ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

Dottorato di ricerca
in Fisica, Ciclo XXX

Settore Concorsuale: 02/A1
Settore Scientifico disciplinare: FIS/01

Inclusive Z boson production and in association with b-jets in proton-proton collisions at 13 TeV with the ATLAS experiment

Presentata da: Camilla Vittori

Coordinatore Dottorato:
Prof.ssa Silvia Arcelli

Supervisore:
Prof.ssa Laura Fabbri

Relatore:
Dott. Benedetto Giacobbe

Esame finale anno 2018

Dedicated to Antonella, Stefano, Luca, and Francesco.

Itaca

Quando ti metterai in viaggio per Itaca
devi augurarti che la strada sia lunga,
fertile in avventure e in esperienze.
Che i mattini d'estate siano tanti
quando nei porti - finalmente e con che gioia -
toccherai terra tu per la prima volta:
negli empori fenici indugia e acquista
madreperle coralli ebano e ambre
tutta merce fina, anche profumi
penetranti d'ogni sorta.

Sempre devi avere in mente Itaca -
raggiungerla sia il pensiero costante.
Soprattutto, non affrettare il viaggio;
fa che duri a lungo, per anni, e che da vecchio
metta piede sull'isola, tu, ricco
dei tesori accumulati per strada
senza aspettarti ricchezze da Itaca.
Itaca ti ha dato il bel viaggio,
senza di lei mai ti saresti messo
sulla strada: che cos'altro ti aspetti?

Costantino Kavafis

Abstract

In this thesis, the measurement of the inclusive Z-boson production cross section at the centre of mass energy of $\sqrt{s} =$13 TeV is first presented. The data analysed were collected by ATLAS during early Run 2 of LHC and correspond to a luminosity of 81 pb^{-1}. The cross section measurement is performed separately in the electron and muon decay channels and in their combination. The ratio with respect to the W cross section is also performed in the same fiducial phase space, and it benefits from cancellation of some experimental uncertainties. Results are compared to NNLO calculations in perturbative QCD.

The Z boson is then studied in associated production with at least one and at least two b-jets, using an enlarged dataset (36 fb^{-1}) collected by ATLAS at the centre of mass energy of $\sqrt{s} =$13 TeV. Integrated cross sections are measured at the particle level for both the electron and muon decay channels. Differential cross sections are measured as a function of the leading b-jet transverse momentum and rapidity, the Z boson rapidity and the angular separation between the Z boson and the b-jet. For events containing two b-jets at particle level, the cross section is measured as a function of the invariant mass and angular separation of the two highest transverse momentum b-jets, and as a function of the kinematics of the sub-leading b-jet. Results are compared to LO and NLO Monte Carlo predictions.

Finally, I present the work performed on the LUCID detector, the official ATLAS luminometer, to analyse calibration data with a ^{207}Bi source, and the activity within the ATLAS muon performance group, to evaluate the muon reconstruction scale factors, using data acquired in Run 2.

Abstract

In questa tesi, è presentata la misura della sezione d'urto di produzione inclusiva del bosone Z all'energia del centro di massa di $\sqrt{s} =$13 TeV. I dati analizzati sono stati raccolti da ATLAS all'inizio del Run 2 di LHC e corrispondono ad una luminosità di 81 pb^{-1}. La sezione d'urto è misurata separatemente nei canali di decadimento elettronico e muonico e successivamente nella loro combinazione. Il rapporto tra le sezioni d'urto dei bosoni Z e W è misurato nello stesso volume fiduciale e benificia della cancellazione di alcune incertezze sperimentali. I risultati sono confrontati con calcoli teorici disponibili al NNLO in QCD perturbativa.

Inoltre, è presentata anche la misura della produzione del bosone Z in associazione ad almeno uno e almeno due b-jets, usando dati raccolti da ATLAS all'energia del centro di massa di $\sqrt{s} =$13 TeV. Le sezioni d'urto integrate sono misurate al particle level per i canali elettronico e muonico. Le sezioni d'urto differenziali sono misurate in funzione del momento trasverso e della rapidità del b-jet più energetico, della rapidità del bosone Z e di una serie di variabili angolari tra il bosone Z e il b-jet. Per gli eventi con due b-jets al particle level, la sezione d'urto è misurata in funzione della massa invariante e delle separazioni angolari tra i due b-jets più energetici e delle variabili cinematiche del secondo b-jet più energetico. I risultati sono confrontati con predizioni Monte Carlo al LO e NLO.

Infine, sono presentati i lavori dedicati all'analisi dei dati di calibrazione del rivelatore LUCID, il luminometro ufficiale di ATLAS, e alla determinazione delle efficienze di ricostruzione dei muoni, all'interno del gruppo di performance, usando i dati acquisiti nel Run 2.

Contents

Introduction

The Large Hadron Collider (LHC) operating at CERN offers the opportunity to study the elementary particles and their interactions at the unprecedented centre of mass energy of $\sqrt{s}$=13 TeV. The wide physics program of LHC is devoted to the discovery of new phenomena and to several precision measurements.

The first analysis I present in the thesis is the measurement of the Z boson cross section, using data collected by the ATLAS detector at the very beginning of Run 2, corresponding to an integrated luminosity L=81 pb^{-1}. Although the Z boson has been discovered in 1983, it is still object of research since the clear signature of its leptonic decays and the high statistics reachable at LHC allow to perform high precision measurements, needed for tests of perturbative QCD and for detector calibration purposes, crucial at the beginning of any new data taking period. The measurement is performed separately in the electron and muon decay channels and then combined. The ratio of the W boson cross section, performed using the same dataset, to the Z boson one, is also evaluated, since it benefits from the partial or total cancellation of several uncertainties. The high precision reached in the ratio allows to test models of perturbative QCD and to constrain proton Parton Distribution Functions (PDFs).

The main subject of this thesis is the measurement of the Z-boson produced in association with b-jets (Z+$\geq$1 b-jet and Z+$\geq$2 b-jets), by using the data collected by ATLAS at the centre of mass energy of $\sqrt{s}$= 13 TeV, corresponding to a luminosity of 36 fb^{-1}. This measurement offers the opportunity to improve the understanding of the main irreducible background to Higgs boson production in association with a vector boson, with the Higgs decaying into a $b\bar{b}$ pair.

This study is of great interest, since, from the theory side, two methods are employed for the generation of these processes, called 4 Flavour Number Scheme (4FNS) and 5 Flavour Number Scheme (5FNS). Since the two calculations differ in including the b-quark in the initial state, a precise measurement of the Z+≥1 b-jet and Z+≥2 b-jets cross sections may in principle help to constrain the b-quark PDF. In this thesis the integrated and differential cross sections of the Z boson produced with b-jets are measured at particle level, the latter as a function of several physics observables, important for the PDF understanding and for testing modelling of different event generators.

The inclusive Z boson and its production in association with b-jets have been measured in the past by the LHC and Tevatron experiments and results have been used as inputs for new PDF sets. The most recent results on these processes are summarised in Chapter 1, together with a description of the Standard Model of particle physics, which governs the well known properties of the Z boson. The ATLAS experiment is introduced in Chapter 2 together with the definitions and reconstruction procedures used to define the physics objects needed in the analysis.

In Chapter 3 the measurement of the inclusive production cross section of the Z boson is presented. The high precision reachable even with low statistics allows to compare the measurement with the most recent theoretical calculations and to constrain proton PDFs.

In Chapter 4 the Z+1 b-jet and Z+2 b-jets selections, based on same flavour and opposite charge leptons (electrons or muons) and one or two b-jets, is described. Among the different background contributions, the estimation of the Z+c-jets and Z+light-jets processes is a key step of the analysis, implemented through a fit on data of the distribution of a b-jet sensitive variable.

In Chapter 5, events passing the reconstruction level selection are corrected for detector distortions and inefficiencies, through the unfolding procedure, which allows to extrapolate measurements to the particle level. The integrated and differential cross sections are presented in Chapter 6, with a detailed discussion on the systematic uncertainties. The particle level measurement are compared to several Monte Carlo predictions, with different flavour schemes for the modelling of Z+1 b-jet and Z+2 b-jets processes and

different precision in the calculations.

My contribution in the presented analyses is the following:

- in the inclusive Z cross section measurement, I personally performed the full study in the muon channel, including the event selection, the background estimation, the efficiency determination and the evaluation of the systematic uncertainties;

- in the Z production cross section measurement in association with b-jets, I contributed to the full analysis chain, both in the muon and electron channels, focusing in particular on the event selection, the background evaluation through the flavour fit, the unfolding procedure and the assessment of the systematic uncertainties, down to the final result.

During the long shut-down between the Run 1 and Run 2 of LHC, many sub-detectors have been redesigned in order to cope for the new conditions at the increased centre of mass energy. In particular, the sensors of LUCID, the ATLAS official luminosity monitor, have been replaced and equipped with a radioactive Bismuth-207 source for a continuous monitoring of their gain-stability. As a technical activity, I directly participated to the choice and the equalisation of LUCID photomultipliers, and to the first measurements using the calibration source, as detailed in Appendix A.

Finally I participated to studies of the ATLAS muon reconstruction performances aimed to the evaluation of the data-driven correction factors ("Scale Factors"), needed to tune Monte Carlo simulations. This work is detailed in Appendix B.

Chapter 1

The Z boson physics

The current description of the elementary particle interactions is based on the Standard Model (SM) of particle physics. Developed in the early 1970s, it provides a description of the electromagnetic, weak and strong fundamental forces. The Standard Model has been tested and validated over a few decades by various experiments. Nevertheless, at the beginning of the 21st century, a wide series of questions are still open:

- the fermion mass hierarchy and the quark flavour mixing in the SM have a very characteristic structure which could originate from unknown physics at not yet explored very high energy scales;

- the observed neutrino flavour oscillation cannot be explained within the SM, as it implies lepton-number violation;

- the nature of the Dark Matter (Dark Energy) which corresponds to $\sim27\%$ ($\sim68\%$) of the observed content of the Universe is still not known;

- the theory of gravity cannot be integrated in the SM description, as no renormalizable quantum field theory of gravity exists.

In the last decades of the 20th century, the Large Hadron Collider (LHC), a proton-proton collider of unprecedented energy, was built to address these open questions. The ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) experiments were designed as general purpose detectors in order to be able to explore the largest variety of experimental signatures.

ATLAS and CMS high detection efficiencies allowed for the observation of a new particle with a mass of 125 GeV (in unit $c=1$), compatible with the SM Higgs boson.

A description of the Standard Model is presented in Section 1.1. At hadron colliders, the calculation of the cross section of the processes of interest is challenging and it is described in Section 1.2. In Section 1.3 the discovery and the main features of the Z boson are illustrated. The Z production mechanism and the theoretical importance of the measurement are explained in Section 1.4. In Sections 1.5 and 1.6 the Z boson production in association with jets and b-jets is discussed. In particular, in Section 1.6 the Z+b-jets theoretical production mechanism and the sensitivity to the b-quark distribution in the colliding protons are addressed, together with a description of the main physics motivations behind the measurements.

1.1   The Standard Model of particle physics

The Standard Model of particle physics is a gauge theory which describes the fundamental particles in nature and their interactions. It is based on the existence of local gauge symmetries that explains the interactions of particles by means of force-mediating fields. The SM Lagrangian reflects the symmetries in the invariance under the transformation of the $\mathrm{SU(3)}_C \times \mathrm{SU(2)}_L \times \mathrm{U(1)}_Y$ gauge group [1]. Each part of the SM Lagrangian represents one of the fundamental forces in nature in terms of quantum fields: the electromagnetic ($\mathrm{U(1)}_Y$), the weak ($\mathrm{SU(2)}_L$) and the strong ($\mathrm{SU(3)}_C$) interactions, which will be addressed in Section 1.1.2, with a description of the spontaneous symmetry breaking mechanism, responsible for the origin of matter.

1.1.1   The fundamental particles

The basic constituents of matter are represented by fermionic fields of spin 1/2-particles, which obey to the Fermi-Dirac statistics and satisfy the Pauli exclusion principle. The symmetry principles postulated in the Standard Model imply the existence of conserved quantities (Noether's theorem), described by quantum numbers [1]. The quantum number associated to $\mathrm{SU(2)}_L$

and U(1)$_Y$ gauge groups are called isospin I and hypercharge Y, respectively. The relation with the electric charge is given by $Q = I_3 + \frac{Y}{2}$, where I_3 is the third component of the weak isospin. The chiral symmetry of the SU(2)$_L$ group leads the associated gauge boson to couple only to the left-handed fermion fields. Fermion fields are arranged into left-handed doublets, carrying a weak isospin $I = 1/2$ and right-handed singlets with $I = 0$.

Fermions are classified in leptons and quarks. Leptons experience the weak and the electromagnetic forces, but not the strong interaction and are organised in three families (isospin doublets):

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

The three families contain massive particles (e, μ and τ) with charge - 1, interacting through both the electromagnetic and weak forces and their corresponding neutrinos, which interact only weakly. Within the Standard Model, neutrinos are neutral massless leptons, in contrast with the experimental evidence of their oscillation, which requires a mass different from zero. The third component of the weak isospin takes the values $+1/2$ for charged leptons and $-1/2$ for neutrinos. Leptons are described by the *leptonic* quantum number L, conserved by all the interactions and each weak doublet has a representative leptonic number: L_e, L_μ and L_τ [2].

Quarks interact through the electromagnetic, weak and strong forces and occur in six different *flavours* labelled *u, d, c, s, t, b*. The *flavour* is conserved by all the forces, with the exception of the weak interaction [2]. Because of the similarity between u- and d-quark masses, these two quarks are grouped in a strong isospin doublet ($I = 1/2$ and $I_3 = \pm 1/2$). Similarly, all quarks are organised in three weak isospin doublets:

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

where the u-like quarks carry $+2/3$ and the d-like quarks $-1/3$ electric charge, respectively. All quarks are described by the *barionic* quantum number B ($+1/3$ for quarks and -1/3 for anti-quarks), conserved by all interactions. The quantum number associated to the SU(3)$_C$ group is called *colour*,

which represents the "charge" of the strong interaction and can assume three possible values, conventionally called *red*, *blue* and *green*. Affected by confinement, quarks do not exist as free particles, but only constrained in barions (mesons with B=0, or hadrons with B=1), which are colour-neutral particles.
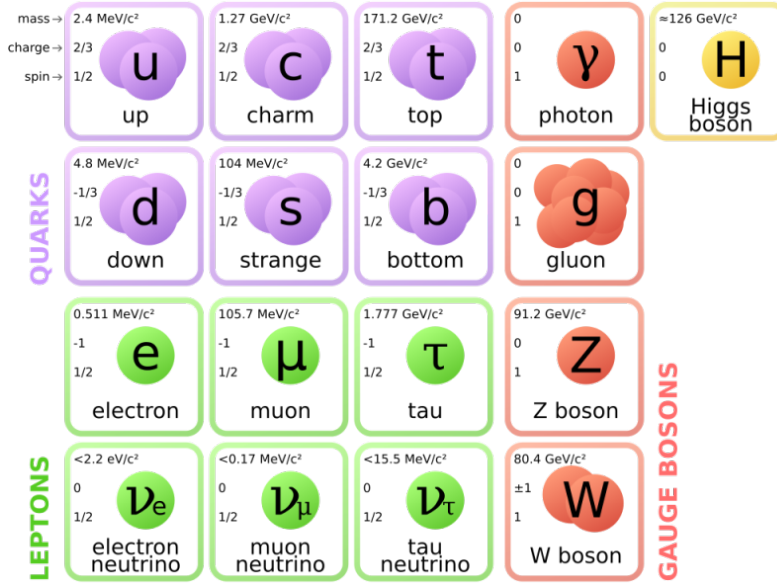
Figure 1.1: The fundamental particles of the Standard Model. The Higgs boson, responsible for the electroweak symmetry breaking, is also shown. For each particle, the mass, charge and spin are reported [3].

1.1.2 The fundamental forces

In the Standard Model, particles interact with each other by coupling with specific fields, whose quanta are spin-1 particles called *bosons*. Bosons emerge from the requirement of local gauge invariance under specific transformations (symmetries) of the fermionic fields. According to this formalism, the structure of the different interaction contributions can be described starting from the free Dirac Lagrangian:

$$\mathcal{L}_{free} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi \tag{1.1}$$

where ψ represents the fermion field with mass m. The global invariance of the Lagrangian of Eq.1.1 under specific transformations implies the conservations of the electromagnetic, weak and strong charges, as ensured by the Noether's theorem. The local gauge invariance allows particles to experience interactions, introducing bosons as necessary components in the covariant derivative. The intensity of the interactions is described by coupling constants which enter in the matrix element of each process.

The electromagnetic interaction

The Quantum Electrodynamics (QED) is the theory which describes the electromagnetic interaction, symmetric with respect to gauge rotation of $U(1)_Y$ group. The electric charge is the responsible for the coupling of charged particles with the field. The QED coupling constant is a dimensionless quantity defined in terms of the electric charge e: $\alpha_e = e^2/4\pi\epsilon_0\hbar c \cong 1/137$. The running behaviour of α_e as a function of the energy involved in the process is explained by the *vacuum polarisation* effect [1]. According to this phenomenon, more energy is needed to probe the real value of a particle charge surrounded by the vacuum medium.
The QED Lagrangian is obtained by requiring the invariance of the Lagrangian of Eq.1.1 under $U(1)_Y$ transformations [1]. Therefore, the partial derivative is substituted by the covariant derivative

$$\partial_\mu \to \mathcal{D}_\mu = \partial_\mu + ieA_\mu(x) \tag{1.2}$$

where the quantum of vector field A_μ is the photon. The resulting QED Lagrangian is:

$$\mathcal{L}_{QED} = \mathcal{L}_{free} - \underbrace{J^\mu A_\mu}_{\text{interaction term}} - \underbrace{\frac{1}{4}F^{\mu\nu}F_{\mu\nu}}_{\text{kinetic term of } A_\mu} \tag{1.3}$$

where J^μ is the electromagnetic current and the last term represents the propagations of free photons, with the Maxwell's electromagnetic tensor $F^{\mu\nu}$. A term with the form $\mathcal{L} = \frac{1}{2}m^2A^\mu A_\mu$ representing the photon mass is not

present in Eq.1.3 as it is forbidden by the gauge invariance: this implies the massless nature of the photon.

The strong interaction

Quantum Chromodynamics (QCD) is the quantum field theory of the strong interactions, symmetric with respect to gauge rotation of $SU(3)_C$ group. The QCD coupling constant α_s has a running behaviour as a function of the energy. In the vacuum, the colour charge of a bare quark is surrounded by quark-antiquark pairs and by coloured gluons, resulting in an enhancement of the strong charge as a function of the distance ("anti-screening" effect). The divergence of the strong coupling at low energy is an indication of *confinement*, the mechanism that explains why the free quark and the gluon degrees of freedom have never been observed. On the other hand, when the transferred momentum is large and quarks are close, the strong interaction is feeble: this property is called *asymphtotic freedom*.

In order to determine the QCD Lagrangian, it is important to consider that each quark field q_a of Eq.1.1 can occur in three colours ($q_a = (q_{a,r}, q_{a,b}, q_{a,g})$). The covariant derivative introduced to guarantee the invariance under $SU(3)_C$ rotations is

$$\mathcal{D}_\mu = \partial_\mu + igt_A \cdot \mathcal{A}_\mu^A(x) \tag{1.4}$$

where $\mathcal{A}_\mu^A$ is the gauge field of the strong interaction, the gluon, and the t_A matrix is a fundamental representation of SU(3). The field strength tensor $G_{\mu\nu}^A$ is expressed in terms of:

$$G_{\mu\nu}^A = [\partial_\mu \mathcal{A}_\nu^A - \partial_\nu \mathcal{A}_\mu^A - gf_{ABC}\mathcal{A}_\mu^B \mathcal{A}_\nu^C] \tag{1.5}$$

and indices A, B and C run over the eight colour degrees of freedom of the gluon field. The third term in Eq.1.5 is a typical feature of a non-abelian theory and generates triplet and quartic gluon self-interactions. The QCD Lagrangian is

$$\mathcal{L}_{QCD} = \sum_{flavours} \bar{q}_a(i\gamma_\mu \mathcal{D}^\mu - m)_{ab}q_b - \frac{1}{4}G_{\mu\nu}^A G_A^{\mu\nu}. \tag{1.6}$$

Similarly to the $\mathcal{L}_{QED}$, $\mathcal{L}_{QCD}$ does not contain quadratic-field terms which should represent the quanta mass: as a consequence, gluons are massless bosons in the SM [1].

The weak interaction

The weak interaction involves all the fundamental particles of the Standard Model. First observed in the β decays, it was originally explained by Fermi as an effective point-like theory. Although valid at low transferred momentum, this approximation does not consider important features of this interaction, like the massive mediators and the parity violation. A quantum field theory based on a V-A (vector-axial) structure and described by the $SU(2)_L$ symmetry group has been introduced, where the label L indicates the coupling with left-handed fermions only. Fermion fields are represented by the left-handed doublets L and the right-handed singlet R, which, for the first family of fermion, becomes:

$$L = \begin{pmatrix} e \\ \nu_e \end{pmatrix}_L \qquad R = e_R$$
$$L = \begin{pmatrix} u \\ d \end{pmatrix}_L \qquad R = u_R, d_R$$

The mediators of the weak interaction are three massive bosons: W^+ and W^- that carry electric charge ($\pm$ 1) and Z which is neutral. These particle were initially predicted without mass, in contrast with experimental evidence of $M_{Z/W} >> 0$. The huge mass of these bosons explains the short range of the weak force.
Weak interactions among quarks coming from different isospin doublets are disfavoured, but not forbidden. The quark mixing is explained considering the weak eigenstates as a rotation of the mass eigenstates through the CKM matrix [2].

The electroweak theory

In the late 1960s, Weinberg, Salam and Glashow unified the electromagnetic and weak interactions in the electroweak theory, described by the

SU(2)$_L$×U(1)$_Y$ symmetry group [4, 5, 6]. Following the same procedure adopted for the QED and QCD theories, the Lagrangian of the left-handed L and right-handed R fermion fields is required to be invariant under global and local transformation of the gauge group. In order to ensure the local invariance, the ∂_μ derivative is replaced by the covariant form:

$$\mathcal{D}_\mu = \partial_\mu + ig\frac{\tau_a}{2}W_\mu^a + i\frac{g'Y}{2}B_\mu \tag{1.7}$$

where W_μ^a and B_μ are the gauge bosons of the SU(2)$_L$ and U(1)$_Y$ groups respectively. The Pauli matrices τ_a (a=1,2,3) and the hypercharge Y represent the generators of such groups. The electroweak Lagrangian can be expressed as

$$\begin{aligned}
\mathcal{L}_{EW} = &-\frac{1}{4}\mathbf{W}_{\mu\nu}\cdot\mathbf{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}\cdot B^{\mu\nu} \\
&+ \bar{L}\gamma^\mu(i\partial_\mu - g\frac{1}{2}\tau\cdot\mathbf{W}_\mu - g'\frac{Y}{2}B_\mu)L + \bar{R}\gamma^\mu(i\partial_\mu - g'\frac{Y}{2}B_\mu)R
\end{aligned} \tag{1.8}$$

where the first line contains the kinetic energy and the self coupling of the $\mathbf{W}_\mu$ fields and the kinetic energy of the B_μ field. In the second line, the fermion kinetic energy $(i\partial_\mu)$ and the fermion interactions with the $W_\mu^{1,2,3}$ $(g\frac{1}{2}\tau\cdot\mathbf{W}_\mu)$ and B_μ $(g'\frac{Y}{2}B_\mu)$ are described. The Lagrangian of Eq.1.8 describes the interaction among massless fermions and massless gauge fields, in contradiction with experimental observations.

The mass terms are generated introducing a SU(2)$_L$ doublet of complex scalar field ϕ, described by the Lagrangian

$$\mathcal{L}_{Higgs} = (\partial_\mu\phi)^\dagger(\partial^\mu\phi) - V(\phi) = (\partial_\mu\phi)^\dagger(\partial^\mu\phi) - \frac{1}{2}\mu^2\phi^\dagger\phi - \frac{1}{4}\lambda(\phi^\dagger\phi)^2 \tag{1.9}$$

where μ is the mass of the scalar field ϕ and λ is a positive dimensionless constant. If $\mu^2 > 0$, the potential exhibits the ground state on the origin ($\phi_{min} = 0$). If $\mu^2 < 0$, the ground state is given for values $\phi_{min} = \pm\sqrt{\frac{-\mu^2}{\lambda}} = \pm v$, called *vacuum expectation value*. Without loss of generality, it is possible to choose $\phi = v = \sqrt{\frac{-\mu^2}{\lambda}}$ and, by performing a perturbation expansion around one of the minima of the potential, in the Lagrangian of Eq.1.9 a real mass term of the form $-\lambda v^2$ appears. The perturbation expansion

must be made around one of the two minima, whose choice determines the **spontaneous symmetry breaking**. The $SU(2)_L \times U(1)_Y$ gauge invariant expression of Eq.1.9 obtained using the covariant derivative is

$$\mathcal{L}_{Higgs} = |(i\partial_\mu - g\frac{1}{2}\tau \cdot \mathbf{W}_\mu - g'\frac{Y}{2}B_\mu)\phi|^2 - V(\phi) \tag{1.10}$$

Choosing the minimum $\phi_{min} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}$, the gauge fields of the previous expression acquire mass and can be expressed as a linear combination of the $SU(2)_L \times U(1)_Y$ fields

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \\ Z_\mu &= \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \\ A_\mu &= \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \end{aligned} \tag{1.11}$$

where θ_W is the Weinberg mixing angle. The $W_\mu^\pm$, Z_μ and A_μ fields can be identified with the $W^\pm$ bosons of mass $m_W = \frac{1}{2}vg$, the Z boson of mass $m_Z = m_W/cos\theta_W$ and the massless photon field. The potential $V(\phi)$ introduces further terms in the Lagrangian of Eq.1.10 which are shown here for a parametrisation of the field ϕ around the ground state:

$$\begin{aligned} \mathcal{L} = &\frac{1}{2}(\partial_\mu h)^2 + [m_W^2 W_\mu^+ W^{-\mu} + m_Z^2 Z_\mu Z^\mu](1 + \frac{h}{v})^2 \\ &- \lambda v^2 h^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4 \end{aligned} \tag{1.12}$$

using $\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+h(x) \end{pmatrix}$. The Lagrangian contains a mass term for the $h(x)$ field (with $m_h = \sqrt{2\lambda v^2}$) which describes a scalar field particle referred to as the **Higgs boson**. The interaction of the gauge bosons $W^\pm$ and Z with the Higgs field explains the generation of their masses.

Another $SU(2)_L \times U(1)_Y$ gauge invariant term is introduced to describe the mass of the fermion fields, generated by coupling with the Higgs field:

$$\mathcal{L} = -(G_1 \bar{L}\phi R + G_2 \bar{L}\phi^* R + \text{hermitian conjugate}) \tag{1.13}$$

with $G_{1,2}$ are the matrices of Yukawa couplings [1].

1.2 Proton-proton interactions at LHC

The calculation of the cross section of physics processes is challenging at hadron colliders such as LHC, because of the complex interplay between perturbative theoretical computations, phenomenological models and use of the experimental data. The calculation of the process at LHC must consider the interacting protons not as fundamental particles, but as compound of quarks and gluons (partons). As a result, the cross section for a process generated by the interaction of two partons with momenta p_1 and p_2 can be expressed as [7]:

$$\sigma(p_1, p_2) = \sum_{i,j=q,\bar{q},g} \int_0^1 dx_1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \hat{\sigma}_{ij} \left(x_1 p_1, x_2 p_2, \alpha_S(\mu_R^2), \frac{Q^2}{\mu_R^2}, \frac{Q^2}{\mu_F^2} \right)$$
(1.14)

where x_1 and x_2 are the proton momentum fractions carried by partons 1 and 2, respectively. $f_i(x_1, \mu_F^2)$ ($f_j(x_2, \mu_F^2)$) is called Parton Distribution Fuction (PDF) and represents the probability for the parton i (j) to carry a fraction x of the proton 1 (2) momentum. PDFs are specific of the type of interacting partons and depend on the QCD factorisation scale μ_F. $\hat{\sigma}_{ij}$ is the interaction cross section between the two partons involved in the interaction: it depends on the strong coupling constant (α_S), whose value is given for a specific energy scale (the renormalisation scale μ_R) and on the transferred momentum Q^2. The sum runs over all contributing parton configurations: i and j represent all the possible combinations of quarks, antiquarks and gluons.

Experimentally, the production cross section determines the number of observed events $N(X)$ for a given process $pp \to X$ with a specific detection efficiency ϵ:

$$N(X) = \epsilon \cdot \sigma(X) \int \mathcal{L} dt$$
(1.15)

where $\int \mathcal{L} dt$ is the integrated luminosity acquired during the measurement (see Appendix A for a detailed description of the luminosity definition and its determination at LHC).

The main ingredients of Eq.1.14 for the cross section determination will be discussed in the following sections. In particular, the strong coupling con-

stant will be described in Section 1.2.1, the partonic cross section $\hat{\sigma}$ will be explained in Section 1.2.2 and the PDF role will be discussed in Section 1.2.3. Finally, the implementation of the cross section calculation in Monte Carlo generators will be described in Section 1.2.4.

1.2.1 The running coupling constant

In a perturbative field theory, the cross section evaluation can be expressed in terms of the coupling constant and visualised through Feynman diagrams. These diagrams can contain loops of particles which result in divergences in the calculation. QCD and QED fulfil the renormalisability requirement of quantum field theory, which allows to absorb the divergences into renormalised definitions of the coupling constants and particle masses [8]. This procedure leads to the introduction of an arbitrary scale, the **renormalisation scale** (μ_R). As a consequence, the coupling constant (α) varies as a function of μ_R, characterizing its *running* behaviour. The generic equation describing the running coupling constant (α) can be formulated as [9]:

$$\alpha(Q^2) = \frac{\alpha(\mu^2)}{1 + \alpha(\mu^2)b_0 ln(\frac{Q^2}{\mu^2})} \tag{1.16}$$

In QED $b_0 = -1/3\pi$ and in QCD $b_0 = (33 - 2n_f)/12\pi$, where n_f is the number of quark flavours; the sign of b_0 determines the direction of the running behaviour (i.e. α increases or decreases) as a function of the energy involved in the process μ^2 and of the transferred momentum Q^2. Q^2 is given by the sum of the four vectors of the interacting parton. In case of the leading-order (LO) resonant scattering $q\bar{q} \to Z$, the energy scale is given by $Q^2 = sx_1x_2 = M_Z^2$, where s represents the centre of mass energy of the collision. Fig.1.2 shows the running behaviour of the strong coupling constant as a function of the energy measured by different experiments [3].

1.2.2 Perturbative QCD

Feynman diagrams help visualising the cross section calculation of a physics process, since each vertex contributes with a term proportional to the cou-
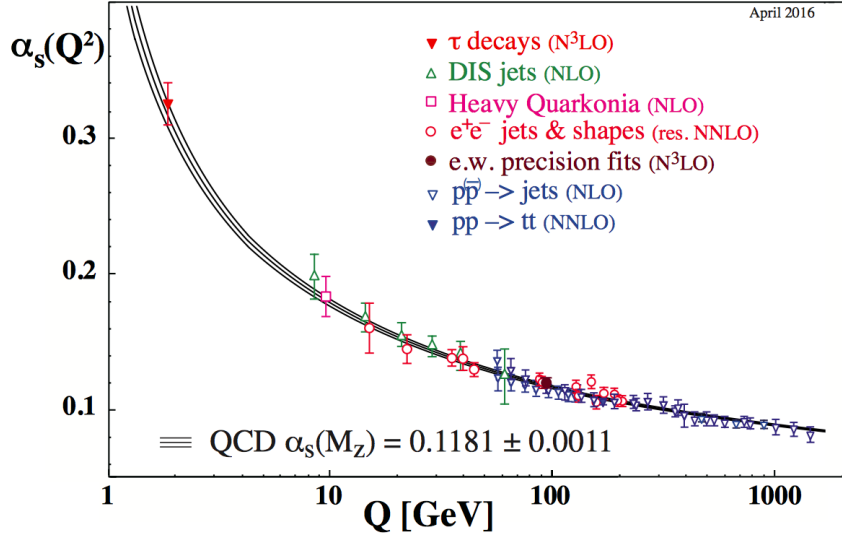
Figure 1.2: Summary of measurements of the strong coupling constant α_s as a function of the energy scale Q.

pling constant of the occurring force. In the specific case of the partonic strong interaction, each vertex amplitude is proportional to α_s and the total partonic cross section can be expressed as [7]:

$$\hat{\sigma} = \alpha_s^k \sum_{m=0}^{\infty} c^{(m)} \alpha_s^m \tag{1.17}$$

where $c^{(m)}$ is the m-th order contribution to the cross section and k represents the number of strong vertices present in the Feynman diagram of the process. Eq.1.17 defines a perturbative calculation which converges if $\alpha_s \ll 1$. The calculation is more and more precise adding higher order terms, but cannot be carried out infinitely, so it is truncated to a given order m, which is the order of the calculation. The cross section $\hat{\sigma}_{ij \to X}$ can be written in a perturbative expansion as:

$$\hat{\sigma}_{ij \to n} = \hat{\sigma}_{ij \to n}^{LO} + \alpha_s \hat{\sigma}_{ij \to n}^{NLO} + \alpha_s^2 \hat{\sigma}_{ij \to n}^{NNLO} + \alpha_s^3 \hat{\sigma}_{ij \to n}^{N^3LO} + \mathcal{O}(\alpha_s^4) \tag{1.18}$$

where the leading-order (LO) is the lowest order in the calculation, followed by the next-to-leading order (NLO), the next-to-next-to-leading-order (NNLO) and so on up to a certain perturbative order.

As anticipated in Section 1.2.1, divergences in the calculation lead to the introduction of the renormalisation scale μ_R. Physical observable, such as the cross section, cannot depend on the choice of this scale, however the truncation of the perturbative calculation produces a residual dependence. Mathematically, this is related to the size of the missing high order terms:

$$\frac{d}{d\mu_R}\hat{\sigma} = \alpha_s^k \frac{d}{d\mu_R}\sum_{m=0}^{\infty} c^{(m)\alpha_s^m} = 0 \tag{1.19}$$

Truncating at $m = n$

$$\frac{d}{d\mu_R}\sum_{m=0}^{m=n} c^{(m)\alpha_s^m} + \frac{d}{d\mu_R}\sum_{m=n+1}^{\infty} c^{(m)\alpha_s^m} = 0 \tag{1.20}$$

and rearranging in

$$\frac{d}{d\mu_R}\sum_{m=0}^{m=n} c^{(m)\alpha_s^m} = \mathcal{O}(\alpha_s^{n+1}) \tag{1.21}$$

due to the choice of μ_R, the variation in the partonic cross section (calculated to order n) gives an estimate of the size of the following term in the perturbative expansion.

1.2.3   Parton Distribution Functions

From Eq.1.14, another important element in the evaluation of the total cross section of a process are the Parton Distribution Functions (PDFs) $f_i(x_1, \mu_F^2)$ and $f_j(x_2, \mu_F^2)$, which represent the probability for a parton (i or j) to carry a fraction x (Bjorken variable) of the total proton momentum. $xf_i(x_1, \mu_F^2)$ and $xf_j(x_2, \mu_F^2)$ are the so-called Parton Density Function, which refer to the probability to find a parton with a momentun fraction between x and $x+dx$. PDFs depend on the **factorisation scale** (μ_F), an arbitrary chosen scale which defines whether a parton emission is treated within the PDF or the partonic cross section $\hat{\sigma}$ of Eq.1.14. A parton emitted with low trans-

verse momentum ($< \mu_F$) is absorbed in the PDF definition and treated as part of the structure of the proton; on the other hand, a parton emitted with high tranverse momentum ($> \mu_F$) is developed in the partonic cross section. Similarly to μ_R, the total cross section cannot depend on the factorisation scale if the perturbative calculation is carried out to all orders.
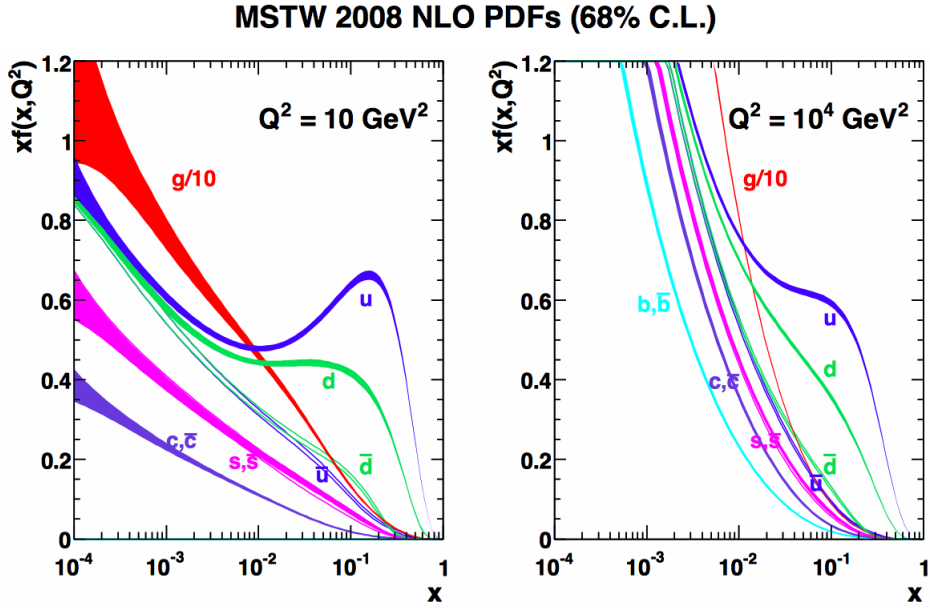
Figure 1.3: The MSTW 2008 NLO PDF set as a function of the Bjorken variable x at two different scales at $Q^2 = 10$ GeV^2 (left) and $Q^2 = 10^4$ GeV^2 (right) [10].

PDFs are non-perturbative objects which cannot be calculated directly and are derived from global fit to experimental data [10]. Generally, a physically motivated ansatz is chosen to parametrise the PDFs as a function of x at a given starting scale. The DGLAP equations allow to evolve the PDFs to arbitrary scales, dealing with parton splitting and emission [11]. The obtained PDFs are used to perform fits on the data, in order to constrain the initial parameters. Global fits are performed by groups such as MSTW, CTEQ and NNPDF using data coming from fixed-target deep inelastic scattering (HERA) and hadron collider (Tevatron and LHC) experiments. An example of a PDF set is shown in Fig.1.3 at two Q^2 scales for the MSTW 2008 NLO PDF. Different experimental data are sensitive to different parton

density functions and probe complementary ranges of the Bjorken variable x and of the partonic center-of-mass energy Q^2. Table 1.1 summarises the data used in the fit for four examples of PDF sets of interest for this thesis [12].

The statistical treatment of the experimental data is the dominant source of uncertainty in the PDF determination, followed by the strategy for the fits, the arbitrary choice of the parametrisation function at the initial scale and the α_s value, used in the process calculation.

	Fixed-target DIS	HERA DIS	Tevatron	LHC
x range	10^{-3}-1	10^{-6}-0.1	10^{-3}-1	10^{-6}-1
Q^2 range [GeV]	0.2-3$\cdot 10^2$	$2\cdot 10^{-2}$-4$\cdot 10^3$	$2\cdot 10^3$-2.5$\cdot 10^5$	10^2-2$\cdot 10^8$
CT14	yes	yes	yes	yes
MMHT14	yes	yes	yes	yes
NNPDF3.0	yes	yes	yes	yes
HERAPDF2.0	no	yes	no	no
ABM12	yes	yes	no	yes

Table 1.1: Brief summary of experimental data used in the global fit for the determination of the PDF sets CT14, MMHT14, NNPDF3.0, HERAPDF2.0 and ABM12. For each experiment, the x-Bjorken and Q^2 ranges are reported [12].

1.2.4 Monte Carlo generators

The calculation of Eq.1.14 provides a prediction of the production cross section of a specific process. However, the full calculation is complex and is dealt within Monte Carlo (MC) generators [13]. In p-p collisions, the interaction takes place among two partons, each belonging to a different proton. The mechanism can originate a resonance, which further decays transferring spin correlations to the decay products. As particles are accelerated, they can emit gluons or photons via bremsstrahlung: the emission associated to the incoming partons is called *Initial State Radiation* (ISR), while the emission associated to outgoing partons is referred to as *Final State Radiation* (FSR). As the proton is made up of a multitude of partons, several parton pairs may collide within one single p-p collision, originating the so called "Multiple

Interactions" (MI). The colliding partons take a fraction of the energy of the incoming hadrons, but much of the initial energy is carried by beam remnants, which continue to travel in the original directions. QCD confinement forces quarks to combine in colourless particles, in a process called *hadronisation*. Hadrons are unstable and decay at various timescales; some are sufficiently long-lived that their decays can be observed in a detector (i.e. B-hadrons).

Monte Carlo generators allow to consider these steps sequentially, and within each step, to define a set of rules that can be used iteratively to build more complex states. A picture of the various simulation stages is shown in Fig.1.4.
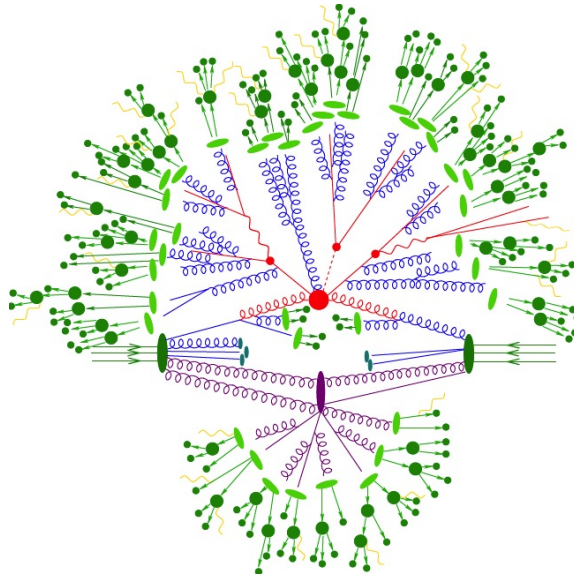
Figure 1.4: Sketch of p-p collision simulation. The hard part of the scattering (red) is simulated by the Matrix Element generators. Parton Showers produce bremsstrahlung (blue) and multiple interaction models simulate secondary interactions (purple). Fragmentation models describe the hadronisation of partons (green), whose decays are simulated by specific tools (dark green). [13]

Matrix Element and Parton Shower

The first element considered in the simulation of the entire process is the direct collision of the incoming partons, often referred to as "hard scattering". The hard scattering is described by Matrix Element (ME) calculations

at fixed order in α_s, starting from the parton momenta described by PDFs using Eq.1.14. ME expands the perturbative calculation in power of α_S to include high order corrections and final states with multiple partons. Given the divergence of the α_s coupling at low energy, perturbative QCD calculations are no more possible. This means that ME cannot be used to explore the internal structure of a jet and it is difficult to match to hadronisation models, which consider soft partons.

The Parton Shower (PS) method is introduced to simulate soft and collinear emissions. It allows to create a random number of gluon or $q\bar{q}$ pair emissions, because it employs kinematic, interference and helicity simplifications. The starting point is to factorise a complex $2 \to n$ process, where n represents a large number of partons in the final state, into a simple $2 \to 2$ process convoluted with showers. PS is based on a sequential stochastic branching to model the splitting and the emission of partons, extended until the parton energy scale reaches values ~ 1 GeV (typically associated to hadron creation). Since both ME and PS have complementary advantages and disadvantages, the combination of the two methods is forthcoming. The fundamental challenge of the combined ME+PS simulation is how to treat the overlapping of soft and collinear regions of phase space. At lowest order in α_s, this problem appears when additional quarks and gluons are included in the ME calculation. The general idea behind the several approaches available for merging [14] is to veto the PS emissions that overlap the ME phase space and apply event weights, based on the probability of the ordering of ME emissions ("Sudakov form factors") involving the splitting functions. In general, ME emissions are preferred to the PS ones, for their more accurate description of hard and angular separated radiations. At NLO, there is a phase space overlap between first real emission with the radiation from the PS.

Hadronisation

At the scale of ~ 1 GeV, the confinement of coloured partons in hadrons (*hadronisation*) is described by non-perturbative models. The combination of PS and hadronisation allows Monte Carlo generators to simulate the collimate showers of hadrons ("jets") observed in the detectors. In this context,

hadronisation refers to the specific model used by a Monte Carlo generator to describe the transition from the partonic to the final state. The two most recent models are based on a isotropic and longitudinal phase space: the "string" and the "cluster" models.

(a) The string model   (b) The cluster model
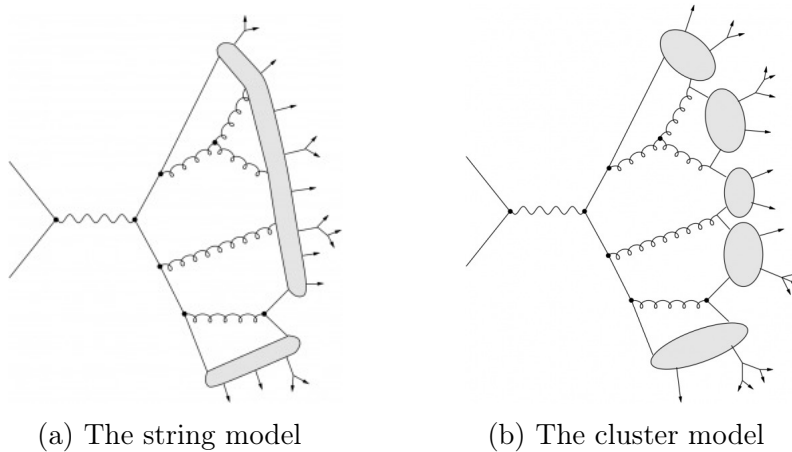
Figure 1.5: (a) The string hadronisation model directly transforms partonic systems in hadrons. (b)The cluster model uses clusters as intermediate steps, with a typical mass of few GeV.

The string model is based on the hypothesis that quarks correspond to the final points of a string and gluons are deviations from it (see Fig.1.5a). All along the string, partons are ordered according to their colour. This model provides a really predictive picture of the primary hadron movement in the space-time. This structure can be applied to multi-hadron configurations. A possible disadvantage is the presence of many parameters linked to quark flavours, which must be tuned with data.

The cluster model is based on confinement features of the PS, which ends with the reconstruction of parton clusters with low mass values. This model starts with a non perturbative splitting of gluons in $q\bar{q}$ pairs, which then originate clusters. The majority of clusters supports the sequential 2-body decay in the phase space. The low-mass spectrum, and the corresponding low transverse momentum, determines the suppression of heavy hadrons. The decay of heavy clusters needs an additional stage characterised by the decay in lighter clusters (see Fig1.5b). Even if the cluster model is less precise compared to

the string one, it provides a complete description of events at high energy when combined with PS, by using a limited number of parameters.

Multiple Interactions and Underlying Events

The underlying event (UE) involves the additional activity from multiple interactions and proton remnants and it is dealt with non-perturbative models, tuned on experimental data.

Multiple simultaneous interactions are modelled by overlaying inelastic p-p events ("minimum bias") on top of the hard scattering. The number of such collisions is stochastic and modelled to match the pileup level in data. This only accounts for the "in-time" pileup: the radiation resulting from collisions that occurred in the same bunch crossing as the primary hard-scattering. "Out-of-time" pileup from bunch crossings different from the primary one is modelled in the same way, but an offset in time is introduced in the simulation to accurately model the processes taking more than the bunch spacing [15].

The interactions among proton remnants involves both 1→2 and 2→2 processes, where the second is statistically more probable. The production of these processes is divergent for $p_T \to 0$ and decreases rapidly for high p_T: therefore in the generation mechanism a lower threshold is introduced ($\sim$2 GeV).

1.3   The Z boson

Since the discovery in 1983 by the UA1 experiment [16], the Z boson has been and still is object of interest in the high energy physics community. In this Section, the Z boson fundamental properties are described, focusing on the reason why after more than 30 years from its discovery, it is still an important ingredient in the particle physics research.

