Measurements of the production cross section of a Z boson in association with high-transverse-momentum jets with the ATLAS detector at the Large Hadron Collider

by

Alexandre Laurier

A thesis submitted to the

Faculty of Graduate and Postdoctoral Affairs

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy in Physics

Department of Physics

Carleton University

Ottawa-Carleton Institute of Physics

Ottawa, Canada

August 31, 2022

CERN-THESIS-2022-132
23/08/2022

Abstract

In this thesis, the measurements of the production cross section of a Z boson produced in association with high-transverse-momentum jets and decaying into a charged-lepton pair are presented. The data used for these measurements are analyzed from proton–proton collisions at the Large Hadron Collider. The dataset corresponds to an integrated luminosity of 139 fb^{-1} collected at a centre-of-mass energy of 13 TeV and recorded by the ATLAS detector between 2015 and 2018. The cross-section measurements are performed separately for the electron and muon decay channels of the Z boson and then combined in order to increase the precision of the final $Z(\rightarrow \ell^+\ell^-)$+jets measurements.

Events with at least one jet with $p_{\mathrm{T}} \geq 500$ GeV/c are selected and populate a high-p_{T} region. In this region, the production of an on-shell Z boson radiated by a quark is enhanced and results in events with a small angle between the Z boson and the associated quark jet such that the two objects are measured to be collinear. A fraction of the events in the high-p_{T} region are also described by Z+1 jet events where the Z boson and the jet recoil against each-other such that the two objects are measured to be back-to-back.

The large dataset has made it possible for the first time to separately study these two event topologies where a Z boson and a jet are either collinear or back-to-back. This is achieved by measuring the angular correlation between the Z boson and the closest jet. The production of a Z boson in association with jets, Z+jets, provides a near inexhaustible amount of information about the physics mechanism for the production of vector bosons at the Large Hadron Collider.

This analysis allows to measure the cross section of Z+jets as a function of characteristic observables. The resulting differential cross sections are compared with

state-of-the-art theoretical predictions. The differential cross sections themselves provide a powerful way to test the Standard Model and in particular quantum chromodynamics. The data are found to agree with the latest next-to-next-to-leading-order and next-to-leading-order theoretical predictions for Z+jets production.

Acknowledgments

I would like to thank my advisor Alain for giving me this opportunity and allowing me to further hone my scientific skills. You were a great advisor, thank you for your patience and support.

My deepest thanks to Ulla for leading the analysis team. You have taught me more about Standard Model physics and analyses than I ever thought possible. In the analysis role you have given me, you have let me flourish in ways I couldn't have imagined.

Yoram, your work invitations to come at the Technion have been my highlights of the past few years. Your hospitality is unequaled. Thanks for your support and guidance on my future after my time in ATLAS.

I can't forget all the friends I have made at Carleton, at CERN, and everywhere in between. You have all been fantastic colleagues and have made the past six years so much better.

I would also like to thank my parents, my family and my friends. You have supported me in the last six years of my life and have kept me sane and motivated to produce the best I had to offer.

Et finalement, je remercie Sophie. Ton soutien incessant dans les dernières années de nos vies m'a permis de grandir en tant que scientifique et de manière plus importante, en tant que personne. La période interminable de mes études a été remplie de belles opportunités et de sacrifices. Merci de m'avoir accompagné dans mon parcours doctoral. Ce ne fut pas toujours facile, mais cet accomplissement est en partie grâce à toi.

Contents

List of Tables

List of Figures

Conventions, List of Symbols, Acronyms and Abbreviations

The following symbols, acronyms and abbreviations are listed by the order in which they appear in this thesis.

SM	Standard Model of particle physics
ATLAS	A Toroidal LHC ApparatuS
CMS	Compact Muon Solenoid
LHC	Large Hadron Collider
eV	Electron-Volt
BSM	Beyond the Standard Model
PDF	Parton distribution function (or Probability distribution function)
VBF	Vector boson fusion
VBF	Vector boson scattering
QCD	Quantum chromodynamics
EW	Electroweak
p_{T}	Transverse momentum
pp	Proton–proton
COM	Centre-of-mass

fb^{-1}	Inverse femtobarn
MC	Monte Carlo
QFT	Quantum Field Theory
U(N)	Unitary group of degree N
SU(N)	Special unitary group of degree N
QED	Quantum electrodynamics
α_s	Strong coupling constant
μ_R	Renormalization scale
μ_f	Factorization scale
GWS	Glashow-Weinberg-Salam
σ	Cross section
CERN	European Organization for Nuclear Research
PS	Proton Synchrotron
$p\bar{p}$	Proton–antiproton
LEP	Large electron–positron
LO	Leading order
pQCD	Perturbative quantum chromodynamics, or pertubative QCD
$\Delta R_{Z,j}^{\text{min}}$	Angular distance between the Z boson and the closest jet
NLO	Next-to-leading Order
α_{EW}	Electroweak coupling constant
IP	Interaction point
η	Pseudorapidity
y	Rapidity
ID	Inner detector
Pixel	Silicon pixel detector

SCT	Semiconductor tracker
TRT	Transition radiation tracker
PV	Primary vertex
IBL	Insertible B-layer
ECal	Electromagnetic calorimeter
HCal	Hadronic calorimeter
FCal	Forward calorimeter
LAr	Liquid argon
$\oplus$	Sum in quadrature
MS	Muon spectrometer
MDT	Monitored drift tubes
CSC	Cathode strip chamber
RPC	Resistive plate chamber
TGC	Thin gap chamber
L1 trigger	Level-1 trigger
HLT	High-level trigger
RoI	Region of interest
NNLO	Next-to-next-to-leading Order
$\mathcal{M}$	Scattering Amplitude
ME	Matrix element
GRL	Good run list
$m_{\ell\ell}$	Invariant dilepton mass
$j1$	Leading jet (or highest-energy jet)
$p_{\mathrm{T},j1}$	Leading jet transverse momentum
$p_{\mathrm{T},\ell\ell}$	Z-boson transverse momentum

N_{jets}	Jet multiplicity
H_{T}	Scalar sum of the transverse momentum of all selected jets and leptons
$r_{Z,j}$	Ratio of Z-boson to closest-jet transverse momenta
S_{T}	Scalar sum of the transverse momentum of all selected jets
L	Integrated Luminosity
Zjj	Z+2 jets

Chapter 1

Introduction

The Standard Model (SM) of particle physics provides the most successful theoretical description of elementary particles and their interactions. From its inception in the 1950's, it has evolved to its current form in the 1970's and has since been thoroughly tested and validated by a myriad of experiments. The discovery of the final particle predicted by the Standard Model, the Higgs Boson – the cornerstone of the model - was announced simultaneously by the ATLAS [1] and CMS [2] experiments at the Large Hadron Collider (LHC) [3] in 2012 [4,5]. The LHC was built to probe the SM at the TeV energy scale, and so far no measurement performed at the LHC experiments has announced a discovery going against the predictions of the SM.

The experimental agreement to the SM at the currently accessible energy regimes presents an interesting conundrum. The SM so far disagrees with none of the measurements performed at the LHC, yet it is known to be an incomplete description of the fundamental laws of the universe: it does not include the description of gravity, the most important force in our day-to-day lives, nor can it explain the non-zero neutrino masses or the observed quantity of dark matter and dark energy in the universe. This provides ample motivation for experimentalists to search for physics Beyond the

Standard Model (BSM) or to look for disagreements between the SM and experimental data in order to arrive at a complete description of the laws of physics of our universe.

The class of processes described by the production of a Z boson in association with jets, Z+jets, provides a near inexhaustible amount of information about the SM physics at the LHC. Some processes such as Z+b jets [6] provide important information on the proton Parton Distribution Function (PDF), whereas others such as Vector-Boson-Fusion (VBF) mediated Z+2 jets production [7] provide stringent tests of the SM by measuring very rare processes that are now just becoming statistically accessible. Furthermore, measurements of inclusive Z+jets production constitute powerful tests of perturbative quantum chromodynamics (QCD) [8, 9] and, in the case of high-energy jets, they provide a way to probe the interplay between QCD and higher-order electroweak (EW) processes [10–13]. Such processes also constitute non-negligible backgrounds in measurements of the Higgs boson properties [14, 15] and in searches for new phenomena [16–18], which often exploit the presence of high-p_{T} jets to enrich a data sample with potential signal of new physics. In those studies, predictions are used to extrapolate Z+jets backgrounds from control regions to the signal regions and to model the distributions of the final discriminants.

This thesis summarizes the analysis and measurements of the production cross section of a Z boson in association with high-transverse-momentum (p_{T}) jets [19] (Z+high-p_{T} jets) using the complete LHC Run–2 dataset performed at ATLAS. The data was collected in proton–proton (pp) collisions at the centre-of-mass (COM) energy of 13 TeV between the years of 2015 and 2018, corresponding to an integrated luminosity of 139 fb^{-1}. The measurements are performed on events that contain a Z-boson candidate reconstructed from an e^+e^- pair or a $\mu^+\mu^-$ pair in association with hadronic jets defined as jets having transverse momentum greater than or equal to 100 GeV/c. This is the first analysis of Z+high-p_{T} jets to be performed by either ATLAS or CMS using the full LHC Run–2 dataset, resulting in new and precise measurements of several key observables in extreme high-energy phase spaces. In this analysis, both the differential and integrated fiducial cross sections for the production of Z+high-p_{T} jets are measured.

These measurements are of great interest to the particle physics community since they probe phase spaces which are known to be severely mismodeled by contemporary Monte Carlo (MC) predictions. In particular, Z+jets theoretical predictions in high p_{T}(Z) and p_{T}(jet) phase spaces tend to overestimate data by upwards of 40%–50% [10, 20]. Furthermore, recent analyses highlight the challenge of MC predictions to model the processes where a W or Z boson is emitted nearby to a jet in W/Z+jets events [21, 22].

Details of the Z+high-p_{T} jets measurements are presented in the following order. The theoretical framework of the SM and the history of the Z boson are discussed in Chapter 2. The experimental apparatus, the Large Hadron Collider and the ATLAS detector, is presented in Chapter 3. In Chapter 4, techniques to model hadron-hadron collisions are briefly described, giving some context on the different theoretical predictions that are used in this analysis. Chapter 5 presents the data and simulated samples, while Chapter 6 discusses the event selection requirements. Chapter 7 elaborates on the analysis strategy of these cross-section measurements. The estimation of the various backgrounds to Z+jets production are discussed in Chapter 8. The different sources of systematic uncertainties are presented in Chapter 9. Chapter 10 presents the cross-section results and comparisons to MC and fixed-order theoretical predictions. Conclusions are provided in Chapter 11.

The ATLAS detector was constructed between 2004 and 2009, and is the cumulative work of thousands of engineers and scientists. Data are taken in a general setting, calibrated and processed, and provided to all the ATLAS collaborators. Simulated data are also supplied to all the collaborators via distributed grid computing [23]. Thus, my contribution to the measurement summarized in this work starts from the data processed and calibrated by the ATLAS collaboration. All code used to perform the data-mining performed in this analysis is fully written by myself in a combination of C++ and Python programming languages.

The work in the Z+high-p_{T} jets analysis [19], summarized in this thesis, is in majority performed by myself. In the rare instances where studies were performed by other team members, I have replicated their work to serve as validation or to extend the results when new data were analyzed. As such, unless stated otherwise, all figures pertaining to the measurements presented in this thesis were created by myself. My contributions include, but are not limited to, the choice and optimization of the event selection criteria, the evaluation of backgrounds, the unfolding procedure, the assessment of systematic uncertainties and the combination of the final results. The publicly available HEPdata entry of this analysis and the corresponding rivet routine are also results of my work. The results summarized in this thesis are accepted for publication in the *Journal of High Energy Physics* and publicly available in Ref. [19].

Chapter 2

Theoretical foundations

In this chapter, the theoretical foundation and experimental context concerning the Z boson, one of the gauge bosons that arise from the unification of the electromagnetic and weak forces, is introduced. A brief outline of the Standard Model of particle physics is given in Section 2.1. An historical overview of the Z boson is presented in Section 2.2. The production of the Z boson in association with jets in proton–proton collisions at the Large Hadron Collider is then put forward in Section 2.3.

2.1 The Standard Model

The Standard Model (SM) of particle physics is a quantum field theory (QFT) that describes the interactions and properties of all currently known particles. In the framework of QFT, elementary particles are represented as excitations of their respective quantum fields. The SM describes three of the four[1] fundamental forces of nature: the electromagnetic, weak, and strong forces. The interactions of these forces are mediated by their corresponding gauge bosons, each having spin one, and

[1]The fourth force, gravity, is rather described by the theory of General Relativity and is not compatible with the SM. Efforts for unifying all four forces have been ongoing for decades, without success.

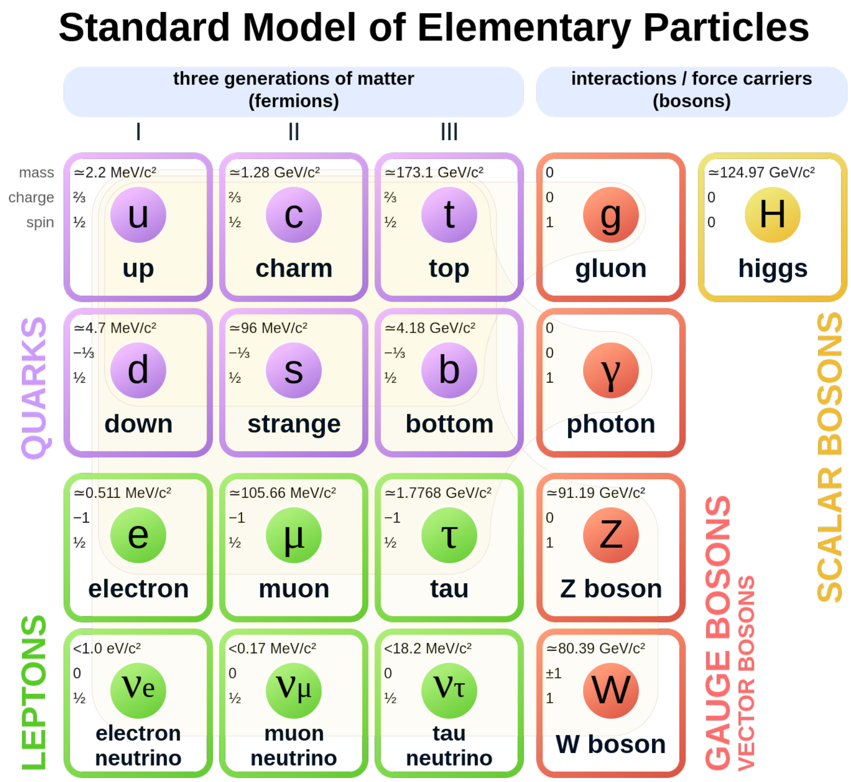


Figure 2.1: Overview of the particle content of the Standard Model of particle physics [24]. For the latest scientific consensus on the masses of particles, please refer to the Particle Data Group [25].

act on the fundamental fermions of spin one-half. The particle content of the SM is summarized in Figure 2.1.

The current form of the SM results from important theoretical efforts in the 1960-70's to describe the theory of the strong, weak and electromagnetic interactions [26–31]. It is based on the internal symmetries of the unitary product group[2]

$$U(1)_Y \times SU(2)_L \times SU(3)_C, \tag{2.1}$$

where $U(1)_Y \times SU(2)_L$ represents the unification of the electromagnetic and weak

[2] A good introduction to Group Theory in the context of particle physics can be found in Ref. [32].

forces, and $SU(3)_C$ the strong force. These groups describe the symmetric transformations of the different interactions that leave the SM Lagrangian invariant. Several conserved quantities arise from the application of Noether's theorem to the guiding local gauge symmetry principles of the SM [33]. These quantities are described by quantum numbers that are associated to the different gauge groups. The quantum number associated to the $SU(3)_C$ gauge group is called colour. It serves as the charge of the strong interaction with the following values (and their opposites): *red*, *green*, and *blue*. The free particles in the SM must be colour-neutral. The quantum numbers associated to the $SU(2)_L$ and $U(1)_Y$ gauge groups are called weak isospin I and weak hypercharge Y, respectively. They are related to the electrical charge Q by $Q = I_3 + \frac{Y}{2}$, where I_3 is the third component of the weak isospin. The Y in $U(1)_Y$ refers to the weak hypercharge, whereas the L index in $SU(2)_L$ signifies that the interaction only affects particles with left-handed chirality. The fundamental fermions that make up all of matter are all classified according to their quantum numbers and whether or not they interact with the strong force. From this point forward, the natural units will be used where the speed of light c and the reduced Planck constant $\hbar$ are equal to 1. In this system, momentum p, mass m and energy E all share the same units of GeV instead of GeV/c, GeV/c^2 and GeV, respectively. Electrical charge will also be given in units of the elementary charge e.

2.1.1 The fundamental fermions

Fermions are particles with half-integer spins that obey Fermi-Dirac statistics. There are twelve types of fundamental fermions fields (and their anti-particles) of spin 1/2 in the SM that make up the basic constituents of matter.

All fermions interact with the electromagnetic or weak forces. The fermions with left-handed chirality are written in doublets with weak isospin $I = 1/2$, whereas the right-handed fermions form singlets with $I = 0$. The fermions come in three families, where each family has the same quantum numbers and differ only in their masses.

Six fermions do not interact with the strong force and are called leptons. They form three lepton families, with one family each for the electron e, the muon μ and the tau τ. They all possess an electrical charge of -1. With their associated neutrinos of charge zero, the ν_e, ν_μ and ν_τ, they form the left-handed doublets:

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}. \tag{2.2}$$

The charged leptons interact with the electromagnetic and weak forces, whereas the neutrinos interact only via the weak force. The charged leptons and neutrinos have weak isospin I of $-1/2$ and $+1/2$, respectively.

In the SM, the neutrinos are massless, whereas the charged leptons are massive with $m_e \ll m_\mu \ll m_\tau$. However, the neutrino masses are known to be non-zero via the measurements of neutrino-flavor oscillations, indicating that the SM is not a complete description of particle physics.

The six remaining fundamental fermions are called quarks and interact with the electromagnetic, weak and strong forces. The left-handed doublets are grouped into three pairs of either up-type or down-type quarks. The up-type quarks, the up u, charm c and top t, have an electrical charge of $+2/3$ and a weak isospin $I = -1/2$. The down-type quarks, the down d, strange s and bottom b, have an electrical charge of $-1/3$ and a weak isospin $I = +1/2$. They form three families within the following left-handed doublets:

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix}. \tag{2.3}$$

Quarks interact with the strong force and carry a colour charge: either *red*, *green*, or *blue*. Due to colour confinement, quarks can not exist as free particles and only form mesons or hadrons that are colour-neutral.

In Feynman diagrams, leptons and quarks are represented with solid straight lines [34].

2.1.2 The fundamental bosons

Bosons are particles with integer spins that obey Bose-Einstein statistics. In the SM, fermions interact with each other by coupling to spin 1 gauge bosons. There are twelve fundamental gauge bosons in the SM that serve as the force carriers and mediate all subatomic interactions.

The $SU(3)_C$ colour symmetry group gives rise to 8 massless gauge bosons called gluons that mediate the strong force and couple only to quarks. The $SU(2)_L \times U(1)_Y$ electroweak symmetry group gives rise to the massless photon γ, which couples with any electrically charged particle, and the massive $W^{\pm}$ and Z^0 bosons that mediate the weak force. The $W^{\pm}$ bosons couple only to left-handed fermions and right-handed anti-fermions.

The spontaneous symmetry breaking of the $SU(2)_L \times U(1)_Y$ groups gives rise to the Higgs boson, a scalar boson of spin 0. This phenomenon gives mass to the W and Z bosons, while leptons gain their masses through Yukawa interactions with the Higgs boson.

In Feynman diagrams, photons and the W and Z bosons are represented with curvy lines, while gluons are drawn as spirals or spring-like lines.

2.1.3 Fields and gauge interactions

The Standard Model describes particles as quantum fields and uses the Lagrangian formalism to describe their propagation and interactions. It uses principles of symmetry, or properties of physical systems to remain unchanged under specific transformations, to describe the interactions between particles. The SM is a gauge theory and follows the principle of local gauge invariance.

The notation in the following section follows that of Ref. [34]. We will also shorthand Lagrangian $\mathcal{L}$ to mean the Lagrangian density.

For a freely propagating fermion ψ with mass m that satisfies the Dirac equation, the equations of motion can be derived from the Lagrangian

$$\mathcal{L}_{\text{Free}} = \bar{\psi}\Big(i\gamma^{\mu}\partial_{\mu} - m\Big)\psi. \tag{2.4}$$

The Lagrangian is invariant under global gauge transformations of the fermion field $\psi \to e^{i\alpha}\psi$. However, the underlying symmetry arguments of the SM require the Lagrangian to be invariant under local gauge transformation. In other words, α should be a function of space-time coordinates x: $\alpha \to \alpha(x)$.

Under this local gauge transformation, the Lagrangian of Equation (2.4) is no longer invariant and instead transforms as

$$\mathcal{L} \to \mathcal{L}_{\text{Free}} + i\bar{\psi}\gamma^{\mu}(\partial_{\mu}\alpha)\psi. \tag{2.5}$$

The solution is to modify the Lagrangian and derivative ∂_{μ} such that the additional term in $\partial_{\mu}\alpha$ is canceled. This is done by adding the gauge field A_{μ} that transforms as $A_\mu \to A_\mu - \frac{1}{e}\partial_\mu \alpha$, where e is a coupling constant. The derivative is also substituted with the covariant derivative such that $\partial_\mu \to D_\mu = \partial_\mu + ieA_\mu$.

The Lagrangian in Equation (2.4) then becomes

$$\mathcal{L} = \bar{\psi}\Big(i\gamma^\mu D_\mu - m\Big)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.6}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. At this point, nothing explicitly prohibits adding a mass term $\frac{1}{2}m' A_\mu A^\mu$ for the gauge field. However, $A_\mu A^\mu$ is not gauge invariant and thus requires the mass m' to be equal to zero for the Lagrangian to stay invariant. The Lagrangian is expanded to obtain

$$\mathcal{L} = \mathcal{L}_{\text{Free}} + e\bar{\psi}\gamma^\mu A_\mu \psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.7}$$

which is locally invariant under the local gauge transformation $\psi \to e^{i\alpha(x)}\psi$ and $A_\mu \to A_\mu - \frac{1}{e}\partial_\mu \alpha$. The Lagrangian can further be expanded into its four separate terms:

$$\mathcal{L}_{\text{QED}} = \underbrace{i\bar{\psi}\gamma^\mu \partial_\mu \psi}_{\text{Kinetic term of } \psi} - \underbrace{m\bar{\psi}\psi}_{\text{Mass term of } \psi} + \underbrace{e\bar{\psi}\gamma^\mu A_\mu \psi}_{\text{Interaction term}} - \underbrace{\frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{\text{Kinetic term of } A_\mu} . \tag{2.8}$$

This is the Lagrangian that describes the theory of quantum electrodynamics (QED). It relates the behaviours of the fermion fields ψ, the photon field A_μ and their couplings. The first and second terms of Equation (2.8) correspond to the kinetic and mass terms of ψ, the third corresponds to the interaction term between ψ and A_μ, and the fourth term is the kinetic term of A_μ. The coupling constant e is the fundamental unit of electric charge, which dictates the strength of the interaction, and $F_{\mu\nu}$ is the field tensor strength.

We have shown that the theory of QED arises from imposing local gauge invariance and creates interactions between the fermion fields ψ through the gauge boson field A_μ, or the photon. This general principle is one of the foundations of quantum field theory and can be extended to a variety of different symmetry groups. This example of the $U(1)_Y$ symmetry group rotation $\psi \rightarrow e^{i\alpha(x)}\psi$ is perhaps the simplest example in QFT. The Lagrangian for the theory of quantum chromodynamics is obtained by requiring a symmetry under the $SU(3)_C$ local gauge transformation and is presented in Section 2.1.4. Section 2.1.5 describes how the symmetry requirement of the $SU(2)_L \times U(1)_Y$ local gauge transformation produces the Lagrangian of the electroweak theory.

2.1.4 Quantum chromodynamics

The theory of quantum chromodynamics (QCD) describes the strong force and its interaction to quarks. It explains the interactions of particles charged under colour, analogous to QED and the behaviour of electrically charged particles. Whereas the transformative properties of the QED vector fields A_μ of Section 2.1.3 were easy to derive without prior knowledge of group theory, the transformation properties of the gluon fields can't easily be guessed and requires some knowledge of the $SU(3)$ symmetry group.

The Lagrangian arises from requiring the local gauge invariance of the $SU(3)_C$ transformation on quark colour fields q. The free Lagrangian of the quark fields is

$$\mathcal{L} = \bar{q}_j\Big(i\gamma^\mu\partial_\mu - m\Big)q_j, \tag{2.9}$$

where q_j represents the quark fields with colour charge $j = red$, *green* or *blue*. The local $SU(3)_C$ local gauge invariance transforms the quark fields as

$$q_j \rightarrow e^{ig_s\alpha\cdot\hat{\mathbf{T}}} q_j, \tag{2.10}$$

where α_a are numerical constants, and $\hat{\mathbf{T}}$ represents the eight generators of the $SU(3)$ group that are related to the Gell-Man matrices by $T^a = \frac{1}{2}\lambda^a$. The eight traceless 3×3 λ matrices form a linearly independent set. By definition they do not commute, satisfying the group algebra $[\lambda_a, \lambda_b] = if_{abc}\lambda_c$, where f_{abc} are the fixed structure constants of the $SU(3)$ Lie group algebra.

The local gauge invariance of Equation (2.9) under Equation (2.10) is conserved by introducing the covariant derivative and eight vector fields with transformation properties such that the $\partial_\mu\alpha_\mu$ terms cancel:

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + ig_s\frac{\lambda_a}{2}G^a_\mu \tag{2.11}$$

and

$$G^a_\mu \rightarrow G^a_\mu - \frac{1}{g_s}\partial_\mu\alpha^a - f^{abc}\alpha^b G^c_\mu, \tag{2.12}$$

where the G^a_μ are the eight gluon fields and g_s is related to the strong coupling constant α_s. The Lagrangian in Equation (2.9) becomes in its expanded form

$$\mathcal{L}_{\text{QCD}} = \underbrace{i\bar{q}_j\gamma^\mu\partial_\mu q_j}_{\text{Kinetic term of } q_j} - \underbrace{m\bar{q}_j q_j}_{\text{Mass term of } q_j} \underbrace{- g_s\ \bar{q}_j\gamma^\mu\frac{\lambda_a}{2}q_j G^a_\mu}_{\text{Interaction term}} - \underbrace{\frac{1}{4}\mathcal{G}^a_{\mu\nu}\mathcal{G}^{\mu\nu}_a}_{\text{Kinetic term of } G^a_\mu}, \tag{2.13}$$

where $\mathcal{G}^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f_{abc} G^b_\mu G^c_\nu$ is the field tensor strength. The third term of the Lagrangian represents the interaction term between the eight gluon fields and the three quark colour fields. The kinetic term of the gluon field, the fourth term in the Lagrangian, includes self-interactions between gluon fields through $g_s f_{abc} G^b_\mu G^c_\nu$ in $\mathcal{G}^a_{\mu\nu}$ and gives rise to gluon-gluon interactions in the SM. The gluons are massless, as demonstrated from the lack of quadratic terms in G^a_μ.

The theory of QCD offers a unique behaviour because it leads to the phenomena of QCD asymptotic freedom and colour confinement. Asymptotic freedom refers to the fact that the coupling constant α_s decreases to zero at a logarithmic rate as the energy scale of the interaction increases and for small distances. On the other hand, colour confinement relates to the large increase in the value of α_s at low energies and large distances. Colour confinement is the reason why quarks and gluons have never been observed directly in isolation: when attempting to separate two quarks, the potential energy increases so that it is energetically favourable to create another quark pair in a process called hadronization.

2.1.5 Electroweak theory

The Glashow-Weinberg-Salam model (GWS) of electroweak theory describes the unification of the electromagnetic and weak interactions through the $SU(2)_L \times U(1)_Y$ gauge group [29, 30]. The derivation of the $SU(2)_L \times U(1)_Y$ Lagrangian is more complex than both the $U(1)_Y$ and $SU(3)_C$ groups presented in Section 2.1.3 and Section 2.1.4, hence this looser approach in describing it.

The electroweak Lagrangian is constructed with a kinetic term and a potential term for a newly introduced complex scalar field ϕ, such that

$$\mathcal{L} = |D_\mu \phi|^2 + V(\phi), \tag{2.14}$$

where D_μ is the covariant derivative and $V(\phi)$ the potential. The scalar field ϕ is defined to transform as a doublet under $SU(2)_L$ and is symmetric under the $SU(2)_L \times$

$U(1)_Y$ gauge transformation

$$\phi \to e^{i\alpha^a \sigma^a/2} e^{i\beta/2}\phi, \tag{2.15}$$

where σ_i are the 2×2 Pauli matrices. In order for the Lagrangian to be invariant under the local gauge transformation, the covariant derivative must be

$$D_\mu = \partial_\mu - i\frac{g'}{2}B_\mu - i\frac{g}{2}W^j_\mu \sigma_j, \tag{2.16}$$

where B_μ and W^j_μ are the field tensor strength of the $U(1)_L$ and three $SU(2)_L$ gauge groups, respectively, while g and g' are coupling constants. The Lagrangian is then expanded and manipulated to simplify some terms before producing the electroweak (EW) Lagrangian

$$\mathcal{L}_{\text{EW}} = \underbrace{|\partial_\mu \phi|^2}_{\text{Kinetic term of }\phi} + \underbrace{\frac{g'^2 + g^2}{2} Z_\mu Z^\mu |\phi|^2}_{\text{Interaction term}} - \underbrace{\frac{g^2}{2} W^+_\mu W^{-\mu} |\phi|^2}_{\text{Interaction term}} + \underbrace{V(\phi)}_{\text{Potential term}}, \tag{2.17}$$

where the gauge field tensor strengths Z_μ, W^+_μ and W^-_μ correspond to the neutral Z boson and the charged W bosons. In this description of the EW Lagrangian, the Z and W bosons are massless due to the lack of quadratic terms in Z_μ and $W^\pm_\mu$. In a similar fashion to the Lagrangian of the electromagnetic interaction, it is impossible to add mass terms for the three bosons while preserving local gauge invariance. However, the Z and W bosons have been measured to have mass, indicating that the universe does not follow a complete $SU(2)_L \times U(1)_Y$ symmetry.

2.1.6 Spontaneous electroweak symmetry breaking

Whereas the unified model of electroweak interactions introduces the W and Z bosons and maintains the invariance of the Lagrangian to local $SU(2)_L \times U(1)_Y$ gauge transformations, the masses of the three associated gauge bosons are explained through spontaneous electroweak symmetry breaking. The $SU(2)_L \times U(1)_Y$ gauge symmetry is spontaneously broken through the Higgs mechanism [35–37] by setting the potential for the $SU(2)_L$ doublet ϕ to

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \frac{\lambda}{2}(\phi^\dagger \phi)^2, \tag{2.18}$$

where μ and λ are coupling constants.

For the potential to have an absolute minimum, λ has to be positive. If μ^2 is negative, then the minimum, or vacuum ground state, is at $\phi = 0$ and the vacuum state is symmetric. However, if the value of μ^2 is positive, then the potential is at a minimum for $\phi = \sqrt{\mu^2/\lambda}$, or at a vacuum expectation value $v = \sqrt{\mu^2/\lambda}$, as visualized in Figure 2.2.

The $SU(2)_L \times U(1)_Y$ symmetry can be spontaneously broken by choosing to express the field ϕ as oscillations around the vacuum expectation value v, such that

$$\phi = \begin{pmatrix} 0 \\ v + h \end{pmatrix}, \tag{2.19}$$

where h represents the Higgs boson field. By substituting ϕ into the EW Lagrangian of Equation (2.17), it becomes

$$\mathcal{L}_{\text{EW}} = (\partial_\mu h)^2 + \frac{g'^2 + g^2}{2} Z_\mu Z^\mu (v+h)^2 - \frac{g^2}{2} W_\mu^+ W^{-\mu} (v+h)^2 + V(v+h). \tag{2.20}$$

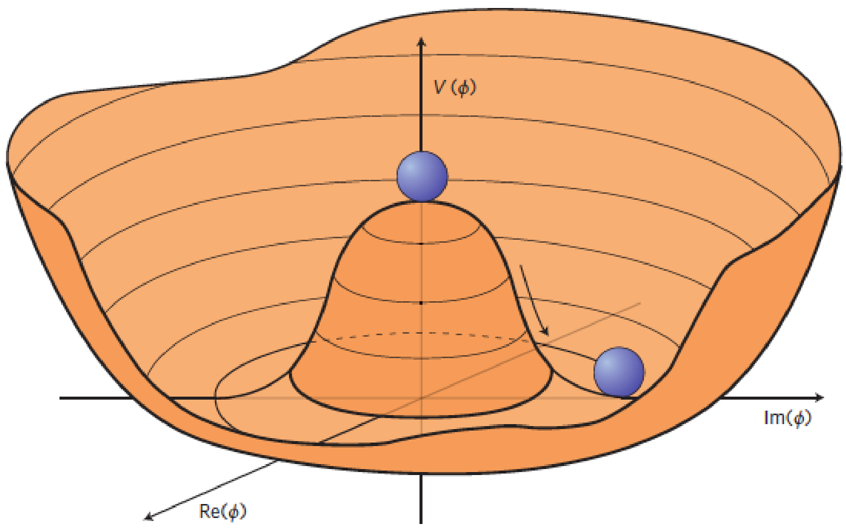


Figure 2.2: Sketch of the scalar potential $V(\phi)$ with $\mu^2 > 0$ and $\lambda > 0$ leading to the spontaneous symmetry breaking of the SM [38].

By expanding the second and third terms, interaction terms between the W, Z and Higgs bosons appear in addition to the following mass terms:

$$v^2 \frac{g'^2 + g^2}{2} Z_\mu Z^\mu \tag{2.21}$$

and

$$v^2 \frac{g^2}{2} W_\mu^+ W^{-\mu}. \tag{2.22}$$

Hence, the W and Z bosons have both gained a mass term with $M_Z = \frac{v}{2}\sqrt{g'^2 + g^2}$ and $M_W = \frac{vg}{2}$. The leptons of the SM do not gain their gain mass through the same symmetry breaking mechanism and instead acquire their mass through Yukawa couplings to the Higgs field.

2.1.7 Proton–proton collisions at hadron colliders

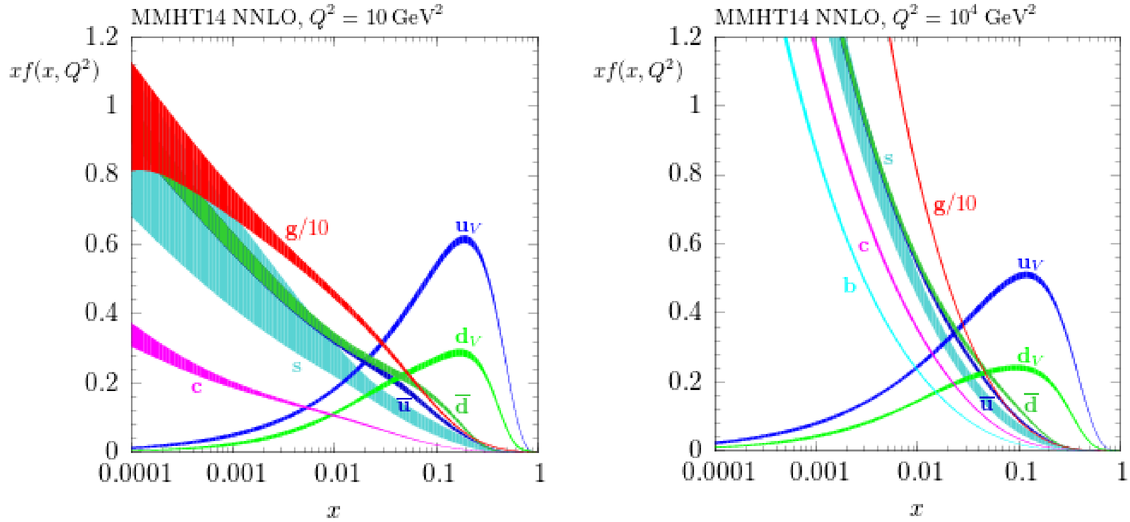
The description of the Standard Model so far described through its Lagrangian formulation does not provide the necessary tools to calculate the cross sections in proton–proton collisions in hadron colliders such as the LHC. The framework so far presented allows for the description of $X \to Y$ physics processes but this is only a small part of the general puzzle of hadron physics.

Hadrons are bound states of an odd number of quarks. The proton is a bound state of the uud valence quarks that interact via the exchange of gluons. Through the interactions of these gluons, short-lived $q\bar{q}$ pairs of all flavors are created within the proton, which are known as sea quarks. In proton–proton collisions, any of these constituents of the proton, called partons, may be involved in the initial state of the hard-scatter interaction.

Due to the composite nature of hadrons, the cross section for two initial-state protons p_1 and p_2 to create the process $p_1p_2 \to X$ is given by

$$\sigma_{p_1p_2\to X} = \sum_{a,b}\int_0^1 dx_1 \int_0^1 dx_2 f_a(x_1,Q^2) f_b(x_2,Q^2)\sigma_{p_ap_b\to X}(x_1x_2s,\alpha_s,\mu_R,\mu_F), \quad (2.23)$$

where x_1 and x_2 are the proton momentum fractions carried by partons 1 and 2, respectively. The Parton Distribution Function (PDF) given by $f_p(x,Q^2)$ describes the probability of a given parton p to hold a fraction x of the proton momentum. PDFs are determined experimentally using data and depend on the transferred momentum squared Q^2 in the interaction. An example of a PDF obtained from the MMHT14 PDF set [39] at different values of Q^2 is shown in Figure 2.3. The cross section of the given $p_ap_b \to X$ process for partons a and b is given by $\sigma_{p_ap_b\to X}$ and depends on the



(a) (b)

Figure 2.3: Proton Parton Distribution Function (PDF) from the MMHT14 PDF set (a) at $Q^2 = 10$ GeV^2 and (b) at $Q^2 = 10^4$ GeV^2 [39].

partonic centre-of-mass energy $\hat{s} = x_1 x_2 s$, the strong coupling constant α_s, and the QCD renormalization and factorization scales μ_R and μ_F.

The renormalization and factorization scales μ_R and μ_F are introduced in QCD cross-section calculations to absorb the ultraviolet and infrared divergences at the high and low energy limits, respectively. In the high energy limit, divergences occur from virtual loops and are absorbed by the renormalization scale μ_R such that the strong coupling constant α_s becomes a function of μ_R: $\alpha_s \to \alpha_s(\mu_R)$. The low energy limit infrared divergences are removed through the factorization scale μ_F below which the hard-scatter processes are absorbed by the PDF instead of being included in the partonic cross section $\sigma_{p_a p_b \to X}$. Stated more directly, partons emitted with transverse momentum higher than μ_F are considered in the partonic cross section and those with transverse momentum lower than μ_F are treated as being a part of the proton structure. By doing so, perturbative QCD calculations are not performed for low energy interactions, a regime outside their range of validity. Partonic cross sections are further discussed in Chapter 4.

2.2 The history of the Z boson

2.2.1 The era of discovery

The CERN Proton Synchrotron (PS) accelerator was designed and built specifically to search for the exchange of intermediate W and Z bosons. The protons of the PS were used to generate beams of muon antineutrinos for fixed target experiments. The first observation of neutral current interactions mediated through the exchange of a Z boson was announced in 1972 in the $\bar{\nu}_\mu + p \to \mu^+ + \Lambda^0$ [40] process and again with the $\bar{\nu}_\mu + e^- \to \bar{\nu}_\mu + e^-$ process in 1973 [41]. These observations were not immediately hailed as the discovery of the Z boson, but rather challenged for the following years. A decade later, the CERN accelerator complex was upgraded to be a 540 GeV proton–antiproton ($p\bar{p}$) collider, called the SPS. The SPS was successful in its goal of observing the direct production of W and Z bosons when, in 1983, the UA1 [42] and UA2 [43] collaborations both measured evidence of the vector bosons [44–47].

Prior to the discovery of the W and Z bosons, the GWS model of electroweak interactions unified the electromagnetic and weak forces. It predicted the masses of the W and Z bosons, which are related to one another, to be

$$M_W^2 = \frac{\sqrt{2}e^2}{8G_F}\frac{1}{\sin^2\theta_w} \tag{2.24}$$

and

$$M_Z = \frac{M_W}{\cos\theta_w}, \tag{2.25}$$

where e is the fundamental unit of charge, G_F the Fermi constant and θ_W the Weinberg angle. With the known values of G_F, e and θ_W, the W and Z bosons masses were estimated to be around $M_W \approx 70$ GeV and $M_Z \approx 82$ GeV.

To produce the Z boson in a $p\bar{p}$ collision and measure it precisely, a quark from one proton must interact and annihilate with an antiquark from the antiproton to produce the $q\bar{q} \to Z \to \ell^+\ell^-$ process, where q is most likely an *up* valence quark, and $\ell = e$ or μ. The partonic centre-of-mass energy is $\sqrt{\hat{s}} = (x_1 p_1 + x_2 p_2) = \sqrt{x_1 x_2 s}$, where x_1 and x_2 are the momentum fraction of the two quarks, p_1 and p_2 the momentum of the proton and antiproton, and $\sqrt{s}$ the centre-of-mass energy of the $p\bar{p}$ collision. To produce an on-shell Z boson, the partonic centre-of-mass energy must be $\sqrt{\hat{s}} \approx M_Z$ so that $\sqrt{s} = M_Z/\sqrt{x_1 x_2}$. At the time of the design of the SPS physics program, it was known from proton PDF measurements that each valence quark held approximately one sixth of the momentum of high-energy protons for a total of approximatly half the momentum, and the other half was carried by gluons [48]. As a consequence, the centre-of-mass energy of the $p\bar{p}$ collision must be $\sqrt{s} \approx 6M_Z$ or about 480 GeV. Thus, it was expected that the designed CERN SPS $p\bar{p}$ collider with $\sqrt{s} = 540$ GeV would have access to the direct production of the Z boson.

In 1983, the UA1 and UA2 experiments observed the $Z \to e^+e^-$ process with a final selection of four and three events, respectively. They produced the Z mass measurement of $M_Z^{UA1} = 95.2 \pm 2.5$ GeV [44] and $M_Z^{UA2} = 91.9 \pm 1.3$ (stat.) $\pm$ 1.4 (syst.) GeV [45]. Both the measured values are consistent with the current state-of-the art measurement of the Z boson mass. Figures 2.4 and 2.5 show the invariant dilepton mass distribution of the UA1 and UA2 results, respectively, for the different steps of the analyses, where the bottom most pads correspond to the final candidate selections. These measurements solidified the existence of the Z boson and the validity of the GWS model of electroweak interactions.

Almost 40 years after its discovery, the Z boson is still the subject of many

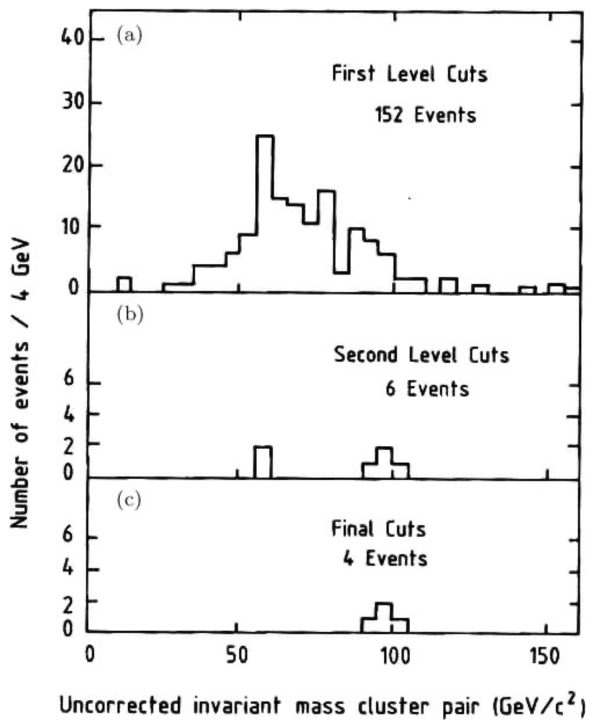


Figure 2.4: Invariant dilepton mass distribution leading to the measurement of $Z \to e^+e^-$ by the UA1 experiment. The bottom pad shows the four event candidates after the final selection cuts [44].

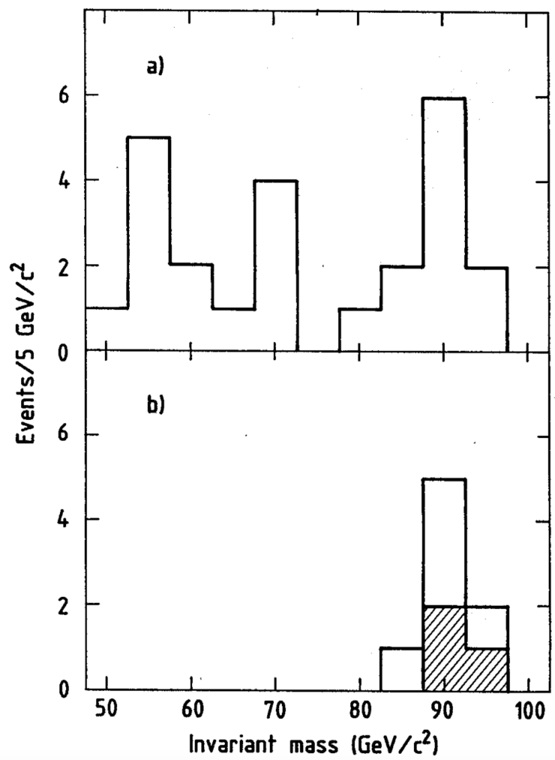


Figure 2.5: Invariant dilepton mass distribution leading to the measurement of $Z \to e^+e^-$ by the UA2 experiment. The three event candidates after the final selection cuts are represented by the hatched area of the bottom pad [45].

