis of quantum chromodynamics. *Phys. Rev.*, D93(3):033006, 2016.

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

[37] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Uncertainties on alpha(S) in global PDF analyses and implications for predicted hadronic cross sections. *Eur. Phys. J.*, C64:653–680, 2009.

[38] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Heavy-quark mass dependence in global PDF analyses and 3- and 4-flavour parton distributions. *Eur. Phys. J.*, C70:51–72, 2010.

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

[40] Andy Buckley, James Ferrando, Stephen Lloyd, Karl Nordstrm, Ben Page, Martin Rfenacht, Marek Schnherr, and Graeme Watt. LHAPDF6: parton density access in the LHC precision era. *Eur. Phys. J.*, C75:132, 2015.

[41] J C Collins, , and D E Soper. The theorems of perturbative qcd. *Annual Review of Nuclear and Particle Science*, 37(1):383–409, 1987.

[42] Thomas Ferbel, editor. *Techniques and Concepts of High-Energy Physics IV*, volume 164. Springer US, 1987.

[43] Torbjörn Sjöstrand and Maria van Zijl. A multiple-interaction model for the event structure in hadron collisions. *Phys. Rev. D*, 36:2019–2041, Oct 1987.

[44] T. Sjostrand and Peter Z. Skands. Multiple interactions and the structure of beam remnants. *JHEP*, 03:053, 2004.

[45] Georges Aad et al. Underlying event characteristics and their dependence on jet size of charged-particle jet events in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Rev.*, D86:072004, 2012.

[46] Peter Zeiler Skands. Tuning Monte Carlo Generators: The Perugia Tunes. *Phys. Rev.*, D82:074018, 2010.

[47] C. Alexa et al. Measurement of Charged Particle Spectra in Deep-Inelastic ep Scattering at HERA. *Eur. Phys. J.*, C73(4):2406, 2013.

[48] Thomas D. Gottschalk. Backwards Evolved Initial State Parton Showers. *Nucl. Phys.*, B277:700–738, 1986.

[49] Torbjorn Sjostrand. A Model for Initial State Parton Showers. *Phys. Lett.*, B157:321–325, 1985.

[50] Mats Bengtsson and Torbjorn Sjostrand. Coherent Parton Showers Versus Matrix Elements: Implications of PETRA - PEP Data. *Phys. Lett.*, B185:435, 1987.

[51] M. Bengtsson and T. Sjöstrand. A comparative study of coherent and non-coherent parton shower evolution. *Nuclear Physics B*, 289:810–846, 1987.

[52] Michelangelo L. Mangano, Mauro Moretti, and Roberto Pittau. Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b} + n$ jets as a case study. *Nucl. Phys.*, B632:343–362, 2002.

[53] F. Caravaglios, Michelangelo L. Mangano, M. Moretti, and R. Pittau. A New approach to multijet calculations in hadron collisions. *Nucl. Phys.*, B539:215–232, 1999.

[54] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber. QCD matrix elements + parton showers. *JHEP*, 11:063, 2001.

[55] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjstrand. Parton fragmentation and string dynamics. *Physics Reports*, 97(23):31 – 145, 1983.

[56] Hans-Uno Bengtsson and Torbjrn Sjstrand. The lund monte carlo for hadronic processes pythia version 4.8. *Computer Physics Communications*, 46(1):43 – 82, 1987.

[57] Yoichiro Nambu. QCD and the String Model. *Phys. Lett.*, B80:372–376, 1979.

[58] Thomas D. Gottschalk. A realistic model for e+e annihilation including parton bremsstrahlung effects. *Nuclear Physics B*, 214(2):201 – 222, 1983.

[59] Thomas D. Gottschalk. An improved description of hadronization in the qcd cluster model for e+e annihilation. *Nuclear Physics B*, 239(2):349 – 381, 1984.

[60] B.R. Webber. A qcd model for jet fragmentation including soft gluon interference. *Nuclear Physics B*, 238(3):492 – 528, 1984.

[61] Thomas D. Gottschalk and Duncan A. Morris. A new model for hadronization and e+e annihilation. *Nuclear Physics B*, 288:729 – 781, 1987.

[62] D. Amati and G. Veneziano. Preconfinement as a Property of Perturbative QCD. *Phys. Lett.*, B83:87–92, 1979.

[63] G. Arnison et al. Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c^2 at the CERN SPS Collider. *Phys. Lett.*, B126:398–410, 1983.

[64] Charalampos Anastasiou, Lance J. Dixon, Kirill Melnikov, and Frank Petriello. High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO. *Phys. Rev.*, D69:094008, 2004.

[65] Ryan Gavin, Ye Li, Frank Petriello, and Seth Quackenbush. FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order. *Comput. Phys. Commun.*, 182:2388–2403, 2011.