1.3.1   The discovery

The Z boson was discovered in 1983 by the UA1 collaboration of the S$p\bar{p}$S collider at CERN [16], whose primary goal was the search of the massive

intermediate bosons of the electroweak interaction. After the discovery of the W bosons in events with single isolated electrons and missing energy [17], the UA2 detector observed eight events interpreted in terms of the reaction

$$\bar{p}p \to Z + \text{anything} \quad \text{with} \quad Z \to e^+e^-$$

in a data sample corresponding to a total integrated luminosity of 131 nb^{-1}. Fig.1.6 shows the distribution of 24 events passing loose requirements based on the energy in the clusters of the electromagnetic calorimeter. Requiring at least one isolated charged track with $p_T > 7$ GeV pointing to the electromagnetic cluster, the number of events was reduced to eight and their distributions as a function of the invariant mass of the dilepton pair is shown in Fig.1.6b. As can be seen from the plot, the eight events are distributed around a mass value of ∼90 GeV. Requiring the track of isolated electron to point to both energy deposits in the electromagnetic calorimeter, only three events survived (shaded band in Fig.1.6b). From these events, the UA2 collaboration measured the mass of the Z boson to be:

$$M_Z = (91.9 \pm 1.3 \pm 1.4)\text{GeV}$$

where the first error accounts for the analysis strategy and the second error for the uncertainty on the energy scale. Assuming the Breit-Wigner distribution of the events, an upper limit on the full width was placed to $\Gamma < 11$ GeV, corresponding to a maximum of ∼50 different neutrino types in the universe. In analogy with the studies performed in the electron channel, in the same year measurements of the Z boson have been carried out in the muon decay mode $Z \to \mu^+\mu^-$ [18]. Three events survived to the full selection chain and were used to measure the Z mass with a value $M_Z = 85.8^{+7.0}_{-5.4}$ GeV, consistent with the value measured in the electron decay channel.

1.3.2   Z boson properties

Precision measurements of the Z-boson resonance were performed by LEP, an electron-positron collider built at CERN in 90's. During a first phase (LEP1) the centre of mass energy was kept in an interval of ± 3 GeV around
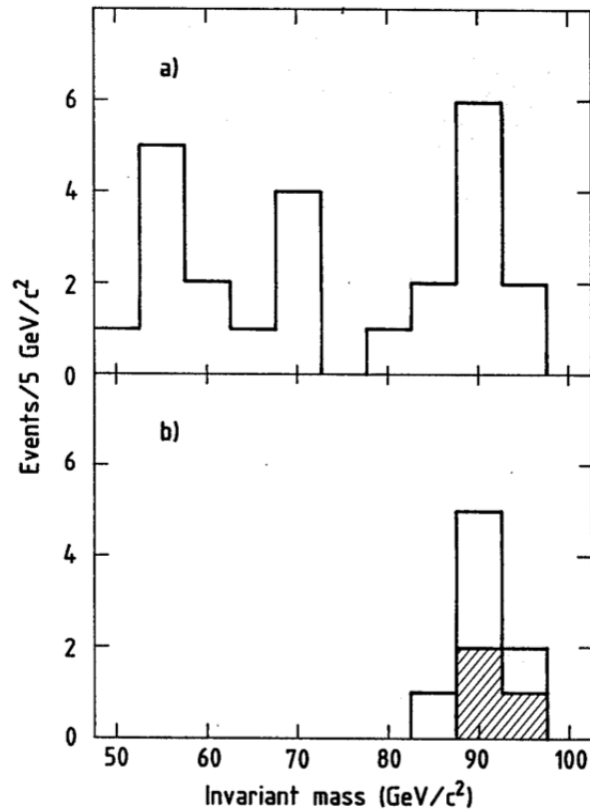
Figure 1.6: Invariant mass distribution of electron pairs in first observed Z candidate events at the UA1 experiment of the S$p\bar{p}$S collider. Fig.a) shows the distribution of 24 events passing requirements on the energy deposits in the electromagnetic clusters. Introducing identification criteria the number of events was further reduced to eight, whose invariant mass distribution peaks around a 90 GeV, as presented in Fig.b). The shaded band in Fig.b) represents the three events surviving the requirement of an isolated track matched with both electromagnetic clusters.

the m_Z; this allowed experiments to collect $\sim 2 \cdot 10^4$ Z decays during 1989-1995 and to perform high precision measurements, which still dominates the world average [3].

The Z boson lineshape parameters, like the mass (m_Z), the total width (Γ_Z) and the partial widths ($\Gamma_{f\bar{f}}$), have been determined from an analysis of the production cross sections of fermionic final states starting from e^+e^- collisions. The shape of the cross section variation around the Z peak can be parametrised by a Breit-Wigner function, with an energy dependent total width. The determination of these parameters was done via an analytic expression of this cross section in term of the parameters themselves

$$\sigma(s) = \sigma_{f\bar{f}}^0 \frac{s\Gamma_Z^2}{(s - m_Z^2)^2 + \frac{s^2}{m_Z^2}\Gamma_Z^2} \qquad (1.22)$$

and fitting the calculated cross section to the measured one, by varying these parameters [3]. In Eq.1.22, $\sigma_{f\bar{f}}^0$ represents the cross section for the process $e^+e^- \to f\bar{f}$ at $\sqrt{s} = m_Z$. If the final state involves the fermionic e^+e^- couple, the above cross section must be integrated to take into account small ($\mathcal{O}(1\%)$) contributions from photons exchange and γ-Z interference.

The world average value of the Z boson mass and total width are [20]:

$$m_Z = (91.1876 \pm 0.0021)\text{GeV}$$
$$\Gamma_Z = (2.4952 \pm 0.0023)\text{GeV}$$

The pole cross section can be expressed as a function of the partial widths of the Z decay modes $\Gamma_{f\bar{f}}$:

$$\sigma_{f\bar{f}}^0 = \frac{12\pi}{m_Z^2} \frac{\Gamma ee\Gamma_{f\bar{f}}}{\Gamma_Z^2} \qquad (1.23)$$

According to the Standard Model, which predicts the Z boson to decay with comparable probability into all species of fermions kinematically allowed, the total width can be expressed as $\Gamma_Z = N_\nu\Gamma_\nu + 3\Gamma_{ee} + \Gamma_{hadrons}$, where N_ν is the number of neutrino families. Measuring the Z decay probabilities in leptons and in hadrons, LEP obtained $N_\nu = 3.27 \pm 0.30$, discarding the hypothesis of a fourth generation of fermions at 98% CL [19].

In Table 1.2 the Z partial widths and the corresponding branching ratios are reported for each decay channel [20].

Decay channel	Partial width ($\Gamma_{f\bar{f}}$) (MeV)	Branching ratio ($\Gamma_{f\bar{f}}/\Gamma_Z$) (%)
e^+e^-	83.91 ± 0.12	3.363 ± 0.004
$\mu^+\mu^-$	83.99 ± 0.18	3.366 ± 0.007
$\tau^+\tau^-$	84.08 ± 0.22	3.370 ± 0.008
invisible	499.0 ± 1.5	20.00 ± 0.06
hadrons	1744.4 ± 2.0	69.91 ± 0.06

Table 1.2: The Z bosons partial widths and branching ratios for each decay mode.

1.4 The inclusive production cross section

The dominating contribution to the Z boson production at LHC is the *Drell-Yan* process, based on the proton-proton scattering at high energy. At the parton level, the annihilation of a couple of quark-antiquark of the same flavour produces a Z boson (or a virtual photon) decaying in a couple of fermions with high transverse momentum. The leading order Feynman diagram of the Drell-Yan production is shown in Fig.1.7.
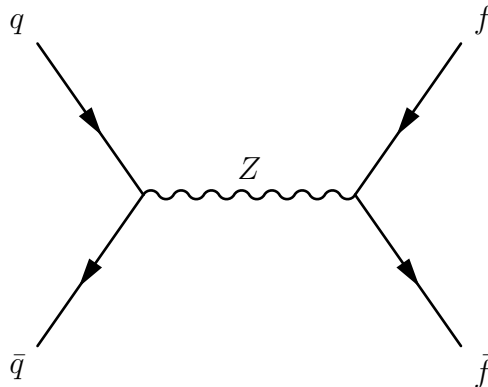

Figure 1.7: Feynman diagram for the Drell-Yan process.

Considering the hadronic structure of the proton, made of two u- and one d-quarks, the Drell-Yan process can be originated by one valence quark and

one antiquark coming from the virtual-sea. Because of the different nature of the initial state quark and antiquark, the valence quark propagates with a larger fraction x of transverse momentum of the proton. At the leading order, the invariant mass of the fermions in the final state is $M_Z^2 \sim x_1 x_2 s$; as a consequence, at the centre of mass energy of $\sqrt{s} = 13$ TeV, the Z boson is produced if $x_1 x_2 > 10^{-6}$.

The partonic cross section for the Z inclusive production is described in Eq.1.22; next-to-leading-order corrections increment the cross section value by about $\sim$20-30%.

1.4.1 State of art of Z boson measurements

The production of the Z vector boson at hadron colliders through the Drell-Yan mechanism is extremely important for physics studies. The process has large production rate and offers a clear signature, given the presence of two high-p_T leptons in the final state. Since the discovery at CERN S$p\bar{p}$S [?], several measurements of the Z boson production cross section at hadron colliders have been performed in $p - \bar{p}$ collisions at Tevatron by the CDF [21] and D0 [22] experiments at centre of mass energy $\sqrt{s}$ =1.8 TeV and $\sqrt{s}$ =1.96 TeV and at LHC by ATLAS and CMS in p-p collisions. These high precision measurements concurred to the determination of the properties of the Z boson. Since the Z boson is considered a Standard Model candle useful for detector calibration and alignment, it is one of the first measurements to be performed after any change in experimental conditions (i.e. centre of mass energy or detector layout).

For the above reasons, it is essential to have accurate theoretical predictions for the vector boson cross sections and the associated kinematic distributions. Cross section calculations are performed at LO as an hard-scatter reaction at the energy scale of the Z mass; significant QCD corrections are then considered at higher order of calculation which increase the value of the cross section itself. Nowadays predictions of the Z boson are generally available up to NNLO in perturbation theory [23].

At LHC, high precision measurements are performed in kinematic domains never explored before and the comparison with accurate theoretical predic-

tions allows to perform challenging tests of perturbative QCD and helps in the determination of the initial condition of PDF evolution. The knowledge of PDFs of the proton mainly comes from deep inelastic scattering (DIS), which cover a broad range of transferred momentum Q^2 and momentum fraction x. As explained in Section 1.2.3, in the region $x \leq 0.01$, PDFs are constrained by the precise measurements performed at HERA, which determine a specific combination of light quark and antiquark distributions. However the flavour composition of the total light sea $x\Sigma = 2x(\bar{u}+\bar{d}+\bar{s})$ has not been established at $x \leq 0.01$ values [24]. In order to achieve this measurement, two fundamental components contribute to complement the information from deep inelastic scattering: the rapidity distribution of the Z boson and the ratio of the W and Z cross sections. The rapidity (y) dependence of Z production in the Drell-Yan process provides constraints on the PDFs of the proton, since it is strongly correlated with the proton momentum fraction x_1, x_2 carried by the partons participating in the hard scattering process:

$$y = \frac{1}{2}ln\frac{x_1}{x_2} \qquad (1.24)$$

The ratio of $(W^+ + W^-)$ and Z cross sections in a leading-order calculation can be expressed as

$$\frac{W^+ + W^-}{Z} \sim \frac{u_v + d_v + 2s}{(V_u^2 \cdot A_u^2)(u_v + s) + (V_d^2 \cdot A_d^2)(d_v + s)} \qquad (1.25)$$

where $u_v(d_v)$ is the up (down) valence quark distribution functions and $V_{u,d}$ and $A_{u,d}$ are the vector and axial-vector weak neutral current couplings of the light quarks. As the numerical values for the Z coupling to the up and down-quarks $V_{u,d}^2 \cdot A_{u,d}^2$ are of similar size, the $W^\pm/Z$ ratio measures a PDF insensitive quantity, providing the flavour composition of the light sea.
ATLAS W/Z cross section measurements performed at $\sqrt{s} = 7$ TeV (corresponding to a luminosity of 35 pb^{-1}) [25] suggested the light quark sea *(u, d, s)* is flavour symmetric. In particular, a QCD study of the results was performed [24], fitting the distribution of the rapidity of the Z boson in two different hypothesises:

   1) the strange quark distribution is suppressed by fixing $\bar{s}/\bar{d} = 0.5$ at the

initial scale of PDF evolution ($Q_0^2 =$1.9 GeV^2 and x=0.013);

2) the strange quark distribution is parametrised assuming $\bar{s} = \bar{d}$ and $xs = x\bar{s}$.

The fit determines the value of:

$$r_s = 0.5\frac{(s + \bar{s})}{\bar{d}} \tag{1.26}$$

to be $r_s = 1.00^{+0.25}_{-0.28}$, evolving the PDF at the scale of the Z mass ($Q^2 = M_Z^2$). This is consistent with the prediction that the light quark sea at low x is flavour symmetric. The Z cross section measurement is compared to NNLO fit results in Fig.1.8, where it is clear that the free $\bar{s}$ fit leads to an improvement in the prediction of the y_Z description.
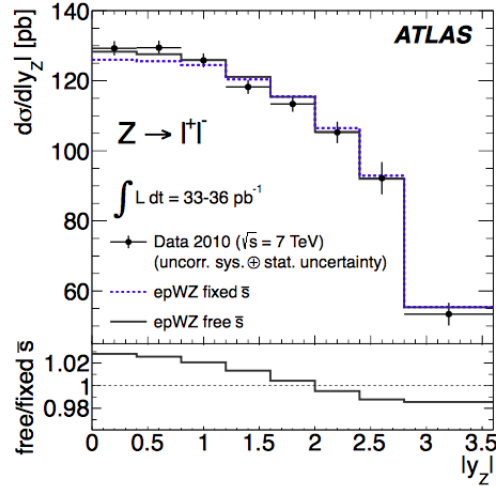
Figure 1.8: Differential $d\sigma/dy_Z$ cross section distribution measured with data collected at $\sqrt{s}$ =7 TeV by the ATLAS experiment. Data are compared with two NNLO fit curves evaluated with free (continuous line) and fixed (dashed line) strangeness, whose ratio is shown in the bottom part of the figure. The error bars take into account the statistical and the uncorrelated uncertainties added in quadrature. Data seem to favour the free strangeness fit.

The enlarged fraction of s-quark in the proton sea leads to a $\simeq$10% decrease of the down and up quark sea densities at the initial scale Q_0^2 because $x\bar{s}$, $x\bar{d}$ and $x\bar{u}$ are tied together at low x by the distribution of data. The

total sea ($x\Sigma$) is correspondingly enhanced by $\simeq 8\%$ as illustrated in Fig.1.9.
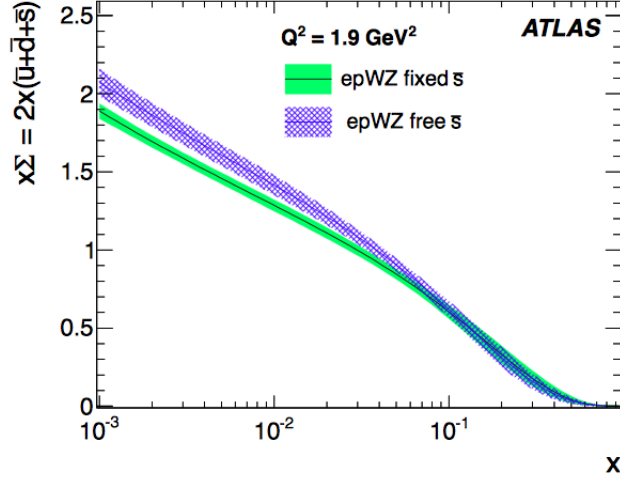
Figure 1.9: Distribution of the light sea quarks, $x\Sigma = 2x(\bar{u} + \bar{d} + \bar{s})$, in the NNLO analysis of ATLAS data with a fixed fraction (green line) and with a fitted fraction (blue line) of strangeness.

The high precision measurements of the Z boson allow to constrain the PDF behaviour in specific kinematic range. The results on the production cross section at $\sqrt{s} = 7$ TeV (with luminosity of 35 pb^{-1} and 4.6 fb^{-1}) and their QCD interpretation [24] have been studied in combination with the DIS data from HERA. From this examination of ATLAS+HERA data, the new PDF set called ATLAS-epWZ16 has been provided, more sensitive to the flavour composition of the sea quark and valence quark distribution at low x [26]. The strange quark distribution in the proton sea has been measured using the total Run 1 statistics, obtaining $r_s = 1.00 \pm 0.07(exp)^{+0.13}_{-0.14}$, with a total uncertainty of about 16%, a reduction with respect to the previous determination by the ATLAS Collaboration [25]. The result is compared with different PDF sets in Fig.1.10: the new ATLAS-epWZ16 shows the best agreement with data, highlighting the hypothesis of symmetric light sea flavour.

Another important measurement which benefits from the high precision reachable in the Z analysis is the ratio of the top pair production cross

Figure 1.10: Determination of the strange-to-down sea quark fraction r_s. The coloured bands represent the measured value and its different uncertainty contributions from experi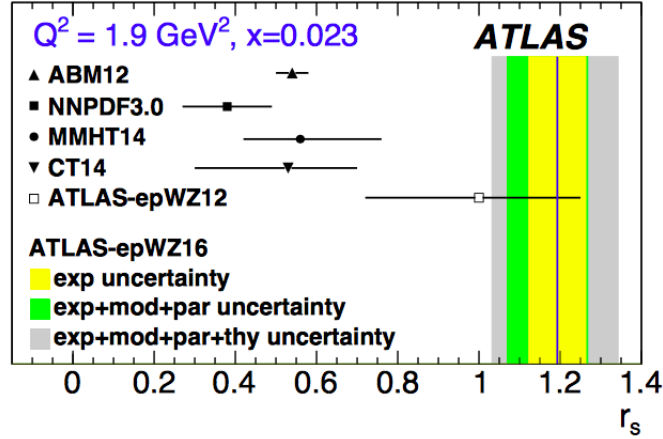mental data, QCD fit and theoretical uncertainties. The measurement has been compared with different NNLO PDF sets. r_s is calculated at the initial scale $Q_0^2 = 1.9$ GeV^2 and at $x=0.023$, corresponding to the point of largest sensitivity at central rapidity of ATLAS data.

section to the Z cross section. Since top-pairs are mainly produced through the gluon-gluon fusion, the $t\bar{t}$ process is sensitive to the gluon distribution within the proton. Given that at a fixed centre of mass energy the luminosity uncertainty cancels, the ratio $t\bar{t}/Z$ has a significant sensitivity to the gluon-to-quark PDF ratio. Considering that the gluon fraction has a stronger dependence on the centre of mass energy than the quark-antiquark one, the measurements of the $t\bar{t}/Z$ ratio has been performed at $\sqrt{s} =7$, 8 and 13 TeV in order to study the gluon PDFs in different x-Bjorken regions [27]. The $t\bar{t}/Z$ measurements at $\sqrt{s} =13$ TeV is compared with different PDF sets in Fig.1.11, showing that the precision reached in the measurement is more accurate than predictions and thus helpful in constraining the gluon PDF.

In order to establish the impact of data on PDF uncertainties, a set of PDFs has been reweighed on the basis of the results of a profile likelihood fit of PDFs on data. Fig.1.12 presents the distribution of the gluon PDF uncertainty as a function of Bjorken-x, showing that the ATLAS $t\bar{t}$ and Z cross section data impose visible constraints on the gluon distribution at $x \sim 0.1$, contributing in reducing the uncertainty.
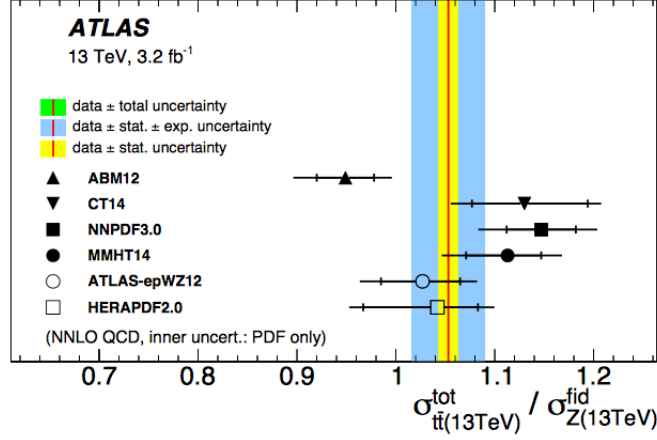
Figure 1.11: $\mathcal{R} = \sigma_{tt^{tot}}/\sigma_{Zfid}$ at $\sqrt{s} =$13 TeV compared to predictions based on different PDF sets. The inner shaded band corresponds to the statistical uncertainty, the middle band to the statistical and experimental systematic uncertainties added in quadrature, while the outer band shows the total uncertainty. The theory predictions are given with the corresponding PDF uncertainties shown as inner bars, while the outer bars include all uncertainties added in quadrature.

1.5   The Z+jets production cross section

As discussed in Section 1.4, the Z boson is produced at LO through the Drell-Yan annihilation of an incoming quark-antiquark pair, which occurs about 65% of times. At NLO the Z boson is produced in association with a jet, a process which involves not only an electroweak vertex, but also a strong interaction vertex. The Feynman diagram of the Z produced with one jet is shown in Fig.1.13: this process contributes roughly 20-30% to the overall Z cross section. Any additional parton in the final state implies supplementary α_S vertex, being therefore suppressed.

The Z+jets production is of extremely high importance within the Standard Model. Since it is a NLO process, it allows to perform studies of perturbative QCD in a wide kinematic range and at different jet multiplicities. From the comparison with theoretical expectations, the high precision reached in the results allows to constrain PDFs, which are the main source of uncertainty in the multi-parton final state modelling. Minimizing the experimental systematic uncertainty to increase the sensitivity of the measurement
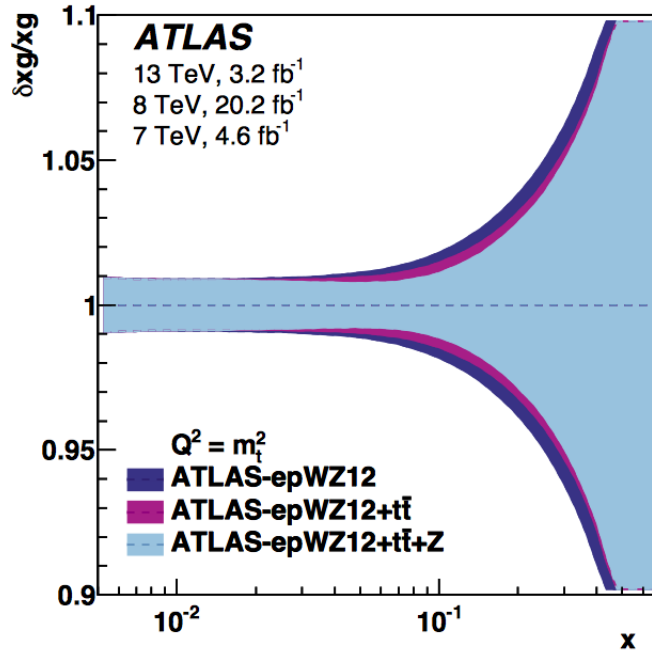
Figure 1.12: Impact of the ATLAS Z-boson and $t\bar{t}$ cross section data on the determination of PDFs. The blue band represents the uncertainty for the ATLAS-epWZ12 PDF set, while the purple and light-blue bands show the uncertainty of the "profiled" PDF after the inclusion of $t\bar{t}$ and $t\bar{t}$+Z measurements, respectively: constraints on the gluon distributions at $x \sim 0.1$ are clearly visible.
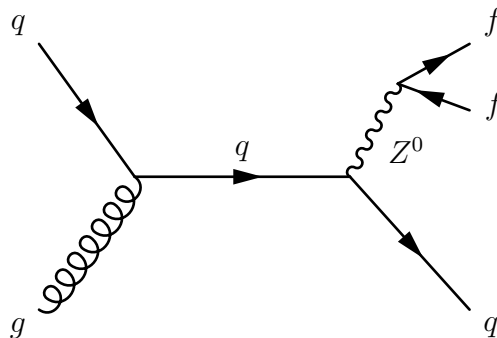
Figure 1.13: Feynman diagram for the NLO process of Z produced in association with one jet.

is therefore of great importance. The Z+jets events also constitute a non-negligible background for studies of the Higgs boson and in searches of new phenomena: typically in these studies, the multiplicities and kinematics of the jets are exploited to achieve a separation of the signal of interest from the SM Z+jets process. These quantities are often simulated with Monte Carlo generators, which are themselves subject to systematic uncertainties and must be validated with data.

The cross section ratio of massive vector bosons, measured in the leptonic final states, in presence of jets is defined by $\mathcal{R}_{\text{jets}} = \sigma_{W+\text{njets}}/\sigma_{Z+\text{njets}}$. Since it benefits from the partial cancellation of the luminosity and PDF approximation uncertainties, this ratio is sensitive to the dynamics of these processes, allowing to perform precise tests of perturbative QCD.

1.5.1 State of art of Z+jets measurements

Previously measured at Tevatron and LHC colliders, the Z+jets differential cross section has been measured by ATLAS with data corresponding to an integrated luminosity of 3.16 fb^{-1}, collected at $\sqrt{s}$ =13 TeV during 2015 [28]. The differential cross section has been measured as a function of many kinematic variables in order to model jet distributions. In Fig.1.14 the cross section has been measured up to a jet multiplicity of seven and compared with different predictions, presented with the PDF and QCD scale uncertainty contributions. In particular, NLO SHERPA 2.2 calculation is compared with the LO simulations from MG5_aMC@NLO CKKWL, MG5_aMC@NLO FXFX and ALPGEN and to the NLO fixed-order predictions from BLACKHAT+SHERPA. The partonic cross section decreases logarithmically as a function of the increasing number of jets involved in the process. Prediction are in agreement with the results for events with up to four jets; for higher multiplicities, jets produced through Parton Shower, involving soft (i.e. non-perturbative) final state emission, should be considered.
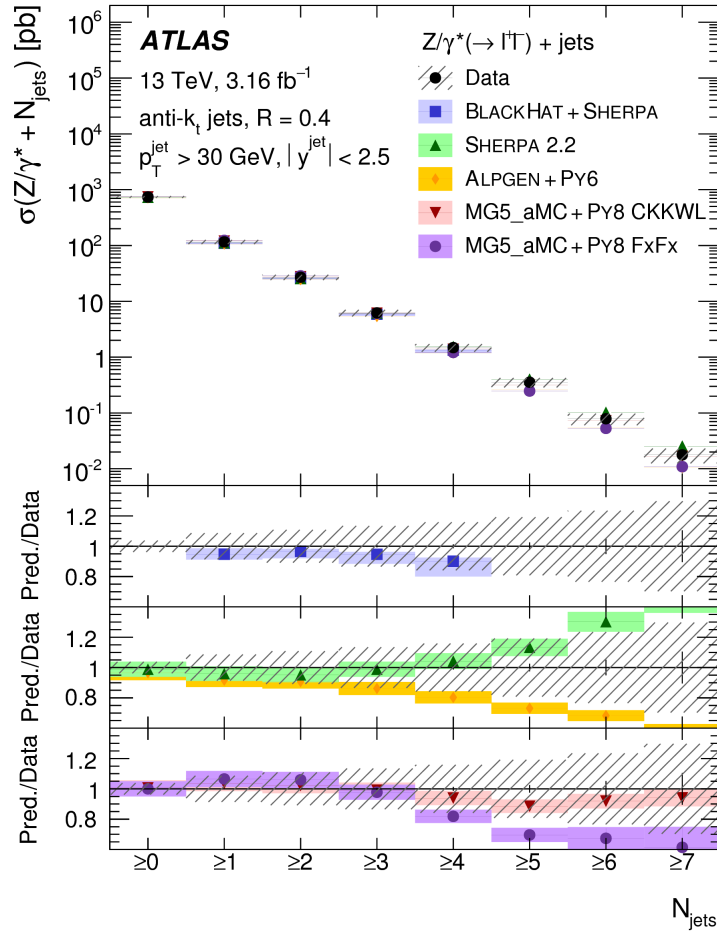
Figure 1.14: Measured Z+jets cross section as a function of the inclusive jet multiplicity. The data are compared to the predictions calculated with different assumptions. The error bars correspond to the statistical uncertainty, and the hatched bands to the statistical and systematic uncertainties (including luminosity) on data added in quadrature.

(a) Z+b 4FNS diagram

(b) Z+b 5FNS diagram

(c) Z+bb: $q\bar{q}$ initial state

(d) Z+bb: gg initial state
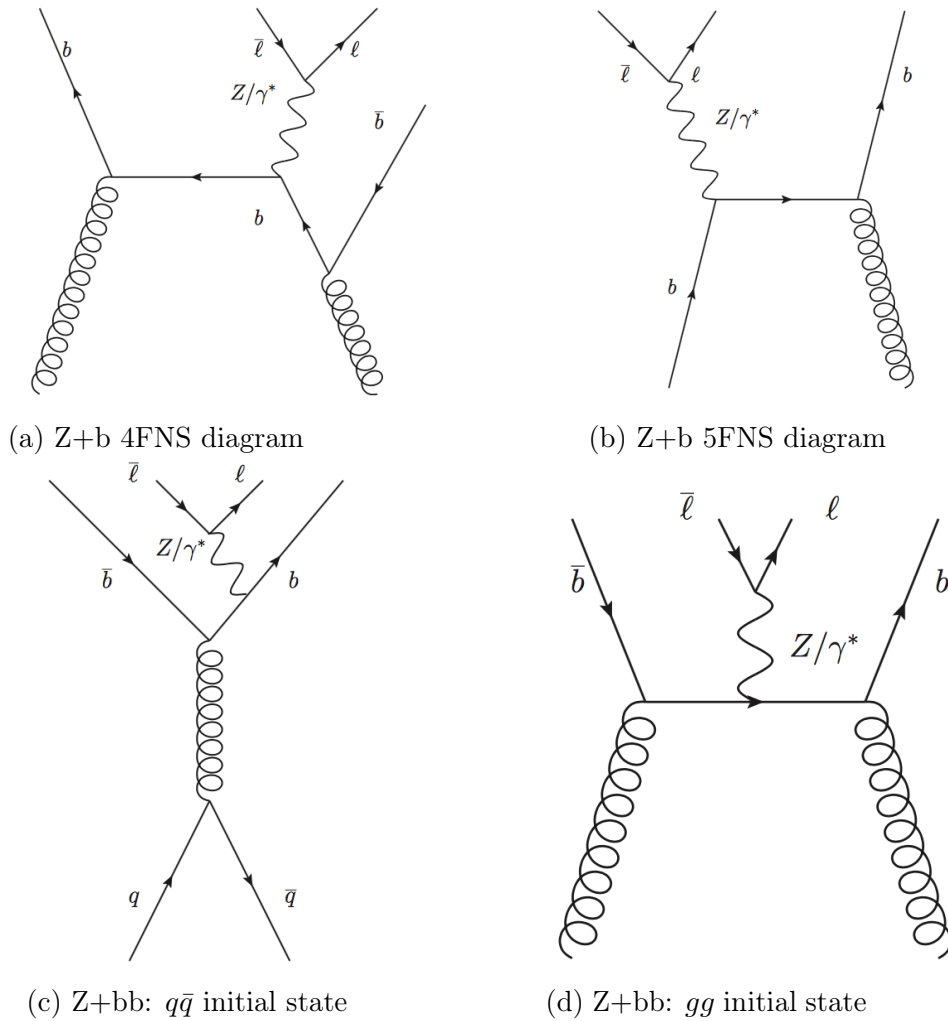
Figure 1.15: Representative Feynman diagrams for Z+b (top) and Z+bb (bottom) production. The leading tree–level diagram for Z+b production in the 4FNS (a) and 5FNS (b) are included. The Z+bb production modes with $q\bar{q}$ (c) or gg (d) initial state are shown. Note that the diagram in (a) is the same as in (d) but with the latter producing two b–quarks inside the acceptance.

1.6   The Z+bjets production cross section

The measurement of the Z boson produced in association with heavy flavour jets is the main subject of this thesis. These processes are a benchmark for Monte Carlo simulations and theoretical calculations, since the predicted cross section rates and distributions reflect both the structure of the proton and the gluon splitting into a pair of b-quarks. For the calculation of the Z+b(b) cross section, two approaches have been developed in parallel, which are referred to as the "four flavour number scheme" (4FNS) and the "five flavour number scheme" (5FNS) [29, 30]. In the 4FNS the b-quarks are produced by hard initial state gluon splitting, described by the QCD matrix element; this process leads to at least one b-jet in the final state. Fig.1.15a represent the Feynman diagram of the process. The 4FNS matrix element treats the b-quark as massive and thus it is not considered in the proton PDF evolution and in the perturbative calculation. On the other hand, in the 5FNS the b-quark PDF is calculated from the gluon density evolution above the b-quark mass threshold ($Q^2 >> m_b^2$) and describing the b-quark as massless. It allows a b-quark density in the initial state and raises the prospect that measurements of heavy flavour production could constrain the b-quark PDF of the proton. An example of Feynman diagram for the Z+b process in the 5FNS is presented in Fig.1.15b. The production of a b-quark pair in association with a Z boson is dominated by two different production modes characterised by quark-antiquark annihilation (Fig.1.15c) or gluon-gluon fusion (Fig.1.15d) in the initial state. In particular, the $q\bar{q}$ initial state includes $b\bar{b}$ if the 5FNS is adopted.

In a calculation to all orders, the 4FNS and 5FNS must give identical results; however, at a given order differences appear between the two. NLO calculations combining the two schemes for initial partons still carry large uncertainties.

The Z+b(b) production is a major background to a large number of process with smaller cross section, such as top-quark, Higgs boson and searches for new physics phenomena. The measurements of the production cross section of the Higgs boson, decaying in a $b\bar{b}$ couple, associated with a Z boson cannot be separated from the Z+bb background. As shown in Fig.1.16 (left),
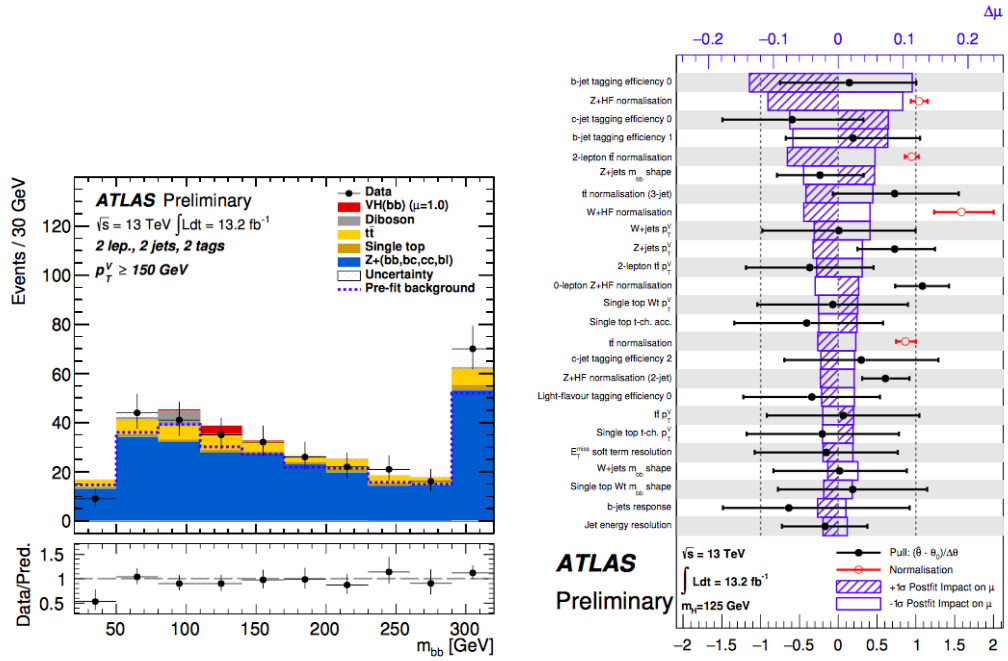
Figure 1.16: Impact of the Z+b(b) production on the ZH($\to b\bar{b}$) measurement. On the left, post-fit distribution of the invariant mass of the $b\bar{b}$ pair (M_{bb}): the Z boson produced with heavy flavour-jets is the major irreducible background. On the right, systematics uncertainty contributions to the fitted signal-strength parameter μ: the normalisation of the Z+b(b) background and the shape of the M_{bb} distributions are among the largest systematics uncertainty [31].

the contamination is still large and multivariate analyses (BDT) are employed to discriminate signal from background events [31]. BDTs are trained in different categories and used to build a global fit on sensitive variables. Among the fitted variables, the invariant mass (M_{bb}) and the angular separation (ΔR_{bb}) of the $b\bar{b}$ couple are considered, which suffer form large modelling uncertainties. Fig.1.16 (right), usually referred to as "ranking plot", presents the signal strength of the ZH process (i.e. the ratio of the measured cross section value with respect to the SM predicted one) together with the main systematic contributions of the various BDT variables: the flavour determination and modelling have a huge impact on the final result. A high precision measurement of the Z+b(b) production can improve the theoretical calculations and thus the modelling.

Other important measurements are the ratios between the Z and W bosons produced in association with one b-jets $\mathcal{R}_b = \sigma_{Z+b}/\sigma_{W+b}$ and two b-jets $\mathcal{R}_b = \sigma_{Z+bb}/\sigma_{W+bb}$. Benefiting from the great reduction of some experimental uncertainties, R_{bb} can probe the difference between the kinematic distributions of the b-jet system recoiling against the W or the Z [32]. At low energy, the difference in vector boson masses translates to a change in momentum transfer between the incoming partons and thus different kinematics of the final state b-jets. At high energy, the vector boson mass difference is small compared to the momentum transfer and the cross section ratio is expected to stabilise around 1. A precise measurement of R_b can therefore be used in the context of searches and to constrain the b-quark PDFs.

1.6.1 State of art of Z+bjets measurements

Measurements of heavy flavour production in association with a Z boson have attracted the interest of the Tevatron and LHC experimental communities in the last decades. The process has been measured in proton-antiproton collisions at $\sqrt{s}$ =1.96 TeV by CDF [33] and D0 [34] experiments and in proton-proton collisions at $\sqrt{s}$ =7 TeV by ATLAS [35] and at $\sqrt{s}$ =7,8 TeV by CMS [36, 37].

The ATLAS cross section for Z+≥1 b-jet and Z+≥2 b-jets are shown in Fig.1.17 and compared with several NLO and LO calculations. Looking at

aMC@NLO, the 5FNS prescription is in agreement with data for the Z+≥1 b-jet cross section measurement, while the 4FNS model underestimates the measured cross section. The situation is reversed for the Z+≥2 b-jets case, where the 4FNS provides a good modelling, while the 5FNS underestimates the cross section, due to the fact that the 5FNS Z+ ≥2 b-jets process is generated only at LO.
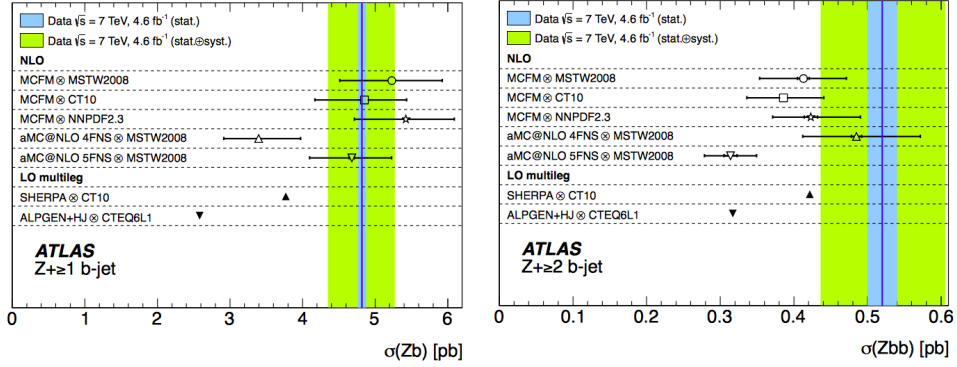
Figure 1.17: Cross section measurement for Z+≥1 b-jet (left) and Z+≥2 b-jets (right). The data are shown as a vertical blue line with the statistical uncertainty (blue band) and the sum in quadrature of statistical and systematic uncertainties (green band). Data are compared to NLO predictions from MCFM and aMC@NLO and to LO predictions from Alpgen+Herwig+Jimmy and Sherpa.

Cross section measurements have been performed as a function of several b-jet and Z kinematic variables. In particular, σ_{Z+b} distributions as a function of Z transverse momentum (p_T^Z) and rapidity (y_Z) are shown in Fig.1.18. Despite all theoretical predictions provide a reasonable description of the shape of data, some discrepancies appear at very high p_T^Z (>200 GeV) and the ratio of the MCFM [38] predictions to data shows a slope as a function of y_Z. Although the Z boson rapidity is the most sensitive to PDFs variable, the large scale theoretical uncertainties prevent from drawing conclusions on the b-quark production scheme.

Fig.1.19 presents the differential σ_{Zbb} distributions as a function of the b-jets invariant mass M_{bb} and angular separation ΔR_{bb} in the $\eta - \phi$ plane. Some evidences of deviations with respect to NLO theoretical predictions appear at low M_{bb} and ΔR_{bb}, in the range where gluon-gluon spitting process is studied.
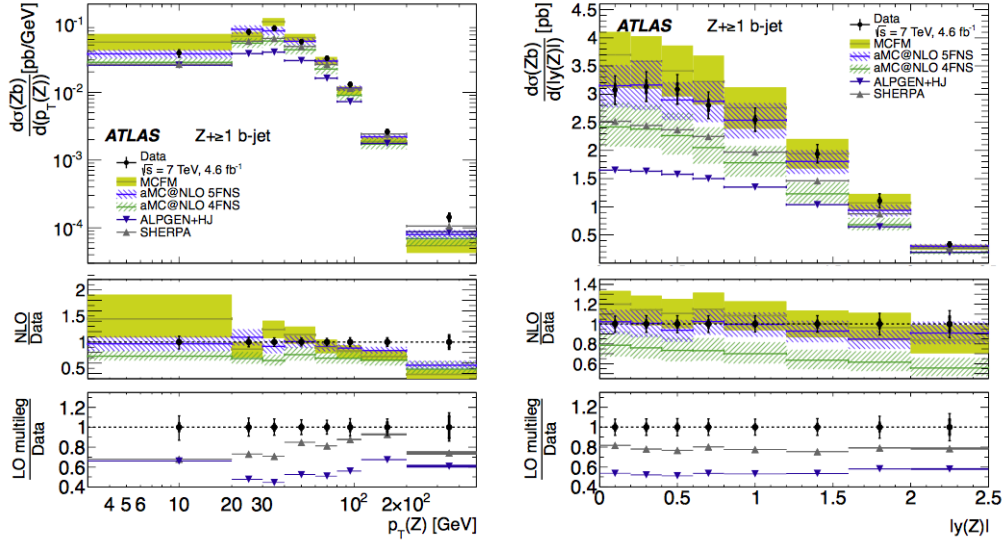
Figure 1.18: Differential σ_{Z+b} cross section as a function of Z boson transverse momentum (left) and rapidity (right). The measured results are compared with NLO predictions from MCFM and aMC@NLO and LO simulations from Alpgen+Herwig+Jimmy and Sherpa.

The results are really interesting to be followed up in Run 2. The main goal of this thesis is indeed to perform differential cross section measurements of the Z+$\geq$1 b-jet and Z+$\geq$2 b-jets processes at the centre of mass energy of 13 TeV, benefiting from the high luminosity and the introduction of new algorithms for b-jet tagging, which can improve the sensitivity to the b-quark PDF of the proton and to gluon splitting.
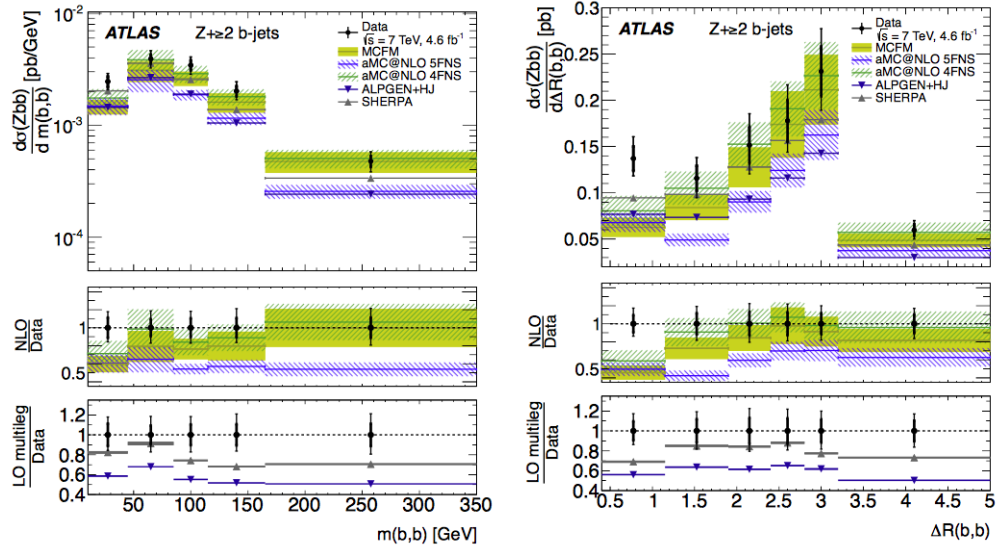
Figure 1.19: Differential σ_{Z+bb} cross section as a function of the invariant mass of the b-jets M_{bb} (left) and angular separation ΔR_{bb} (right). The measured results are compared with NLO predictions from MCFM and aMC@NLO and LO simulations from Alpgen+Herwig+Jimmy and Sherpa.

Chapter 2

The ATLAS experiment and particle reconstruction

Founded in 1954, the Conseil Européen pour la Recherche Nucléaire (CERN) is the world largest research centre for particle physics. Located near Geneva, it was one of the first European joint project and nowadays it has 22 member nations that cooperate to probe the fundamental structure of the universe. The complex and large scientific instrumentation involves accelerators to boost beams of particles up to very high energy and detectors to collect and observe the results of the collisions.

The Large Hadron Collider (LHC) is the newest accelerator complex at CERN. Built in a circular tunnel of 27 km and situated from 50 to 175 m under ground, it hosts four detectors along its ring: ATLAS, CMS, LHCb and ALICE. In Section 2.1, a description of the LHC complex is given, completed in Section 2.2 by an overview of the physics requirements that guided its construction. The two measurements I present in this thesis have been performed using data collected by ATLAS, whose features are presented in Section 2.3.

In Section 2.4, the physics objects reconstructed in ATLAS and used in the analyses I present in this document are described: tracks, electrons, muons, jets, b-jets and transverse energy.

2.1   LHC

LHC was built to address the challenges of the post-LEP (1989-2000) and post-Tevatron (1987-2011) era, exploring the TeV scale. It consists of a 27-kilometer ring where beams of particles travel in opposite directions, kept separated in two ultra-high vacuum chambers at a pressure of 10^{-10} torr, and collide in four points, where the four experiments are hosted. The proton-proton (p-p) collision mode is the primary operational setup, but LHC is also designed to perform heavy ion collisions.
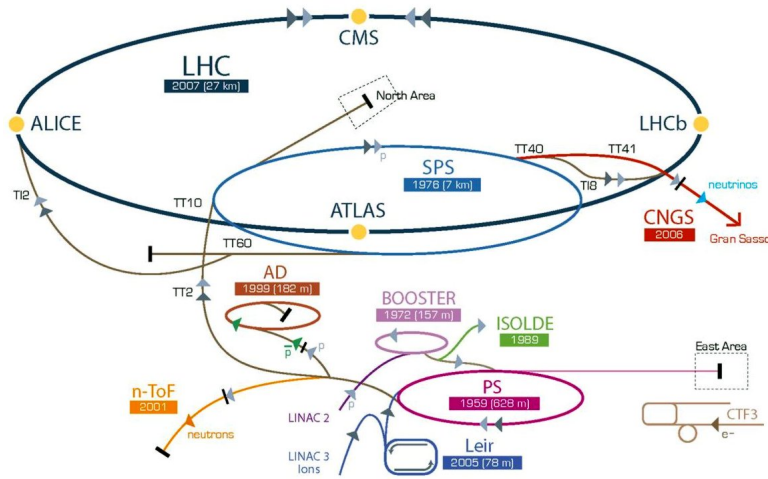
Figure 2.1: Schematic layout of the CERN acceleration complex and experiments.