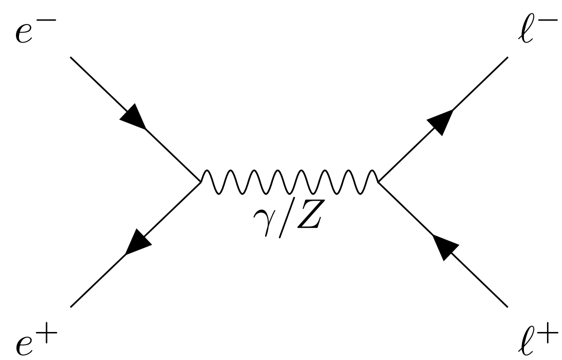


Figure 2.6: Leading-order s-channel Z boson mediated lepton-pair production at the LEP collider. For e^+e^- final states, the t-channel diagram also contributes where the electrons scatter against the Z boson.

studies and still takes an important role in modern particle physics research.

2.2.2 The era of precision

The era of precision physics for the Z boson came shortly after its discovery, when in the 1980's the Large Electron-Positron (LEP) collider was built at CERN. It is still the largest lepton collider ever built, reaching a maximum centre-of-mass energy of 209 GeV. Four experiments were placed around the LEP facility: ALEPH [49], DELPHI [50], L3 [51] and OPAL [52]. Since the colliding leptons are not composite particles like hadrons, where each parton carries an indeterminate fraction of the hadron momentum, the entirety of the centre-of-mass energy is carried by the two interacting leptons. The leptonic Z-boson decay channels were studied, corresponding to the $e^+e^- \to Z \to \ell^+\ell^-$ process with $\ell = e$ or μ, and resulted in very clean collision products and leading to very precise measurements. The leading-order (LO) representative Feynman diagram for the process studied at LEP are shown in Figure 2.6.

From the beginning of its operation in 1989 until 1995, LEP operated at a centre-of-mass energy set primarily to match the Z boson mass within a small window of a few GeV. This allowed the different LEP experiments to collect approximately

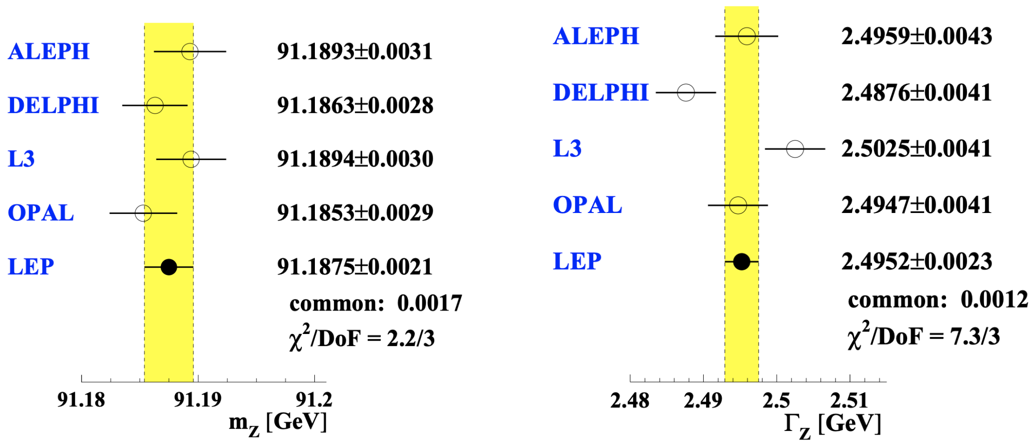
Year	Centre-of-mass energy range [GeV]	Integrated luminosity [pb^{-1}]
1989	88.2 – 94.2	1.7
1990	88.2 – 94.2	8.6
1991	88.5 – 93.7	18.9
1992	91.3	28.6
1993	89.4, 91.2, 93.0	40.0
1994	91.2	64.5
1995	89.4, 91.3, 93.0	39.8

Figure 2.7: Summary of the yearly integrated luminosity provided by LEP to its experiments between 1989–1995 alongside the different centre-of-mass energies [53].

twenty million Z boson events, leading to several high precision measurements of the Z boson properties. These measurements are still unmatched in precision to this day. The yearly amount of data provided by LEP to its experiments for the years of 1989–1995 are provided alongside the different centre-of-mass energies in Figure 2.7. A precision on the centre-of-mass energy of approximately 2 MeV was achieved by LEP and allowed for the very precise sweep around the Z boson mass resonance.

The measurements performed at LEP pushed the limits of precision results in high-energy physics. The most striking measurement of the Z boson properties is that of its mass $M_Z = 91.1875 \pm 0.0021$ GeV, corresponding to a precision of 23 parts-per-million. The equally impressive measurement of the Z boson decay width produced $\Gamma_Z = 2.4952 \pm 0.0023$ GeV, corresponding to a precision of approximately 1 part-per-thousand. The measurements of M_Z and Γ_Z are summarized in Figure 2.8, where the measurements from the four LEP experiments are presented along with their combination.

The LEP experiments performed precision measurements not only on the Z boson mass and decay width, but also on many other different key properties of



(a) (b)

Figure 2.8: Measurements from the LEP experiments [53] (a) of the Z boson mass and (b) of the total Z boson decay width. The combined measurements from the four experiments are also shown.

the SM. The wide array of measurements performed by these experiments set the standards of today's precision measurements, and will most likely not be beat in precision for the foreseeable future. Although the LHC provides an abundance of data at high centre-of-mass energies, its inherent design as a hadron collider makes the experimental conditions difficult for precision measurements on the same level as LEP. Instead, the LHC opens the door to different types of Z boson measurements with its high centre-of-mass energies at 13 TeV and large amount of provided collision data. It allows for the investigations of processes corresponding to Z bosons produced in association with jets and other rare Z production mechanisms.

2.3 Z boson in association with jets at the LHC

The LHC offers the perfect experimental conditions to test both QCD and EW interactions through processes involving the production of a Z boson in association with jets. Through the large inclusive production cross section of $\sigma_{pp\rightarrow Z+X}$, the LHC pro-

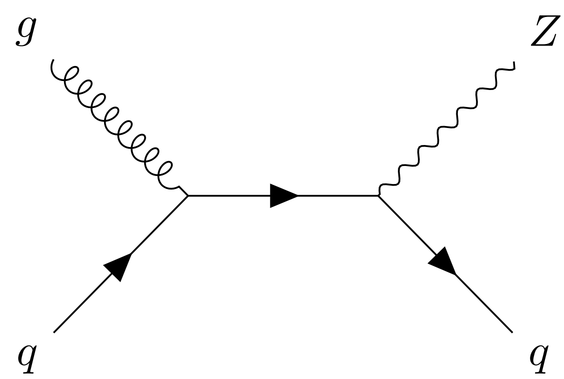


Figure 2.9: Leading-order Feynman diagram for the production of a Z boson in association with jets. The leftmost qgq vertex is described by QCD, while the rightmost qqZ vertex involves the EW interaction.

duces a very large amount of data that can be used to carefully study the properties of the Z boson, the proton PDFs and myriad of different Z+jets processes. The leading order Feynman diagram describing the Z+1 jet production is shown in Figure 2.9.

The production of Z+jets offers a powerful platform to test QCD and EW interactions because it provides a near inexhaustible amount of information about the SM physics at the LHC. Figure 2.10 shows an example of a higher-order Z+jets process where the Z boson is produced in associated with several jets. This diagram can be described as a leading-order Z+1 jet process with several higher-order corrections in the form of QCD gluon radiation. Every such quark and gluon radiation provides a further test of perturbative QCD (pQCD) but also suppresses the process by a factor of α_s. The proton PDFs can be studied through the two initial state partons, whereas the EW interactions can be probed via the vertices around the Z boson.

The Z+jets process often constitutes a non-negligible background in several measurements and in searches for new physics [14–18]. In these studies, the kinematic properties of jets, including their transverse momentum, masses and multiplicity, help separate the signal of interest from a multitude of backgrounds. The precise

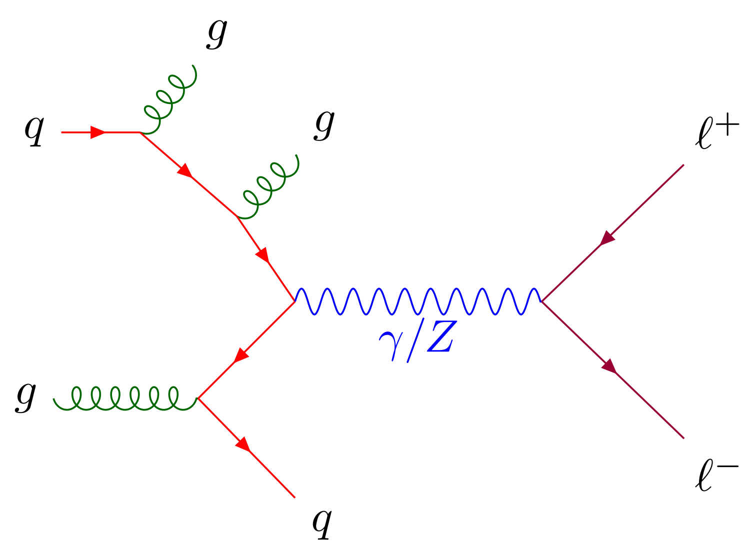


Figure 2.10: Feynman diagram for the production of a Z boson in association with jets. The figure shows the sensitivity of Z+jets processes to higher initial energy states through additional radiation above the leading-order Z boson production. The outgoing quark and gluons further hadronize and shower to become jets, while the Z boson decays to two opposite-charge leptons. The resulting final state consists of jets and two leptons.

understanding of Z+jets, including the relationships between the Z boson and the jets, is an essential tool to exploit the full potential of the data provided by the LHC.

2.3.1 State-of-the-art measurements

The Z+jets process has been extensively studied since the first construction of high centre-of-mass energies hadron colliders, starting with the Tevatron Collider in 1983. Measurements of differential cross sections of observables in inclusive Z+jets production have been performed at many centre-of-mass energies in pp collisions, such as for $\sqrt{s} = 1.96$ TeV [54, 55], and later at the LHC for $\sqrt{s} = 7$ TeV [56–59], $\sqrt{s} = 8$ TeV [60–62] and $\sqrt{s} = 13$ TeV [20, 63, 64].

Measurements of inclusive Z+jets production are important, but do not illustrate the intricacies of specific phenomena involving the Z boson. The Z+b jets measurements [6, 65] provide important information on the relationship between the Z boson and the b quark. They provide invaluable insight about the proton PDF of heavy quarks in the case of $Z + 1$ b-quark and for gluon splitting $g \rightarrow b\bar{b}$ in $Z + 2$ b-quark events. These measurements can be used in subsequent analyses that use myriad of W/Z+jets measurements to determine the proton PDF [66], amongst other key SM quantities.

Other processes, such as Vector-Boson-Fusion mediated Z+2 jets production [7], provide stringent tests to the SM by measuring very rare interactions in which the couplings between the multiple gauge bosons can be carefully investigated. Measurements involving several W and Z bosons in the final state [67–69] also provide important information on gauge boson coupling constants and are an important source of backgrounds in several analyses such as those involving the Higgs boson.

The ATLAS collaboration has performed a myriad of measurements on different processes around the production of W and Z bosons, all of which have found to be consistent with the predictions of the SM, as shown in Figure 2.11.

2.3.2 Z boson in association with high-p_{T} jets

The measurements of the production of Z+high-p_{T} jets serve as benchmarks for Monte Carlo predictions since the process is sensitive to higher-order pQCD, next-to-leading order EW corrections, and their interplay [10–13]. Previous publications [21, 22] have highlighted the difficulty of contemporary Monte Carlo predictions to model the Z+high-p_{T} jets phase space, and especially in the cases where the Z boson and a jet are nearby each other such that their angular separation is small.

Measurement of the production of a Z boson in association with jets with $p_{\mathrm{T}} \geq$ 100 GeV using the full ATLAS Run–2 dataset are performed for the first time and sum-

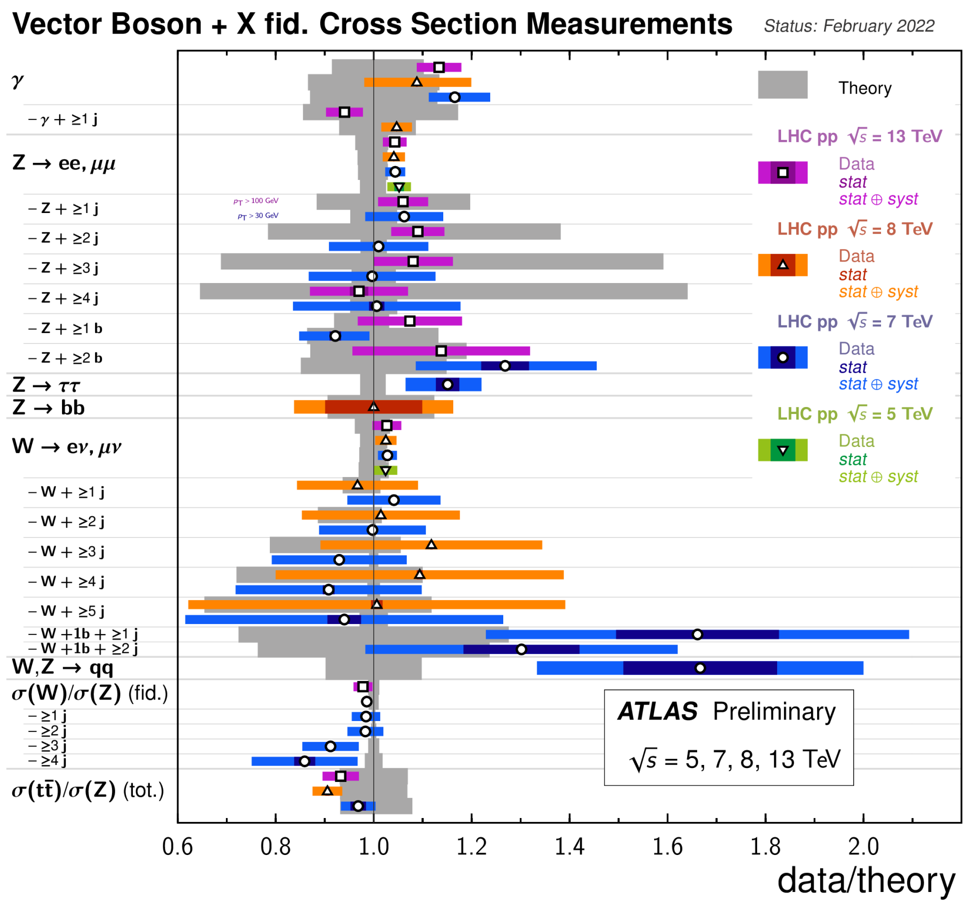


Figure 2.11: Ratios of the data over the predictions of several different vector boson (W, Z, γ) production cross sections measured by ATLAS [70].

marized in this thesis, unlocking the rare and extreme phase space of Z+high-p_{T} jets production. The expected production cross section for Z+high-p_{T} jets where the Z boson decays to a charged lepton pair is $\mathcal{O}(14)$ picobarns. A further selection cut is applied so that the leading jet p_{T} is greater than 500 GeV ($p_{\mathrm{T},j1} \geq 500$ GeV). This selection enhances two processes of great interest: 1) that of the Z+1 jet production where the Z boson recoils against the only jet such that they are measured to be back-to-back, and 2) that of the radiation of a real on-shell Z boson from dijet events. In the case of the real Z-boson radiation in dijet events, due to kinematic rules, the Z boson is emitted nearby the jet so that their angular separation is small and they are measured to be collinear. The representative Feynman diagrams of these Z+1 jet and Z+2 jets processes are shown in Figure 2.12. One key observable for these processes is the angular separation between the Z boson and the nearest jet $\Delta R^{\mathrm{min}}_{Z,j}$. The measurements of Z+high-p_{T} jets performed so far at the LHC [21, 22, 71] studied the correlations between the Z boson and the jets using partial Run–2 datasets. This limits the extent of the accuracy of the Z+high-p_{T} jets study due to low statistics. The use of the full Run–2 dataset in this analysis unlocks more extreme phase spaces and allows for the first time the complete study of the individual Z+1 jet and Z+2 jets processes detailed above. From this point forward, Z+high-p_{T} jets is used exclusively to refer to this analysis where all jets possess $p_{\mathrm{T}} \geq 100$ GeV.

The high-p_{T} Z+1 jet process is of great interest due to its sensitivity to virtual next-to-leading order (NLO) EW corrections, represented by the Feynman diagrams of Figure 2.13. The virtual NLO EW correction is described with an exchange of an internal Z boson line between fermion lines, as shown in Figure 2.13b. These corrections contribute negatively to the Z+1 jet cross section and grow as the square of the transferred momentum Q^2 by the virtual Z boson. In very energetic Z+1 jet

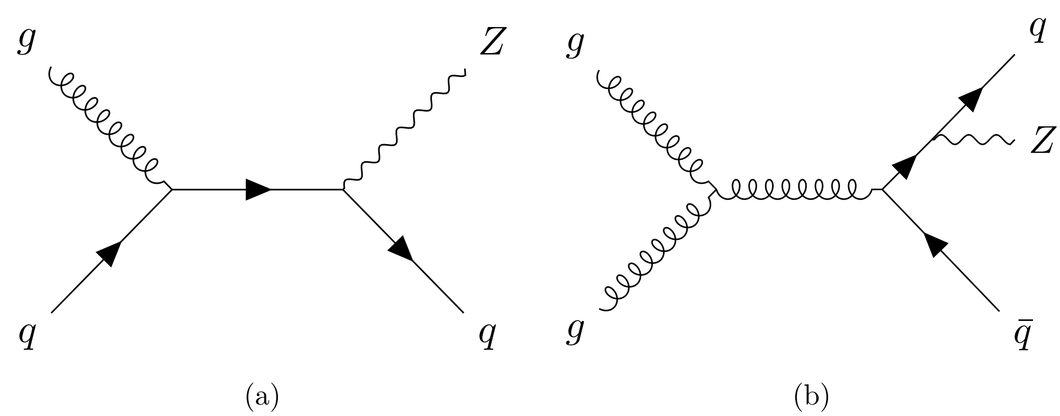
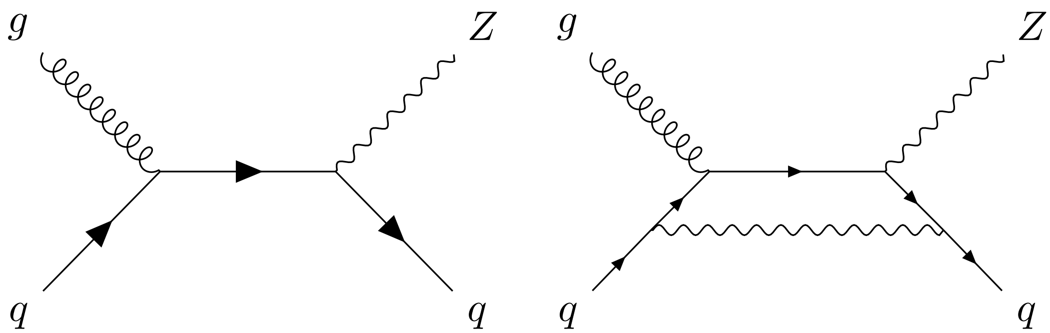


Figure 2.12: Representative Feynman diagrams for the production of a Z boson in association with high-p_{T} jets. (a) The Z+1 jet events are expected to populate the *back-to-back* region where the Z boson is balanced against a single high-p_{T} jet. (b) In dijet events, the Z boson is expected to be radiated from the quark leg, with kinematics leading to small values of the angular distance between the Z boson and the closest jet, $\Delta R^{\mathrm{min}}_{Z,j}$, and therefore populating the *collinear* region.

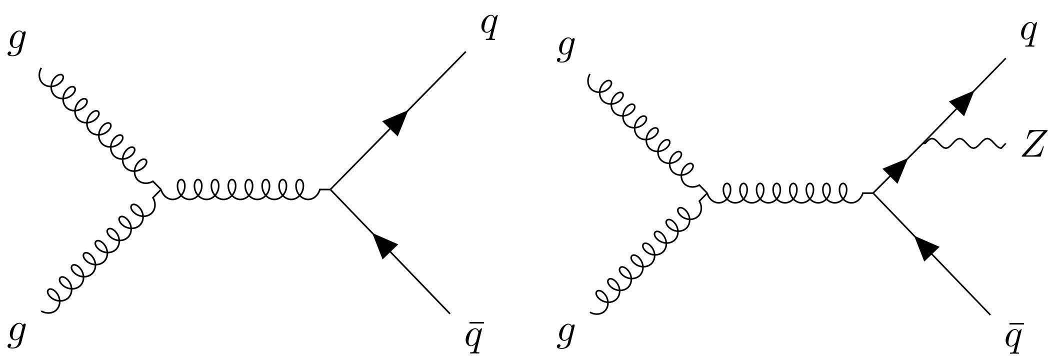
events, such as those where p_{T}(Z) and p_{T}(jet) approach one TeV, the virtual EW corrections can reduce the cross section by as much as 20%. The phase space with such a high energy Z boson and jet has so far been statistically limited and only recently has more precise measurements been accessible with the full LHC Run–2 dataset. Recent studies performed by ATLAS have outlined the possible effects and mismodelling of these virtual EW corrections [72], motivating the need of precise measurements of Z+high-p_{T} jets using the full Run–2 dataset.

The measurement of high-p_{T} Z+2 jets probes the effect of high-energy EW corrections to QCD processes, represented by the Feynman diagrams of Figure 2.14. The EW corrections take the form of real Z boson radiation from an in- or outgoing quark leg, as shown in Figure 2.14b. Naively, the emission of the Z boson in dijet events is suppressed by a factor of α_{EW}, the electroweak coupling constant, due to the additional EW vertex. However, the presence of Sudakov logarithmic factors in collinear emissions [73–78] of the Z boson instead lead to corrections going



(a) LO Z+1 jet production (b) NLO virtual EW correction

Figure 2.13: Representative Feynman diagrams for (a) the LO production of Z+1 jet and (b) the NLO virtual EW correction to Z+1 jet production. The NLO virtual EW corrections contribute negatively to the cross section and increase in size with Q^2 of the virtual Z boson.



(a) Dijet (b) Z+2 jets

Figure 2.14: Representative Feynman diagrams for (a) dijet production and (b) Z+2 jets production. The Z+2 jets diagram corresponds to the final state radiation of the Z boson from a quark in dijet events. When compared to the dijet production process, the Z+2 jets process is suppressed by an order of α_{EW} due to the presence of the additional electroweak vertex.

as $\alpha_{\text{EW}}\ ln^2\left(\frac{p_{\text{T},j1}}{M_Z}\right)$ such that at high jet p_{T}, the real radiation of a Z boson from dijet events becomes a significant contributor to the Z+jets production cross section. Using these corrections, the expected ratio of the cross sections between the Z+2 jets and Z+1 jet diagrams in Figures 2.12b and 2.12a, respectively, equals $\alpha_s\ ln^2\left(\frac{p_{\text{T},j1}}{M_Z}\right)$. Hence, for large enough jet p_{T}, *e.g.* $p_{\text{T},j1} \geq 500$ GeV, the Z+2 jets and Z+1 jet processes have similar cross sections and can be measured simultaneously. Recent analyses highlight the challenge of MC predictions to model this process in W/Z+jets events [21, 22], further motivating this measurement.

In summary, this chapter aims to give an introduction of the relevant physics content in the SM, provides a historical review of Z+jets production and motivates this Z+high-p_{T} jets analysis. More details about the experimental apparatus are given in the next chapter as the logical step towards the description and execution of the analysis strategy for the cross-section measurements presented in this thesis.

Chapter 3

The ATLAS experiment at the Large Hadron Collider

In this chapter, the experimental apparatus used for the measurements of this analysis is introduced. Located at the European Organization for Nuclear Research (CERN), the Large Hadron Collider is the world's largest subatomic particle accelerator. It accelerates protons to nearly the speed of light to observe the results of their collisions. Four experiments are situated along the collider's 27 km circumference ring, including the ATLAS detector. The Large Hadron Collider accelerator complex and the ATLAS detector are described in Sections 3.1 and 3.2, respectively.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the world's largest and most energetic particle accelerator. As a hadron–hadron collider, it is capable of accelerating and colliding protons or heavy ions. This analysis only discusses and considers proton–proton (pp) collisions. The LHC spans a 27 km circumference ring located approximately

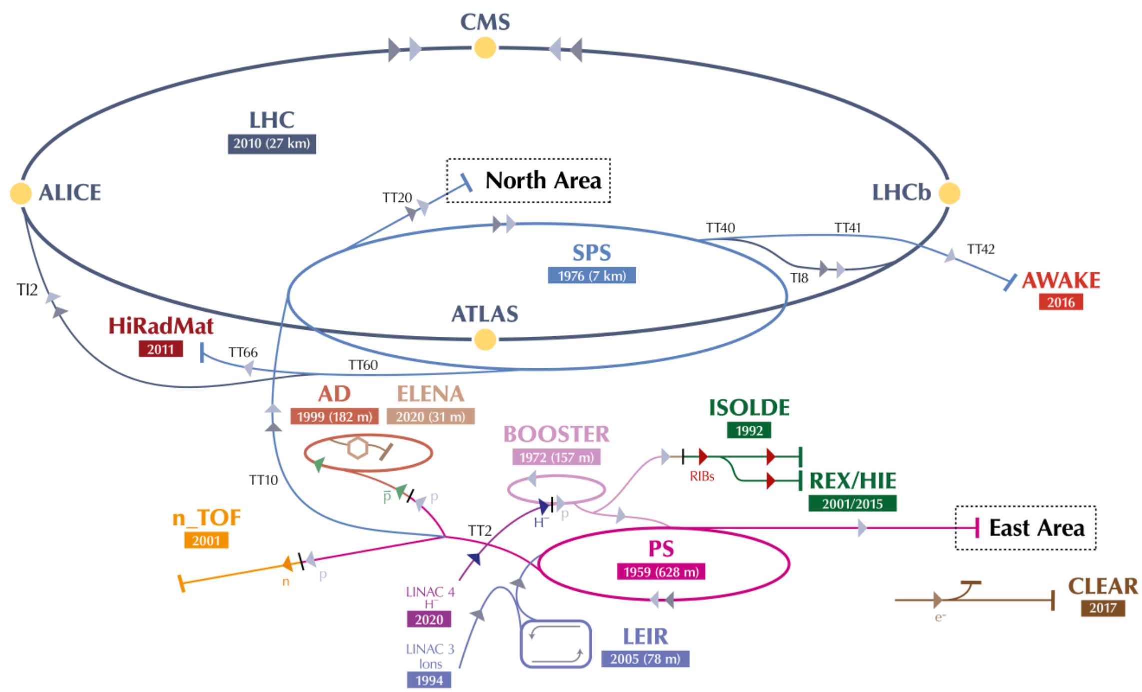


Figure 3.1: Diagram of the CERN accelerator complex. The LHC is the largest ring of a complex chain of particle accelerators. The four interaction points are shown as yellow circles [79].

100 m underground at the main CERN facility. It is the final step in a chain of five different subatomic particle accelerators. Protons are produced by ionizing hydrogen gas, accelerated by four different accelerators to an energy of 450 GeV and finally injected into the LHC ring. The LHC further accelerates two proton beams to energies above a few TeV in opposite directions before colliding them at four interaction points distributed along its ring. At each interaction point, a detector is housed that records the collisions: ATLAS, CMS, LHCb and ALICE. A diagram of the LHC within the CERN accelerator complex is shown in Figure 3.1.

The LHC was designed to accelerate proton beams to a centre-of-mass energy of $\sqrt{s} = 14$ TeV with an instantaneous luminosity of $L = 10^{34}$ cm^{-2}s^{-1}. At these energies, the protons travel at speeds just a few m/s slower than the speed of light.

Protons are not accelerated individually and instead are packed into bunches. The bunches contain approximately 10^{11} protons each and fill the LHC accelerator, separated in time by only 25 ns, corresponding to a distance of approximately 7.5 m. The luminosity depends on several variables, including the number of protons in each bunch and the intensity of the beams.

During the second stage of its physics program between 2015 and 2018, so called Run–2, the LHC operated at a centre-of-mass energy of $\sqrt{s} = 13$ TeV and achieved a peak instantaneous luminosity of $L = 2.1 \times 10^{34}$ cm^{-2}s^{-1} [80, 81]. During this time, the ATLAS detector collected 147 fb^{-1} from the 156 fb^{-1} of pp collisions delivered by the LHC. Of the data collected by ATLAS, 139 fb^{-1} were deemed to be of good quality for physics analyses, as shown in Figure 3.2. As a consequence of the high luminosity and large pp cross section, several pp interactions occur in a single bunch crossing. During the LHC Run–2, the average number of interactions per bunch crossing, $\langle\mu\rangle$, was 33.7. The distribution of the mean number of interactions per bunch crossing is shown in Figure 3.3.

3.2 The ATLAS detector

ATLAS is a general purpose detector designed to record and reconstruct the pp collisions produced at the LHC. It is built following a symmetric cylindrical configuration centred around the interaction point (IP), providing an almost 4π solid angle coverage. It consists of several subdetectors designed to optimize the detection of the particles produced in the pp collisions. Closest to the beamline, the inner detector is used to reconstruct the trajectory and momentum of charged particles in the region nearest to the IP. Electromagnetic and hadronic calorimeters surround the inner detector and capture photons and electrons, and hadrons, respectively. The inner

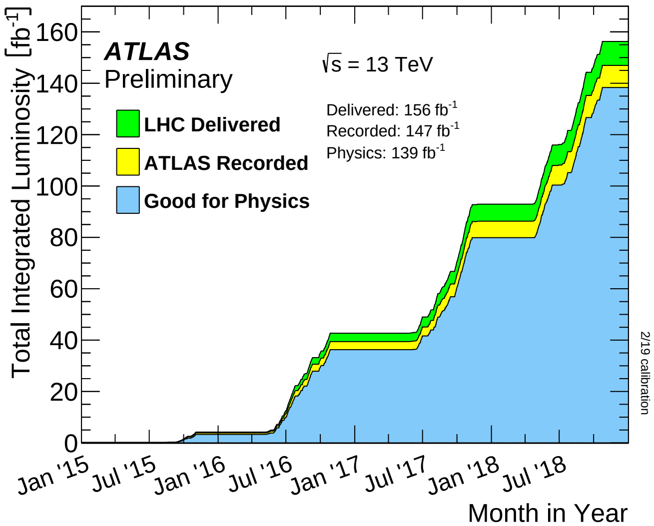


Figure 3.2: Cumulative integrated luminosity versus time delivered by the LHC in green and recorded by ATLAS in yellow. The data suitable for physics analyses are also shown in blue. The data corresponds to stable beams for *pp* collisions at 13 TeV centre-of-mass energy during Run–2 [1].

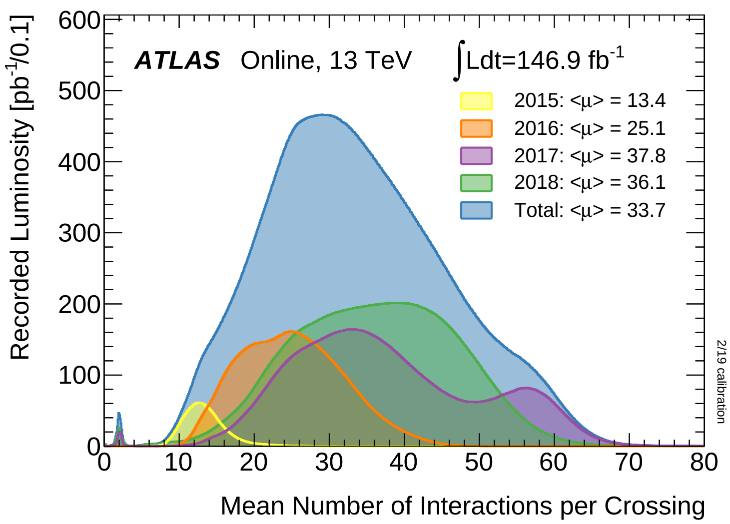


Figure 3.3: Distribution of the mean number of *pp* interactions per bunch crossing, μ, in *pp* collisions collected at ATLAS during the LHC Run–2 [1].

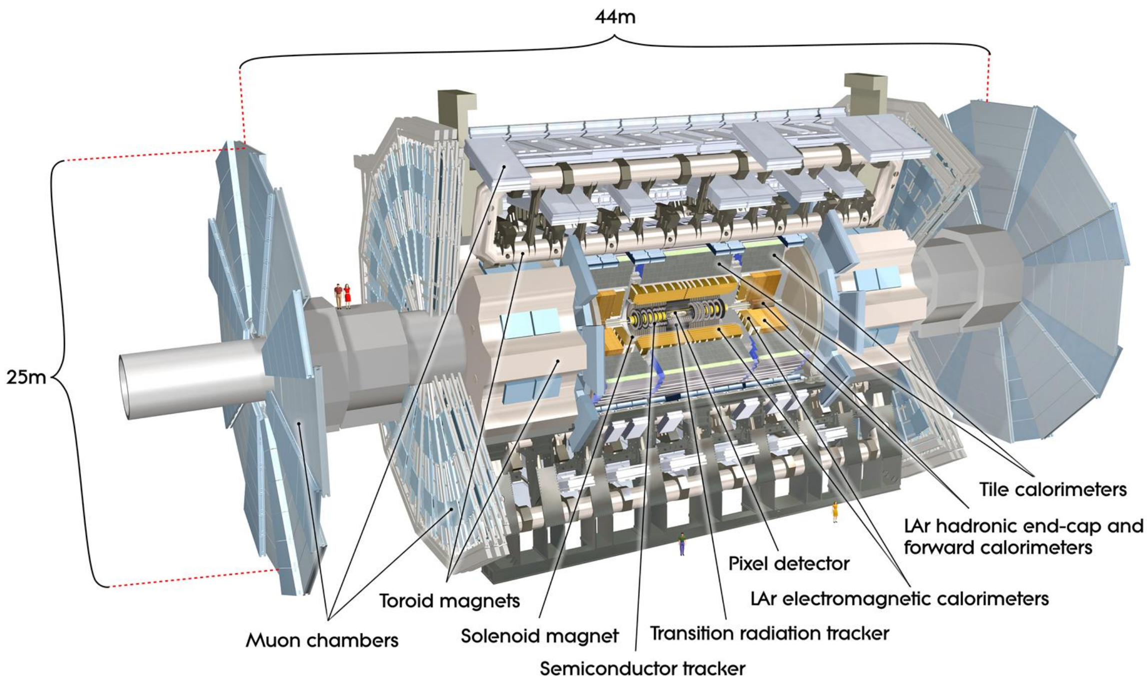


Figure 3.4: Diagram of the different detector subsystems of the ATLAS detector [82].

detector and calorimeters are enclosed by the muon spectrometer, which is designed to reconstruct the passage of muons. The subdetector systems are arranged into a central barrel and two end-cap regions. A diagram representing the ATLAS detector and its subsystems is shown in Figure 3.4. A two-level trigger system is used that selects events to be read out and saved to disk.

The ATLAS experiment uses a right-handed coordinate system measured from the IP. The z-axis is defined along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, while the y-axis points upwards. The azimuthal and polar angles ϕ and θ are measured around the z-axis and the x-axis, respectively. The pseudorapidity is defined as $\eta = -\ln\tan(\theta/2)$. When dealing with massive jets and particles, the rapidity $y = (1/2)\ln[(E + p_z)/(E - p_z)]$ is used, where E is the jet/particle energy and p_z is the z-component of the jet/particle momentum. The angular distance is defined as $\Delta R \equiv \sqrt{(\Delta y)^2 + (\Delta\phi)^2}$.

The full 4π solid angle coverage of the detector is not only important to capture as many particles as possible, but also is crucial to the detection of neutrinos. The neutrino does not interact with the ATLAS detector and passes by undetected. Therefore, it is only deduced from imbalances in transverse momentum and energy. This technique can be used for any hypothetical particles that are believed to not interact with ordinary matter and can only produce trustworthy results if as many of the detectable particles are measured.

3.2.1 The inner detector

The inner detector (ID) is the tracking system of the ATLAS detector and covers the pseudorapidity region $|\eta| < 2.5$. It is immersed in a magnetic field of 2 T created by a superconducting solenoid. The magnetic field is aligned such that charged particles are deflected only in the ϕ direction. The ID measures the transverse momentum of charged particles via their trajectory's curvature. The trajectory of a charged particle is referred to as a track. The 1.6 m barrel section of the ID consists of three subdetector systems arranged in concentrical shells around the IP. The silicon pixel detectors (Pixel) are closest to the beamline, followed by the semiconductor tracker (SCT) and finally by the transition radiation tracker (TRT). A diagram of the barrel section of the ID and its subsystems is shown in Figure 3.5. The end-cap IDs are mounted on disks and are placed perpendicular to the beamline for 2.8 m on both sides of the IP, as shown in Figure 3.6

By extrapolating the tracks to their origin point near the IP, interaction vertices are reconstructed. In particular, the point where the pp collision occurs is called the primary vertex (PV). Secondary vertices arise where long-lived particles decay and

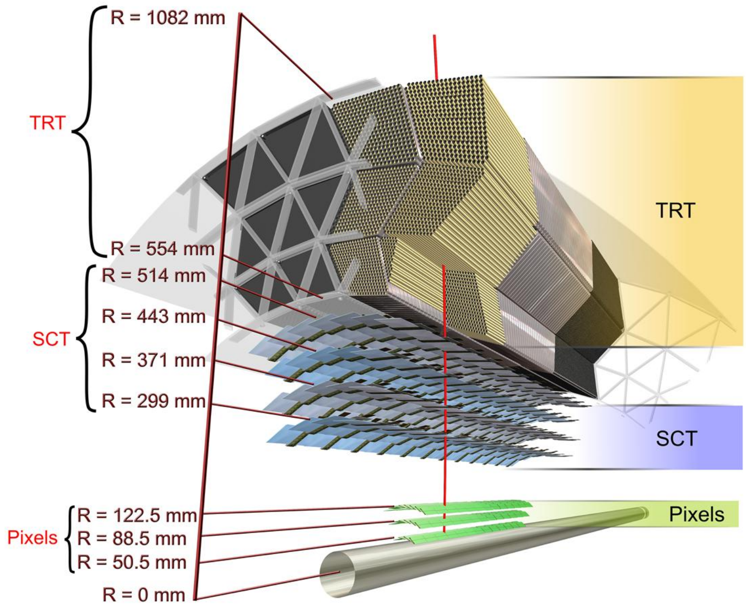


Figure 3.5: The barrel section of the ATLAS inner tracker detector. The distances to the interaction point are also shown [83].

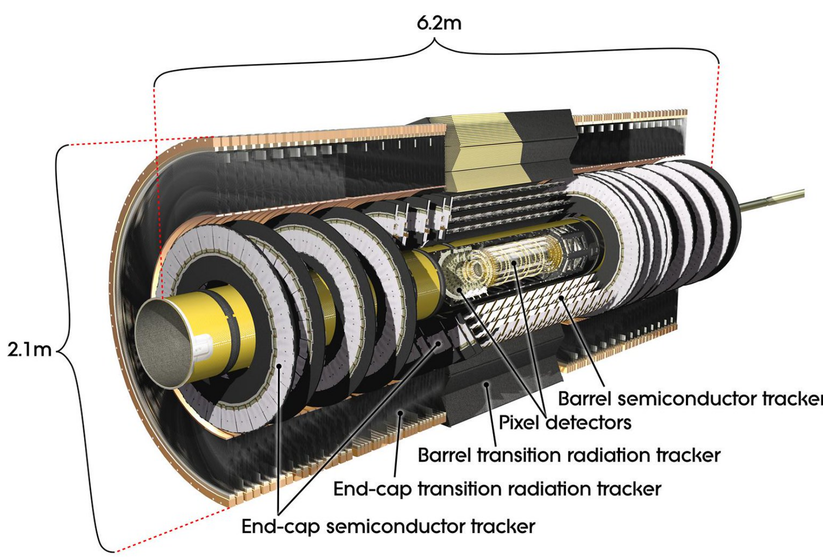


Figure 3.6: Overview of the ATLAS inner tracker detector. The dimension of the inner detector, including the barrel and both end-caps, is also shown [83].

are often used to identify mesons or baryons containing a b quark.

The Pixel subdetector [84] is located in the innermost layers of the ATLAS detector. It provides measurements of the three spatial coordinates of charged particles over the pseudorapidity region $|\eta| < 2.5$. The Pixel detector has the highest granularity of the ID due to its proximity to the IP and is pivotal for the reconstruction of primary and secondary vertices with a high resolution. It is made of silicon pixels with a size of $50 \times 400\ \mu\text{m}^2$ providing spacial resolution in the range of $10\ \mu\text{m} - 100\ \mu\text{m}$. The insertable B-layer (IBL) [85] was added to the Pixel detector before the start of Run–2 to improve the resolution of the inner tracker, to better resolve interaction vertices and to overcome radiation damage. Its pixel size is approximately half that of the original three layers of the Pixel detector.

The SCT [86] surrounds the Pixel detector. In the barrel region, it is made of four layers with two silicon microstrip sensors each. The sensors measure 80 μm in width. In the end-cap regions, nine disc layers are installed leading to the full coverage of the pseudorapidity region $|\eta| < 2.5$. sensors. In each layer, the two sensors are rotated by 40 mrad such that it allows for the measurement of a second spatial coordinate in the barrel and the end-cap regions.

The outermost detector of the ID is the TRT that covers the pseudorapidity region $|\eta| < 2.0$. It consists of straw tube proportional counters. The TRT is designed so that different dielectric indices of the gas and the materials of the straw tubes lead to x–ray transition radiation when charged particles pass through the detector. The radiation then ionizes the gas and creates a measurable electrical signal. Since the energy of transition radiation is related to the speed of the emitting particle, particles can be identified according to their masses. For transition radiation, given the same energy, a particle with low mass will emit much higher energy radiation than its heavier counterpart. When compared to the Pixel and SCT detectors, the 4 mm width of the straw tubes leads to a relatively low resolution of only 130 μm. However, the large number of hits per track, typically $35-40$, compensates for the lower hit resolution.

In the range $2.5 < |\eta| < 2.7$, particle detection relies on the calorimeter and muon spectrometer subsystems.

3.2.2 The calorimeter system

The calorimeter system is designed to provide the energy measurement of electrons, photons and hadrons by fully absorbing them. It covers the pseudorapidity region $|\eta| < 4.9$ and, like the ID, it is also arranged in a cylindrical configuration. The calorimeter system is built of sampling calorimeters that consist of alternative active and passive layers. Particles from the *pp* collision interact with the passive layers and produce cascade showers that further propagate through the calorimeter. The active layers absorb the resulting particles and measure the energy deposits. The calorimeter system is calibrated to infer the total shower energy to account for the missing energy deposits in the passive layers. The electromagnetic calorimeter (ECal) is placed closest to the IP and is surrounded by the hadronic calorimeter (HCal). They cover the central region of $|\eta| < 1.7$ in the barrel, and the region of $1.5 < |\eta| < 3.2$ in the end-caps. Forward calorimeters (FCal) are installed in the forward region of $3.1 < |\eta| < 4.9$. The calorimeter system is shown in Figure 3.7.

The ECal is an liquid argon (LAr) calorimeter [88] that uses LAr as the active material and lead as the passive absorber. It is designed to absorb and measure the energy of electrons and photons. It has a thickness of 22 radiation lengths X_0 in the barrel region, and 24 X_0 in the end-caps. The radiation length is a material

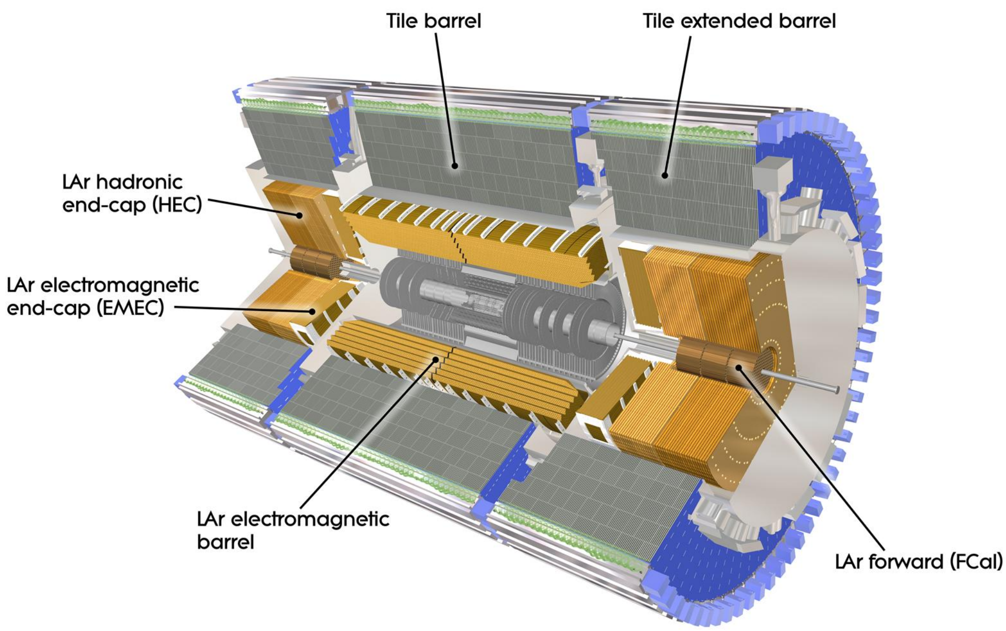


Figure 3.7: Overview of the ATLAS calorimeter subsystems [87].

characteristic that is related to the energy loss of particles interacting with it. An electron loses all but $1/e$ of its energy after one radiation length. The radiation length is also approximately the mean free path of a photon before it decays through pair production. The thickness of the ECal is designed to fully contain electromagnetic showers in the calorimeter system. It is separated into three layers. The inner layer consists in very fine strips to allow a high degree of angular resolution. The middle layer constitutes the thickest part of the ECal and is designed to contain the shower. The third layer is comparably thin and coarse and is designed to measure photon and electron showers with energies $E \geq 50$ GeV. The relative energy resolution of the ECal is $\sigma_E/E = 10\%/\sqrt{E} \oplus 1.2\%$, where E is in GeV, where $\oplus$ represents a sum in quadrature.