[66] Ryan Gavin, Ye Li, Frank Petriello, and Seth Quackenbush. W Physics at the LHC with FEWZ 2.1. *Comput. Phys. Commun.*, 184:208–214, 2013.

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

[68] J Butterworth, E Dobson, U Klein, B Mellado Garcia, T Nunnemann, J Qian, D Rebuzzi, and R Tanaka. Single Boson and Diboson Production Cross Sections in pp Collisions at $\sqrt{s}$=7 TeV. Technical Report ATL-COM-PHYS-2010-695, CERN, Geneva, Aug 2010.

[69] Erik Gerwick, Tilman Plehn, Steffen Schumann, and Peter Schichtel. Scaling Patterns for QCD Jets. *JHEP*, 10:162, 2012.

[70] S.D. Ellis, R. Kleiss, and W.J. Stirling. W's, Z's and jets. *Physics Letters B*, 154(5):435 – 440, 1985.

[71] F.A. Berends, W.T. Giele, H. Kuijf, R. Kleiss, and W.J. Stirling. Multijet production in w, z events at pp colliders. *Physics Letters B*, 224(1):237 – 242, 1989.

[72] F.A. Berends, H. Kuijf, B. Tausk, and W.T. Giele. On the production of a W and jets at hadron colliders. *Nuclear Physics B*, 357(1):32 – 64, 1991.

[73] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber. HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 01:010, 2001.

[74] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, 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. *JHEP*, 07:079, 2014.

[75] Carlo Oleari. The POWHEG-BOX. *Nucl. Phys. Proc. Suppl.*, 205-206:36–41, 2010.

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

[77] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.

[78] Lyndon Evans and Philip Bryant. LHC machine. *Journal of Instrumentation*, 3(08):S08001, 2008.

[79] Stephen Myers. *The LEP Collider, from design to approval and commissioning*. John Adams' Lecture. CERN, Geneva, 1991. Delivered at CERN, 26 Nov 1990.

[80] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[81] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008.

[82] A. Augusto Alves, Jr. et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

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

[84] LHC. 2012 - Progress. `http://lhc-commissioning.web.cern.ch/lhc-commissioning/progress2012.htm`, 2012. [Online; accessed 03-February-2016].

[85] LHC. LuminosityPublicResults. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults`, 2016. [Online; accessed 03-February-2016].

[86] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[87] ATLAS magnet system: Technical design report. 1997.

[88] ATLAS inner detector: Technical design report. Vol. 1. 1997.

[89] A. Airapetian et al. ATLAS calorimeter performance Technical Design Report. 1996.

[90] M. Aharrouche et al. Energy linearity and resolution of the atlas electromagnetic barrel calorimeter in an electron test-beam. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 568(2):601 – 623, 2006.

[91] P. Adragna et al. Testbeam studies of production modules of the ATLAS tile calorimeter. *Nucl. Instrum. Meth.*, A606:362–394, 2009.

[92] ATLAS muon spectrometer: Technical design report. 1997.

[93] ATLAS high-level trigger, data acquisition and controls: Technical design report. 2003.

[94] A. Airapetian et al. ATLAS: Detector and physics performance technical design report. Volume 1. 1999.

[95] Stefano Frixione and Bryan R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 06:029, 2002.

[96] *ATLAS Computing: technical design report*. Technical Design Report ATLAS. CERN, Geneva, 2005.

[97] S. Agostinelli et al. Geant4 - a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003.

[98] J. Allison et al. Geant4 developments and applications. *Nuclear Science, IEEE Transactions on*, 53(1):270–278, Feb 2006.

[99] Elzbieta Richter-Was, D Froidevaux, and Luc Poggioli. ATLFAST 2.0 a fast simulation package for ATLAS. Technical Report ATL-PHYS-98-131, CERN, Geneva, Nov 1998.

[100] Rene Brun and Fons Rademakers. ROOT - An Object Oriented Data Analysis Framework. *Nucl. Inst. and Meth. in Phys. Res. A*, 389:81–86, 1997. See also http://root.cern.ch/.

[101] G.J. Alner et al. The UA5 high energy pp simulation program. *Nuclear Physics B*, 291:445 – 502, 1987.

[102] J M Butterworth and J R Forshaw. Photoproduction of multi-jet events at hera: a monte carlo simulation. *Journal of Physics G: Nuclear and Particle Physics*, 19(10):1657, 1993.

[103] J. M. Butterworth, Jeffrey R. Forshaw, and M. H. Seymour. Multiparton interactions in photoproduction at HERA. *Z. Phys.*, C72:637–646, 1996.

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

[105] Steffen Schumann and Frank Krauss. A Parton shower algorithm based on Catani-Seymour dipole factorisation. *JHEP*, 03:038, 2008.