The acceleration process takes advantage from the whole CERN complex, whose final step takes place in the LHC ring, as depicted in Fig.2.1. Protons are obtained by ionising hydrogen atoms and are first accelerated up to 50 MeV by LINAC, a linear accelerator. After this first acceleration stage, circular systems are used to increase particle energy. Beams are injected in order into the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), reaching energy of 1.4 GeV, 25 GeV and 450 GeV, respectively, before being delivered to LHC. PS and SPS provide protons in bunches, separated by 25 ns; a maximum of 2808 bunches, containing about 10^{11} protons each, can be injected into the LHC

ring. Inside LHC, proton bunches are kept circulating by using 1232 NbTi superconducting magnets and are focalised by 392 additional quadrupole magnets. The superconducting magnets are cooled down to 1.9 K with liquid helium. The magnet systems use a twin bore design to bend particles in both beams simultaneously. 16 radiofrequency cavities with a maximum electric field of 5.5 MV/m are used to accelerate particles from 450 GeV to a maximum of 7 TeV. The ensemble of all bunches gives rise to 10^9 collisions per second, assuming a total proton-proton cross section of 10^{-25} cm^2 at the LHC energy. The most important LHC parameters are reported in Table 2.1 [39].

	Design	2012	2015	2016
Beam energy (TeV)	7	4	6.5	6.5
Dipole magnetic filed (T)	8.33	∼6.3	∼8.0	∼8.0
Number of protons per bunch	$1.15 \cdot 10^{11}$	$1.7 \cdot 10^{11}$	$1.2 \cdot 10^{11}$	$1.18 \cdot 10^{11}$
Maximum number of proton bunches	2808	1374	2244	2076
Number of bunch places		3564		
Bunch separation (ns)	25	50	50-25	25
Peak luminosity (cm^{-2}s^{-1})	10^{34}	$7.7 \cdot 10^{33}$	$5.0 \cdot 10^{33}$	$1.4 \cdot 10^{34}$
Integrated recorded luminosity (fb^{-1})	80-120	21.3	3.9	36.5
Average pile-up interactions	19	37	15	25

Table 2.1: An overview of design goal specifications of LHC, with performance-related parameters during LHC operations in 2012, 2015 and 2016.

During the most recent data taking period (Run 2), LHC accelerated protons up to an energy of 6.5 TeV, corresponding to a centre of mass energy of 13 TeV, with a maximum luminosity of $1.4 \cdot 10^{34}$ cm^{-2}s^{-1}. The scheduled periods of active work of LHC and the shutdown periods are represented in Fig.2.2, together with future developments and upgrades, like the Phase-2 (High Luminosity-LHC from 2018) [40].

The goal of the LHC technical program is to increase the number of collisions, together with the integrated luminosity delivered to the detectors. A more detailed explanation of the luminosity concept is described in Appendix A, while in Fig.2.3 the total luminosity delivered by LHC and recorded by ATLAS during 2015 and 2016 data taking periods is shown [41]. As a consequence of the high delivered luminosity, the number of p-p interactions in
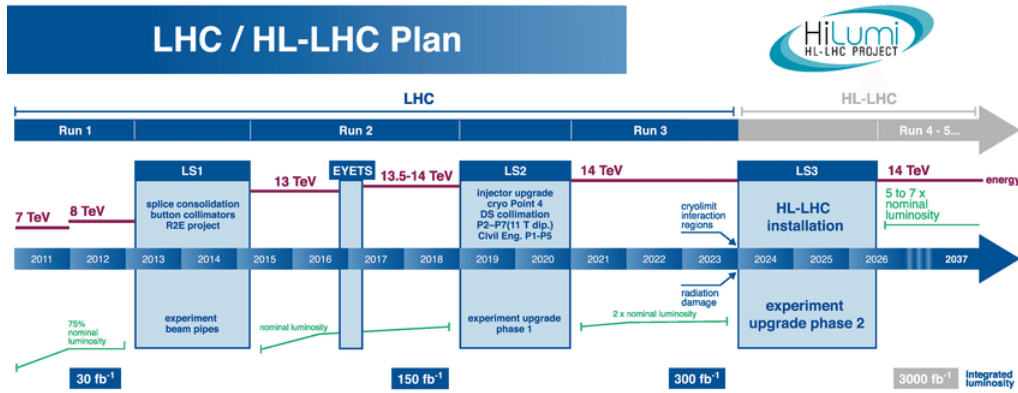
Figure 2.2: Time schedule of LHC, involving active periods and technical shut down, from Run 1 to last upgrade to High Luminosity LHC.

a single bunch crossing ("pile-up") is much larger than one. Fig.2.4 shows the average number of interaction per bunch crossing in 2015 and 2016. The pile-up constitutes a challenge for detectors that must resolve the properties for individual collision.
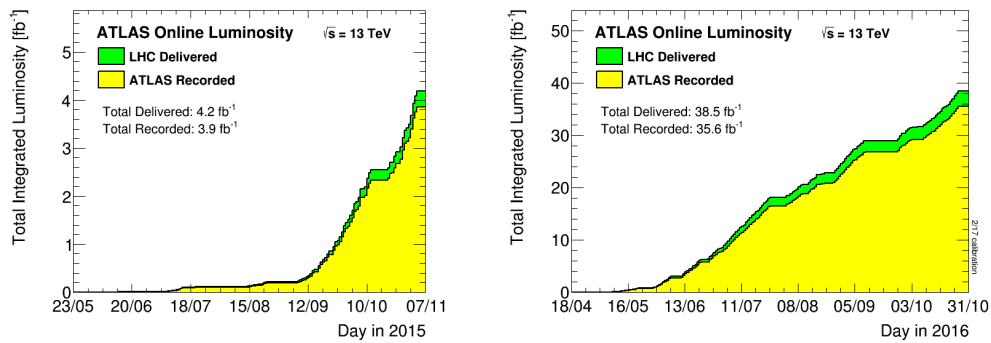
Figure 2.3: Integrated delivered (green) and recorded (yellow) luminosity versus time during stable beams of pp collisions in 2015 (left) and 2016 (right).

2.2 Physics requirements

LHC aims to observe rare physics processes that may be generated at the TeV scale. The high luminosity and the large cross sections enable to
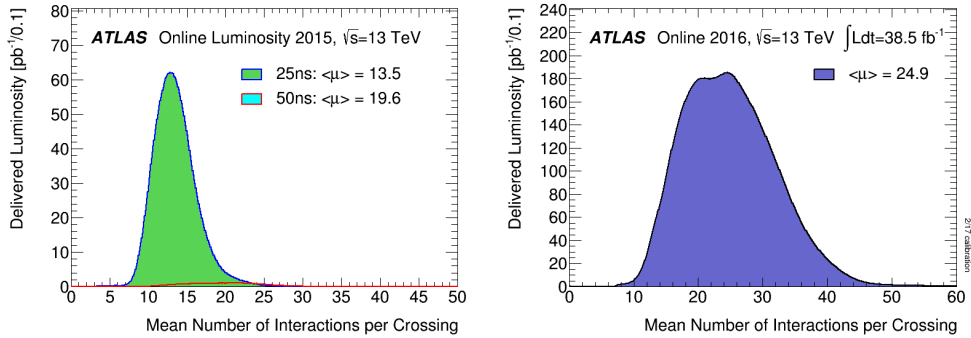
Figure 2.4: Mean number of interaction per bunch crossing during 2015 (left) and 2016 (right) p-p collisions.

increase the production rate of all physics channels. On the other hand, the nature of the p-p collisions leads the QCD jet production to increase as a function of energy, making the reconstruction of rare processes challenging. The design of ATLAS and CMS detectors and the choice of their subsystems is guided by the need to perform high precision measurement of electroweak interactions, to test QCD, to verify the consistency of the Standard Model and to discover new processes.

In order to address the previous physics goals, the LHC detectors must fulfil the following requirements:

- fast response and radiation hard electronics;

- high granularity to handle the numerous signals detected and to separate overlapping events;

- large acceptance for high efficiency particle detection and missing energy measurement;

- tracking system with good momentum resolution for charged particles and large reconstruction efficiency for particle detection and primary (secondary) vertices reconstruction;

- good electromagnetic calorimetry for electron and photon identification, complemented with hadronic calorimetry for jets and transverse energy measurements;

- good muon identification and momentum resolution, with unambiguously determination of the charge;

- highly efficient trigger on both high and low transverse-momentum objects with sufficient background rejection;

- precise measurement of the luminosity for cross section measurements.

The structure of the four experiments situated along the LHC ring is designed to cope with different physics program:

- **ATLAS (A Toroidal LHC ApparatuS)** is a multi-purpose experiment designed to perform precise measurements of the Standard Model and to discover signatures of new physics;

- **CMS (Compact Muon Solenoid)** is a multi-purpose experiment that pursues the same physics goals as ATLAS, using different and complementary technologies;

- **LHCb** placed in the forward direction, investigates the physics of B mesons and CP violation, using a completely different asymmetric design, covering the forward direction for 20 meters along the beam pipe, in order to efficiently reconstruct the B meson decays;

- **ALICE (A Large Ion Collider Experiment)** is dedicated to the study of the quark gluon plasma produced in heavy ion collisions in dedicated LHC runs.

2.3   The ATLAS detector

With a total length of 42 m, a diameter of 22 m and a weight of 7000 t, ATLAS is the largest among the LHC experiments [42]. It is of cylindrical shape, symmetric with respect to the interaction point (IP), with a almost 4π coverage and concentric layers of sub-detectors arranged around the beam pipe. The detector is organised in a central barrel and two end-caps. The overall ATLAS layout is portrayed in Fig.2.5 and the performances of the main sub-systems are summarised in Table 2.2.

Detector component	Required resolution (GeV)	Obtained resolution in 2015 (GeV)	η coverage Measurement	Trigger
Tracker	$\sigma_{p_T}/p_T = 0.05\ \%p_T \oplus 1\%$	$\sigma_{p_T}/p_T = 0.038\%p_T \oplus 1.5\%$	± 2.5	
EM Calorimeter	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.2\%$	± 3.2	± 2.5
Hadronic Calorimeter				
barrel and end-caps	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward (LAr)	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$		$3.1 < |\eta| < 4.9$	$3.1 < |\eta| < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1\mathrm{TeV}$	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1\mathrm{TeV}$	± 2.7	± 2.4
combined with tracker		$\sigma_{p_T}/p_T = 7\%$ at $p_T = 1\mathrm{TeV}$	± 2.7	± 2.4

Table 2.2: General performance of the ATLAS subdetectors, both design and as achieved in 2015.
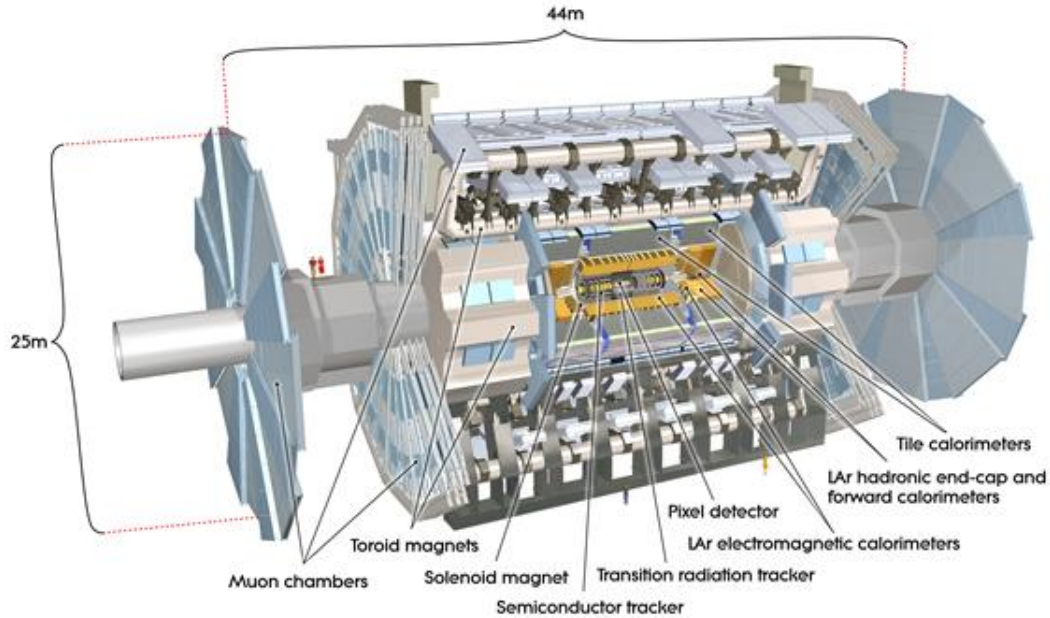
Figure 2.5: The ATLAS detector layout.

The ATLAS detector technologies are designed to properly reconstruct particles according to their interaction with the materials, which is deeply linked to the nature of the particles themselves. A complete representation of the particles reconstructed and identified in ATLAS is presented in Fig.2.6. All charged objects, such as protons, electrons, muons and $\pi^{\pm}$, deposit a part of their energy in the innermost section of ATLAS, defined by the presence of the tracking system (Inner Detector). A magnetic field bends particles and allows to reconstruct their momentum and to measure their charge from the bending direction. All particles, neutral ones included, deposit their energy in the calorimeters. ATLAS has two different technologies for calorimetry, specifically developed to measure the energy of electrons, photons and hadrons. The only exception is given by muons, which cross all the sub-systems depositing only a small fraction of the total energy, before being stopped in dedicated spectrometer, placed after the hadronic calorimeter. Neutrinos are not detected in ATLAS, since they do not interact with any of the sub-systems mentioned previously. Their presence is deduced by

the momentum balance in the recorded event. A brief description of the AT-
LAS subdetectors follows in the next paragraphs of this Section, focusing on
the different technologies and efficiencies. A complete description of particle
reconstruction, is presented in to Section 2.4.
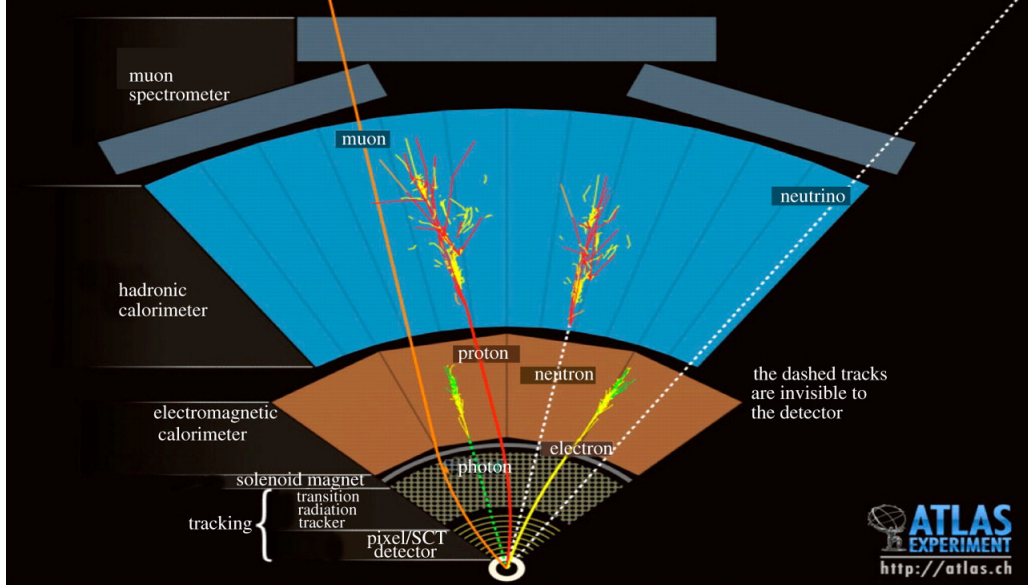
Figure 2.6: Different reconstruction of particles in ATLAS according to their
different interactions with materials.

2.3.1   The coordinate system

ATLAS uses a right-handed coordinate system with the origin in the nom-
inal interaction point (IP) and the z-axis directed along the beam-pipe. The
x-axis points to the centre of the LHC ring and the y-axis points skywards.
The side-A of the detector is defined for positive z and side-C for negative
z. Given the detector symmetry, cylindrical coordinates (r, ϕ, θ) are used,
being ϕ the azimuthal angle around the z-axis and θ the polar angle.
The rapidity y is defined as:

$$y = \frac{1}{2} \log \frac{E + p_z}{E - p_z} \quad .$$

(2.1)

In case of particles with a mass negligible with respect to the energy, y corresponds to the pseudorapidity η:

$$\eta = -\log[\tan(\frac{\theta}{2})] \quad . \tag{2.2}$$

The transverse momentum p_T, the transverse energy E_T and the missing transverse energy E_T^{miss} are defined in the x-y plane. ΔR is the distance in the $\eta - \phi$ space defined as: $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

2.3.2 The tracking system

Particle tracking is performed in ATLAS by the Inner Detector (ID), the detector closest to the interaction point. It is composed by concentric layers of detecting material and has a total radius of 1.2 m and length of 6.2 m [43]. This corresponds to a pseudorapidity region of $|\eta| < 2.5$, beyond which particle detection must rely solely on the calorimeters and muon spectrometer. The main purpose of the ID is to reconstruct charged tracks and momenta. Since the ID is immerse in a 2 T magnetic field generated by a central solenoid, from particle bending it is possible to achieve the measurements of the charge and momentum. By extrapolating tracks to their origin point in the beam pipe, interaction vertices are reconstructed: in particular, the primary vertex corresponds to the vertex where p-p collisions take place, while secondary vertices correspond to the position of heavy particle decay. The ID provides a transverse impact parameter resolution of ~ 35 (~ 10) μm for pions with $p_T =$ 5 (100) GeV and a transverse momentum resolution of about 4% for 100 GeV muons.

The ID is made of several sub-detectors, characterised by different technologies (see Fig.2.7). Starting from the inner layer: the Insertable B-Layer, the Pixel and the Silicon microstrip of the Semi Conductor Trackers, used in conjunction with the Transition Radiation Tracker. The main sub-systems devoted to particle tracking are described in detail in the following sections.
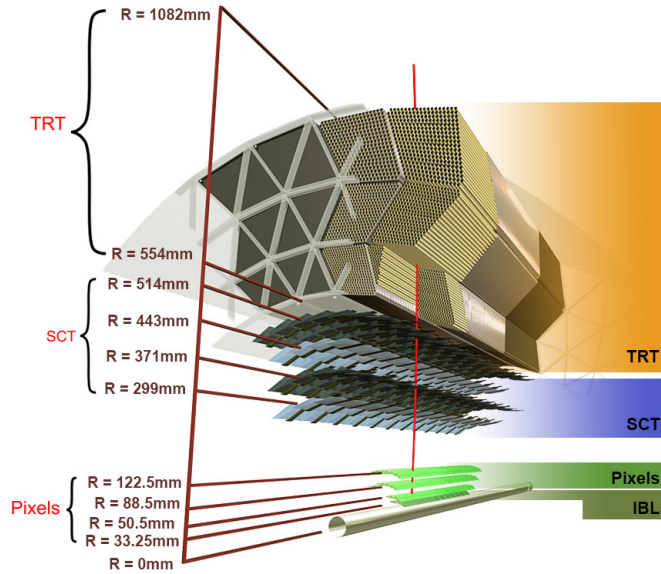
Figure 2.7: An illustration of the ATLAS Inner Detector, including the new insertable B-layer (IBL) added between Run 1 and Run 2. The red line indicates the trajectory of a hypothetical particle with p_T = 10 GeV and η = 0.3.

Insertable B-Layer (IBL)

The IBL is the sub-detector closest to the beam pipe. It is made of a single cylindrical layer of silicon pixels. It is placed between the thinner Beryllium beam pipe and the inner Pixel layer. The radiation length at η=0 is 1.54% of X_0. Due to its position close to the beam pipe, the IBL pixel sensors have a small size (50×250 μm) and are radiation hard, to reduce occupancy at high luminosity. It provides a full ϕ coverage, with a precision in the coordinate measurements of 23 μm. Inserted during the LS2, the IBL improves the quality of the impact parameter measurement, of the vertex reconstruction and of the b-tagging performance. In particular, in case of the b-tagging, the IBL reduced the probability of mis-tagging particles by a factor two [44, 45].

The pixel detector

Due to its position close to the barrel where the particle density is at its highest, the Pixel detector has a thin granularity. It is composed by three layers of silicon pixels 50μm×400μm (ϕ×z) in size, in both the barrel and
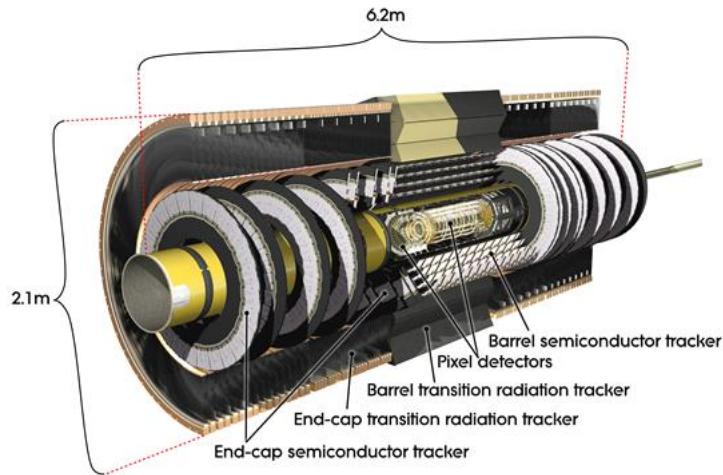
Figure 2.8: View of the ATLAS Inner Detector. The IBL is missing in the picture since inserted during the last long shutdown.

the end-cap positions. The innermost barrel layer is located at R=51 mm and the outermost at R=123 mm from the z-axis, while the first end-cap disk is positioned at $|z|$=495 mm and the last at $|z|$=650 mm. The dimensions are chosen in order to maximise the probability that a particle crossing one layer will cross also the other two. The system covers a total active area of about 1.7 m^2 with an approximate radiation length of 0.1 X$_0$ at $\eta = 0$. The intrinsic precision in the measurements is 10 μm for the R-ϕ plane and 115 μm for the z(R)-coordinate in the barrel (end-caps).
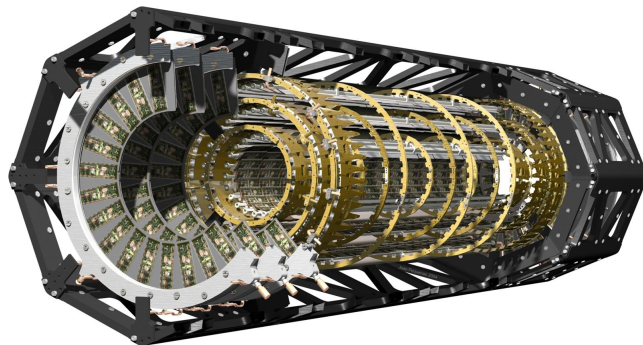
Figure 2.9: Cutaway of the ATLAS Pixel Detector.

SemiConductor Tracker (SCT)

Placed in the intermediate range of the ID, the SCT employs the same semiconductor technology as the Pixel detector, with the difference in the use of silicon microstrips instead of pixels (120 mm×60 mm in ϕ×z). It is composed by four layers in the barrel (300 mm<R<510 mm) and nine in each end-cap (850 mm<z<2730 mm). In each module, strip planes are glued back to back, with the inner strip aligned along z in the case of the barrel and radially in the case of the end-caps. Each of the 4088 modules incorporates two layers with strips rotated by 40 mrad for a more accurate position measurement: a single strip can indicate the transverse location of a hit (ϕ and R in the barrel, for example) and the 40 mrad stereo angle adds the determination of the longitudinal coordinate (z in the barrel). The radiation length at $\eta = 0$ is approximately 0.1 X_0. The intrinsic measurement accuracies per module are 17 μm for the R-ϕ plane and 580 μm for the z(R)-coordinate in the barrel (end-caps).

Transition Radiation Tracker (TRT)

Positioned in the outermost part of the ID, the TRT is a combination of a tracker (based on the straw tubes) and a Transition Radiation detector for the pattern recognition. The Transition radiation detector allows to discriminate between light and heavy particles from the transition radiation that particles emit according to their speed, passing through several layers of material with different refraction indices. High relativistic particles (typically electrons) can be recognised by a wider emitted radiation with respect to the other particles. A single TRT component is made of Polyimide drift (straw) tubes of 4 mm diameter that contains the anodes, tungsten wires gold plated, directly connected to the front-end electronics and kept at ground potential. The gap between the straw and the wire is filled by a mixture of gases. The passage of ionizing particles induces a low amplitude signal on the anodes. At the same time, some particles crossing polypropylene fibres cause transition radiation emission, absorbed by the Xenon present in the gas mixture; this last process leads to a high amplitude signal in the TRT electronics that can be distinguished from low amplitude ionization signal.

A large number of hits is provided by the straw tubes of the TRT, typically 36 for tracks up to $|\eta| = 2.0$. The TRT only provides R-ϕ information, for which it has an intrinsic measurement accuracy of 130 μm per straw. In the barrel region, straws are parallel to the beam axis and are 144 cm long, with their wires divided into two halves, approximately at $|\eta| = 0$. In the end-cap region, the straws are arranged radially in wheels. The radiation length at $|\eta| = 0$ is approximately 0.2 X_0.

2.3.3   The magnet system

A strong magnetic field is crucial to provide sufficient bending power for accurate track reconstruction and momentum measurement. ATLAS uses a hybrid system of four large superconducting magnets, unique in the history of particle detectors, that gives to ATLAS its original shape. Fig.2.10 shows the general layout of the ATLAS magnetic system [43], consisting of:
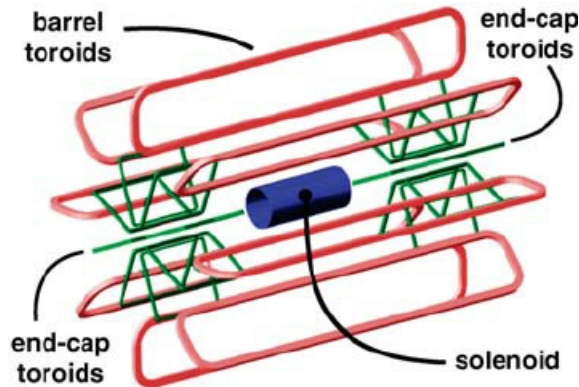
Figure 2.10: The ATLAS magnet system.

- a **solenoid**, 5.3 m long and 2.5 m of diameter, with the axis parallel to the beam direction. It surrounds the Inner Detector providing a 2 T magnetic field. The layout was designed to keep the material thickness in front of the calorimeter as low as possible: the solenoid contributes a total of ∼0.86 radiation lengths (X_0) at normal incidence;

- a **toroid** system, providing a 4 T field, mostly orthogonal to muon trajectories. As visible in Fig.2.10, it is composed by eight Barrel Toroids

(BT) 25 m long (inner core 9.4 m, outer diameter 20.1 m) and two
End-Cap Toroids (ECT), 5 m long (inner core 1.64 m, outer diameter
10.7 m). The toroid structure is designed to minimize the uncertainty
on the momentum measurements due to multiple scattering. The BT
bends particles in the region $|\eta| < 1$, while ECTs provide particle bend-
ing in $1.4 < |\eta| < 2.7$. In the transition region, $1 < |\eta| < 1.4$, magnetic
deflection is provided by a combination of barrel and end-cap.

2.3.4 The calorimetry system

The main task of the calorimeters is to fully contain and measure the
energy of the incident particles. The only exception is given by muons, as
explained in Section 2.3.5.
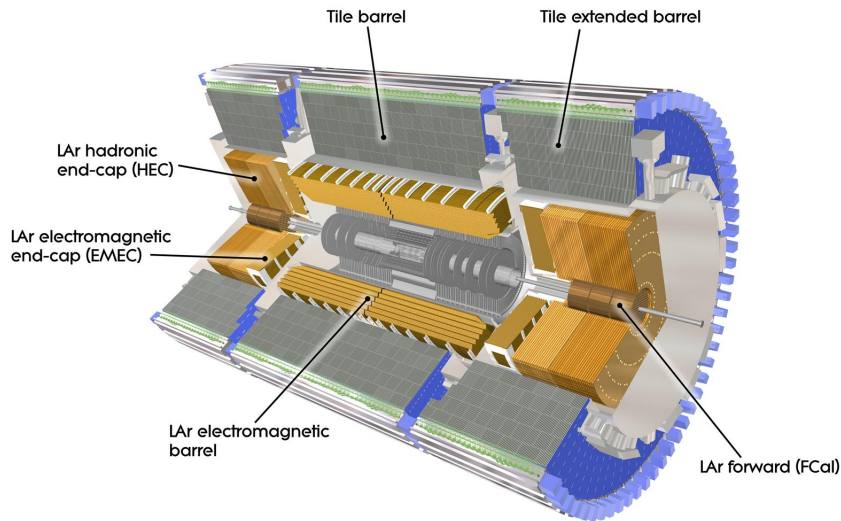
Figure 2.11: Overall overview of the ATLAS calorimetry system.

A view of the ATLAS calorimetry system is shown in Fig.2.11. Different
techniques are used in order to address the wide range of physics processes
of interest. The calorimeters cover a range of $|\eta| < 4.9$ [43]. The central parts
have high granularity to enable the pointing back to the primary vertex, al-
lowing the reconstruction of the interacting particle direction. The ATLAS

calorimeter system consists of two separate sub-systems: the electromagnetic and the hadronic calorimeters, structured in three cryostats, one barrel and two end-caps. Surrounding the ID, the electromagnetic calorimeter uses liquid argon (LAr) as active material and lead as absorber to measure the energy of electrons and photons. The hadronic calorimeter employs the LAr technology and a copper absorber in the forward region and active scintillating tiles with absorbing steel in the barrel (TileCal), to contain the shower of jets. The forward calorimeter (FCAL) uses LAr and copper in the electromagnetic part and tungsten in the hadronic part.

The main requirement for a calorimeter in order to obtain a good transverse energy resolution is hermeticity. The large radius of the toroidal magnet muon system allows the calorimeter to have enough thickness to achieve a good containment for jets. The total thickness of the electromagnetic calorimeter is more than 22 radiation length (X_0) in the barrel and more than 24 X_0 in the end-caps. The total thickness of the hadronic calorimeter is 9.7 interaction length (λ) in the barrel and 10 λ in the end-caps, adequate to fully contain high energy jet showers. The total interaction length of the entire system is $\sim$10 λ, with a peak values of 11 λ at η=0. The main features of the electromagnetic and hadronic systems are described in the following paragraphs.

The electromagnetic calorimeter

The LAr electromagnetic calorimeter is a sampling calorimeter divided into a barrel ($|\eta|$ <1.475) and two end-caps (1.375< $|\eta|$ <3.2) regions. In front of the cryostat wall, an additional pre-sampler detector, consisting of an active layer of liquid argon of 1.1 cm thickness, provides a correction to the energy loss in the material in front of the calorimeter ($|\eta|$ <1.8). The barrel calorimeter is composed by two parts with a 6 mm gap at z=0; the end-caps (EMEC) are divided into two coaxial wheels, covering up to $|\eta|$ <2.5 and 3.2 respectively. The accordion geometry of the electrodes, in both the barrel and end-cap regions, guarantees a complete azimuthal coverage. Fig.2.12 shows a module of the barrel calorimeter and highlights the sampling structure, which is organised in three layers [43]:

- the inner layer (4.3 X_0) consists of strips with $\Delta\eta$=0.0031 to allow separation between charged and neutral pions;

- the middle layer (16 X_0) is segmented in squared towers $\Delta\eta \times \Delta\phi = 0.025\times0.025$;

- the outer layer is made by $\Delta\eta \times \Delta\phi = 0.050\times0.025$ towers, specifically designed to measure showers of electrons and photons with energy E> 50 GeV.

The overall segmentation allows for high precision spatial measurements, providing a pointing geometry that helps in identifying photons coming from a primary vertex. The LAr calorimeter allows to achieve an energy resolution, independent on η for photons of 100 GeV better than 1.5% ($\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.2\%$).
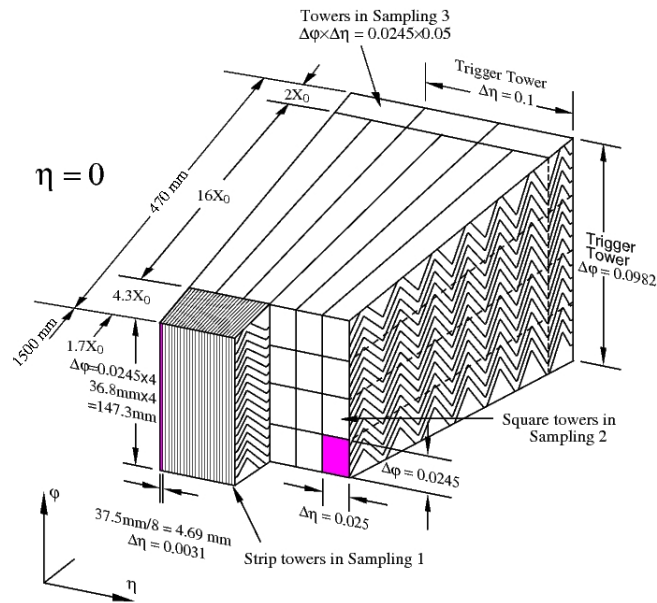

Figure 2.12: Sketch of a LAr barrel module. The granularity in η and ϕ cells of each of the three layers and of the towers is also shown.

The hadronic calorimeter

The hadronic calorimeter is placed directly outside the EM calorimeter and it is compounded by a barrel and two end-caps.

The barrel region (TileCal) is a sampling calorimeter using iron as the absorber and scintillating tiles as active material, with a granularity of $\Delta\eta \times \Delta\phi$ = 0.01×0.01. It is divided into a central ($|\eta|$ <1) and two extended barrels (1< $|\eta|$ <1.7). TileCal is segmented in three layers, approximately 1.5, 4.1 and 1.8 λ thick for the central barrel, and 1.5, 2.6 and 3.3 λ for the extended barrel. The total detector thickness at the outer edge of the tile-instrumented region is 9.7 λ at $\eta = 0$. Gap scintillators are placed along the internal edge of the extended barrel to partially recover the energy loss between the two barrel regions, where the readout of the EM calorimeter is placed.

The hadronic end-cap (HEC) covers the pseudorapidity region 1.5< $|\eta|$ <3.2. It is located in the same cryostat as the EMEC and the forward calorimeter (FCal) and its absorbing material is copper. Each hadronic end-cap is composed of two independent wheels with $\sim$2 m outer radius; each wheel is segmented longitudinally in two parts.

TileCal and HEC measure jet energies with a resolution $\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 2.5\% \oplus \frac{5\%}{E}$, where the energy E is in GeV.

Forward calorimeter

The Forward Calorimeter (FCal) is an electromagnetic and hadronic calorimeter, made using an intrinsic radiation hard technology. The FCal is approximately 10 interaction length deep, and consists of three modules in the end-cap: the first is optimised for electromagnetic measurements and uses copper as passive material, while the two others, with tungsten as passive material, measure the energy of the hadronic interactions with a energy resolution is $\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%$, where the energy E is in GeV.

2.3.5 The muon system

The muon spectrometer is designed to reconstruct muons, which travel in the detector much more than the other charged particles generated in the collisions, since they hardly interact with materials, they radiate bremsstrahlung far less than the electrons and are long-lived particles. The features mentioned above explain the reason why the muon systems is located in the outermost part of ATLAS, within the toroidal magnetic field and covers the

pseudorapidity region $|\eta|$ <2.7. The direction of the magnetic field causes muons to bend in the $R-z$ plane, orthogonal to the solenoid field in the ID. An illustration of the muon spectrometer can be found in Fig.2.13.
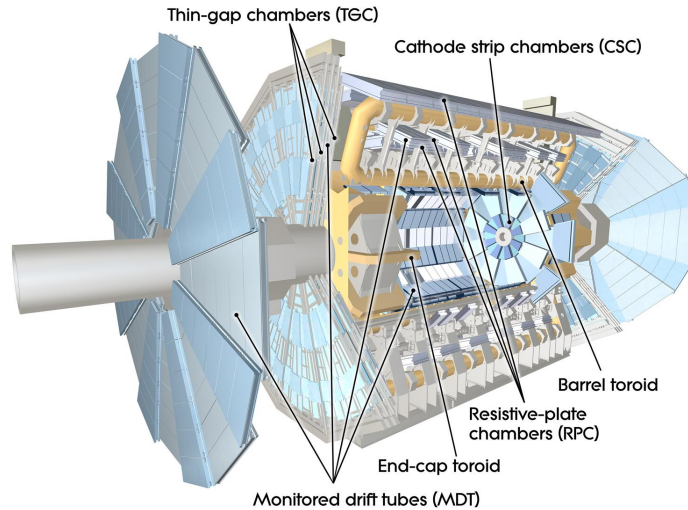
Figure 2.13: Cutaway view of the ATLAS Muon System.

The muon spectrometer is composed of several sub-detectors, grouped according to two main features:

a) chambers providing precision measurement of the momentum;

b) chambers with fast response for on-line triggering and coarser granularity.

The type a) sub-detectors are: the Monitored Drift Tube (MDT) chambers and the Cathode Strip Chambers (CSC). The **MDT**s form the bulk of the muon spectrometer, with several layers of drift tubes, operated with Ar/CO_2 gas at 3 bar pressure. The MDT layout is projective, increasing the layer dimensions and chamber sizes as a function of the distance from the interaction point. They cover the pseudorapidity region $|\eta|$ <2.7, except for the innermost end-cap layer which covers $|\eta|$ <2.0. MDTs are complemented by **CSC**s, characterized by a better time resolution. CSCs are multi-ware proportional chambers, containing cathode planes segmented in strips along the orthogonal direction to allow for measurements in both the bending and

the transverse planes. They are composed by two disks, with eight chambers each. All relevant parameters of the Muon Spectrometer sub-detectors are summarised in Table 2.3.

Type	Function	Chamber resolution in			Hit/track		Number of	
		z/R (μm)	ϕ (mm)	time (ns)	barrel	endcap	chambers	channels
MDT	tracking	35 (z)	-	-	20	20	1088	339k
CSC	tracking	40 (R)	5	7	-	4	32	30.7k
RPC	trigger	$10 \cdot 10^3$ (z)	10	1.5	6	-	544	359k
TGC	trigger	$2\text{-}6 \cdot 10^3$ (R)	3-7	4	-	9	3588	318k

Table 2.3: Main parameters of the four sub-detectors constituting the ATLAS Muon Spectrometer.

The detectors of type b) provide a ϕ coordinate measurement, completing the MDT precision. Benefiting from a very fast response, they provide the on-line event selection (trigger), identifying the bunch crossing originated an event with 99% accuracy. In the barrel region ($|\eta| < 1.05$) they consist of Resistive Plate Chambers (RPC) and in the end-caps ($1.05 < |\eta| < 2.4$) of Thin Gap Chambers (TGC), whose relevant parameters are highlighted in Table 2.3. The **RPC** sub-system has planes arranged in three cylindrical layers around the z-axis; in order to measure the η and ϕ coordinates, each plane is further composed of two layers. The **TGC** sub-detector complements the MDTs, measuring the azimuthal coordinate. TGCs are segmented radially into one end-cap and forward region. The TGC wire groups measure the bending plane coordinate and the radial strips provide the ϕ measurements. The precision of the momentum measurement for a high-p$_T$ muon track depends on the resolution of the sagitta, namely the deviation in the R-z plane with respect to a straight line. For a high-momentum track (p$_T$ $\sim$1 TeV), the typical sagitta is around 500 μm. The MDTs provide a momentum measurement with $\sigma_{p_T}/p_T \sim 10\%$ resolution for 1 TeV muons, and 2-3% for lower momenta. In case of low-momentum muons, it is necessary to complement the measurement of the MS with information from the Inner Detector (see Appendix B).

2.3.6 The forward detectors

Additional detectors are displaced from the ATLAS structure to provide a coverage in the high rapidity region [46]. Fig.2.14 illustrates the position of the forward detectors with respect to their distance from the IP: the detector closest to ATLAS is LUCID (Luminosity measurement Using Cherenkov Detectors), followed by ZDC (Zero Degree Calorimeter), AFP (ATLAS Forward Proton) and ALFA (Absolute Luminosity For ATLAS).
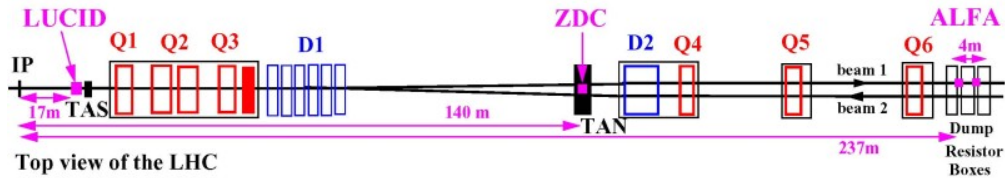
Figure 2.14: Forward detector positions along the beam line with respect to the ATLAS interaction point (IP).

Placed ± 17 m from the IP, **LUCID** is a detector built to measure luminosity. LUCID is the official ATLAS luminosity monitor since the beginning of Run 2 and is described in detail in Appendix A.

ZDC [47] is located ± 140 m from the interaction point, where the beam pipe is divided into two separated tubes (see Fig.2.14) and embedded in the TAN (Target Absorber Neutral). ZDC is designed to detect neutrons from heavy-ion collisions with $|\eta| > 8.3$ and extract from the number of spectator (forward) neutrons the centrality of the interaction. It is also used to reduce the beam-halo and beam-gas backgrounds by requiring a coincidence between the two ZDC modules, symmetric around the IP. Its time resolution is approximately 100 ps, allowing the determination of the IP with a 3 cm precision in z coordinate. Each side of the detector contains four modules: one electromagnetic and three hadronic calorimeters. The EM module has 11 tungsten plates, normal to the beam axis, each vertically extended by 290 mm steel plates. Quartz rods enter the plates parallel to the z direction, forming a 8×12 matrix. They are read out by phototubes, to capture the

Cherenkov light emitted by the incident particles. The hadronic modules have a similar layout, with fewer readout channels.

AFP [48] is designed to measure forward protons with Roman Pots for the study of soft and hard diffractive processes. It consists of 2 stations on each side: a tracker is placed at 205 m and another tracker with a ToF (Time of Flight) detector is located 217 m from the IP. The tracker is a pixel detector with high granularity 10 μm$\times\mu$m (x-y) and radiation hard. The ToF detector measures the time of flight of the colliding particles with a time resolution $\leq$30 ps.

ALFA [49] is located at $\pm$240 m from the interaction point. It is a system of detectors dedicated to the measurement of elastic scattering at very small angles (down to Coulomb-Nuclear interference region) for the total cross section determination and diffractive studies. The angles needed to perform such measurements ($\sim$3 μrad) are smaller than the beam divergence: therefore, special beam condition are required (large values of the betatron function at the IP, $\beta^* \sim$[90m , 2.5 km]). ALFA is composed by scintillating fibre trackers and it is placed inside Roman Pots, which allow the sub-detectors to be as close as possible to the beam ($\sim$1 mm). There are two Roman Pots on each side at 4 m distance from each other. ALFA spatial resolution is 25$\pm$3 μm.

2.3.7   Trigger and data acquisition system

The trigger system is an essential component of the ATLAS experiment, since it is the responsible for deciding whether or not save a given collision among the large number of data collected. ATLAS uses a complex Trigger and Data Acquisition System (collectively TDAQ) to select the events and reduce the event rate from 40 MHz to $\sim$1 kHz, the maximum for data storage. The TDAQ system [50] consists of a hardware Level-1 and a software-based High Level Trigger (HLT) levels: each level refines the selection made in the previous step by adding additional information and selection requirements (see Fig.2.15).

The initial selection is made by the Level-1 trigger, which reduces the event rate from the LHC bunch crossing of $\sim$40 MHz to $\sim$100 kHz, with a
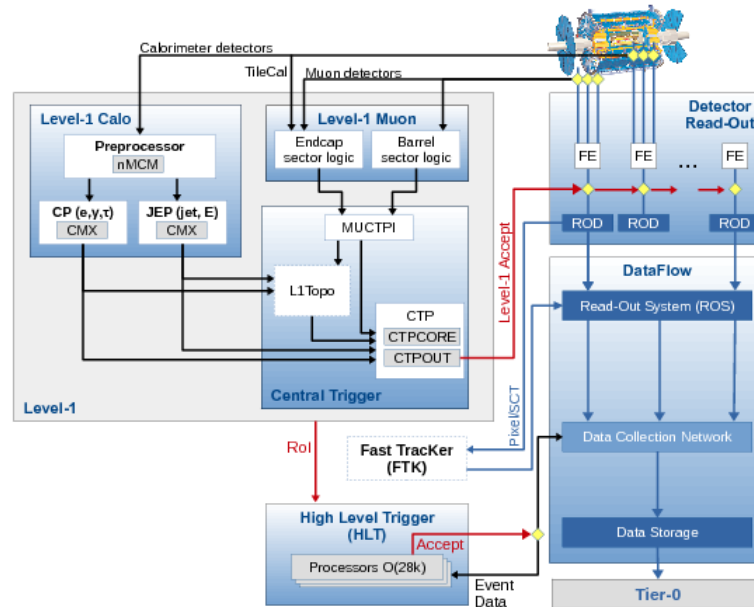
Figure 2.15: Scheme of the ATLAS Trigger and Data Acquisition System.

decision time of 25 μs. The Level-1 trigger uses custom electronics to determine the Regions of Interest (RoIs) in the detector, taking as input coarse granularity calorimeter and muon detector information. The RoIs represent the position of the triggering object in the $\eta - \phi$ plane and are sent to the HLT, which performs a full event reconstruction by using sophisticated algorithms adapted to on-line selection algorithms. The HLT is the responsible for the final physics selection for the following off-line analyses.

The HLT reduces the rate from Level-1 output to approximately 1 kHz, with a processing time of $\sim$200 ms and an event size of about 2 Mbyte. In Run 1 the HLT has separate Level-2 and Event Filter computer clusters. In Run 2, the system has been merged into a single event processing HLT form, in order to reduce the complexity and allow the dynamic source sharing between algorithms. This arrangement reduces algorithm duplication and results in a more flexible HLT.

2.4   Object reconstruction

In ATLAS different types of particles must be reconstructed, such as electrons, muons, taus, photons, jets and neutrinos, exploiting the technologies described in Section 2.3. Since ATLAS is a complex system, made of several components of different technologies, a particle is reconstructed using signals deposited in many different parts of the detector. The first step in the reconstruction is to build the low-level objects representing individual particles: in the Inner Detector, tracks are reconstructed from spacepoint hits and in the calorimeters, cells clusters are formed. From tracks and calorimeter clusters, electrons, muons, taus, photons and jets are constructed. All particles mentioned above are then used to measure the total transverse energy, and through the energy balance, the so-called "missing transverse energy", is attributed to neutrinos. In the following paragraphs, a descriptions of track, electron, muon, jets and missing transverse energy reconstruction is given.

2.4.1   Tracks and primary vertex

Tracks are reconstructed from individual particle interaction with the detector and reveal the passage of charged particles through the ID. Tracks are used both to identify the particles produced in the collision and to locate the primary vertex, by extrapolating their path to the beam line.
Tracks are reconstructed from *spacepoints*, which are three dimensional representations of detector measurements. ATLAS uses two different algorithms for tracking: the "inside-out" and the "outside-in", whose names refer to the order in which spacepoints are connected to form tracks. First the inside-out method seeds the track with silicon hits (pixels and SCT hits) before extrapolating outwards and find a coincident hit in the TRT; then the outside-in procedure uses TRT hits to seed tracks, which are extrapolated back towards the IP to identify coincident silicon hits. Pixel spacepoints are simple to determine due to the high granularity of the pixel detector. This is the reason why, as explained in Section 2.3.2, two SCT microstrips are glued together at a small stero angle to allow the determination of the coordinate orthogonal to the beam direction. A detailed explanation of the algorithm steps used to construct tracks is reported in [51].