The barrel region of the HCal, TileCalo [89], uses scintillating tiles as the active material and iron as the passive absorber. In the end-caps, the HCal consists of LAr calorimeters. It is designed to absorb and measure the energy of hadronic jets, with a thickness of 10 interaction lengths λ. Similar to X_0, the interaction length is a material characteristic that represents the mean distance travelled by a hadronic particle before interacting with the medium. The HCal is designed to contain the hadronic showers in the calorimeter system. The TileCalo is separated into three layers, sharing the same design logic as the ECal. The first layer has the highest granularity, the second is the thickest and the third is designed to measure the highest energy jets. The relative energy resolution of the HCal is $\sigma_E/E = 50\%/\sqrt{E} \oplus 2.5\% \oplus 5\%/E$, where E is in GeV.

The FCal consists of a LAr calorimeter made from radiation hard technology so that it can withstand the high rate environment of the forward region. The front of the FCal consists an ECal designed to contain electromagnetic showers. It uses copper as passive material. The second and third components of the FCal use tungsten as passive material and provide energy measurements of hadronic showers. The relative energy resolution of the FCal is $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$, where E is in GeV.

3.2.3 The muon spectrometer

The muon spectrometer (MS) [90] is the largest and outermost detector system of ATLAS. Covering the pseudorapidity region $|\eta| < 2.7$, it is designed to detect muons, which are not captured by the calorimeters like the electrons, photons and hadrons. The MS consists of four types of gas-based detectors and a system of toroid magnets that creates a magnetic field up to 2 T. The magnet system is designed to deflect muons in the η direction, orthogonal to their bending direction from the ID solenoid. The barrel region is assembled from three cylindrical layers and the end-caps consist of four wheels, covering the regions of $|\eta| < 1.05$ and $1.05 < |\eta| < 2.7$, respectively.

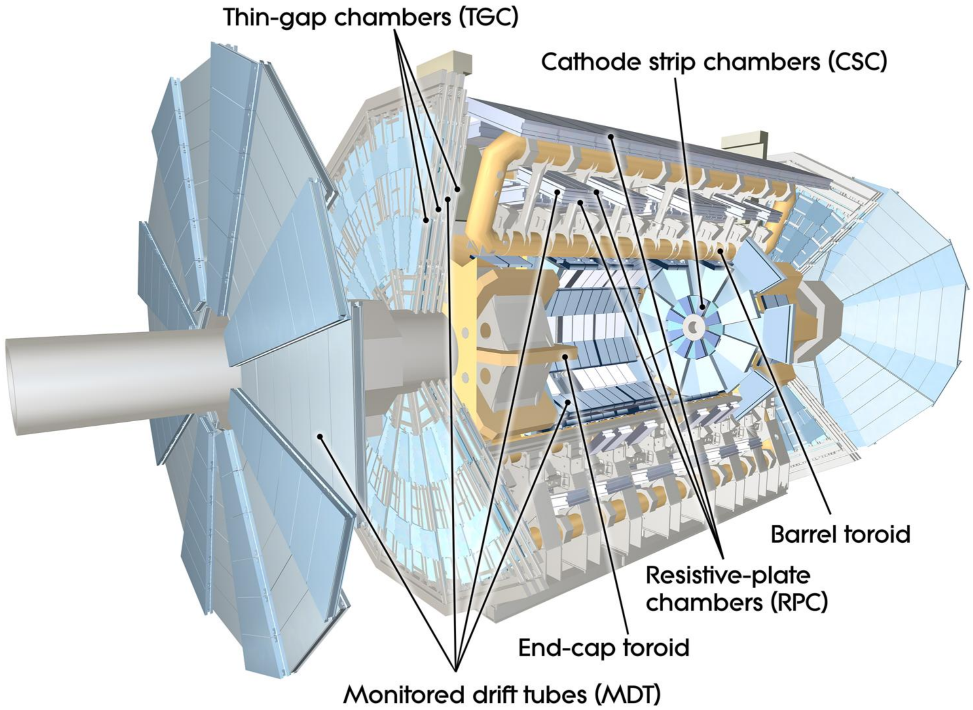


Figure 3.8: Overview of the ATLAS muon spectrometer subsystems [91].

The MS consists of several subdetectors with the goals of either providing precision tracking or fast triggering capabilities. A diagram of the MS and its subsystems is shown in Figure 3.8.

The Monitored Drift Tubes (MDT) and the Cathode Strip Chambers (CSC) form the precision tracking subdetectors. The MDTs are installed over the full pseudorapidity range $|\eta| < 2.7$ of the MS. In the forward-most region $2.0 < |\eta| < 2.7$, where the rate of particles is highest, the CSCs complement the MDTs due to their better time resolution. The subdetectors responsible for the fast triggering capabilities, and to some extent tracking, are the resistive plate chambers (RPC) and thin gap chambers (TGC). Both subdetector technologies benefit from very fast trigger response, in the order of the nanoseconds. The trigger measurements are provided in the barrel region by RPCs for $|\eta| < 1.05$ and in the end-caps by the TGCs for

$1.05 < |\eta| < 2.4$. The MS does not provide triggering capabilities for $|\eta| > 2.4$.

For high-energy muons, the measurement their momentum depends on the measurement of the track sagitta, or the bending of the track in respect to a straight line. Because of this, the relative resolution on the muon momentum p is proportional to itself: $\sigma_p/p \sim p$. This lead to rising momentum resolution as the momentum increases. The MS provides a relative momentum resolution of $\sigma_p/p \sim 10\%$ for 1 TeV muons and 2%–4% for lower momenta. In the case of low-energy muons, the combined MS + ID measurements are necessary to properly resolve the transverse momentum.

3.2.4 The trigger system

The ATLAS trigger system [92] is designed to reduce the rate of incoming data by several orders of magnitude so that it can be realistically stored to permanent storage. During Run–2, the rate of collisions was approximately 40 MHz. This high rate leads to a data storage requirement in the order of a petabyte per second, which is unsustainable for modern data-handling solutions. Furthermore, a fraction of pp collisions produce uninteresting events that can simply be discarded. The Run–2 trigger system uses predefined algorithms to reduce the rate of stored events from 40 MHz to approximately 1 kHz. It consists of the level-1 trigger (L1) followed by the high-level trigger (HLT). The HLT further refines the selection made from the L1 trigger, based on additional event selection requirements.

The L1 trigger is a hardware based system that reduces the event rate to 70 kHz–100 kHz. Until this trigger decision is taken, all the detector readout information are kept in buffers. Hence, the quick decision-time of 25 μs from the L1 trigger is of upmost importance. Because of this short time requirement, the L1 trigger uses a reduced detector granularity and resolution from the calorimeter and muon spectrometer to decide if a given event is discarded or not. By performing a simplified event reconstruction, it searches for signatures of muons, electrons, photons and jets. If the signatures are above a certain energy or p_{T} threshold, they determine region of interests (RoI) in the detector. Events which pass RoI requirements are fully read out and transferred to the HLT.

The HLT consists of a software based system that runs on a processing farm where events are further analyzed using the full detector readout information. The HLT further refines the decision from the L1 trigger by allowing the reconstruction of tracks in the inner detector and additional muon spectrometer segments. It further reduces the event rate from the L1 trigger output down to approximately 1 kHz. The average processing time per event is below 300 ms. Events saved to disk occupy on average approximately one megabyte of memory. The HLT provides the final trigger selection for the data samples utilized for off-line physics analyses.

To summarize, this chapter provides a description of the ATLAS experiment at the LHC used in the measurements of Z+high-p_{T} jets production. More details about the theoretical techniques used in the description of proton–proton collisions are presented in the next chapter. The tools and analysis strategy are then further outlined starting in Chapter 5.

Chapter 4

Modelling of hadron-hadron collisions

The cross section of two initial-state protons p_1 and p_2 to create the process $p_1 p_2 \to X$ was previously introduced in Equation (2.23). It includes the complete description of possible processes between proton constituents that can produce a given final state. The equation can be described as two distinct problems: 1) the parametrization of the proton structure into PDFs $f(x, \mu_F)$, and 2) the calculation of the partonic cross sections. The proton PDFs are parameterized using fits to experimental data, while the partonic cross-section must be calculated from first principles.

The calculation of the cross section for a given process is very complex and depends on several key parameters. In the field of particle physics, this problem is overcome with Monte Carlo (MC) event generators. MC event generators allow for the simulation of subatomic particle collisions and provide the fully detailed final states at the level of observable particles. The generated events can also be passed through a faithful simulation of the experimental apparatus. This is of crucial importance as it allows for a complete description of the detector response to incident particles and also for the experimental data to be interpreted on the same footing as the MC theoretical predictions.

In practice at the LHC, the MC event generation is split up into different steps prior to reaching the final state of observable particles. First, the hard scattering process is calculated in a fixed-order perturbative expansion. The expansion is based on Feynman diagrams to describe a given process of interest, *e.g.* Z+1 jet production, and dictates the accuracy of the simulation via the order of the calculation. Since the computational time of calculating the perturbative expansion of final states with $\mathcal{O}(10s)$ of particles is too high, the further emissions of particles are instead described by the parton showering. In a second step, the parton showering describes the process by which the initial parton loses energy via soft and collinear parton radiation down to the cut-off energy scale of $\mathcal{O}(1\ \text{GeV} - 2\ \text{GeV})$. Third, the underlying event, or multiple interaction, describes further interactions of the proton constituents that did not take part in the hard scattering. Then, in a fourth step, the hadronization models the formation of hadrons out of the available partons. The unstable hadrons finally undergo decays to lighter and longer-lived particles. The steps undertaken by MC event generators in modelling a pp collision are shown schematically in Figure 4.1.

Fixed-order cross-sections are described in Section 4.1. Parton showering is introduced in Section 4.2. Hadronization is discussed in Section 4.3. The methods of matching and merging are explored in Section 4.4. A brief discussion about virtual EW corrections is included in Section 4.5.

4.1 Fixed-order cross sections

Fixed-order cross sections have a three-fold importance to this analysis. First, the discussion of different Z+jets production processes is facilitated with the use of Feyn-

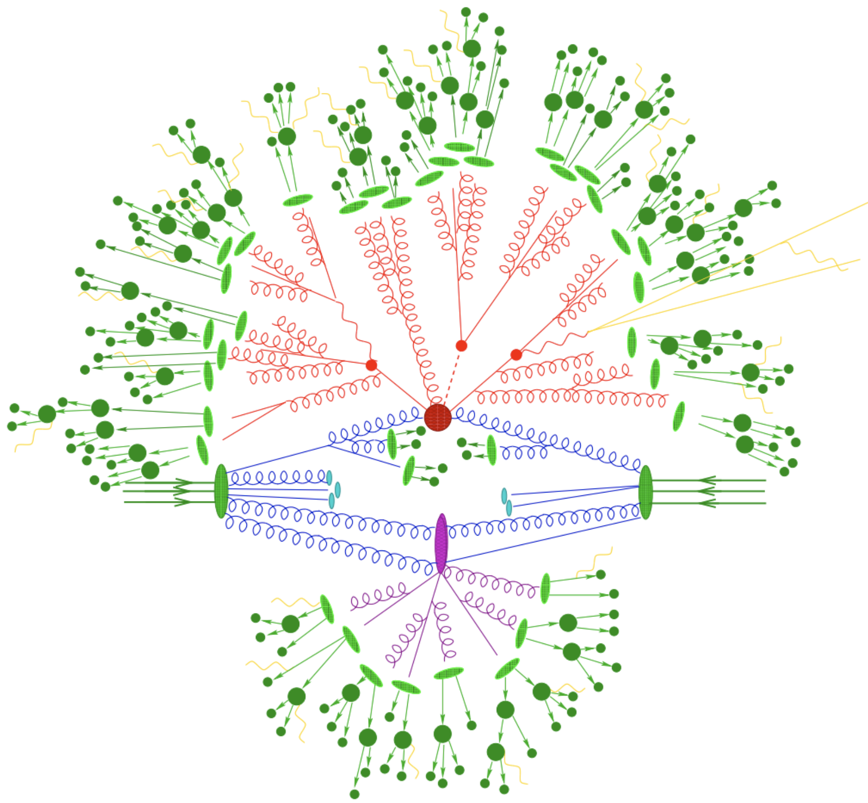


Figure 4.1: Sketch of the simulation of a *pp* collision. The incoming protons are located in the centre of the figure and are identified as green arrows. Interacting partons from the protons are shown in blue. The hard scattering process, represented by the red circles, is predicted with MC event generators. The parton showering simulation produces the outgoing partons in red. A secondary hard-scatter is shown in purple and consists of multiple parton interactions. The simulation of parton hadronization is shown in light green and is followed by parton decay in dark green. The radiation of photons is shown in yellow [93].

man diagrams, a visual representation related to the fixed-order cross section. Second, all the MC simulated event samples were produced at a given fixed order in the cross section, as described in Chapter 5. Lastly, the cross-section measurements are compared to fixed-order calculations from the NNLOJET program [94, 95], as discussed in Chapter 5. As a result, the formalism behind the order of a fixed-order cross section is introduced in Section 4.1.1. Then, a more precise formulation of fixed-order calculations is given in Section 4.1.2.

4.1.1 Partonic cross sections

The partonic cross section $\sigma_{p_a p_b \to X}$, first introduced in Chapter 2, describes the cross section for the production of the final state X from the incoming partons p_a and p_b. It is calculated as a perturbative series of the strong coupling constant α_s:

$$\sigma = \alpha_s^k \sum_{n=0}^{\infty} c^{(n)} \alpha_s^n, \tag{4.1}$$

where k is the minimum number of strong vertices in the tree-level Feynman diagram and the coefficients $c^{(n)}$ correspond to the n-th order contribution to the cross section. Each increasing order of n introduces additional virtual-loop contributions and real emissions of fermions and bosons. At the energy scale of the LHC, the strong coupling constant is small, $\alpha_s \approx 0.118 \ll 1$, such that the perturbative expansion defines a converging series. The $c^{(n)}$ coefficients are classified by their order n: $n = 0$ terms are called leading order (LO), $n = 1$ for next-to-leading order (NLO), $n = 2$ for next-to-next-to-leading order (NNLO) and so on. Using this notation, Equation (4.1) becomes

$$\sigma = \sigma^{LO} + \alpha_s \sigma^{\mathrm{NLO}} + \alpha_s^2 \sigma^{\mathrm{NNLO}} + \ldots \quad . \tag{4.2}$$

Although the accuracy of the calculation increases with the addition of higher order terms, theoretical predictions truncate the sum to a given order N that defines the order of the calculation. For Z+1 jet production, the LO process contains one strong and EW vertex such that $\sigma^{LO} \sim \mathcal{O}(\alpha_s \alpha_{\text{EW}})$ and is described by the Feynman diagram of Figure 2.9.

4.1.2 Fixed-order calculations

A more precise formulation of fixed-order calculations is presented in this section to give some context on the NNLOJET theoretical predictions considered in this analysis. The exact perturbative expansion of the cross section σ for a given process is written as

$$\sigma = \sum_{k=0}^{\infty} \int_{N+k} d\Phi_{N+k} \left| \sum_{l=0}^{\infty} \mathcal{M}^{(l)}_{N+k}(\Phi_{N+k}) \right|^2 . \tag{4.3}$$

The sum and integrals over the different PDFs that appear in pp collisions are omitted for the sake of clarity. The indices N, k and l represent the number of particles in the final state, the number of additional particle emissions, and number of loops in the Feynman diagram, respectively. Furthermore, the integral is taken over final phase space described by $N + k$ particles, Φ_{N+k}. The scattering amplitude $\mathcal{M}^{(l)}_{N+k}(\Phi_{N+k})$ includes all possible scattering amplitudes describing phases spaces with $N + k$ final particles and l loops. The real-valued product of scattering amplitudes $|\mathcal{M}|^2$ is known as the matrix element (ME). Differential cross sections are obtained through the same calculation, with the exception that the integral is not performed over the variable X:

$$\frac{d\sigma}{dX} = \sum_{k=0}^{\infty} \int_{N+k} d\Phi_{N+k} \left| \sum_{l=0}^{\infty} \mathcal{M}^{(l)}_{N+k}(\Phi_{N+k}) \right|^2 \delta\Big(X - X(\Phi_{N+k})\Big). \tag{4.4}$$

In practice however, the perturbative expansion is truncated to a certain order and not carried over an infinite number of terms. Hence, the cross section calculations no longer yield exact results, but rather approximations. The LO cross section is written as

$$\sigma^{\mathrm{LO}} = \int_N d\Phi_N \mathcal{B}(\Phi_N), \tag{4.5}$$

where $\mathcal{B}(\Phi_N)$ corresponds to the Born-level matrix element

$$\mathcal{B}(\Phi_N) = \left|\mathcal{M}_N^{(0)}(\Phi_N)\right|^2. \tag{4.6}$$

The Born contributions are defined as having no loops in the Feynman diagrams, *i.e.* $l = 0$. The cases where the LO process includes loops, *e.g.* in gluon fusion Higgs production, are not considered here for the sake of simplicity.

The NLO cross section includes one more power in α_s and is

$$\sigma^{\mathrm{NLO}} = \int_N d\Phi_N \Big(\mathcal{B}(\Phi_N) + \mathcal{V}(\Phi_N)\Big) + \int_{N+1} d\Phi_{N+1} \mathcal{R}(\Phi_{N+1}), \tag{4.7}$$

where

$$\mathcal{V}(\Phi_N) = 2\mathcal{R}e\Big(\mathcal{M}_N^{(0)^\dagger} \mathcal{M}_N^{(1)}\Big) \tag{4.8}$$

and

$$\mathcal{R}(\Phi_{N+1}) = \left|\mathcal{M}_{N+1}^{(0)}\right|^2. \tag{4.9}$$

The contributions from $\mathcal{R}$ and $\mathcal{V}$ are called real and virtual emission corrections, respectively. The virtual correction contains the contribution from including one internal loop $l = 1$, whereas the real emissions describe final states with an additional emitted particle and no loops. The higher-order cross sections are described using the same formalism by adding contributions with increasing orders of α_s. The accuracy of the cross section increases with order of the expansion, but so does the complexity of the calculation. As α_s is small, the increase in accuracy of the prediction grows slowly, while the complexity of the calculation instead grows exponentially.

In principle, the Born (or leading order) cross section is finite. Depending on the studied process, it may instead be divergent due to different types of singularities that arise. Moreover, these singularities can also appear in higher-order terms of the perturbative expansion. In many cases, the configurations leading to singularities are caused by the presence of soft, collinear or massless partons in the final state. Applying jet selection cuts in the matrix element (ME) phase space that require a minimum momentum and minimum separation from the associated partons can often avoid these problematic configurations.

The fixed-order calculations from NNLOJET considered in this analysis correspond to NNLO and NLO predictions of the Z+1 jet process, where the Z boson decays to either an electron or muon pair, as described in Chapter 5. The resulting cross sections correspond to Born-level partons, *i.e.* the leptons, quarks and gluons have not been subject to further radiation, decay or hadronization.

From this point forward, the above fixed-order cross-section calculation method is simply referred as a ME calculation.

4.2 Parton showers

Following the ME calculations above, MC event generators then go through parton showering to introduce further radiation of the resulting partons. Since none of the predictions considered in this analysis include EW parton showering, this section only describes QCD parton showers.

The appearance of jets in QCD interactions is owed to the running behaviour of the strong coupling constant α_s. The value of α_s increases exponentially as the energy scale of the interaction decreases. For this reason, as a given parton loses energy, *e.g.* via QCD radiation $q \to qg$ or $g \to q\bar{q}$, the likelihood of it undergoing further energy-loss interactions increases. As a result, cones of partons form around quarks and gluons produced in the ME and hadronize to form the jets that are measured in particle physics experiments.

It is theoretically possible to calculate the matrix elements for $2 \to N$ processes with $N = \mathcal{O}(10 - 30)$. However, the calculation is very time consuming and resource intensive such that it is impossible in practice. For this reason, parton showering programs are introduced to separate the $2 \to N$ process into a simple $2 \to n$ process followed by many $1 \to 2$ parton splittings, where $n \sim \mathcal{O}(2 - 6)$. Hence, the parton shower provides the link between the high energy partons from the fixed-order cross sections and the hadronization models that are applied to describe the formation of hadrons at energy scales in the order of 1 GeV–2 GeV.

The process of parton showering is modelled sequentially by evolving high-energy partons from the ME calculation to continuously lower energy scales. The energy loss is mediated by a random number of gluon and $q\bar{q}$ pair emissions. The parton emissions are described by splitting functions, while the probability that a parton undergoes an emission is characterized by Sudakov form factors. The evolution of the parton is tracked via the ordering variable Q. The ordering variable can be chosen from several valid choices, such as the transverse momentum of the mother parton, the opening angle between the two partons and many more. The sequence of parton showering is repeated until the parton energy scale reaches the cut-off renormalization scale $\mu_R = 1$ GeV–2 GeV, where QCD can no longer be described with perturbation theory and instead hadronization programs take over.

The SHERPA 2.2.11 prediction [72] considered in this analysis presents a unique treatment of the parton shower evolution histories, as described in Chapter 5. The parton showering program requires that the ordering variable Q must decrease between each splitting, a property referred to as ordered-histories. Such requirement leads to increasingly softer (low-energy) and collinear partons emissions. This novel approach is credited with improving the performance of the SHERPA 2.2.11 theoretical predictions in high jet-p_{T} and high Z-boson p_{T} phase spaces [72]. The ordering variable Q is defined as the '$\mathrm{k_t}$' scale of the splittings, a variable linked to the separation between the partons.

4.3 Hadronization

The process of hadronization describes the transition between the partonic final states from the parton shower to the hadronic final states comprised of baryons and mesons. It is applied after the parton showering lowers the energy of indiviual partons down to the cut-off range of a few GeV, the energy scale where QCD can no longer be modelled by perturbation theory. The confinement of partons into hadrons is then applied to the low-energy partons. Hadronization programs are described by non-perturbative models and are tuned to the data. The most recent and common hadronization models are the Lund string model [96, 97] and the cluster model [98]. The combination of parton showering and hadronization provides the complete description of the particle jets that are measured in particle physics experiments.

4.4 Matching and merging

Matrix element (ME) calculations provide accurate descriptions of high-energy $2 \to N$ processes, with N being generally small. When the number of final state particles increases, the complexity of the calculation becomes large and is the limiting factor when computing cross sections. ME calculations are also sensitive to divergences that arise from soft and collinear parton emissions. In contrast, parton showers generate additional emissions through probabilistic methods, often performing well only for soft and collinear emissions. However, due to the nature of parton showers, they provide results with worst accuracy than ME calculations in some regions of phase space. The logical progression of combining ME calculations and parton showers is used to take advantage of the strengths of both methods simultaneously. In some cases, the parton shower can replicate the same contributions from higher-order ME calculations and special consideration must be taken as not to double count these effects. For example, a parton-shower emission on top of the LO Born cross section is also described by the real emission corrections in the NLO cross section:

$$\sigma^{\mathrm{LO}} + \text{Parton shower} \approx \sigma^{\mathrm{NLO}}. \tag{4.10}$$

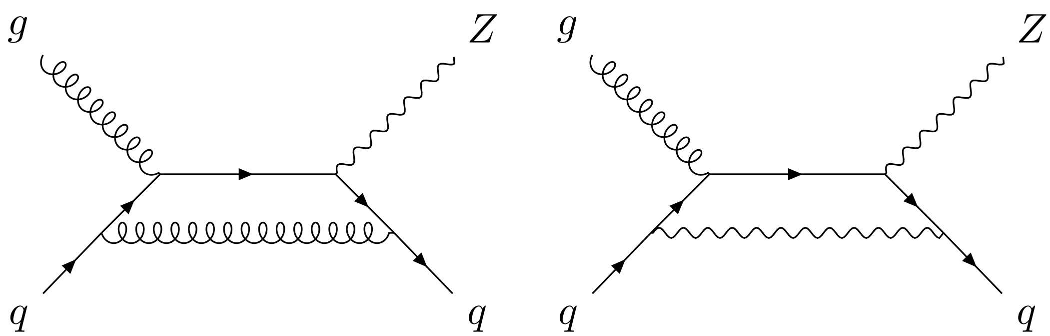
As such, the MC event generator must carefully consider and remove the duplicated contributions in order to produce accurate cross sections. The procedures that combine higher-order ME calculations with parton showers with this in consideration are called matching and merging.

The process of matching refers to the use of an higher-order cross section from the ME calculation to improve the parton shower. This can be done with several methods, such as correcting the parton shower to match the precise ME calculation, or by assigning some contributions to either the ME or the parton shower. The process of merging describes the combination of NLO and LO ME calculations for various particle multiplicities prior to undergoing through the parton shower.

An example of merging can be explained with the combination of the parton shower with the NLO ME calculations for $Z + 0/1/2$ jets production. First, the MEs are calculated separately for the NLO $Z + n$ jets processes, with $n \leq 2$. The merging procedure then combines the results prior to undergoing parton showering. Without merging, the ME calculations each go through the parton shower before being combined and results in the double counting of several cross-section contributions. This example describes the typical merging applied in the SHERPA 2.2.11 Z+jets MC event sample used in this analysis.

Various matching and merging schemes exist to combine the ME calculations with the parton shower. Three of these procedures are used in the different MC predictions in these measurements: 1) the CKKW-L merging [99,100], 2) the MEPS@NLO matching [101–104], and 3) FxFx merging [105].

The CKKW-L merging consists in correcting the ME calculations and parton shower by introducing the merging scale μ_{MS}. Partons emitted below the merging scale are modelled by the parton shower, while those above are modelled in the ME calculation. The MEPS@NLO matching computes shower subtraction terms that cancel the double counted parton shower contributions. The FxFx merging combines MEPS@NLO matched ME calculations of different parton multiplicities to perform NLO accurate multijet merging. Through the use of a merging scale, it dictates if a parton emission from a NLO n-jet ME is replaced with the contribution from the NLO $n + 1$-jet ME.



(a) NLO virtual strong correction (b) NLO virtual EW correction

Figure 4.2: Representative Feynman diagrams for virtual NLO corrections to the production of Z+1 jet. (a) The NLO virtual strong correction of order $\mathcal{O}(\alpha_s^2\alpha_{\mathrm{EW}})$ is shown alongside (b) the NLO virtual EW correction of order $\mathcal{O}(\alpha_s\alpha_{\mathrm{EW}}^2)$.

4.5 Virtual electroweak corrections

Electroweak corrections to Z+jets production can have an important impact in regions of phase space where the Z boson has large transverse momentum. These corrections arise from the presence of EW Sudakov type factors of the form $log(s/m_Z^2)$ [78], leading to an impact as large as several tens of percent at the TeV scale. As such, these corrections are crucial for the accurate modelling of Z+jets production in high-p_{T}(Z) phase spaces.

When considering Z+jets production, strong and EW contributions have to be understood as an expansion in both the strong and weak coupling constants α_s and α_{EW}, respectively. Hence, the order of the respective production cross sections are be described in terms of $\mathcal{O}(\alpha_s^M\alpha_{\mathrm{EW}}^N)$, where M and N are the order of the strong and EW expansions, respectively. This is simply demonstrated with Z+1 jet production as an example. The Z+1 jet process is described at LO as $\mathcal{O}(\alpha_s\alpha_{\mathrm{EW}})$, while the virtual NLO strong and EW corrections are in the order of $\mathcal{O}(\alpha_s^2\alpha_{\mathrm{EW}})$ and $\mathcal{O}(\alpha_s\alpha_{\mathrm{EW}}^2)$, respectively. The corresponding NLO virtual strong and EW corrections are shown in Figure 4.2.

When discussing the NLO strong and EW corrections, it is useful to use the following formalism for the NLO cross section:

$$\begin{aligned} \sigma^{\text{NLO}}_{\text{STRONG}} &= \sigma^{\text{LO}} + \delta\sigma^{\text{NLO}}_{\text{STRONG}} \\ \sigma^{\text{NLO}}_{\text{EW}} &= \sigma^{\text{LO}} + \delta\sigma^{\text{NLO}}_{\text{EW}}, \end{aligned} \tag{4.11}$$

where, σ^{LO} is the LO cross section and $\delta\sigma^{\text{NLO}}$ is the NLO correction.

Naively, the most simple approach to combining NLO strong and EW corrections is through a direct sum, in the so called additive scheme, so that

$$\sigma^{\text{NLO}}_{\text{STRONG+EW}} = \sigma^{\text{LO}} + \delta\sigma^{\text{NLO}}_{\text{STRONG}} + \delta\sigma^{\text{NLO}}_{\text{EW}}. \tag{4.12}$$

Alternatively, the higher-order strong and EW corrections can be combined in a way that includes possible interplay between strong and EW processes in the multiplicative scheme:

$$\sigma^{\text{NLO}}_{\text{STRONG}\times\text{EW}} = (\sigma^{\text{LO}} + \delta\sigma^{\text{NLO}}_{\text{STRONG}}) \times (1 + \delta\sigma^{\text{NLO}}_{\text{EW}}). \tag{4.13}$$

A third combination scheme is possible, named the exponentiated scheme:

$$\sigma^{\text{NLO}}_{\text{STRONG}\times e^{\text{EW}}} = (\sigma^{\text{LO}} + \delta\sigma^{\text{NLO}}_{\text{STRONG}}) \times e^{\delta\sigma^{\text{NLO}}_{\text{EW}}}. \tag{4.14}$$

Generally, predictions of hadron-hadron collisions are calculated only as a perturbative series of the strong coupling constant α_s. Hence, higher-order EW corrections are not included in these predictions. However, virtual EW corrections were made available recently [78, 106] and are included in the Sherpa 2.2.11 prediction studied in this analysis, as described in Chapter 5. The EW corrections modelled by the SHERPA 2.2.11 generator only contain virtual-loop terms since applying the full NLO EW corrections is not currently possible in the current multi-jet merging procedure. The impact of all these EW corrections are further discussed in Chapter 10 and Appendix A. Quantum interference between strong and EW Z+jets production processes are also relevant and their effects are exposed in Chapter 8.

In summary, this chapter aims to give a general description of the physics content needed for an accurate modelling of proton–proton collisions where a Z boson is produced in association with jets. The SHERPA 2.2.11 event generator discussed above provides the principle tool utilized in this analysis for the description of Z+high-p_{T} jets events. It corresponds to the state-of-the-art MC prediction for the production of W/Z+jets at ATLAS. More details about the simulated event samples are presented in the next chapter as the logical step towards event selection, background estimation and finally the measurement of differential cross sections.

Chapter 5

Data and simulated event samples

For the study of Z+high-p_{T} jets production, candidate events are selected from the data recorded by the ATLAS detector, where the Z boson decays to an electron or muon pair: $Z \to \ell^+\ell^-$ with $\ell = e, \mu$. Out of its possible decay modes, charged leptonic final states provide the cleanest signature of the Z boson at hadron colliders. In this chapter, the dataset and MC simulated event samples used for the Z+high-p_{T} jets cross-section measurements are presented. The dataset and MC simulated event samples are described in Section 5.1 and 5.2, respectively. Fixed-order theoretical predictions of Z+high-p_{T} jets production are introduced in Section 5.3.

5.1 Dataset

The data used for this analysis were collected by ATLAS in pp collisions at the centre-of-mass energy of 13 TeV between the years of 2015 and 2018. They correspond to the full LHC Run–2 dataset and amount to an integrated luminosity of 139 fb^{-1} $\pm 1.7\%$ [107]. Compared to previous measurements described in Section 2.3, this dataset provides an increase of statistical precision by a factor of 2 to 5. This dataset enables for the first time the precise measurement of the production cross section of Z+high-p_{T} jets. The mean number of pp interactions per bunch crossing, including hard scattering and other interactions in the same and neighbouring bunch crossings (pile-up), is $\langle\mu\rangle = 33.7$.

5.2 Simulated Monte Carlo event samples

In particle physics, Monte Carlo (MC) simulations are a crucial component of data mining and physics analyses. In these cross-section measurements, MC event samples are utilized to validate the event reconstruction, estimate the background, model the Z+jets process and evaluate the detector response. MC generators are used to simulate events of the Z+jets signal process and seven background processes that can mimic the signature of strong Z+jets production. The generated samples are processed using the GEANT4-based ATLAS detector simulation [108, 109] and the same event-reconstruction algorithms are used for both the MC samples and the data.

5.2.1 Strong Z+jets signal samples

As outlined in Section 2.3, contemporary MC predictions have difficulties modelling high-energy Z+jets phase spaces, especially where p_{T}(Z) or p_{T}(jet) is large, and where the Z boson and a jet are nearby. In order to compare the behaviours of different strong Z+jets MC predictions, four different event samples are produced and subsequently compared to the final cross-section measurements of Chapter 10. The four simulated signal samples are summarized in Table 5.1.

The production of Z bosons in association with jets was simulated with the ATLAS configuration of SHERPA 2.2.11 [72], which includes matrix elements for up to five partons at LO and up to two partons at NLO. They are calculated with the Comix [110] and OpenLoops [111–113] libraries and matched with the Sherpa parton shower [114] using the MEPS@NLO prescription [101–104] with a set of tuned parameters ('tune') developed by the Sherpa authors. In contrast to Sherpa 2.2.1 [115], used in this analysis and in previous ATLAS publications, it includes a modified Catani–Seymour subtraction scheme [116], the Hessian NNPDF 3.0 NNLO PDF set [117] is used, and an analytic enhancement technique has been introduced [72]. The cross section in the phase space with large $p_{\mathrm{T}}(\text{jet})$ or $p_{\mathrm{T}}(\mathrm{Z})$ has been considerably reduced relative to the prediction from the previous Sherpa version by switching to an improved matching scheme with a different treatment of unordered histories [104]. Sherpa 2.2.11 is also the only sample used in this analysis which includes NLO virtual EW corrections [112, 113], as presented in Section 4.5. The samples were produced using three options for the combination of NLO EW and QCD corrections: an additive, a multiplicative, and an exponential scheme. The nominal prediction is derived via the additive scheme; a systematic uncertainty band is derived from the envelope of all schemes. In contrast to virtual EW corrections, EW parton showers are not included in any of the generators used in this analysis. The Sherpa 2.2.11 Z+jets samples are used for the nominal unfolding of the data distributions, to estimate the systematic uncertainties and in comparisons with the cross-section measurements. These samples are state-of-the-art predictions and are joining the roster of ATLAS recommended MC predictions for the LHC Run–3.

A second Z+jets sample, referred to as MG5_aMC+Py8 FxFx [72], was produced by using the MadGraph5_aMC@NLO 2.6.5 [118] program to generate matrix elements at NLO accuracy in QCD for up to three additional partons in the final state. The NNPDF 3.1 NNLO set [117] was used in the generation. The parton showering and subsequent hadronization was performed using PYTHIA 8.240 [119] with the A14 tune [120] and the NNPDF 2.3 LO PDF set [121]. The jet multiplicities were merged using the FxFx prescription [105]. This prediction is compared with the unfolded cross-section measurements. It is also a state-of-the-art sample and is taking place in the lineup of ATLAS MC predictions for the LHC Run–3.

A third sample of Z+jets events was produced with the SHERPA 2.2.1 [115] generator using NLO matrix elements for up to two partons, and LO matrix elements for up to four partons, calculated with the Comix and OPENLOOPS libraries. They were matched with the SHERPA parton shower using the MEPS@NLO prescription with a set of tuned parameters developed by the SHERPA authors. The MC replica version of the NNPDF 3.0 NNLO set of PDFs was used. The SHERPA 2.2.1 Z+jets sample is used in comparisons with the unfolded cross-section measurements, as it was used as a standard prediction in previous ATLAS Run–2 publications.

The fourth Z+jets sample, referred to as MG5_aMC+PY8 CKKWL, was generated using LO-accurate matrix elements with up to four final-state partons calculated by MADGRAPH5_aMC@NLO 2.2.2 [118]. The ME calculation employed the NNPDF 3.0 NLO PDF set and was interfaced to PYTHIA 8.186 [122] for the modelling of the parton shower, hadronization, and underlying event. The overlap between matrix element and parton shower emissions was removed using the CKKW-L merging procedure [99, 100]. The A14 tune of PYTHIA8 was used with the NNPDF 2.3 LO PDF set. This sample is used to validate the unfolding method and in comparisons with the unfolded cross-section measurements.

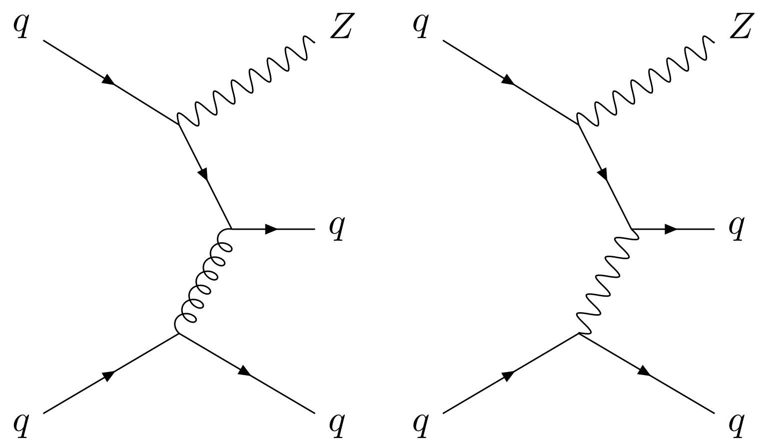
Table 5.1: Summary of the simulated Z+jets signal samples. For every process the name of the generator used is indicated in the second column. The third column reports the order of the QCD calculation in the matrix elements, where np denotes the number of real parton emissions. The first two predictions correspond to state-of-the-art MC generators, whereas the bottom two are standard predictions used in previous ATLAS Run–2 measurements.

Process	Generator	Order pQCD	References
Strong Z+jets	SHERPA 2.2.11	0–2p NLO, 3–5p LO	[72, 101–104, 110–117]
	MG5_aMC+PY8 FXFX	0–3p NLO	[72, 105, 117–121]
Strong Z+jets	SHERPA 2.2.1	0–2p NLO, 3–4p LO	[101–104, 110–115]
	MG5_aMC+PY8 CKKWL	0–4p LO	[99, 100, 118, 122]

5.2.2 Background samples

Seven background processes are simulated and used to estimate the background that mimic the strong Z+jets signal. The simulation of these samples is described below and summarized in Table 5.2.

The EW Z+2 jets (or Zjj) process is defined by the t-channel exchange of a weak boson and at tree level is calculated at $\mathcal{O}(\alpha_{\text{EW}}^4)$ when including the decay of the Z boson [7]. In contrast, the strong Zjj process, which is covered by the Z+jets samples, has no weak boson exchanged in the t-channel and at tree level is calculated at $\mathcal{O}(\alpha_{\text{EW}}^2\alpha_s^2)$ when including the decay of the Z boson. Representative Feynman diagrams of the strong and EW Zjj processes are shown in Figure 5.1. In this analysis, the measurements target only the strong Z+jets production and treat the EW Zjj process as a source of background. The EW Zjj samples were produced in the vector-boson fusion (VBF) approximation with HERWIG 7.1.5 [123, 124] at NLO accuracy in the strong coupling, using VBFNLO 3.0.0 [125] to provide the loop amplitude. The MMHT2014LO PDF set [126] was used along with the default set of tuned parameters for parton showering, hadronization and the underlying event. It is found that the interference between strong and EW Z+2 jets production is present in



(a) QCD mediated Z+2 jets (b) EW mediated Z+2 jets

Figure 5.1: Representative Feynman diagrams for the production of Z+2 jets. (a) The strong Zjj production is described by the exchange of a gluon and (b) the EW Zjj production is described with the exchange of a Z boson.

the fiducial phase space of these measurements, but it only corresponds only to a small number of events. Hence, to account for the interference between the strong and EW Zjj production processes, a uniform modelling uncertainty is applied. The relative importance of the interference effects, as described in Chapter 8, are determined from simulation with MadGraph5_aMC@NLO 2.9.5.

Event samples of $Z(\rightarrow \tau\tau)$+jets and W+jets processes were produced with the Sherpa 2.2.1 generator using the same configuration as the Sherpa 2.2.1 Z+jets sample.

The $t\bar{t}$ event samples were modelled using the PowhegBox v2 [127–130] generator at NLO with the NNPDF 3.0 NLO PDF set and the h_{damp} parameter set to 1.5 m_{top} [131], where m_{top} is the top quark mass. The h_{damp} parameter is a resummation damping factor and one of the parameters that control the matching of Powheg matrix elements to the parton shower, and thus effectively regulates the high-p_{T} radiation against which the $t\bar{t}$ system recoils. The events were interfaced to Pythia 8.230 [119] to model the parton shower, hadronization, and underlying event, with parameters set according to the A14 tune using the NNPDF 2.3 LO set of PDFs. The $t\bar{t}$ sample is normalized to the cross-section prediction at NNLO accuracy, including the resummation of next-to-next-to-leading logarithmic soft-gluon terms calculated with Top++ 2.0 [132–138]. The $t\bar{t}$ background in this analysis is derived with a data-driven method as described in Chapter 8 utilizing these MC event samples as an intermediate step.

Single top quark production in the s-channel, in the t-channel, and in association with a W boson (tW) was modelled using the PowhegBox v2 generator at NLO in QCD with the five-flavor scheme and the NNPDF 3.0 NLO set of PDFs. The diagram-removal scheme [139] was used to remove interference and overlap with $t\bar{t}$ production. The tW cross section is corrected to the theory prediction at approximate NNLO accuracy [140, 141], where the s- and t-channel cross sections are corrected to the prediction at NLO accuracy [142, 143].

Samples of diboson final states (VV) were produced with the Sherpa 2.2.1, or Sherpa 2.2.2 generator depending on the process, including off-shell effects and Higgs boson contributions where appropriate. Here, V stands for vector bosons $V = W$ or Z. Fully leptonic final states and semileptonic final states, where one boson decays leptonically and the other hadronically, were generated using matrix elements at NLO accuracy in QCD for up to one additional parton and at LO accuracy for up to three additional parton emissions. The matrix element calculations were matched and merged with the Sherpa parton shower as detailed above for Sherpa 2.2.1, and the NNPDF 3.0 NNLO set of PDFs was used.

The production of $V+\gamma$ final states was simulated with the Sherpa 2.2.8 [115] generator. Matrix elements at NLO QCD accuracy for up to one additional parton and LO accuracy for up to three additional parton emissions were matched and

Table 5.2: Summary of the simulated Z+jets background samples. For every process the name of the generator used is indicated in the second column. The third column reports the order of the QCD calculation in the matrix elements, where np denotes the number of real parton emissions.

Process	Generator	Order pQCD	References
EW Zjj	HERWIG 7.1.5, VBFNLO 3.0.0	NLO	[123–125]
Strong $Z(\to\tau\tau)$+jets	SHERPA 2.2.1	0–2p NLO, 3–4p LO	[101–104, 110–115]
W+jets	SHERPA 2.2.1	0–2p NLO, 3–4p LO	[101–104, 110–115]
$t\bar{t}$	POWHEGBOX v2 + PYTHIA 8.230	NLO	[127–130]
Single top (t-, Wt-, s-channel)	POWHEGBOX v2 + PYTHIA 8.230	NLO	[127–130]
$Z/W(\to qq)Z(\to \ell\ell)$	SHERPA 2.2.1	0–1p NLO, 2–3p LO	[101–104, 110–115]
$W(\to \ell\nu)Z(\to qq)$	SHERPA 2.2.1	0–1p NLO, 2–3p LO	[101–104, 110–115]
$W^{\pm}(\to qq)W^{\mp}(\to \ell\nu)$	SHERPA 2.2.1	0–1p NLO, 2–3p LO	[101–104, 110–115]
$\ell\ell\nu\nu, \ell\ell\ell\nu, \ell\ell\ell\ell$	SHERPA 2.2.2	0–1p NLO, 2–3p LO	[101–104, 110–115]
$V(\to \ell\ell) + \gamma$	SHERPA 2.2.8	0–1p NLO, 2–3p LO	[101–104, 110–115]

merged with the SHERPA parton shower as detailed above for SHERPA 2.2.1, and the NNPDF 3.0 NNLO set of PDFs was used.

5.3 Fixed-order theoretical predictions

Two additional Z+jets theoretical predictions were calculated with the NNLOJET program [94, 95], which computes fixed-order parton-level predictions for inclusive jet processes at higher orders in QCD. They are summarized in Table 5.3. The NLO and NNLO predictions, referred to as NNLOJET@NLO and NNLOJET@NNLO, respectively, were calculated as higher-order corrections to the parton-level LO process of Z+1 jet production. The NNLOJET@NNLO prediction corresponds to a fixed-order calculation of Z + up to three jets, whereas the NNLOJET@NLO predicts only Z + up to two jets. The NNPDF 3.1 NNLO set was used with a central scale choice of $\mu_0 = \frac{1}{2}(E_{\mathrm{T},Z} + \sum_{i\in\mathrm{partons}} p_{\mathrm{T},i})$ with $E_{\mathrm{T},Z} = \sqrt{m_{\ell\ell}^2 + p_{\mathrm{T},Z}^2}$. These samples are pure QCD predictions and correspond to Born-level partons. These predictions are used as complimentary theoretical comparisons to the final cross-section measurements of Chapter 10.

Table 5.3: Summary of the fixed-order Z+jets theoretical predictions. For every process the name of the program used is indicated in the first column. The second column reports the order of the QCD calculation in the matrix elements, where np denotes the number of real parton emissions.

Process	Program	Order pQCD	References
Strong Z+jets	NNLOjet@NNLO	1p NNLO	[94, 95]
	NNLOjet@NLO	1p NLO	[94, 95]

In summary, this chapter introduces the dataset on which the Z+high-p_{T} jets cross-section measurements are performed. Furthermore, the simulated event samples that describe the numerous background processes and the Z+jets theoretical predictions are enumerated. Details about the detector-level event selection requirements applied to the data and MC simulated events are presented in the next chapter.