[106] Stefan Hoeche, Frank Krauss, Steffen Schumann, and Frank Siegert. QCD matrix elements and truncated showers. *JHEP*, 05:053, 2009.

[107] Torbjörn Sjöstrand and Maria van Zijl. A multiple-interaction model for the event structure in hadron collisions. *Phys. Rev. D*, 36:2019–2041, Oct 1987.

[108] Jan-Christopher Winter, Frank Krauss, and Gerhard Soff. A Modified cluster hadronization model. *Eur. Phys. J.*, C36:381–395, 2004.

[109] D.R Yennie, S.C Frautschi, and H Suura. The infrared divergence phenomena and high-energy processes. *Annals of Physics*, 13(3):379 – 452, 1961.

[110] Stefan Hoche, Frank Krauss, Marek Schonherr, and Frank Siegert. NLO matrix elements and truncated showers. *JHEP*, 08:123, 2011.

[111] Keith Hamilton and Paolo Nason. Improving NLO-parton shower matched simulations with higher order matrix elements. *JHEP*, 06:039, 2010.

[112] The ATLAS collaboration. Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data. 2014.

[113] M. Aharrouche et al. Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam. *Nucl. Instrum. Meth.*, A614:400–432, 2010.

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

[115] Expected electron performance in the ATLAS experiment. Technical Report ATL-PHYS-PUB-2011-006, CERN, Geneva, Apr 2011.

[116] Georges Aad et al. Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data. *Eur. Phys. J.*, C72:1909, 2012.

[117] Georges Aad et al. Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data. *Eur. Phys. J.*, C74(7):2941, 2014.

[118] The ATLAS collaboration. Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data. 2014.

[119] Georges Aad et al. Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data. *Eur. Phys. J.*, C74(10):3071, 2014.

[120] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-kt jet clustering algorithm. *JHEP*, 04:063, 2008.

[121] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J.*, C72:1896, 2012.

[122] C. Cojocaru et al. Hadronic calibration of the atlas liquid argon end-cap calorimeter in the pseudorapidity region in beam tests. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 531(3):481 – 514, 2004.

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

[124] Georges Aad et al. Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Eur. Phys. J.*, C75:17, 2015.

[125] Georges Aad et al. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. 2016.

[126] T Barillari, E Bergeaas Kuutmann, T Carli, J Erdmann, P Giovannini, K J Grahn, C Issever, A Jantsch, A Kiryunin, K Lohwasser, A Maslennikov, S Menke, H Oberlack, G Pospelov, E Rauter, P Schacht, F Span, P Speckmayer, P Stavina, and P Strzenec. Local Hadronic Calibration. Technical Report ATL-LARG-PUB-2009-001-2. ATL-COM-LARG-2008-006. ATL-LARG-PUB-2009-001, CERN, Geneva, Jun 2008. Due to a report-number conflict with another document, the report-number ATL-LARG-PUB-2009-001-2 has been assigned.

[127] Jet global sequential corrections with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. Technical Report ATLAS-CONF-2015-002, CERN, Geneva, Mar 2015.

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

[129] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Catchment Area of Jets. *JHEP*, 04:005, 2008.

[130] N Makovec. Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data. Technical Report ATLAS-COM-CONF-2012-018, CERN, Geneva, Feb 2012. .

[131] Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS. Technical Report ATLAS-CONF-2015-037, CERN, Geneva, Aug 2015.

[132] Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s} = 8\ TeV$. Technical Report ATLAS-CONF-2015-017, CERN, Geneva, Apr 2015.

[133] Measurement of the production cross section for Z/γ^* in association with jets in pp collisions at s = 7 TeV with the ATLAS Detector. Technical Report ATLAS-CONF-2011-001, CERN, Geneva, Feb 2011.

[134] New ATLAS event generator tunes to 2010 data. Technical Report ATL-PHYS-PUB-2011-008, CERN, Geneva, Apr 2011.

[135] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann. HATHOR: HAdronic Top and Heavy quarks cross section calculatoR. *Comput. Phys. Commun.*, 182:1034–1046, 2011.

[136] M. Beneke, P. Falgari, S. Klein, and C. Schwinn. Hadronic top-quark pair production with NNLL threshold resummation. *Nuclear Physics B*, 855(3):695 – 741, 2012.

[137] N. Davidson, G. Nanava, T. Przedzinski, E. Richter-Was, and Z. Was. Universal Interface of TAUOLA Technical and Physics Documentation. *Comput. Phys. Commun.*, 183:821–843, 2012.

[138] G. D'Agostini. A multidimensional unfolding method based on bayes' theorem. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 362(23):487 – 498, 1995.

[139] Tim Adye. Unfolding algorithms and tests using RooUnfold. In *Proceedings of the PHYSTAT 2011 Workshop, CERN, Geneva, Switzerland, January 2011, CERN-2011-006, pp 313-318*, pages 313–318, 2011.