Tracks are reconstructed with a resolution in the transverse momentum measurement $\sigma_{p_T}/p_T = 0.05\% \times p_T/GeV \oplus 1\%$. The reconstruction algorithm fits five parameters: d_0, z_0, ϕ, θ and p/q, where d_0 and z_0 are the impact parameters to the beam line in the transverse and the longitudinal planes, respectively, q is the track charge and p is the track momentum.
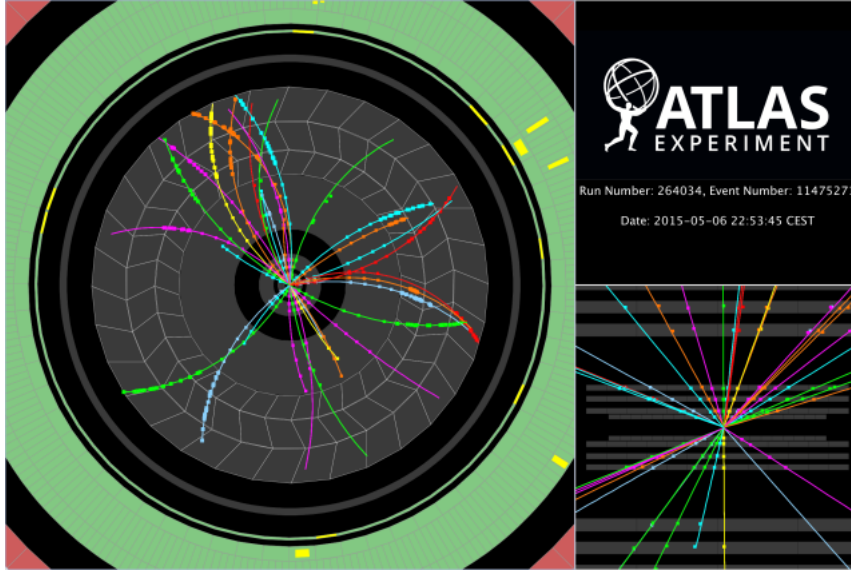
Figure 2.16: Display of a proton-proton event collision recorded by ATLAS on 6th May 2015 at 900 centre of mass energy. Tracks are reconstructed from hits in the ID, including the IBL, shown as a small ring in the left-hand azimuthal view, and in the innermost layers in the right hand longitudinal view.

Due to the large hit multiplicity, fake track candidates (tracks not associated with the trajectory of an actual particle) are present and must be rejected. Moreover, if n$\sim \mathcal{O}(10)$ particles are produced within the ΔR$<$0.02 radius jet, for example, then multiple particles can hit the same pixel and one of the resulting tracks can be lost. The fitting procedure and the requirement of good quality criteria allow to discard many hits and to correctly reconstruct particle directions. Fig.2.16 shows an event display from the early Run 2 data, where tracks are reconstructed from all three ID sub-systems including the IBL [52].

Primary Vertex Reconstruction

The **primary vertex** is the location where the p-p interaction takes place. In general, vertices are reconstructed by extrapolating tracks to the beam pipe and looking for intersection of multiple tracks [51]. First of all, tracks must satisfy certain quality criteria, such as $p_T > 400$ MeV, ≥ 9 silicon pixel hits with no pixel holes (when particle passes through inert material or dead modules) and $|d_0| < 4$ mm. The constrain on d_0 prevents the reconstruction of vertices outside the beam pipe. Then, vertex finding is seeded using the z coordinate of each extrapolated track with the beamline. Tracks within 7σ of the seed are candidates for the vertex reconstruction.

To determine the vertex position iterative χ^2-fits are made, where in each iteration outlying tracks have lower weights, in order to minimise the χ^2. The vertex with the highest sum of tracks p_T^2 is taken as the primary vertex, while the other vertices are assumed to be pile-up vertices. The overall vertex reconstruction efficiency is above 80% for beam crossings with few vertices and 50% for crossing with more than 40 interactions.

2.4.2   Electrons

In ATLAS electrons are reconstructed using a combination of the Inner Detector tracks and Electromagnetic Calorimeter clusters. The information from the ID define the direction at the interaction point, while the EM clusters infer the energy of the electron.

The reconstruction algorithm begins finding the seed-clusters as longitudinal towers in the $\eta \times \phi$ plane of the EM middle layer, with total cluster energy above 2.5 GeV [53]. Then the reconstruction methodology involves track reconstruction from the ID, taking into account energy loss due to bremsstrahlung, the number of precision hits in the SCT and the track-cluster EM matching.

The Particle IDentification (PID) determines whether the reconstructed electron candidates are signal-like objects or background-like objects, such as hadronic jets or converted photons. The PID algorithms use quantities related to the electron cluster and track measurements, including calorimeter shower shapes, information from the TRT, track-cluster matching related

quantities, track properties and variables measuring bremsstrahlung effects. In Run 2, several changes in the input variables used for electron PID have been introduces: taking advantage of the IBL, the number of hits in the innermost pixel layer is used to discriminate between electrons and converted photons. The baseline PID algorithm used in Run 2 analyses is the likelihood-based (LH) method. It is a multivariate (MVA) technique that simultaneously determines several properties of the candidate when a selection decision is made. The LH procedure uses the signal and background probability density function of the discriminating variables. Based on these functions, an overall probability is calculated for the object to be signal or background.

Three working points are typically provided for the electron PID: ordered by background rejection power, they are referred to as *Loose*, *Medium* and *Tight*. Each working point includes the tighter one. The PID operating points are optimised in several bins of η and E_T, since the distributions of electron showers depend on the amount of material the electron crosses and on their energy. The performance of the LH identification algorithm is illustrated in Fig.2.17 for electrons from Z decay: the signal efficiencies for electron candidates with $E_T= 25$ GeV are in range from 78% (tight) to 92% (loose) and increase with E_T [53].

In addition to the ID criteria described above, most analyses require electrons to fulfil isolation requirements, to further discriminate between signal and background. The isolation variable quantifies the energy of the particles produced around the electron candidate and allow to distinguish prompt electrons from heavy resonance decays (i.e. $W \to e\nu$ and $Z \to ee$) from non-isolated candidates, such as electrons coming from converted photons produced in (heavy) hadron decays and light hadrons misidentified as electrons. The discriminating variables designed for this purpose are:

- the *calorimetric* isolation energy $E_T^{cone0.2}$, defined as the sum of transverse energy of topological calibrated clusters in EM calorimeter, within a cone of ΔR=0.2 around the candidate electron cluster;

- the *track* isolation $p_T^{varcone0.2}$, defined as the sum of transverse momentum of all tracks, satisfying quality requirements [53], within a cone
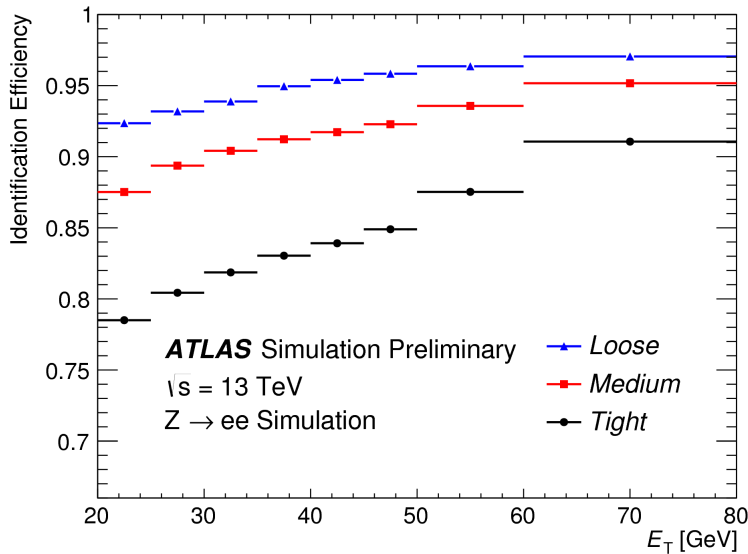
Figure 2.17: Electron identification from $Z \to ee$ decays for Loose, Medium and Tight operating points. The efficiencies are obtained using Monte Carlo simulations and are measured comparing the reconstructed electrons to Monte Carlo truth electrons.

of $\Delta R = \min(0.2, 10\text{ GeV}/E_T)$ around the candidate electron track and originated from the reconstructed primary vertex of the hard collision.

A variety of selection requirements on the quantities $E_T^{cone0.2}/E_T$ and $p_T^{varcone0.2}/E_T$ have been defined to select isolated electron candidates. The resulting operating points can be divided into two classes:

- efficiency target operating points: various requirements are used in order to obtain a given isolation efficiency, which can be either constant or a function of E_T. In the inclusive Z cross section measurement presented in Chapter 3, the *Gradient* isolation operating point is chosen. In the *Gradient* isolation, electrons must satisfy p_T dependent isolation requirements, using both tracking detector and calorimeter information. The isolation efficiency is at least 90% for $p_T > 25$ GeV, reaching 99% at 60 GeV.

- fixed requirement operating points: the upper thresholds on the isolation variables are constant. In the Z+b(b)-jets analysis presented in

this thesis, the *FixedCutTight* criterion ($p_T^{varcone20/p_T}$ <0.06 and E_T^{cone20}/p_T <0.06) is used.

.

2.4.3 Muons

Muon reconstruction is performed independently in the ID and MS, then the information from individual sub-detectors is combined to form the muon tracks used in the physics analyses. Muon reconstruction in the MS begins searching for hit patterns inside each muon chamber to form segments. In each MDT chamber and nearby trigger chamber, the Hough transform pattern recognition algorithm [54] is used to search for hits aligned to a trajectory in the bending plane of the detector. The MDT segments are reconstructed by performing a straight-line fit to the hits found in each layer. Segments in the CSC detectors are built using a separate search in the η and ϕ detector planes. Muon tracks are then built by fitting together hits from segments in different layers.
The ID-MS muon reconstruction is performed according to various algorithms based on the information from the ID, MS and calorimeters [55]. Four muon types are defined according to which sub-systems are used in the reconstruction (see Fig.2.18):

- *Combined* (CB) muon: track reconstruction is performed independently in the ID and MS, and a combined track is formed with a global fit that uses the hits from both the ID and the MS sub-systems;

- *Segment-Tagged* (ST) muon: a track in the ID is classified as a muon if, once extrapolated to the MS, it is associated with at least one local track segment in the MDT or CSC chambers;

- *Calorimeter-Tagged* (CT) muon: a track in the ID is identified as a muon if it can be matched to an energy deposit in the calorimeter compatible with a minimum-ionizing particle;

- *Extrapolated or Standalone* (ME) muons: the muon trajectory is reconstructed based only on the MS track and a loose requirement on

compatibility with origin from the IP.
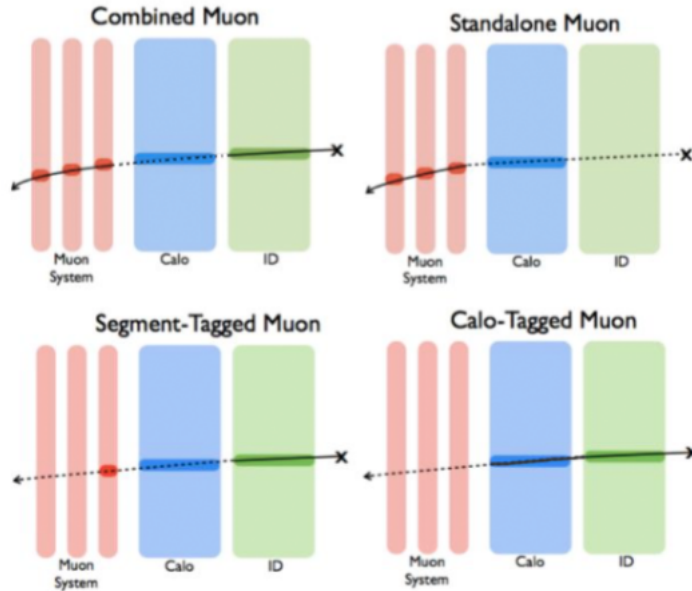
Figure 2.18: Schematic view of the four reconstruction types of candidate muons, defined according to the detection in different subdetectors.

Muon identification is performed by applying quality requirements that suppress background, mainly from pion and kaon decays, and select prompt muons with high efficiency and robust momentum measurement. Four muon identification operating points are provided to address the specific needs of different physics analyses: *Loose*, *Medium*, *Tight* and *High-p_T*. Loose, Medium and Tight are inclusive categories, such that muons identified with tighter requirements are also included in the looser ones.

- The *Loose* identification criteria are designed to maximise the reconstruction efficiency and provide good-quality muon tracks. All muon types are used: CT and ST muons are restricted to the $|\eta| <$0.1 region; in $|\eta| <$2.5, about 97% are combined muons, approximately 1.5% are CT and the remaining 1% are ST muons.

- The *Medium* identification operating point provides the default selection for muons in ATLAS. This selection minimises the systematic un-

certainties associated with muon reconstruction and calibration. Only CB and ME tracks are used. The former are required to have ≥ 3 hits in at least two MDT layers, except for tracks in the $|\eta| < 0.1$ region, where tracks with at least one MDT layer but no more than one MDT hole layer are allowed. The latter are required to have at least three MDT/CSC layers, and are employed only in the $2.5 < |\eta| < 2.7$ region to extend the acceptance outside the ID geometrical coverage.

- *Tight* muons are selected to maximise the purity of muons at the cost of loosing efficiency. Only CB muons with hits in at least two stations of the MS and satisfying the Medium selection criteria are considered.

- The *High-p_T* selection aims to maximise the momentum resolution for tracks with transverse momentum above 100 GeV. CB muons passing the Medium selection and having at least three hits in three MS stations are selected. This last requirement while reducing the reconstruction efficiency by about 20 %, improve the p_T resolution of muons above 1.5 TeV by approximately 30 %.

The muon reconstruction efficiency is obtained with the tag-and-probe method for muons in the region $|\eta| < 2.5$, using J/Ψ and Z decays for low- ($<$10 GeV) and high-p_T muons respectively. The procedure is detailed in Appendix B, since I directly worked on the low-p_T muon reconstruction efficiency for a dataset corresponding to 3.2 fb^{-1}. The reconstruction performance for different operational points is shown in Fig.2.19 as a function of the pseudorapidity for high-p_T muons from Z decay.

Muons originated from the decay of heavy particles, such as W, Z or Higgs bosons, are produced isolated from other particles. The measurement of the muon isolation, namely the detector activity around a muon candidate, is therefore a powerful tool for background rejection in many physics analyses. As for electron, track-based and calorimeter-based isolation variables are defined. Track-based isolation variable $p_T^{varcone30}$ is defined as the scalar sum of the transverse momenta of the tracks with $p_T > 1$ GeV in a cone of size $\Delta R = \min(10 \text{ GeV}/p_T^\mu, 0.3)$ around the muon. The cone size is p_T dependent
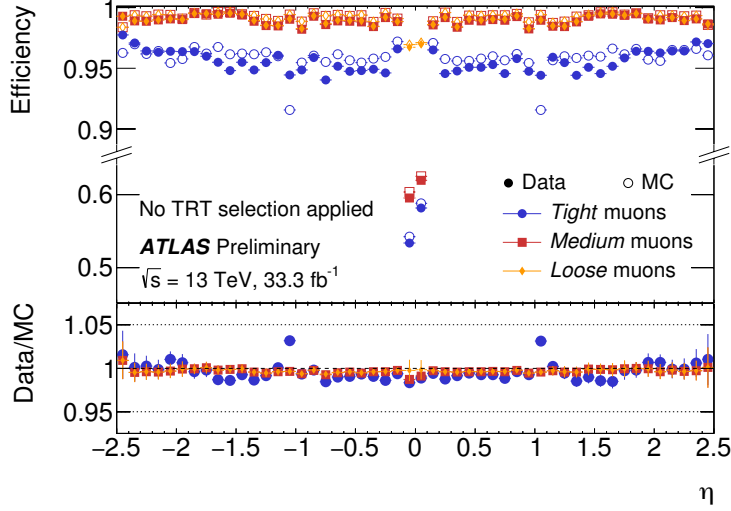
Figure 2.19: Muon reconstruction efficiencies for the Loose, Medium and Tight identification operational points measured in $Z \to \mu\mu$ events as a function of the rapidity for muon with $p_T > 10$ GeV.

in order to improve the performance for muons produced in the decay of particles with a large transverse momentum. The calorimeter-based isolation variable $E_T^{topocone20}$ is defined as the sum of the transverse energy of topological clusters in a cone of size $\Delta R{=}0.2$ around the muon, after subtracting contributions from pile-up and underlying events.

The isolation selection criteria are determined using the relative isolation variable, defined as the ratio of the track- or calorimeter-based isolation variables to the transverse momentum of the muon. A full description of the isolation working points can be found in [55]. In particular, in the Z inclusive cross section measurement (see Chapter 3), the *Gradient* isolation criteria is used: $p_T^{varcone30}/p_T^{\mu}$ and $E_T^{topocone20}/p_T^{\mu}$, corresponding to an isolation efficiency of $\geq 95(99)\%$ for muons with $p_T{=}25$ (60) GeV. In the Z+b(b)-jets cross section measurements the *FixedCutTightTrackOnly* is chosen, based on track information only ($p_T^{varcone30}/p_T^{\mu} < 0.06$).

2.4.4 Jets

Hadronic collisions in ATLAS produce a variety of particles, including large number of quarks and gluons. These particles immediately hadronise, producing a collimated shower of particles with a net momentum equal to the initiating quark or gluon one. This shower of particles is called "jet" and is extremely common in LHC collisions. Within the terminology of event reconstruction in a detector, "jet" takes an additional meaning. Track jets are the ensemble of ID tracks by charged particles, composing the jet. However, more often, "jet" refers to the calorimeter object. As described in Section 2.3.4, the calorimeter purpose is to measure the energy of particles, and in this case of jets. Instead of tracks, cone-shaped groups of energy deposits are reconstructed. A jet may have tracks in the ID corresponding to the passage of the particles belonging to the jet, as well as tracks in the MS if a muon is produced as part of the jet or the hadronic calorimeter fails to absorb all the energy of the jet.

The jet reconstruction begins with the formation of clusters, performed by a topological clustering algorithm, from a high energy seed cell [56]. Seed cells are required to have an energy significance at least 4σ above the noise level, defined as the quadratic sum of electronic and pile-up signals. Neighbour cells with energy significance higher than 2σ with respect to the noise are iteratively added to build topo-clusters. Topo-clusters group neighbour cells in the $\eta-\phi$ plane with significant energy deposit, in order to reconstruct jets. After the topo-cluster reconstruction, a splitting algorithm divides clusters in energy categories using a local energy maxima criterion.

The main strategy for jet reconstruction is to determine if the separation of two clusters is large enough, with respect to their energy, to be the results of different showers rather than part of the same jet. The anti-k$_t$ algorithm [57] sequentially recombines clusters with the following procedure:

- for each cluster i evaluate the distance d_{ij} with each other cluster j

$$d_{ij} = min(p_{T,i}^{2k}, p_{T,j}^{2k})\frac{\Delta R_{ij}^2}{R^2} \qquad (2.3)$$

where ΔR_{ij}^2 is the angular distance between clusters i and j in the $\eta-\phi$

plane, R is the cone radius of the jet and k is a parameter of the anti-k_t algorithm fixed equal to -1;

- for each cluster i, evaluate the distance from the beam:

$$d_{iB} = p_{T,i}^{2k} \quad ; \tag{2.4}$$

- find the minimum distance between d_{ij} and d_{iB};

- if the minimum value is d_{ij}, then combine i and j into a single pseudo-jet and repeat from the first step. Otherwise consider i as a final state and do no consider it in further iterations.

The anti-k_t algorithm favours the jet clusterization around hard seeds, rather than soft particles (k_t-algorithm, k=1) or energy independent clustering (Cambridge-Achen, k=0). The anti-k_t is an infrared and collinear safe procedure (IRC), which means that the set of hard jets remains unchanged even in case of a collinear splitting or the addition of a soft term gluon [57]. In the Z+b(b)-jets analysis presented in this thesis, jets are reconstructed using the anti-k_t algorithm.

A correct energy measurement is crucial in jet reconstruction. This is not-trivial due to the lower response of the hadronic with respect to the electromagnetic calorimeter, the energy deposited outside the calorimeter and outside the jet, pile-up effects and noise thresholds. Jet calibration combines several complementary corrections and calibrations [58, 59, 60]:

- a correction to the jet origin, which changes only the jet direction to point to the primary vertex and does not affect the energy of the jet;

- a pile-up correction, obtained by in situ measurements depending on the number of reconstructed primary vertices, the jet pseudorapidity and the bunch spacing.

- energy and η calibration, which involves corrections for energy loss in inert material or deposited outside the jet cone.

- a residual in situ calibration is applied only to jets in collision data and it is determined from collisions events by comparing the energy of the reconstructed jet to a reference object such as a photon, a Z or another jet.

In order to distinguish jets from pile-up, ATLAS has developed three different track-based tagging approaches [61]. In the following, only the algorithm used in the Z+b-jets analysis is described. The Jet Vertex Tagger (JVT) is a new discriminant derived from variables based on the sum of transverse momentum of jet ID tracks and the number of primary vertices. The reconstruction algorithm is a multivariate (MVA) technique called k-Nearest Neighbourhood, which evaluates the probability for a jet to arise from a hard-scatter vertex. The relative probability for a jet to be of "signal-type" is computed as the ratio of the number of hard-scatter jets to the number of hard-scatter plus pile-up jets found in a local neighbourhood of the jet. The $Z \to \mu\mu$+jets events are used to measure the jet efficiency for JVT in data, exploiting a tag-and-probe procedure [61] and calibrations are provided in bin of Z boson transverse momentum.

According to the JVT efficiency and fake reduction, several operating points are defined: *Loose*, *Medium* and *Tight*. In the Z+b(b)-jets, the *Medium* calibration point is used, which corresponds to require the JVT discriminant larger than 0.59 for jets with 20 GeV$< p_T <$60 GeV and $|\eta| <$2.4. The efficiency of the Medium JVT tagger is illustrated in Fig.2.20 as a function of the jet p_T.

b-jets

b-jets are jets originated from the fragmentation and hadronization of a b-hadron. They play a key role in many analyses within the ATLAS program, including Standard Model precision measurements, Higgs studies and searches. b-hadrons have a collection of unique properties which can be exploited for their identification. An illustration of the production of a b-jet is given in Fig.2.21. Within the Standard Model, the decay of a b-hadron is suppressed, resulting in a longer flight path before decaying in the inner detector, compared to c- and light-hadrons. In the laboratory frame, this
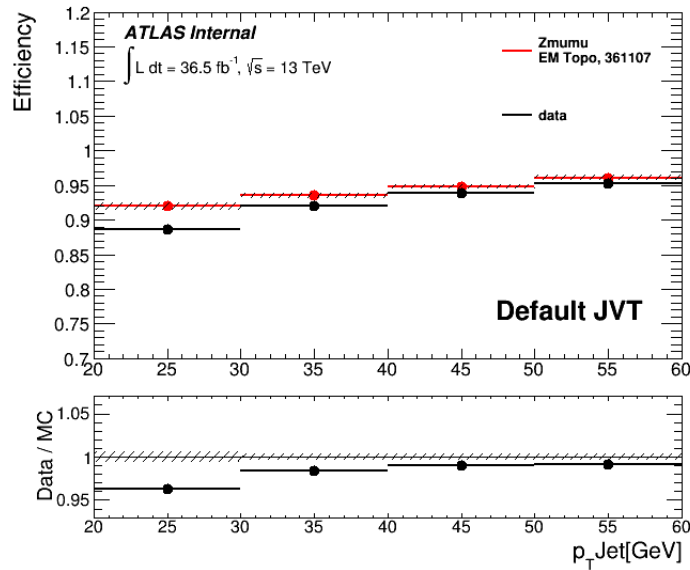
Figure 2.20: Efficiency of the JVT discriminant as a function of the jet transverse momentum p_T for the Medium operating point in Z+jets events.

means that a b-hadron will travel of the order of a few millimetres before decaying. This originates a displaced secondary vertex which is a key signature of b-hadron decay. In addition, the b-quark is the heaviest quark that can form hadrons, which results in a a heavy hadronic signature. Finally, the decay of a b-hadron produces on average a higher charged track multiplicity than lighter hadrons.

A new b-jet tagging algorithm, referred to as MV2, was developed for Run 2 [62]. It is based on a boosted decision tree (BDT) approach. All algorithms which provide the input variables for MV2 exploit the relatively long b-hadron lifetimes: a likelihood-based combination of the transverse (d_0) and longitudinal (z_0) impact parameter significances; the presence of a secondary vertex and related properties; and the reconstruction of the b-hadron decay chain by searching for a common direction connecting the primary vertex to both the bottom and the charm decay vertices. The BDT is trained to discriminate b-jets from light-jets. In the Z+b(b)-jets analysis, the MV2c10 algorithm is chosen for b-tagging, which allows for a 10% charm-jet contami-
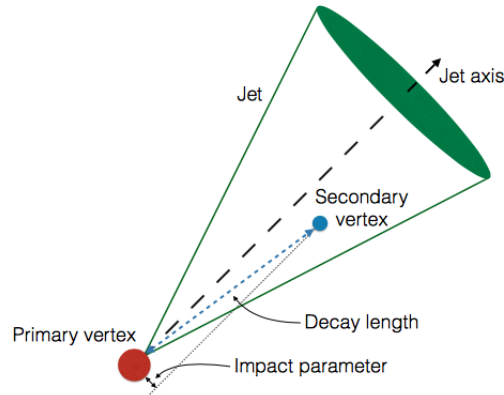
Figure 2.21: Ilustration of the production of a b-jet.

nation. The MV2c10 discrimination performances is shown in Fig.2.22 (left).
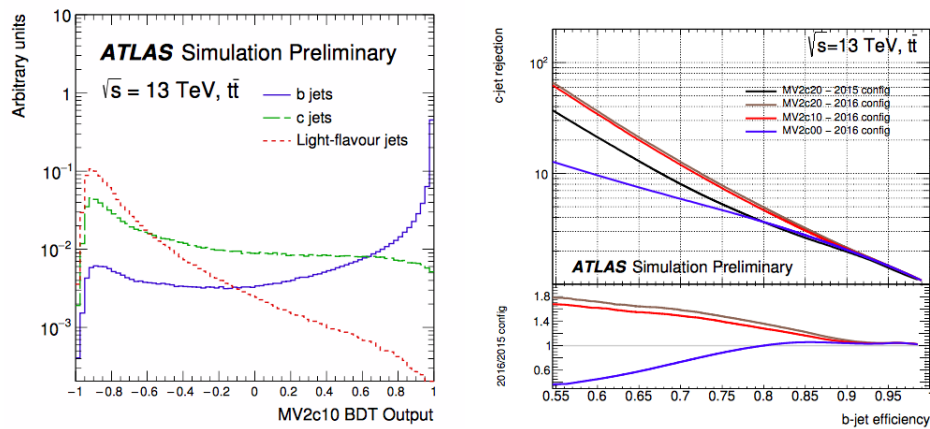
Figure 2.22: b-, c- and ligh-jets composition as selected by the Mv2c10 algorithm on $t\bar{t}$ events (left)[63]. c-jet rejection versus b-jet efficiency for 2015 and 2016 configurations of the MV2 b-tagging algorithm evaluated on $t\bar{t}$ events (right)[63].

The b-jet tagging efficiency is measured for jets in the pseudorapidity range $|\eta| < 2.5$ and with transverse momentum $p_T > 20$ GeV for several operating points. Four operating points are defined corresponding to 60, 70, 77 and 85% b-jet tagging efficiencies.

The performance of Run 2 MV2 has seen an improvement of 10% with respect to the Run 1 multivariate b-tagging algorithm (MV1) [63], due primarily to the inclusion of the IBL, which is dedicated to secondary vertex identification. Between 2015 and 2016, retraining of the b-tagging classifier has further improved the c-jet rejection by around 40% at the 77% efficiency operating point (see Fig.2.22(right)).

Calibrations of b-tagging algorithm imply a correction factor applied to simulated data to match the b-tagging efficiency measured in data. This correction factor is derived both for true b-jets (i.e. correctly identified), as well as for light and c-jet incorrectly b-tagged. Studies on b-tagging efficiency are performed on top pair samples since they are enriched in b-jets, benefiting from the high production rate at LHC. There are different methods for b-tagging calibration [64]:

- the *tag and counting* method fits the multiplicity of b-tagged jets in $t\bar{t}$ events. According to the Standard Model decay of the top-quark, the majority of the events are expected to contain exactly two real jets. Actually, the number of reconstructed b-jets will not necessary be equal to two, since b-jets can be lost outside the detector acceptance, or additional generated by gluon-gluon splitting or c- and ligh-jets can be tagged as b-jets.

- the *kinematic selection* method measures the b-tagging rate of the leading jets. It relies on the knowledge of the flavour composition of the $t\bar{t}$ signal and background samples, and extracts the b-jet tagging efficiency by measuring the fraction of b-tagged jets in data.

- the *kinematic fit* method uses a fit of the $t\bar{t}$ event topology to extract a highly purified sample of b-jets, where the b-tagging efficiency is evaluated. Based on a χ^2 fit, the procedure exploits the masses of the two top quarks and W bosons as constraints.

- the *combinatorial likelihood* method improves the precision offered by the kinematic selection, by exploiting the kinematic correlations between the jets in the event. It relies on an a priori knowledge of the flavour composition of the $t\bar{t}$ samples.

c-jet calibration uses a sample enriched in c-jets selected by reconstructing the decay chain of the $D^* \rightarrow D_0(K\pi)\pi^0$. The yields before and after applying a b-tagging requirement come from a certain fraction of true b-jets correctly selected with the corresponding efficiency, and a certain fraction of spurious c-jets. By fitting the proper lifetime of the events, the contamination from b-hadron decays can be estimated and used to extract the c-tagging efficiency. The light-jet tagging efficiency is evaluated in enriched regions built by inverting the sign of the impact parameter significance. The inclusive distribution of the tagged light-jets has a similar shape to the mis-tagged light-jets, and once the contamination from heavy flavour and long-lived particles is accounted for, it can be used to extract the light-jet tagging efficiency.

2.4.5 Missing energy

The ATLAS geometry and coverage allows to reconstruct most of the particles produced in the collisions. The conservation of the momentum in the plane transverse to the beam axis $(x - y)$ implies that the vector sum of transverse momenta of the collision products should be zero. To measure the momentum balance in the event the missing transverse energy (E_T^{miss}) quantity is used. E_T^{miss} is defined as the negative vectorial sum of energy in the transverse plane of all objects reconstructed in the event [65].
Since neutrinos do not interact with the ATLAS detector, in a Standard Model event a large E_T^{miss} is a signature of events with one or more high-p_T neutrinos. However E_T^{miss} can result from other sources, such as beam halo muons, cosmic muons and possible physics beyond SM.
The E_T^{miss} is computed as:

$$\begin{cases} E_x^{miss} = -\sum_{i=1}^{N_{cell}} E_i \sin\theta_i \cos\phi_i \\ E_y^{miss} = -\sum_{i=1}^{N_{cell}} E_i \sin\theta_i \sin\phi_i \end{cases} \tag{2.5}$$

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \tag{2.6}$$

where E_i, θ_i and ϕ_i are the calorimeter cell energy, the polar angle and the azimuthal angle respectively. The sum excludes cells flagged as noisy by the

ATLAS database and only includes cells belonging to topo-clusters in order to suppress noise.

Chapter 3

Inclusive Z cross section measurements

The production of the electroweak vector boson Z at hadron colliders provides a benchmark for the understanding of strong (QCD) and electroweak (EW) processes. The relatively large production cross section at the LHC energy and the decay in two leptons result in a clear experimental signature of the Z boson and, therefore, in a relatively easy identification. These features provide the possibility to perform high precision measurements, sensitive to the sea-quark distributions inside the colliding protons, as explained in Chapter 1. Given that the production cross section depends on the Parton Distribution Functions (PDFs) and on the underlying strongly interacting particles, a precise measurement offers also the possibility to test models of parton dynamics. Moreover, the measurement of the Z production cross section is one of the first analyses performed at the beginning of any new data taking period, because it offers the possibility to verify detector calibration and alignment and to perform luminosity check through the *Z-counting* technique [66].

In this Chapter the measurement of the Z production cross section performed using data collected at very beginning of LHC Run 2 is presented [67]. In Section 3.1 the definition of the fiducial and total cross sections is explained. In Section 3.2 the dataset used is presented, while Monte Carlo samples employed for signal and background simulation are summarised in Section 3.3.

An explanation of the theoretical predictions and their associated uncertainties is presented in Section 3.4. Section 3.5 contains a description of the event selection in order to enhance the Z signal over the background contributions, reduced through the methodology summarised in Section 3.6. In Section 3.7 the list of calibration and scale corrections applied to physics objects is presented. The data/MC comparison is performed in the description of kinematic distributions in Section 3.8. The efficiency C_Z and acceptance factor A_Z evaluation is described in Section 3.9 and the results of the measurements with the associated uncertainties are presented in Section 3.10. The analysis has been performed in parallel to the measurement of the W cross section, not presented in this thesis, in order to measure the W/Z cross section ratio that benefits from the cancellation of some experimental uncertainties, as described in detail in Section 3.11.

3.1 Fiducial and total cross section definition

The total cross section of the Z boson production is defined as:

$$\sigma_Z^{tot} \times BR(Z \to l^+l^-) = \frac{N^{obs} - B}{A_Z \cdot C_Z \cdot L} \qquad (3.1)$$

where the branching ratio (BR) is the probability for the Z boson to decay in a specific lepton channel (in this thesis $l=\mu$ or e). N_{data}^{obs} is the number of observed events that have been reconstructed by the ATLAS detector and that have passed the selection described in Section 3.5 (*detector level*). B represents the number of background events, estimated with Monte Carlo generators, that have passed the same selection. L is the total integrated luminosity referred to the analysed dataset. A_Z is the acceptance factor, which takes into account the geometrical and kinematic acceptance of the detector. A_Z is evaluated from the Monte Carlo truth (*truth level*) and it is defined as:

$$A_Z = \frac{N_{truth}^{MC\,fiducial\,volume}}{N_{truth}^{MC\,total}} \qquad (3.2)$$

where $N_{truth}^{MC\,total}$ is the total number of generated events, while $N_{truth}^{MC\,fiducial\,volume}$ represents the number of truth events in the fiducial volume of the detector,

defined by the following criteria:

- leptons with transverse momentum $p_T^l > 25$ GeV;

- leptons with pseudorapidity $|\eta^l| < 2.5$;

- invariant mass of the two same-flavour-opposite-charged leptons 66 GeV $< m_{ll} < 116$ GeV.

In Eq.3.1, C_Z represents the probability of reconstructing an event, if all the relevant physics objects characterising the event are in the detector acceptance and have passed the selection requirements. It can be expressed as:

$$C_Z = \epsilon_{trig} \cdot \epsilon_{reco} \cdot \epsilon_{iso} \qquad (3.3)$$

where ϵ_{trig}, ϵ_{reco} and ϵ_{iso} are, respectively, the lepton trigger, reconstruction and isolation efficiencies (see Section 3.5). C_Z is evaluated as the ratio of the number of events passing the detector level selection ($N_{selected}^{MC\,pass\,all\,cuts}$) to the number of events inside the fiducial volume at truth level ($N_{truth}^{MC\,fiducial\,volume}$):

$$C_Z = \frac{N_{selected}^{MC\,pass\,all\,cuts}}{N_{truth}^{MC\,fiducial\,volume}} \qquad . \qquad (3.4)$$

Since C_Z represents a correction from detector to truth level, it allows to perform measurements at *particle level*, namely corrected for detector resolutions and efficiencies.

The total production cross section measurement is performed in the full phase-space, including regions outside the detector acceptance. Therefore, it heavily relies on the extrapolation in an unmeasured region, described only by the Monte Carlo modelling, which is sensitive to purely theoretical uncertainties (i.e. the PDF). For this reason, experimentally the cross section is first measured in the fiducial volume described above, and then corrected to the full phase-space by means of the acceptance factor A_Z. The fiducial cross section is therefore defined as:

$$\sigma_Z^{fid} \times BR(Z \to l^+ l^-) = \frac{N_{data}^{obs} - B}{C_Z \cdot L} \qquad (3.5)$$

In both the definitions of Eq.3.4 and Eq.3.2, leptons are considered at *born* level, namely before they emit photons via QED Final State Radiation (FSR).

3.2   Data samples

The data used in this analysis have been collected at the very beginning of Run 2, between June 13th and July 16th 2015, during Data Periods A4 and C2-C5. The centre of mass energy was 13 TeV and LHC circulated beams with 50 ns bunch spacing. In period A4 the peak instantaneous luminosity was $L = 1.5 \cdot 10^{32}$ cm^{-1}s^{-1} and the mean number of pileup events was $< \mu >$=18.9. In periods C2-C5 the peak instantaneous luminosity was $L = 17.2 \cdot 10^{32}$cm^{-1}s^{-1} and the mean number of pileup events was $< \mu >$=19.2. The used dataset corresponds to a total L of 81±4 pb^{-1} (for a detailed discussion of the luminosity determination in ATLAS see Appendix A). Basic data quality criteria are applied in order to select data with fully operational detectors.

3.3   Monte Carlo samples

Monte Carlo (MC) simulations are used to evaluate the detector efficiency and background events and, therefore, are crucial ingredients for the cross section determination.

Signal events containing a Z boson decaying in the muon or electron channels are simulated with the POWHEG Monte Carlo [68, 69]. POWHEG is a fixed order NLO generator which produces the Z boson in association with up to one parton, using the CT10 PDFs set [70]. The POWHEG Monte Carlo program is interfaced with PYTHIA v.8.1 [71] (AZNLO CTEQL1 tune [72]) for the simulation of the non perturbative processes: parton shower, hadronisation and decays. The EvtGen v.1.2.0 tool [73] is used for the simulations of the b- and c-flavours and Photos++ [74] is used for QED emission of electroweak processes.

The POWHEG+PYTHIA program is also used to produce the electroweak backgrounds, whose main contributions are given by W boson decays and diboson processes.

The distributions of the top quark production (both $t\bar{t}$ and single top) are processed with the POWHEG-BOX v2 generator [69] and PYTHIA v.6.4 (Perugia tune) [75].

QCD multijet has been generated with PYTHIA v.8, whose events containing bb and cc quarks final states with PYTHIA8B [71].

All generated samples are normalised to the NNLO cross section predicted with the FEWZ program [76] using the MSTW2008NNLO PDF set [77]. The only exception is given by top sample, whose expected contribution is normalized to the NNLO+NNLL with the Top++2.0 program [78].

To simulate the effects of additional proton-proton collisions in the same and nearby bunch crossings (underlying events), additional events are generated using soft QCD processes of PYTHIA with A2 tune [79] and the MSTW2008LO PDF. Moreover all Monte Carlo samples are reweighted so that the mean number of pileup events matches the observed distribution in data. The samples are then processed with GEANT4 [80] to simulate the ATLAS detector and are produced within the ATLAS Framework (release 20.1.4.4) [81].

The full list of generators used in this measurement is presented in Table 3.1. For a detailed description of the generation mechanism see Chapter 1.

3.4 Theoretical cross sections

The theoretical cross sections are used for comparison with measurements and are computed by two programs: DYNNLO 1.5 [82] and FEWZ 3.1 [83] provide NNLO and NLO QCD calculations respectively. DYNNLO is used to evaluate the central value of the QCD predictions, while FEWZ is used for all systematics variations, such as scale variations and α_s running behaviour. The reason to use two separate programs originated from studies

Physics process	Generator	Theo unc.	$\sigma \cdot$ BR (pb)	Order	Reference
$Z \to e^+e^-$	POWHEG BOX + PYTHIA8	5%	1890	NNLO	FEWZ/CT14
$Z \to \mu^+\mu^-$	POWHEG BOX + PYTHIA8	5%	1890	NNLO	FEWZ/CT14
$Z \to \tau^+\tau^-$		5%	1890	NNLO	FEWZ/CT14
$W^+ \to l^+\nu$	POWHEG BOX + PYTHIA8	5%	11548	NNLO	FEWZ/CT14
$W^- \to l^-\bar{\nu}$		5%	8546	NNLO	FEWZ/CT14
$t\bar{t}$	POWHEG BOX + PYTHIA6	6%	831	NNLO+NNLL	TOP++v2.0
single top		6%	137.5	NLO	POWHEG
ZZ		5%	14.7	NLO	MCFM/CT10
WZ	SHERPA2.1.1	7%	42.4	NLO	MCFM/CT10
WW		5%	108.7	NLO	MCFM/CT10

Table 3.1: Simulated signal and background samples used in both electron and muon channel. For each physic process the production cross section, multiplied by the branching ratio, is specified with corresponding theoretical uncertainty. All Monte Carlo processes are given at NNLO with the exception of diboson and single top background contributions which are given at NLO.

of 2011 W,Z cross sections, where FEWZ showed a 1% bias in the fiducial predictions. On the other hand, PDF variations with respect to a central value are not affected.

Calculations are performed using the following PDF sets: CT14NNLO [84], CT10NNLO [85], NNPDF3.0 [86], MMHT14NNLO68CL [87], ABM12LHC [88], HERAPDF2.0 [89] and ATLAS-epWZ12NNLO [90]. The nominal renormalisation (μ_R) and factorisation (μ_F) scales are set to the dynamic value of the lepton invariant mass m_{ll}.

The NLO EW corrections are provided by FEWZ 3.1 taking the fundamental parameters from [91], consistently with the predictions used in the analysis at $\sqrt{s}$ =7 TeV. The production cross section of the Z boson decaying into two leptons is calculated at the *born level* to match the definition of the measured cross section in data. The following EW corrections are included in the calculations: virtual QED and weak corrections, real Initial State Radiation (ISR) and its interference with the real Final State Radiation (FSR). The PHOTOS program [74] is used to correct data for the FSR. EW corrections correspond approximatively to -0.2% and -0.35% in the fiducial and total phase space, respectively.

In Table 3.2, the DYNNLO central values and the FEWZ variations of the fiducial and total cross sections are summarised for various PDFs.

The uncertainties on the calculations are dominated by the limited knowledge of the proton PDFs and, in general, are derived from the following sources:

- **PDF**: it is evaluated by comparing the nominal CT14NNLO set with other three NNLO PDFs: NNPDF3.0, MMHT14NNLO68CL and ABM12LHC. The PDF uncertainty of CT14NNLO was rescaled from 90% CL to 68% CL.

- **Scales**: it is estimated by varying the renormalisation and factorisation scales by a factor two, in the constraint $0.5 \leq \mu_R/\mu F \leq 2$.

- α_S: the uncertainty on the strong coupling constant is estimated according to CT14NNLO PDF prescription, varying α_S by ± 0.001 to cor-

| | DYNNLO 1.5 | | FEWZ 3.1 | |
PDF	σ_Z^{tot} [pb]	σ_Z^{fid} [pb]	σ_Z^{tot} [pb]	σ_Z^{fid} [pb]
CT14nnlo	1892	745.4	1890_{-50}^{+45}	751_{-25}^{+20}
CT10nnlo	-	-	1929_{-51}^{+43}	769_{-25}^{+20}
NNPDF3.0	1861	736.8	1861 ± 40	740 ± 16
MMHT14nnlo	1910	756.3	1909_{-27}^{+31}	761_{-13}^{+15}
ABM12LHC	1915	763.9	1914 ± 23	769 ± 10
HERAPDF2.0	1982	781.9	1982_{-34}^{+57}	781.9_{-17}^{+26}
ATLAS-epWZ12	1971	788.4	1971_{-31}^{+34}	788.4_{-16}^{+16}

Table 3.2: Summary of the total σ_Z^{tot} and fiducial σ_Z^{fid} calculations for $Z \to ll$ using various PDF sets, computed using DYNNLO 1.5 (left) and FEWZ 3.1 (right). For the DYNNLO 1.5 program, only central values are shown. FEWZ 3.1 calculations are reported with the variations corresponding to the PDF uncertainties at 68% CL. For both programs, no EW corrections are applied.

respond to 68% CL. The uncertainty corresponds to $\pm0.9\%$ for both the fiducial and the total Z cross sections.

- **Beam energy**: the uncertainty due to a 1% on E_{beam} knowledge is estimated to be 1.1%. The uncertainty amounts to $\pm1.1\%$ for the total and to $_{-1.0}^{+0.8}\%$ for the fiducial cross sections.

- **Intrinsic theoretical uncertainties**: they are related to the limitations of NNLO calculations, the non perturbative parameters and the comparison between DYNNLO and FEWZ. These uncertainties are $\sim0.2\%$ for the fiducial and $\sim0.4\%$ for the total cross sections and thus can be neglected.

The NNLO QCD+NLO EW calculations for the total and fiducial cross sections with CT14NNLO PDF are:

$$\sigma_{tot}^Z = 1888_{-50}^{+45}(\text{PDF})_{-21}^{+14}(\text{scale}) \pm 26(\text{other})\text{pb}$$
$$\sigma_{fid}^Z = 742_{-25}^{+20}(\text{PDF})_{-5}^{+3}(\text{scale}) \pm 11(\text{other})\text{pb}$$

The first uncertainty is related to the variations of the PDFs, the second to the QCD scales and the third to an estimate of all remaining systematics mentioned previously, added in quadrature. Since the values of QCD scale uncertainties contain a statistical component, the scale uncertainties given before have been replaced with a flat and symmetric 1.1% uncertainty for each prediction. This number is derived from an envelope of the observed variations.

The W fiducial and total cross section are specified in the following for completeness, since they are important ingredient in the ratio of fiducial cross section calculations.

$$\sigma_{tot}^{W+} = 11540_{-309}^{+318}(\text{PDF})_{-154}^{+107}(\text{scale}) \pm 160(\text{other})\text{pb}$$
$$\sigma_{tot}^{W-} = 8543_{-237}^{+214}(\text{PDF})_{-108}^{+86}(\text{scale}) \pm 120(\text{other})\text{pb}$$
$$\sigma_{tot}^{W} = 120083_{-539}^{+525}(\text{PDF})_{-262}^{+193}(\text{scale}) \pm 280(\text{other})\text{pb}$$

$$\sigma_{fid}^{W+} = 4423_{-139}^{+127}(\text{PDF})_{-49}^{+20}(\text{scale}) \pm 70(\text{other})\text{pb}$$
$$\sigma_{fid}^{W-} = 3396_{-110}^{+89}(\text{PDF})_{-26}^{+25}(\text{scale}) \pm 51(\text{other})\text{pb}$$
$$\sigma_{fid}^{W} = 7819_{-247}^{+214}(\text{PDF})_{-75}^{+45}(\text{scale}) \pm 117(\text{other})\text{pb}$$

The comparison of the predicted cross section values with the obtained measurements is presented in Sections 3.10-3.11.

3.5 Event selection

The cross section measurement of the Z boson production relies on the selection of the events where the Z boson decays into a pair of opposite charge and same flavour leptons. In order to select those events, the final objects must fulfil the requirements described in detail in Table 3.3. The particle reconstruction in ATLAS and the description of the identification and isolation criteria are extensively described in Chapter 2.

The event selection is based on:

- **Trigger**

Preselection: GRL + $\geq$ 1 primary vertex with $\geq$ 2 tracks	
Electon Selection	
Trigger	HLT_e24_lhmedium_iloose_L1EM20VH OR e60_lhmedium
p_T	> 25 GeV
η	$|\eta| < 2.47$ AND $1.37 < |\eta| < 1.52$
ID	MediumLH
Isolation	Gradient Isolation
Muon Selection	
Trigger	HTL_mu20_iloose_L1MU15 OR HTL_mu50
p_T	> 25 GeV
η	$|\eta| < 2.4$
ID	Medium
Isolation	Gradient Isolation
Z boson Selection	
Lepton Features	exactly 2 leptons same flavor opposite charge
Mass Window	66 GeV $< m_{ll} <$ 116 GeV

Table 3.3: Description of the event selection for the Z production cross section measurement.

Data are acquired with triggers that require at least one electron with $p_T > 24$ GeV (HLT_e24_lhmedium_iloose_L1EM20VH) or one muon with $p_T > 20$ GeV (HTL_mu20_iloose_L1MU15). To recover possible efficiency losses at high momenta, additional electron and muon triggers with looser identification are employed with thresholds of $p_T > 60$ GeV for the electron channel (e60_lhmedium) and $p_T > 50$ GeV for the muon channel (HTL_mu50).

- **Vertex**
  Events are required to have at least one primary vertex, with at least two associated tracks with $p_T > 400$ MeV each.

- **Electrons**
  Electrons are identified with the *Medium Likelihood* criterion. They are required to have a transverse momentum larger than 25 GeV and to fall within the fiducial region of pseudorapidity $|\eta| < 2.47$. The calorimeter crack region between $1.37 < |\eta| < 1.52$ is excluded. Electrons must fulfil the *Gradient Isolation* operating point (see Chapter 2).

- **Muons**
  Muons are required to pass the *Medium* identification criterion and to have $p_T > 25$ GeV in the pseudorapidity region $|\eta| < 2.4$. Muons must pass the *Gradient Isolation* operating point (see Chapter 2).

  For both electrons and muons, the *Gradient Isolation* is tuned in such a way that the isolation cut efficiency is around 90% for leptons of $p_T > 25$ GeV and around 99% for leptons of $p_T > 60$ GeV.