Chapter 6

Event selection

This analysis focuses on the measurements of the production cross section of a Z boson in association with high-transverse-momentum jets and decaying into a charged-lepton pair. Final states containing e^+e^- or $\mu^+\mu^-$ pairs are exploited to easily identify the Z boson and study its properties. The cross-section measurements of Z+high-p_{T} jets production rely on the selection criteria imposed on data to select a clean sample of signal events. To select these events, trigger, isolation, identification and kinematic cuts are applied on the reconstructed objects of the ATLAS detector.

This chapter describes the detector-level event selection criteria applied to data and MC samples that are used to measure high purity Z+high-p_{T} jets events. The requirements on the considered events are presented in Section 6.1. The object selection requirements for electrons, muons and jets are discussed in Section 6.2. The final detector-level event selection is introduced in Section 6.3. The five signal regions and the different measured observables are outlined in Section 6.4.

6.1 Event requirements

The data are required to satisfy criteria that ensure that the proton beams and the ATLAS detector were in good operating condition [144]. The data events that follow these criteria form the standard ATLAS good run list (GRL). The GRL is the product of data quality monitoring that filters out data compromised by anomalous operating conditions. The dataset size of 139 fb^{-1} is calculated from the contents of the GRL.

Events must also have a primary vertex, defined as the vertex with the highest sum of track p_{T}^2, with at least two associated tracks with $p_{\mathrm{T}} \geq 500$ MeV [145]. They are selected using triggers [146–148] that require the presence of at least two electrons or two muons, or the combination of at least one electron and one muon. The efficiencies of these triggers plateau above 95% for $p_{\mathrm{T}}(e, \mu) \geq 25$ GeV. Triggers selecting at least two electrons or muons are chosen to suppress backgrounds that are not from Z boson production. The trigger demanding at least one muon and electron is used to validate the $t\bar{t}$ background in its own control region, as described in Chapter 8.

6.2 Object selection

The particle reconstruction algorithms in ATLAS combine information from the ID, calorimeter and MS systems described in Chapter 3 to identify the most likely physical objects responsible for the observed signals in the detector. Through the application of carefully chosen event cleaning techniques and selection criteria, those objects are extracted so that they are uniquely associated to particles and jets.

6.2.1 Electrons

Electron candidates are reconstructed from inner-detector tracks that come from the primary vertex and are matched to clusters of energy deposits in the EM calorimeter. To fulfil the primary-vertex condition, the electron must pass two requirements: 1) its track's transverse impact parameter significance must satisfy $|d_0|/\sigma(d_0) < 5.0$, where d_0 is the transverse impact parameter and $\sigma(d_0)$ its uncertainty; and 2) the longitudinal impact parameter z_0 must satisfy $|z_0 \sin\theta| < 0.5$ mm, where θ is the angle of the track to the beamline. Electron candidates must satisfy the 'Medium' likelihood-based identification requirements [149] based on EM shower shapes, track quality, and track–cluster matching. The 'Medium' requirement provides an electron identification efficiency of approximately 88%. Electron candidates must also satisfy the 'PflowLoose' [149] isolation requirement. The 'PflowLoose' constraint prioritizes efficiency at the cost of a lower background rejection relative to the tighter isolation working points. Electrons are selected if they pass all the above criteria, have $p_{\mathrm{T}} \geq$ 25 GeV and fall within the coverage of the inner tracker: $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$. The gap in η is caused by a crack in the calorimeter between the barrel and endcap regions that leads to a poor energy resolution.

6.2.2 Muons

Muon candidates are identified by matching inner-detector tracks from the primary vertex to either full tracks or track segments reconstructed in the muon spectrometer. The muon must fulfil the two primary-vertex requirements: 1) the transverse impact parameter significance must satisfy $|d_0|/\sigma(d_0) < 3.0$; and 2) the longitudinal impact parameter must satisfy $|z_0 \sin\theta| < 0.5$ mm, where d_0, $\sigma(d_0)$, z_0 and θ are as defined above for the electrons. Muons must pass 'Medium' identification requirements [150, 151] based on quality criteria applied to the inner-detector and muon-spectrometer tracks. The 'Medium' requirement provides a muon identification efficiency of over 96% for muons with $p_{\mathrm{T}} \geq 25$ GeV, while keeping identification related systematic uncertainties small. Muon candidates with $p_{\mathrm{T}} \geq 300$ GeV must satisfy tighter identification requirements in the muon spectrometer in order to improve the muon-p_{T} resolution. The tighter identification restriction enforces a higher number of hits in the muon spectrometer's precision layers. Muon candidates must also satisfy the 'PflowLoose' isolation requirement, built from tracking and calorimeter information, with a muon-p_{T}-dependent variable cone size ΔR [151]. This constraint provides the highest efficiency for muons with $p_{\mathrm{T}} \geq 25$ GeV. Muons are selected if they pass all the above criteria, have $p_{\mathrm{T}} \geq 25$ GeV and fall within $|\eta| < 2.4$.

6.2.3 Jets

Jets of hadrons are reconstructed using a particle-flow algorithm [152] based on noise-suppressed topological clusters in the calorimeter. Energy deposited in the calorimeter by charged particles is subtracted and replaced by the momenta of tracks which are matched to those topological clusters. The jets are clustered using the anti-k_t [153] algorithm implemented in the FastJet package [154] with a radius parameter $R = 0.4$. They are further calibrated according to in situ measurements of the jet energy scale [155]. All selected jets are required to have a calibrated $p_{\mathrm{T}} \geq 100$ GeV and $|y| < 2.5$.

6.2.4 Overlap removal

The electrons, muons and jets are reconstructed and identified independently. In some cases, objects are reconstructed by several algorithms and identified as multiple objects simultaneously. One common example is the high energy electron that leaves clusters in the hadronic calorimeter and is reconstructed both as an electron and a jet. An overlap-removal procedure is then applied to uniquely identify these objects in an event. For the lepton–jet overlap removal, the p_{T} requirement on jets is relaxed so that preselected jets with $p_{\mathrm{T}} \geq 30$ GeV and $|y| < 2.5$ are considered. The overlap removal is performed sequentially. First, preselected jets with a high probability to have been initiated by an electron or a radiated photon such that ΔR between the jet and a lepton is smaller than 0.2 are removed. In a second step, leptons closer than $\Delta R = 0.4$ to any remaining jet are removed.

The angular separation between two objects is defined as

$$\Delta R \equiv \sqrt{(\Delta y)^2 + (\Delta \phi)^2}, \tag{6.1}$$

where Δy and $\Delta \phi$ are the pseudorapidity and azimuthal angle differences, respectively.

6.3 Detector-level event selection

The Z+high-p_{T} jets process is defined by the presence of two specific objects: the Z boson decaying to charged lepton pairs $(e^+e^-, \mu^+\mu^-)$ and high-p_{T} jets. Events are selected if they satisfy the criteria defined above and the following kinematic requirements on the Z boson and jets. These criteria define the detector-level event selection.

Events are selected if they contain exactly two reconstructed same-flavor and opposite-charge leptons ($\ell = e, \mu$). The dilepton invariant mass $m_{\ell\ell}$ is required to be 71 GeV $\leq m_{\ell\ell} \leq$ 111 GeV. The selected events must also contain at least one jet. The invariant mass constraint within 20 GeV of the Z boson mass $m_Z = 91$ GeV

Table 6.1: Summary of the Z+high-p_{T} jets objects and event selection requirements. Only the kinematic cuts applied to the objects and to the event selection are shown.

Physics objects
Electrons
$p_{\mathrm{T}} \geq 25$ GeV
$\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.47$
Muons
$p_{\mathrm{T}} \geq 25$ GeV
$\lvert\eta\rvert < 2.4$
Jets
$p_{\mathrm{T}} \geq 100$ GeV
$\lvert y\rvert < 2.5$

Event selection
***Z* boson**
Exactly 2 same-flavor and opposite-charge leptons ℓ
$\ell = e, \mu$
71 GeV$\leq m_{\ell\ell} \leq 111$ GeV
Jets
At least 1 jet

suppresses background events and ensures the production on an on-shell Z boson. The objects and event selection are summarized in Table 6.1

6.4 Signal regions and observables

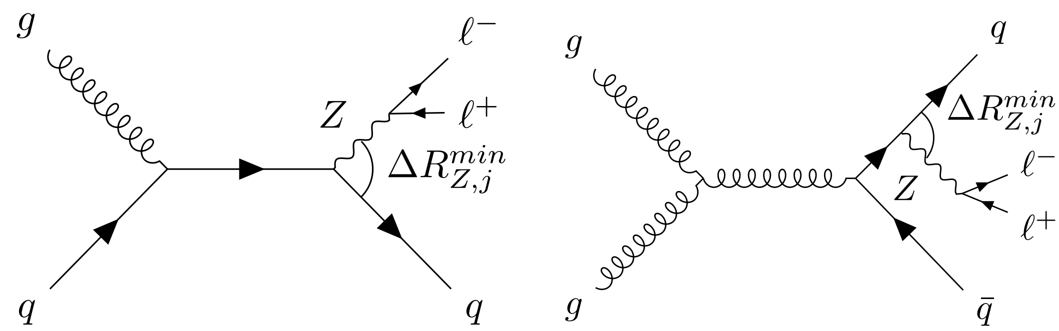
This analysis presents the measurements of Z+high-p_{T} jets production in thirteen different observables split over five signal regions. The different observables and signal regions are described below.

In Z+jets measurements, the transverse momentum of the leading jet ($p_{\mathrm{T},j1}$) and the transverse momentum of the Z boson ($p_{\mathrm{T},\ell\ell}$) are fundamental observables that probe perturbative QCD over a range of energy scales. The jet multiplicity (N_{jets}) also provides an important probe of QCD. The scalar sum of the transverse momentum of all selected jets and leptons H_{T}, defined as

$$H_{\text{T}} = p_{\text{T},\ell_1} + p_{\text{T},\ell_2} + \sum_{\substack{i\ \in \\ \text{selected jets}}} p_{\text{T},j_i}, \tag{6.2}$$

supplies complimentary information on the modelling of the event as a whole. The phase space with at least one jet is labelled as the *inclusive* region. In the *inclusive* region, the measured observables are $p_{\text{T},j1}$, $p_{\text{T},\ell\ell}$, N_{jets} and H_{T}.

The *high-*p_{T} region is a subset of the *inclusive* and is selected by requiring the presence of a jet with $p_{\text{T}} \geq 500$ GeV. This region, described in Section 2.3.2, is predicted to be composed of two characteristic topologies: 1) the hard scatter of a Z boson against a jet in the back-to-back topology as shown in Figure 6.1a, and 2) the soft radiation of a Z boson from a jet in the collinear topology as visualized in Figure 6.1b. The two topologies are expected to produce different distributions in three key observables. The first key observable to study the collinear and back-to-back topologies is the angular separation between the Z boson and the closest jet: $\Delta R^{\text{min}}_{Z,j}$. In the collinear topology, the Z boson is radiated near a jet such that $\Delta R^{\text{min}}_{Z,j}$ is small, whereas in the back-to-back topology the Z boson and the jet travel in opposite directions such that $\Delta R^{\text{min}}_{Z,j} \sim \pi \approx 3.14$. To further study these topologies, the *high-*p_{T} region is split by measuring $\Delta R^{\text{min}}_{Z,j}$ to create the *collinear* and *back-to-back* regions. The *collinear* and *back-to-back* regions, both subsets of the *high-*p_{T} region, are constructed by requiring $\Delta R^{\text{min}}_{Z,j} \leq 1.4$ and $\Delta R^{\text{min}}_{Z,j} \geq 2.0$, respectively. The second key observable to distinguish between the collinear and back-to-back topology is the ratio of the Z-boson p_{T} to the closest-jet p_{T}, defined as



(a) Z+1 jet (b) Z+2 jets

Figure 6.1: Representative Feynman diagrams for the production of a Z boson in association with high-p_{T} jets. In these diagrams, the angular distance between the Z boson and the closest jet are identified and labelled $\Delta R^{\mathrm{min}}_{Z,j}$. (a) The Z+1 jet events are expected to populate the *back-to-back* region where the Z boson is balanced against a single high-p_{T} jet. (b) In dijet events, the Z boson is expected to be radiated from the quark leg, with kinematics leading to small values of the angular distance between the Z boson and the closest jet, $\Delta R^{\mathrm{min}}_{Z,j}$, and therefore populating the *collinear* region.

$$r_{Z,j} \equiv \frac{p_{\mathrm{T},\ell\ell}}{p_{\mathrm{T}}(\text{closest jet})}.$$

In collinear Z-boson emissions, conservation laws lead to small values of $p_{\mathrm{T},\ell\ell}$ such that $p_{\mathrm{T},\ell\ell} \ll p_{\mathrm{T}}(\text{closest jet})$ and result in very small values of $r_{Z,j}$. In the back-to-back Z+1 jet hard scatter, the Z boson is balanced against the jet such that their p_{T} are approximately equal so that $r_{Z,j}$ is approximately equal to 1.0. Finally, the third key observable is the jet multiplicity N_{jets}. The *collinear* region is expected to be dominated by Z+2 jets events ($N_{\mathrm{jets}} = 2$), whereas the *back-to-back* region is predicted to be described mostly by Z+1 jet events ($N_{\mathrm{jets}} = 1$). In the *high-p_{T}* region, the measured observables are $\Delta R^{\mathrm{min}}_{Z,j}$, $r_{Z,j}$ and N_{jets}. In the *collinear* and *back-to-back* regions, restricted by $\Delta R^{\mathrm{min}}_{Z,j} \leq 1.4$ and $\Delta R^{\mathrm{min}}_{Z,j} \geq 2.0$, respectively, the $r_{Z,j}$ and N_{jets} observables are measured.

An alternative phase space is selected by applying a cut on the scalar sum of the p_{T} of the selected jets S_{T}:

$$S_{\mathrm{T}} = \sum_{\substack{i\ \in \\ \text{selected jets}}} p_{\mathrm{T},j_i}. \tag{6.3}$$

The *high*-S_{T} region is defined by requiring $S_{\mathrm{T}} \geq 600$ GeV. This alternative region probes high-energy events but does not depend on the presence on a single very energetic object. In contrast to the *high*-p_{T} region, this region is expected to have a large fraction of events with higher jet multiplicities. The $\Delta R^{\mathrm{min}}_{Z,j}$ and N_{jets} observables are measured in the *high*-S_{T} region.

The definitions of the five signal regions and the observables measured within them are summarized in Table 6.2. This chapter provides the detector-level definitions of electrons, muons and jets that are used to define the measurements of Z+high-p_{T} jets. Details on the analysis strategy are presented in the next chapter and naturally lead to the background estimation and finally the measurements of differential cross sections.

Table 6.2: Summary of the Z+high-p_{T} jets signal regions and observables measured within them.

Region	**Definition**	**Observables**
Inclusive	$Z + \geq 1$ jet	$p_{\mathrm{T},\ell\ell}$ $p_{\mathrm{T},j1}$ H_{T} N_{jets}
High-p_{T}	*Inclusive* and $p_{\mathrm{T},j1} \geq 500$ GeV	$\Delta R^{\mathrm{min}}_{Z,j}$ $r_{Z,j}$ N_{jets}
Collinear	*High-p_{T}* and $\Delta R^{\mathrm{min}}_{Z,j} \leq 1.4$	$r_{Z,j}$ N_{jets}
Back-to-back	*High-p_{T}* and $\Delta R^{\mathrm{min}}_{Z,j} \geq 2.0$	$r_{Z,j}$ N_{jets}
High-S_{T}	*Inclusive* and $S_{\mathrm{T}} \geq 600$ GeV	$\Delta R^{\mathrm{min}}_{Z,j}$ N_{jets}

Chapter 7

Analysis strategy

In experimental particle physics, the outcome of a measurement is the result of an underlying physical process that is convoluted with the response of a detector. Therefore, unless an effort is taken to decouple detector effects from the measurements, these measurements fundamentally represent detector-dependent observables. The unbiased, or particle-level, quantities are obtained by deconvoluting away the detector effects through an unfolding procedure. In this chapter, the methodology for the measurements of the Z+high-p_{T} jets production cross section is presented.

The differential cross-section measurement strategy is introduced in Section 7.1. The particle-level selection that defines the fiducial volume of these measurements is presented in Section 7.2. The unfolding of detector effects is described in Section 7.3. The detector response to Z+high-p_{T} jets events is explored in Section 7.3.2. The validation of the unfolding procedure is exposed in Section 7.3.3.

7.1 Measurement strategy

The goal of this analysis is to measure differential and integrated cross sections of Z+high-p_{T} jets production. However, cross sections are not directly measured and instead are derived from the results of counting experiments. In particle physics, measurements are based on counting signal events that fall within a particular region of phase space, called the fiducial volume. The fiducial volume of this analysis broadly corresponds to the phase space described by Z+high-p_{T} jets production that can be measured with the ATLAS detector.

Instead of publishing the number of signal events in a given dataset, the fiducial cross section

$$\sigma = \frac{N^{\text{signal,fid}}}{L} \tag{7.1}$$

is relayed, where $N^{\text{signal,fid}}$ is the number of signal events in the fiducial volume and L is the integrated luminosity of the dataset. Since cross-section measurements are often performed for a given observable X described by a binned distribution, the differential fiducial cross section is instead reported and is defined for a given bin i as

$$\frac{d\sigma_i}{dX_i} = \frac{N_i^{\text{signal,fid}}}{L\Delta X_i}, \tag{7.2}$$

where ΔX_i is the bin width. However, the fiducial volume defined in the measurement is not perfectly replicated by the detector due to several factors including response effects. Instead, the number of events measured by the detector N is related to the number of events in the fiducial region N^{fid} such that

$$N^{\text{fid}} = N\mathcal{R}, \tag{7.3}$$

where $\mathcal{R}$ is a complex function that broadly describes the detector response. Furthermore, background processes mimic the signature of the signal process such that the measured number of events N^{meas} is instead given by

$$N^{\text{meas}} = N^{\text{signal}} + N^{\text{bg}}, \tag{7.4}$$

where N^{bg} is the expected number of background events.

Equation (7.2) for the differential fiducial cross section in bin i of the given observable X now becomes

$$\frac{d\sigma_i}{dX_i} = \frac{\mathcal{R}(N^{\text{meas}} - N^{\text{bg}})}{L\Delta X_i}. \tag{7.5}$$

The differential fiducial cross section of Equation (7.5) encompasses the analysis strategy for the cross-section measurements of Z+high-p_{T} jets production. Three unknowns need to be evaluated: 1) the measured data N^{meas}, 2) the estimated background N^{bg}, and 3) the detector response function $\mathcal{R}$. The data is measured at the detector level to produce kinematic distributions of the different observables as described in Chapter 6. The different backgrounds are described and estimated in Chapter 8. The response function is introduced in Section 7.3. Once all three unknowns are evaluated, the fiducial cross sections can be derived.

7.2 Fiducial selection

Although the detector-level selection used in this analysis has been described in Section 6.3, it is not the fiducial definition in which the cross-section measurements are performed. The fiducial selection is applied to particle-level objects and allows for the cross-section measurements to be decoupled from the ATLAS detector through an unfolding procedure. MC samples are used to model the relationship between the underlying particle-level events and the associated simulated detector-level reconstructed events. Therefore, detector-independent distributions can easily be compared with theoretical predictions or results from other experiments.

The cross-section measurements presented in this analysis are performed within the fiducial acceptance region defined by the following requirements:

- Two same-flavor, opposite-charge leptons with $p_{\mathrm{T}} \geq 25$ GeV and $|\eta| < 2.5$,
- 71 GeV $\leq m_{\ell\ell} \leq$ 111 GeV,
- At least one jet, where jets must have $p_{\mathrm{T}} \geq 100$ GeV and $|y| < 2.5$.

The *inclusive*, *high-*p_{T}, *collinear*, *back-to-back*, and *high-*S_{T} signal fiducial regions are defined in analogy to the detector-level definitions defined in Section 6.4.

Particle-level leptons are used to define the fiducial volume. They correspond to 'dressed' electrons and muons. The dressing procedure corrects for effects of final-state photon radiation. Dressed lepton are defined as the four-vector combination of a prompt lepton (that does not originate from the decay of a hadron or a τ-lepton, or from a photon conversion) and all prompt photons within a surrounding cone of size $\Delta R = 0.1$. Particle-level jets are found by applying the anti-k_t algorithm with radius parameter $R = 0.4$ to final-state particles with decay length $c\tau > 10$ mm, excluding dressed Z-boson decay products.

The same overlap procedure from the detector-level selection, defined in Section 6.2.4, is applied at the particle level. First, jets with $p_{\mathrm{T}} \geq 30$ GeV within $\Delta R = 0.2$ of a dressed lepton are removed. Second, leptons within $\Delta R = 0.4$ of the remaining jets are removed. This overlap removal is applied at particle level in order to best match the detector response, especially in the *collinear* region where the detector is not able to discriminate easily between nearby objects. Ideally, the overlap removal would be loosened in the fiducial volume of the measurement to allow for a more precise measurement of the *collinear* region, where jets and leptons are often very nearby. However, because of the difficulty in resolving very nearby objects within the ATLAS detector, the extrapolation from the detector-level measurement to the particle-level fiducial volume cannot be perfect and makes it necessary to apply the same overlap removal at both the detector and particle level.

This analysis includes the fixed-order theoretical cross-section predictions from NNLOJET, as discussed in Section 5.3. These predictions correspond to Born-level leptons, where the electrons and muons don't undergo the effects of final state radiation, and thus do not perfectly replicate the above fiducial selection. Hence, to translate between the Born- and dressed-level leptons, scale factors are computed with SHERPA 2.2.11 and applied to the predictions from NNLOJET. Additionally, the particle-level overlap removal procedure is not perfectly replicated in the NNLOJET program due to technical limitations. To translate between the different overlap removal procedures, scale factors are consequently computed with SHERPA 2.2.11 and applied to the NNLOJET predictions.

7.3 Unfolding of detector effects

For an ideal detector, the measurement of a quantity x would yield the observable x such that the reconstructed (or detector-level) distribution would equal exactly the true (or particle-level) distribution. In reality, the measurement of the quantity x also includes a convolution with the response of the detector. Detector effects such as its finite resolution smears the measured quantity x. Others, like its limited acceptance and reconstruction efficiency, result in a number of particles that are partially or not at all measured. Additionally, pile-up can produce false signatures of signal events, or fake events, at the detector level. In consequence, these detector effects blur the measurements of the given true value x. The process of correcting for the detector efficiency, pile-up, and resolution effects is called unfolding. It is used to correct for detector acceptance and extrapolate the measured event yields to particle-level event yields.

The problem of unfolding is particularly challenging in particle physics experiments. The sheer size and complexity of the detectors make the modelling of the detector effects complex. Furthermore, the simulated signal process must accurately describe the data for the detector response modelling to be accurate. The guiding principle behind unfolding is the presence of a detector response function that can be applied to the true distribution of an observable in order to produce the measured distribution. Mathematically, this can be expressed through the use of probability density functions (PDF). The PDF of a given measurement x, $f_{\text{meas}}(x)$, is expressed as

$$f_{\text{meas}}(x) = \int R(x|y) f_{\text{true}}(y) dy, \tag{7.6}$$

where $f_{\text{true}}(y)$ is the true PDF of the observable y and $R(x|y)$ is the detector response function. Since the cross-section measurements in this analysis correspond to binned distributions, the formalism is applied to binned distributions instead of continuous probability density functions. Without loss of generality, $f_{\text{meas}}(x)$ and $f_{\text{true}}(y)$ are replaced by the expected number of detector-level and particle-level events x and y, respectively. Equation (7.6) is then transformed into the matrix equation

$$x_i = \sum_{j=1}^{k} R_{ij} y_j, \tag{7.7}$$

where i and j represent the detector- and particle-level bins, respectively. Here, k is the total number of bins and R_{ij} is the response matrix describing the detector effects between the i-th and j-th bins of the distribution of the y observable. In the presence of background processes, the total number of expected events in bin i is equal to

$$x_i = \sum_{j=1}^{k} R_{ij} y_j + b_i, \tag{7.8}$$

where b_i is the expected number of background events in the i-th bin. In this formulation, the background contribution is estimated directly at the detector-level and the response matrix only describes the detector effects caused by the signal process. To compute the expected number of particle-level events, Equation (7.8) can be transformed to

$$y_j = \sum_{i=1}^{k} R_{ij}^{-1} (y_i - b_i). \tag{7.9}$$

Equation (7.9) describes the unfolding process. By inverting the response matrix and applying it to the background-subtracted data, the true spectrum of the y observable is extracted. The computation of the response matrix is not a trivial endeavour, nor is its inversion. This matrix equation can be solved using a myriad of techniques, each with their advantages and disadvantages. The unfolding method used in this analysis is performed with the iterative Bayesian unfolding, as detailed below.

With the formalism introduced in Section 7.1, the differential fiducial cross section of a given observable X in the j-th bin becomes

$$\frac{d\sigma_j}{dX_j} = \frac{\sum_{i=1}^{k} R_{ij}^{-1} (N_i^{\text{meas}} - N_i^{\text{bg}})}{L \Delta X_j}. \tag{7.10}$$

The response matrix R_{ij} is estimated using simulated event samples. In order to enter the response matrix, events must satisfy the detector-level or particle-level selections described in Section 6 and Section 7.2. Events that pass both selections provide crucial information on the bin-by-bin migration due to detector resolution. The events that only pass the particle-level selection are used to compute the detector efficiency corrections of the response matrix. Moreover, events that only pass the detector-level selection are used to estimate the rate of unmatched signal events.

7.3.1 Iterative Bayesian unfolding

An iterative unfolding technique [156] with two iterations, as implemented in the RooUnfold package [157], is used to unfold the background-subtracted data to the particle level, thereby accounting for the impact of detector efficiencies and resolution.

The iterative technique is driven by the Bayesian theorem. The probability density function of a given measurement x, $f_{\text{meas}}(x)$, as a function of the true PDF of the observable y, $f_{\text{true}}(y)$, and the conditional probability of measuring x given y, $R(x|y)$, was previously defined in Equation (7.6). The equation can be inverted to obtain

$$f_{\text{true}}(y) = \int R(y|x) f_{\text{meas}}(x) dx. \tag{7.11}$$

This equation describes the unfolding of data previously introduced where the measurement of x is unfolded to the true distribution of y. It is crucial to point out that $R(y|x)$ is actually a function of $f_{\text{true}}(y)$ itself. As a result, the solution to Equation (7.11) requires some prior knowledge of $f_{\text{true}}(y)$. This type of problem frequently arises in many fields of physics, labelled as non-linear problems, and can often be solved with an iterative method. In the iterative method, the problem is first solved by using an initial estimate of its parameters. The solution is then used to compute new parameter values and the problem is solved again. This is repeated through several iterations until certain conditions on the solution are met.

In the iterative unfolding technique, an initial estimate of $R(y|x)$, evaluated using a MC simulation, is used to compute the solution $f_{\text{true}}(y)$. The solution is then used to calculate a new $R(y|x)$ and the procedure is repeated through a number of unfolding iterations. By increasing the number of iterations, the bias from an incorrect modelling of the response matrix is decreased, but at the cost of an increase in statistical uncertainty. For well understood processes such as Z+jets, the response matrix is mostly stable such that it does not change much between iterations.

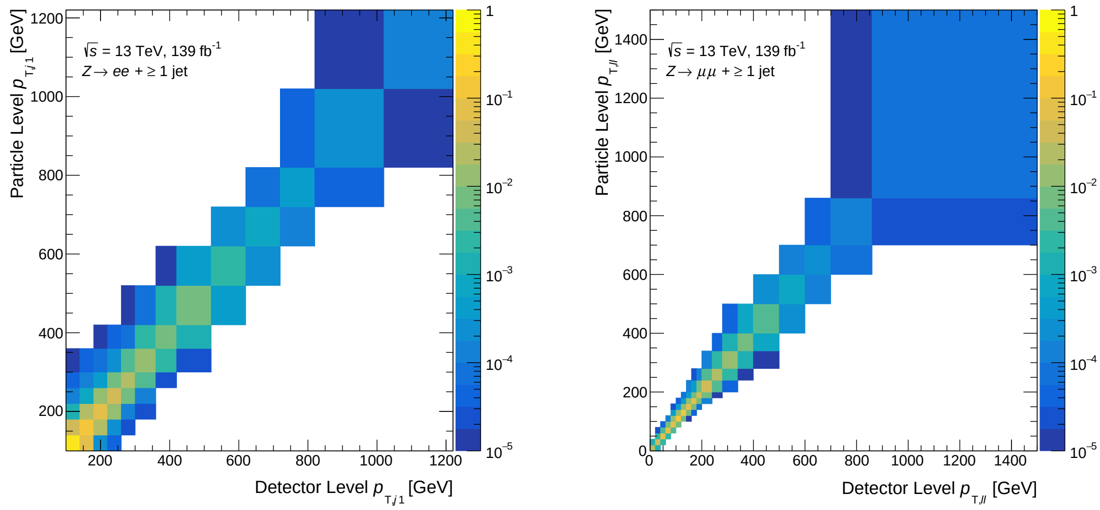
In this analysis, the optimal number of unfolding iterations is found to be two. Two iterations provided minimal bias and low statistical uncertainties when compared to higher number of iterations. The size of the bias is included as a source of uncertainty on the final differential cross sections, as discussed in Chapter 9.

The electron and muon Z-boson decay channels are unfolded separately to best account for the differences in detector performance between the channels, reducing the unfolding bias. Afterwards, the unfolded results are combined to measure the production cross section for a Z boson decaying into a single charged-lepton flavor: $Z(\rightarrow \ell^+\ell^-)$. The combination is done to increase the precision of the final Z+high-p_{T} jets measurements.

7.3.2 Detector response

For this analysis, the response matrix is computed using simulated event samples from Sherpa 2.2.11 to model the detector response to Z+high-p_{T} jets events. The response matrix captures effects that stem from: 1) bin-by-bin migration, 2) detector efficiencies, and 3) events that pass the detector-level selection but not the particle-level selection, known as ‘unmatched events’.

Three distributions can be extracted from the response matrix to characterize

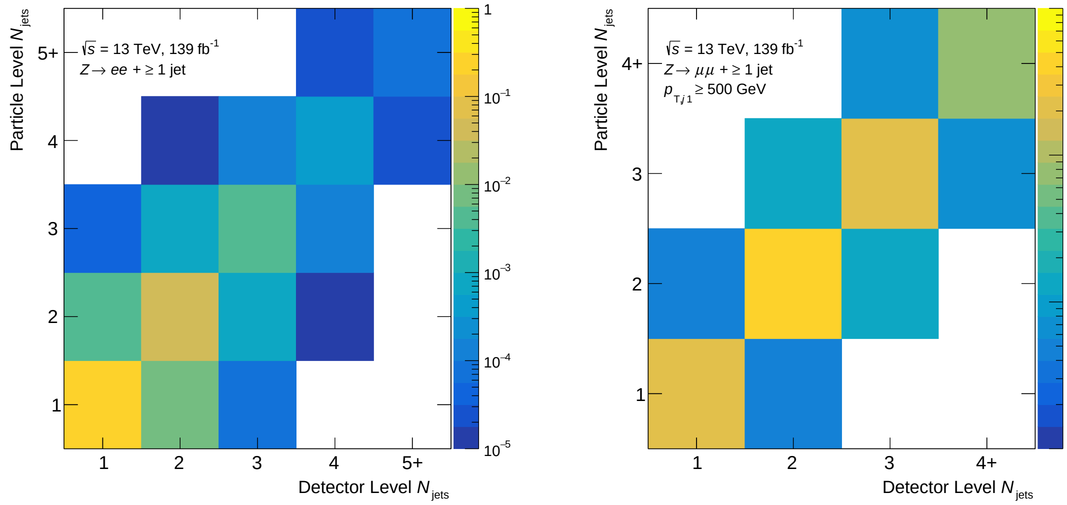


(a) (b)

Figure 7.1: Migration matrix obtained with Sherpa 2.2.11 in the *inclusive* region (a) of the transverse momentum of the leading jet $p_{\mathrm{T},j1}$ for the electron channel and (b) of the transverse momentum of the Z boson $p_{\mathrm{T},\ell\ell}$ for the muon channel. The number of entries (colour coded) is normalized to unity.

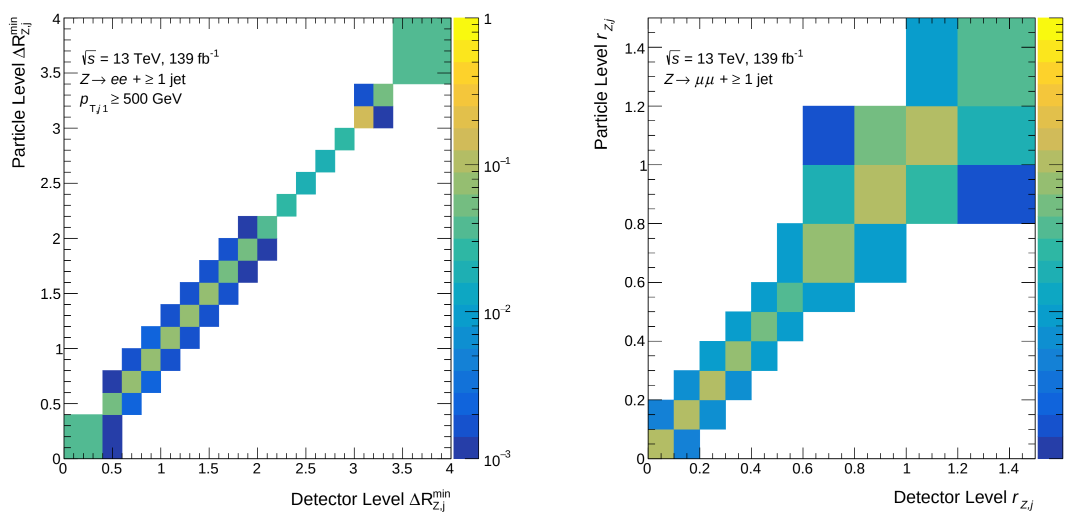
these effects. The first distribution, the migration matrix, describes the bin-by-bin migration of events due to the limited detector resolution. It only includes events that pass both the detector- and particle-level selections. A given bin (i, j) in the migration matrix is related to the probability that a true object in bin j is reconstructed in bin i of the distribution. The migration matrices for a few selected observables are shown in Figures 7.1–7.3. The migration matrices for these cross-section measurements follow a very diagonal behaviour for every observable and Z-boson decay channel. Diagonal migration matrices indicate that the binning is well optimized and the detector response is well understood such that detector-level quantities are reconstructed in the same particle-level bins.

The second distribution describes the detector-level reconstruction efficiency. It corresponds to the fraction of events in a given particle-level bin i that also pass the detector-level selection. The reconstruction efficiency distributions for a few se-



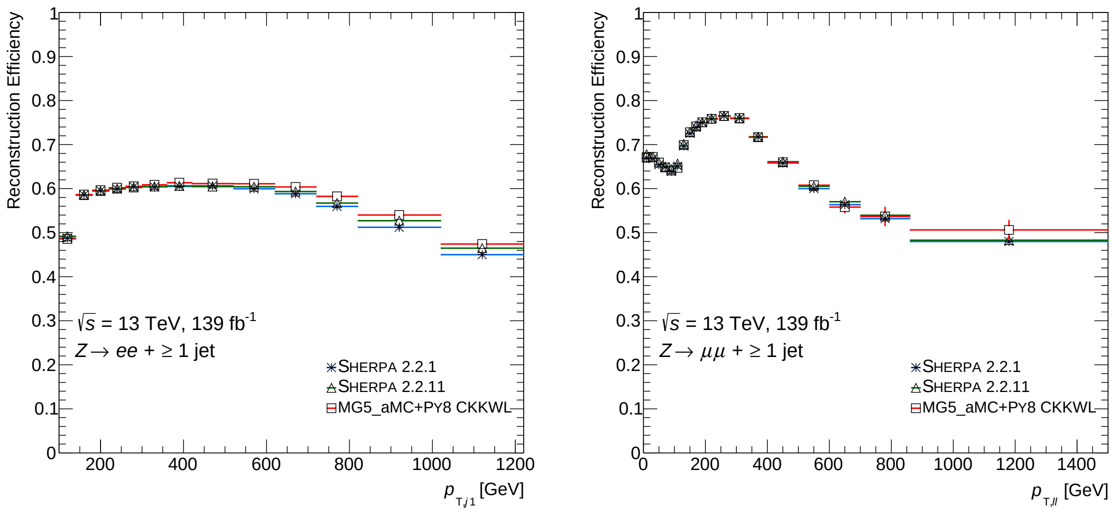
(a) (b)

Figure 7.2: Migration matrix obtained with SHERPA 2.2.11 of the jet multiplicity N_{jets} (a) in the *inclusive* region for the electron channel and (b) in the *high*-p_{T} region for the muon channel. The number of entries is normalized to unity.



(a) (b)

Figure 7.3: Migration matrix obtained with SHERPA 2.2.11 in the *high*-p_{T} region (a) of angular distance between the Z boson and the closest jet $\Delta R_{Z,j}^{\text{min}}$ for the electron channel and (b) of the ratio of Z-boson to closest-jet transverse momenta $r_{Z,j}$ for the muon channel. The number of entries is normalized to unity.

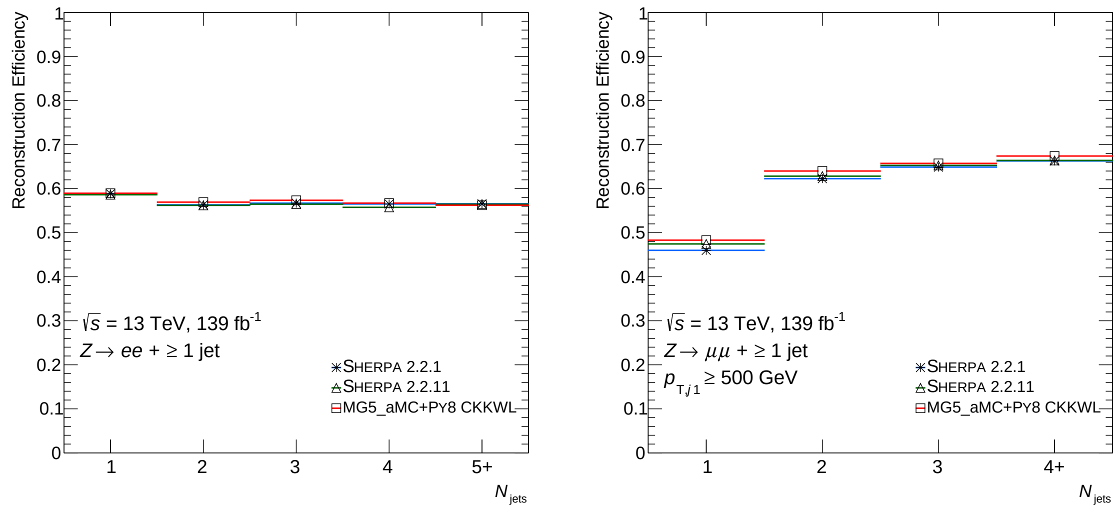


(a) (b)

Figure 7.4: Detector-level reconstruction efficiency in the *inclusive* region as a function (a) of the transverse momentum of the leading jet $p_{\mathrm{T},j1}$ for the electron channel and (b) of the transverse momentum of the Z boson $p_{\mathrm{T},\ell\ell}$ for the muon channel. The reconstruction efficiency from the SHERPA 2.2.1 and MG5_aMC+PY8 CKKWL Z+jets samples are also included to validate the performance of SHERPA 2.2.11.

lected observables are shown in Figures 7.4–7.6. Reconstruction efficiencies typically fall between 40% and 80% and the distributions follow slowly increasing or decreasing trends. The efficiencies are mostly due to the detector-level trigger, isolation and identification requirements and do not point to severely misunderstood detector performances. The reconstruction efficiencies in this analysis are typical of Z+jets measurements.

The third distribution describes the fraction of unmatched events. It corresponds to the fraction of events in a given detector-level bin i that do not pass the particle-level selection. The fraction of unmatched events for a few selected observables are shown in Figures 7.7–7.9. In the *inclusive* region, the fraction of unmatched events is explained by events where the particle-level $p_{\mathrm{T},j1}$ is just below 100 GeV but the detector-level jet is reconstructed just above that selection threshold. This be-



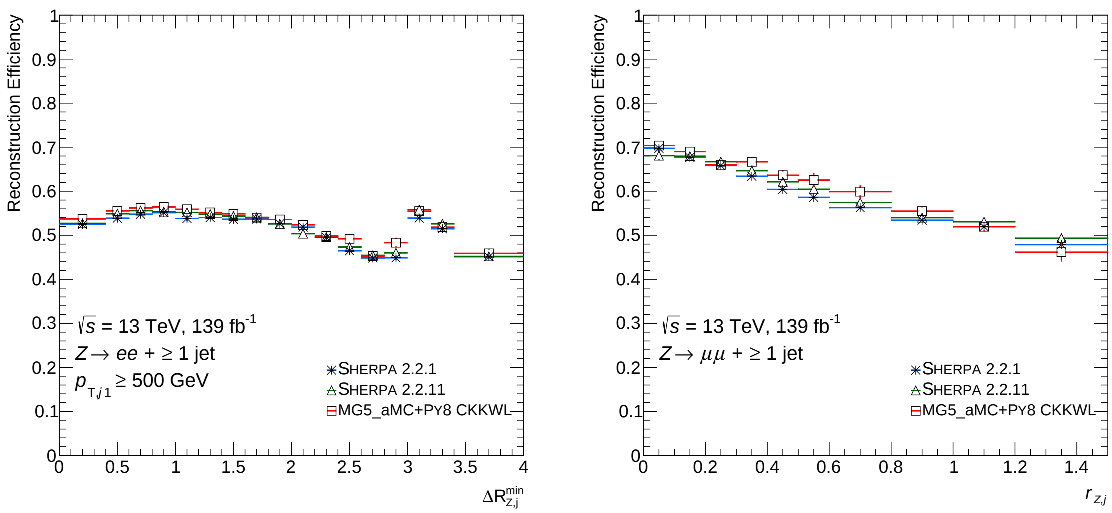
(a) (b)

Figure 7.5: Detector-level reconstruction efficiency as a function of jet multiplicity N_{jets} (a) in the *inclusive* region for the electron channel and (b) in the *high*-p_{T} region for the muon channel. The reconstruction efficiency from the Sherpa 2.2.1 and MG5_aMC+Py8 CKKWL Z+jets samples are also included to validate the performance of Sherpa 2.2.11.

haviour is observed specifically in Figure 7.7a where the first bin in $p_{\text{T},j1}$ has a large fraction of unmatched events. The same phenomenon causes the generally larger event of unmatched events in the *high*-p_{T} region, where a high-p_{T} jet is found only at the detector level.

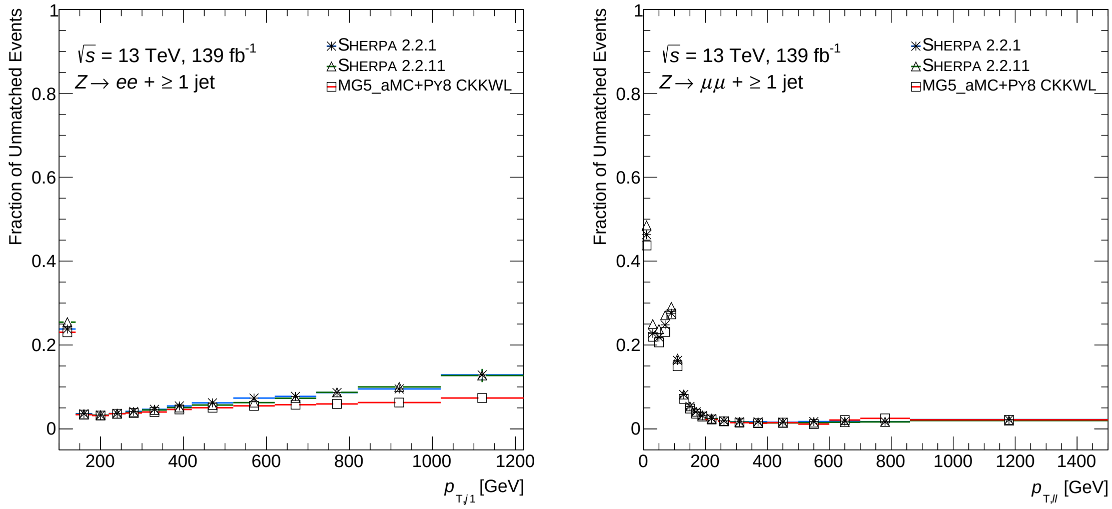
7.3.3 Validation of the unfolding procedure

This section presents the closure tests that serve as validation to the correct implementation of the unfolding procedure. The closure test presented here consists in unfolding the detector-level distributions from Sherpa 2.2.11 with the response matrix evaluated with Sherpa 2.2.11. By definition of the unfolding procedure, the resulting distribution should be exactly equal to the particle-level prediction from Sherpa 2.2.11.