- **Z boson**
  Candidate events are required to contain exactly 2 leptons, satisfying the previous cuts, with same flavour and opposite charge, with an in-

variant mass in the window 66 GeV$< m_{ll} <$116 GeV.

A total number of 35009 events pass all the requirements in the electron channel and 44898 events in the muon channel. In Tables 3.4 and 3.5 the Z candidates at each step of the chain are summarised, in the electron and muon channel respectively.

Requirements	$Z \to e^+e^-$	ϵ_{rel} (%)	ϵ_{tot} (%)
Trigger	141600	100	100
2 Medium ID electrons	42680	30.1	30.1
Isolation	36900	86.5	26.1
Same flavour-opposite charge e^+e^- pairs	36370	98.6	25.7
66 GeV $< m_{ee} <$ 116 GeV	35009	96.3	24.7

Table 3.4: Number of $Z \to e^+e^-$ candidate events after the application of each criterion of the selection chain. The relative efficiency ϵ_{rel} of each cut is calculated with respect to the previous one. The total efficiency ϵ_{tot} is given with respect to the number of events that have passed the trigger requirement.

Requirements	$Z \to \mu^+\mu^-$	ϵ_{rel} (%)	ϵ_{tot} (%)
Trigger	445400	100	100
2 Medium ID muons	59300	13.3	13.3
Isolation	46910	79.1	10.5
Same flavour-opposite charge $\mu^+\mu^-$ pairs	46880	99.9	10.5
66 GeV $< m_{\mu\mu} <$ 116 GeV	44898	95.8	10.1

Table 3.5: Number of $Z \to \mu^+\mu^-$ candidate events after the application of each criterion of the selection chain. The relative efficiency ϵ_{rel} of each cut is calculated with respect to the previous one. The total efficiency ϵ_{tot} is given with respect to the number of events that have passed the trigger requirement.

3.6 Background contributions

The clear signature of the Z-boson decay into two leptons, reconstructed as mentioned in the previous section, makes its identification relatively easy.

The background contributions are expected to be small and are mainly due to:

- W decays, where the neutrino missing energy is misidentified as a lepton;

- $t\bar{t}$ processes, where the fully leptonic decay involves two same-flavour leptons in the final state;

- diboson production.

All backgrounds are simulated with Monte Carlo generators (see Section 3.3) and normalised to the cross sections in Table 3.1.

Additional source of backgrounds are due to the semileptonic decays of heavy quarks, the misidentification of hadrons as leptons and the production of electrons from internal conversions. These processes are referred to as "multijet" background and their contribution is determined separately in the electron and muon channels.

In the electron channel, the multijet background is estimated using events generated by Monte Carlo simulations and reconstructed using isolation criteria, relaxed with respect to the nominal selection. This procedure indicates a contribution of less than 0.1%, consistent with studies performed with data collected at the centre of mass energy $\sqrt{s} =$7 TeV, and considering that in the 13 TeV analysis tighter lepton selection criteria are used.

In the muon channel, the multijet contribution is estimated through a data-driven technique. The strategy consists in performing a fit on the transverse impact parameter (d_0) in a region enriched of jets, obtained by inverting the isolation cut and requiring $|d_0| > 0.1$ mm for one of the candidates (Control Region). Then, the tag-and-probe method is employed to obtain the d_0 distribution of the multijet background, using the candidate in the control region as "tag" and the second one, within the signal region, as "probe". The invariant mass distribution of the selected muons does not show the Z-boson peak, which is a clear signal of the purity of the multijet sample (Fig.3.1). A fit on data is performed, leaving the multijet normalisation as free parameter. The fit estimates a multijet background in the muon channel around (0.06 ± 0.04)% and the result is presented in Fig.3.2.
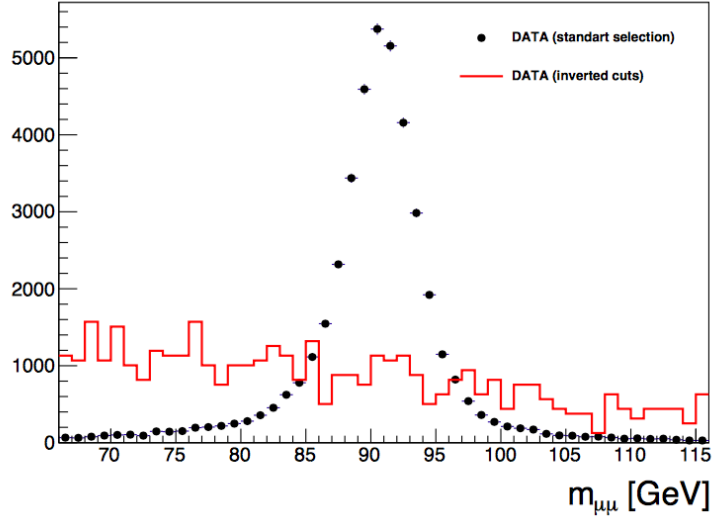
Figure 3.1: The invariant mass distributions of the two muon-candidates using the inversion of the isolation cuts and requiring $|d_0| > 0.1$ mm for one of the two muons (red line) and standard selection (black points). The Z peak disappears when inverting the requirements, as expected in a region enriched by background contributions.
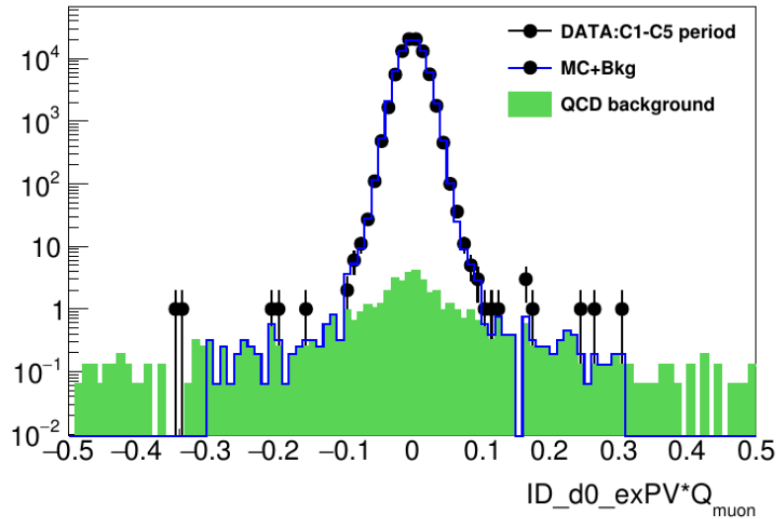

Figure 3.2: Fit of the d_0 distribution for $Z \to \mu^+\mu^-$ events to estimate multijet contribution. The template is built using the tag-and-probe method, considering as "tag" the muons satisfying the $|d_0| > 0.1$ mm cut and as "probe" the other.

3.7 Corrections and scale variations

In general, the sum of signal and background Monte Carlo samples is expected to describe data. Moreover, a proper estimation of C_Z requires reliable efficiencies for triggering, reconstruction and isolation. For these reasons, several event-by-event corrections are applied to simulated samples to match the data. Corrections to physics objects are given by the so-called "Scale-Factors" (SFs), defined as the ratio of efficiency in data ϵ_{data} with respect to efficiency in Monte Carlos ϵ_{MC}: SF$= \epsilon_{data}/\epsilon_{MC}$.

The other corrections applied at this stage take into account lepton calibration and number of pileup vertices in the event. The pileup re-weighting makes the distribution of MC events as a function of the number of reconstructed primary vertices similar to the distribution observed in data.

Scale Factors, calibrations and Pileup reweighting are provided by specific tools within the ATLAS Framework, with associated statistical and systematic uncertainties. The uncertainty on the signal and background events is evaluated by varying within $\pm 1\sigma$ the nominal value of corrections and scale-factors. In the following, a list of all the ATLAS tools employed in the analysis and the corresponding corrections and uncertainties is presented.

Muon trigger scale factor (MUON_EFF_TrigUncertainty) corrects for differences in data and Monte Carlo in the trigger requirement. The statistics and systematics variations are calculated separately: for the statistical component, a toy model is used [67], while variations of $\pm 1\sigma$ are employed to determine the systematic components.

Muon identification scale factor corrects for efficiency in the muon identification. Variations of $\pm 1\sigma$ provide both statistical (MUON_EFF_STAT) and systematics (MUON_EFF_SYS) uncertainties.

Muon calibration is a correction which takes into account the smearing of the ID and MS tracks and of the energy scale. Variation of $\pm 1\sigma$ are used to evaluate the associated systematics: MUONS_ID, MUONS_MS and MUONS_SCALE.

Electron calibration computes the energy of electrons and photons and corrects for simulations, gain corrections, scale factors (see [67] for the full list of corrections). All the physical effects are summed in quadrature. The associated systematics involves scale (EG_SCALE_ALL) and resolution (EG_RESOLUTION_ALL) variations.

Electron reconstruction, identification, trigger and isolation
The ElectronEfficiencyCorrections tool provides reconstruction, identification, trigger and isolation scale factors, and the associated systematic uncertainties. The statistical component of the uncertainty is evaluated with a toy model [67]. As explained in Table 3.3, triggers are applied to single lepton, which means that the trigger cut is fulfilled if at least one of the two leptons satisfied the requirement. The trigger efficiency is therefore given by a combination of the single lepton trigger efficiency, calculated with the following formula:

$$\epsilon_{e_{1,2}}^{\text{triggMC}} = \frac{\epsilon_{e_1}^{\text{triggdata}} + (1 - \epsilon_{e_1}^{\text{triggdata}})\epsilon_{e_2}^{\text{triggdata}}}{\epsilon_{e_1}^{\text{triggMC}} + (1 - \epsilon_{e_1}^{\text{triggMC}})\epsilon_{e_2}^{\text{triggMC}}} \ . \tag{3.6}$$

The same methodology is followed in the muon channel, but the calculation is provided directly from the Muon trigger scale factor tool.

Pileup Reweighting
Monte Carlo samples are generated with a generic spectrum of average interactions per bunch-crossing μ, therefore, the simulation should be corrected for actual values. This has an impact on the data/MC comparison of several kinematic variables sensitive to the amount of secondary interaction of the p-p collisions. Different methods have been investigated in the contest of the $W \rightarrow \mu\nu$ to define the best strategy for both W and Z analyses (see [67]). The selected method is based on a reweighting of the μ distribution to match the corresponding distributions in data after that μ values are scaled by a factor 1/1.16. This scaling has been evaluated within the ATLAS Tracking Group in order to take into account the fraction of inelastic activity differently described in MC with respect to data.

The calculation of the systematic uncertainty associated to this correction must take into account that the variation of the pileup scaling factor has an impact on the lepton isolation and identification efficiencies, whose contributions are subtracted to avoid double counting.

The previous corrections and calibrations have an impact on the C_Z systematic uncertainty. The contribution of each systematic component to C_Z is summarised in Table 3.7 and Table 3.8 for the electron and muon channels, respectively. In the following Section, the Monte Carlo comparison with data is presented for several physics observables. In the following plots, the band of uncertainty corresponds to the sum in quadrature of the previously described corrections and scale variations.

3.8 Kinematic distributions

In this Section the kinematic distributions at *detector level* of Z events passing the full selection chain are shown. From Fig.3.3 to Fig.3.7 the $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ kinematic distributions are presented one next to the other to allow the comparison between the channels. The shaded uncertainty band in these plots contains the contributions described in Section 3.7, mainly due to:

- lepton energy and momentum scale and resolution;

- lepton trigger efficiency;

- lepton reconstruction, identification and isolation efficiencies;

- uncertainty in the cross section calculation for electroweak and top background contributions;

- statistical uncertainty due to limited Monte Carlo size;

while the luminosity systematics ($\pm 5\%$) is not included. The systematic uncertainties are combined in the shaded band and the statistical uncertainty is shown on the data points. The background contributions have been estimated from Monte Carlo simulations and are represented with different

colors in the plots. A very good description of the background-subtracted data is provided by the Monte Carlo in all variables.
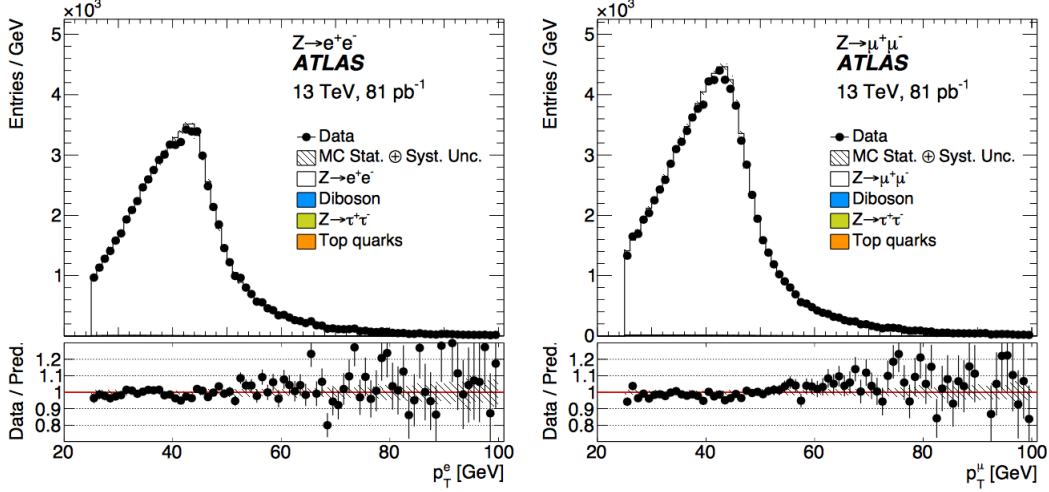

Figure 3.3: Transverse momentum distributions of the two leptons from the $Z \to e^+ e^-$ (left) and $Z \to \mu^+ \mu^-$ (right) selections.

3.9 Background evaluation and C_Z and A_Z determination

In order to measure the fiducial cross section of Eq.3.5, the number of data subtracted by the background contributions, the efficiency correction C_Z and the geometrical factor A_Z are required.

3.9.1 Data and background events

The data do not contain only Z-candidates, but also background processes, whose contribution is simulated with Monte Carlo generators. Table 3.6 summarises the number of observed candidates in $Z \to e^+ e^-$ and $Z \to \mu^+ \mu^-$ channels and include the number of the background expected events from electroweak and top quark processes. The multijet contribution has been evaluated to be negligible (less than 0.1%) in Section 3.6. The first
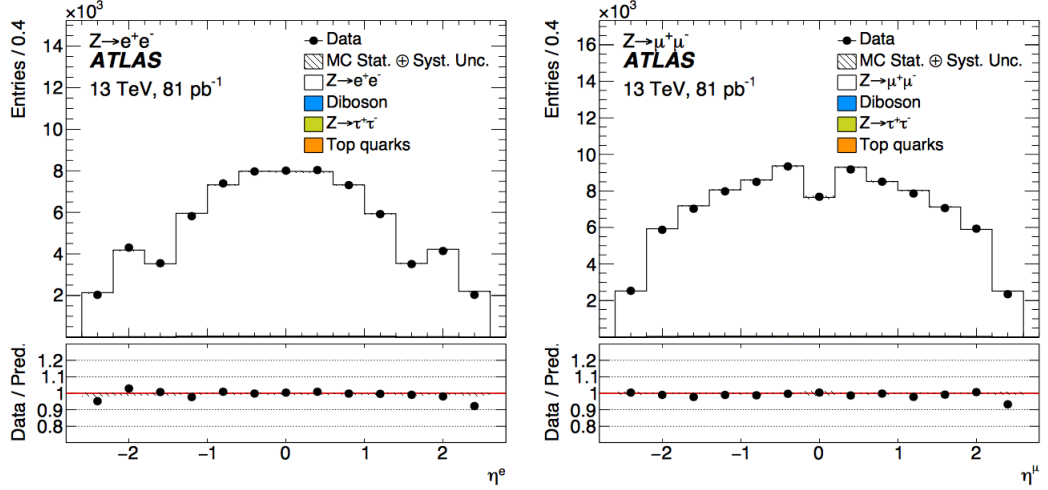
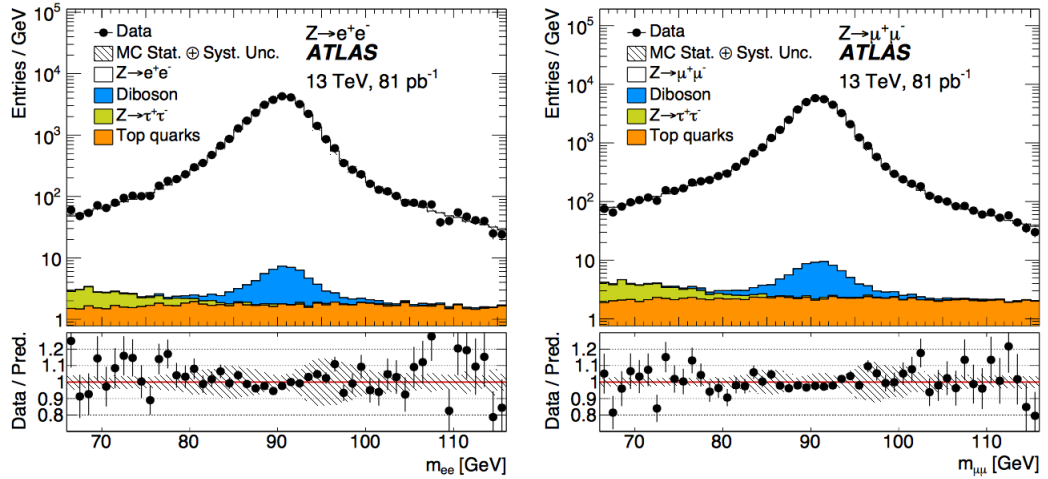Figure 3.4: Pseudorapidity distributions of the two leptons from the $Z \to e^+e^-$ (left) and $Z \to \mu^+\mu^-$ (right) selections.

Figure 3.5: Dilepton mass distribution after the $Z \to e^+e^-$ (left) and $Z \to \mu^+\mu^-$ (right) selections in a logarithmic scale.

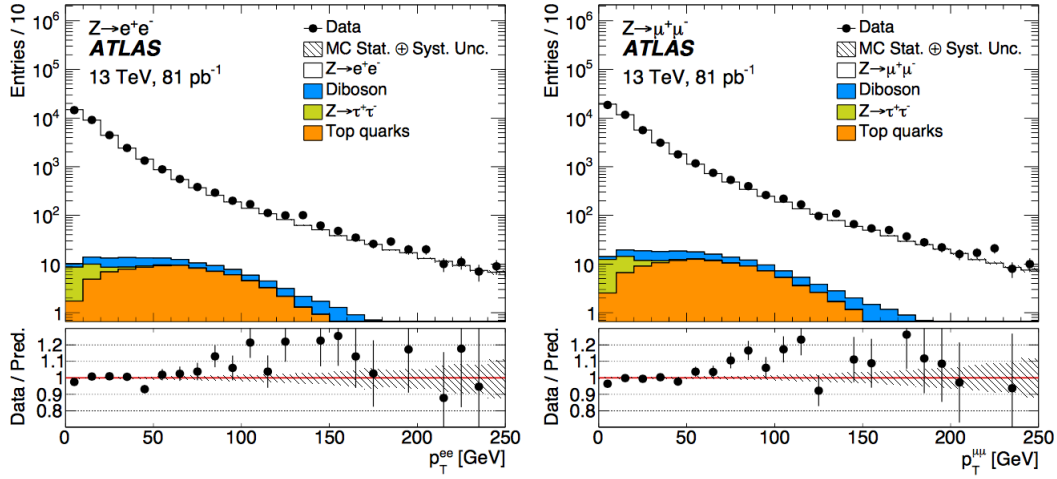Figure 3.6: Z boson transverse momentum distribution after the $Z \to e^+e^-$ (left) and $Z \to \mu^+\mu^-$ (right) selections.

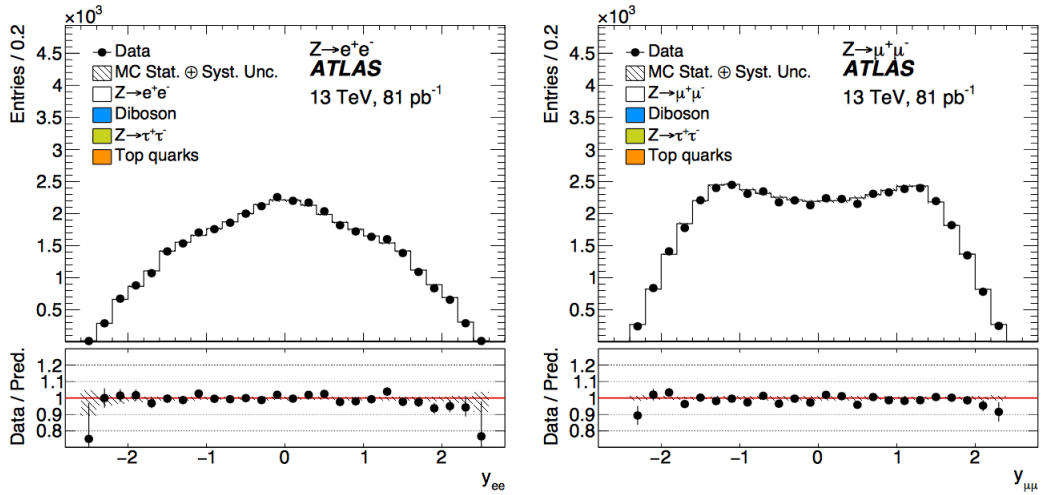Figure 3.7: Z boson rapidity distribution after the $Z \to e^+e^-$ (left) and $Z \to \mu^+\mu^-$ (right) selections.

component of the uncertainty is due to statistics (mainly on data sample) and the second to systematics, according to the corrections and scale factors described in Section 3.7. The luminosity determination uncertainty of 5% contributes to the systematic uncertainty.

Channel	N^{obs}_{data}	EW+top Background	$N^{selected}_{data}$ (Data - Background)
		value $\pm$ (stat.) $\pm$ (syst.)	value $\pm$ (stat.) $\pm$ (syst.)
$Z \to ee$	35009	$143.9 \pm 1.0 \pm 7.5$	$34865.1 \pm 187.1 \pm 7.5$
$Z \to \mu\mu$	44898	$191.3 \pm 1.2 \pm 9.8$	$44706.7 \pm 211.9 \pm 9.8$

Table 3.6: Number of observed Z candidate events in the electron and muon channel. In the second column, the number of background due to electroweak and top processes (EW+top) is shown with the statistical and systematic uncertainties. The number of background-subtracted events is evaluated and reported in the third column with its statistical and systematic uncertainty components.

3.9.2 C_Z and its uncertainty

The systematics uncertainties on C_Z are due to the corrections applied to physics objects (listed in Section 3.7), to the order of approximation of the Monte Carlo generator, to the PDF choice and to the analysis strategy. The contributions to the systematic uncertainties in the determination of the correction factor C_Z are summarised in Table 3.7 for the electron channel and in Table 3.8 for the muon channel.

In the **electron channel**, the main contribution on uncertainties arises from the electron reconstruction and identification. The *MediumLH* electron identification efficiency is calculated with respect to all reconstructed electrons: therefore additional effects must be taken into account if an electron within the fiducial volume fails to be reconstructed as a candidate electron. Another uncertainty is given by the electron charge misidentification: it occurs when an electron early radiates photons in the detector, which consequently convert and are reconstructed as high-p_T tracks. A particle with reconstructed charge opposite to the parent electron may then accidentally

be associated with the calorimeter cluster. The probability to correct identify the charge of the candidate electron is evaluated with the *tag-and-probe* method. The "probe" is identified as the ensemble of di-electron pairs without any requirement on the reconstructed sign of the track. The probe can be anywhere within the inner detector acceptance. The "tag" is confined to the barrel region of the calorimeter ($|\eta|$ <0.8), where the charge reconstruction efficiency is close to 100%. It is found that the sample where both electron-candidates are reconstructed with the same sign has a small background (below 10%). Based on this study, a 0.15% uncertainty is assigned for the effect of the opposite charge requirement, covering potential background contamination from the same-sign sample (0.14%) and the small difference between data and MC (0.01%).

In the **muon channel** the main contributions arise from the reconstruction and isolation efficiencies; other effects come from the inefficiencies for selected collisions with a reconstructed primary vertex, as well as the muon scale and resolution.

In both channels, Monte Carlo samples are reweighed to the medium number of pileup events observed in data ($< \mu >$). The pileup uncertainty is estimated as the difference between the distributions obtained applying or the nominal scale factor or the scale factor varied according to the tracking-CP group [67].
Another effect with an impact on the final C_Z value is the choice of the PDF set used at the generation stage. The Z signal is simulated using the CT10NLO PDF set, which has 26 vectors of variations.
The obtained values of the correction factors C_Z are:

$$C_{Z \to e^+ e^-} = 0.5518 \pm 0.0003 \text{(stat)} \pm 0.0055 \text{(syst)}$$
$$C_{Z \to \mu^+ \mu^-} = 0.7111 \pm 0.0003 \text{(stat)} \pm 0.0075 \text{(syst)}$$

where the systematic uncertainties is given by the sum in quadrature of the contributions listed in Table 3.7 and Table 3.8 for the electron and muon

channel, respectively.

Process	$Z \to e^+e^-$	
Uncertainty on $\delta C_Z/C_Z$	Up (%)	Down (%)
Statistics	0.05	-0.05
EG_RESOLUTION_ALL	-0.02	0.01
EG_SCALE_ALL	0.22	-0.23
EL_EFF_ID_COMBMCTOY	0.48	-0.48
EL_EFF_Iso_COMBMBTOY	0.29	-0.29
EL_EFF_Reco_TotalCorrUncertainty	0.77	-0.76
EL_EFF_TRIG_COMBMCTOY	0.05	-0.05
Pileup	0.01	-0.01
Opposite charge	-0.15	0.15
PDF	0.14	-0.08
Total	1.00	-0.99

Table 3.7: Summary of the different terms contributing to the uncertainty on C_Z in the electron channel.

Process	$Z \to \mu^+\mu^-$	
Uncertainty on $\delta C_Z/C_Z$	Up (%)	Down (%)
Statistics	0.05	-0.05
MUONS_ID	-0.05	-0.01
MUONS_MS	-0.01	0.00
MUONS_SCALE	-0.07	0.06
MUON_EFF_STAT	0.61	-0.61
MUON_EFF_SYS	0.64	-0.64
MUON_EFF_TrigSystUncertainty	0.17	-0.17
MUON_EFF_TrigStatTOYUncertainty	0.10	-0.10
MUON_ISO_STAT	0.49	-0.48
MUON_ISO_SYS	0.22	-0.21
Pileup	-0.01	-0.03
PDF	0.02	-0.01
Total	1.05	-1.05

Table 3.8: Summary of the different terms contributing to the uncertainty on C_Z in the muon channel.

3.9.3 A_Z and its uncertainty

The geometrical acceptance A_Z is used to extrapolate the cross section measurement from the fiducial to the full phase-space (see Eq.3.1). A_Z is calculated using the fixed order DYNNLO 1.5 prediction for the central value, FEWZ 3.1 for the PDF variations and the CT14NNLO PDF set for the baseline value. The central values of A_Z are provided in Table 3.9. The statistical uncertainties resulting from these evaluation are negligible.

PDF	A_Z
CT14NNLO	$0.393^{+0.004}_{-0.005}$
NNPDF3.0	0.395
MMHT14nnlo	0.395
ABM12LHC	0.398

Table 3.9: Summary of geometrical acceptance factor A_Z for $Z \to ll$ events using various PDF sets. The uncertainty on the top row represents the variation for the CT14NNLO PDF, while for the other PDFs only central value are given.

The systematic uncertainties on the acceptance are due to the following contributions:

- **PDF**: the uncertainty of CT14NNLO PDF has been rescaled from 90% to 68% CL. Moreover, a study on the change of PDF has been performed to take into account possible effect of extra PDF uncertainty. The envelope is estimated with four different NNLO PDFs: CT10NNLO, NNPDF3.0, MMH14NNLO68cl and ABM12LHC (see Table 3.9). This is the dominant contribution on A_Z systematics.

- **Scale**: the scale uncertainty has been determined by varying the renormalization and factorisation scales by a factor of two.

- α_s: the related uncertainty is estimated by varying α_s of $\pm 0.001\%$ to correspond to 68% CL.

- **Comparison with POWHEG+PYTHIA**: the difference between fixed-order predictions and Monte Carlo simulations is taken as additional uncertainty.

- **Parton Showers and Hadronisation**: this uncertainty arises from the difference in acceptance calculated with POWHEG using CTEQ 6.6 PDF set, but different models (namely HERWIG or PYTHIA) for the simulation of the parton shower and the hadronisation.

The sum in quadrature of the previous terms gives the final uncertainty on the geometrical acceptance A_Z. Given that the fiducial phase space has the same definition in the electron and muon channel, the resulted A_Z has the same value in both lepton modes:

$$A_Z = 0.393 \pm 0.007.$$

3.10 Cross section measurement

All the necessary elements to measure the fiducial cross section and to extrapolate it to the full phase space have been described in Section 3.9. The fiducial and total cross section measurements are presented in Table 3.10 for the electron and muon channel.

	Electron channel value $\pm$ stat $\pm$ syst $\pm$ lumi	**Muon Channel** value $\pm$ stat $\pm$ syst $\pm$ lumi
$N_{data}^{selected}$	$34865.1 \pm 187.1 \pm 6.9 \pm 3.0$	$44706.7 \pm 211.9 \pm 9.0 \pm 4.0$
C_Z	$0.5518 \pm 0.0003\,^{0.0055}_{-0.0055}$	$0.7111 \pm 0.0003\,^{0.0075}_{-0.0075}$
σ_Z^{fid} (pb)	$780.8 \pm 4.2 \pm 7.7 \pm 16.4$	$777.0 \pm 3.7 \pm 8.2 \pm 16.3$
A_Z	0.393 ± 0.007	0.393 ± 0.007
σ_Z^{tot} (pb)	$1986.9 \pm 10.7 \pm 40.5 \pm 41.7$	$1977.1 \pm 9.4 \pm 40.9 \pm 41.5$

Table 3.10: Results of the fiducial and the total cross sections for the production of the Z boson in the electron and muon decay channels. The number of background-subtracted observed events ($N_{data}^{selected}$), the correction factors (C_Z) and the acceptance factors (A_Z) are also reported in both channels.

The results in the electron and muon channels are expected to agree, according to the lepton universality predicted by the Standard Model. The ratio of the measurements preformed in the electron and muon channels is calculated taking into account all the correlated systematics. In Fig.3.8, the

ratio is shown for both the Z and the W bosons, the latter not being discussed in this thesis. A good agreement with Standard Model expectations within the systematic uncertainties is visible, in particular in the Z sector.
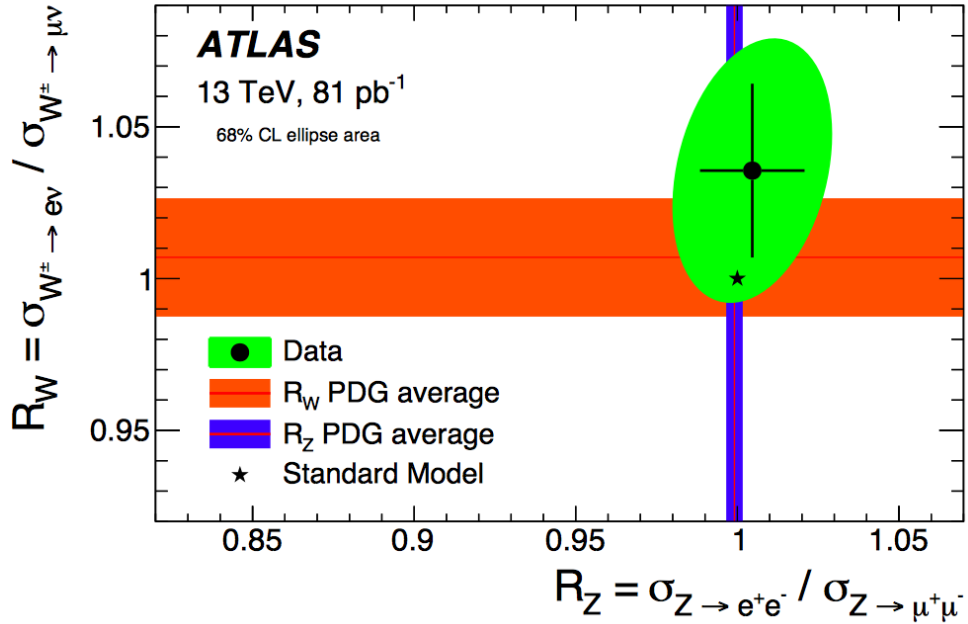

Figure 3.8: Ratio of production cross sections obtained in the electron and muon decay channels for the Z- (x axis) and W- (y axis) bosons, compared to the Standard Model expectations (back star). The bands represent the PDG average bands as from previous measurements.

3.10.1   Cross section combinations

As the single channel cross sections are in good agreement, the measurements have been combined, taking into account the correlated uncertainties. The combination is done for the W^+, W^- and Z fiducial cross sections simultaneously with the HERAverager tool. The combination employs individual sources of systematics (see Section 3.7), symmetrized around the nominal distribution. Correlations among uncertainties depend on the analysis (whether Z or W) and on the leptonic channel and are summarised in the following:

- lepton identification, trigger and reconstruction are uncorrelated between the two lepton channels;

- the missing E_T reconstruction affects only the W analysis;

- PDFs have an impact on all the measurements;

- electron identification, trigger and isolation and muon trigger statistical uncertainties are estimated using a toy MC method: these sources are not fully correlated [67] and the correlation is stronger for W^+ vs W^- compared to W vs Z.

- background contributions are treated as uncorrelated between W and Z and totally correlated between different channels in the same analysis. They are due to two different sources: theoretical cross section determination and common luminosity uncertainty.

- the common normalization uncertainty due to luminosity is excluded.

The combination result indicates a good compatibility of the measurements ($\chi^2/N_{d.f.} = 3.0/3$).

The combined fiducial and total cross section measurements of the Z boson production are summarised in Table 3.11. W cross sections are indicated for completeness. There is a sizeable reduction of uncertainty compared to individual electron and muon channel measurements, since most of the uncertainties are uncorrelated. The results both on the fiducial and total cross sections are in agreement with the calculations presented in Section 3.4 within $\sim 1\sigma$.

The HERAverager method has been cross-checked using the Bayesian Analysis Toolkit (BAT) [67]. BAT allows to evaluate the probability densities applying the Bayes formula and using the Markov Chain Monte Carlo (MCMC) technique to sample the posterior probability density. The two methods provide compatible results.

The measured cross sections are compared to the theoretical predictions presented in Section 3.4 and using four different PDF sets: CT14NNLO, NNPDF3.0, MMHT14NNLO68CL and ABM12LHC. The experimental uncertainty is comparable to the theoretical precision, whose largest contribu-

Fiducial Cross Sections	
Channel	value $\pm$ stat $\pm$ syst $\pm$ lumi (pb)
Z	$778.6 \pm 2.8 \pm 5.6 \pm 16.4$
W^-	$3497.9 \pm 6.2 \pm 72.7 \pm 73.5$
W^+	$4530.8 \pm 7.1 \pm 90.7 \pm 95.1$
$W^\pm$	$8028.7 \pm 9.5 \pm 160.7 \pm 168.6$
Total Cross Sections	
Channel	value $\pm$ stat $\pm$ syst $\pm$ lumi (pb)
Z	$1981.2 \pm 7.0 \pm 38.1 \pm 41.6$
W^-	$8788.7 \pm 15.7 \pm 239.3 \pm 184.6$
W^+	$11829.8 \pm 18.6 \pm 320.7 \pm 248.4$
$W^\pm$	$20639.3 \pm 24.4 \pm 555.6 \pm 433.4$

Table 3.11: Results for the fiducial and total cross sections of the Z- and W-boson production, in the combined electron and muon channel.

tion is given by the PDF uncertainty Fig.3.9. In Fig.3.10, the cross sections are compared to the theoretical predictions and to the previous measurements performed by ATLAS, CMS, UA1 and UA2 (at the centre of mass energy $\sqrt{s}$=0.63 TeV at the CERN Sp$\bar{p}$S) and CDF and D0 experiments (at $\sqrt{s}$=1.8 TeV and $\sqrt{s}$= 1.96 TeV at the Tevatron collider). All measurements are in good agreement with the theoretical predictions and their behaviour as a function of the centre of mass energy is well described.

3.11   W/Z cross section ratio

The W/Z cross section ratios has been measured in the past by the ATLAS and CMS collaborations [92]. The ratios of the cross section allow to perform high precision measurements, benefiting from the cancellation of many experimental uncertainties. The ratios are sensitive to the quark distributions inside protons: the ratio of W^+ to W^- (not discussed here) is sensitive to the difference between up- and down-valence quarks at low-Bjorken-x (u_v-d_v), while the $W^\pm$ to Z ratio constraints the strange-quark distribution in
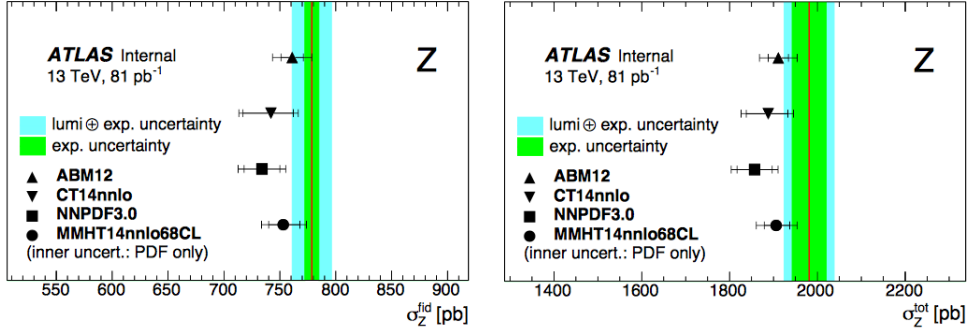
Figure 3.9: The measured fiducial (left) and total (right) cross sections (straight line) are compared with predictions produced with four PDFs: CT14NNLO, NNPDF3.0, MMHT14NNLO68CL and ABM12LHC. The green (cyan) band corresponds to the total experimental uncertainty without (with) the luminosity contribution. The inner error bar of the predictions represents PDF uncertainty, while the outer bar corresponds to the sum in quadrature of all the uncertainties.
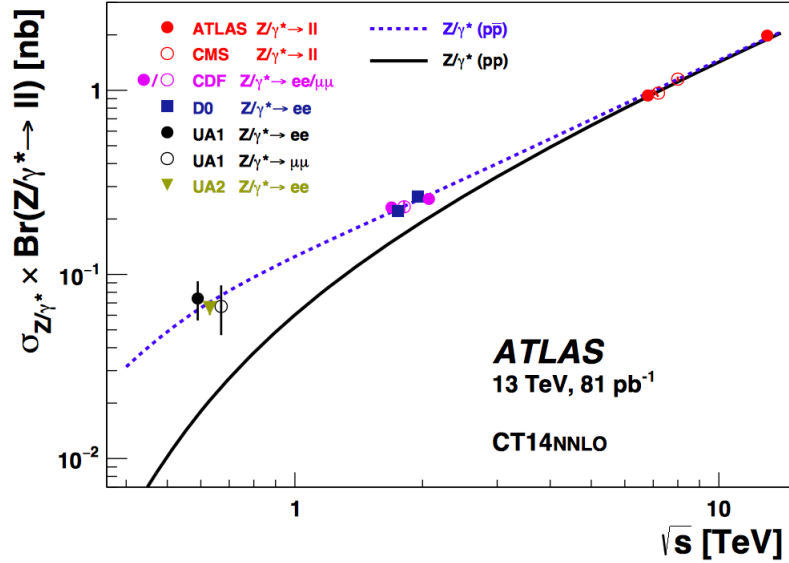

Figure 3.10: The measured value of $\sigma_{Z/\gamma^*} \times BR(Z/\gamma^* \to l^+l^-)$ where the electron and muon channels have been combined. The predictions are shown for both proton-proton (line) and proton-antiproton (dashed line) collisions as a function of the centre of mass energy $\sqrt{s}$. The CDF and D0 measurements are presented at the centre of mass energies $\sqrt{s} = 1.8$ TeV and $\sqrt{s} = 1.96$ TeV. All data points are displayed with the total uncertainty, except for the theoretical one, which is not shown.

the proton.

In the ratios, the systematic uncertainties are uncorrelated between the electron and muon channels, with the exception of the luminosity uncertainty. On the other hand, there is a significant correlation between $W^{\pm}$ and Z uncertainties for the same flavour measurement.

The ratio can be performed using separate electron and muon cross section results first and combining them later or taking ratios of already combined measurements. Both approaches have been tried as a cross check and the results for the two different strategies agree within the uncertainties:

$$
\begin{aligned}
\text{e} - \text{channel} \qquad & \sigma_{W^{\pm}}/\sigma_Z = 10.59 \pm 0.06 \pm 0.32 \\
\mu - \text{channel} \qquad & \sigma_{W^{\pm}}/\sigma_Z = 10.27 \pm 0.05 \pm 0.21 \\
\text{combined} \qquad & \sigma_{W^{\pm}}/\sigma_Z = 10.31 \pm 0.04 \pm 0.20
\end{aligned}
$$

The first component of uncertainty refers to statistics and the second to systematics. The ratio obtained in the electron and in the muon channels are in good agreement, while the ratio obtained using the combined result has a reduced uncertainty. Fig.3.11 shows the comparison of the measured $\sigma_{W^{\pm}}/\sigma_Z$ ratio with theory predictions. The measurement accuracy is not competitive with most of the predictions based on different PDF sets but it agrees reasonably with all of them. In Fig.3.12 the $\sigma_{W^{\pm}}/\sigma_Z$ ratio is compared with CT10NNLO, NNPDF3.0 and MMHT14NNLO68CL PDFs as a function of the centre of mass energy. The result obtained at the centre of mass energy $\sqrt{s} = 7$ TeV has been included in the comparison, after correcting to the 13 TeV fiducial phase space, by applying a factor of 1.0156 [93]. There is an excellent agreement between measurements and predictions at 7 TeV, while at 13 TeV predictions appear more scattered and only MMHT14NNLO68CL and CT14NNLO PDF set agree with data.
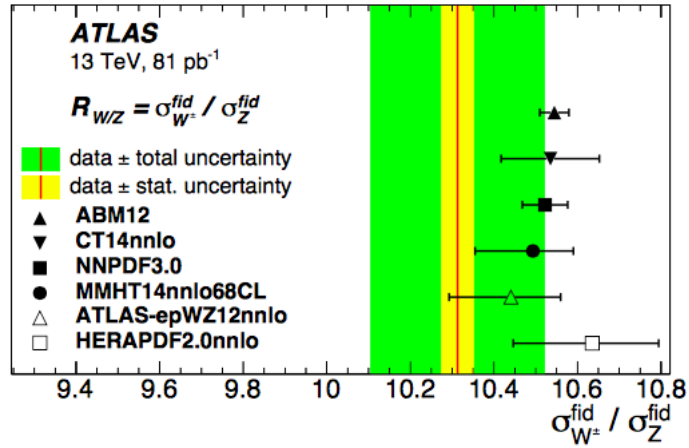
Figure 3.11: Ratio of $W^{\pm}$ and Z boson production fiducial cross sections compared with different PDF sets. The yellow shaded band corresponds to the statistical uncertainty with the green outer band refers to the statistical and systematic contributions added in quadrature. The theoretical predictions are given with the PDF uncertainty only.
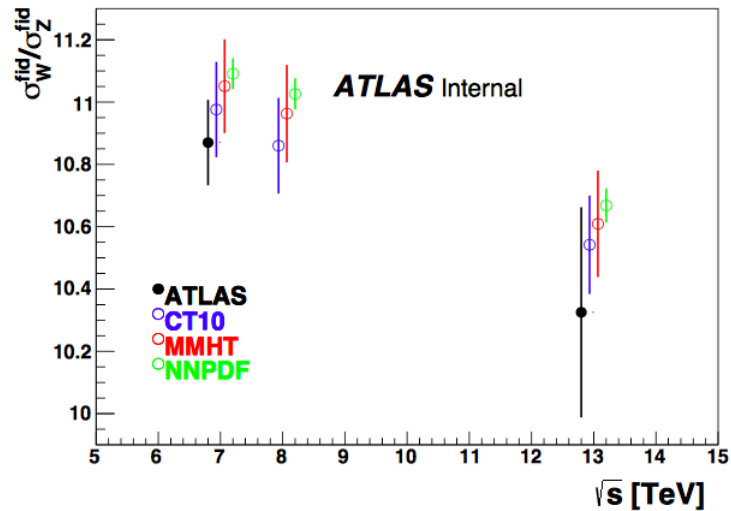

Figure 3.12: Prediction of the fiducial W and Z cross section ratio for three PDFs: CT10NNLO, NNPDF3.0 and MMHT14NNLO68CL and compared to measured value for the 13 TeV data (reported in Table 3.10) and to the ATLAS 2010 results for the 7 TeV (corrected for the 13 TeV phase space).

3.12   Conclusions

The cross section of the Z boson has been measured by ATLAS at the very beginning of Run 2 of LHC, using data collected between June and July 2015 at $\sqrt{s} = 13$ TeV. The measurement has been performed in the muon and the electron channels, obtaining a good compatibility. The combined results reaches a precision $<1\%$ (excluding a 5% of luminosity uncertainty) and allows comparisons with the most recent QCD calculations at NNLO. The ratio of the W and Z cross sections is in good agreement with various PDF sets and allow to test model of dynamics of the colliding partons, benefiting from the total or partial cancellation of several experimental uncertainties.

Chapter 4

Z+b-jets cross section measurement

The measurements of the integrated and differential cross sections of the Z boson produced in association with 1 or 2 b-jets is the main goal of this thesis. The Z+b-jets cross section measurements are performed in the muon and electron decay channels, by using the dataset collected during 2015 and 2016 at a centre of mass energy of $\sqrt{s} = 13$ TeV, corresponding to a luminosity 36.5 fb^{-1}. As described in Chapter 1, this process is sensitive to the presence of b-quarks in the initial state. Two schemes are employed in the calculations: the 4FNS, which considers PDFs of gluons and of the first two generations of quarks in the proton, and the 5FNS, which allows a b-quark PDF in the initial state. The comparison between the two schemes raises the prospect that measurements of Z+b-jets production could constrain the b-quark PDF in the proton. Furthermore the Z+b-jets is one of the major backgrounds to a large number of processes with smaller cross sections, like the SM Higgs boson production, with b-jets in the final state (for example $H \rightarrow b\bar{b}$).

The strategy of the Z+b-jets integrated and differential cross section measurements is described in this chapter and in the next one. The measurement procedure consists in a first step based on the study of the *detector* (or *reconstruction*) *level* objects, where the candidate signal events in data are selected and the background contributions are estimated and subtracted. The recon-

struction level measurements are presented in this chapter. In Chapter 5, the
signal is extrapolated to the *particle* (or *hadron*) *level*, through the unfolding
procedure.

After a description of the data and the Monte Carlo (MC) samples used in
the analysis (Sections 4.2 and 4.3), the main steps of the measurement are
presented:

- the event selection (Section 4.4);

- the estimation of the systematic uncertainties at detector level (Section
  4.5);

- the background reduction (Section 4.6), with a particular emphasis on
  the estimation of the Z+c-jets and Z+light-jets contributions, by means
  of a procedure called "flavour fit";

- the data to Monte Carlo comparison (Section 4.7);

- the unfolding procedure to obtain the final cross section results at par-
  ticle level and the estimation of the associated uncertainties (this is the
  content of Chapter 5).

4.1   Cross section definition

The differential cross section measurement foresees two separate steps:

1. reconstruction (detector) level: data and Monte Carlo events are se-
   lected based on trigger, reconstruction and analysis based cuts. At this
   stage, both data and Monte Carlo events are affected by the distor-
   tions introduced by the detector (efficiency, resolution and bin-by-bin
   migration).

2. particle (hadron) level: in order to correct data for the detector-related
   distortions, an unfolding procedure (see Section 5.2.1) is performed to
   reconstruct the unbiased differential distributions. The unfolding pro-
   cedure relies on the comparison between reconstructed and generated
   (so called "truth") Monte Carlo quantities.