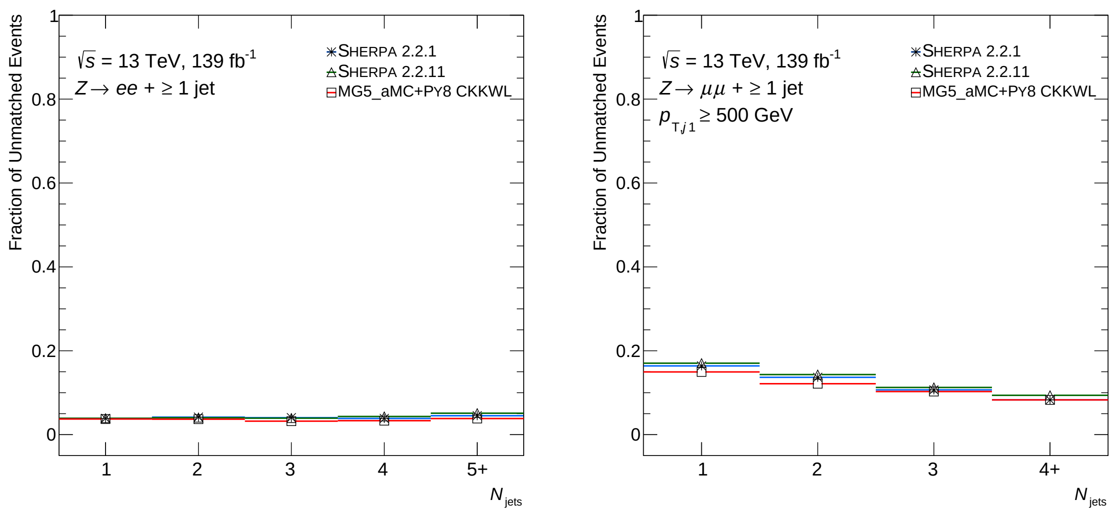
(a) (b)

Figure 7.6: Detector-level reconstruction efficiency in the *high*-p_T region as a function (a) of angular distance between the Z boson and the closest jet $\Delta R^{\mathrm{min}}_{Z,j}$ for the electron channel and (b) of the ratio of Z-boson to closest-jet transverse momenta $r_{Z,j}$ for the muon channel. The reconstruction efficiency from the Sherpa 2.2.1 and MG5_aMC+Py8 CKKWL Z+jets samples are also included to validate the performance of Sherpa 2.2.11.



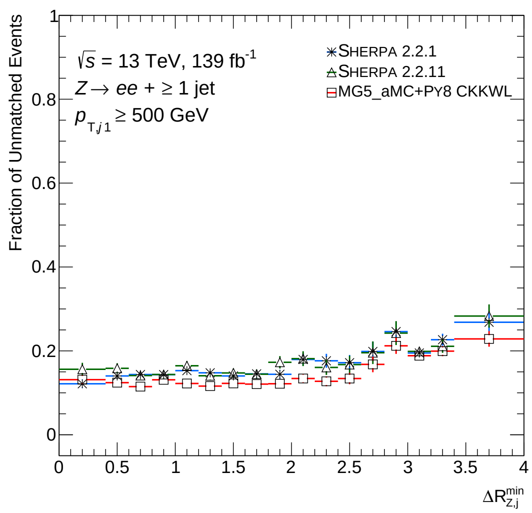
(a) (b)

Figure 7.7: Fraction of unmatched events in the *inclusive* region as a function (a) of the transverse momentum of the leading jet $p_{\mathrm{T},j1}$ for the electron channel and (b) of the transverse momentum of the Z boson $p_{\mathrm{T},\ell\ell}$ for the muon channel. The distributions from the Sherpa 2.2.1 and MG5_aMC+Py8 CKKWL Z+jets samples are also included to validate the performance of Sherpa 2.2.11.

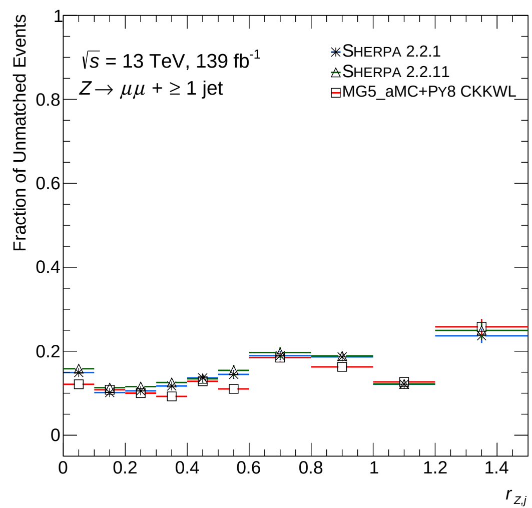


(a) (b)

Figure 7.8: Fraction of unmatched events as a function of jet multiplicity N_{jets} (a) in the *inclusive* region for the electron channel and (b) in the *high-*p_{T} region for the muon channel. The distributions from the Sherpa 2.2.1 and MG5_aMC+Py8 CKKWL Z+jets samples are also included to validate the performance of Sherpa 2.2.11.



(a)



(b)

Figure 7.9: Fraction of unmatched events in the *high*-p_{T} region as a function (a) of angular distance between the Z boson and the closest jet $\Delta R^{\mathrm{min}}_{Z,j}$ for the electron channel and (b) of the ratio of Z-boson to closest-jet transverse momenta $r_{Z,j}$ for the muon channel. The distributions from the SHERPA 2.2.1 and MG5_aMC+PY8 CKKWL Z+jets samples are also included to validate the performance of SHERPA 2.2.11.

Figures 7.10–7.12 show the resulting closure tests for a few selected observables. The unfolded distribution are compared to the particle-level distributions, labelled 'Truth'. For all the observables, in all the bins, the two distributions are exactly equal and validate the implementation of the unfolding procedure.

In summary, this chapter provides the analysis and unfolding strategy for the cross-section measurements of Z+high-p_{T} jets. Data are unfolded with a Bayesian unfolding method, correcting for detector effects, and results in fiducial, or particle-level, cross-section measurements. Details about the background estimation are presented in the next chapter prior to introducing the different sources of uncertainties and finally the results of the fiducial cross-sections measurements.

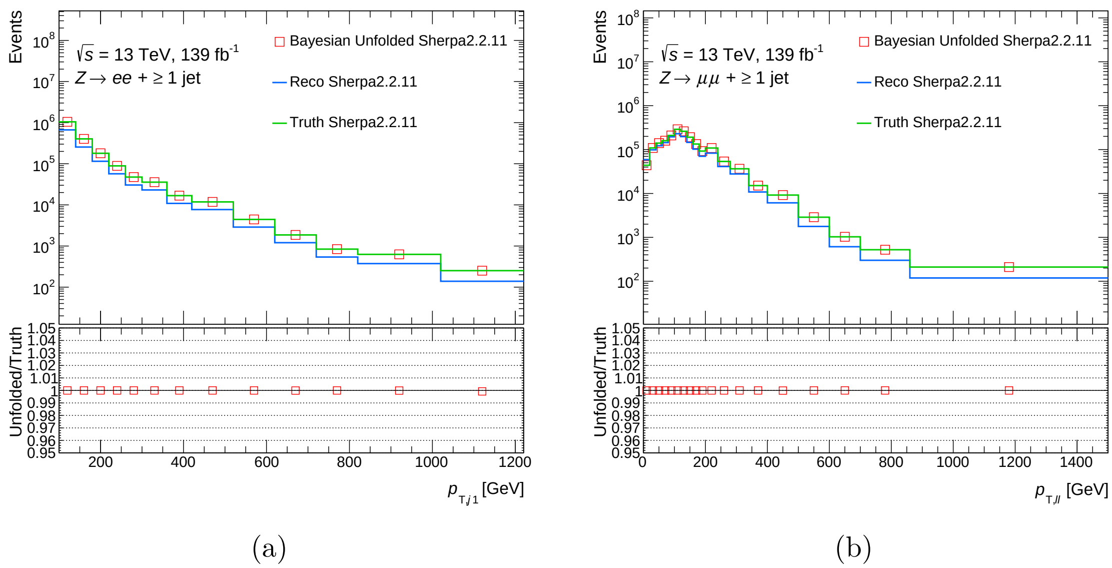
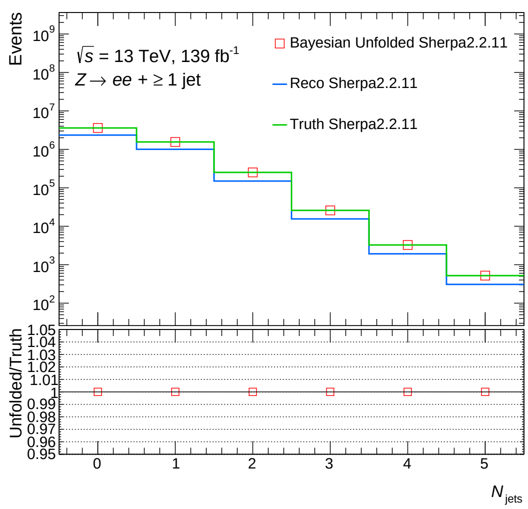
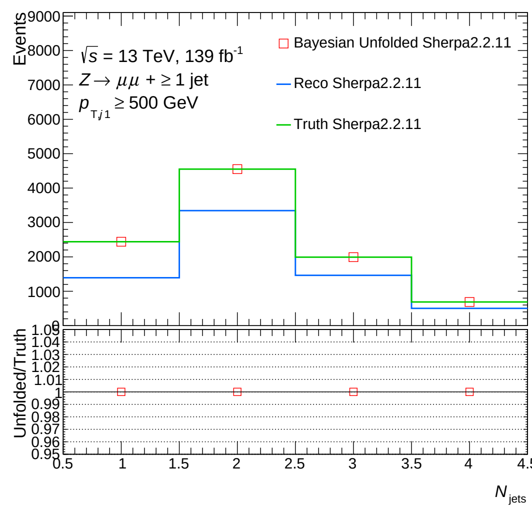


Figure 7.10: Unfolding closure test distributions in the *inclusive* region (a) of the transverse momentum of the leading jet $p_{\mathrm{T},j1}$ for the electron channel and (b) of the transverse momentum of the Z boson $p_{\mathrm{T},\ell\ell}$ for the muon channel. By definition, the particle-level distributions ('Truth') should be exactly equal to the unfolded distribution (red squares). The Bayesian unfolding method is used for the cross-section measurements in this analysis.

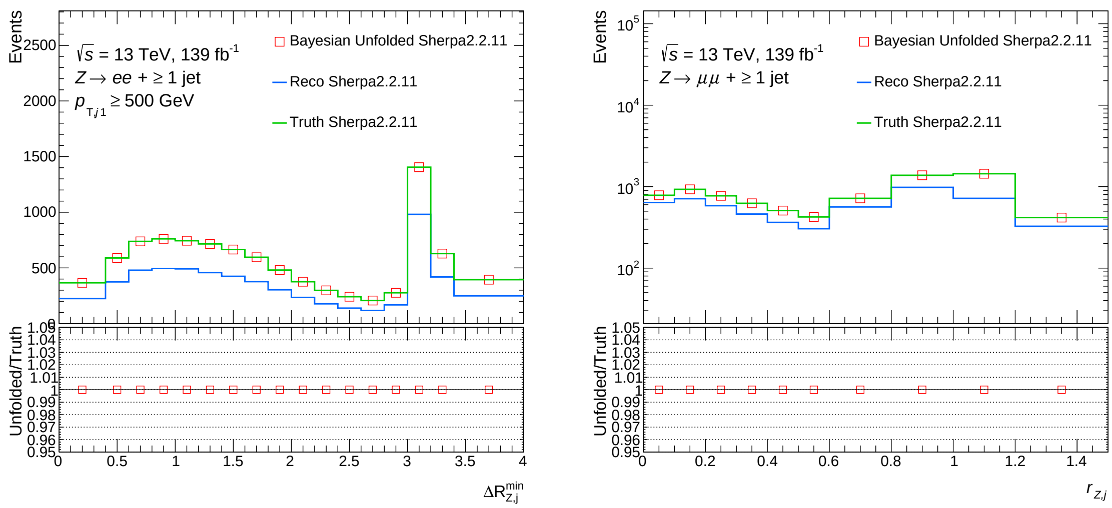


(a)



(b)

Figure 7.11: Unfolding closure test distributions of jet multiplicity N_{jets} (a) in the *inclusive* region for the electron channel and (b) in the *high-*p_{T} region for the muon channel. By definition, the particle-level distributions ('Truth') should be exactly equal to the unfolded distribution (red squares). The Bayesian unfolding method is used for the cross-section measurements in this analysis.



(a) (b)

Figure 7.12: Unfolding closure test distributions in the *high*-p_{T} region (a) of angular distance between the Z boson and the closest jet $\Delta R_{Z,j}^{\mathrm{min}}$ for the electron channel and (b) of the ratio of Z-boson to closest-jet transverse momenta $r_{Z,j}$ for the muon channel. By definition, the particle-level distributions ('Truth') should be exactly equal to the unfolded distribution (red squares). The Bayesian unfolding method is used for the cross-section measurements in this analysis.

Chapter 8

Background estimation

The clean signature of the Z boson decaying into two opposite-charge leptons makes the selection of strongly produced Z+jets events relatively straightforward. There are few background processes that can mimic the two lepton final state with dilepton invariant mass around that of the Z boson, 71 GeV $\leq m_{\ell\ell} \leq$ 111 GeV. As such, the background contributions are expected to be small and are mainly from $t\bar{t}$, diboson and EW Zjj production. Backgrounds from single-boson, diboson and single-top-quark processes are estimated using the MC samples described in Section 5.2. The backgrounds from $t\bar{t}$ production and 'multijet events' are estimated from data.

Sections 8.1 and 8.2 describe the data-driven evaluations of the $t\bar{t}$ and multijet backgrounds, respectively. Section 8.3 introduces the different background processes that are estimated using Monte Carlo event samples. The quantum mechanical interference between the strong Zjj signal and the EW Zjj background is further detailed in Section 8.3.1. The number of measured and predicted events are then counted and summarized in Section 8.4. Finally, the kinematic distributions of all the detector-level measured observables, as well as the predicted signal and backgrounds, are shown in Section 8.5.

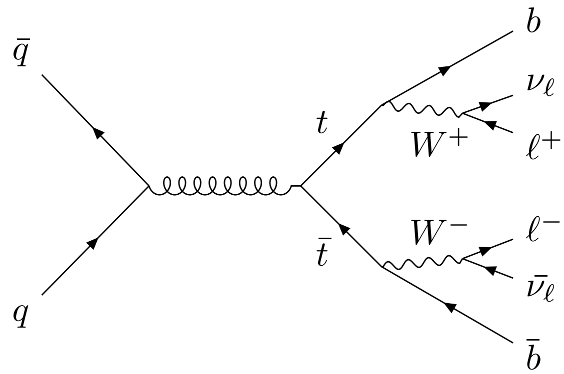
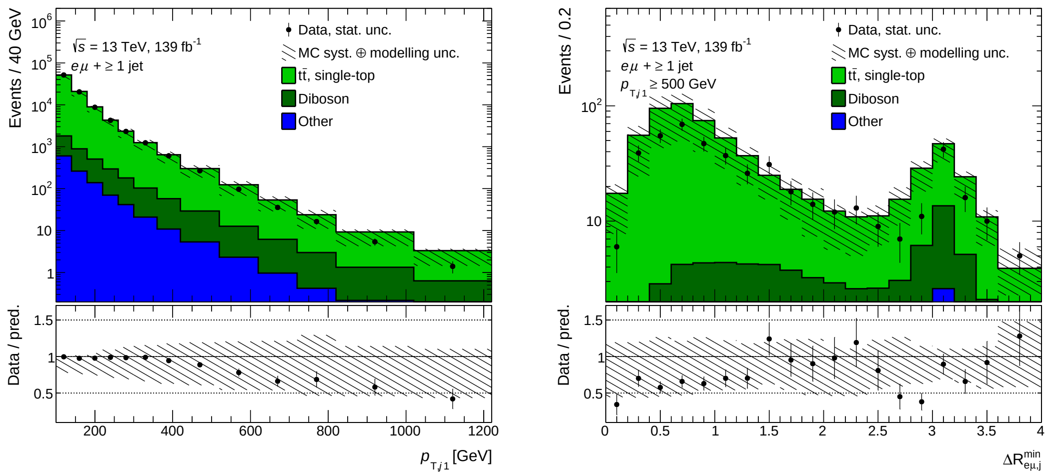


Figure 8.1: Representative Feynman diagram for the $t\bar{t}$ background of Z+jets production. The W bosons decay leptonically to same flavor leptons and the b quark and anti-quark hadronize to produce jets.

8.1 $t\bar{t}$ background

The $t\bar{t}$ processes are a dominant background in this analysis. Produced via gluon fusion or t-channel interactions with a cross section of about 1000 pb at $\sqrt{s} = 13$ TeV, they become a problematic source of background when both W bosons decay leptonically and into the same lepton flavor, as represented in the Feynman diagram of Figure 8.1. Although the $t\bar{t}$ processes can be predicted using MC samples, previous analyses show that the predicted cross sections overestimate the number of $t\bar{t}$ events in phase spaces corresponding to dilepton final states with high-p_{T} jets [158, 159]. Furthermore, the modelling uncertainties of the MC prediction grow to nearly 50% in events with $p_{\mathrm{T}}(\text{jet}) \geq 500$ GeV. These behaviours of the MC $t\bar{t}$ predictions would have a large impact in the final Z+high-p_{T} jets cross-section measurements. The large MC mismodelling of data and the large associated uncertainties are circumvented by estimating the $t\bar{t}$ contribution using data.

A dedicated $t\bar{t}$-enriched control region is constructed with the same event selection as the signal region described in Chapter 6, but with $e^{\pm}\mu^{\mp}$ final states instead of the same-flavor e^+e^- or $\mu^+\mu^-$ pairs. This control region contains only percent-level



(a) (b)

Figure 8.2: Detector-level distributions in the $t\bar{t}$ control region (a) of the transverse momentum of the leading jet $p_{\mathrm{T},j1}$ in the *inclusive* region and (b) of angular distance between the $e\mu$ dilepton system and the closest jet $\Delta R^{\min}_{e\mu,j}$ in the *high*-p_{T} region. In the figure, the $t\bar{t}$ prediction is estimated with MC simulated event samples. All backgrounds and the signal samples are stacked to produced the figures. Z+jets, W+jets and $V+\gamma$ processes are combined and labelled 'Other'. The hatched band corresponds to the sum in quadrature of the instrumental and theoretical uncertainties described in Chapter 9 for the predicted distributions. The theoretical uncertainties include only those from the $t\bar{t}$ prediction. The data statistical uncertainty is displayed with error bars.

contributions from strong Z+jets, W+jets, diboson, EW Zjj, single-top and strong $Z(\to \tau\tau)$+jets events. Hence, in this control region, the $t\bar{t}$ processes are the dominant contribution to the selected data. Figure 8.2 shows the event distributions of $p_{\mathrm{T},j1}$ and $\Delta R^{\min}_{e\mu,j}$ in the $t\bar{t}$ control region, where $\Delta R^{\min}_{e\mu,j}$ is defined as the angular separation between the $e\mu$ dilepton system and the closest jet. In the figure, the $t\bar{t}$ prediction is estimated with MC simulated event samples and carries theoretical uncertainties as described in Chapter 9. A significant disagreement is observed between data and simulation for the $p_{\mathrm{T},j1}$ distribution, especially in the *high*-p_{T} region. The data-driven estimation of the $t\bar{t}$ background is in part motivated by this disagreement.

Fully leptonically decaying $t\bar{t}$ processes will produce one of three leptonic final states: e^+e^-, μ^+, μ^- or $e^\pm\mu^\mp$. Due to lepton universality, every top-quark leptonic decay mode is as likely as the others. As a result, the $e^\pm\mu^\mp$ final states will occur exactly at twice the rate of the e^+e^- and $\mu^+\mu^-$ final states. By using data to estimate the $t\bar{t}$ contribution in $e^\pm\mu^\mp$ + jets events, the $t\bar{t}$ background in the $\ell^+\ell^-$ final states can be extracted.

Theoretically the rate of events between $e^\pm\mu^\mp$ and $\ell^+\ell^-$ states differ by exactly a factor two, but experimentally it is not so simple: the response of the ATLAS detector to electrons and muons is not equal. The trigger, reconstruction, isolation and identification requirements significantly differ between both leptons flavors and cause important differences in the reconstruction efficiencies. These efficiency differences are overcome through the use of simulated $t\bar{t}$ event samples to model the detector response to $t\bar{t}$ events. The simulated $t\bar{t}$ event samples are used to compute $ee/e\mu$ and $\mu\mu/e\mu$ scale factors, defined as the ratio of $\ell^+\ell^-$ to $e^\pm\mu^\mp$ events. The scale factors are computed for each bin of every observable in the signal region.

The prediction for the $t\bar{t}$ distributions in the $\ell^+\ell^-$ signal region is obtained by multiplying the corresponding measured $t\bar{t}$ distributions in the control region by the $ee/e\mu$ and $\mu\mu/e\mu$ scale factors. This method is commonly used in Z+jets measurements [57]. The $t\bar{t}$ distribution in the control region corresponds to the background-subtracted data.

The method is mathematically described by the following equation:

$$N_{\text{data}}^{\ell^+\ell^-} = N_{\text{data}}^{e^\pm\mu^\mp} \times \underbrace{\frac{N_{\text{MC}}^{\ell^+\ell^-}}{N_{\text{MC}}^{e^\pm\mu^\mp}}}_{\text{Scale factor}} , \tag{8.1}$$

where N corresponds to the number of $t\bar{t}$ events in a bin of a given observable. The data and MC subscripts indicate if the events are evaluated using data or MC predictions, and the leptons indicate the signal and control regions.

This procedure is repeated once for every considered source of instrumental uncertainty, as described in Chapter 9. Due to this data-driven approach, the $t\bar{t}$ uncertainties are therefore naturally covered within the different instrumental uncertainties.

The resulting $t\bar{t}$ background constitutes the leading background in the *inclusive* region and the sub-leading background in the *high-*p_{T}, *collinear*, *back-to-back* and *high-*S_{T} regions.

8.2 Multijet background

In the context of this analysis, the multijet background originates from all-hadronic and QCD-only processes. These multijet events result from semileptonic decays of heavy quarks, hadrons misidentified as leptons, and electrons from photon conversion. The multijet production from QCD-only processes has a large cross section at the LHC and is expected to be present in many Z+jets measurements. The high jet-p_{T} and invariant dilepton mass requirements of this analysis are expected to suppress the multijet background, but it is nonetheless investigated.

Multijet events are assessed with a data-driven approach in a multijet-enriched control region. Since the multijet events stem mostly from objects misidentified as leptons, the control region is defined by relaxing or inverting the lepton selection requirements associated with isolation, identification and charge defined in Section 6.2. The transverse impact parameter significance is loosened to enhance the presence of heavy flavor jets. The isolation and identification criteria are then inverted or loosened to select leptons with low likelihoods of being electrons and muons. The same-sign requirement is also dropped to suppress the Z boson production processes.

The multijet background is estimated using a template fit of the $m_{\ell\ell}$ distribution in the *inclusive* Z+high-p_{T} jets region. The template is extracted from the multijet control region. The sub-percent contributions to the multijet template that do not originate from the multijet background are evaluated using simulation and subtracted. The fit is performed over the range 51 GeV $\leq m_{\ell\ell} \leq$ 151 GeV. The contribution from multijet events is then estimated in the invariant-mass interval of the signal region (71 GeV $\leq m_{\ell\ell} \leq$ 111 GeV). The resulting fraction of multijet events is at the sub-percent level and so is neglected and not considered as a source of background.

8.3 Monte Carlo predicted backgrounds

The estimation of strong Z+high-p_{T} jets background processes with MC simulation is discussed in this section. MC generated event samples perform well in describing data and carry reasonable uncertainties. This is the case for the predictions of the EW Zjj, diboson, W+jets, $V+\gamma$, strong $Z(\rightarrow \tau\tau)$+jets and single-top-quark backgrounds.

8.3.1 Electroweak Zjj

The EW Z+2 jets, or Zjj, process is a significant background in this analysis. It is estimated with simulated event samples generated with Herwig, as described in Section 5.2. In measurements of $Z(\rightarrow \ell^+\ell^-)$-enriched phase spaces [7], the Herwig predictions provided the best description of data between the three MC generators often used in ATLAS. Due to the generator's good performance, Herwig event samples are used to describe the EW Zjj background. Although the production cross section of EW Zjj is only about 4×10^{-2} pb at $\sqrt{s} = 13$ TeV, EW Zjj events correspond to real Z boson production and do not simply mimic the strong Z+high-p_{T} jets production.

Although the EW Zjj events generated with HERWIG describe data well, the MC event samples do not account for the quantum mechanical interference between strong and EW Z+2 jets production. In order to properly describe the EW Zjj background, a careful study of the interference is performed with MadGraph and an additional modelling uncertainty is applied on the estimated cross section of EW Zjj production.

8.3.1.1 Matrix elements and interference

The fixed-order cross section was described in Chapter 4. In order to describe the process behind quantum mechanical interference, the formalism is once again used here. Without loss of generality, the cross section σ of a specific $A + B \rightarrow X$ process is related to the scattering amplitude $\mathcal{M}$ by

$$\sigma = \int d\Phi |\mathcal{M}|^2, \tag{8.2}$$

where the integral is taken over final phase space momenta Φ.

The total scattering amplitude for a given process is given by the linear sum of all the possible scattering amplitudes such that

$$\mathcal{M} = \sum_i \mathcal{M}_i. \tag{8.3}$$

Equivalently, the total scattering amplitude can instead be written as the sum of scattering amplitudes mediated by the strong and EW interactions

$$\mathcal{M} = \mathcal{M}_{\text{STRONG}} + \mathcal{M}_{\text{EW}}. \tag{8.4}$$

Using this notation, the cross section of a given process described through both strong and EW interactions is

$$\begin{aligned}\sigma = \int d\Phi|\mathcal{M}|^2 &= \int d\Phi\Big(|\mathcal{M}_{\text{STRONG}} + \mathcal{M}_{\text{EW}}|^2\Big) \\ &= \int d\Phi\Big(|\mathcal{M}_{\text{STRONG}}|^2 + |\mathcal{M}_{\text{EW}}|^2 + \mathcal{M}_{\text{STRONG}}\mathcal{M}^{\dagger}_{\text{EW}} + \mathcal{M}_{\text{EW}}\mathcal{M}^{\dagger}_{\text{STRONG}}\Big).\end{aligned} \tag{8.5}$$

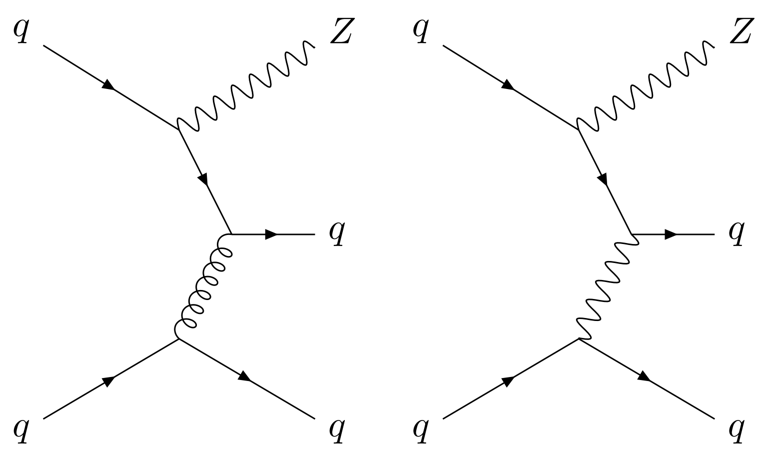
And with the general notation introduced in Equation (8.2), the cross section becomes

$$\sigma = \underbrace{\sigma_{\text{STRONG}}}_{\text{strong cross section}} + \underbrace{\sigma_{\text{EW}}}_{\text{EW cross section}} + \underbrace{\int d\Phi 2\mathcal{R}e\Big(\mathcal{M}^{\dagger}_{\text{STRONG}}\mathcal{M}_{\text{EW}}\Big)}_{\text{Interference}}, \tag{8.6}$$

where σ_{STRONG} and σ_{EW} are the cross sections for the strong and EW processes, respectively. The products of the strong and EW scattering amplitudes here are interpreted as the interference between the processes and may have a non-zero contribution to the total cross section. In some production processes, the interference terms are very small and can be ignored. However, for Z+2 jets production, the interference terms can become as large as the EW cross section in some phase spaces, including those described by the VBF mediated Zjj production [7]. The representative Feynman diagrams describing the strong and EW Zjj production are shown in Figure 8.3, where the two diagrams differ by the internal gauge boson line. The ATLAS collaboration provides MC predictions for both the strong Z+jets and EW Zjj processes, corresponding to σ_{STRONG} and σ_{EW}, with no consideration for the possible interference.

8.3.1.2 Evaluation of the Zjj interference

The interference between the strong and EW mediated Zjj production is determined from simulation using MadGraph5_aMC@NLO 2.9.5. Two set of samples were



(a) Strong Zjj production (b) EW Zjj production

Figure 8.3: Representative Feynman diagrams for the production of Z+2 jets. (a) The strong Zjj production is described by the exchange of a gluon and (b) the EW Zjj production is described with the exchange of a Z boson.

produced: 1) one consisting only of the EW Zjj process, and 2) one consisting only of the interference between the strong and EW mediated Zjj production. These samples are compared at the particle level after applying the event selection requirements of this analysis. In order to be insensitive to the differences between the MadGraph5_aMC@NLO and Herwig generators, the relative size of the interference is obtained by comparing the predictions of MadGraph5_aMC@NLO only. The relative difference between the EW Zjj production and the interference is used to evaluate the additional modelling uncertainty on the EW Zjj predictions from Herwig 7.1.5.

Figure 8.4 shows the differential cross sections of the EW Zjj and the interference the interference between strong and EW mediated Zjj as a function of $p_{\mathrm{T},j1}$, p_{T}(Z) and jet multiplicity in the *inclusive* region, and of $\Delta R^{\mathrm{min}}_{Z,j}$ in the *high-*p_{T} region. The figure also compares the EW Zjj predictions of Herwig 7.1.5 to those from MadGraph5_aMC@NLO. Both generators produce very similar distributions, sharing the same observable shapes and having similar absolute differential cross sections. Since the predictions of both generators are in agreement with one another, rough conclusions drawn from MadGraph5_aMC@NLO can be applied to Herwig. The impact from the interference is largest in events with large $p_{\mathrm{T},j1}$, reaching 40% of the EW Zjj cross section in the *collinear* region. The interference has a small impact on $p_{\mathrm{T}}(\mathrm{Z})$, always staying below 20% and falling to 0% at $p_{\mathrm{T}}(\mathrm{Z}) \approx 200$ GeV. The following conservative modelling uncertainty is chosen to cover all possible effects of interference on the EW Zjj cross section: a flat 25% uncertainty is applied, except in the *collinear* region where an uncertainty of 40% is applied instead, motivated by Figure 8.4d. Since the EW Zjj background corresponds to approximately 1%–3% of the total measured event yields, these added modelling uncertainties have a small impact on the uncertainties of the final strong Z+high-p_{T} jets measurements.

The package containing all the required tools and instructions to evaluate the effects of interference is located on gitlab: `https://gitlab.cern.ch/alaurier/interference-studies` [160].

8.3.2 Diboson, single-diboson and single-top-quark

The generation of the background events described in this section are explored in Section 5.2.

The third significant background in this analysis comes from diboson processes. Diboson backgrounds originate from semi-leptonic and fully-leptonic decays of WW, WZ and ZZ final states. The production cross section of diboson events is approximately 100 pb at $\sqrt{s} = 13$ TeV for events containing a Z boson. The dominant semi-leptonic processes are $W(\rightarrow qq)Z(\rightarrow \ell\ell)$ and $Z(\rightarrow qq)Z(\rightarrow \ell\ell)$, contributing to over 50% of the total diboson events. The following semi-leptonic processes are considered and are found to be negligible contributions to the diboson background:

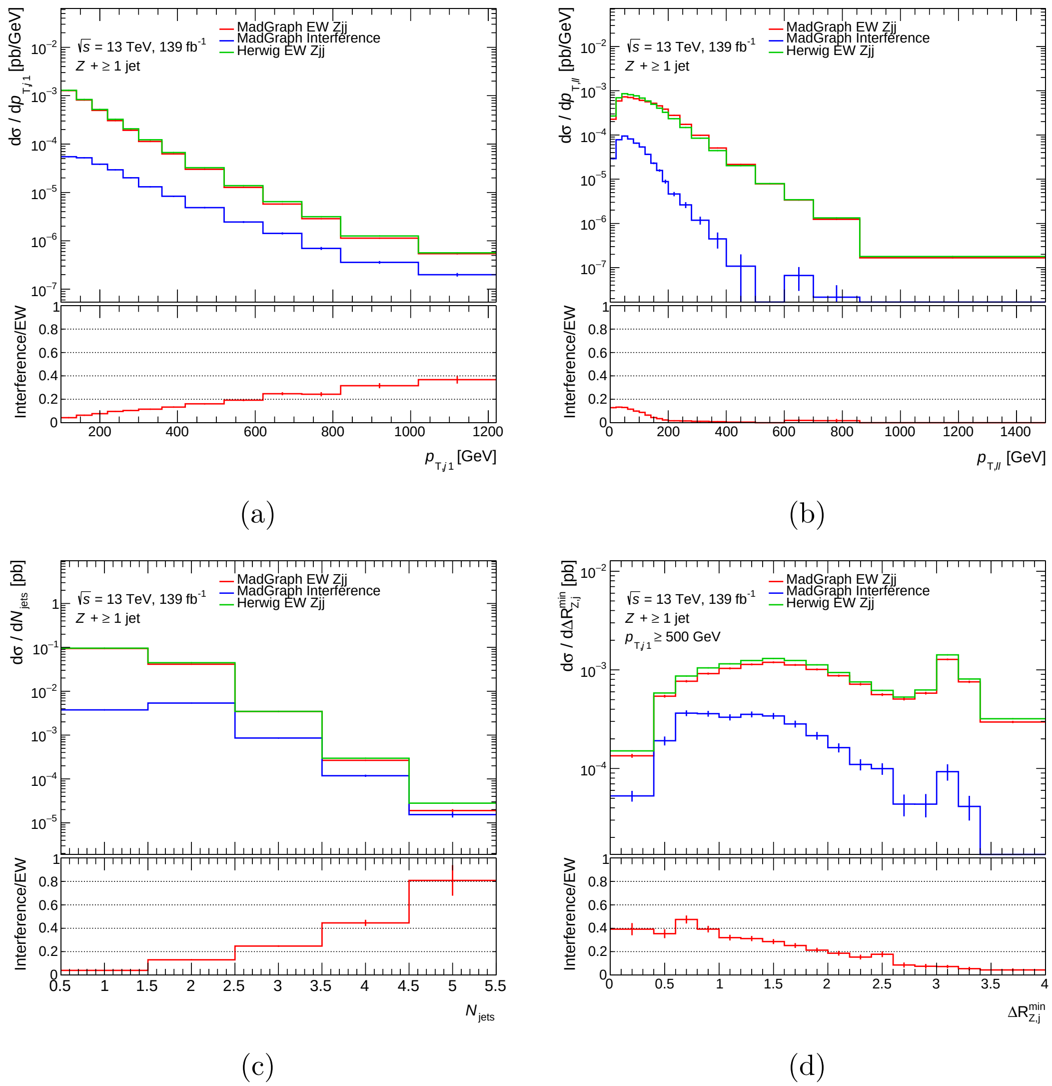


Figure 8.4: Differential cross sections of the EW Zjj and the interference between strong and EW mediated Zjj, in the *inclusive* region as a function (a) of the transverse momentum of the leading jet $p_{\mathrm{T},j1}$, (b) of the transverse momentum of the Z boson $p_{\mathrm{T},\ell\ell}$, (c) of the jet multiplicity N_{jets}, and in the *high*-p_{T} region as a function (d) of angular distance between the Z boson and the closest jet $\Delta R^{\min}_{Z,j}$. The predictions from HERWIG are included to confirm the validity of the MADGRAPH5_aMC@NLO predictions. The MC statistical uncertainty are depicted as error bars.

$(W \to \ell\nu)Z(\to qq)$ and $W^{\pm}(\to \ell\nu)W^{\mp}(\to qq)$. The fully leptonic decays producing $\ell\ell\nu\nu$, $\ell\ell\ell\nu$ and $\ell\ell\ell\ell$ final states represent the second largest contribution to the diboson background. In fully leptonic decays, at least one lepton is reconstructed as a high-p_{T} jet in order to pass the detector-level event selection. Diboson processes are the leading background in the *high-*p_{T}, *collinear*, *back-to-back* and *high-*S_{T} regions, and the sub-leading background in the *inclusive* region. The semi-leptonic and fully-leptonic decays of the WW, WZ and ZZ diboson production are well modelled by the SHERPA MC generator, as recently studied by ATLAS [161–165], and motivates the use of SHERPA to predict the diboson background.

The W+jets, $V+\gamma$, strong $Z(\to \tau\tau)$+jets and single-top-quark backgrounds are also modelled using simulated event samples. These processes contribute at a sub-percent level to the total measured event yield.

8.4 Event yields of selected events

A summary of the composition and relative importance of the backgrounds in the various signal regions is given in Table 8.1 for the electron and muon Z-boson decay channels. The latter table lists the expected number of events from MC predictions for the strong Z+jets signal, as well as all the other sources of backgrounds. The total number of events predicted can then be compared with the number of events in data. The uncertainty on the respective number of events is computed as the sum in quadrature of the statistical and instrumental systematic uncertainties described in Section 9.1.1. The overall purity of the strong Z+jets selections (defined as the expected fraction of signal events after the final selection) ranges from 94% in the *inclusive* region to 87%, 86%, 87% and 88% in the *high-*p_{T}, *collinear*, *back-to-back* and *high-*S_{T} regions, respectively. Backgrounds are dominated by $t\bar{t}$, diboson and

EW Zjj processes, with fractions of 2%–5% , 2%–6% and 1%–5%, respectively. The fraction of events arising from strong $Z(\to \tau\tau)$+jets, W+jets, $V+\gamma$, and multijet backgrounds is below the percent level in all signal regions.

An interesting feature is observed: the measured event yield is higher for muons in the *collinear* region and higher for electron in the *back-to-back* region. The ATLAS detector reconstruction algorithms and overlap removal procedure have issues separating electrons very close to nearby jets, which is a main signature of the *collinear* region. If jets and electrons are close by, the reconstructed jet axis is often shifted towards the electron as the jet incorporates parts of the electron radiation, until the overlap removal passes the threshold of $\Delta R \leq 0.2$ and removes the jet. As a consequence of losing the jet collinear to the Z boson, the event is no longer identified as being in the *collinear* region and will most likely be assigned in the *back-to-back* region. For the muon Z boson decay channel, the jet axis is not shifted unless the muon radiates a photon, which is rare.

8.5 Kinematic distributions of selected events

In this section, some kinematic distributions at the detector level of Z+high-p_{T} jets events passing the full selection criteria from Chapter 6 are shown. These distributions compare the data event yields to the sum of the predicted background and signal processes for some key observables. The event yield distributions are measured separately for each of the electron and muon Z-boson decay channels. The uncertainty band in these figures correspond to the sum in quadrature of the instrumental uncertainties described in Section 9.1.1. Moreover, the data statistical uncertainty is displayed with error bars. In general, the Sherpa 2.2.11 predictions agree with the data, but do not describe it precisely in the full range of the measurements. The

Table 8.1: Event yields in the different Z+jets signal regions for the electron and muon channels. Uncertainties correspond to the statistical and instrumental systematic uncertainties described in Section 9.1.1 added in quadrature. The Z+jets prediction is computed with SHERPA 2.2.11. The number of multijet events is negligible and not included.

$Z \to e^+e^-$	*Inclusive*	*High-p_T*	*Collinear*	*Back-to-back*	*High-S_T*
Strong Z+jets	1 171 000 ± 49 000	6 150 ± 310	2 520 ± 120	2 520 ± 150	18 300 ± 800
$t\bar{t}$	43 400 ± 1 300	209 ± 16	136 ± 13	47.2 ± 7.5	917 ± 41
Diboson	19 530 ± 750	428 ± 29	183 ± 16	167 ± 16	1 008 ± 53
EW Zjj	13 270 ± 500	312 ± 23	102 ± 11	135 ± 14	789 ± 43
Single-top	2 430 ± 160	27.9 ± 5.5	14.0 ± 3.8	9.8 ± 3.2	54.2 ± 8.2
Strong $Z(\to \tau\tau)$+jets	515 ± 37	4.6 ± 4.2	1.6 ± 2.1	2.2 ± 1.7	10.6 ± 6.2
W+jets	93 ± 16	3.4 ± 1.9	0.3 ± 0.6	2.9 ± 1.7	3.4 ± 1.9
$V+\gamma$	1 413 ± 83	14.2 ± 4.3	6.5 ± 2.6	5.1 ± 2.3	34.1 ± 7.3
Total predicted	1 252 000 ± 51 000	7 150 ± 350	2 970 ± 130	2 890 ± 170	21 100 ± 880
Data	1 312 145	7 539	2 955	3 231	21 746
$Z \to \mu\mu$	*Inclusive*	*High-p_T*	*Collinear*	*Back-to-back*	*High-S_T*
Strong Z+jets	1 537 000 ± 63 000	6 700 ± 300	2 950 ± 130	2 420 ± 120	23 110 ± 920
$t\bar{t}$	55 400 ± 1 300	209 ± 16	142 ± 12	39.1 ± 6.6	1 058 ± 41
Diboson	24 160 ± 870	438 ± 27	198 ± 16	157 ± 14	1 149 ± 55
EW Zjj	17 020 ± 580	328 ± 22	113 ± 12	134 ± 13	915 ± 45
Single-top	3 110 ± 190	29.1 ± 5.5	13.6 ± 3.8	11.2 ± 3.5	70.0 ± 9.2
Strong $Z(\to \tau\tau)$+jets	460 ± 33	3.5 ± 4.0	1.1 ± 2.3	1.8 ± 1.5	8.8 ± 5.4
W+jets	128 ± 14	1.9 ± 1.4	0.3 ± 0.5	1.5 ± 1.3	2.7 ± 2.0
$V+\gamma$	1 273 ± 90	2.5 ± 2.4	0.0 ± 0.7	2.2 ± 1.5	22.4 ± 5.5
Total predicted	1 638 000 ± 64 000	7 710 ± 330	3 420 ± 140	2 770 ± 140	26 300 ± 1 000
Data	1 673 057	7 896	3 372	3 059	26 567

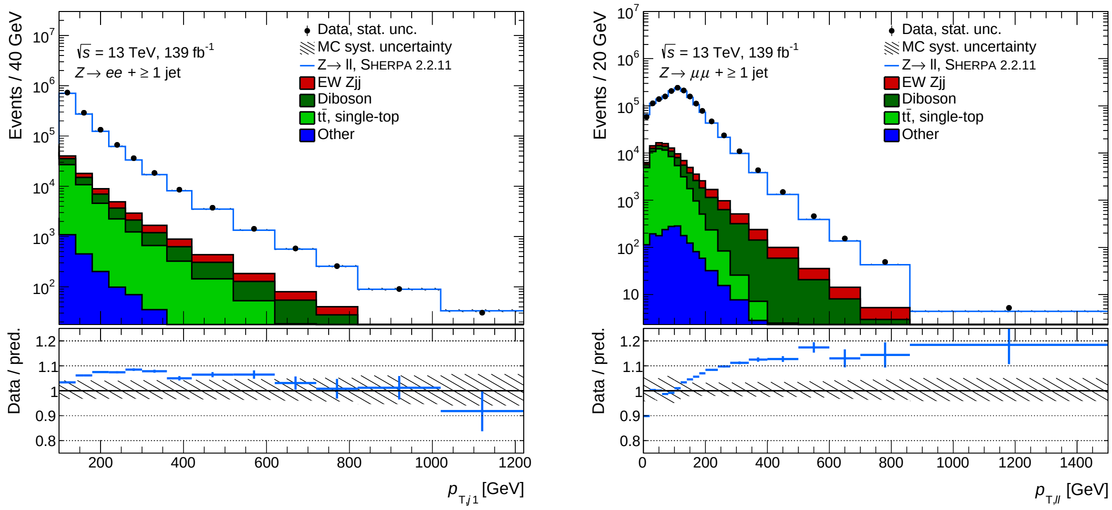
distributions are discussed in more detail when examining the final differential cross sections measurements of Chapter 10.

Figure 8.5 shows the data and predicted detector-level event yields in the *inclusive* region as a function of $p_{T,j1}$ for the electron channel and as a function of $p_{T,\ell\ell}$ for the muon channel. The $p_{T,j1}$ distribution (Figure 8.5a) is represented by a quickly falling distribution as the leading-jet p_T increases. Throughout the range of this observable, the sum of predicted signal and backgrounds describes the data well. The distribution of Z-boson p_T (Figure 8.5b) has a maximum near $p_{T,\ell\ell} = 100$ GeV, an effect partially caused by the jet-p_T selection threshold set at $p_T(\text{jet}) \geq 100$ GeV. As a matter of fact, the *inclusive* region is populated by a majority of Z+1 jet events. In these events, the Z boson is balanced against the jet such that the $p_{\mathrm{T},\ell\ell}$ distribution is described by a majority of events where $p_{\mathrm{T},\ell\ell} \approx p_{\mathrm{T},j1}$. In other Z+jets measurements, where the selected jets have $p_{\mathrm{T}} \geq 25$ GeV, the maximum in the $p_{\mathrm{T},\ell\ell}$ distribution is instead found at lower values closer to the jet-p_{T} threshold [57, 63]. The sum of predicted signal and backgrounds describe the data adequately in the full range of Z-boson p_{T}.

Figure 8.6 shows the data and predicted detector-level event yields as a function of jet multiplicity in the *inclusive* region for the electron channel and in the *high-p_{T}* region for the muon channel. The *inclusive* region is dominated with Z+1 jet events and the distribution quickly falls off with the increasing number of jets. In contrast, the *high-p_{T}* region is instead populated by a majority of Z+2 jets events.

Figure 8.7 shows the data and predicted detector-level event yields in the *high-p_{T}* region as a function of $\Delta R^{\mathrm{min}}_{Z,j}$ for the electron channel and as a function of $r_{Z,j}$ for the muon channel. In the $\Delta R^{\mathrm{min}}_{Z,j}$ distribution (Figure 8.7a), SHERPA 2.2.11 describes the data within uncertainties at the exception of the peak at $\Delta R^{\mathrm{min}}_{Z,j} \sim \pi$. The large number of events at $\Delta R^{\mathrm{min}}_{Z,j} \sim \pi$ populate the *back-to-back* region and are described by Z+1 jet events, where the Z boson and jet are travelling in opposite directions. The *collinear* region, where $\Delta R^{\mathrm{min}}_{Z,j} \leq 1.4$, is described well by SHERPA 2.2.11. The distribution of $r_{Z,j}$ in the *high-p_{T}* region (Figure 8.7b) exhibits two maxima. The first is found at small values of $r_{Z,j}$ and are populated by *collinear* events, where the Z boson is emitted by the nearest jet. The second maximum is found around $r_{Z,j} = 1$ and corresponds to *back-to-back* Z+1 jet events, where the Z boson p_{T} is approximately equal to the jet p_{T}.