In general, a differential cross section as a function of a physics observable X ca be defined as:

$$\frac{d\sigma_i}{dX_i} = \frac{M_{ij}^{-1}(N_j^{obs} - B_j)f_j}{L \cdot \Delta X_i}$$ (4.1)

where

- label i (j) indicates the particle (reconstruction) level bin of the differential distribution;

- N_j^{obs} (B_j) are the number of data (background) events selected in bin j at the reconstruction level. The difference between the two gives the number of Z+b-jets signal events in bin j;

- M_{ij} is the unfolding (migration) matrix describing the bin-to-bin migration effects and allowing to obtain a particle level measurement from the detector level one. It describes the fact that events reconstructed in bin j could have been produced in bin i. The elements of the matrix are obtained using the Monte Carlo, by comparing reconstruction to the truth quantities;

- f_j is the *fake correction* factor defined as:

$$f_j = \frac{N_j^{matched}}{N_j} \, .$$ (4.2)

    where $N_j(N_j^{matched})$ are the number of Monte Carlo events reconstructed (generated and reconstructed) in bin j. f_j accounts for events that have passed the reconstruction level selection, but not the particle level one.

- L and ΔX_i are the total integrated luminosity, which provides the overall normalisation, and the width of bin i.

Differential cross sections presented in this thesis are all performed in a fiducial volume and can be grouped in the three following categories:

1. jet multiplicities and jet properties, of primary importance for test of QCD predictions and MC modelling:

    - as a function of the inclusive b-jets multiplicities;

- as a function of the transverse momentum (p_T) of the leading b-jet for events with $N_{b-jets} \geq 1$;

- as a function of the rapidity (y) of the leading b-jet for events with $N_{b-jets} \geq 1$.

2. di-jets distributions, important for Higgs and new-physics topologies:

- as a function of $\Delta\Phi_{bb}$, the difference of the Φ coordinates of the two leading b-jets for events with $N_{b-jets} \geq 2$;

- as a function of Δy_{bb}, the difference of the rapidity of the two leading b-jets for events with $N_{b-jets} \geq 2$;

- as a function of ΔR_{bb}, the radial distance of the two leading b-jets for events with $N_{b-jets} \geq 2$;

- as a function of the invariant mass M_{bb} of the two leading b-jets for events with $N_{b-jets} \geq 2$;

3. boson and boson-jet observables, for PDF understanding:

- as a function of the Z rapidity (y_Z) for events with $N_{b-jets} \geq 1$;

- as a function of the Δy(Z,b), the difference in rapidity between the Z boson and the leading b-jet for events with $N_{b-jets} \geq 1$.

A list of ancillary differential cross section measurements are performed in order to verify the overall consistency of the analysis strategy. They are:

1. jet properties

- as a function of the transverse momentum p_T of the second leading b-jet for events with $N_{b-jets} \geq 2$;

- as a function of the rapidity y of the second leading b-jet for events with $N_{b-jets} \geq 2$.

2. boson-jet distributions:

- as a function of the $\Delta\Phi$(Z,b) between the Z boson and the leading b-jet for events with $N_{b-jets} \geq 2$;

- as a function of the $\Delta R(Z,b)$ between the Z boson and the leading b-jet for events with $N_{b-jets} \geq 2$.

The listed variables allow to investigate all the relevant kinematic features of the processes of interest.

In this thesis the integrated cross sections for Z+$\geq$1 b-jet and Z+$\geq$2 b-jets processes are also presented.

4.2 Data sample

The dataset used in this analysis was recorded during 2015 and 2016, when LHC circulated 6.5 TeV proton beams with 25 ns bunch spacing. In the 2015 data taking, the peak delivered instantaneous luminosity is $\mathcal{L}$=5$\times$10^{33} cm^{-2}s^{-1} and the mean number of p-p interactions per bunch crossing corresponds to μ=13. The peak delivered luminosity for 2016 p-p collisions is $\mathcal{L}$=1.38$\times$10^{34} cm^{-2}s^{-1} with $< \mu >$=25. In Fig.4.1, the peak luminosity recorded during 2015 and 2016 data taking periods is shown as a function of time.
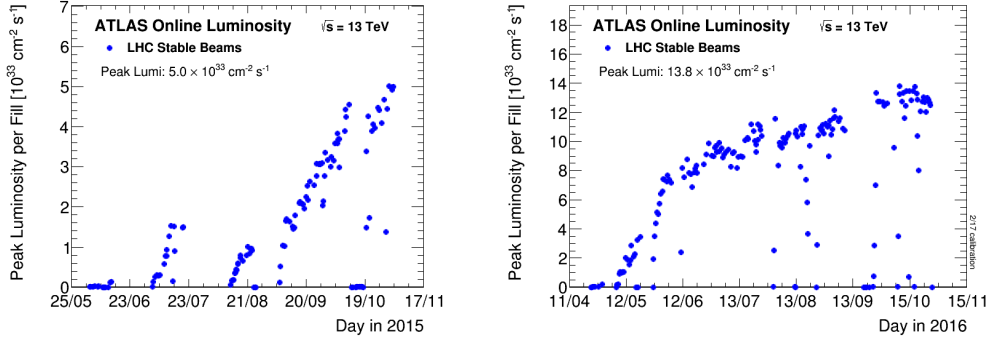

Figure 4.1: The peak instantaneous luminosity delivered to ATLAS during stable beams for p-p collisions at a centre of mass energy $\sqrt{s} = 13$ TeV is shown for each LHC fill in 2015 (left) and 2016 (right) [41].

Only events recorded during stable beam conditions, with all the ATLAS sub-systems fully operational (Good Run List, GRL), are considered. For the 2015+2016 dataset, the events passing the basic quality requirements

correspond to a total luminosity L = 36.5 fb^{-1}. The uncertainty on the total integrated luminosity is $\pm 2.1\%$, derived from special VdM scans, as described in Appendix A. The total integrated luminosity contribution per year and its uncertainty are summarised in Table 4.1. For the distributions of the total integrated luminosity as a function of time see Fig.2.3.

Year	L [fb^{-1}]
2015	$3.16 \pm 2.1\%$
2016	$33.3 \pm 2.2\%$
Total	$36.5 \pm 2.1\%$

Table 4.1: Integrated luminosity divided per year with the associated systematic uncertainty. The total integrated luminosity is shown at the bottom of the table.

4.2.1   Derivation Framework

The data format must contain all the relevant information of triggered events and must be suitable for a fast access and easy usage. A reduction of the amount of recorded data is therefore performed, in order to select only interesting events and speed up the process of managing data files. In Run 2, the reduction of the acquired data (referred to as "derivations") is made by the Derivation Framework, which provides off-line tools to select interesting events. For 2015 and 2016 datasets, derivations have been performed starting from the output of general reconstruction framework ATHENA, in a format called xAOD and providing an output with the same general format, but with reduced size called DxAOD (Derived-xAOD).

Fig.4.2 shows the derivation procedure, based on three operations. The full set of reconstructed events is first reduced through the *skimming* procedure, which consists on the removal of not interesting events. It is followed by the *thinning*, which removes not useful reconstructed objects from an event, but keeping the rest of the event, and finally by the *slimming*, which deletes not necessary information from objects. Derivations have been defined by individual teams according to the specific analysis needs. In this thesis, the STDM4 DxAODs are used: produced for Standard Model purposes, they are

characterised by the presence of at least one lepton in each event, passing single lepton triggers [94]. The skimming and thinning requirements developed for the Z+b-jets analysis are explained in Section 4.4. Then, by means of the CxAOD Framework, a calibration is applied to physics objects and the events are saved in a format similar to DxAODs called CxAODs (Calibrated xAOD).
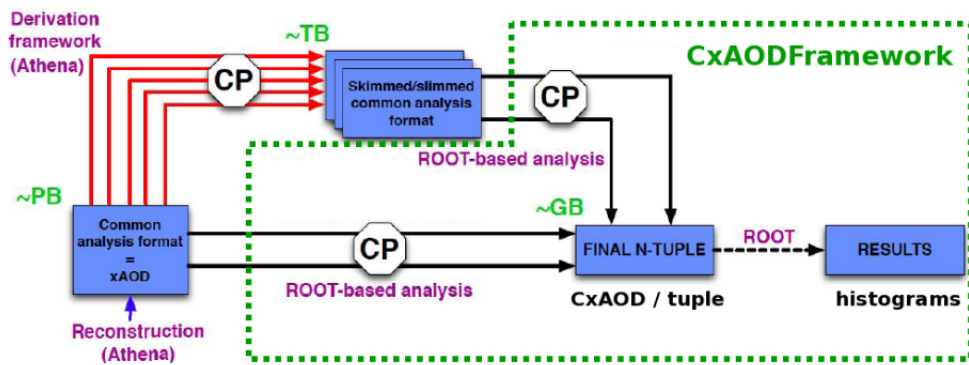

Figure 4.2: Scheme of the Derivation Framework used by the ATLAS Collaboration in Run 2. Data are reconstructed with ATHENA and derived by the Derivation Framework in DxAOD samples with smaller size. CxAODs contain information of the physics objects and their calibrations.

4.3 Monte Carlo samples

In the cross section measurements of the Z+b-jets, Monte Carlo (MC) samples are needed for the signal process simulation and to unfold data from reconstruction to particle level. In addition, Monte Carlo generators are used to estimate the background sources, whose contributions is explained in Section 4.6.

All MC samples are normalised to the highest-order available cross section calculation.

Process	Generator	PDF	PS	Normalisation	Unc.
Z+b-jets Z+c-jets Z+light-jets	Sherpa	NNPDF3.0	Sherpa	NNLO	5%
Z+b-jets Z+c-jets Z+light-jets	Alpgen	CTEQ6L1	Pythia 6	NNLO	5%
Z+b-jets Z+c-jets Z+light-jets	Madgraph	NNPD2.3	Pythia 8	NNLO	5%
$t\bar{t}$	POWHEG		Pythia6	NNLO+NNLL	6%
Single top-s Single top-t Single top-Wt	POWHEG		Pythia 6	NLO+NNLL	6%
Diboson	Sherpa	CT10	Sherpa	NLO	6%
W+jets	Sherpa	NNPDF3.0	Sherpa	NNLO	5%
W+jets	Alpgen	CTEQ6L1	Pythia 6	NNLO	5%
W+jets	Madgraph	NNPD2.3	Pythia 8	NNLO	5%
$Z(\to \tau\tau)$+jets	Sherpa	NNPDF3.0	Sherpa	NNLO	5%
$Z(\to \tau\tau)$+jets	Alpgen	CTEQ6L1	Pythia 6	NNLO	5%
$Z(\to \tau\tau)$+jets	Madgraph	NNPD2.3	Pythia 8	NNLO	5%

Table 4.2: A summary of signal and background Monte Carlo samples and the generators used in the simulation.

Signal processes

To describe the Z+jets events, including the Z+b-jets signal, the Sherpa 2.2.1 MC is used [95]. Sherpa is a Parton Shower (PS) Monte Carlo generator simulating additional parton emissions. The merging of multi-parton matrix elements with PS is achieved using the CKKW matching procedure [96], so to divide the phase space for parton emissions into the hard region of jet production and the softer regime of jet evolution. Then Matrix Elements (ME) for different parton multiplicities are used to describe the production of corresponding number of jets and Sudakov factors are used to introduce leading higher order corrections to ME. The ME and PS matching is extended to NLO accuracy using the MEPS@NLO prescription [97] with a matching scale of 20 GeV. Matrix elements are generated up to two additional parton

emissions at NLO accuracy and up to four additional parton emissions at LO accuracy. All additional jets are produced by PS. SHERPA 2.2.1 uses the 5 FNS with massless b- and c-quarks in the ME and massive quarks in PS. The PDF set used is NNPDF3.0nnlo [86] with $\alpha_S = 0.018$.

Alternative samples for the Z+jets production at 13 TeV are produced using ALPGEN v2.14 [98], a LO ME generator, interfaced with Pythia v.6.426 [75] to model PS with the Perugia2011C tune [99], using the nominal PDF set CTEQ6L1 [100]. Matrix elements are calculated up to five additional partons, using the MLM prescription [101] with a merging scale of 20 GeV. The MLM merging procedure provides the overlap removal between partons jet production from ME and from PS. Processes characterised by the Z boson produced with up to five partons are generated in separate samples and, since no distinction is made about the flavour of the parton used in the calculation, they are inclusive. Predictions follow a 4FNS for the production of heavy-flavour jets. Explicit ME evaluation, with massive treatment of the heavy flavour quarks, is also possible for Z+bb and Z+cc. Each of the previous samples is produced separately, with up to five partons, and a heavy-flavour overlap removal (HFOR) procedure is used to combine the inclusive and the Z+b-jets(c-jets) samples. The HFOR is needed to avoid the double counting of phase space between different samples, because heavy flavour quarks are produced by both the gluon-gluon splitting in the PS and by matrix element calculations. HFOR requires the removal of events containing two heavy flavour quarks arising from the PS with an angular separation $\Delta R(q, \bar{q}) > 0.4$, as well as arising from the ME and with $\Delta R(q, \bar{q}) < 0.4$. Only quark anti-quark pairs or quarks belonging to the same weak doublet are considered in this removal, while quarks produced by multi-parton interactions or originating directly from PDFs are not included. HFOR rejects from 5% to 30% of generated events. In this analysis, the ALPGEN samples are used for comparison with SHERPA in both the flavour fit and the cross section measurements.

Another generator used to simulate Z+b-jets signal is MADGRAPH 5_AMC @NLO v2.2.2 (referred to as MADGRAPH in the following). It is interfaced with Pythia v8.186 [71] for the modelling of PS and underlying event, using

the CKKW-L matching procedure, with a merging scale of 30 GeV. The LO
ME generates up to four partons, while all the other jets are simulated by PS.
The A14 tune [79] of Pythia 8 is used and the EvtGen v1.2.0 program [73] is
used for bottom and charm hadron decays. MADGRAPH provides the 5FNS
for the calculations of heavy flavour jets, with massless b- and c-quarks in the
matrix element, and massive quarks in the Pythis 8 shower. The NNPD2.3
LO PDF set is used, with α_s=0.13 [86]. In this analysis, the MADGRAPH
samples are used for closure tests in the flavour fit and for comparison with
unfolded results.

Top background

The top background production (both single top quark and $t\bar{t}$ pairs) is
generated with POWHEG-Box v2 [68]. This generator is combined with
Pythia v.6.4 for PS using Perugia 2012 tune. The $t\bar{t}$ sample is normalised to
the cross section calculated at NNLO+NNLL with the Top++2.0 program
[78]. The POWHEG matching between ME and PS is set to the top quark
mass (172.5 GeV).

Diboson background

Another source of background is given by the diboson processes, with
one of the bosons decaying hadronically and the other leptonically. Diboson
events are generated using SHERPA v2.1.1 generator with CT10 PDF set.
The matrix element generates ZZ, WW and WZ processes with zero or one
parton at NLO and up to three partons at LO. The merging between ME
and PS is driven by the MEPS@NLO prescription.

W+jets background

W+jets background processes are simulated with SHERPA 2.2.1 Monte
Carlo generator, with the same features used for the signal modelling. Other
simulations come from ALPGEN and MADGRAPH and are used for compari-

son.

Pileup

Multiple p-p collisions are simulated with the soft QCD processes of Pythia, using A2 tune [79] and the MSTW2008LO PDF. The pileup distributions of the Monte Carlo samples have been reweighted so that the distribution matches the observed pileup in data.

All Z+jets and W+jets samples are normalised to the inclusive NNLO Z and W cross sections respectively, as provided by FEWZ program [83].
The generated signal and background processes are passed through the full ATLAS detector simulation using GEANT 4 and are then processed through the same reconstruction software of the data. In Table 4.2, for each physics process the Monte Carlo used is specified, with a summary of the PDF set used, the parton shower method and the normalisation order used.

4.4 Event Selection

The Z+b-jets analysis is based on physics requirements, needed to identify with high efficiency the signal events and to the reject background. The selection of the events of interest is described in detail in this Section, while for object reconstruction, I refer to Chapter 2.

4.4.1 Preselection

Events passing the GRL are required to have at least one primary vertex with at least two associated tracks. The xAODs samples are reduced to CxAODs by applying the following selection cuts:

- only events with at least 1 jet ($\geq$1 jet) and one lepton (electron or muon) are stored (skimming);

- electrons with at least loose identification, p_T >7 GeV and $|\eta|$ <2.47 (thinning);

- muons with at least loose identification, $p_T >7$ GeV and $|\eta| <2.5$ (thinning);

- jets must satisfy $p_T >20$ GeV and $|y| <4.5$ (thinning).

4.4.2 Physics object selections

Lepton Selection - Electrons		

Triggers	2015
	`HLT_e24_lhmedium_L1EM18VH`
	`HLT_e24_lhmedium_L1EM20VH`
	`HLT_e60_lhmedium`
	`HLT_e120_hloose`
		2016
		`HLT_e24_lhtight_nod0_ivarloose`
		`HLT_e26_lhtight_nod0_ivarloose`
		`HLT_e60_lhmedium_nod0`
		`HLT_e60_lhmedium`
		`HLT_e140_lhloose_nod0`
		`HLT_e300_lhloose_nod0`

ID Isolation p_T η d_0, z_0	Tight LH FixedCutTight >27 GeV $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$ $|d0signBL| <5$, $|z0BL*sin\theta| < 0.5$ mm	
Lepton Selection - Muons		

Trigger	2015
	`HLT_mu20_iloose_L1MU15`
	`HLT_mu40`
		2016
		`HLT_mu24_iloose`
		`HLT_mu24_ivarmedium`
		`HLT_mu26_ivarmedium`
		`HLT_mu40`
		`HLT_mu50`

ID Isolation p_T η d_0, z_0	Medium FixedCutTightTrackOnly >27 GeV $|\eta| < 2.4$ $|d0signBL| <3$, $|z0BL*sin\theta| < 0.5$ mm	

Table 4.3: Overview of the selection requirement for electrons and muons at the detector level.

Electrons

Electron candidates are reconstructed from isolated energy deposits in the elctromagnetic calorimeter, matched to a track in the inner detector. Candidates must satisfy at least one of the triggers reported in Table 4.3, which correspond to a set of requirements based on the p_T, the identification, the isolation and the transverse parameter (i.e. `HLT_e24_lhtight_nod0_ivarloose` trigger selects events with $p_T >$24 GeV, "lhtight" identification, "ivarloose" isolation and no requirement on the transverse parameter "nod0").
Electrons are identified by the likelihood algorithm at the tight working point (Tight LH) and are required to be isolated, fulfilling the FixedCutTight criterion (see Chapter 2). Electrons must have a transverse momentum larger than 27 GeV and fall in the pseudorapidity region $|\eta| <$2.47, excluding the calorimeter crack region 1.37$< |\eta| <$1.52.
Electrons tracks are associated with the primary vertex, using constraints on the transverse and longitudinal parameters. In particular, the transverse impact parameter is evaluated considering information of the B-layer (IBL) and its significance must be $|d0sigBL| <$5; the longitudinal impact parameter is corrected by the reconstruction position of the primary vertex and must be $|z0BL * sin\theta| <$0.5 mm, where θ is the polar angle.

Muons

Events containing muons are demanded to pass the triggers listed in Table 4.3. Muon candidates are reconstructed in the region $|\eta| <$2.4, by matching tracks in the muon spectrometer with tracks in the inner detector. Muons are identified by the medium algorithm and must satisfy the FixedCutTight-TrackOnly isolation (see Chapter 2). Muons are required with a transverse momentum larger than 27 GeV. Tracks are associated to the primary vertex with constraints on the transverse and longitudinal impact parameters. Considering the information from the IBL, the transverse impact parameter significance must be $|d0sigBL| <$3. The longitudinal parameter, corrected by the position of the primary vertex, must be $|z0BL * sin\theta| <$0.5 mm.

Missing transverse energy

The missing transverse energy (E_T^{miss}) is reconstructed from momentum imbalance, as the sum of negative vectors of selected high-p_T calibrated objects (electrons, muons and jets) and of the soft-event contribution, which is reconstructed from tracks or calorimeter cell clusters not associated to the hard objects. The pileup present in data degrades the resolution of the calorimeter-based measurement of the missing transverse momentum: for this reason a track-based measurement of the soft objects is used as a default. As discussed in Chapter 2, the E_T^{miss} quantity is linked to the presence of neutrinos in the event. Since in this analysis neutrinos are not present in the final state, a cut on E_T^{miss} is required to reduce the top background, as detailed in Section 4.6.

Jets and b-jets

The jet and b-jet selection requirements are specified in Table 4.4. Jets are reconstructed with the anti-k$_T$ algorithm with radius parameter R=0.4, by using topological clusters in the calorimeter. Jets are calibrated using the E_T- and η-dependent simulation based scheme presented in Chapter 2. Jets must have a p_T larger than 20 GeV and fall within $|y|$ <4.5. In order to reduce the pileup, a significant amount of tracks are required to have origin compatible with the primary vertex. The discriminant for the jet vertex tagger (JVT) is required to be larger than 0.59 for jets with 20< p_T <60 GeV and $|\eta|$ <2.4. Jets must pass the jet cleaning requirement (LooseBad operative point), which allows to remove fake contaminations with high efficiency [103].
Jets are considered as originating from a b-quark (b-tagged) if they fulfil the following requirement: the MV2c10 weight has to be larger than 0.65, corresponding to a 77% efficiency for b-originated jets (see Chapter 2). The b-tagging efficiency is measured trough pseudo-continuous calibrations, which make use of the MV2c10 tagger. The distribution of this variable is divided into five bins, defined by the operating point requirements. The pseudo-continuous calibration allows for separation of jets in five classes, according to the tightest working point with which they are tagged. b-jets are required

to have a transverse momentum p_T >20 GeV and to fall in the rapidity region $|y|$ <2.5.

Jet Selection	
Algorithm	anti-k_t with R=0.4
pT	>20 GeV
y	$|y|$ <2.5
JVT	Medium for p_T <60 GeV or p_T <20 GeV and $|\eta|$ <4.5
Overlap Removal	$\Delta R(lep, jet) > 0.4$ && lepton-jet OR
b-jets Selection	
Algorithm	MV2c10 > 0.65
pT	$p_{T>}$20 GeV
y	$|y| < 2.5$

Table 4.4: Overview of the jet and b-jets selection criteria applied at detector level.

The overlap removal

The overlap removal ensures that the same particle leaving its signature in the detector, is not reconstructed as two different objects in the same event. This procedure is applied to calibrated leptons and jets that have passed the quality and kinematic selections described previously. The overlap removal is a sequential algorithm that acts differently in the two lepton channels considered in this analysis. The overlap removal between the electron and the jet (electron-jet OR) is based on two separate steps:

1) remove the non b-tagged jet that duplicates an electron in an inner cone in ΔR(jet, e)<0.2;

2) remove the isolated electron near the surviving jets, considering in this case a dynamic scale where the distance is a function of the tranverse momentum of the electron ΔR(jet, e)<min(0.4, 0.04+10 GeV/p_T^e).

The overlap removal between the muon and the jet (muon-jet OR) is also based on two steps;

1) remove non b-tagged jets in ΔR(jet, μ)<0.2. In order to account for the large number of jets misidentified as muons ("fake"), another requirement is added based on the number of tracks associated to the jets. According to this, the jet is removed if the number of associated tracks with p_T >500 MeV is <3 or p_T^{μ}/p_T^{jet} >0.7;

2) remove muons within a dynamic distance from the surviving jets ΔR(jet, μ)<min(0.4, 0.04+10 GeV/p_T^{μ}).

The driving principle behind the ΔR(jet, lepton)<0.2 cut on jets is that electron clusters plus bremsstrahlung should be contained in such one. However if the electron clusters are still very close to the area of well reconstructed jet, then we should consider the object in the other collection. In the muon case, the track-counting within an overlying jets allows to remove the ambiguity between what is just a muon and what is a jet containing a semi-leptonic decay.

In addition, jets that have passed the previous procedure are removed if within ΔR(jet, lepton)<0.4 from the selected leptons.

4.4.3   Z+jets selection

The Z boson is reconstructed by requiring exactly two leptons with the same flavour and opposite charge. At least one of the two selected leptons is required to pass the "trigger matching", thus to match the lepton that has triggered the event. In addition, events containing two leptons of the same flavour produced with a third lepton with different flavour are vetoed. The considered leptons have passed all the cuts of the selection chain and the overlap removal requirement. In order to select Z boson events, the invariant mass of the 2 leptons in the final state is required to be 76 GeV< m_{ll} <106 GeV. In this analysis, the Z boson is always considered in associated production to at least one jest: this is due the skimming procedure (see Section 4.4.1). In Table 4.5, the criteria used to select Z candidate events with at least one jet are presented at reco level. For completeness, the cut on the missing transverse energy is indicated, since events not passing the requirement are discarded, but it will be explained more in detail in Section 4.6.

Z+≥jets Selection	
Vertex	1 primary vertex with ≥2 tracks
Leptons	exactly 2 leptons with same flavour and opposite charge
Mass window	76 GeV< m_{ll} <106 GeV
E_T^{miss}	<60 GeV
Jets	at least 1 jet

Table 4.5: Selection requirements for Z+jets events at detector level. Leptons and jets must fulfil the criteria detailed in Table 4.3 and Table 4.4, respectively.

4.4.4 Z+b-jets selection

The main goal of this thesis is the measurement of the cross section of the Z boson produced in association to 1 or 2 b-jets. In order to define these two signal regions, additional requirement must be introduced.

	Z+1 b-jets Selection	Z+2 b-jets Selection
Data	exactly 1 b-jet veto on additional jets	exactly 2 b-jets veto on additional jets
MC	flavour label Z+b exactly 1 b-jet veto on additional jets	flavour label Z+bb exactly 2 b-jets veto on additional jets

Table 4.6: Definition of Z+1 b-jet and Z+2 b-jets signal regions, starting from Z+jets event selection.

In data, events are required to have exactly one b-jets (or two b-jets), passing the selection in Table 4.4. A veto on additional jets with p_T >20 GeV and $|y|$ <2.5 is also applied, as summarised in Table 4.6.

In the simulation samples, a flavour label is assigned to each jet of the event that have passed the selection of Table 4.4, in order to identify among jets, which ones come from a b- or c-hadron decay. The label is assigned combining Monte Carlo reconstruction and truth information, by matching jets at reconstruction level to weakly decaying b- and c-hadrons, at truth level, with p_T >5 GeV in a cone of radius ΔR=0.3 [104]. Therefore if a b-hadron is found within the cone, the jet is labelled as a b-jet; if no b-hadron is found,

the search is repeated for c-hadrons. If no match is found for b-hadrons and c-hadrons, the jet is labelled as light-jet. This procedure is exclusive since the hadron is matched only to the closest jet. In this analysis, the flavour tagging labelling of jets is used to characterise the event. For each event the number of jets flavour-labelled as b-, c- and light- are counted and according to these numbers, the Z+jets events are classified as follows:

- Z+b, with exactly one jet flavour labelled as b-jet;

- Z+bb, with more than one b-jet;

- Z+c, with exactly one jet flavour labelled as c-jet;

- Z+cc, with more than one c-jet;

- Z+l, not containing any b- and c-jets.

The flavour composition of the Z+jets processes simulated by SHERPA is displayed in Fig.4.3. In all the events passing the Z selection, independent of the flavour labelling, jets are required to pass the b-tagging algorithm (see Section 4.4.2). Hence, for example, a flavour-labelled light-jet can be defined as a b-jet or a c-jet, by the b-tagging algorithm. A detailed discussion on the estimation of the correct assignment of jet flavours is presented in Section 4.6.3.

In the Z+1 b-jet channel, only events flavour-labelled as "Z+b" are selected as signal. These events must have passed the Z selection and must contain exactly one b-jet. A veto is imposed on additional jets with p_T >20 GeV and $|y|$ <2.5. In the Z+2 b-jets analysis, only events flavour classified as "Z+bb" are considered. The events must have exactly two b-jets, vetoing all additional jets. In Table 4.6, the selection of the Z+1 b-jet and Z+2 b-jets analysis is summarised for both data and Monte Carlo.

In Table 4.7 and Table 4.8 the cutflow for data and Monte Carlo signal events, as generated by SHERPA, is presented for the electron and muon channel, respectively.
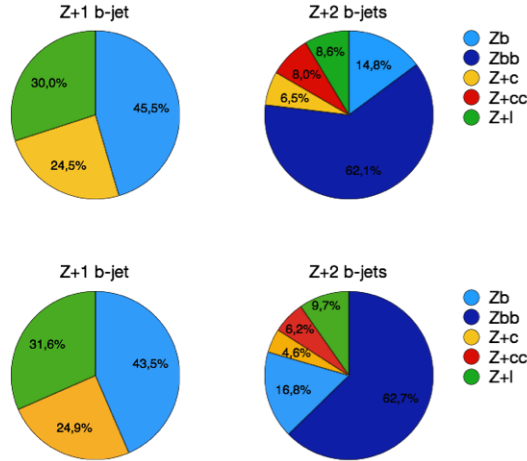
Figure 4.3: Flavour composition of Z+jets SHERPA 2.2.1 sample for the electron (top) and muon channel (bottom) for the Z+1 b-jet (left) and Z+2 b-jets (right) analyses.

Cutflow in the electron channel						
Cut	Data		Z+b (MC)		Z+bb (MC)	
Z+jets	3827605	N/A	139304	N/A	17225	N/A
Z+≥1 b-jet	266525	7.0%	99710	71.6%	16900	97.9%
Z+1 b-jet (veto jets)	111775	2.9%	4899	35.2%	0	0.0%
Z+≥2 b-jets	19955	0.5%	2351	1.7%	9542	55.3%
Z+2 b-jets (veto jets)	7580	0.2%	823	0.6%	3487	20.2%

Table 4.7: Number of candidate signal events that pass the selection in the electron channel. The cutflow is reported for data and for flavour classified "Z+b" and "Z+bb" MC events as generated by SHERPA 2.2.1 (Z+jets sample). In correspondence of each cut, the total efficiency is shown.

Cutflow in the muon channel						
Cut	Data		Z+jets 1B		Z+jets NB	
Z+jets	6890606	N/A	230574	N/A	26821	N/A
Z+≥1 b-jet	454890	6.6%	163883	71.1%	26170	97.6%
Z+1 b-jet (veto jets)	193562	2.8%	82220	35.7%	0	0.0%
Z+≥2 b-jets	31687	0.5%	3834	1.7%	14776	55.1%
Z+2 b-jets (veto jets)	11863	0.2%	1428	0.6%	5293	19.7%

Table 4.8: Number of candidate signal events that pass the selection in the muon channel. The cutflow is reported for data and for flavour classified "Z+b" and "Z+bb" MC events as generated by SHERPA 2.2.1 (Z+jets sample). In correspondence of each cut, the total efficiency is shown.

4.5   Systematic uncertainties

Systematic uncertainties are in general originated from an imperfect knowledge of the detector, an approximated theoretical modelling and the analysis procedure. In this section, a description of the systematic components related to object reconstruction is given. These uncertainty sources have an impact in the flavour fit and, therefore, in the cross section measurements.

Each systematic uncertainty is evaluated by varying the corresponding variable distribution by 1 standard deviation ($\pm 1\sigma$). This provides two different distributions of the event yields, representing the fluctuation $+1\sigma$ (up) and -1σ (down) with respect to the nominal distribution. All systematics are applied to Monte Carlo signal and background components. In case both up and down variations change the distribution in the same direction with respect to the nominal one, the resulting variation is symmetrised, namely it is assumed to be of the same size in the two directions. In the propagation of these uncertainties through the analysis, possible asymmetric variations are maintained separated.

4.5.1   Lepton

Uncertainties associated to electrons and muons arise from the trigger, reconstruction, identification and isolation efficiency, as well as scale and res-

olutions. Trigger performances are estimated on $Z \to ee$ and $Z \to \mu\mu$ events, while reconstruction, identification and isolation efficiencies are provided by Scale Factors (SF) corrections in Monte Carlo to match data, together with the corresponding statistical and systematics uncertainties. The variation within $\pm 1\sigma$ of the SFs is propagated to the analysis and provide the corresponding uncertainties. Momentum scale and resolution are obtained with the calibration procedure explained in Chapter 2. The total systematic associated to electrons is $\sim 0.2\%$ and to muons is $\sim 1.4\%$.

4.5.2   Jets

Jet Energy Scale

The Jet Energy Scale (JES) is a calibration of the energy of the reconstructed jet, tuned in order to correspond to the energy of the associated stable particle in ATLAS. This calibration ensures the correct measurement of the average energy across the whole detector. Several in-situ techniques have been developed to determine the uncertainty on the jet energy measurement [105]. The standard jet calibration and the corresponding uncertainty are determined for isolated jets in an inclusive jet data sample. The uncertainties are derived by combining information from test-beam data, LHC collisions and simulation. The full description of JES uncertainties contains 73 nuisance parameters factorised in 8 independent sources: in-situ analyses (Z+jet balance, gamma+jet balance, and multi-jet balance), high-p_T jets, pile-up, non-closure of fast simulation (AFII) calibration, flavour composition, flavour response, b-jets and η intercalibration. Each nuisance parameter is taken as a function of the jet kinematics and pileup conditions, using a reduction of the 73 components to 21. The 21 parameters characterise the JES uncertainty and are propagated through the analysis and then combined in quadrature in the final results. The JES uncertainty is one of the largest source of systematics: 2.9% (3.2%) in the electron and 2.7% (3.7%) in the muon channels for the Z+1 b-jet (Z+2 b-jets) detector level measurements.

Jet Energy Resolution

The Jet Energy Resolution (JER) is determined by exploiting the transverse momentum balance in events containing jets with large p_T [106]. The uncertainty on JER is derived by comparing Data and Monte Carlo using in-situ methods. An extra p_T smearing is added to jets, based on their kinematics, to account for possible underestimate of the jet energy resolution in simulations. Jets in Monte Carlos are oversmeared 5 times using different random seed: when errors are combined, an average of the absolute uncertainty is taken and symmetrised to form the final error band. The JER contribution is 2.7% (4.9%) in the electron channel and 3.5% (3.4%) in the muon channel for the Z+1 b-jet (Z+2 b-jets) analysis.

4.5.3   B-jets

The uncertainty related to b-tagging is implemented as event-by-event weight systematics (scale factors). The individual systematics for b-tagging are of the order of 40 elements per flavour (b, c, τ and light). In an effort to alleviate the number of variations, an analysis on the principal components has been performed, in order to reduce the number of systematic variations, resulting in three categories: loose, medium and tight. In this analysis the medium reduction scheme is applied which comprises 3 variations for b-jets (B0-B2), 4 for c-jets (C0-C3) and 5 for light-jets (L0-L4). Two additional uncertainties are not included in the reduction scheme and must be treated separately: the p_T extrapolation outside the kinematic ranges covered by these measurements (totally correlated with the other components) and extrapolation to Run 2 of the c- and light-jet measurements performed with Run 1 data. The uncertainties on b-jets are the largest systematic contribution at reconstruction level: 6.9% (8.2%) in the electron and 7.1% (10.1%) in the muon channels for the Z+1 b-jet (Z+2 b-jets) analysis.

4.5.4   Missing transverse energy

Systematic uncertainties associated to the missing transverse energy are evaluated as MC-to-data corrections. Soft terms and jets contribute to the

total uncertainty, provided as $\pm 1\sigma$ variations. MET systematics ranges from 0.7% in the muon channel to 1.4% in the electron mode.

4.5.5 Luminosity

The uncertainty on the combined integrated luminosity corresponding to 2015+2016 dataset is 2.1%. A detailed description of the procedure used to estimate such uncertainty can be found in [107].

4.6 Background estimation

The reconstructed events passing the full reconstruction selection can contain contributions coming from processes different from the one of interest. Even if reduced by the event selection, backgrounds mimic the Z+b-jets final state and must be estimated and subtracted from the observed data (see Eq.4.1).

The W+jets, diboson, $Z \to \tau\tau$+jets and single top processes are simulated using the Monte Carlo samples described in Section 4.3. Some of events generated by these background sources can pass the full selection chain if a jet (or lepton) is misidentified as a lepton (or a jet) or if physics objects fall outside the detector acceptance. These electroweak backgrounds are small (<1% for Z+1 b-jet and <2% in Z+2 b-jets), so no additional cuts are needed.

The electroweak production of top pairs is one of the dominant sources of background, since it directly provides the same final state as the signal. For this reason, a detailed study has been performed using Monte Carlo in order to optimise the criterion for its reduction. These studies are presented in Section 4.6.1.

From the previous analysis performed at the centre of mass energy $\sqrt{s} =7$ TeV, the multijet background source is expected to be negligible. Multijet is estimated through data-driven technique, based on fit on data distributions in enriched regions, as discussed in Section 4.6.2.

The largest background contributions are due to the Z+c-jets and Z+light jets processes. Their normalisation is extracted from dedicated fit on data (flavour fit) and explained in detail in Section 4.6.3.

4.6.1 $t\bar{t}$ rejection studies

The leptonic decay of the $t\bar{t}$ processes, characterised by both W boson decaying to leptons, are the second main source of background, after c- and light-jet misidentification. Even if the $t\bar{t}$ contribution is reduced by the invariant mass cut, it is still large at high jet p_T. In order to reduce it, dedicated studies on the missing transverse energy quantity have been performed. A large amount of E_T^{miss} in the final state can be due to the presence of neutrinos, produced in the decay of the W bosons from top decay, which are not present in the Z+b-jets process. For this reason the E_T^{miss} observable has a relevant discrimination power in the separation of the signal events from the top-pair ones. The $t\bar{t}$ and signal distributions are shown in Fig.4.4 and Fig.4.5 for Z+1 b jet and Z+2 b-jets, in the electron and muon channels, respectively, normalised to their theoretically predicted cross sections.
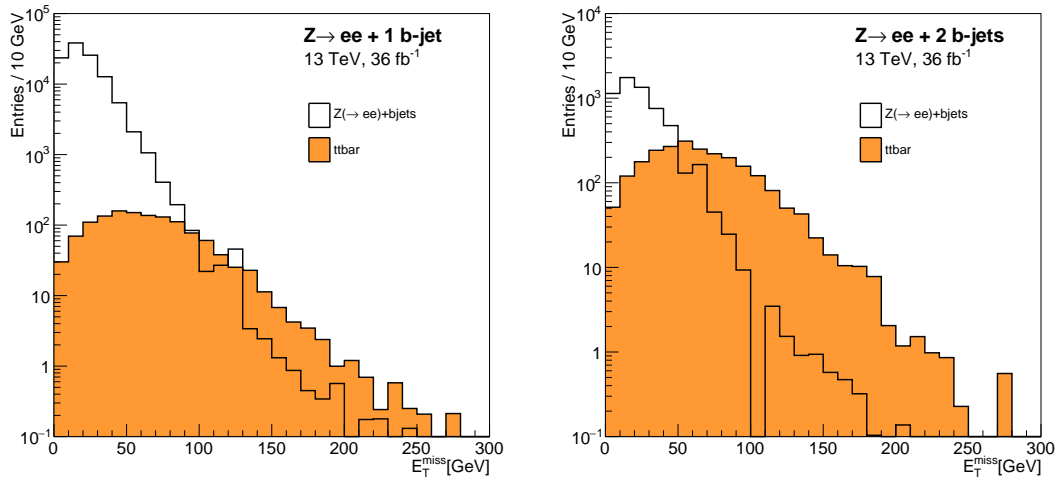
Figure 4.4: E_T^{miss} Monte Carlo distributions of $t\bar{t}$ and signal events for the $Z \to ee$+1 b-jet (left) and $Z \to ee$+2 b-jets (right) processes. Samples are normalised to their theoretical predicted cross section.

From Fig.4.4 and Fig.4.5, it is evident that the $t\bar{t}$ can be reduced by requiring the E_T^{miss} smaller than a certain value, denoted here as $E_T^{miss}|_{cut}$. The efficiency of the cut is defined as:

$$\epsilon_S = \frac{N_S|_{cut}}{N_S^{tot}} \tag{4.3}$$
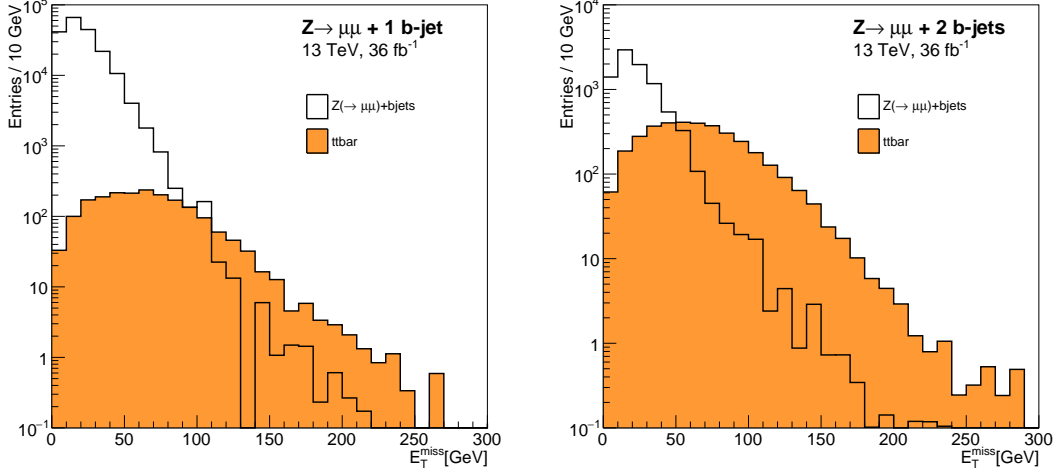
Figure 4.5: E_T^{miss} Monte Carlo distributions of $t\bar{t}$ and signal events for the $Z \to \mu\mu+1$ b-jet (left) and $Z \to \mu\mu+2$ b-jets (right) processes. Samples are normalised to their theoretical predicted cross section.

where $N_S|_{cut}$ is the number of signal (S) events passing the cut and N_S^{tot} is the total number of signal events. The choice of the $E_T^{miss}|_{cut}$ must be driven by the maximisation of the signal efficiency together with background rejection, defined as:

$$1 - \epsilon_B = 1 - \frac{N_B|_{cut}}{N_B^{tot}} \tag{4.4}$$

where ϵ_B represents the efficiency of the cut on the $t\bar{t}$ sample. The signal efficiency and the background rejection are evaluated for each value of the E_T^{miss} distributions in Fig.4.4 and Fig.4.5. The effects of the different choice of the $E_T^{miss}|_{cut}$ are shown in the ROC curves of Fig.4.6 and Fig.4.7, where the background rejection is plotted as a function of the signal efficiency.

The optimal value of the $E_T^{miss}|_{cut}$ giving the best compromise between efficiency in signal selection and background rejection is chosen using the significance $S/\sqrt{S+B}$, where S and B are the number of the resulting signal and background events. The significance curve as a function of the E_T^{miss} is shown in Fig.4.8 and Fig.4.9 for the two lepton channels. According to this study, a value of the $E_T^{miss} < 60$ GeV is chosen, as best compromise between Z+1 b-jet and Z+2 b-jets, which rejects 52% (50%) of $t\bar{t}$ background still
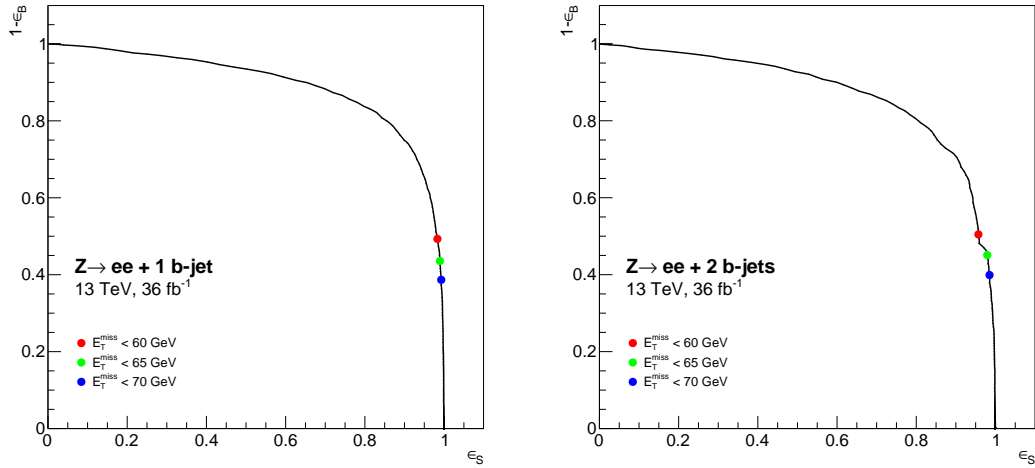
Figure 4.6: Evolution of $t\bar{t}$ background rejection and signal efficiency parametrised as a function for different E_T^{miss} cuts for Z+1 b-jet (left) and Z+2 b-jets (right) in the electron channel. Three sample values of the E_T^{miss} cut are displayed in the curve for illustration.
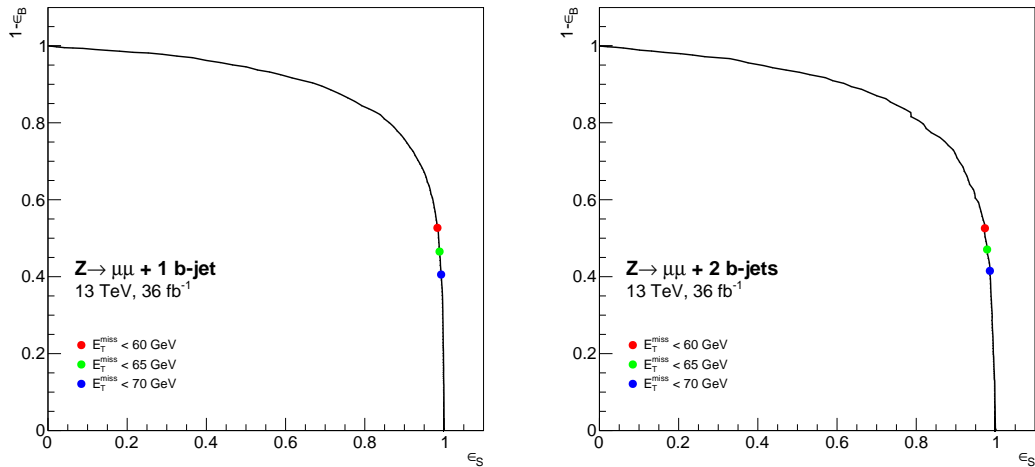
Figure 4.7: Evolution of $t\bar{t}$ background rejection and signal efficiency parametrised as a function for different E_T^{miss} cuts for Z+1 b-jet (left) and Z+2 b-jets (right) in the muon channel. Three sample values of the E_T^{miss} cut are displayed in the curve for illustration.

retaining a 98%(95%) of signal events for Z+1 b-jet (Z+2 b-jets).
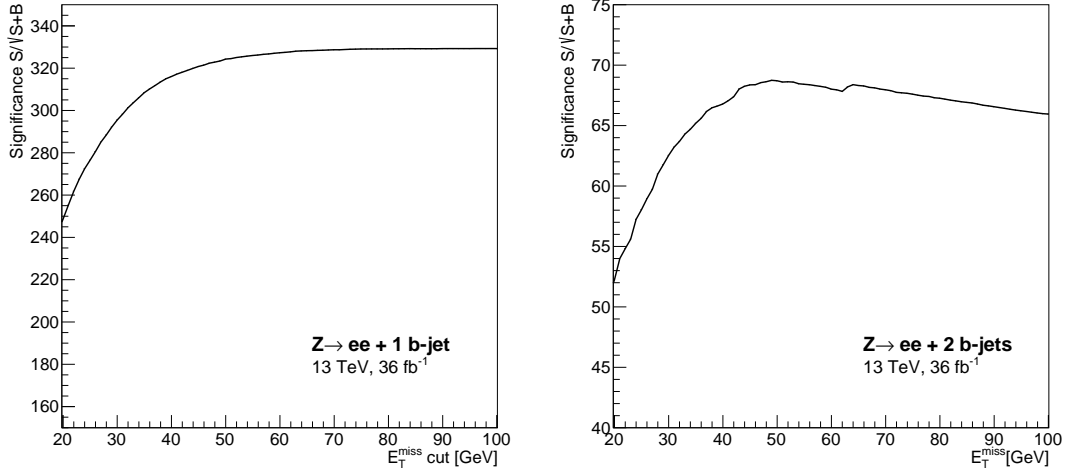
Figure 4.8: Significance distribution as a function of the E_T^{miss} for Z+1 b-jet (left) and Z+2 b-jets (right) in the electron channel.