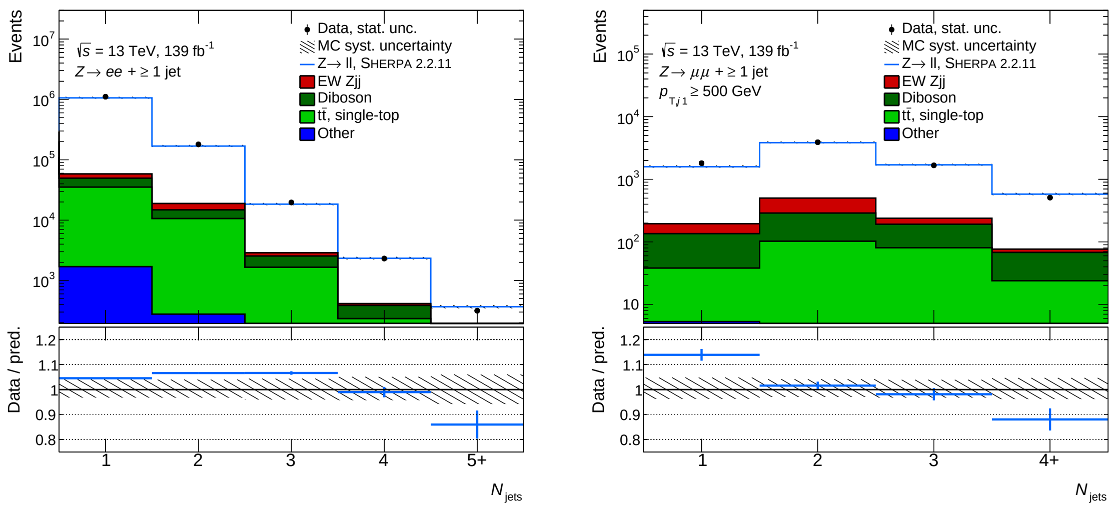
Due to the good performance of SHERPA 2.2.11 in modelling the data, it is



(a) (b)

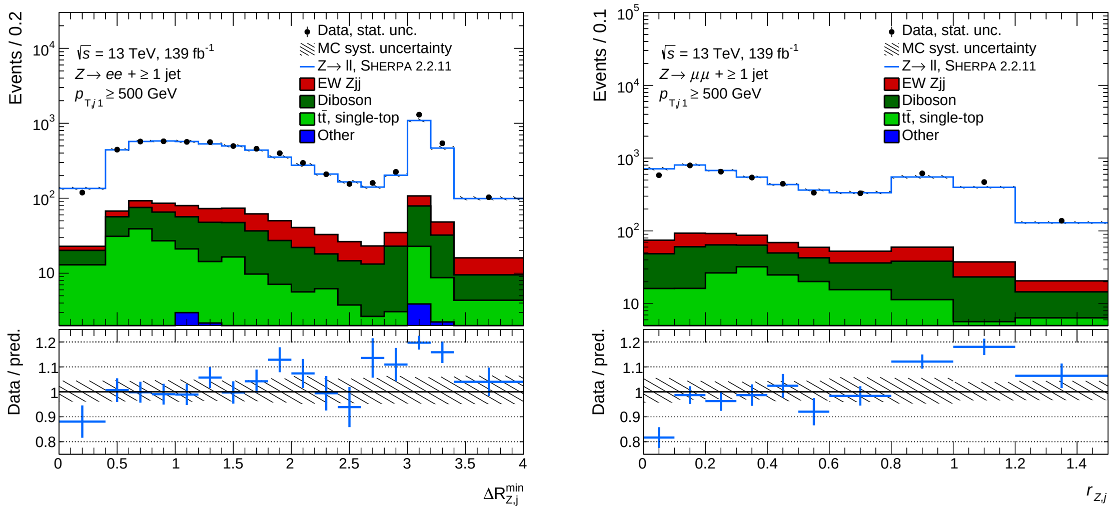
Figure 8.5: Detector-level distributions in the *inclusive* region (a) of the transverse momentum of the leading jet $p_{\mathrm{T},j1}$ for the electron channel and (b) of the transverse momentum of the Z boson $p_{\mathrm{T},\ell\ell}$ for the muon channel. The signal and background samples are stacked to produce the distributions. The W+jets, strong $Z(\to \tau\tau)$+jets and $V+\gamma$ processes are combined and labelled 'Other'. The bottom panel shows the ratio of the data to the total prediction. The hatched band corresponds to the sum in quadrature of the instrumental uncertainties described in Section 9.1.1 for the predicted distributions. The data statistical uncertainty is displayed with error bars.

used to describe the response of the ATLAS detector to Z+high-p_{T} jets events in the unfolding procedure described in Section 7.3 and for the results of the particle-level differential cross sections presented in Chapter 10. Before providing the analysis results, the systematic uncertainties on the cross-section measurements will be presented in Chapter 9.



(a) (b)

Figure 8.6: Detector-level distributions of jet multiplicity N_{jets} (a) in the *inclusive* region for the electron channel and (b) in the *high*-p_{T} region for the muon channel. The signal and background samples are stacked to produce the distributions. The W+jets, strong $Z(\rightarrow \tau\tau)$+jets and $V+\gamma$ processes are combined and labelled 'Other'. The bottom panel shows the ratio of the data to the total prediction. The hatched band corresponds to the sum in quadrature of the instrumental uncertainties described in Section 9.1.1 for the predicted distributions. The data statistical uncertainty is displayed with error bars.



(a) (b)

Figure 8.7: Detector-level distributions in the *high*-p_{T} region (a) of angular distance between the Z boson and the closest jet $\Delta R_{Z,j}^{\mathrm{min}}$ for the electron channel and (b) of the ratio of Z-boson to closest-jet transverse momenta $r_{Z,j}$ for the muon channel. The signal and background samples are stacked to produce the distributions. The W+jets, strong $Z(\rightarrow \tau\tau)$+jets and $V+\gamma$ processes are combined and labelled 'Other'. The bottom panel shows the ratio of the data to the total prediction. The hatched band corresponds to the sum in quadrature of the instrumental uncertainties described in Section 9.1.1 for the predicted distributions. The data statistical uncertainty is displayed with error bars.

Chapter 9

Systematic uncertainties

Systematic uncertainties on the cross-section measurements arise from several independent sources. These include uncertainties describing the detector's performance, uncertainties on the background estimation, uncertainties on the unfolding procedure and finally the uncertainty on the integrated luminosity of the dataset. The resulting cross-section measurements are compared to theoretical predictions that also carry their own modelling uncertainties.

In this chapter, the different sources of uncertainties are discussed. The different experimental uncertainties are introduced in Section 9.1. The breakdown of the total uncertainties on the fiducial cross-section measurements of Z+high-p_{T} jets production is provided at the end of Section 9.1. The theoretical uncertainties associated with the different MC and fixed-order predictions are then discussed in Section 9.2.

9.1 Experimental uncertainties

The different sources of experimental systematic uncertainties on the differential cross-section measurements are presented here. Sources of experimental errors on cross-section measurements stem from instrumental, signal and background event modelling, and unfolding uncertainties. To evaluate the systematic uncertainty of a given source, the corresponding quantity is varied by one standard deviation ($\pm$ 1σ). The shift is applied to the MC events as it affects the reconstructed observables associated with the selected electrons, muon and jet objects. The uncertainties are therefore propagated to the measured cross sections by varying the subtracted background and by recomputing the event yield as prescribed by Equation (7.5). Variations are also applied to the MC inputs for the unfolding procedure (migration matrix, fraction of unmatched events, reconstruction efficiency, etc.).

Experimental uncertainties common to both electron and muon Z-boson decay channels are treated as being correlated over kinematic regions, distributions of observables, decay channels and between the signal and background processes. The experimental uncertainties specific to one leptonic decay channel are independent so they are not correlated between channels.

9.1.1 Instrumental uncertainties

Instrumental uncertainties specific to each leptonic final state ($\mathbf{Z \rightarrow e^+e^-}$ and $\boldsymbol{Z \rightarrow \mu^+\mu^-}$): Systematic uncertainties on the lepton-candidate selection are related to the reconstruction, identification, isolation, and trigger [149, 150]. Uncertainties on the lepton calibrations can cause changes in the acceptance, owing to the migration of events across the p_{T} threshold and the $m_{\ell\ell}$ boundaries. The uncertainties on the electron energy scale and resolution are taken into account [166], as are those related to the muon momentum scale, inner-detector and muon-spectrometer resolution, and sagitta-bias correction [150].

Instrumental uncertainties common to the electron and muon final states: Systematic uncertainties associated with jet reconstruction are addressed via jet-energy scale (JES) variations in a 29-nuisance-parameter scheme and jet-energy resolution (JER) variations in a 13-nuisance-parameter scheme [167, 168]. The JES and JER systematic uncertainties originally form over 130 nuisance parameters, but are reduced down to 42 by combining several smaller components. This reduction scheme maintains the correlation information between sources of error, while reducing the computational cost of estimating the systematic uncertainties. To account for differences in detector response to quark and gluon initiated jets, jet flavor composition and response uncertainties are also considered. Systematic uncertainties related to jet reconstruction form the leading source of uncertainty on these cross-section measurements.

Furthermore, imperfect modelling of the effects of pile-up leads to acceptance changes for different jet multiplicities. To assess this uncertainty, the average number of pile-up interactions is varied in simulation. The uncertainty on the combined 2015–2018 integrated luminosity is obtained using the LUCID-2 detector [81]. Hence, each simulated event sample has its luminosity varied by the recommended 1.7% [80] prior to being subtracted from the measured data. Since the backgrounds are small, this has a negligible impact on the total uncertainty. The uncertainty on the integrated luminosity is also carried through the calculation of the cross section as introduced in Chapter 7 resulting in a flat 1.7% uncertainty. The uncertainties on the $t\bar{t}$ background are naturally covered within all these different instrumental uncertainties via the data-driven background estimation method described in Chapter 8.

The instrumental uncertainties quoted in Sections 8.4-8.5 consist of the sum in quadrature of the different instrumental uncertainties described above.

9.1.2 Uncertainties associated with the unfolding procedure

Systematic uncertainties account for possible residual biases in the unfolding procedure, such as those due to the modelling of the signal events or the finite bin width used in each distribution. The limited size of a simulation sample can also create biases in the distribution used in the unfolding procedure. The following uncertainties from the unfolding procedure are considered:

1. The statistical uncertainties of the MC inputs to the unfolding procedure. They are propagated to the unfolded cross sections with a 'toy' simulation method based on 1000 ensembles (pseudo-experiments) of unfolding inputs. The Sherpa 2.2.11 MC samples contain several billion events and provide sufficient statistical precision in the Z+high-p_{T} jets fiducial volume at the particle level. As a result, the unfolding uncertainty due to MC statistical effects is generally below 1%.

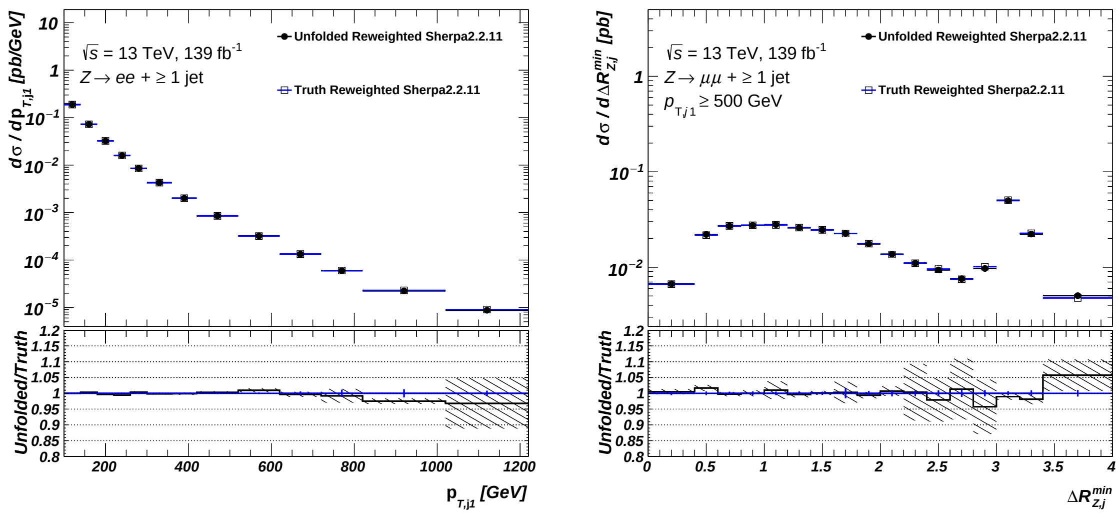
2. The effects of the mismodelling of the data by the MC simulation on the results of the unfolding procedure. The approach consists in estimating the bias introduced by the MC simulation on the unfolded data. This uncertainty is derived by reweighting the Sherpa 2.2.11 Z+jets MC simulation at particle level for each unfolded observable, such that the MC simulation distribution matches the background-subtracted data at the reconstruction level. Hence, the reweighted MC describes data at the detector level and also carries particle-level information. The reweighted MC simulation is unfolded with the non-reweighted response matrix and the uncertainty is obtained by comparing the unfolded result against the reweighted distribution at particle level, a so-called closure test. This bias of the unfolding can be reduced by modifying the number of

iterations in the procedure, as discussed in Section 7.3. Each iteration in the unfolding procedure adjusts the response matrix according to prior knowledge, correcting for incorrect modelling of data. Two unfolding iterations are chosen to minimize this systematic uncertainty. Furthermore, the bias is small due to the good performance of SHERPA 2.2.11 in modelling the data in the fiducial phase space of Z+high-p_{T} jets events. Since SHERPA 2.2.11 models the data well, the reweighted sample is very similar to the nominal one so that the expected bias is small. This systematic uncertainty ranges within 0%–2%, with typical values smaller than 1%. Figure 9.1 shows the results of this closure test for $p_{\mathrm{T},j1}$ in the *inclusive* region for the electron channel and $\Delta R_{Z,j}^{\mathrm{min}}$ in the *high-*p_{T} region for the muon channel. The distributions show that the unfolding procedure introduces a minimal bias and results in a very small systematic uncertainty. In comparison, the same study performed with SHERPA 2.2.1 Z+jets samples yields uncertainties in the range of 0%–10% with an average of approximately 5% in the *high-*p_{T} region. This analysis indeed clearly shows that SHERPA 2.2.11 is a much better generator at describing Z+jets events and it is supersedes SHERPA 2.2.1.

3. An additional uncertainty is derived to account for more subtle differences between the SHERPA 2.2.11 and MG5_aMC+PY8 CKKWL generators (*e.g.* hadronization models, additional soft objects, distributions in other kinematic dimensions) that are not accounted for by the previous method. This uncertainty covers the possible bias introduced with the choice of a given MC prediction to model the response matrix in the unfolding. This additional uncertainty is not standard in ATLAS analyses and is motivated by the concern that the unfolding bias uncertainty introduced above is underestimated. An

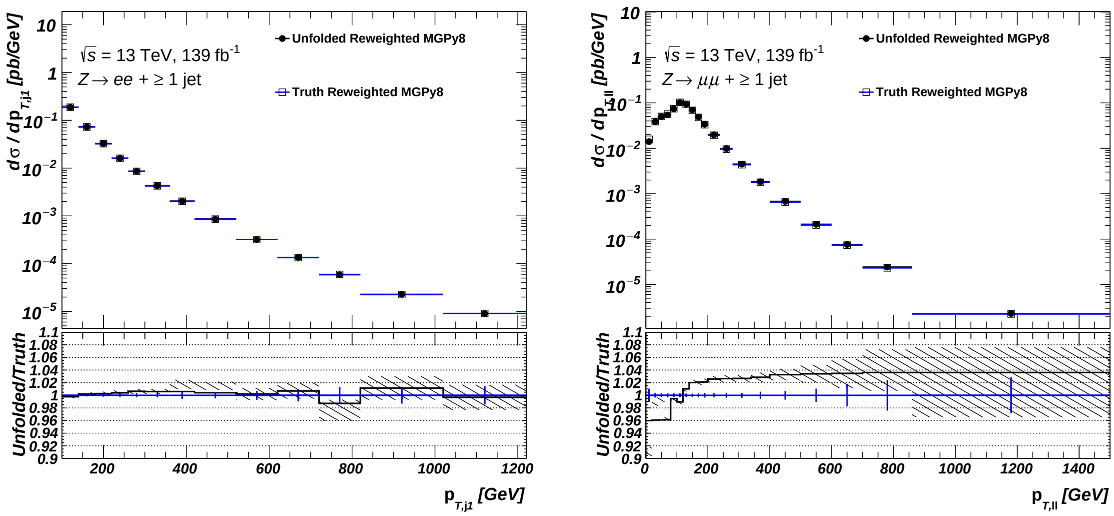
additional closure test is performed where the MG5_aMC+Py8 CKKWL samples are first reweighed to match Sherpa 2.2.11 particle-level distributions for each observable in turn and subsequently unfolded with Sherpa 2.2.11. The uncertainty is evaluated by comparing the unfolded result and the reweighted distribution at particle level. The MG5_aMC+Py8 CKKWL generator is chosen over Sherpa 2.2.1 and MG5_aMC+Py8 FxFx specifically because it corresponds to a LO prediction that is very different from Sherpa 2.2.11 and poorly describes data. The choice of this generator provides a very conservative systematic uncertainty that covers any potentially missed effects from the previous method. Nonetheless, the resulting uncertainty is generally small and does not have a large impact in the final cross-section measurements. This systematic uncertainty is evaluated to be within 0%–4%. Typically, this uncertainty is below 1%, except for the $p_{\mathrm{T},\ell\ell}$ and H_{T} observables which do reach 4%. Figure 9.2 shows the results of this closure test in the *inclusive* region for $p_{\mathrm{T},j1}$ for the electron channel and $p_{\mathrm{T},\ell\ell}$ for the muon channel. In most cases, the results of the closure test are covered by statistical variations. However, relative differences of up to 4% are observed for the $p_{\mathrm{T},\ell\ell}$ observable, which are not covered by statistical variations and may point to a real bias introduced in the unfolding procedure.

The total unfolding uncertainty corresponds to the sum in quadrature of the above three sources. The total unfolding uncertainty is typically below 2%. For $p_{\mathrm{T},\ell\ell}$ and H_{T}, the unfolding uncertainty stemming from the additional closure test raises the unfolding uncertainty above 3% in the first bins of the distributions. The $\Delta R^{\mathrm{min}}_{Z,j}$ and $r_{Z,j}$ observables are also affected by low MC statistical precision in some bins and can increase the total unfolding uncertainty to up to 5%.



(a) (b)

Figure 9.1: Closure test for the unfolding uncertainty stemming from a possible bias introduced in the unfolding procedure. The distributions are shown as a function (a) of the transverse momentum of the leading jet $p_{\mathrm{T},j1}$ in the *inclusive* region for the electron channel and (b) of angular distance between the Z boson and the closest jet $\Delta R_{Z,j}^{\mathrm{min}}$ in the *high-*p_{T} region for the muon channel. The relative difference between the unfolded and particle-level (labelled 'Truth') distributions of the reweighted SHERPA 2.2.11 samples is taken as a source of unfolding uncertainty. The MC statistical uncertainty of the unfolded distribution is shown as a hatched band, while the statistical uncertainty on the particle-level distribution is shown as error bars.



(a) (b)

Figure 9.2: Additional closure test for the unfolding uncertainty stemming from a possible bias introduced in the unfolding procedure. The distributions in the *inclusive* region are shown as a function (a) of the transverse momentum of the leading jet $p_{\mathrm{T},j1}$ for the electron channel and (b) of the transverse momentum of the Z boson $p_{\mathrm{T},\ell\ell}$ for the muon channel. The relative difference between the unfolded and particle-level (labelled 'Truth') distributions of the reweighted MG5_aMC+PY8 CKKWL samples is taken as a source of unfolding uncertainty. The MC statistical uncertainty of the unfolded distribution is shown as a hatched band, while the statistical uncertainty on the particle-level distribution is shown as error bars.

9.1.3 Signal and background modelling uncertainties

Distribution-shape changes stemming from the prescribed variations in the PDF, QCD scales and EW uncertainties on the SHERPA 2.2.11 Z+jets simulation, as described in Section 9.2, are propagated to the fiducial cross sections via changes in the detector response. The response matrix is adjusted to reflect the varied migration matrix and associated unmatched-events and efficiency corrections. Although the associated uncertainties on the predictions can be large, their effect on the cross-section measurement is minimized by the iterative unfolding technique. The systematic uncertainties on the modelling of background MC samples are propagated to the fiducial cross section via the background subtraction in the signal regions. The background-modelling uncertainty comes mainly from the diboson and EW Zjj backgrounds. The $t\bar{t}$ background is insensitive to MC background modelling uncertainties due to its estimation with a data-driven approach, as described in Chapter 8.

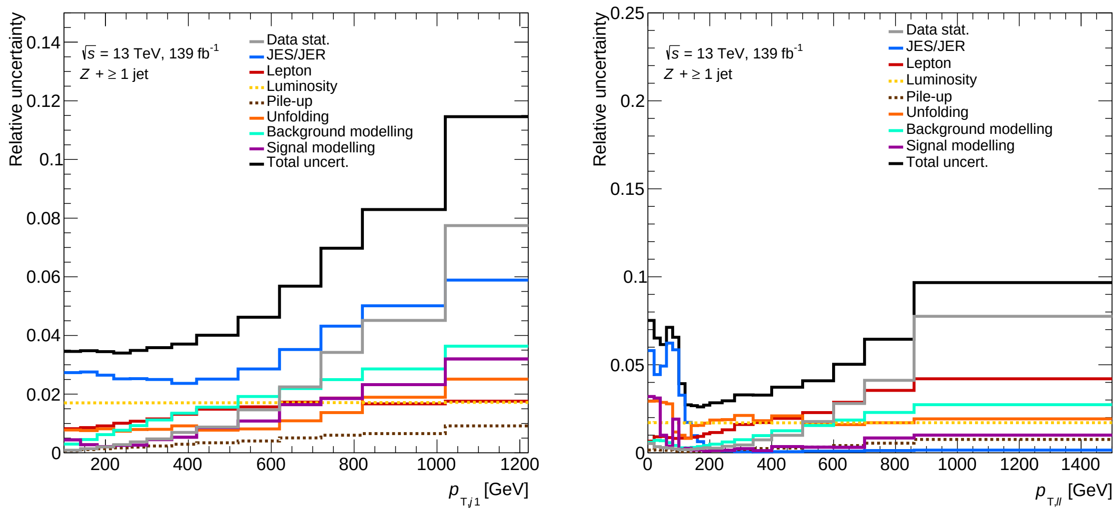
9.1.4 Total systematic uncertainty

The total uncertainties on the unfolded differential cross sections for the combined Z+high-p_{T} jets production measurement are summarized here and shall be used in the result section of the next chapter. The total systematic uncertainty is defined as the sum in quadrature of all the experimental systematic uncertainties enumerated in the previous sections (9.1.1, 9.1.2 and 9.1.3). This total systematic uncertainty is quoted in Table 9.1. Moreover, the total uncertainty is defined as the sum in quadrature of the data statistical uncertainty with the total systematic uncertainty. Table 9.1 shows the breakdown of the total relative statistical and experimental systematic uncertainties on the measured integrated fiducial cross sections for Z+high-p_{T} jets production in the five signal regions. In the *inclusive* and *high-*S_{T} regions, jet reconstruction uncertainties dominate with a relative contribution of 2.6% and 2.8%, respectively. In the *high-*p_{T} region and its *collinear* and *back-to-back* subregions, the jet reconstruction and data statistical uncertainties are largest, with relative contributions of 3.2% and 2.1% in the *high-*p_{T} region, respectively, 2.8% and 2.9% in the *collinear* region, and 3.6% and 2.7% in the *back-to-back* region. The background modelling uncertainties are driven by the diboson and EW Zjj backgrounds.

Table 9.1: Relative statistical and systematic uncertainties (in %) on the measured integrated cross sections for Z+high-p_{T} jets production in the five signal regions, computed by integrating over the respective jet-multiplicity distribution. The total systematic uncertainty is defined as the sum in quadrature of all experimental systematic uncertainties. The total uncertainty is defined as the sum in quadrature of the data statistical uncertainty with the total systematic uncertainty.

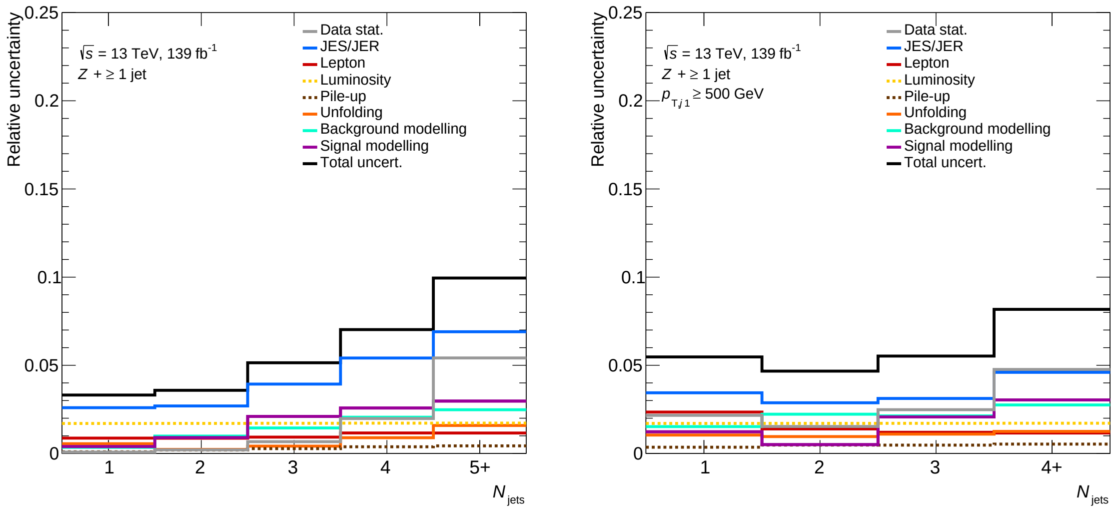
Uncertainty source [%]	*Inclusive*	*High-*p_{T}	*Collinear*	*Back-to-back*	*High-*S_{T}
JES/JER	2.6	3.2	2.8	3.6	2.8
Lepton	0.9	1.6	1.4	2.0	1.1
Luminosity	1.7	1.7	1.7	1.7	1.7
Pile-up	0.1	0.4	0.4	0.4	0.4
Unfolding	0.5	1.0	1.1	1.4	0.8
Background modelling	0.5	2.0	2.0	1.9	1.7
Signal modelling	0.5	1.2	1.1	1.1	1.1
Total syst. uncertainty	3.4	4.8	4.4	5.3	4.2
Data stat. uncertainty	0.1	2.1	2.9	2.7	1.2
Total uncertainty	3.4	5.3	5.3	5.9	4.4

In Figures 9.3–9.9, the total fractional uncertainties are shown for the thirteen observables measured in the Z+high-p_{T} jets analysis. Each source of uncertainty is shown individually and them summed in quadrature to create the total uncertainty. The leading source of experimental uncertainty on the full range of observables are the jet reconstruction (JES/JER) systematic uncertainties, confirming the results shown in Table 9.1. For the uncertainties on $p_{\mathrm{T},\ell\ell}$, the jet systematic uncertainties unexpectedly dominate for $p_{\mathrm{T},\ell\ell} \leq 100$ GeV and is caused by the migration of the leading-jet p_{T} above and below the $p_{\mathrm{T}}(\text{jet}) \geq 100$ GeV selection requirement. The data statistical uncertainty has a leading impact in the *high-p_{T}*, *collinear*, *back-to-back* and *high-S_{T}* regions, sometimes approaching an uncertainty of 10% for some bins. This behaviour indicates that even with the full LHC Run–2 dataset, the Z+high-p_{T} jets measurements are statistically limited. The impact of the unfolding uncertainty is well behaved and is generally within 2%.



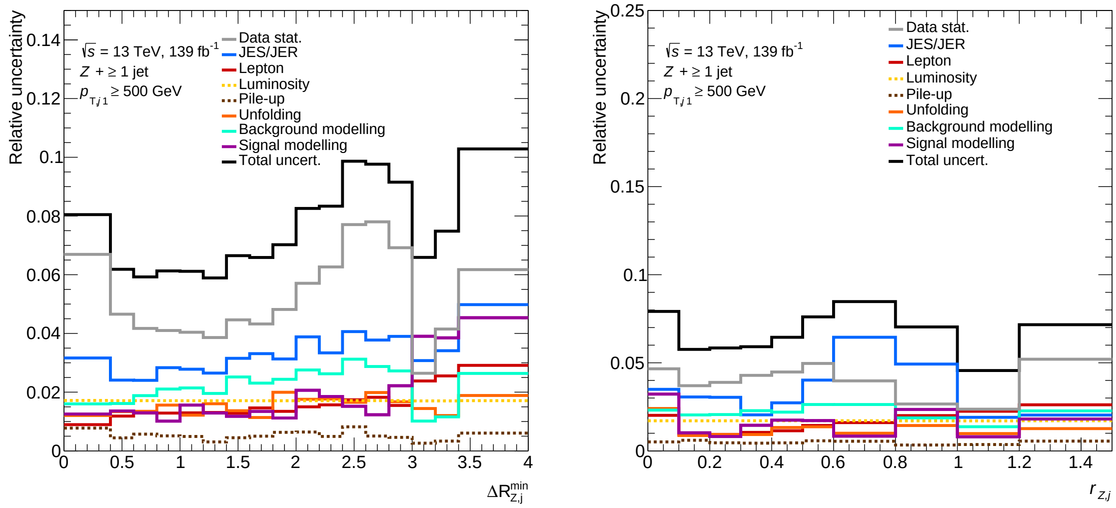
(a) (b)

Figure 9.3: Fractional uncertainties on the measured differential fiducial cross section in the *inclusive* region as a function (a) of the transverse momentum of the leading jet $p_{\mathrm{T},j1}$ and (b) of the transverse momentum of the Z boson $p_{\mathrm{T},\ell\ell}$.



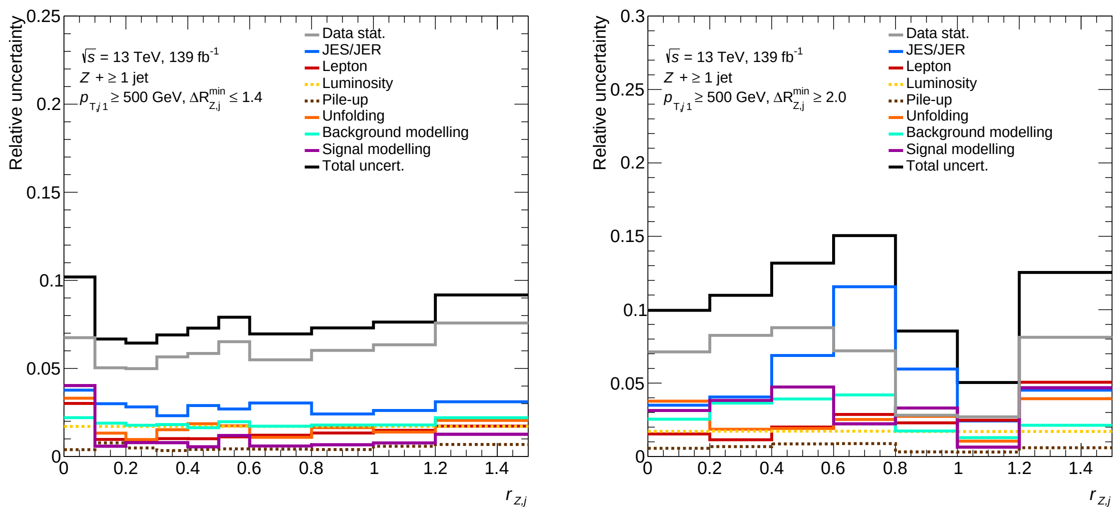
(a) (b)

Figure 9.4: Fractional uncertainties on the measured differential fiducial cross section as a function of the jet multiplicity N_{jets} (a) in the *inclusive* region and (b) in the *high*-p_{T} region. In each case, the last bin is inclusive in higher jet multiplicities.



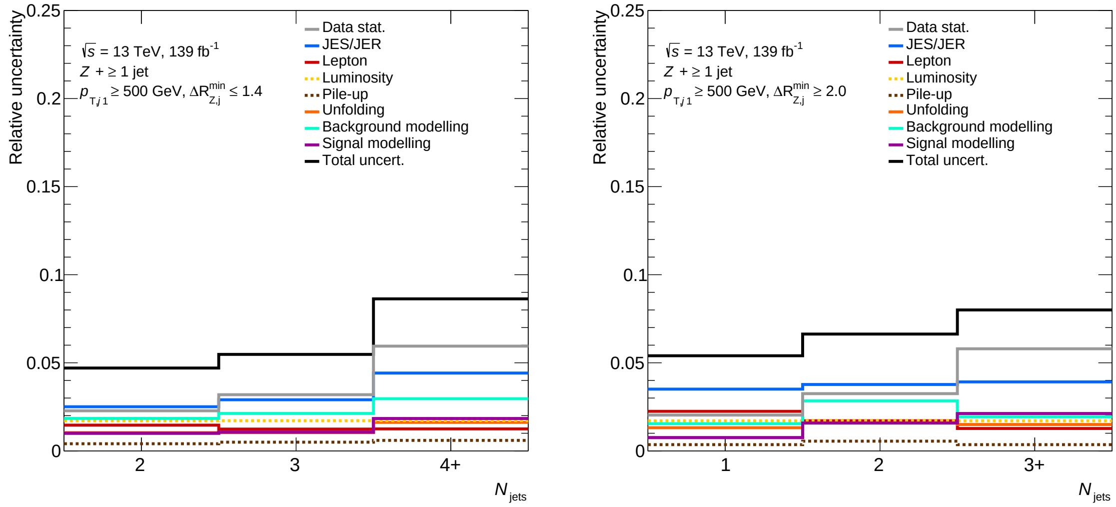
(a) (b)

Figure 9.5: Fractional uncertainties on the measured differential fiducial cross section in the *high*-p_{T} region as a function (a) of angular distance between the Z boson and the closest jet $\Delta R^{\mathrm{min}}_{Z,j}$ and (b) of the ratio of Z-boson to closest-jet transverse momenta $r_{Z,j}$.



(a) (b)

Figure 9.6: Fractional uncertainties on the measured differential fiducial cross section as a function of the ratio of Z-boson to closest-jet transverse momenta $r_{Z,j}$ (a) in the *collinear* $\Delta R^{\mathrm{min}}_{Z,j} \leq 1.4$ region and (b) in the *back-to-back* $\Delta R^{\mathrm{min}}_{Z,j} \geq 2.0$ region.



(a) (b)

Figure 9.7: Fractional uncertainties on the measured differential fiducial cross section as a function of the jet multiplicity N_{jets} (a) in the *collinear* $\Delta R^{\text{min}}_{Z,j} \leq 1.4$ region and (b) in the *back-to-back* $\Delta R^{\text{min}}_{Z,j} \geq 2.0$ region. In each case, the last bin is inclusive in higher jet multiplicities.

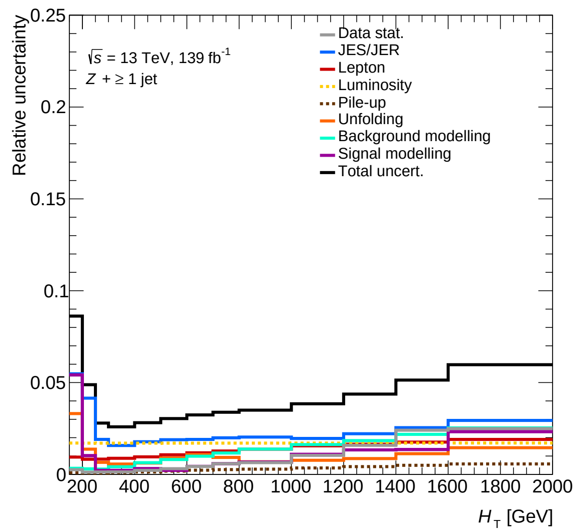
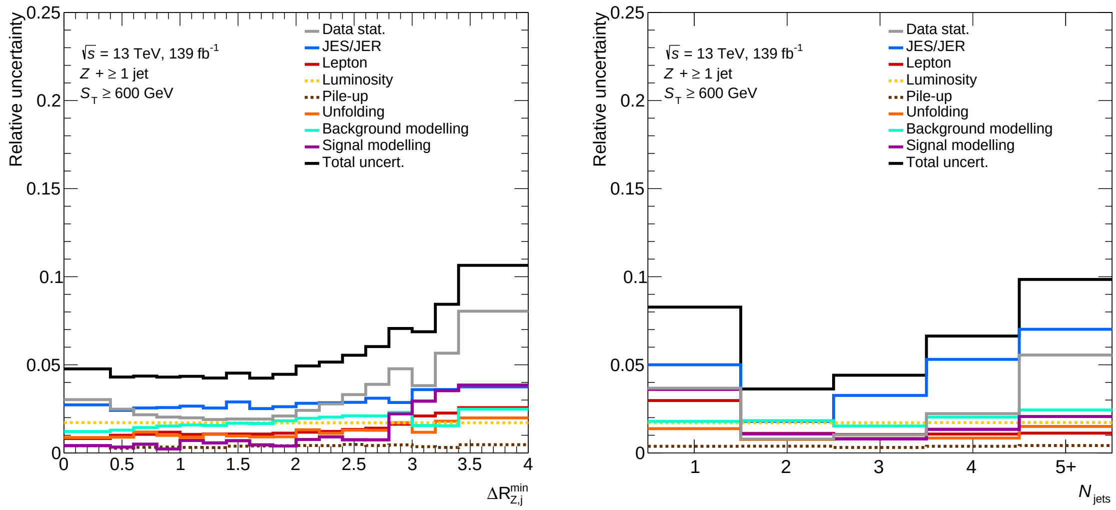


Figure 9.8: Fractional uncertainty on the measured differential fiducial cross section as a function of the transverse momenta scalar sum H_{T} in the *inclusive* region.



(a) (b)

Figure 9.9: Fractional uncertainties on the measured differential fiducial cross section in the *high*-S_T region as a function (a) of angular distance between the Z boson and the closest jet $\Delta R^{\mathrm{min}}_{Z,j}$ and (b) of the jet multiplicity N_{jets}.

9.2 Theoretical uncertainties

Uncertainties in the modelling of MC generator and fixed-order predictions are relevant not only when unfolding the data, but also in the comparisons with the fiducial cross-section measurements. It is of upmost importance for the investigation of new physics at the energy frontier with LHC data to fine-tune the modelling of key interactions such as $pp \to Z$+jets + X processes. This analysis serves as a benchmark for improving MC predictions in the so far never measured Z+high-p_T jets phase space. Modelling uncertainties affect either the event kinematics or the event weight, resulting in different predicted observable distributions for both signal and backgrounds. Systematic uncertainties on the estimated backgrounds manifest themselves by changing the MC predicted event yields that are subtracted from the data prior to unfolding. Uncertainties on the modelling of the Sherpa 2.2.11 Z+jets signal samples have an impact on the response matrix in the unfolding procedure and lead to a source of uncertainty on the resulting fiducial cross sections. The different Z+jets MC and fixed-order predictions also carry systematic uncertainties that need to be considered when reporting conclusions on their individual performances and ultimately their agreement with data.

Modelling uncertainties are taken into account by varying the PDFs, the QCD scales and, in the case of Sherpa 2.2.11, the virtual EW corrections. Moreover, the EW Zjj prediction also takes into account interference effects as introduced in Section 8.3.1.

9.2.1 Parton distribution function

The proton parton distribution function (PDF) uncertainties stem from three possible effects: 1) the uncertainty around the central value of the PDF, 2) the choice of the PDF itself, and 3) the effect of varying the value of α_s.

The first uncertainty source arises from the variations around the central predicted distribution of the PDF [169]. For this analysis, each PDF is modelled from 100 different uncorrelated parameters that can be varied by one standard deviation to create a further 100 PDF sets. The associated uncertainty is evaluated following different recipes according to the PDF of the MC program. For most of the predictions, the uncertainty is calculated from the standard deviation of the set of 100 PDF distributions

$$\delta_{\sigma}^{\mathrm{PDF}} = \sqrt{\frac{1}{100}\sum_{k=1}^{100}(\sigma^{(k)} - \sigma^{(0)})^2}, \tag{9.1}$$

where δ_σ is the uncertainty, the sum is over the 100 sets, $\sigma^{(k)}$ is the prediction of the k-th set, and $\sigma^{(0)}$ is the nominal prediction. The Sherpa 2.2.11 prediction instead follows a slightly different procedure to evaluate the uncertainty around the 100 sets, as determined by its PDF provider. Due to differences in what the 100 PDF sets represent, the Sherpa 2.2.11 PDF uncertainty is instead computed as

$$\delta_\sigma^{\mathrm{PDF}} = \sqrt{\sum_{k=1}^{100} (\sigma^{(k)} - \sigma^{(0)})^2}. \quad (9.2)$$

The second type of PDF uncertainty compares the distributions obtained with the nominal predictions to those obtained with the MMHT2014NNLO [126] and the CT14NNLO [170] PDFs. It covers the possible bias introduced with the choice of a specific PDF set.

The third type of PDF uncertainty is only included in the Sherpa Z+jets and diboson predictions. This uncertainty compares the results of varying the strong coupling constant α_s in the PDF by one standard deviation [117]. This variation of α_s is also propagated to the hard-scatter calculation of the MEs. Using the current value for the strong coupling constant of $\alpha_s = 0.1180 \pm 0.0015$, the uncertainty is computed by comparing the distributions so that

$$\delta_\sigma^{\alpha_s} = \left| \frac{\sigma(\alpha_s = 0.1195) - \sigma(\alpha_s = 0.1165)}{2} \right|, \quad (9.3)$$

where $\sigma(\alpha_s)$ is the distribution obtained by setting α_s in the PDF to the values $\alpha_s = 0.1195$ and $\alpha_s = 0.1165$.

The Z+jets and diboson predictions from Sherpa include the three sources of PDF uncertainties, whereas MG5_aMC+Py8 FxFx only includes the first two. The different considered PDF uncertainties are added in quadrature to produce the total PDF uncertainty. The total of PDF uncertainties generally ranges between 0%–2%. Due to computational limitations in the NNLOjet program, its predictions do not include PDF uncertainties.

9.2.2 QCD scale

The effect of QCD scale uncertainties is evaluated by varying the renormalization (μ_R) and factorization (μ_F) scales introduced in Section 2.1.7. The QCD scales are varied by factors of two to provide a total of five sets of predictions with the following pairs of μ_R and μ_F:

$$\Big(\mu_R, \mu_F\Big), \Big(0.5 \times \mu_R, \mu_F\Big), \Big(2 \times \mu_R, \mu_F\Big), \Big(\mu_R, 0.5 \times \mu_F\Big), \Big(\mu_R, 2 \times \mu_F\Big). \tag{9.4}$$

The QCD scale uncertainty is defined as the envelope of the five predictions around the central value. This QCD scale uncertainty is the dominating systematic uncertainty on the different theoretical predictions of Z+high-p_{T} jets. In the SHERPA 2.2.11 predictions, the uncertainty ranges between 10% and 60% with average values near 25%. Meanwhile, the corresponding range in MG5_aMC+PY8 FXFX is between 5% and 20%. The differences between these two generators and their uncertainties are further explored in Ref. [72]. In the NNLOJET predictions, the QCD scale uncertainties are typically in the range of 5%–10%.

9.2.3 NLO virtual electroweak corrections

The prediction from SHERPA 2.2.11 also considers uncertainties related to NLO virtual EW corrections [72]. The uncertainty is defined as the envelope of the predictions obtained with the additive, multiplicative and exponentiated EW correction schemes introduced in Section 4.5. The resulting uncertainties amount to a maximum of 5% where the EW corrections are largest: in the *back-to-back* region with large $p_{\mathrm{T},\ell\ell}$ and $\Delta R_{Z,j}^{\mathrm{min}} \approx \pi$.

9.2.4 Theoretical uncertainties on background processes

The theoretical uncertainty on the diboson predictions used for the data background subtraction are computed with the PDF and scale uncertainties described above. Moreover, the EW Zjj prediction includes the effects of PDF, QCD scale and interference uncertainties, which amount to normalization uncertainties of 2%, 9% and 25%–40% respectively. The effects of PDF and QCD scale uncertainties on the single-top predictions amount to a total normalization uncertainty of about 4%, primarily from the normalization to theory predictions at NNLO and NLO accuracies.

The predicted $t\bar{t}$ background is insensitive to background theoretical uncertainties due to the data-driven estimation method described in Section 8.1. However, the MC $t\bar{t}$ predictions include the three PDF uncertainties and the QCD scale uncertainties described above. These uncertainties are used only in the detector-level distributions of the $t\bar{t}$ control-region and are not used in the data-driven $t\bar{t}$ background estimation, as described in Chapter 8.

In summary, this chapter provides the description of the total uncertainties on the unfolded differential cross sections for the combined Z+high-p_{T} jets production measurement. In addition, the theoretical uncertainties on the different Z+jets predictions are outlined. These uncertainties are shown in the resulting fiducial cross-section measurements of the next chapter.

Chapter 10

Unfolded results

This chapter provides the final results of this analysis. The results of the Z+high-p_{T} jets production measurements at particle level, referred to as fiducial cross sections, are presented alongside the comparison to several theoretical predictions. Each fiducial volume corresponds to a phase space region: *inclusive*, *high-*p_{T}, *collinear*, *back-to-back* and *high-*S_{T}.

The differential fiducial cross-section measurements are performed separately for the electron and muon decay channels of the Z boson. For each of the two decay channels, the estimated background are subtracted from the detector-level data and the resulting distributions are unfolded using the method described in Chapter 7. The particle-level unfolded distributions correspond to measurements in the fiducial region. By unfolding the lepton channels separately, the respective response matrices describe more accurately the detector response to the given leptonic final state. The background estimation is taken from the MC and data-driven methods provided in Chapter 8. Finally, the error on individual fiducial cross-section measurements follows the prescription detailed in Chapter 9.

The combination of the electron and muon Z-boson decay channels is described in Section 10.1. The resulting integrated fiducial cross sections are presented in Section 10.2. Then, the differential fiducial cross sections of Z+high-p_{T} jets production are shown and discussed in Section 10.3. The conclusion follows in Chapter 11.

10.1 Combination of Z-boson decay channels

The compatibility between the fiducial cross-section results of both Z-boson decay channels is first tested. According to lepton universality, the fiducial cross-section distributions for the electron and muon decay channels should be compatible within uncertainties. The resulting cross-section measurements are indeed found to be compatible between channels. As a result, the unfolded results are combined to measure the production cross section for a Z boson decaying into a single charged-lepton flavor, $Z(\rightarrow \ell^{+}\ell^{-})$, in order to increase the precision of the final Z+high-p_{T} jets measurements. The combination is done with the Best Linear Unbiased Estimate (BLUE) method [171]. The BLUE method provides an unbiased distribution estimate given two measurements, including their uncertainties and correlations. For this analysis, where the uncertainties are treated as being Gaussian, the BLUE method is equivalent to performing a maximum likelihood fit. In the combination, all systematic uncertainties except the lepton-related instrumental uncertainties are treated as correlated between the electron and muon channels. The data and MC statistical uncertainties are treated as uncorrelated between channels.

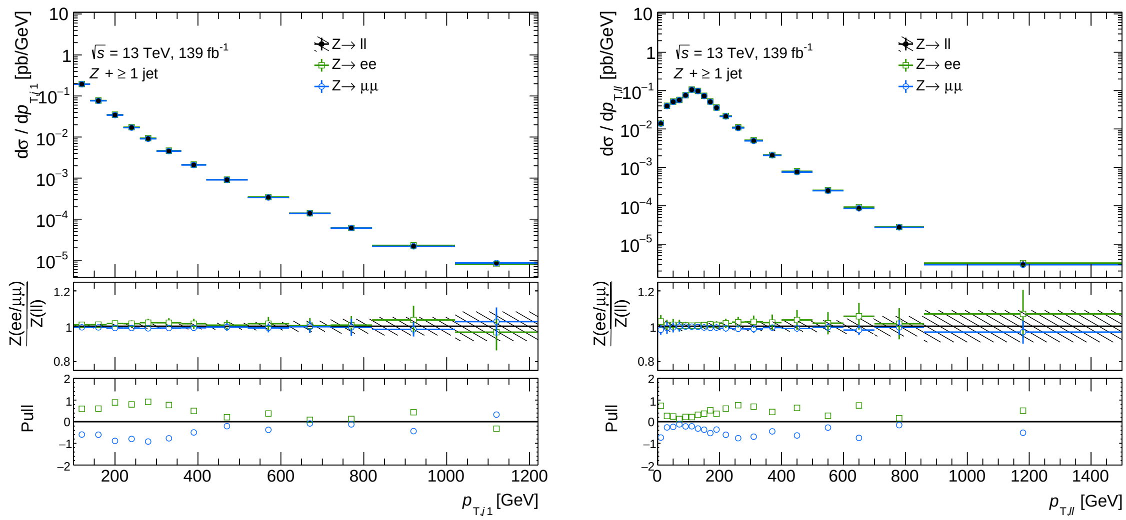
The resulting combined measurements are consistent with both individual decay channels for the full set of observables. Differential fiducial cross-section distributions of the combined and individual lepton channel measurements are compared for some selected observables in Figures 10.1–10.3. In these figures, the hatched band represents the sum in quadrature of statistical and experimental systematic uncertainties on the combined result that are uncorrelated between decay channels. As such, the uncertainty bands are not artificially made larger by correlated uncertainties, *e.g.* by the jet reconstruction (JES/JER) systematic uncertainties, and the compatibility can be more carefully studied. The coloured error bars represent the sum in quadrature of the statistical and uncorrelated systematic uncertainties on a single decay-channel. The pull p in a given bin i is defined as

$$p_i = \frac{x_i - \langle x \rangle_i}{\sigma(x_i)}, \tag{10.1}$$

where x is the measurement in the individual decay channel and $\langle x \rangle$ is the combined measurement. The uncertainty $\sigma(x)$ is the sum in quadrature of the statistical and uncorrelated systematic uncertainties of the individual decay channel measurement. The pulls are well behaved and usually within unity, indicating that the combination is compatible to both decay channel measurements. This good agreement serves as a validation of the cross-section measurement strategy of this analysis. As a result, the final fiducial cross-section measurements of Z+high-p_{T} jets production of the combined electron and muon channels are presented in the following sections.

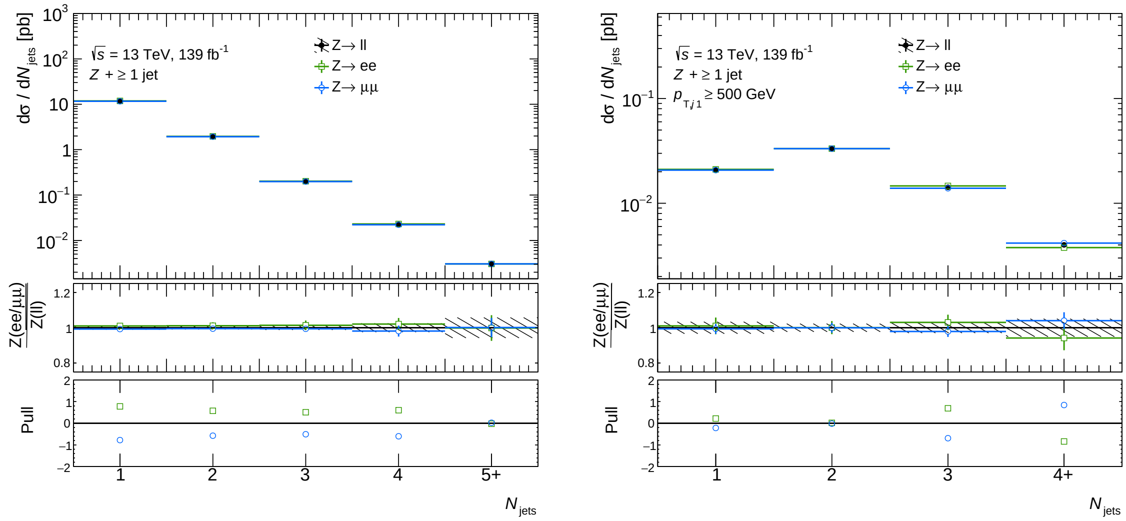
10.2 Integrated fiducial cross sections

Integrated fiducial cross sections for Z+high-p_{T} jets production are evaluated in the *inclusive*, *high-p_{T}*, *collinear*, *back-to-back* and *high-S_{T}* signal regions by summing over the respective unfolded jet-multiplicity distributions. The fiducial cross sections are compared with the predictions from Sherpa 2.2.1, Sherpa 2.2.11, MG5_aMC+Py8 FxFx, and with NNLO and NLO predictions from NNLOjet in Tables 10.1 and 10.2. On the measurements, the statistical and systematic uncertainties as described in



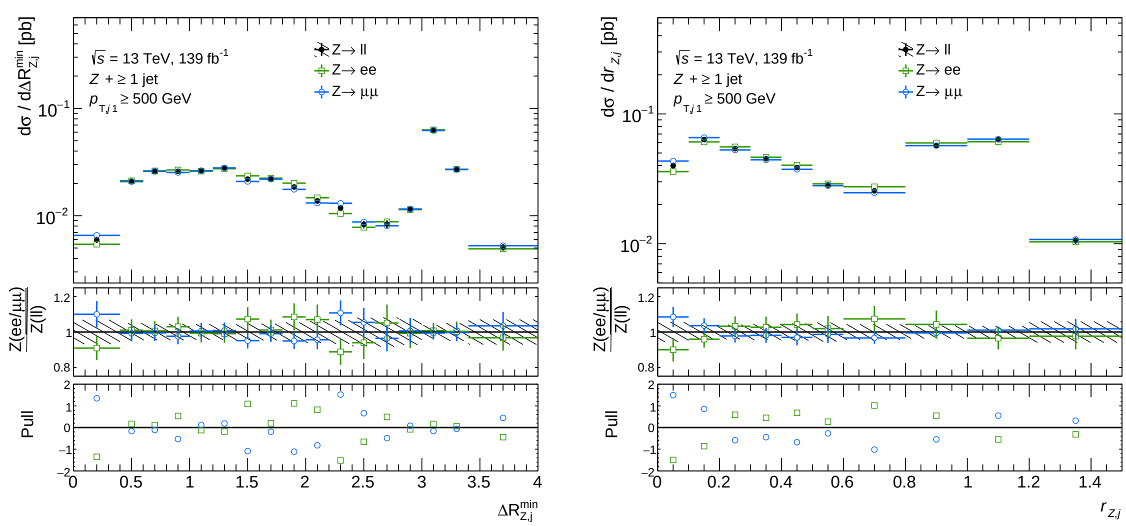
(a) (b)

Figure 10.1: Comparison of the differential fiducial cross section in the *inclusive* region for the two Z-boson decay channels and their combination as a function (a) of the transverse momentum of the leading jet $p_{\mathrm{T},j1}$ and (b) of the transverse momentum of the Z boson $p_{\mathrm{T},\ell\ell}$. The hatched band represents the sum in quadrature of statistical and experimental systematic uncertainties on the combined result which are uncorrelated between decay channels. The coloured error bars represent the sum in quadrature of the statistical and uncorrelated systematic uncertainties on a single decay-channel. The pull is defined as the difference between the individual and the combined measurements over the sum in quadrature of the statistical and uncorrelated systematic uncertainties of the individual decay channel measurement.



(a) (b)

Figure 10.2: Comparison of the differential fiducial cross section for the two Z-boson decay channels and their combination as a function of the jet multiplicity N_{jets} (a) in the *inclusive* region and (b) in the *high-*p_{T} region. In each case, the last bin is inclusive in higher jet multiplicities. Further details on the legend, ratio panel and pull are provided in the caption of Figure 10.1.



(a) (b)

Figure 10.3: Comparison of the differential fiducial cross section in the *high-*p_{T} region for the two Z-boson decay channels and their combination as a function (a) of angular distance between the Z boson and the closest jet $\Delta R^{\mathrm{min}}_{Z,j}$ and (b) of the ratio of Z-boson to closest-jet transverse momenta $r_{Z,j}$. Further details on the legend, ratio panel and pull are provided in the caption of Figure 10.1.

Section 9.1.4 are shown separately. The modelling uncertainties of the different predictions include the uncertainties from the variations of PDF, QCD scales and, in the case of Sherpa 2.2.11 only, virtual EW corrections, as described in Sections 9.2.1, 9.2.2, and 9.2.3, respectively. The cross sections predicted by the three MC generators and the NNLOJet predictions agree with the measured values within the theory uncertainties. The uncertainties on the integrated fiducial cross-section measurements are clearly dominated by the systematic uncertainties. They are smaller than the uncertainties on the predictions, except for NNLOJet@NNLO, which matches or exceeds the precision of the measurements in some kinematic regions. These results are also depicted in Figure 10.4, where the data total uncertainty corresponds to the sum in quadrature of statistical and systematic uncertainties, as described in Section 9.1.4. The uncertainties on the predictions are found by summing in quadrature the uncertainties from the variations of the PDF (excluding NNLOJet predictions), QCD scales and, for Sherpa 2.2.11, virtual EW contributions, as explained in Sections 9.2.1, 9.2.2, and 9.2.3, respectively. When the virtual EW corrections are removed from the Sherpa 2.2.11 prediction, the integrated cross sections for the *inclusive*, *high-*p_{T}, *collinear*, *back-to-back* and *high-*S_{T} regions increase by 0.065%, 6.9%, 3.8%, 11% and 3.0%, respectively. However, the theoretical predictions from Sherpa 2.2.11 are not precise enough to meaningfully interpret the importance of NLO virtual EW corrections. Because of this, the importance of NLO virtual EW corrections on theoretical predictions are instead studied with the very precise NNLOJet@NNLO predictions in Appendix A.

The integrated fiducial cross-sections measurements provide limited information on the performance of the different theoretical predictions. Although the predicted values are consistent with the measured data, the associated modelling uncertainties are in some cases too large to meaningfully interpret the ability of the different theoretical predictions to describe Z+high-p_{T} jets production. For example, the predictions from SHERPA are very imprecise with uncertainties reaching upwards of 50%. Instead, the differential fiducial cross-section distributions provide much more important information about the individual performance of different theoretical predictions in modelling Z+high-p_{T} jets events.

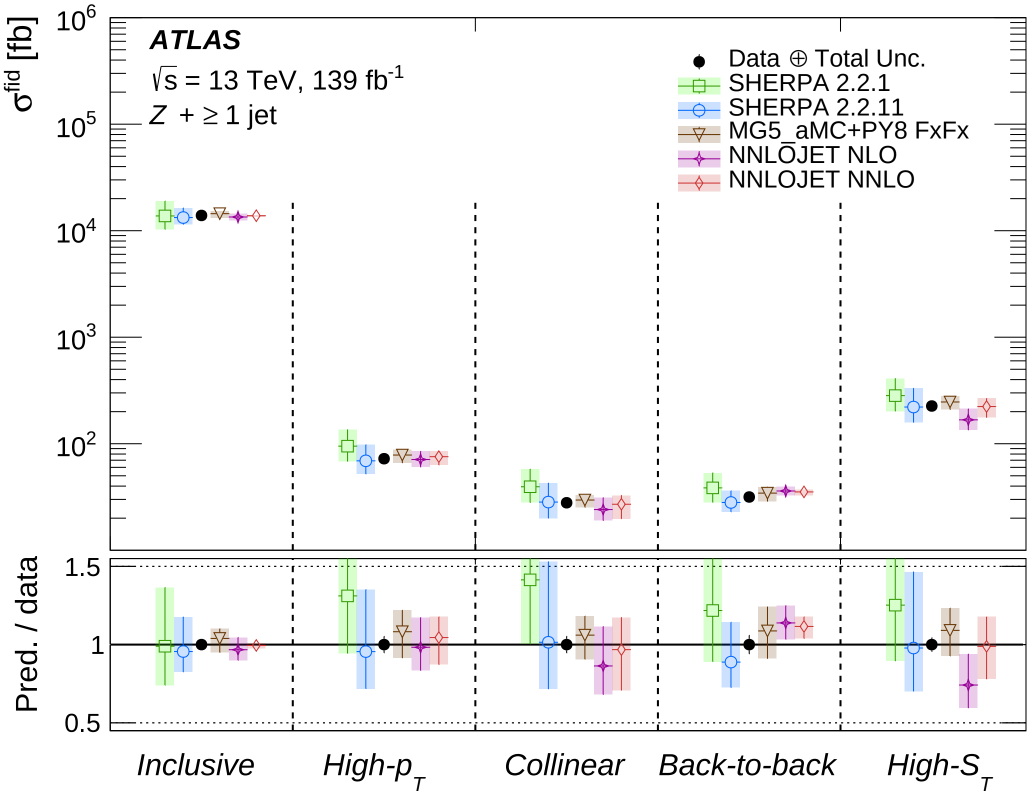


Figure 10.4: Summary of integrated fiducial cross-section results for Z+high-p_{T} jets production [19]. The unfolded fiducial cross sections are shown with black data points and the error bars represent the total uncertainty. The total uncertainty corresponds to the sum in quadrature of the data statistical and total systematic uncertainties, as defined in Section 9.1.4. Data are compared with predictions from MC generators and fixed-order calculations. The uncertainties on the predictions are found by summing in quadrature the uncertainties from the variations of the PDF (excluding NNLOJET predictions), QCD scales and, for SHERPA 2.2.11, virtual EW contributions, as explained in Sections 9.2.1, 9.2.2, and 9.2.3, respectively.

Table 10.1: Measured integrated fiducial cross sections for Z+high-p_{T} jets production in the *inclusive*, *high-p_{T}*, *collinear* and *back-to-back* signal regions and predictions from Sherpa 2.2.1, Sherpa 2.2.11, MG5_aMC+Py8 FxFx, NNLOjet@NNLO and NNLOjet@NLO. The statistical and systematic uncertainties as described in Section 9.1.4 are shown separately for the data. The modelling uncertainties on the predictions include the uncertainties from the variations of the PDF, QCD scales, and, in the case of Sherpa 2.2.11 only, virtual EW corrections, as described in Sections 9.2.1, 9.2.2, and 9.2.3, respectively.

Inclusive Z+jets					
Data	13.90	± 0.01 (stat)	± 0.47 (syst)		pb
Sherpa 2.2.1	13.8	$^{+0.5}_{-0.5}$ (PDF)	$^{+5.2}_{-3.4}$ (Scale)		pb
Sherpa 2.2.11	13.3	$^{+0.2}_{-0.2}$ (PDF)	$^{+3.1}_{-1.8}$ (Scale)	$\pm \leq 0.1$ (EW)	pb
MG5_aMC+Py8 FxFx	14.5	$^{+0.1}_{-0.1}$ (PDF)	$^{+0.8}_{-1.2}$ (Scale)		pb
NNLOjet@NNLO	13.83		$^{+0.18}_{-0.27}$ (Scale)		pb
NNLOjet@NLO	13.5		$^{+1.1}_{-0.9}$ (Scale)		pb
High-p_{T}: $p_{\mathrm{T},j1} \geq 500$ GeV					
Data	72.3	± 1.5 (stat)	± 3.5 (syst)		fb
Sherpa 2.2.1	95	$^{+4}_{-3}$ (PDF)	$^{+40}_{-26}$ (Scale)		fb
Sherpa 2.2.11	69	$^{+2}_{-1}$ (PDF)	$^{+28}_{-17}$ (Scale)	$^{+2}_{-2}$ (EW)	fb
MG5_aMC+Py8 FxFx	78	$^{+4}_{-1}$ (PDF)	$^{+9}_{-12}$ (Scale)		fb
NNLOjet@NNLO	76		$^{+10}_{-12}$ (Scale)		fb
NNLOjet@NLO	71		$^{+14}_{-11}$ (Scale)		fb
Collinear: *High-p_{T}* and $\Delta R^{\mathrm{min}}_{Z,j} \leq 1.4$					
Data	27.9	± 0.8 (stat)	± 1.2 (syst)		fb
Sherpa 2.2.1	39	$^{+2}_{-1}$ (PDF)	$^{+18}_{-11}$ (Scale)		fb
Sherpa 2.2.11	28	$^{+1}_{-1}$ (PDF)	$^{+14}_{-8}$ (Scale)	$\pm \leq 1$ (EW)	fb
MG5_aMC+Py8 FxFx	29.6	$^{+1.3}_{-0.3}$ (PDF)	$^{+3.1}_{-4.3}$ (Scale)		fb
NNLOjet@NNLO	27.0		$^{+5.7}_{-7.2}$ (Scale)		fb
NNLOjet@NLO	24.1		$^{+7.0}_{-5.1}$ (Scale)		fb
Back-to-back: *High-p_{T}* and $\Delta R^{\mathrm{min}}_{Z,j} \geq 2.0$					
Data	31.6	± 0.8 (stat)	± 1.7 (syst)		fb
Sherpa 2.2.1	38	$^{+2}_{-1}$ (PDF)	$^{+15}_{-10}$ (Scale)		fb
Sherpa 2.2.11	28.1	$^{+0.6}_{-0.3}$ (PDF)	$^{+7.9}_{-4.9}$ (Scale)	$^{+1.4}_{-1.4}$ (EW)	fb
MG5_aMC+Py8 FxFx	34.4	$^{+1.6}_{-0.3}$ (PDF)	$^{+4.6}_{-5.6}$ (Scale)		fb
NNLOjet@NNLO	35.3		$^{+1.9}_{-2.4}$ (Scale)		fb
NNLOjet@NLO	36.0		$^{+3.5}_{-3.3}$ (Scale)		fb

Table 10.2: Measured integrated fiducial cross sections for Z+high-p_{T} jets production in the *high*-S_{T} signal region and predictions from SHERPA 2.2.1, SHERPA 2.2.11, MG5_aMC+PY8 FXFX, NNLOJET@NNLO and NNLOJET@NLO. The statistical and systematic uncertainties as described in Section 9.1.4 are shown separately for the data. The modelling uncertainties on the predictions include the uncertainties from the variations of the PDF, QCD scales, and, in the case of SHERPA 2.2.11 only, virtual EW corrections, as described in Sections 9.2.1, 9.2.2, and 9.2.3, respectively.

High-S_{T}: $S_{\mathrm{T}} \geq 600$ GeV					
Data	226.0	$\pm$ 2.6 (stat)	$\pm$ 9.5 (syst)		fb
SHERPA 2.2.1	280	$^{+10}_{-10}$ (PDF)	$^{+130}_{-80}$ (Scale)		fb
SHERPA 2.2.11	220	$^{+10}_{-10}$ (PDF)	$^{+110}_{-60}$ (Scale)	$\pm \leq 10$ (EW)	fb
MG5_aMC+PY8 FXFX	247	$^{+10}_{-2}$ (PDF)	$^{+30}_{-37}$ (Scale)		fb
NNLOJET@NNLO	223		$^{+43}_{-47}$ (Scale)		fb
NNLOJET@NLO	168		$^{+45}_{-33}$ (Scale)		fb

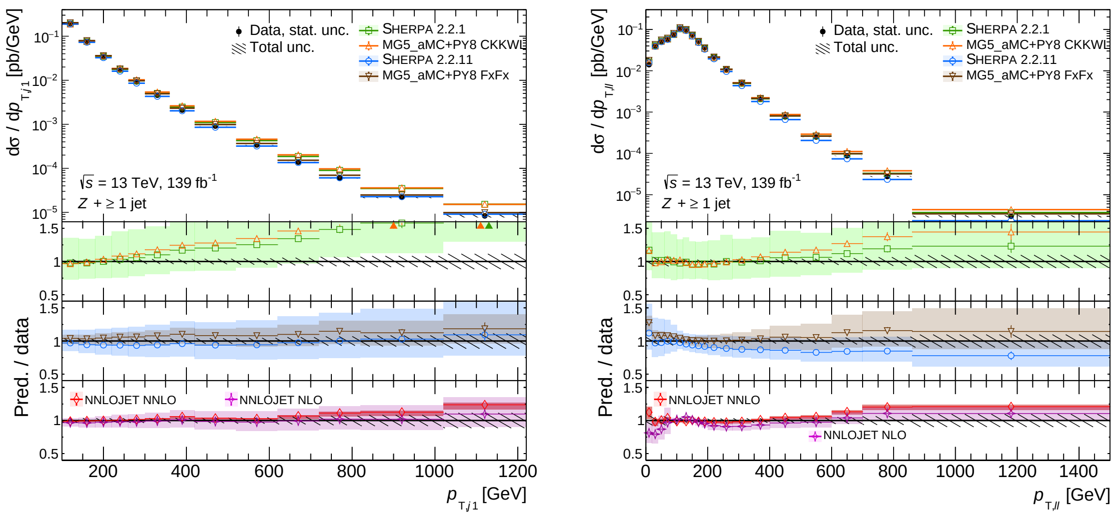
10.3 Differential fiducial cross sections

Differential fiducial cross sections are measured and compared in Figures 10.5–10.11 with predictions from SHERPA 2.2.1, MG5_aMC+PY8 CKKWL, and the next-generation MC generators SHERPA 2.2.11 and MG5_aMC+PY8 FXFX, and with NLO and NNLO Z+jets calculations from NNLOJET. The data statistical uncertainty is displayed with error bars and the hatched bands correspond to the total uncertainty defined as the sum in quadrature of the data statistical and total systematic uncertainties, as defined in Section 9.1.4. In the ratio panels, the MC statistical uncertainty is displayed with coloured error bars and the theoretical uncertainties are shown as coloured bands. The uncertainties on the predictions are found by summing in quadrature the uncertainties from the variations of the PDF (excluding NNLOJET predictions), QCD scales and, for SHERPA 2.2.11, virtual EW contributions, as explained in Sections 9.2.1, 9.2.2, and 9.2.3, respectively. The MG5_aMC+PY8 CKKWL prediction is included as it was often used as a benchmark in previous

ATLAS Run–2 analyses. However, the MG5_aMC+PY8 CKKWL prediction is not shown with uncertainties as they are in the order of $\mathcal{O}(100\%)$ due to the LO order nature of the matrix element calculation. In general, NNLOJET@NNLO and MG5_aMC+PY8 FXFX provide the most precise predictions.

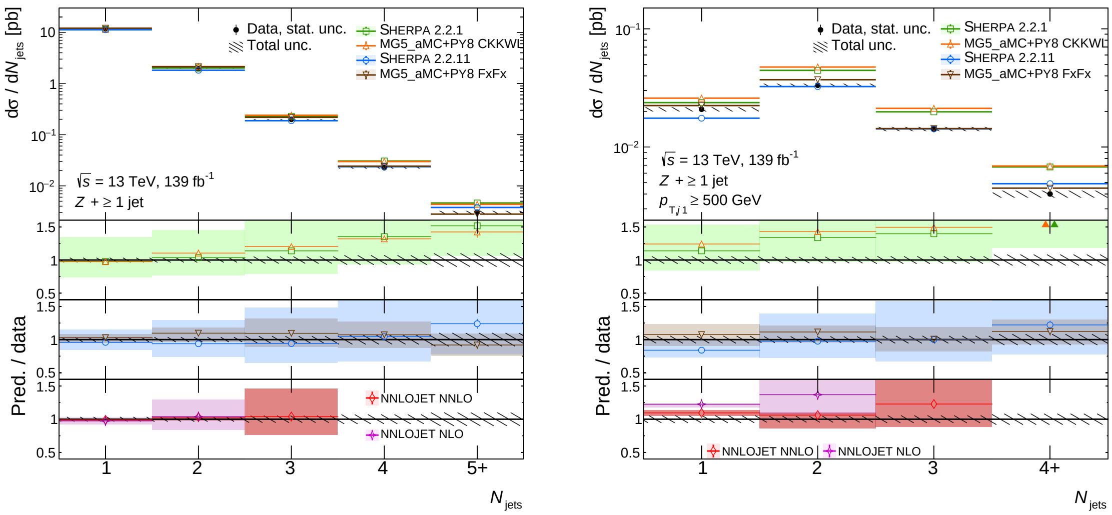
The Z-boson and jet transverse momenta (two correlated quantities) are fundamental observables of the Z+jets process and probe perturbative QCD over a wide range of scales. Moreover, understanding the kinematics of jets in events with vector bosons produced in association with several jets is essential for the modelling of backgrounds for other SM processes and searches beyond the SM. Figure 10.5 shows the differential fiducial cross section as a function of $p_{\mathrm{T},j1}$ and $p_{\mathrm{T},\ell\ell}$. The high-$p_{\mathrm{T},\ell\ell}$ region is dominated by the back-to-back topology and receives significant negative corrections due to EW effects. In contrast, events with a high-p_{T} jet typically result in both back-to-back and collinear topologies. The SHERPA 2.2.1 and MG5_aMC+PY8 CKKWL generators predict a harder $p_{\mathrm{T},j1}$ distribution than seen in the data, resulting in an overestimation of the cross section for high $p_{\mathrm{T},j1}$. Although they agree with data within uncertainties, their central predictions perform very poorly for high $p_{\mathrm{T},j1}$. In contrast, SHERPA 2.2.11 and MG5_aMC+PY8 FXFX show significantly better modelling of the $p_{\mathrm{T},j1}$ spectrum and are also in good agreement with the measured $p_{\mathrm{T},\ell\ell}$ spectrum. The smaller cross section from SHERPA 2.2.11 relative to SHERPA 2.2.1 in the *high-*p_{T} region is attributed to the improved matching scheme with a different treatment of unordered histories [72]. The prediction from MG5_aMC+PY8 FXFX models the data more precisely owing to the inclusion of NLO matrix elements with three partons. The $p_{\mathrm{T},\ell\ell}$ and $p_{\mathrm{T},j1}$ distributions show the significant improvement of the central predictions in SHERPA 2.2.11 and MG5_aMC+PY8 FXFX over SHERPA 2.2.1 and MG5_aMC+PY8 CKKWL. This improvement is not visible in the integrated fiducal cross-section measurements shown above and highlights the importance of the differential distributions of this analysis that benefits from the increased statistics of the full Run–2 dataset. The NNLOJET@NNLO predictions describe the data very precisely, except for very large values of $p_{\mathrm{T},\ell\ell}$ (and $p_{\mathrm{T},j1}$), where negative NLO virtual EW corrections of 10%–20% are expected and the QCD-only calculation overestimates the data.

Jet-multiplicity distributions provide an excellent probe of QCD. Whereas the Z+1 jet bin is most sensitive to PDF effects, those with higher jet multiplicities are more sensitive to perturbative QCD effects [172]. Jet-multiplicity distributions also probe the validity of predictions in the presence of jet vetoes, which are frequently used in searches that require a specific number of jets in their selection criteria. Figure 10.6 shows the differential fiducial cross section as a function of the jet multiplicity in the *inclusive* and *high-*p_{T} regions. As expected [172], the jet multiplicity in the *inclusive* phase space follows a downward staircase pattern, whereas in the *high-*p_{T} phase space, the cross section increases between 1-jet and 2-jet events followed by a downward pattern for higher jet multiplicities. Whereas SHERPA 2.2.1 and MG5_aMC+PY8 CKKWL tend to overestimate the cross section for higher jet multiplicities, SHERPA 2.2.11 and MG5_aMC+PY8 FXFX agree with the data for both the *inclusive* and *high-*p_{T} regions. The latter generator offers a higher level of precision. Although the four MC predictions agree with data within uncertainties, the central predictions from SHERPA 2.2.1 and MG5_aMC+PY8 CKKWL do not describe the data for events with a large number of jet. Once again, this poor modelling of the data was not visible in the integrated fiducial cross-section measurements. The NNLOJET predictions agree with the data in the *inclusive* and *high-*p_{T} phase spaces at high precision when it is expected from the order of the calculation, and at lower precision



(a) (b)

Figure 10.5: Differential fiducial cross section in the *inclusive* region as a function (a) of the transverse momentum of the leading jet $p_{\mathrm{T},j1}$ and (b) of the transverse momentum of the Z boson $p_{\mathrm{T},\ell\ell}$. The unfolded data are shown with the black points where the statistical uncertainty is given as an error bar and the total uncertainty as a hatched region. The total uncertainty is given by the sum in quadrature of the data statistical and total systematic uncertainties, as defined in Section 9.1.4. The data are compared with predictions from MC generators and fixed-order calculations. In the ratio panels, the error bars correspond to the statistical uncertainty of the prediction and solid triangles indicate that the prediction is outside the vertical-axis range, whereas the total uncertainty of the unfolded data is represented as the hatched region. The uncertainties on the theoretical predictions, dominated by the QCD scale uncertainties, are shown only in the ratio panels except for MG5_aMC+PY8 CKKWL, which is not included. They are found by summing in quadrature the uncertainties from the variations of the PDF (excluding NNLOJET predictions), QCD scales and, for SHERPA 2.2.11, virtual EW contributions, as explained in Sections 9.2.1, 9.2.2, and 9.2.3, respectively.



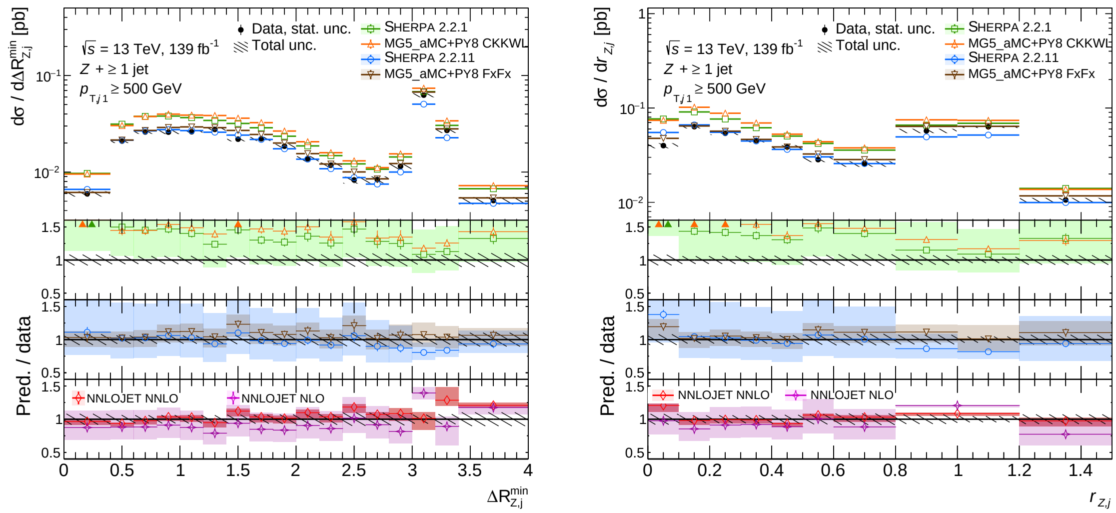
(a) (b)

Figure 10.6: Differential fiducial cross section as a function of the jet multiplicity N_{jets} (a) in the *inclusive* region and (b) in the *high-p_{T}* region. In each case, the last bin is inclusive in higher jet multiplicities. Further details on the legend and ratio panels are provided in the caption of Figure 10.5.

when the order of the calculation is exceeded.

The angular distance between the Z boson and the closest jet, $\Delta R^{\text{min}}_{Z,j}$, and the ratio of the Z-boson transverse momentum to the closest-jet transverse momentum, $r_{Z,j}$, provide a way to distinguish between the presence of collinear Z-boson emission and back-to-back topologies. In the *high-p_{T}* selection, the *collinear* region is sensitive to logarithmic enhancements in production proportional to $\alpha_s \ln^2(p_{\text{T},j1}/m_Z)$, whereas the *back-to-back* region receives non-negligible virtual EW corrections. Figure 10.7 shows the differential fiducial cross sections as a function of $\Delta R^{\text{min}}_{Z,j}$ and $r_{Z,j}$ in the *high-p_{T}* region. Both distributions show an accumulation of events at low and high values of these two quantities: *collinear* events populate the figures at values $\Delta R^{\text{min}}_{Z,j} \leq 1.4$ and $r_{Z,j} \leq 0.4$, whereas the *back-to-back* events are observed with $\Delta R^{\text{min}}_{Z,j} \approx \pi$ and $r_{Z,j} \approx 1.0$. The collinear events are expected to be dominated by diagrams corresponding to the EW radiation of a Z boson from one of the quark-jets in a dijet event. Consequently, they are expected to correspond to the accumulation of events with low values of $r_{Z,j}$. This hypothesis is validated in Figure 10.8a, which shows the measurement of the $r_{Z,j}$ distribution for the subset of *collinear* events defined by $\Delta R^{\rm min}_{Z,j} \leq 1.4$, where only the accumulation of low-$r_{Z,j}$ events is observed. In contrast, the measurement of the $r_{Z,j}$ distribution for the *back-to-back* selection defined by $\Delta R^{\rm min}_{Z,j} \geq 2.0$ is populated by events with $r_{Z,j} \approx 1$ as is shown by Figure 10.8b. The jet multiplicity is also measured separately for the *collinear* and *back-to-back* topologies as shown in Figure 10.9. It is found that the *collinear* region is dominated by dijet events, whereas the *back-to-back* region is dominated by Z+1 jet events. These distributions provide for the first time detailed measurements of real Z-boson emission from high-$p_{\rm T}$ jets and hard Z+1 jet back-to-back emissions in a high jet-$p_{\rm T}$ phase space.

Figures 10.7–10.9 show that while still marginally in agreement with data within modelling uncertainties of up to 50%, SHERPA 2.2.1 central values increasingly overestimate the cross section with decreasing values of $\Delta R^{\rm min}_{Z,j}$. A similar trend is observed for MG5_aMC+PY8 CKKWL. The modelling is improved by the state-of-the-art generators MG5_aMC+PY8 FXFX and SHERPA 2.2.11. Good agreement of MG5_aMC+PY8 FXFX and SHERPA 2.2.11 with data even for very collinear events indicates that resummation of the additional logarithmic terms, *e.g.* via EW radiation in parton showers, is not needed at the present level of theoretical and experimental precision in the experimentally accessible kinematic regime. The $\Delta R^{\rm min}_{Z,j}$ distribution in Figure 10.7a provides important information on the generator predictions that was missing from the integrated cross-section results. In the *collinear* region, where $\Delta R^{\rm min}_{Z,j} \leq 1.4$, the central predictions of SHERPA 2.2.1 and MG5_aMC+PY8 CKKWL describe the distribution of data very poorly. This behaviour points to an

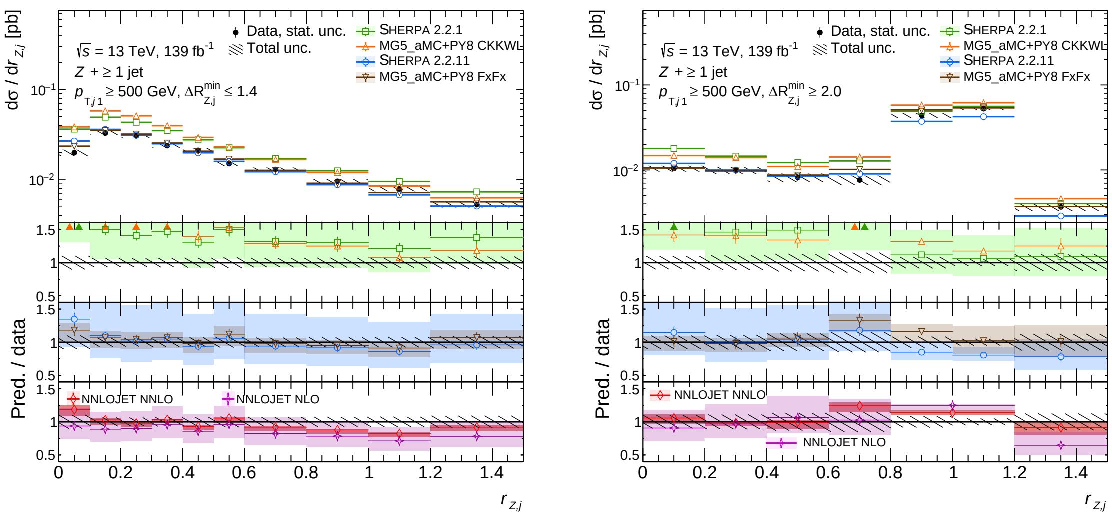


(a) (b)

Figure 10.7: Differential fiducial cross section in the *high-*p_{T} region as a function (a) of angular distance between the Z boson and the closest jet $\Delta R^{\mathrm{min}}_{Z,j}$ and (b) of the ratio of Z-boson to closest-jet transverse momenta $r_{Z,j}$. For the NNLOJET predictions, the bins around $r_{Z,j} = 1$ are merged to be insensitive to the singularity in the fixed-order calculation. Further details on the legend and ratio panels are provided in the caption of Figure 10.5.

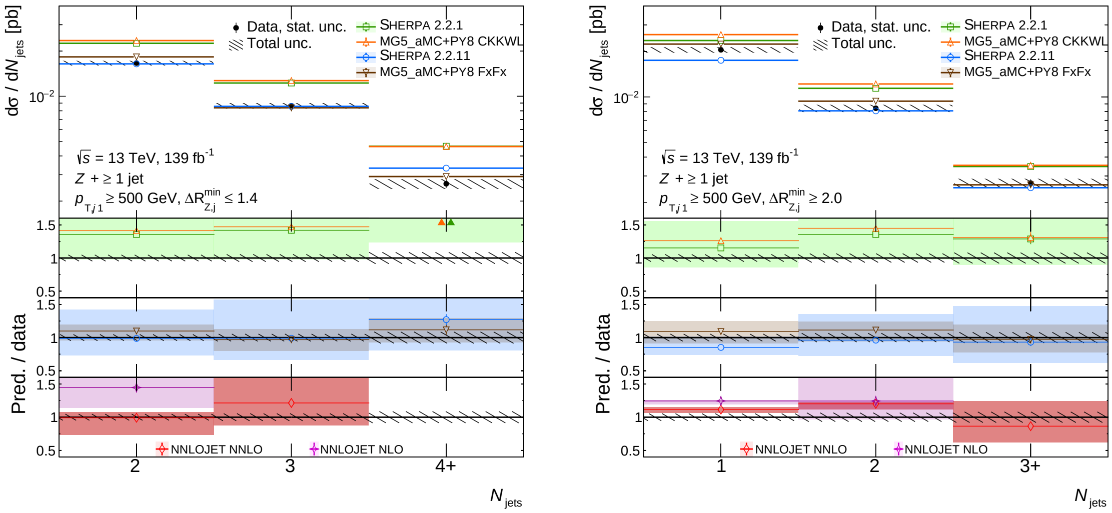
underlying problem with the generators and not to a cross-section normalization issue. The NNLOJET@NNLO prediction models the data distribution in both the *collinear* and the *back-to-back* regions with high precision. The good performance in the *collinear* phase space is remarkable, as this region contains a large fraction of events with at least three jets, which are simulated only at LO in NNLOJET@NNLO and not at all by NNLOJET@NLO. The QCD-only calculation overestimates the cross section for exact *back-to-back* events in the *high-*p_{T} region, dominated by high-$p_{\mathrm{T},\ell\ell}$ events, consistent with the pattern observed in Figure 10.5.

An alternative way to select a high-energy phase space is by requiring a large value of H_{T} or S_{T}. The former is often used as a dynamical scale choice, whereas the latter is more suited to selecting a phase space similar to the *high-*p_{T} region. Figure 10.10 shows the differential fiducial cross section as a function of H_{T}. The central



(a) (b)

Figure 10.8: Differential fiducial cross section as a function of the ratio of Z-boson to closest-jet transverse momenta $r_{Z,j}$ (a) in the *collinear* $\Delta R^{\min}_{Z,j} \leq 1.4$ region and (b) in the *back-to-back* $\Delta R^{\min}_{Z,j} \geq 2.0$ region. For the NNLOJET predictions, the bins around $r_{Z,j} = 1$ are merged to be insensitive to the singularity in the fixed-order calculation. Further details on the legend and ratio panels are provided in the caption of Figure 10.5.



(a) (b)

Figure 10.9: Differential fiducial cross section as a function of the jet multiplicity N_{jets} (a) in the *collinear* $\Delta R^{\text{min}}_{Z,j} \leq 1.4$ region and (b) in the *back-to-back* $\Delta R^{\text{min}}_{Z,j} \geq 2.0$ region. In each case, the last bin is inclusive in higher jet multiplicities. In data, no events with exactly one jet are selected in the *collinear* region. Further details on the legend and ratio panels are provided in the caption of Figure 10.5.

values from SHERPA 2.2.1 increasingly overestimate the cross section with increasing H_{T}, while still marginally agreeing with the data within modelling uncertainties of up to 50%. The prediction from MG5_aMC+PY8 CKKWL shows a similar trend. In contrast, the state-of-the-art predictions from SHERPA 2.2.11, MG5_aMC+PY8 FXFX and NNLOJET@NNLO model the data well, the latter with higher precision. Figure 10.11 shows the differential fiducial cross section as a function of $\Delta R^{\text{min}}_{Z,j}$ and jet multiplicity in the *high-*S_{T} region. The *high-*S_{T} region probes high-energy events, where the energy is typically shared by several jets. Compared to the *high-*p_{T} region, the jet multiplicity distribution is shifted towards higher values, and the *back-to-back* peak in the $\Delta R^{\text{min}}_{Z,j}$ distribution is suppressed relative to events where a jet is in closer proximity to the Z boson. As in the *high-*p_{T} region, SHERPA 2.2.1 is marginally consistent with the data within a large theoretical uncertainty, with the overestimation

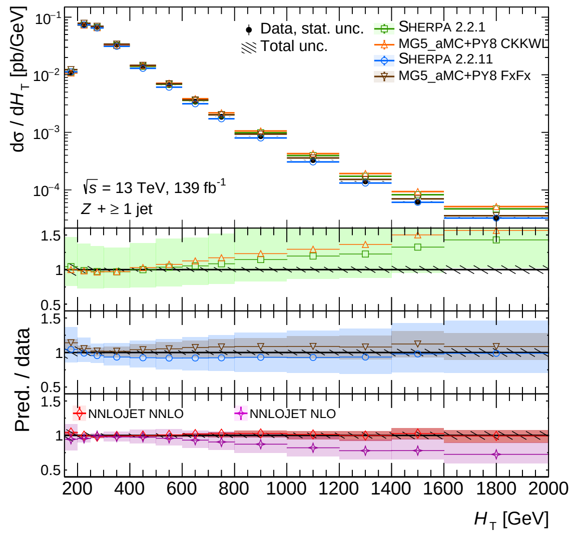
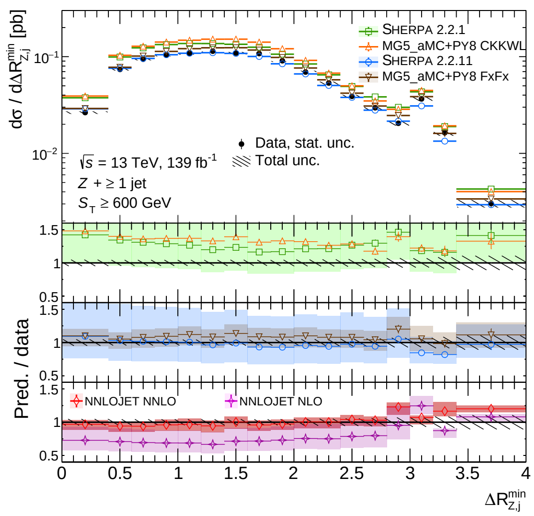


Figure 10.10: Differential fiducial cross section as a function of the transverse momenta scalar sum H_{T} in the *inclusive* region. Further details on the legend and ratio panels are provided in the caption of Figure 10.5.

of the central values most pronounced in the *collinear* region and for higher jet multiplicities, with MG5_aMC+Py8 CKKWL showing a similar trend. Both Sherpa 2.2.11 and MG5_aMC+Py8 FxFx are consistent with the data, the latter again at higher precision. The NNLOjet@NNLO prediction models this region well and with high precision, even though the fraction of events with at least three jets is larger than in the *high-*p_{T} phase space.