$t\bar{t}$ modelling

The E_T^{miss} cut can introduce biases in the modelling of the $t\bar{t}$ sample. With the only purpose to verify the stability of the modelling and validate the chosen cut, data and Monte Carlo distributions are compared in a region enriched of $t\bar{t}$ background. Since in the Z+b-jets selection (see Table 4.5) events with additional jets not b-tagged are vetoed, only the dileptonic decay of the top pairs contribute as a background. The dileptonic decay is characterised by the presence of two opposite charge leptons, neutrinos and b-jets, without any specific constraint on the lepton flavour. Therefore the validation region is built selecting events passing the nominal selection, with the exception of the lepton flavour requirement: only opposite flavour leptons with opposite charge are selected.

Fig.4.10 shows the comparison between data and Monte Carlo samples in the validation region. All distributions are well modelled (within 20%) by Monte Carlo samples, whose largest contribution is given by $t\bar{t}$, suggesting the E_T^{miss} cut does not introduce any mismodelling. The cut is therefore incorporated in the full event selection chain.
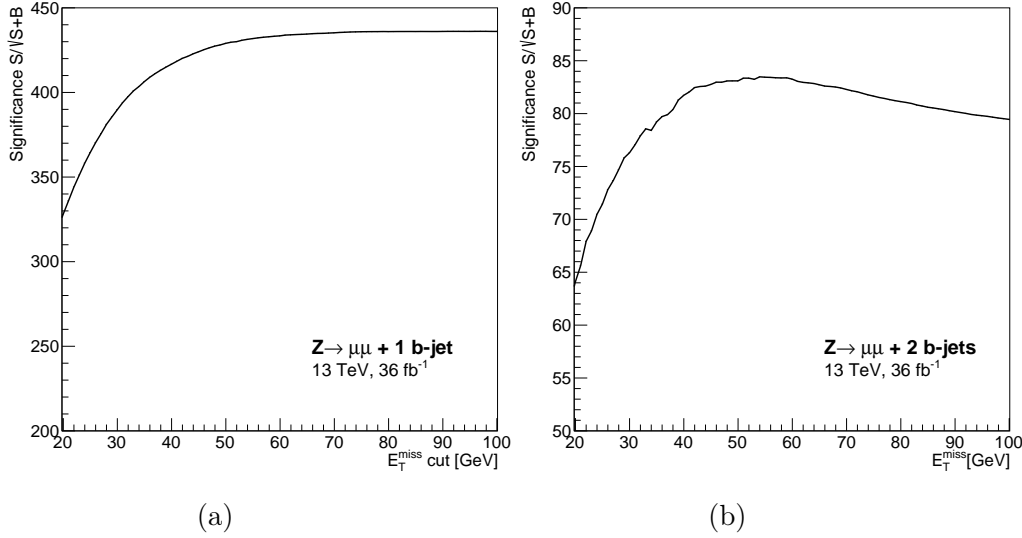
Figure 4.9: Significance distribution as a function of the E_T^{miss} for Z+1 b-jet (left) and Z+2 b-jets (right) in the muon channel.

4.6.2   Multijet data driven estimation

Multijet backgrounds to Z+b-jets production arise primarily from jets which are reconstructed as leptons (dominant in the electron channel) or real leptons coming from heavy flavour decays within jets (dominant in the muon channel). For the estimate of the multijet (or QCD) background a data-driven technique is used. This choice is due to the large uncertainty on the misidentification rate of jets. The selection of the events performed in this analysis is efficient in rejecting the QCD background, but due to the very large cross section of QCD jet production at LHC, residual events originated from different processes can contaminate the signal and the effect must be quantified.

The procedure used to estimate multijet background is based on two main steps, which are detailed in the following sections for the electron and muon channels.

- **Multijet enriched selection**. A special event selection is defined to enrich the data sample of QCD jets or fake leptons. This selection allows to build a specific kinematic region, called Control Region (CR),
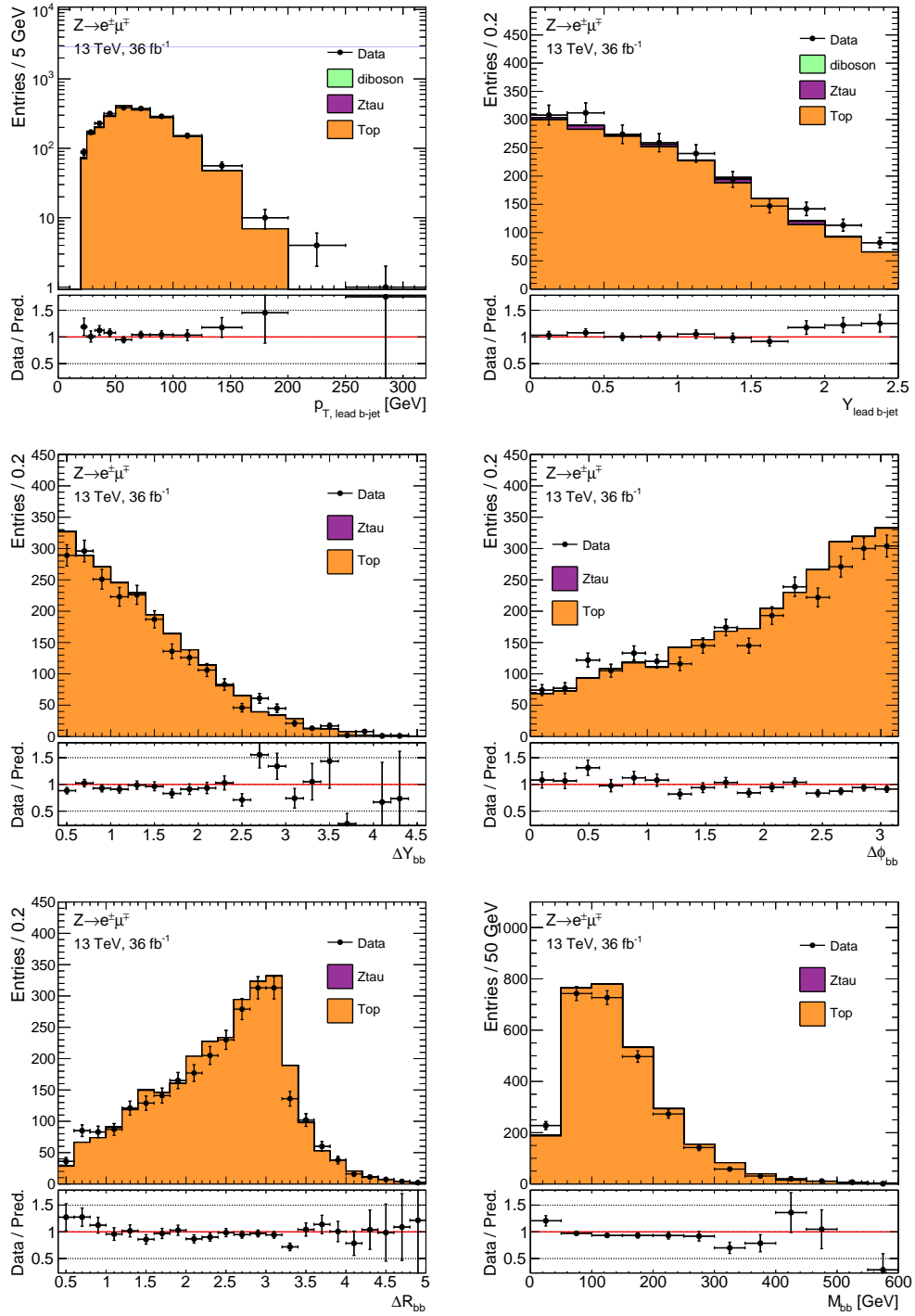
Figure 4.10: Data and Monte Carlo comparison in the $Z \to e^{\pm}\mu^{\mp}$+1 b-jet (top) and $Z \to e^{\pm}\mu^{\mp}$+2 b-jets (middle and bottom) regions. These regions are enriched by dileptonic $t\bar{t}$ background in order to validate the modelling.

where the shape of the multijet background can be extracted for a specific discriminating variable. The control region is designed to be as close as possible to the nominal selection, but with some cuts carefully changed to reduce the efficiency of signal selection and enrich the multijet sample. The template of the multijet background is given by the data events in the CR, subtracted by all processes already known and thus estimated with Monte Carlo samples in the same region.

- **Estimation of the multijet normalisation**. The number of multijet events contributing to the nominal selection is estimated with a maximum-likelihood template fit to the discriminating variable. In this analysis the dilepton invariant mass is chosen as discriminating variable. The extrapolation of the multijet contribution evaluated in the control region to the Signal Region (SR) (see Section 4.4) is driven by the realistic assumption that the multijet shape, thus the template, does not change between the two regions, only the absolute yield remaining to be quantified.

The procedure described above can be parametrised as follows:

$$N_{SR}^{data} = \alpha \text{Zjets}_{SR}^{MC} + \beta B_{SR}^{MC} + \gamma(N_{CR}^{data} - \alpha \text{Zjets}_{CR}^{MC} - \beta B_{CR}^{MC}). \qquad (4.5)$$

The number of events that pass the nominal selection (N_{SR}^{data}) is compared with Monte Carlo predictions, which are compounded by the signal (Zjets_{SR}^{MC}) and background (B_{SR}^{MC}) processes. Any missing component for the Monte Carlo to describe data is assigned to the multijet background. Each Monte Carlo sample is normalised to the cross section of the specific process it describes: α is the normalisation of the signal Zjets sample and β is the normalisation of the sum of the background components. Both α and β are allowed to float in the fit by at most $\pm 5\%$ around their nominal values. As this discussion is independent of the region we are looking at, α and β are assumed to describe the contribution of the known backgrounds also in the CR so that $N_{CR}^{data} - \alpha \text{Zjets}_{CR}^{MC} - \beta B_{CR}^{MC}$ represents the multijet background contribution in the CR. γ is the free normalisation factor, also called Scale Factor (SF), which extrapolates the number of multijet events from the con-

trol to the signal region. The maximum-likelihood fit is therefore performed leaving only γ as fully free parameter, while, as said, α and β are allowed to float by at most $\pm5\%$ from their nominal value.

The solidity of this procedure must be carefully checked due to the arbitrariness of the template selection (CR) and the choice of the discriminant variable to fit, and systematic uncertainties must be evaluated to take this into account.

In the following, the fit and the checks are discussed separately for the muon and the electron channels.

Multijet in the muon channel

In the muon channel, the multijet enriched control region (CR) can be obtained by requiring two same sign muons, namely:

- the invariant mass range is 70 GeV$< m_{\mu\mu} <$140 GeV;

- muons must have the same charge.

Starting from this selection of muons, the multijet background is estimated for two separate jet multiplicity selections corresponding to the Z+1 b-jet and Z+2 b-jets analyses: $\mu^{\pm}\mu^{\pm}$+1jet (not b-tagged) and $\mu^{\pm}\mu^{\pm}$+2jets (not b-tagged). In Fig.4.11 the comparison between data and Monte Carlo samples in the CR is presented in the $\mu^{\pm}\mu^{\pm}$+1 jet and $\mu^{\pm}\mu^{\pm}$+2 jets cases. Both selections are clearly enriched by multijet events, as can be seen by the difference between data and Monte Carlo (which does not include multijet contribution).

Apart from the lepton selection, the CR used to extract the multijet do not correspond to SR since jets are not b-tagged. For this reason the multijet fraction (γ) is not extrapolated directly in the signal regions, as explained in Eq.4.5, but in intermediate ones, called Normalisation Regions (NR). NR are defined by the same jet multiplicities as the control regions, but using nominal selection for leptons (opposite sign). In order to obtain the yield of multijet in the signal region, events in the normalisation region are required to pass the b-jet selection and the invariant mass cut. This is possible under the assumption that the shape of the multijet template is unchanged. Table
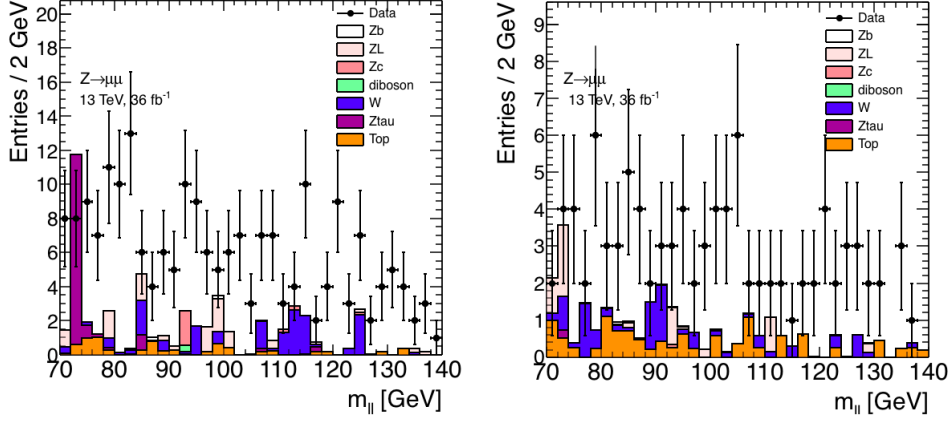
Figure 4.11: Data and Monte Carlo comparison in the region enriched by multijet events for different jet multiplicities: Z+1 jet (left) and Z+2 jets (right).

Signal Region	Normalisation Region	Control Region
$\mu^+\mu^-$+1 b-jets	$\mu^+\mu^-$ (no mass cut)+1 jet	$\mu^\pm\mu^\pm$ (no mass cut)+1 jet
$\mu^+\mu^-$+2 b-jets	$\mu^+\mu^-$ (no mass cut)+2 jets	$\mu^\pm\mu^\pm$ (no mass cut)+2 jets

Table 4.9: Correspondence among the signal, normalisation and control regions studied for the data driven multijet estimation in the Z+1 b-jet and Z+2 b-jets processes.

4.9 summarises the correspondence between the control, normalisation and signal regions.

The contributions of the multijet backgrounds to the NR are shown in Fig.4.12 for the two jet selections, together with the corresponding extracted template: apart from the different statistics, the shape of the multijet background is similar in the two selections. The multijet template is multiplied by the γ scale factor obtained from the fit in each jet selection and added to the simulated processes, for comparison with data. In both $\mu^\pm\mu^\pm$+1 jet and $\mu^\pm\mu^\pm$+2 jets normalisation regions, the plot is dominated by the signal Z+jets Monte Carlo process, followed by the sum of all simulated backgrounds. The multijet is not visible in the two plots, since its contribution is compatible with zero (see Fig.4.12). Data are well modelled, meaning the data-driven procedure does not introduce any bias in the mass distribution.
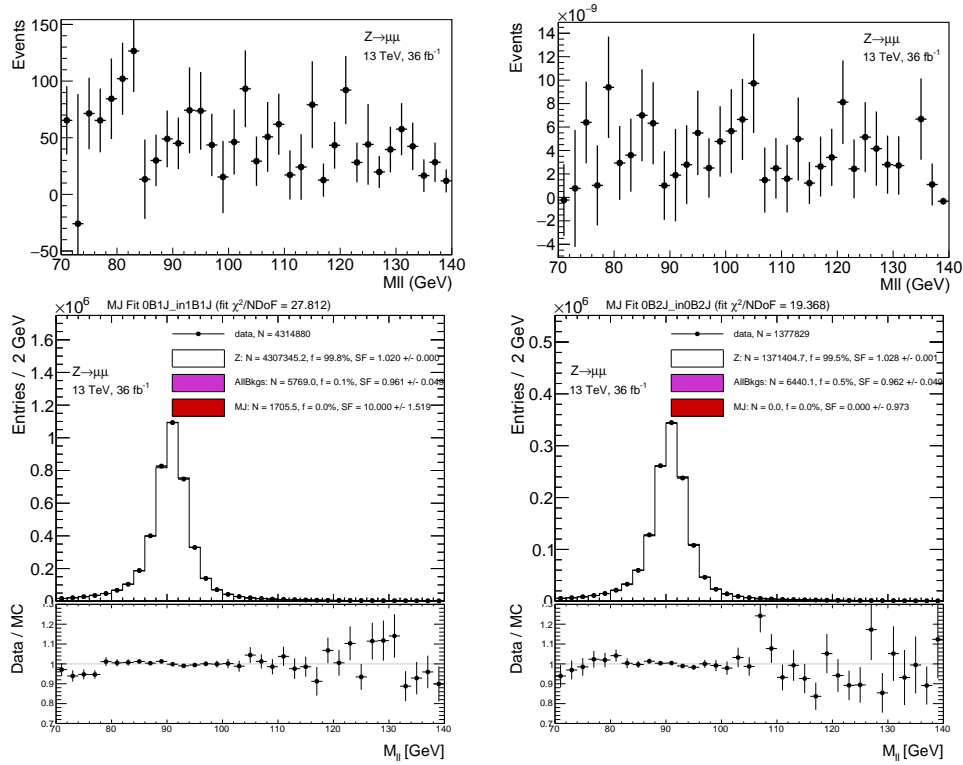
Figure 4.12: The template shape exctracted in the control region (top) for different jet multiplicities: Z+1 jet (left) and Z+2 jets (right). The multijet contribution in the muon channel estimated by a fit template on data in the mass range 70 GeV< $m_{\mu\mu}$ <140 GeV (bottom) for the described jet mulitplicity selections.

The multijet contribution is negligible also in the signal region. The measured values of γ factors are 10.0 ± 1.5 for the Z+1 b-jet channel, while it is compatible with 0 for the Z+2 b-jets one. The uncertainty on the gamma normalisation comes directly from the fit, by estimating which variation of the parameter changes the log-likelihood by one unity.
The obtained results confirm that the event selection used in the analysis is highly efficient in reducing the QCD background. For this reason this background source will not be propagated further in the analysis procedure.

Multijet in the electron channel

In analogy with the muon channel, the estimation of the multijet in the electron channel begins with the choice of the enriched control region. Several selections have been studied with the common requirement of two same sign electrons without the mass cut, but differing in the following requirements on the electrons:

- both electrons fail the isolation cut;

- the isolation requirement is not applied;

- one electron satisfies the nominal identification quality and the other one fails it;

- one electron fails the identification cut and the isolation is not applied;

- one electron fails the identification and one electron fails the isolation.

None of such CR show a sizeable multijet contribution. The invariant mass distribution in the enlarged window (70 GeV$< m_{ee} <$140 GeV) is still dominated by the Z+jets Monte Carlo sample and by the other backgrounds. An example of the multijet background study in the electron channel is given for the control region where electrons are both required to fail the isolation cut and to have same sign, being all the others of similar shape. In Fig.4.13, the data and Monte Carlo comparison is given as a function of the invariant mass of the two electrons in the enlarged mass window for two different jet selections: $e^{\pm}e^{\pm}+1$ jet (left) and $e^{\pm}e^{\pm}+2$ jets (right).
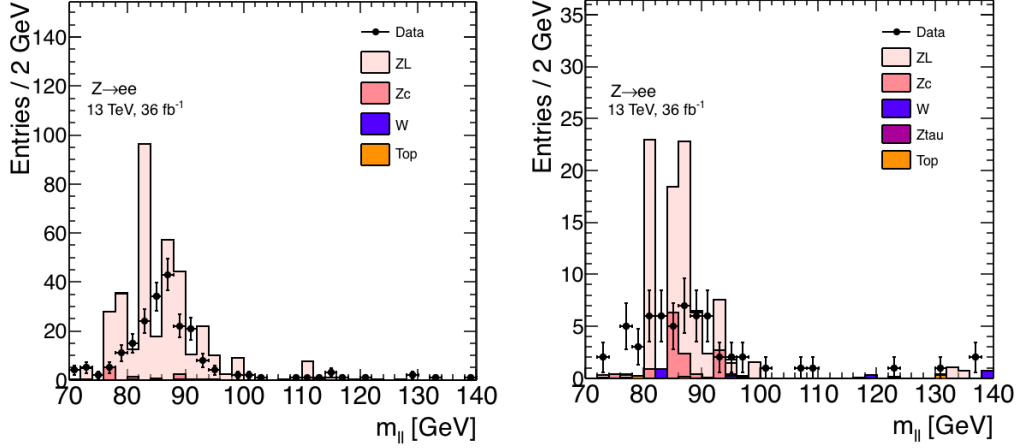
Figure 4.13: Data and Monte Carlo comparison in the region enriched by multijet events for different jet multiplicities: Z+1 jet (left), Z+2 jets (right). In this region electrons are both required to fail the isolation and to have same sign.

The dominant presence of the Z+jets events when requiring two same sign electrons is due to electron charge misidentification: high-p_T electrons have a large bending radius and thus charge misidentification is easier, by using the information of the Inner Detector, than for muons which are mostly reconstructed by the muon spectrometer.

In the electron channel, additional multijet enriched selections for the control regions are not possible because of the trigger used to acquire the data, which already contains loose requirements on the identification and the isolation, preventing from further loosening the requirements on the electron quality.

Despite the negligible number of events in the CRs, the fit on data converges in all cases and the results are all stable and comparable. The multijet contribution in the signal region is compatible with zero and will not be propagated in the rest of the analysis.

4.6.3   The flavour fit

In Fig.4.14 the distributions of Z+1 b-jet and Z+2 b-jets are shown as a function of the leading b-jet p_T and di-b-jets invariant mass M_{bb}, respectively, for data and Monte Carlo events passing the selection of Table 4.5. For the

Monte Carlo, the various jet-flavour components are shown for the SHERPA generator. As can be noted, the agreement is qualitatively satisfactory, but subjected to uncertainties, such as the dependence on the used generator both in the shape, flavour composition and b-tagging efficiency, all of them impacting the signal and background estimation. The agreement between the sum of the flavour contributions predicted by the various Monte Carlo and data is within 30%.
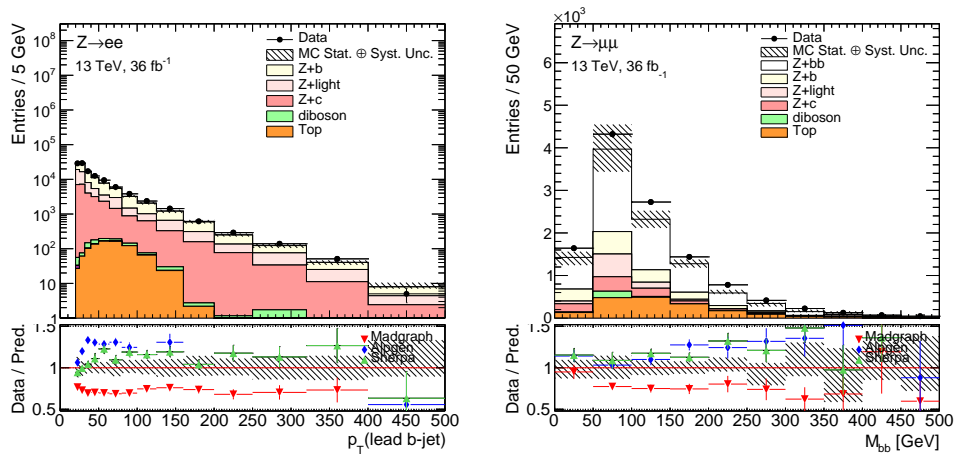
Figure 4.14: Data and Monte Carlo comparison for the distributions of the leading b-jet p_T in the $Z(\to ee)+1$ b-jet analysis (left) and of the di-bjets invariant mass M_{bb} in the $Z(\to \mu\mu)+2$ b-jets one (right), respectively. The comparison to data is shown for SHERPA samples (top) and the ratio Data/MC is evaluated for the SHERPA, MADGRAPH and ALPGEN generators (bottom). The systematics are evaluated by using SHERPA (shaded band) as a sum in quadrature of the contribution explained in Section 4.5.

A fit of the different components to the data, called "flavour fit", is therefore used to correct the flavour contributions predicted by the various generator and provide the best description of the data. The goal is to correctly evaluate the normalisation of the flavour components by extracting scale factors, which will be propagated to reconstruction level distributions (see Section 4.7) and to the unfolding procedure (see Chapter 5). The Z+b component (in the Z+1 b-jet) and the Z+bb (in the Z+2 b-jets analyses) are therefore not scaled, since they represent the signal processes, measured using data.

A profile likelihood fit is used to extract the normalisation of the different components of the Z+jets sample in the Z+1 b-jet and Z+2 b-jets processes. In the Z+1 b-jet analysis, the discriminating variable used to extract the flavour composition is the tagging weight of the identified b-jet (MV2c10), because its discriminating power is remarcable, as can be seen in Fig.4.15 (top), where the Z+b component is enhanced increasing the b-tagging weight. Three flavour components are identified by the truth labelling procedure in the Z+1 b-jets process: $Z + b$ (signal), $Z + c$ and $Z + l$, where l stands for light.

In the Z+2 b-jets analysis, the sum of the weights of the two b-jets is considered as a discriminating variable. Here five components are individuated from the flavour classification: $Z + bb$ (signal), $Z + b$, $Z + cc$, $Z + c$ and $Z + l$. The shapes and the normalisation of single top-quark, $t\bar{t}$, diboson and W+jets processes are estimated from Monte Carlo simulations, while multijet is not considered since it is negligible, as previously discussed. The fit is performed separately in the two regions which have respectively 3 (Z+1 b-jet) and 6 bins (Z+2 b-jets). In both cases, the fit is tested in the electron and muon channels, and the compatibility of the results in the combined electron and muon channel is also verified.

Theoretical fit model

The binned maximum likelihood $\mathcal{L}(\mu, \theta)$ used to fit the Monte Carlo to data is defined as:

$$\mathcal{L}(\mu, \theta) = \prod_{i=1}^{N} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \qquad (4.6)$$

where s_i is the predicted number of signal (Z+b and Z+bb) events in bin i and b_i is the predicted Z+jets backgrounds (Z+c and Z+l for the Z+1 b-jet analysis and Z+b, Z+cc, Z+c and Z+l for Z+2 b-jets). μ is the signal strength parameter, i.e. the scale factor resulting from the fit and providing the rescaling of the signal with respect to the Monte Carlo predictions, and n_i is the number of measured events in bin i. The $\mathcal{L}$ function depends on the parameter θ (nuisance parameter), describing the systematic uncertain-

ties on the signal and background events predicted in each bin and resulting from the uncertainties associated to the lepton and b-jet selection. Nuisance parameters associated the systematic sources listed in Section 4.5 are treated as uncorrelated and are distributed according to specific probability distribution functions [35].

The fit is performed by minimising the logarithm of the maximum likelihood estimator: $L(\hat{\mu}, \hat{\theta})$. It represents the unconditional likelihood, characterised by having both $\hat{\mu}$ and $\hat{\theta}$ as maximum likelihood estimators. The output of the fit is the scale factor $\hat{\mu}$ corresponding to its minimum.

Flavour fit in the Z+1 b-jet analysis

In the Z+1 b-jet analysis, the Z+b, Z+c and Z+l flavour components are extracted from the fit. The samples considered for this purpose are Z+jets process and all the other Monte Carlo backgrounds. Z+b, Z+c and Z+l are floating freely in the fit, and the systematics are taken accordingly to the description in Section 4.5. The majority of the reconstruction level systematics take into account both the shape and the normalisation component, and a log-normal constraint of 5% is considered for the non floating backgrounds as the theoretical cross section uncertainties. No limit on the systematic variation is applied.

The fit is performed using SHERPA for the Z+jets simulation. Ancillary measurements are developed with MADGRAPH and ALPGEN, in order to test the closure of the obtained scale factors, by looking at the yield of each fitted process.

Pre-fit and post-fit distributions of the MV2c10 observable are shown in Fig.4.15 for the muon and electron channels using SHERPA. MADGRAPH and ALPGEN plots are shown in Appendix C. The MV2c10 bins are first mapped in same-size bins, coherently with the format needed to perform the likelihood fit detailed in Section 4.6.3. The combined result (bottom raw of Fig.4.15) assumes correlations of all the nuisance parameters considered in the fit: scale factors are extracted from a fit to 6 bins, instead of the 3 bins used in the single-channel fits.

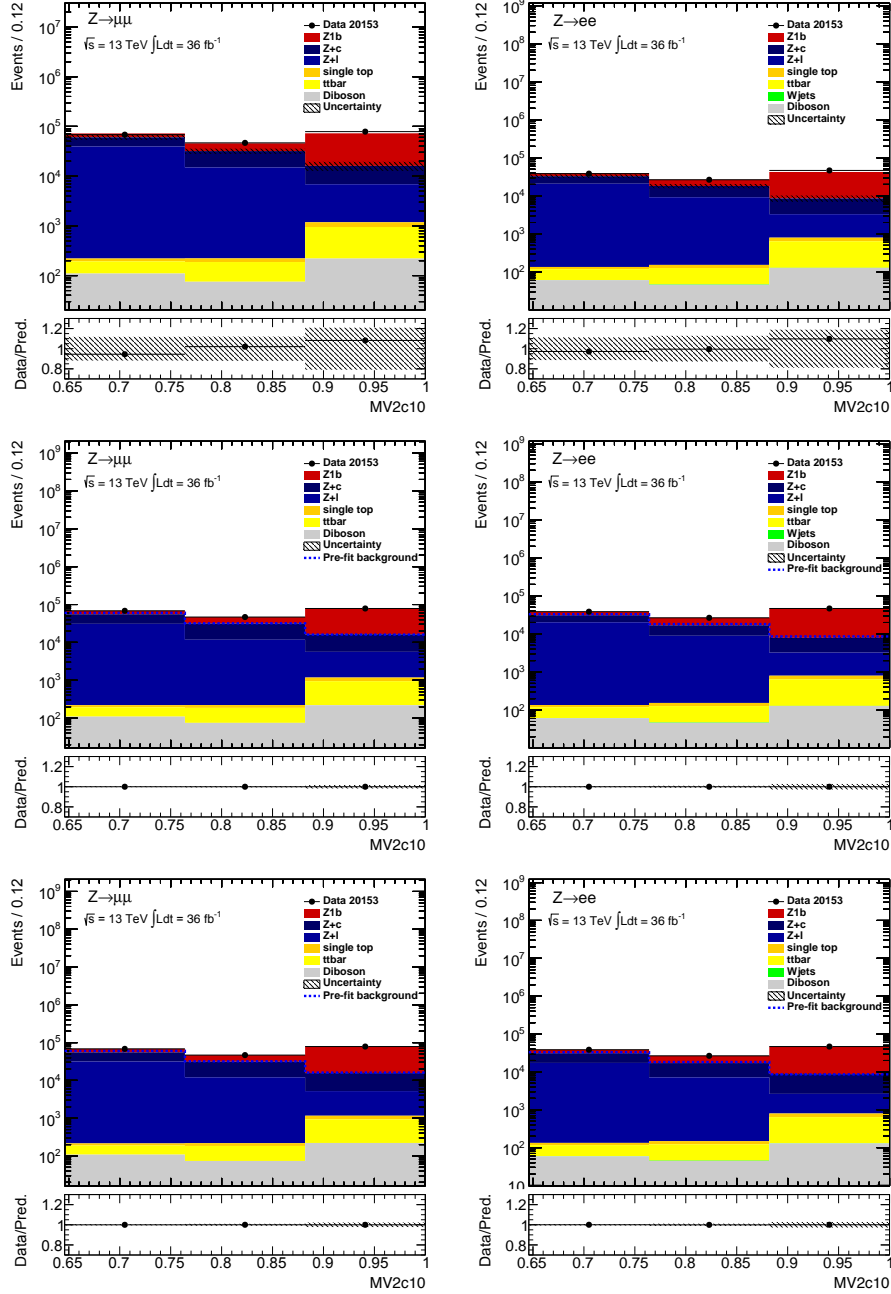Scale factors for the different generators are reported in Table 4.10. De-

Figure 4.15: Pre- (top row) and post-fit (middle and bottom row) plots for the muon (left) and electron (right) channels, using the SHERPA generator as baseline for the Z+jets samples in the Z+1 b-jet analysis. The fits shown in the middle row are performed separately in the electron and muon channels, while the ones shown in the bottom row are performed combining the two channels together.

spite the large spread of values, scale factors are consistent in each generator between electron and muon channels and their combination, within the quoted errors.

		e-channel	μ-channel	Combination
SHERPA	Z+b (μ)	1.14 ± 0.08	1.10 ± 0.09	1.13 ± 0.09
	Z+c	0.9 ± 0.4	1.2 ± 0.3	1.1 ± 0.3
	Z+l	1.0 ± 0.3	0.8 ± 0.2	0.8 ± 0.2
MADGRAPH	Z+b (μ)	0.57 ± 0.05	0.64 ± 0.05	0.55 ± 0.04
	Z+c	1.1 ± 0.2	0.9 ± 0.2	1.03 ± 0.17
	Z+l	0.7 ± 0.2	0.8 ± 0.2	0.69 ± 0.17
ALPGEN	Z+b (μ)	1.7 ± 0.1	1.7 ± 0.2	1.69 ± 0.15
	Z+c	1.2 ± 0.2	1.2 ± 0.2	1.16 ± 0.21
	Z+l	0.7 ± 0.2	0.8 ± 0.2	0.79 ± 0.18

Table 4.10: The scale factors derived for the three different Z+jets generators in the three cases of electron only, muon only, and combined fit in the Z+1 b-jet analysis. μ corresponds to the Z+b signal scale factor and it is compared with the other two flavour components of Z+c and Z+l.

The pre-fit yields are summarised in Table 4.11 for the muon and electron channels. They can be compared with the yield obtained from the fit in Tables 4.12-4.13, where the yields are reported both using μ resulting from single channel and combined channels fit (see Table 4.10). Ideally one would expect both results to agree, being this agreement a proof of consistency between the electron and muon channels. This is the case for both lepton modes. Moreover, one would expect that, no matter which generator one uses, the post-fit yields agree among generators independently of the used fit procedure. This is the case for muon mode in the single lepton fit, while some tensions can be observed in the electron mode already in the single channel fit. This tension propagates to the combined channel fits.

A deeper understanding of the fit results can be obtained by looking at the pulls and the correlations of the nuisance parameters. The pull is defined

Process	Sherpa	Madgraph	Alpgen
		Muon channel	
Diboson		404.77 ± 81.4	
$\bar{t}$		923.05 ± 103.3	
Single top		296.29 ± 26.9	
Z+b	82220.3 ± 4548.5	137558.9 ± 7744.3	49856.4 ± 2955.6
Z+l	59696.6 ± 12503.1	56008.3 ± 11657.8	51980.8 ± 11224.9
Z+c	47134.1 ± 2988.3	64138.6 ± 4444.8	55041.9 ± 3706.9
data		193562	
		Electron channel	
Diboson		235.4 ± 24.8	
Wjets		1.0 ± 0.1	
$t\bar{t}$		653.3 ± 73.1	
Single top		199.1 ± 16.5	
Z+b	48998.7 ± 2619.2	85202.2 ± 4218.7	30482.2 ± 1690.2
Z+l	32353.5 ± 6819.0	32269.9 ± 6807.0	28714.1 ± 6131.7
Z+c	26418.0 ± 1749.4	37421.8 ± 2230.0	32366.6 ± 2161.4
data		111775	

Table 4.11: Pre-fit yields for the different processes for the muon and electron channels for the three possible choices of Z+jets generators in the Z+1 b-jet analysis.

Process	SHERPA	MADGRAPH	ALPGEN
Muon channel fit only			
Diboson	404.8 ± 81.5	404.3 ± 82.0	405.3 ± 81.8
W+jets		3.7 ± 1.2	13.0 ± 1.8
$t\bar{t}$	923.0 ± 106.0	923.0 ± 109.0	924.6 ± 105.7
Single top	296.3 ± 28.2	296.2 ± 28.9	296.6 ± 27.8
Z+b	90463.7 ± 5265.9	87880.4 ± 5613.0	85019.5 ± 5391.1
Z+l	46670.5 ± 10273.2	47224.5 ± 10519.7	42806.8 ± 10212.8
Z+c	54800.0 ± 12987.0	56825.8 ± 13995.4	64097.4 ± 12945.1
Combined muon and electron channel fit			
Diboson	402.5 ± 81.8	513.1 ± 100.7	429.1 ± 88.7
W+jets		3.9 ± 1.1	13.4 ± 1.8
$\bar{t}$	900.1 ± 110.9	1079.2 ± 89.6	929.1 ± 124.6
Single top	290.0 ± 28.8	338.1 ± 25.3	300.1 ± 31.8
Z+b	91690.6 ± 5065.2	79658.1 ± 5432.9	84717.8 ± 5126.8
Z+l	47249.3 ± 9151.3	41324.5 ± 8862.3	42099.9 ± 8877.8
Z+c	53050.1 ± 11997.6	70579.3 ± 12057.5	65012.9 ± 11421.9
data	193562		

Table 4.12: Post-fit yields for the different processes for the muon channel for the three possible choices of Z+jets generators in the Z+1 b-jet analysis, in the case of the combined electron and muon fit, and in the standalone muon fit cases.

Process	SHERPA	MADGRAPH	ALPGEN
	Electron channel fit only		
Diboson	235.4 ± 26.4	235.6 ± 29.1	235.5 ± 26.1
W+jets	1.00 ± 0.08	2.4 ± 0.6	16.8 ± 3.1
$t\bar{t}$	653.2 ± 80.6	653.5 ± 87.0	653.5 ± 77.2
Single top	199.104 ± 19.3	199.1 ± 18.7	199.2 ± 17.6
Z+b	55981.6 ± 2896.2	48611.7 ± 3381.3	51114.2 ± 2954.4
Z+l	31390.3 ± 8962.3	22872.5 ± 6007.3	21153.6 ± 4714.8
Z+c	23316.5 ± 10850.8	39203.1 ± 8191.0	38408.0 ± 6306.4
	Combined electron and muon channel fit		
Diboson	232.3 ± 26.8	253.3 ± 22.4	225.5 ± 28.5
W+jets	0.98 ± 0.08	3.9 ± 1.00	16.7 ± 2.9
$t\bar{t}$	645.4 ± 82.6	734.8 ± 72.6	641.4 ± 90.6
Single top	200.0 ± 19.8	200.9 ± 16.1	198.2 ± 18.0
Z+b	54827.0 ± 2923.7	48521.9 ± 3118.8	51212.8 ± 2977.0
Z+l	25981.3 ± 5126.4	22745.0 ± 4937.5	22509.6 ± 4762.3
Z+c	29876.8 ± 6877.7	39393.8 ± 6909.8	37034.7 ± 6365.8
data	111775		

Table 4.13: Post-fit yields for the different processes for the electron channel for the three possible choices of Z+jets generators in the Z+1 b-jet analysis, in the case of the combined electron and muon fit, and in the standalone electron fit cases.

as:

$$P = \frac{\theta_i - \hat{\theta}}{\sigma_{\theta_i}} \quad , \tag{4.7}$$

where θ_i represents the ith nuisance parameter value pre-fit and $\hat{\theta}$ its best value after performing the fit. The difference is expressed in σ_{θ_i} unity, which represents the uncertainty on the post-fit nuisance parameter θ. If no deviation is present, post-fit nuisance parameter is expected to have the same value as the pre-fit one and the pull to be centred to zero with an associated uncertainty of $\pm 1\sigma$. If the uncertainty on the nuisance parameter is smaller, the systematic is over constrained by the fit.

As already anticipated, in the case of the Z+1 b-jet analysis, the discriminating variable used to extract the normalisation of Z+b, Z+c and Z+l processes is the b-tagging weight (MV2c10), which is distributed in three bins. This means that the single channel (electron or muon) fit procedure in this case corresponds to a system of three equations with three unknowns, with zero degrees of freedom. As a consequence, one expects no pulls of the nuisance parameters in the single lepton fit. This is the case, looking at the pull plots for the electron and muon channels presented in Fig.4.16 for SHERPA samples. Analogous results are obtained for MADGRAPH and ALPGEN, which are shown in Appendix C.

For the combined channel instead, the fit is performed in six bins (three for each lepton channel) and the unknowns are still the three normalisation of Z+b, Z+c and Z+l, leaving three degrees of freedom. The combined fit pulls for SHERPA are shown in the top plot of Fig.4.16. No strong pull on the systematics is observed in the combination between the two channels, reflecting a satisfactory stability in the fit. In general, the behaviour is different among the three generators, reflecting the large variety of scale factors (see Table 4.10). The first two points in Fig.4.16, refer to the values of the Z+c and Z+l scale factors, which have no pull since they are the parameters extracted from the fit itself.

The correlations are shown in Fig.4.17 only for the nuisance parameters which have at least a correlation larger 50% using SHERPA generator. SHERPA presents some correlation or anticorrelation between the estimated normalisation scale factors (Z+c and Z+l) and some other systematics, that
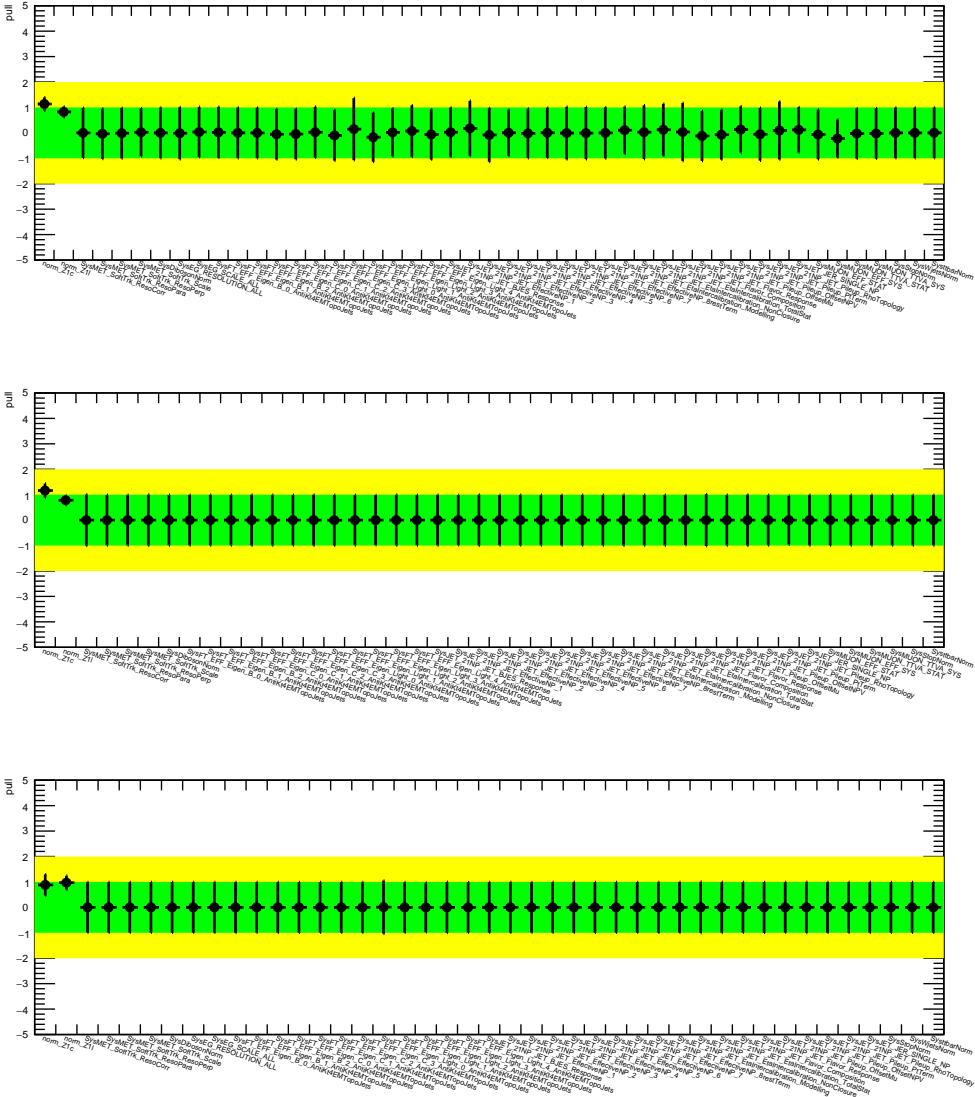
Figure 4.16: Pulls of all the nuisance parameters included in the fit using the SHERPA generator (the MC statistic is not shown) in the Z+1 b-jet analysis. Pulls are shown for the combined lepton channel fit (top), muon-only fit (middle) and electron-only fit (bottom).

can vary if looking at the single channel or combined fit. In particular,

Z+b tends to be anticorrelated with Z+c, correlated to the flavour tagging uncertainty (B0 and C0 components, see Section 4.5), and the jet flavour composition parameter;

Z+c tends to be anticorrelated with Z+b, Z+l, flavour tagging uncertainty (C3 and L2 components), and the MC statistical uncertainty in the second bin of the MV2c10 distirbution;

Z+l tends to be anticorrelated with Z+c, flavour tagging uncertainty (L3), the MC statistical uncertainty in the first bin of the MV2c10 distribution, and correlated with the MC statistical uncertainty of the second bin of the distribution, and Z+b.

The correlations between the different Z+jets process normalisations are not surprising, neither correlations with the flavour tagging uncertainties. The correlations with the MC statistical uncertainty are present in the single channel fits only and need some further investigations. The combined fit presents much fewer correlations.
The correlation of nuisance parameters for fits using MADGRAPH and ALPGEN are summarised in Appendix C. In general, the three generators agree in the nuisance parameters correlations.

Flavour fit in the Z+2 b-jets region

In the Z+2 b-jets analysis, the normalisation of Z+bb, Z+b, Z+c, Z+cc and Z+l samples should in principle be extracted independently from the fit. The statistics in this region is smaller compared to the Z+1 b-jet selection, while SHERPA, MADGRAPH and ALPGEN differences in the modelling of the flavour categories are larger. For these reasons, the Z+b, Z+c, Z+cc and Z+l backgrounds are taken as one single component in the fit, further indicated as Z+jj, and is treated as free parameter in the fit. The Z+bb component is completely free to float, since it represent the signal in the Z+2 b-jets region. Most of the reconstruction level systematics have both a shape and a normalisation component, and a log-normal constraint of 5% is considered for each of the non-floating backgrounds. Since top process has a small yield
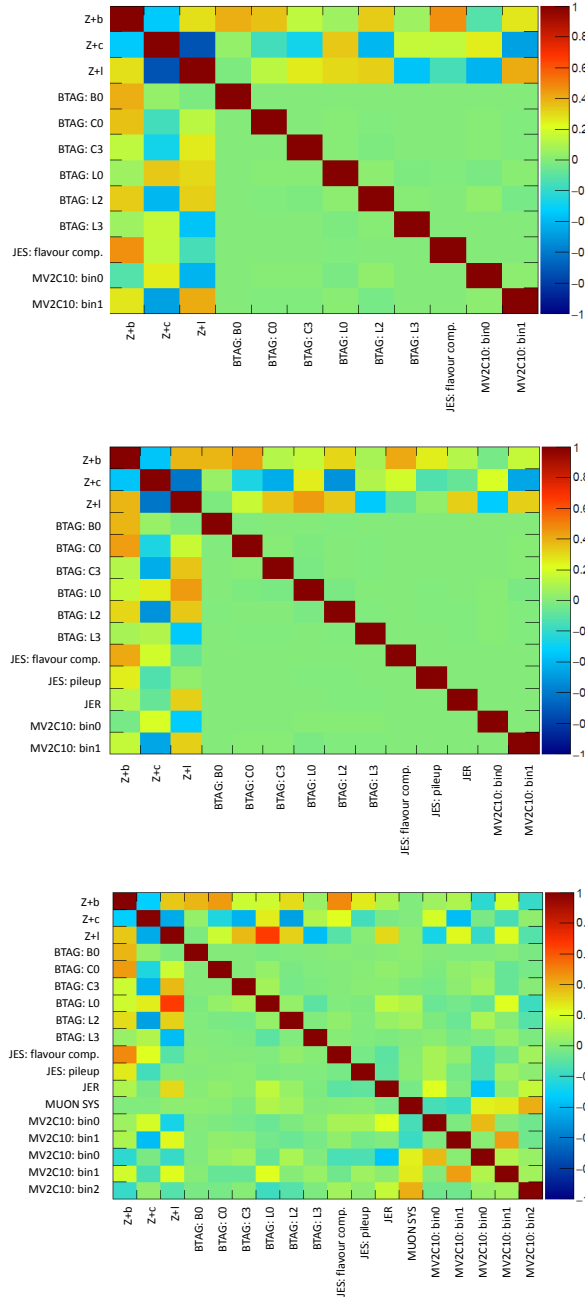
Figure 4.17: Correlations of the nuisance parameters included in the fit using the SHERPA generator. Only parameters with correlation larger than 50% are shown for electron-only (top), muon-only (middle) and combined (bottom) fits.

and its shape is similar to the $t\bar{t}$ one, the two samples are summed together and considered in the fit as a single component, called "top backgrounds". No limit on the systematic variations is applied.