(a) (b)

Figure 10.11: Differential fiducial cross section in the *high*-S_{T} region as a function (a) of angular distance between the Z boson and the closest jet $\Delta R^{\min}_{Z,j}$ and (b) of the jet multiplicity N_{jets}. In the jet-multiplicity distribution, the last bin is inclusive in higher jet multiplicities. Further details on the legend and ratio panels are provided in the caption of Figure 10.5.

Chapter 11

Conclusion

This thesis presents the measurements of the production fiducial cross sections for a Z boson produced in association with high-transverse-momentum jets and decaying into a charged-lepton pair. The data were collected from pp collisions at $\sqrt{s} = 13$ TeV at the LHC between 2015 and 2018. The dataset was recorded with the ATLAS detector and corresponds to an integrated luminosity of 139 fb^{-1}. Measurements are performed on events that contain a Z-boson candidate reconstructed from an e^+e^- pair or a $\mu^+\mu^-$ pair in association with hadronic jets defined as jets having transverse momentum greater than or equal to 100 GeV.

The results from this analysis are publicly available in Ref. [19]. The cross-section measurements are collected in the following dedicated HEPdata [173] record: Ref. [174]. Additionally, a rivet analysis routine is provided and can be found in the rivet repository [175].

The analysis focuses on selections with very high leading-jet p_{T} ($p_{\mathrm{T},j1} \geq 500$ GeV) and very high S_{T} ($S_{\mathrm{T}} \geq 600$ GeV), which are used to study two populations of events – collinear events and back-to-back events – with distinct patterns in distributions of $\Delta R^{\mathrm{min}}_{Z,j}$, $r_{Z,j}$, and jet multiplicity. This is the first time that a measurement is performed that explicitly focuses on understanding the contributions of real Z-boson emission from high-p_{T} jets, or collinear events, in Z+jets processes. Moreover, this is also the first measurement that focuses on measuring events with very large Z-boson and leading-jet p_{T} corresponding to back-to-back Z+1 jet events.

The data distributions are unfolded to the particle level and compared with state-of-the-art generator predictions and fixed-order calculations. The integrated fiducial cross section for Z+high-p_{T} jets production are measured in the *inclusive*, *high-p_{T}*, *collinear*, *back-to-back* and *high-S_{T}* signal regions for the first time using the full statistical precision of the LHC Run–2 dataset. These integrated fiducial cross-section measurements are compared to five different theoretical predictions, all of which are in agreement with the observed values in data. The *high-p_{T}* region requires the presence of a jet with $p_{\mathrm{T}} \geq 500$ GeV to enhance the production of *collinear* and *back-to-back* events. The *collinear* region is a subset of the *high-p_{T}* region and is designed to select events that correspond to the real Z-boson emission from a high-p_{T} jet in dijet events. Due to kinematic rules, the Z boson and the closest jet are expected to be measured with a small separating angle $\Delta R^{\mathrm{min}}_{Z,j}$ and with small values of the ratio of Z-boson to closest-jet transverse momenta $r_{Z,j}$. In contrast, the *back-to-back* region is defined as a subset of the *high-p_{T}* region with $\Delta R^{\mathrm{min}}_{Z,j} \geq 2.0$ to enhance the selection of Z+1 jet events where the Z boson and jet are travelling in opposite directions such that $\Delta R^{\mathrm{min}}_{Z,j} \sim \pi$ and $r_{Z,j} \sim 1.0$.

The measured integrated fiducial cross sections are:

$$\sigma^{inclusive} = 13.90 \pm 0.01(\text{stat.}) \pm 0.47(\text{syst.})\ \text{pb},$$

$$\sigma^{high\text{-}p_{\mathrm{T}}} = 72.3 \pm 1.5(\text{stat.}) \pm 3.5(\text{syst.})\ \text{fb},$$

$$\sigma^{collinear} = 27.9 \pm 0.8(\text{stat.}) \pm 1.2(\text{syst.})\ \text{fb},$$

$$\sigma^{back\text{-}to\text{-}back} = 31.6 \pm 0.8(\text{stat.}) \pm 1.7(\text{syst.})\ \text{fb and}$$

$$\sigma^{high\text{-}S_\mathrm{T}} = 226.0 \pm 2.6(\text{stat.}) \pm 9.5(\text{syst.})\ \text{fb}.$$

These results consist of some of the smallest Z boson production cross sections measured by ATLAS so far, as shown in Figure 11.1. The measurements exceed the precision of the currently available theoretical predictions from MC generators. The NNLO prediction from the NNLOJET fixed-order calculation instead matches or exceeds the precision of the measurements in some kinematic regions. In general, a meaningful interpretation of the MC generators' performances is difficult due to their unsightly large modelling uncertainties, sometimes reaching upwards of 50%. A better understanding of the causes behind these modelling uncertainties is needed in order to improve the precision of the next-generation MC generators and to match the currently accessible experimental uncertainties.

Differential fiducial cross-section measurements are also performed and compared to six theoretical predictions. These measurements provide crucial information on the performance of the different theoretical predictions in modelling a range of key distributions in the Z+high-p_T jets process. The predictions from both SHERPA 2.2.1 and MG5_aMC+PY8 CKKWL overestimate the cross sections for large values of $p_{\mathrm{T},j1}$, H_T and S_T. The predictions from MG5_aMC+PY8 FXFX, with matrix elements for up to three partons at NLO, offer a significant improvement over MG5_aMC+PY8 CKKWL (which models up to four partons at LO) and in general match the data with high precision. Similarly, SHERPA 2.2.11, with the addition of a fifth parton at LO in the matrix element, the addition of NLO virtual EW corrections, and a different treatment of unordered histories in the parton shower, shows

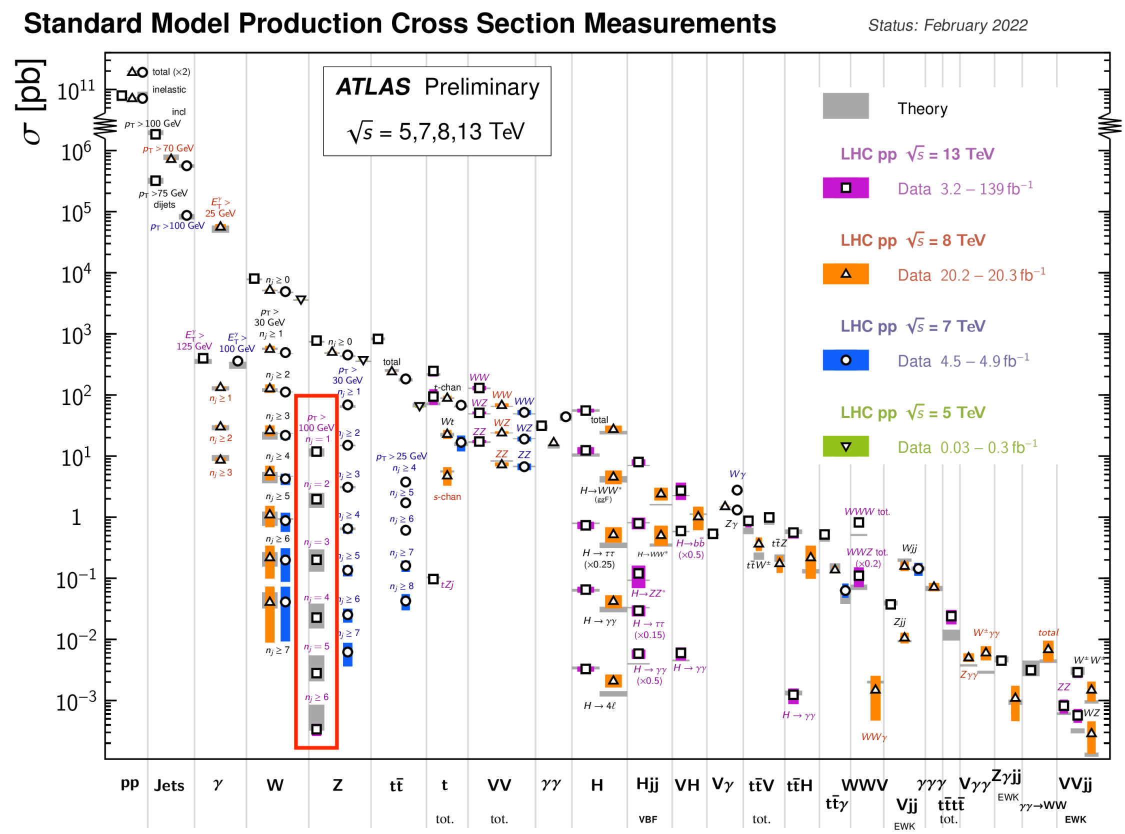


Figure 11.1: Summary of several Standard Model total and fiducial production cross-section measurements from the ATLAS experiment [70]. The measurements are corrected for branching fractions and compared to the corresponding theoretical expectations. The fiducial cross-section measurements of Z+high-p_{T} jets production presented in this thesis are highlighted in the red square as a function of jet multiplicity for the *inclusive* region.

a significant improvement over Sherpa 2.2.1 and agrees well with the data. The NNLO calculations at fixed order from NNLOjet describe the data cross sections at a very high level of precision, including in high-p_{T} regions where a sizeable fraction of the events have more than two jets. The calculation exhibits a harder $p_{\mathrm{T},\ell\ell}$ spectrum than the data in a region where larger negative EW corrections are expected. The generally poor performance of Sherpa 2.2.1 and MG5_aMC+Py8 CKKWL in describing high jet p_{T}, *collinear* and *back-to-back* events expresses the need of future measurements to use state-of-the-art predictions that can properly model the Z+high-p_{T} jets process.

The predicted signatures of collinear Z boson radiation and Z+1 jet back-to-back scattering are found to be in agreement with data. Indeed, the *collinear* region is measured to be dominated by Z+2 jets events. The collinear emission of a Z boson from a quark-jet is also described by events where the angle between the Z boson and the closest jet is small, $\Delta R^{\mathrm{min}}_{Z,j} \leq 1.4$. Moreover, in these events, the ratio of Z-boson to closest-jet transverse momenta $r_{Z,j}$ is small and generally falls below values of 0.4. Furthermore, the *back-to-back* region is measured to be dominated by Z+1 jet events with most likely values of $\Delta R^{\mathrm{min}}_{Z,j} \sim \pi$ and $r_{Z,j} \sim 1.0$. This analysis presents the first ever measurements of the jet multiplicity, $\Delta R^{\mathrm{min}}_{Z,j}$ and $r_{Z,j}$ observables in *collinear* and *back-to-back* events.

These measurements have an important impact on the future of MC generators and fixed-order calculations that model Z+high-p_{T} jets events. The measured differential fiducial cross sections greatly exceed the precision of the theoretical predictions against which they are compared. Hence, the limiting factor for future measurements that are sensitive to Z+high-p_{T} jets processes is the precision of the theoretical predictions. These results will serve two important purposes: 1) they provide important information on the ability of different MC generators to describe the shape of data distributions, and 2) they act as benchmarks for the future development of next-generation event generators and fixed-order calculations. In analyses where Z+jets processes constitute a background, the background distributions are often estimated via template fits. Hence, the ability of generators to model the shape of a given distribution is much more important than predicting the correct integrated cross section. Furthermore, these measurements inform theorists on possible underlying problems, or missing higher-order corrections, that have an impact on the ability of a prediction to model the data. One such crucial correction is the logarithmic enhancement in collinear emissions of the Z boson. In contrast, due the low precision of SHERPA 2.2.11, the importance of NLO virtual electroweak corrections can not be meaningfully interpreted at this time. Novel MC generators would give the ability to study more precisely processes with Z, W or Higgs bosons produced in association with jets. The High-Luminosity LHC program calls for a detailed investigation of rare electroweak processes such as Vector Boson Fusion (VBF) and Vector Boson Scattering (VBS), where it is critical to understand with unprecedented precision the strong Z/W/H+jets backgrounds.

The precise measurement of Z+high-p_{T} jets events, and especially *collinear* and *back-to-back* events, has just now been unlocked with the current centre-of-mass energy and full Run–2 dataset provided by the LHC. These rare production processes are relevant for many Higgs boson measurements and searches for new physics at the LHC, *e.g.* where the production of a Z boson collinear to a jet can become a significant background for the rare sought after signatures. As the centre-of-mass energy and rate at which data is provided by the LHC increase, the Z+high-p_{T} jets process will become more prevalent in the results of high-energy proton–proton collisions.

Although the precision of the measurements presented in this thesis will be surpassed soon, these results provide invaluable information on the modelling of Z+high-p_{T} jets events that will be used for the physics analyses of, and hopefully beyond, the LHC Run–3.

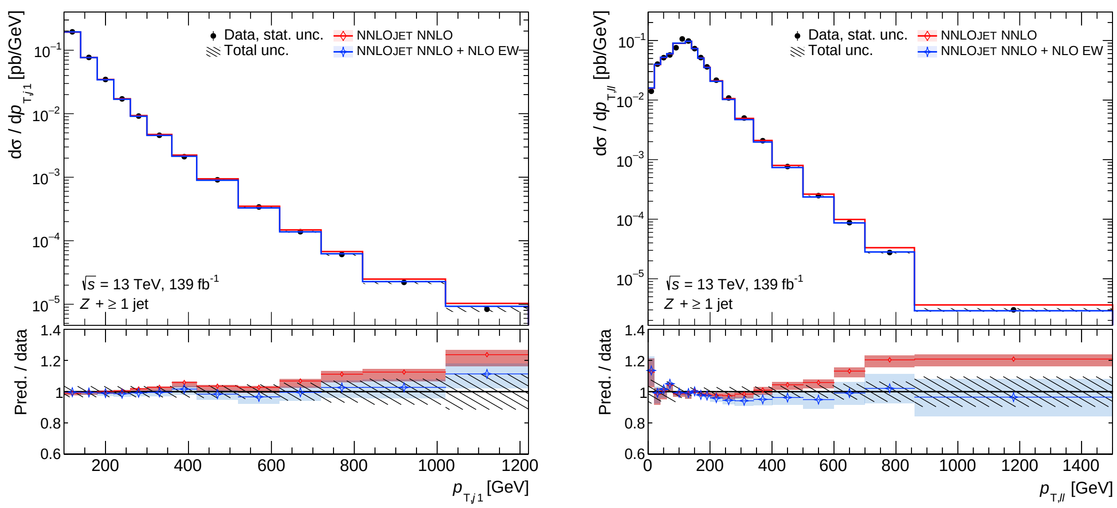
Appendix A

NNLOJET@NNLO predictions with NLO virtual EW corrections

The predictions from NNLOjet@NNLO are the only ones presented in this analysis that are precise enough to be sensitive to NLO virtual EW corrections. To study the importance of these corrections on theoretical predictions, their relative size and uncertainties are extracted from the Sherpa 2.2.11 predictions and applied to those from NNLOjet@NNLO. Differential fiducial cross-sections measurements are compared in Figures A.1–A.2 with predictions from NNLOjet@NNLO and NNLOjet@NNLO supplied with NLO virtual electroweak corrections from Sherpa 2.2.11. The data statistical uncertainty is displayed with error bars and the hatched bands correspond to the total uncertainty defined as the sum in quadrature of the data statistical and total systematic uncertainties, as defined in Section 9.1.4. In the ratio panels, the MC statistical uncertainty is displayed with coloured error bars and the theoretical uncertainties are shown as coloured bands. The uncertainties on the predictions are found by summing in quadrature the uncertainties from the variations of the QCD scales and, if applicable, virtual EW contributions, as explained in Sections 9.2.2 and 9.2.3, respectively. The EW corrections are largest in phase spaces with large $p_{\mathrm{T},\ell\ell}$ and $p_{\mathrm{T},j1}$, as observed in Figure A.1. In the tail of the $p_{\mathrm{T},j1}$ and $p_{\mathrm{T},\ell\ell}$ distributions, the nominal NNLOJET@NNLO predictions overestimate data but near perfect agreement is observed with the application of NLO virtual EW corrections.

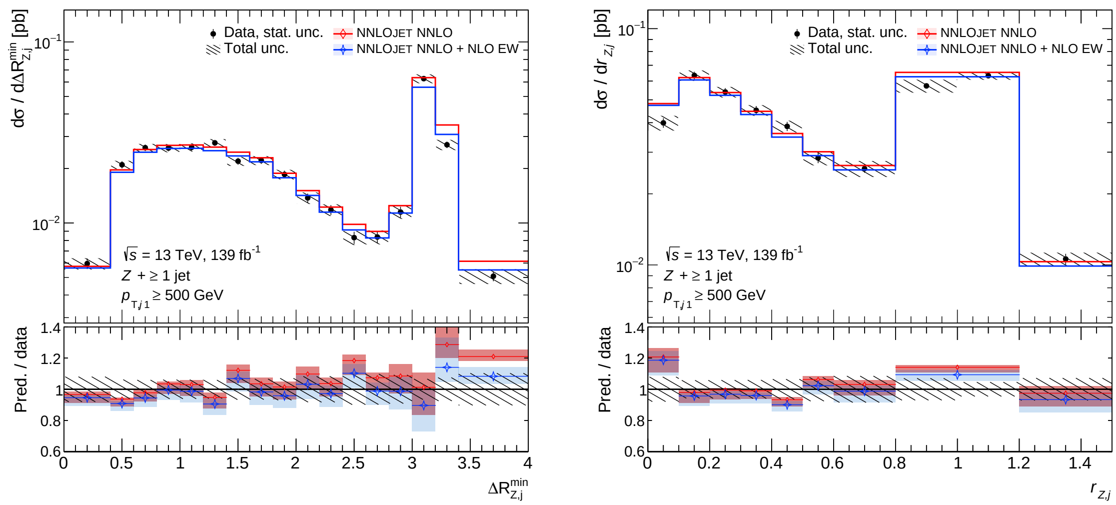
Figure A.2 compares the distributions for the $\Delta R^{\mathrm{min}}_{Z,j}$ and $r_{Z,j}$ observables. As expected, the impact of the EW corrections is small except for the *back-to-back* peaks at $\Delta R^{\mathrm{min}}_{Z,j} \sim \pi$ and $r_{Z,j} \sim 1.0$ that are described by Z+1 jet events where the Z-boson p_{T} is very large. At the exact back-to-back peak $\Delta R^{\mathrm{min}}_{Z,j} = \pi$ of the $\Delta R^{\mathrm{min}}_{Z,j}$ distribution, the EW corrections increase the disagreement between the prediction and data. Discussions with NNLOJET authors have pointed to possible bin migration issues since the following bin of $3.2 \leq \Delta R^{\mathrm{min}}_{Z,j} \leq 3.4$ instead overestimates the data. Furthermore, in the $r_{Z,j}$ distribution, the EW corrections improve the prediction in the *back-to-back* peak, $r_{Z,j} \sim 1.0$, while having a negligible impact in the other bins.

These results suggest that NLO virtual EW corrections are important to predictions of Z+jets processes where the Z-boson has large transverse momentum. The corrections have an impact of almost 25% for $p_{\mathrm{T},\ell\ell} \sim 1$ TeV events and correct the NNLOJET@NNLO prediction from overestimating to almost perfectly describing data. However, since the corrections are supplied from the SHERPA 2.2.11 predictions and not directly calculated within NNLOJET, these conclusions are sensitive to the differences between the generator and fixed-order calculation programs. To confirm these results, the NNLOJET program would need to include these corrections in its calculations. Although SHERPA 2.2.11 is too imprecise to meaningfully interpret the importance of these corrections, its central predictions with virtual EW corrections generally give a better agreement with the data, as shown in Chapter 10.



(a) (b)

Figure A.1: Differential fiducial cross section in the *inclusive* region as a function (a) of the transverse momentum of the leading jet $p_{\mathrm{T},j1}$ and (b) of the transverse momentum of the Z boson $p_{\mathrm{T},\ell\ell}$. The unfolded data are shown with the black points where the statistical uncertainty is given as an error bar and the total uncertainty as a hatched region. The total uncertainty is given by the sum in quadrature of the data statistical and total systematic uncertainties, as defined in Section 9.1.4. The data are compared with predictions (in red) from NNLOJET@NNLO and (in blue) NNLOJET@NNLO supplied with NLO virtual electroweak corrections from SHERPA 2.2.11. In the ratio panels, the total uncertainty of the unfolded data is represented as the hatched region. The red uncertainty bands corresponds to the scale variations of the NNLOJET@NNLO prediction, while the blue uncertainty band also contains the envelope of the different virtual EW correction schemes from SHERPA 2.2.11, as explained in Sections 9.2.2, and 9.2.3, respectively.



(a) (b)

Figure A.2: Differential fiducial cross section in the *high*-p_{T} region as a function (a) of angular distance between the Z boson and the closest jet $\Delta R^{\mathrm{min}}_{Z,j}$ and (b) of the ratio of Z-boson to closest-jet transverse momenta $r_{Z,j}$. For the predictions, the bins around $r_{Z,j} = 1$ are merged to be insensitive to the singularity in the fixed-order calculation. The unfolded data are shown with the black points where the statistical uncertainty is given as an error bar and the total uncertainty as a hatched region. The total uncertainty is given by the sum in quadrature of the data statistical and total systematic uncertainties, as defined in Section 9.1.4. The data are compared with predictions (in red) from NNLOJET@NNLO and (in blue) NNLOJET@NNLO supplied with NLO virtual electroweak corrections from SHERPA 2.2.11. In the ratio panels, the total uncertainty of the unfolded data is represented as the hatched region. The red uncertainty bands corresponds to the scale variations of the NNLOJET@NNLO prediction, while the blue uncertainty band also contains the envelope of the different virtual EW correction schemes from SHERPA 2.2.11, as explained in Sections 9.2.2, and 9.2.3, respectively.

References

[1] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.

[2] CMS Collaboration, *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004.

[3] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.

[4] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1, [arXiv:1207.7214].

[5] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30, [arXiv:1207.7235].

[6] ATLAS Collaboration, *Measurements of the production cross-section for a Z boson in association with b-jets in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **07** (2020) 044, [arXiv:2003.11960].

[7] ATLAS Collaboration, *Differential cross-section measurements for the electroweak production of dijets in association with a Z boson in proton–proton collisions at ATLAS*, *Eur. Phys. J. C* **81** (2021) 163, [arXiv:2006.15458].

[8] D. J. Gross and F. Wilczek, *Asymptotically Free Gauge Theories. I*, *Phys. Rev. D* **8** (1973) 3633.

[9] H. D. Politzer, *Asymptotic freedom: an approach to strong interactions*, *Phys. Rep.* **14** (1974) 129.

[10] M. Rubin, G. P. Salam, and S. Sapeta, *Giant QCD K-factors beyond NLO*, *JHEP* **09** (2010) 084, [arXiv:1006.2144].

[11] J. R. Christiansen and S. Prestel, *Merging weak and QCD showers with matrix elements*, *Eur. Phys. J. C* **76** (2016) 39, [arXiv:1510.01517].

[12] R. Boughezal, C. Focke, and X. Liu, *Jet vetoes versus giant K-factors in the exclusive Z + 1-jet cross section*, *Phys. Rev. D* **92** (2015) 09400, [arXiv:1501.01059].

[13] J. R. Christiansen and T. Sjöstrand, *Weak gauge boson radiation in parton showers*, *JHEP* **04** (2014) 115, [arXiv:1401.5238].

[14] ATLAS Collaboration, *Measurement of the associated production of a Higgs boson decaying into b-quarks with a vector boson at high transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett. B* **816** (2021) 136204, [arXiv:2008.02508].

[15] ATLAS Collaboration, *Measurement of the production cross section for a Higgs boson in association with a vector boson in the $H \to WW^* \to \ell\nu\ell\nu$ channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett. B* **798** (2019) 134949, [arXiv:1903.10052].

[16] ATLAS Collaboration, *Search for squarks and gluinos in final states with one isolated lepton, jets, and missing transverse momentum at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **81** (2021) 600, [arXiv:2101.01629].

[17] ATLAS Collaboration, *Searches for electroweak production of supersymmetric particles with compressed mass spectra in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, *Phys. Rev. D* **101** (2020) 052005, [arXiv:1911.12606].

[18] CMS Collaboration, *Search for a heavy vector resonance decaying to a Z boson and a Higgs boson in proton–proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **81** (2021) 688, [arXiv:2102.08198].

[19] ATLAS Collaboration, *Cross-section measurements for the production of a Z boson in association with high-transverse-momentum jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [arXiv:2205.02597].

[20] ATLAS Collaboration, *Measurements of the production cross section of a Z boson in association with jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **77** (2017) 361, [arXiv:1702.05725].

[21] ATLAS Collaboration, *Measurement of W boson angular distributions in events with high transverse momentum jets at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *Phys. Lett. B* **765** (2017) 132, [arXiv:1609.07045].

[22] CMS Collaboration, *Measurements of the differential cross sections of the production of Z+jets and γ+jets and of Z boson emission collinear with a jet in pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **05** (2021) 285, [arXiv:2102.02238].

[23] I. Bird, *Computing for the Large Hadron Collider*, *Annual Review of Nuclear and Particle Science* **61** (2011), no. 1 99–118.

[24] "Standard model of elementary particles." https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

[25] P.A. Zyla et al. (Particle Data Group), *Review of Particle Physics*, *PTEP* **2020** (2020), no. 8 083C01, and 2021 update.

[26] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, *Phys. Lett. B* **47** (1973) 365–368.

[27] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343–1346.

[28] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346–1349.

[29] S. L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579–588.

[30] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, *Phys. Lett.* **13** (1964) 168–171.

[31] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[32] A. Zee, *Group Theory in a Nutshell for Physicists.* Princeton University Press, USA, 2016.

[33] S. Weinberg, *The Making of the standard model*, *Eur. Phys. J. C* **34** (2004) 5–13, [hep-ph/0401010].

[34] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory.* Addison-Wesley, Reading, USA, 1995.

[35] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508–509.

[36] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321–323.

[37] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585–587.

[38] J. Ellis, M. K. Gaillard, and D. V. Nanopoulos, *A Historical Profile of the Higgs Boson*, pp. 255–274. 2016. [arXiv:1504.07217].

[39] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J. C* **75** (2015), no. 5 204, [arXiv:1412.3989].

[40] F. J. Hasert et al., *Search for Elastic ν_μ Electron Scattering*, *Phys. Lett. B* **46** (1973) 121–124.

[41] F. J. Hasert et al., *Observation of Neutrino Like Interactions without Muon or Electron in the Gargamelle Neutrino Experiment*, *Nucl. Phys. B* **73** (1974) 1–22.

[42] A. Astbury et al., *A 4π Solid Angle Detector for the SPS Used as a Proton-Antiproton Collider at a Centre of Mass Energy of 540 GeV*, .

[43] M. Banner et al., *Proposal to Study Antiproton-Proton Interactions at 540 GeV CM Energy*, .

[44] G. Arnison et al., *Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider*, *Phys. Lett. B* **126** (1983) 398–410.

[45] P. Bagnaia et al., *Evidence for $Z^0 \to \mathrm{e}^+\mathrm{e}^-$ at the CERN $\bar{p}p$ Collider*, *Phys. Lett. B* **129** (1983) 130–140.

[46] G. Arnison et al., *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, *Phys. Lett. B* **122** (1983) 103–116.

[47] M. Banner et al., *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, *Phys. Lett. B* **122** (1983) 476–485.

[48] A. J. Buras and K. J. F. Gaemers, *Simple Parametrizations of Parton Distributions with q**2 Dependence Given by Asymptotic Freedom*, *Nucl. Phys. B* **132** (1978) 249–267.

[49] ALEPH Collaboration, *ALEPH: A detector for electron-positron annnihilations at LEP*, *Nucl. Instrum. Meth. A* **294** (1990) 121–178. [Erratum: Nucl.Instrum.Meth.A 303, 393 (1991)].

[50] DELPHI Collaboration, *The DELPHI detector at LEP*, *Nucl. Instrum. Meth. A* **303** (1991) 233–276.

[51] L3 Collaboration, *The Construction of the L3 Experiment*, *Nucl. Instrum. Meth. A* **289** (1990) 35–102.

[52] OPAL Collaboration, *The OPAL detector at LEP*, *Nucl. Instrum. Meth. A* **305** (1991) 275–319.

[53] S. Schael et al., *Precision electroweak measurements on the Z resonance*, *Phys. Rept.* **427** (2006) 257–454, [hep-ex/0509008].

[54] CDF Collaboration, *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** (2008) 102001, [arXiv:0711.3717].

[55] D0 Collaboration, *Measurement of differential Z/γ^* + jet + X cross sections in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96-TeV*, *Phys. Lett. B* **669** (2008) 278–286, [arXiv:0808.1296].

[56] ATLAS Collaboration, *Measurement of the production cross section for Z/γ^* in association with jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *Phys. Rev. D* **85** (2012) 032009, [arXiv:1111.2690].

[57] ATLAS Collaboration, *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** (2013) 032, [arXiv:1304.7098].

[58] CMS Collaboration, *Jet Production Rates in Association with W and Z Bosons in pp Collisions at $\sqrt{s} = 7$ TeV*, *JHEP* **01** (2012) 010, [arXiv:1110.3226].

[59] CMS Collaboration, *Measurements of jet multiplicity and differential production cross sections of Z + jets events in proton–proton collisions at $\sqrt{s} = 7$ TeV*, *Phys. Rev. D* **91** (2015) 052008, [arXiv:1408.3104].

[60] CMS Collaboration, *Comparison of the Z/γ^* + jets to γ + jets cross sections in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **10** (2015) 128, [arXiv:1505.06520].

[61] CMS Collaboration, *Measurements of the differential production cross sections for a Z boson in association with jets in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **04** (2017) 022, [arXiv:1611.03844].

[62] LHCb Collaboration, *Measurement of forward W and Z boson production in association with jets in proton-proton collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **05** (2016) 131, [arXiv:1605.00951].

[63] CMS Collaboration, *Measurement of differential cross sections for Z boson production in association with jets in proton–proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **78** (2018) 965, [arXiv:1804.05252].

[64] CMS Collaboration, *Study of Z bosons plus jets using variables sensitive to double-parton scattering in pp collisions at* 13 *TeV*, *JHEP* **10** (2021) 176, [arXiv:2105.14511].

[65] CMS Collaboration, *Measurement of the production cross section for* $Z + b$ *jets in proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, [arXiv:2112.09659].

[66] ATLAS Collaboration, *Determination of the parton distribution functions of the proton from ATLAS measurements of differential* $W^{\pm}$ *and Z boson production in association with jets*, *JHEP* **07** (2021) 223, [arXiv:2101.05095].

[67] ATLAS Collaboration, *Measurements of* $W^+W^- + \geq 1$ *jet production cross-sections in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, *JHEP* **06** (2021) 003, [arXiv:2103.10319].

[68] ATLAS Collaboration, *Observation of electroweak production of two jets and a Z-boson pair with the ATLAS detector at the LHC*, [arXiv:2004.10612].

[69] ATLAS Collaboration, *Measurement of the* $Z(\rightarrow \ell^+\ell^-)\gamma$ *production cross-section in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, *JHEP* **03** (2020) 054, [arXiv:1911.04813].

[70] ATLAS Collaboration, *Standard Model Summary Plots February 2022*, .

[71] CMS Collaboration, *Event shapes and azimuthal correlations in Z + jets events in pp collisions at* $\sqrt{s} = 7$ *TeV*, *Phys. Lett. B* **722** (2013) 238, [arXiv:1301.1646].

[72] ATLAS Collaboration, *Modelling and computational improvements to the simulation of single vector-boson plus jet processes for the ATLAS experiment*, [arXiv:2112.09588].

[73] W. J. Stirling and E. Vryonidou, *Electroweak corrections and Bloch-Nordsieck violations in 2-to-2 processes at the LHC*, *JHEP* **04** (2013) 155, [arXiv:1212.6537].

[74] A. Denner, S. Dittmaier, T. Kasprzik, and A. Muck, *Electroweak corrections to W + jet hadroproduction including leptonic W-boson decays*, *JHEP* **08** (2009) 075, [arXiv:0906.1656].

[75] A. Denner, S. Dittmaier, T. Kasprzik, and A. Muck, *Electroweak corrections to dilepton + jet production at hadron colliders*, *JHEP* **06** (2011) 069, [arXiv:1103.0914].

[76] A. Denner, S. Dittmaier, T. Kasprzik, and A. Mück, *Electroweak corrections to monojet production at the LHC*, *Eur. Phys. J. C* **73** (2013), no. 2 2297, [arXiv:1211.5078].

[77] A. Denner, L. Hofer, A. Scharf, and S. Uccirati, *Electroweak corrections to lepton pair production in association with two hard jets at the LHC*, *JHEP* **01** (2015) 094, [arXiv:1411.0916].

[78] S. Kallweit, J. M. Lindert, P. Maierhöfer, S. Pozzorini, and M. Schönherr, *NLO QCD+EW predictions for V + jets including off-shell vector-boson decays and multijet merging*, *JHEP* **04** (2016) 021, [arXiv:1511.08692].

[79] E. Mobs, "The CERN Accelerator Complex. Complexe des Accélérateurs du CERN." General Photo, Jul, 2016.

[80] ATLAS Collaboration, "Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC." ATLAS-CONF-2019-021, 2019.

[81] G. Avoni et al., *The new lucid-2 detector for luminosity measurement and monitoring in atlas*, *JINST* **13** (2018), no. 07 P07017.

[82] J. Pequenao, "Computer generated image of the whole ATLAS detector." Mar, 2008.

[83] J. Pequenao, "Computer generated image of the ATLAS Inner Detector." Mar, 2008.

[84] ATLAS Collaboration, *ATLAS Pixel Detector: Technical Design Report*, 1998.

[85] ATLAS Collaboration, *ATLAS Insertable B-Layer: Technical Design Report*, 2010.

[86] ATLAS Collaboration, *Operation and performance of the ATLAS semiconductor tracker in LHC Run 2*, *JINST* **17** (2021) P01013, [arXiv:2109.02591].

[87] J. Pequenao, "Computer generated image of the ATLAS calorimeter." Mar, 2008.

[88] ATLAS Collaboration, *ATLAS liquid argon calorimeter: Technical design report*, 12, 1996.

[89] ATLAS Collaboration, *ATLAS tile calorimeter: Technical design report*, 12, 1996.

[90] ATLAS Collaboration, *ATLAS muon spectrometer: Technical design report*, 6, 1997.

[91] J. Pequenao, "Computer generated image of the ATLAS muons subsystem." Mar, 2008.

[92] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, *Eur. Phys. J. C* **77** (2017) 317, [arXiv:1611.09661].

[93] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007, [arXiv:0811.4622].

[94] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, A. Huss, and T. A. Morgan, *Precise QCD Predictions for the Production of a Z Boson in Association with a Hadronic Jet*, *Phys. Rev. Lett.* **117** (2016) 022001, [arXiv:1507.02850].

[95] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, A. Huss, and T. A. Morgan, *The NNLO QCD corrections to Z boson production at large transverse momentum*, *JHEP* **07** (2016) 133, [arXiv:1605.04295].

[96] B. Andersson, *The Lund model*, vol. 7. Cambridge University Press, 7, 2005.

[97] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, *Phys. Rept.* **97** (1983) 31–145.

[98] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, *Nucl. Phys. B* **238** (1984) 492–528.

[99] L. Lönnblad, *Correcting the Colour-Dipole Cascade Model with Fixed Order Matrix Elements*, *JHEP* **05** (2002) 046, [hep-ph/0112284].

[100] L. Lönnblad and S. Prestel, *Matching tree-level matrix elements with interleaved showers*, *JHEP* **03** (2012) 019, [arXiv:1109.4829].

[101] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *A critical appraisal of NLO+PS matching methods*, *JHEP* **09** (2012) 049, [arXiv:1111.1220].

[102] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, *JHEP* **04** (2013) 027, [arXiv:1207.5030].

[103] S. Catani, F. Krauss, B. R. Webber, and R. Kuhn, *QCD Matrix Elements + Parton Showers*, *JHEP* **11** (2001) 063, [hep-ph/0109231].

[104] S. Hoeche, F. Krauss, S. Schumann, and F. Siegert, *QCD matrix elements and truncated showers*, *JHEP* **05** (2009) 053, [arXiv:0903.1219].

[105] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*, *JHEP* **12** (2012) 061, [arXiv:1209.6215].

[106] S. Kallweit, J. M. Lindert, P. Maierhöfer, S. Pozzorini, and M. Schönherr, *NLO electroweak automation and precise predictions for W+multijet production at the LHC*, *JHEP* **04** (2015) 012, [arXiv:1412.5157].

[107] ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, *Eur. Phys. J. C* **73** (2013) 2518, [arXiv:1302.4393].

[108] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823, [arXiv:1005.4568].

[109] GEANT4 Collaboration, S. Agostinelli, et al., GEANT4 – *a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.

[110] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039, [arXiv:0808.3674].

[111] F. Buccioni, J.-N. Lang, J. M. Lindert, P. Maierhöfer, S. Pozzorini, H. Zhang, and M. F. Zoller, *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019), no. 10 866, [arXiv:1907.13071].

[112] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601, [arXiv:1111.5206].

[113] A. Denner, S. Dittmaier, and L. Hofer, COLLIER*: A fortran-based complex one-loop library in extended regularizations*, *Comput. Phys. Commun.* **212** (2017) 220–238, [arXiv:1604.06792].

[114] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani–Seymour dipole factorisation*, *JHEP* **03** (2008) 038, [arXiv:0709.1027].

[115] E. Bothmann et al., *Event generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019), no. 3 034, [arXiv:1905.09127].

[116] S. Catani, S. Dittmaier, M. H. Seymour, and Z. Trocsanyi, *The Dipole formalism for next-to-leading order QCD calculations with massive partons*, *Nucl. Phys. B* **627** (2002) 189–265, [hep-ph/0201036].

[117] R. D. Ball et al., *Parton distributions for the LHC run II*, *JHEP* **04** (2015) 040, [arXiv:1410.8849].

[118] 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** (2014) 079, [arXiv:1405.0301].

[119] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159, [arXiv:1410.3012].

[120] ATLAS Collaboration, "ATLAS Pythia 8 tunes to 7 TeV data." ATL-PHYS-PUB-2014-021, 2014.

[121] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244–289, [arXiv:1207.1303].

[122] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852–867, [arXiv:0710.3820].

[123] M. Bahr, S. Gieseke, M. Gigg, D. Grellscheid, K. Hamilton, et al., *Herwig++ physics and manual*, *Eur. Phys. J. C* **58** (2008) 639–707, [arXiv:0803.0883].

[124] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016), no. 4 196, [arXiv:1512.01178].

[125] J. Baglio et al., *VBFNLO: A parton level Monte Carlo for processes with electroweak bosons – Manual for Version 2.7.0*, 2011.

[126] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J. C* **75** (2015), no. 5 204, [arXiv:1412.3989].

[127] S. Frixione, G. Ridolfi, and P. Nason, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, *JHEP* **09** (2007) 126, [arXiv:0707.3088].

[128] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, [hep-ph/0409146].

[129] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070, [arXiv:0709.2092].

[130] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, [arXiv:1002.2581].

[131] ATLAS Collaboration, "Studies on top-quark Monte Carlo modelling for Top2016." ATL-PHYS-PUB-2016-020, 2016.

[132] M. Beneke, P. Falgari, S. Klein, and C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, *Nucl. Phys. B* **855** (2012) 695–741, [arXiv:1109.1536].

[133] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason, *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, *Phys. Lett. B* **710** (2012) 612–622, [arXiv:1111.5869].

[134] P. Bärnreuther, M. Czakon, and A. Mitov, *Percent level precision physics at the Tevatron: First genuine NNLO QCD corrections to $q\bar{q} \to t\bar{t} + X$*, *Phys. Rev. Lett.* **109** (2012) 132001, [arXiv:1204.5201].

[135] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, *JHEP* **12** (2012) 054, [arXiv:1207.0236].

[136] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, *JHEP* **01** (2013) 080, [arXiv:1210.6832].

[137] M. Czakon, P. Fiedler, and A. Mitov, *Total top-quark-pair-production cross section at hadron colliders through $O(\alpha_S^4)$*, *Phys. Rev. Lett.* **110** (2013) 252004, [arXiv:1303.6254].

[138] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 2930, [arXiv:1112.5675].

[139] S. Frixione, E. Laenen, P. Motylinski, C. White, and B. R. Webber, *Single-top hadroproduction in association with a W boson*, *JHEP* **07** (2008) 029, [arXiv:0805.3067].

[140] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-*, *Phys. Rev. D* **82** (2010) 054018, [arXiv:1005.4451].

[141] N. Kidonakis, *Top Quark Production*, in *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013)*, pp. 139–168. [arXiv:1311.0283].

[142] P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mlbitz, P. Rieck, and P. Uwer, HatHor *for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, *Comput. Phys. Commun.* **191** (2015) 74–89, [arXiv:1406.4403].

[143] 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** (2011) 1034–1046, [arXiv:1007.1327].

[144] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, *JINST* **15** (2020) P04003, [arXiv:1911.04632].

[145] ATLAS Collaboration, "Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV." ATL-PHYS-PUB-2015-026, 2015.

[146] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, *Eur. Phys. J. C* **80** (2020) 47, [arXiv:1909.00761].

[147] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, *JINST* **15** (2020) P09015, [arXiv:2004.13447].

[148] ATLAS Collaboration, *The ATLAS inner detector trigger performance in pp collisions at 13 TeV during LHC Run 2*, *Eur. Phys. J. C* **82** (2022), no. 3 206, [arXiv:2107.02485].

[149] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, *JINST* **14** (2019) P12006, [arXiv:1908.00005].

[150] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **76** (2016) 292, [arXiv:1603.05598].

[151] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **81** (2021) 578, [arXiv:2012.00578].

[152] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, *Eur. Phys. J. C* **77** (2017) 466, [arXiv:1703.10485].

[153] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063, [arXiv:0802.1189].

[154] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, *Eur. Phys. J. C* **72** (2012) 1896, [arXiv:1111.6097].

[155] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **81** (2020) 689, [arXiv:2007.02645].

[156] G. D'Agostini, *A multidimensional unfolding method based on bayes' theorem*, *Nucl. Instrum. Meth. A* **362** (1995), no. 2 487 – 498.

[157] T. Adye, *Unfolding algorithms and tests using RooUnfold*, in *Proceedings, 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding (PHYSTAT 2011)*, pp. 313–318. [arXiv:1105.1160].

[158] ATLAS Collaboration, *Measurements of top-quark pair differential cross-sections in the $e\mu$ channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, *Eur. Phys. J. C* **77** (2017) 292, [arXiv:1612.05220].

[159] CMS Collaboration, *Measurements of $t\bar{t}$ differential cross sections in proton–proton collisions at $\sqrt{s} = 13$ TeV using events containing two leptons*, *JHEP* **02** (2019) 149, [arXiv:1811.06625].

[160] "Interference studies git package by Alexandre Laurier." https://gitlab.cern.ch/alaurier/interference-studies.

[161] ATLAS Collaboration, *Measurement of ZZ production in the $\ell\ell\nu\nu$ final state with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **10** (2019) 127, [arXiv:1905.07163].

[162] ATLAS Collaboration, *Measurement of fiducial and differential W^+W^- production cross-sections at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **79** (2019) 884, [arXiv:1905.04242].

[163] ATLAS Collaboration, *Measurement of $W^\pm Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **79** (2019) 535, [arXiv:1902.05759].

[164] ATLAS Collaboration, *$ZZ \to \ell^+\ell^-\ell'^+\ell'^-$ cross-section measurements and search for anomalous triple gauge couplings in 13 TeV pp collisions with the ATLAS detector*, *Phys. Rev. D* **97** (2018) 032005, [arXiv:1709.07703].

[165] ATLAS Collaboration, *Measurement of $WW/WZ \to \ell\nu qq'$ production with the hadronically decaying boson reconstructed as one or two jets in pp collisions at $\sqrt{s} = 8$ TeV with ATLAS, and constraints on anomalous gauge couplings*, *Eur. Phys. J. C* **77** (2017) 563, [arXiv:1706.01702].

[166] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 639, [arXiv:1902.04655].

[167] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **96** (2017) 072002, [arXiv:1703.09665].

[168] ATLAS Collaboration, *Jet energy resolution in proton–proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, *Eur. Phys. J. C* **73** (2013) 2306, [arXiv:1210.6210].

[169] J. Butterworth et al., *PDF4LHC recommendations for LHC run 2*, *J. Phys. G* **43** (2016) 023001, [arXiv:1510.03865].

[170] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev. D* **93** (2016), no. 3 033006, [arXiv:1506.07443].

[171] L. Lyons, D. Gibaut, and P. Clifford, *How to combine rorrelated estimates of a single physical quantity*, *Nucl. Instrum. Meth. A* **270** (1988) 110.

[172] E. Gerwick, T. Plehn, S. Schumann, and P. Schichtel, *Scaling Patterns for QCD Jets*, *JHEP* **10** (2012) 162, [arXiv:1208.3676].

[173] E. Maguire, L. Heinrich, and G. Watt, *HEPData: a repository for high energy physics data*, *J. Phys. Conf. Ser.* **898** (2017), no. 10 102006, [arXiv:1704.05473].

[174] ATLAS Collaboration, "Cross-section measurements for the production of a Z boson in association with high-transverse-momentum jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector (Version 1)." HEPData (collection), 2022. https://doi.org/10.17182/hepdata.114865.v1.

[175] C. Bierlich et al., *Robust Independent Validation of Experiment and Theory: Rivet version 3*, *SciPost Phys.* **8** (2020) 026, [arXiv:1912.05451].