Pre-fit and post-fit plots in the muon and electron channels are shown in Fig.4.18 using the SHERPA generator. The distribution are mapped in same size bins, coherently with the format needed to perform the likelihood fit. The combined electron and muon fit (bottom of Fig.4.18) assumes correlations of all nuisance parameters considered, therefore the background normalisations are extracted from a fit to 12 bins, instead of the 6 bins used for single lepton channel. The pre-and post-fit distributions obtained by using MADGRAPH or ALPGEN for the Z+jets simulation are reported in Appendix C.

The scale factors obtained from the different fits are reported in Table 4.14. For each generator, the normalisation factor of the Z+bb signal are compatible within errors across all the single lepton channels and their combination, even if MADGRAPH shows larger discrepancies between electron and muon channels. The scale factors for Z+jj background present a very good agreement across the three different Monte Carlo calculations.

		e-channel	μ-channel	Combination
SHERPA	Z+bb (μ)	1.28 ± 0.12	1.38 ± 0.13	1.3 ± 0.1
	Z+jj	0.8 ± 0.2	1.1 ± 0.2	0.98 ± 0.19
MADGRAPH	Z+bb (μ)	0.56 ± 0.05	0.70 ± 0.05	0.63 ± 0.05
	Z+jj	1.06 ± 0.19	0.91 ± 0.17	1.04 ± 0.16
ALPGEN	Z+bb (μ)	1.27 ± 0.11	1.36 ± 0.14	1.33 ± 0.10
	Z+jj	0.77 ± 0.16	0.81 ± 0.13	0.78 ± 0.12

Table 4.14: The scale factors derived for the three different Z+jets generators in electron only, muon only and combined fits in the Z+2 b-jets analysis. The Z+bb signal scale factor is compared with the Z+jj background, corresponding to the sum of Z+b, Z+c, Z+cc and Z+l contributions.

The pre-fit yields for the muon and electron channels are reported in Table 4.15. They can be compared with the yields obtained from the fit, which are summarised in Tables 4.16-4.17. Reasonable compatibility is observed in the post-fit results among the three generators, even if some differences persist
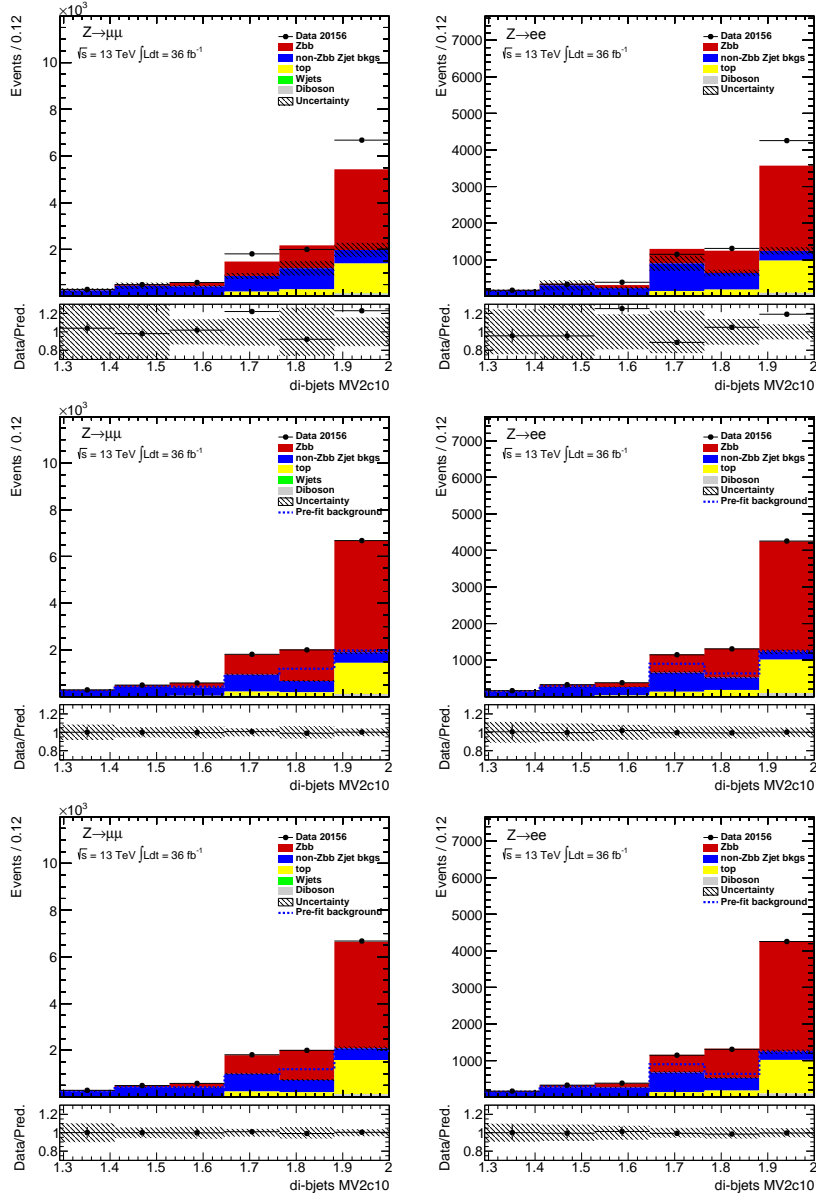
Figure 4.18: Pre- (top row) and post-fit (middle and bottom row) plots for the muon (left) and electron (right) channels, using the SHERPA generator as baseline for the Z+jets samples in the Z+2 b-jets anlysis. The fits shown in the middle row are performed separately in the electron and muon channels, while the ones shown in the bottom row are performed combining the two channels together.

Process	SHERPA	MADGRAPH	ALPGEN
Muon channel			
Diboson		184.2 ± 22.5	
Top		1763.3 ± 170.4	
Z+bb	5606.0 ± 401.4	10527.1 ± 765.3	5293.5 ± 408.4
Z+jj	3116.9 ± 621.9	2970.5 ± 360.3	3208.9 ± 837.12
data		11863	
Electron channel			
Diboson		125.6 ± 13.3	
Top		1208.4 ± 104.9	
Z+bb	3733.1 ± 294.7	7343.0 ± 543.9	3486.6 ± 249.5
Z+jj	1908.0 ± 320.4	1887.8 ± 281.8	2107.0 ± 473.0
data		7580	

Table 4.15: Pre-fit yields for the different processes for the muon and electron channels for the SHERPA, MADGRAPH and ALPGEN in the Z+2 b-jets analysis.

and can be explained by the different models used for the description of the heavy flavour component: ALPGEN is a LO generator using the 4FNS, while MADGRAPH and SHERPA employ the 5FNS being the first LO and the second NLO generators.

The pull plots for the muon, electron and combined fit using SHERPA are shown in Fig.4.19, while MADGRAPH and ALPGEN plots are in Appendix C. In the Z+2 b-jets analysis, the fit is performed using the MV2c10 variable for the di-bjets system, distributed in six bins to extract the normalisation of Z+bb and Z+jj samples. Therefore, compared to the Z+1 b-jet region, the fit has four degrees of freedom. Looking at Fig.4.19, no strong pull are observed, since all nuisance parameters differ from their pre-fit values less than 1σ, which is an indication of the stable behaviour of the fit. In the combined fit, SHERPA presents the most stable pulls, compared to MADGRAPH and ALPGEN, in particular in the b-tagging calibration parameters.

Correlations of the nuisance parameters in single lepton and combined fit are shown in Fig.4.20 using SHERPA. As in the case of the Z+1 b-jet, correlations are displayed only for the nuisance parameters presenting at least

Process	SHERPA	MADGRAPH	ALPGEN
Combined electron and muon channel fit			
Diboson	196.3 ± 19.2	201.3 ± 19.6	174.5 ± 18.4
Top	1901.8 ± 162.9	1968.7 ± 159.1	1746.6 ± 154.1
Z+bb	6935.2 ± 272.0	6702.2 ± 268.6	7374.5 ± 256.6
Zjj	2792.0 ± 252.8	2921.7 ± 187.0	2541.8 ± 169.2
Muon channel only fit			
Diboson	178.6 ± 20.5	181.4 ± 19.1	171.5 ± 20.7
Top	1729.5 ± 174.7	1765.8 ± 159.6	1704.2 ± 172.3
Z+bb	7304.9 ± 436.3	7218.8 ± 311.4	7509.7 ± 364.5
Zjj	2647.3 ± 367.0	2692.9 ± 223.2	2476.1 ± 220.5
data	11863		

Table 4.16: Post-fit yields for the different processes for the combined and muon channel fits, for the three Z+jets generators, in the Z+2 b-jets mode.

Process	SHERPA	MADGRAPH	ALPGEN
Combined electron and muon channel fit			
Diboson	133.3 ± 14.2	128.2 ± 12.8	124.6 ± 11.1
Top	1242.7 ± 97.0	1178.5 ± 89.9	1160.6 ± 91.3
Z+bb	4506.6 ± 166.6	4593.6 ± 161.1	4831.6 ± 147.9
Zjj	1731.5 ± 109.4	1742.1 ± 116.3	1486.2 ± 87.9
Electron channel only fit			
Diboson	131.3 ± 15.9	138.1 ± 15.8	128.2 ± 13.8
Top	1247.5 ± 105.1	1269.2 ± 111.9	1228.0 ± 111.0
Z+bb	4510.1 ± 186.9	4201.1 ± 235.4	4738.5 ± 163.8
Zjj	1690.3 ± 142.6	1972.2 ± 182.2	1484.3 ± 115.0
data	7580		

Table 4.17: Post-fit yields for the different processes for the combined and electron channel fits, for the three Z+jets generators, in the Z+2 b-jets mode.
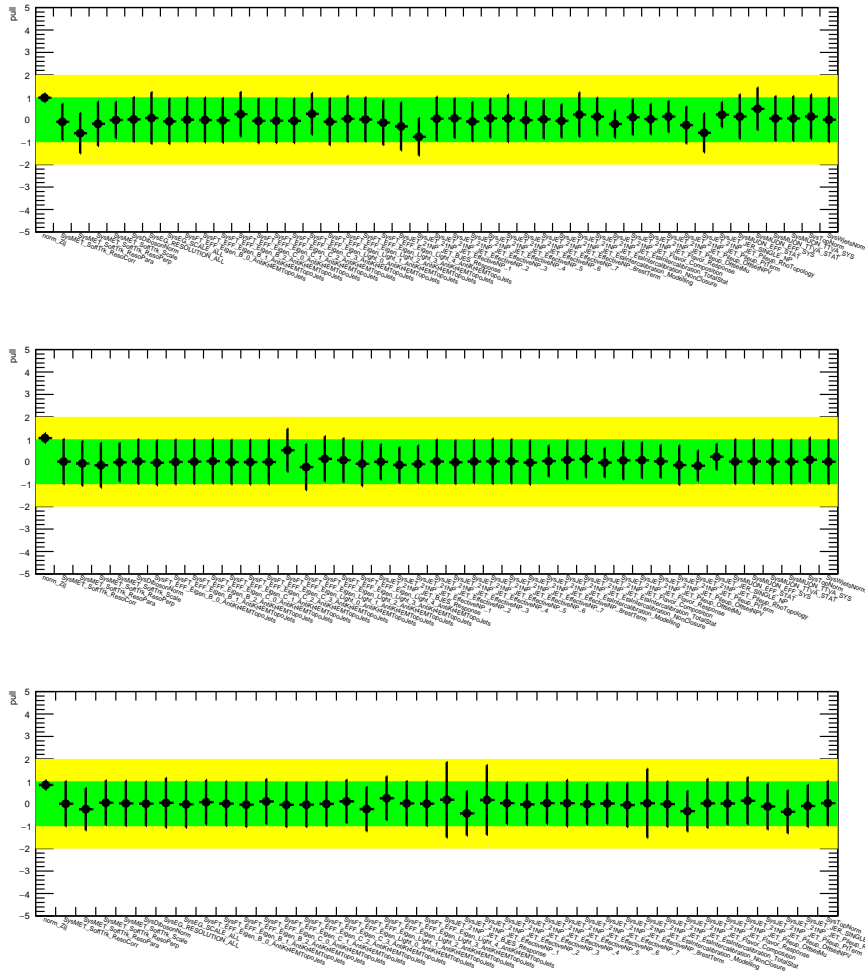
Figure 4.19: Pulls of all the nuisance parameters included in the fit (MC statistics excluded) with SHERPA as preferred signal generator in the Z+2 b-jets analysis. The top plot refers to the combined muon and electron fit, the middle plot to the muon only fit, and the bottom plot to the electron only fit.

a correlation larger than 50%. In both single lepton and combined fits, the
Z+bb signal has strong correlation ($\sim$70%) with the B0 component of the
b-jets calibration. The Z+jj background normalisation presents a evident
correlation with the L0 parameter of the systematics associated to the b-
tagging uncertainty. The correlations of Z+jets processes with the flavour
composition of these samples is not surprising. The nuisance parameter cor-
relations for fit results performing using MADGRAPH and ALPGEN generators
as baseline are reported in Appendix C.

4.7   Data and MC comparison

In this section, the detector level kinematic plots for some of the main
variables used in the analysis are shown separately in the electron and muon
channel. Only events passing the selection defined in Section 4.4 are repre-
sented in the plots and are divided in three different categories:

- events containing the Z boson produced in association with jets (Z+jets)
  are shown in Section 4.7.1. In the following, this region is referred to
  as "inclusive", since the number of b-jets are not counted yet.

- plots for events corresponding to the Z+1 b-jet and Z+2 b-jets analyses
  are reported in Section 4.7.2.

For each of the following distributions, the level of agreement between Data
and Monte Carlo samples is quoted. In each distribution the signal Monte
Carlo and the $Z \to \tau\tau$+jets and $W \to l\nu$+jets are modelled with SHERPA
(green) generator, while in the ratio of data with respect to predictions
(Data/Pred.) also MADGRAPH (red) and ALPGEN (blue) simulations are
shown. In the Data/MADGRAPH and Data/ALPGEN ratios, all Z+jets and
W+jets signal and background processes are substituted to SHERPA ones.
The uncertainty on the data (shaded band) is shown only for SHERPA and
corresponds to the sum in quadrature of al the contributions described in
Section 4.5. The luminosity uncertainty is not considered in the total uncer-
tainty and therefore it is not displayed in the following plots.
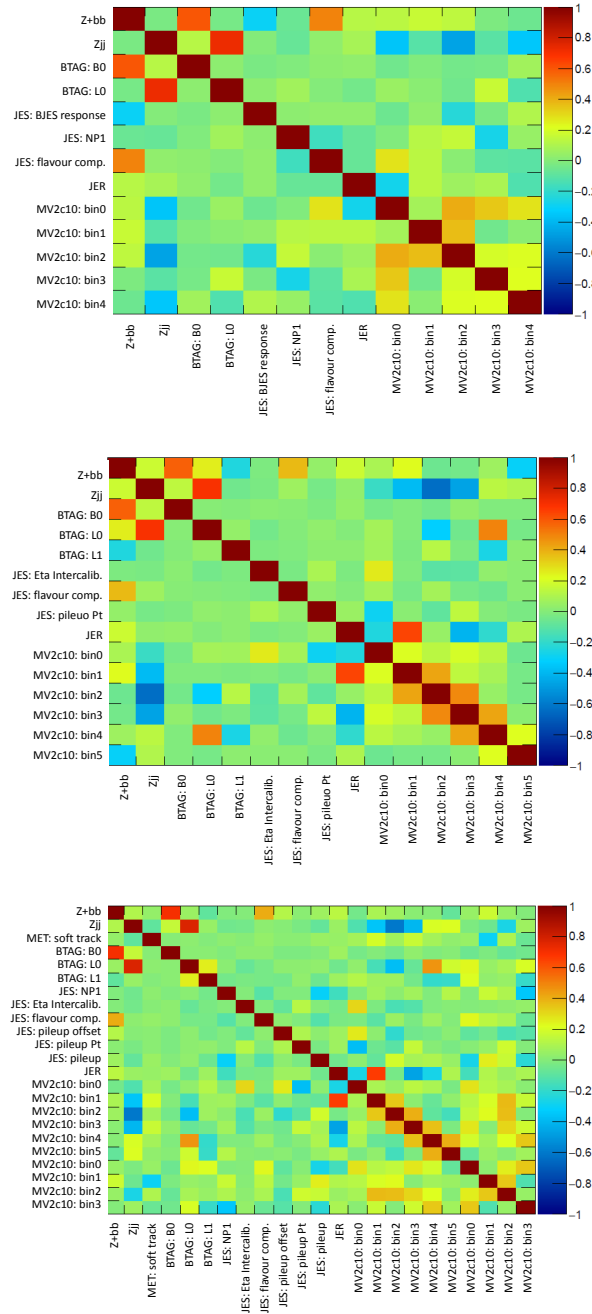
Figure 4.20: Correlations of the nuisance parameters in the Z+2 b-jets analysis. SHERPA is the signal generator used in the electron only (top), muon only (middle) and combined (bottom) fits. Only parametrs with a correlation with an absolute value larger than 50% are shown.

4.7.1   Inclusive jet plots

In the inclusive region, the Z boson is required to have at least one jet, without applying the b-tagging algorithm. Although this region does not correspond to the Z+1 b-jet and Z+2 b-jets analysis, it is of extreme interest for the understanding of the kinematic of the jets and the Monte Carlo modelling. Fig.4.21 shows the distribution of the p_T and y of the leading jet. While the y is well described by all Monte Carlo generators, some tensions appear for p_T >400 GeV.
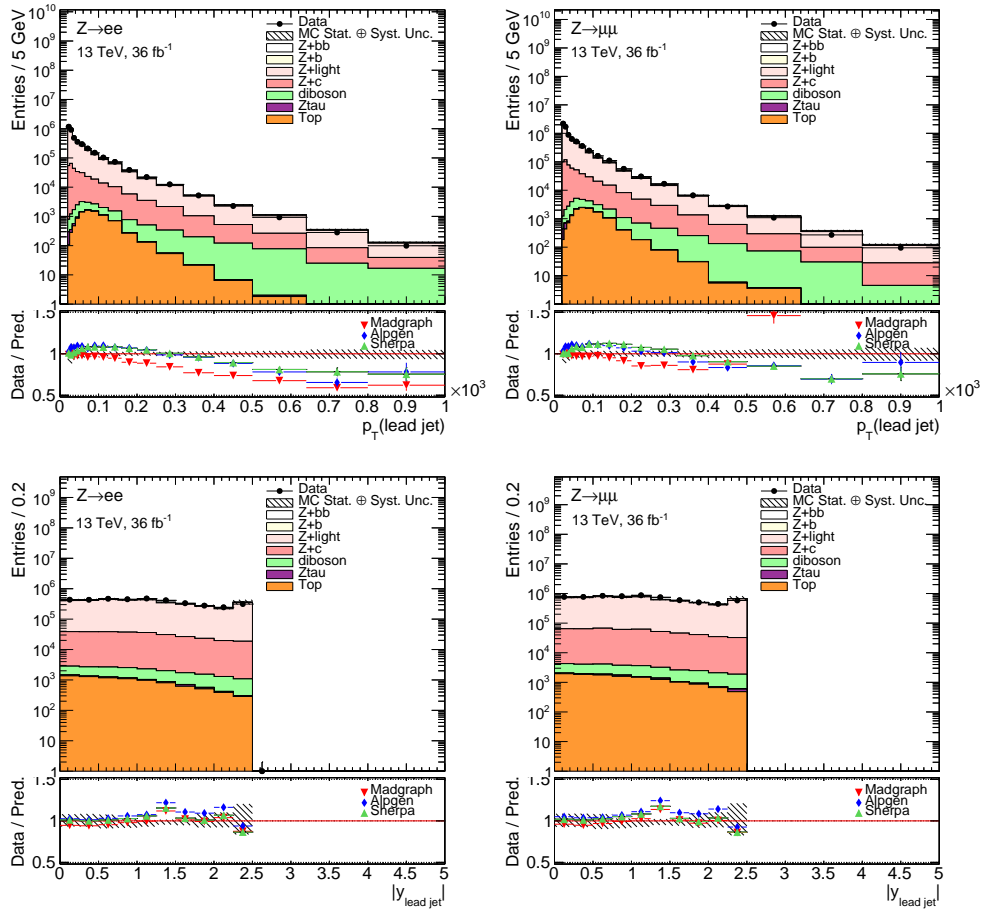
Figure 4.21: Data and Monte Carlo comparison for the leading jet p_T (top) and y (bottom). SHERPA generator is used to model the Z+jets processes and the ratio Data/Pred. is shown for SHERPA (green), MADGRAPH (red) and ALPGEN (blu). The systematic band is evaluated using SHERPA and does not include the luminosity uncertainty.

Fig.4.22 shows the distributions of the inclusive jet multiplicity for events with at least one jet. SHERPA shows the best agreement with data, while MADGRAPH and ALPGEN show evident discrepancies for events with at least 4 jets, where the parton emission is described by Parton Shower.
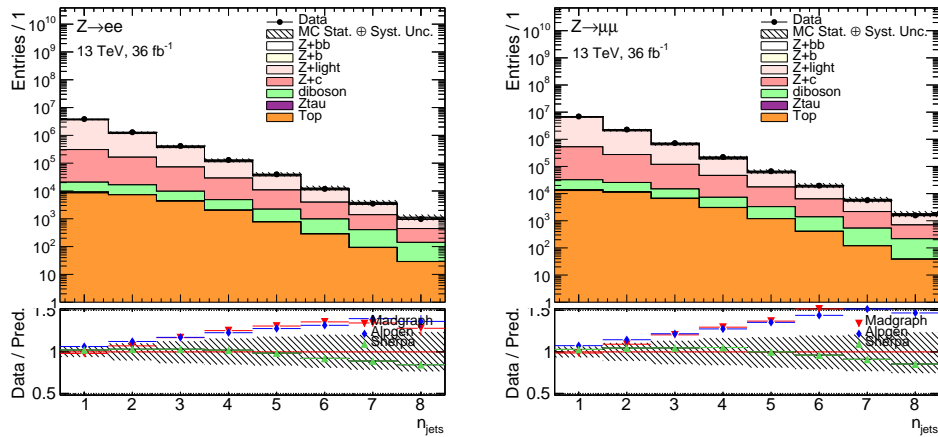
Figure 4.22: Data and Monte Carlo comparison for the inclusive jet multiplicity. SHERPA generator is used to model the Z+jets processes (top) and the ratio Data/Pred. is shown for SHERPA (green), MADGRAPH (red) and ALPGEN (blu) (bottom). The systematic band is evaluated using SHERPA and does not include luminosity uncertainty.

In Fig.4.23 the distribution of events as a function of the exclusive multiplicity of the b-jets is shown. The content of the first (second) bin of this variable is the number of events passing the Z+1 b-jet (Z+2 b-jets) analysis of Table 4.6. This means jets are required to pass the b-tagging algorithm and simulated jets are required to fulfil the truth flavour labelling. Consequently each of the Z+bb, Z+b, Z+c, Z+cc and Z+l yield is normalised for the corresponding scale factor obtained separately in the Z+1 b-jet and Z+2 b-jets regions. The scale factor extracted from the combined fit is used. As expected, after correcting for the overall normalisation with the flavour fit procedure, the description of the b-jet multiplicity provided by SHERPA, MADGRAPH and ALPGEN is very good in both lepton channels. The uncertainty does not include the error associated to the various scale factors.
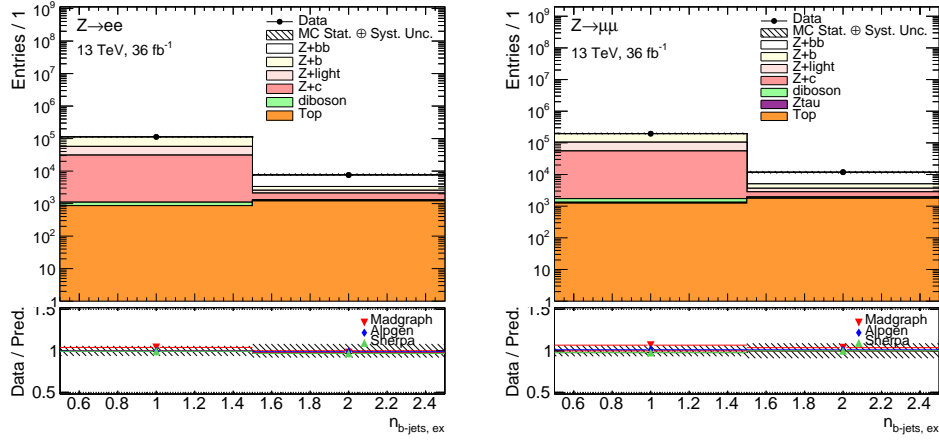
Figure 4.23: Data and Monte Carlo comparison for the exclusive b-jet multiplicity. The first and second bins correspond to the events passing the Z+1 b-jet and Z+2 b-jets selections, respectively. The yields of the $Z(\to ee/\mu\mu)$+jets processes are scaled according to the combined fit results in the two regions. SHERPA generator is used to model the Z+jets processes (top) and the ratio Data/Pred. is shown for SHERPA (green), MADGRAPH (red) and ALPGEN (blu) (bottom). The systematic band is evaluated using SHERPA and does not include luminosity uncertainty.

4.7.2   Z+1,2 b-jets plots

In the Z+1 b-jets, the yields of the Z+b signal events and Z+c and Z+l processes are scaled by the scale factors obtained in the lepton combined fit. The uncertainty associated to the scale factors is not incorporated in the systematic band, since it is evaluated starting from reconstruction level systematics.

Fig.4.24 shows the distribution of the Z boson rapidity (y). The agreement between data and predictions is very good for all generators. The same feature can be observed in the distributions of the leading b-jet p_T and y in Fig.4.25 for SHERPA, while some tensions is visible in MADGRAPH and ALPGEN. Fig.4.26 presents the Z-to-leading b-jet related observables. All distributions are well modelled by SHERPA, while MADGRAPH and ALPGEN are less satisfactory, in particular in the $\Delta\Phi(Z,b)$ distribution. In general, the agreement between data and Monte Carlo is similar in electron and muon channels.

In the Z+2 b-jets analysis, the yields of the Z+bb signal events and Z+b,
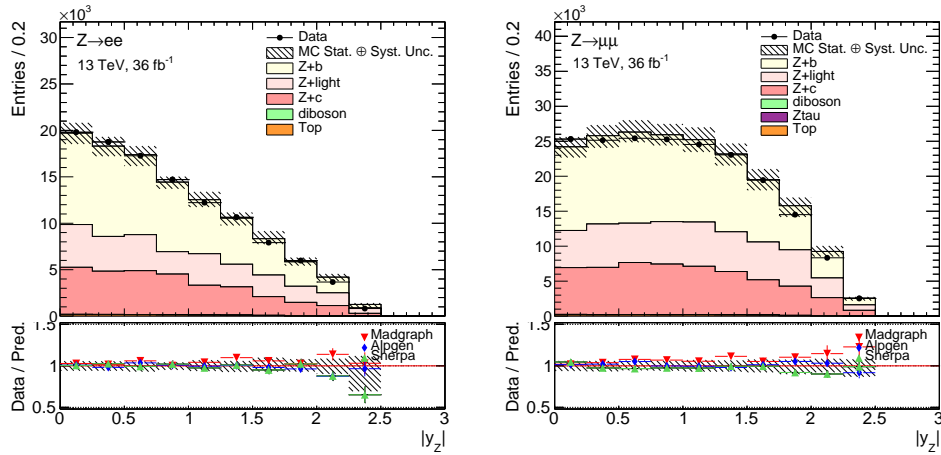
Figure 4.24: Data and Monte Carlo comparison for the Z y in the Z+1 b-jet analysis. The normalisation of Z+b, Z+c and Z+l is given by the combined fit result. SHERPA generator is used to model the Z+jets processes (top) and the ratio Data/Pred. is shown for SHERPA (green), MADGRAPH (red) and ALPGEN (blu) (bottom). The systematic band is evaluated using SHERPA does not involve luminosity uncertainty.

Z+cc, Z+c and Z+l background processes are scaled by the scale factors obtained in the lepton combined fit.

Fig.4.27 shows the distributions of the sub-leading b-jet p_T and y, presenting a good agreement between data and all Monte Carlo generators, especially in the rapidity. The same feature is visible in the di-bjets observables M_{bb} and ΔR_{bb} in Fig.4.28 and Δy_{bb} $\Delta \Phi_{bb}$ in Fig.4.29. As for Z+1 b-jet, the comparison in the electron and muon channels shows a good consistency.
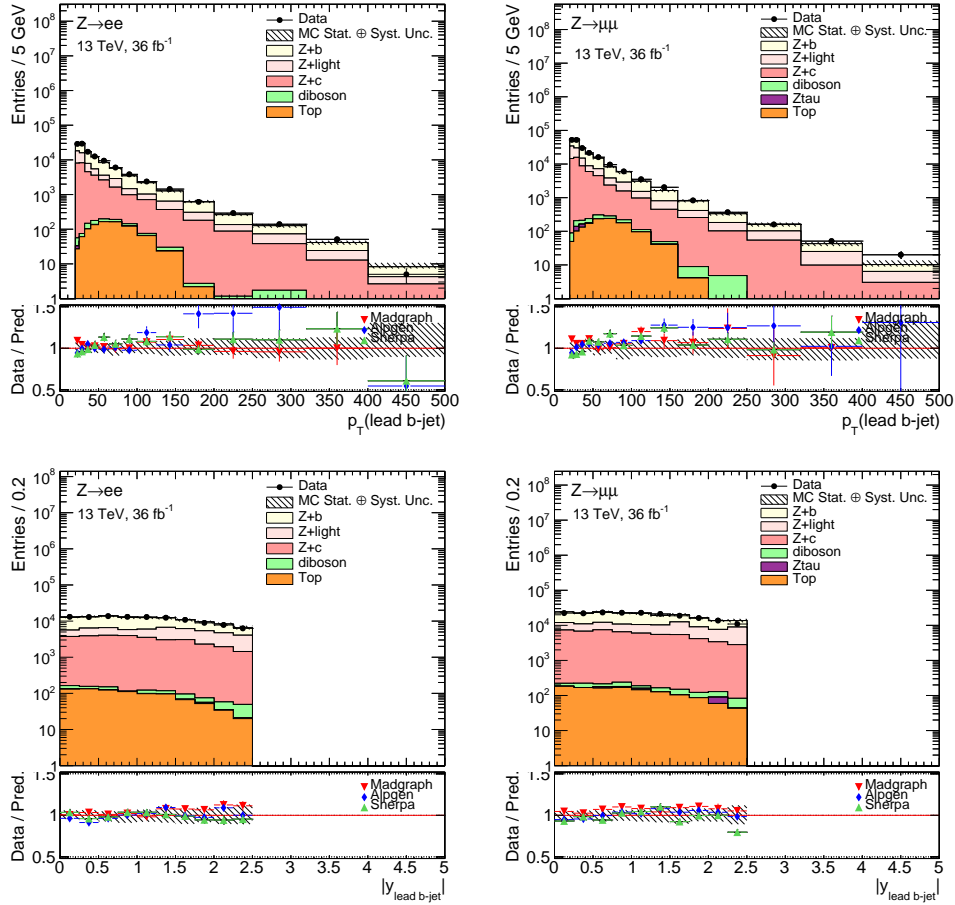
Figure 4.25: Data and Monte Carlo comparison for the leading b-jet p_T (top) and y (bottom). The normalisation of Z+b, Z+c and Z+l is given by the combined fit result. SHERPA generator is used to model the Z+jets processes (top) and the ratio Data/Pred. is shown for SHERPA (green), MADGRAPH (red) and ALPGEN (blu) (bottom). The systematic band is evaluated using SHERPA and does not include luminosity uncertainty.
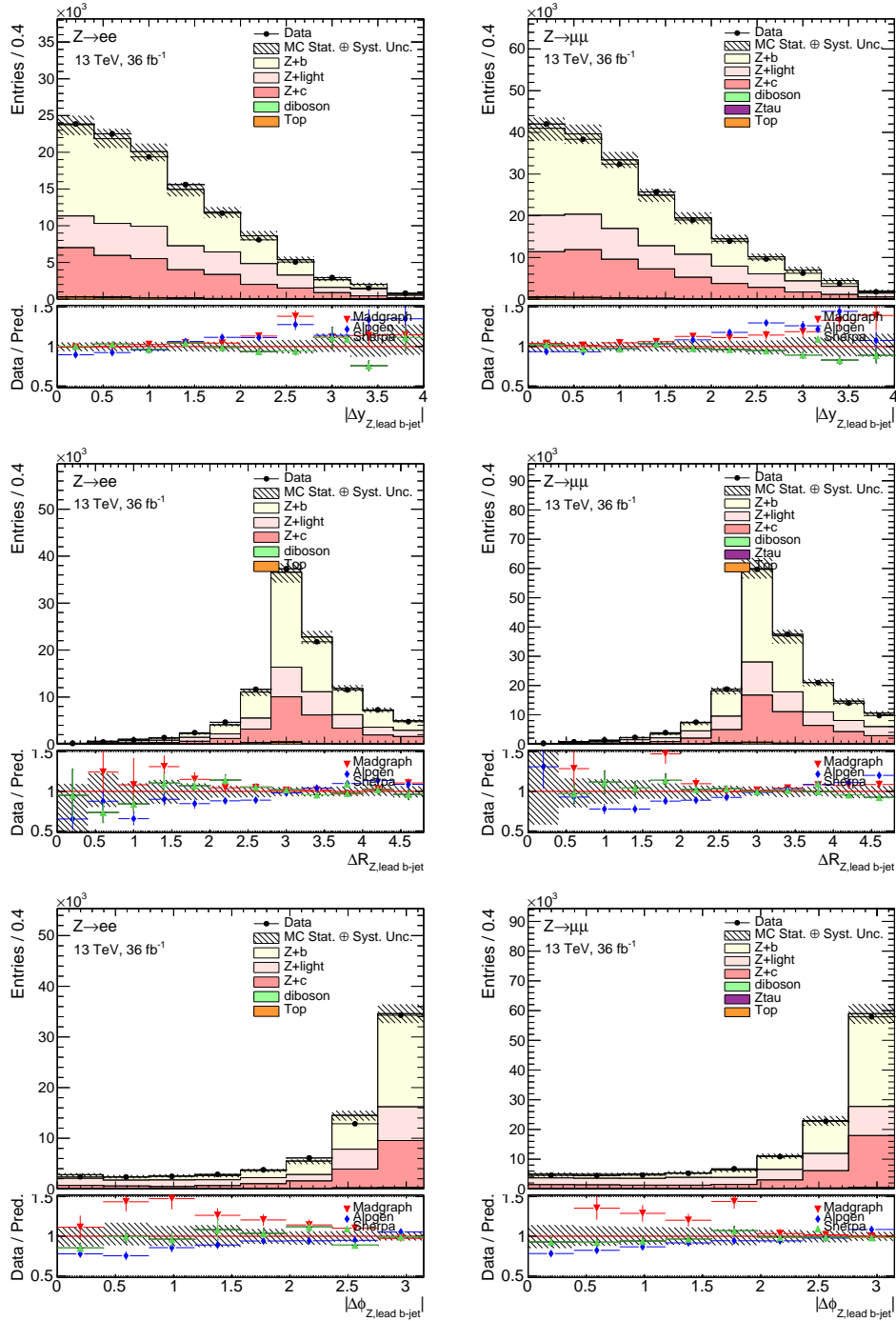
Figure 4.26: Data and Monte Carlo comparison for the Z-leading b-jet related quantities: Δy (top), ΔR (middle) and $\Delta\Phi$ (bottom). The normalisation of Z+b, Z+c and Z+l is given by the combined fit result. SHERPA generator is used to model the Z+jets processes (top) and the ratio Data/Pred. is shown for SHERPA (green), MADGRAPH (red) and ALPGEN (blu) (bottom). The systematic band is evaluated using SHERPA and does not include luminosity uncertainty.
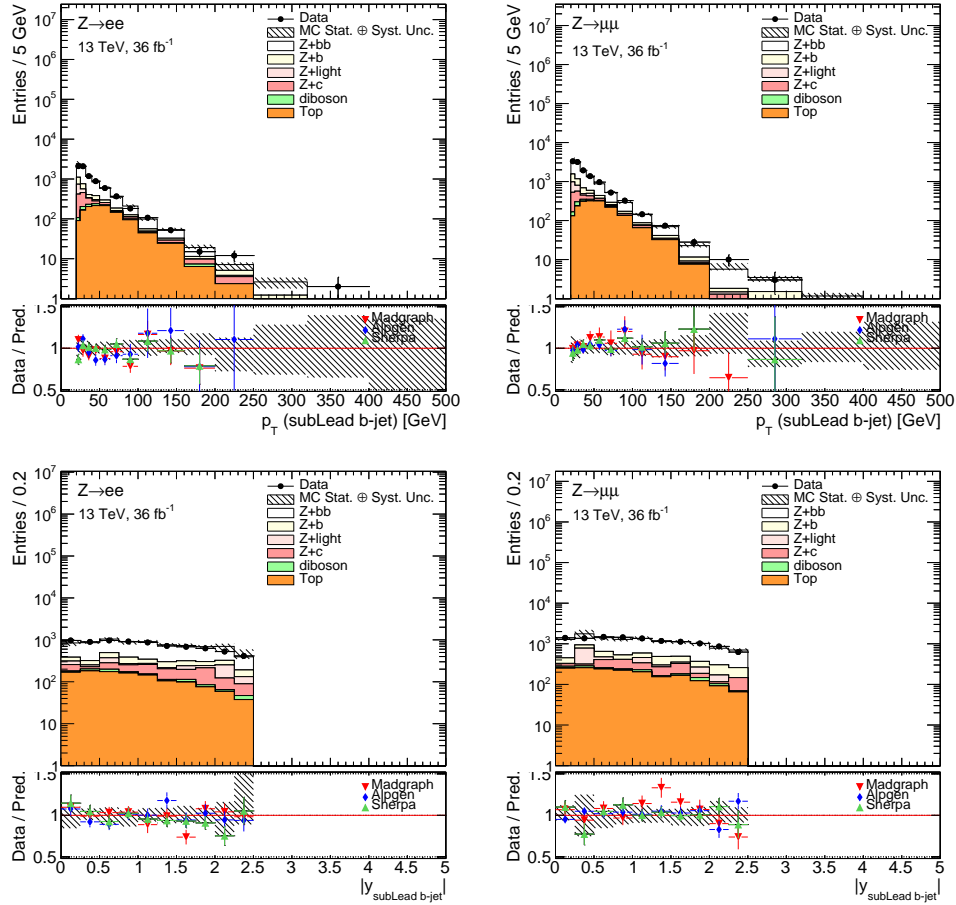
Figure 4.27: Data and Monte Carlo comparison for the second leading b-jet p_T (top) and y (bottom). The normalisation of Z+bb, Z+b, Z+cc, Z+c and Z+l is given by the combined fit result. SHERPA generator is used to model the Z+jets processes (top) and the ratio Data/Pred. is shown for SHERPA (green), MADGRAPH (red) and ALPGEN (blu) (bottom). The systematic band is evaluated using SHERPA and does not include luminosity uncertainty.
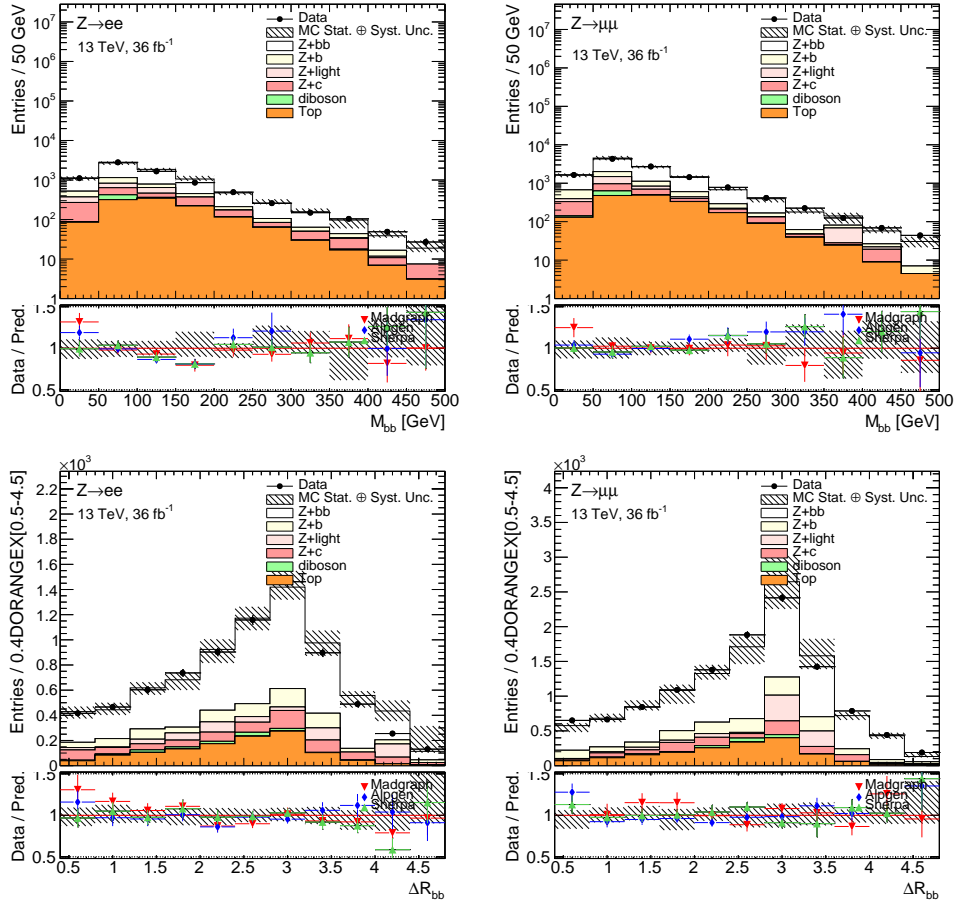
Figure 4.28: Data and Monte Carlo comparison for the di-bjets M_{bb} (top) and ΔR_{bb} (bottom). The normalisation of Z+bb, Z+b, Z+cc, Z+c and Z+l is given by the combined fit result. SHERPA generator is used to model the Z+jets processes (top) and the ratio Data/Pred. is shown for SHERPA (green), MADGRAPH (red) and ALPGEN (blu) (bottom). The systematic band is evaluated using SHERPA and does not include luminosity uncertainty.
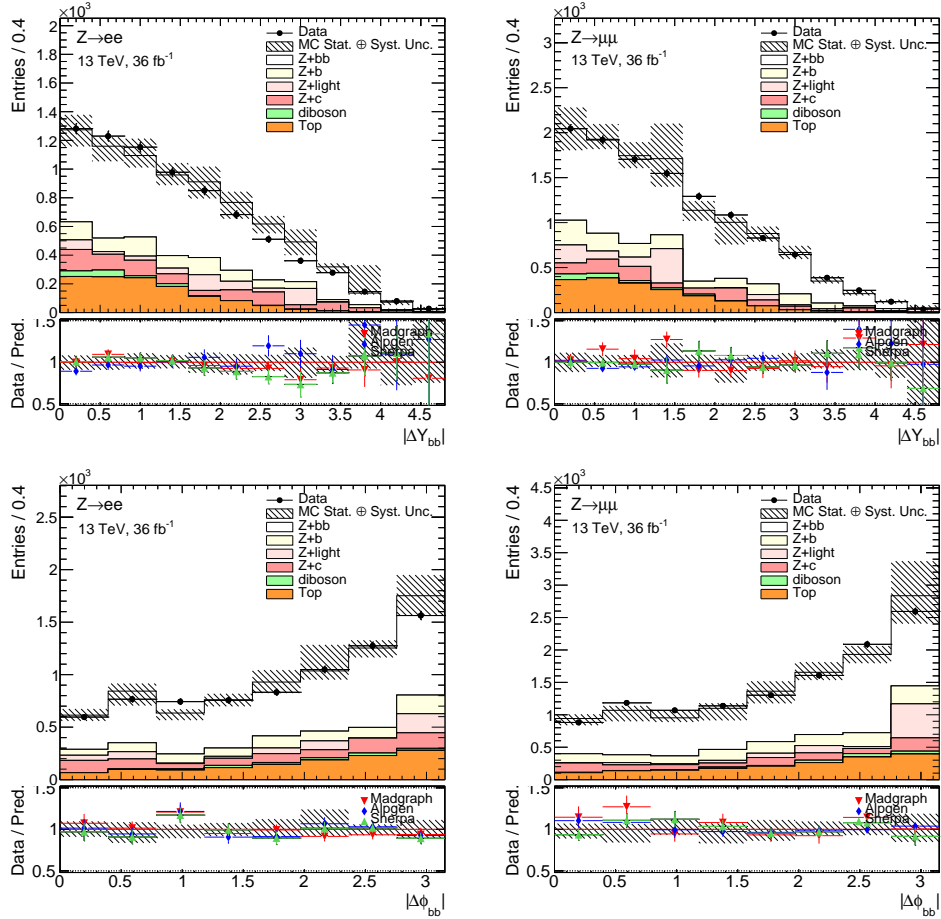
Figure 4.29: Data and Monte Carlo comparison for the di-bjets Δy_{bb} (top) and $\Delta \Phi_{bb}$ (bottom). The normalisation of Z+bb, Z+b, Z+cc, Z+c and Z+l is given by the combined fit result. SHERPA generator is used to model the Z+jets processes (top) and the ratio Data/Pred. is shown for SHERPA (green), MADGRAPH (red) and ALPGEN (blu) (bottom). The systematic band is evaluated using SHERPA and does not include luminosity uncertainty.

Chapter 5

Cross section measurements

The Z+b-jets analysis presented in Chapter 4 leads to the selection of events with a Z boson decaying in muon or electron pairs in association with one or two b-jets, reconstructed in ATLAS (detector level). The distributions of events, including both signal and background, as a function of several physics observables are extracted from data and compared to Monte Carlo simulations.

In this Chapter, the methodology employed for the measurement of the integrated and differential cross sections at particle level is described. The definition of the particle level objects is presented in Section 5.1: in particular, the particle level signal regions are inclusive in the number of b-jets (Z+$\geq$1 b-jet and Z+$\geq$2 b-jets), allowing, as detailed in the text, the measurement of inclusive cross sections. The extrapolation from the exclusive distributions (at reconstruction level) to the inclusive (at particle level) is performed through the unfolding procedure. In Section 5.2 the explanation of the unfolding methodology and the necessary ingredients is given and the tests performed to verify the closure and the stability of the full procedure are described in Section 5.3. The systematics associated to the cross section measurement strategy are described in Section 5.4.

The measured variables and the corresponding binning choice are presented in Table 5.1.

Physics observable	Signal region	Range
Z $\lvert y\rvert$	Z+$\geq$1 b-jet	[0, 2.25]
leading b-jet p_T [GeV]	Z+$\geq$1 b-jet	[20, 500]
leading b-jet $\lvert y\rvert$	Z+$\geq$1 b-jet	[0, 2.25]
$\lvert\Delta y(Z,b)\rvert$	Z+$\geq$1 b-jet	[0, 3.2]
$\Delta R(Z,b)$	Z+$\geq$1 b-jet	[0.8, 4.8]
$\Delta\Phi(Z,b)$ [rad]	Z+$\geq$1 b-jet	[0.4, π]
sub-leading b-jet p_T [GeV]	Z+$\geq$2 b-jets	[20, 300]
sub-leading b-jet $\lvert y\rvert$	Z+$\geq$2 b-jets	[0, 2.25]
M_{bb} [GeV]	Z+$\geq$2 b-jets	[0, 500]
ΔR_{bb}	Z+$\geq$2 b-jets	[0.4, 4]
$\lvert\Delta y_{bb}\rvert$	Z+$\geq$2 b-jets	[0, 4]
$\Delta\Phi_{bb}$ [rad]	Z+$\geq$2 b-jets	[0, π]

Table 5.1: Summary of the measured differential distributions for the Z+1 b-jet and Z+2 b-jets analyses.

5.1 Particle level definition

The distributions of the reconstructed number of events as a function of physics observables must be corrected for detector effects to the particle level, by using Monte Carlo samples. The correction procedure is called "unfolding" and consists in converting the reconstruction level observables in data into particle level quantities, allowing the comparison with theoretical calculations. As detailed in Chapter 4, particle level calculations are formulated in terms of jets, dressed leptons (leptons after radiating a photon) and isolated photons, built out of the final state particles (hadrons, leptons and photons) simulated by Monte Carlo generators. The latter indeed adopt procedures similar to reconstruction algorithms used by experiments, to cluster final state particles into jets, dressed leptons and isolated photons. The unfolding methodology accounts for inefficiency and resolution effects both in jets and Z boson selections.

The cross section measurements are performed in the same fiducial kinematic region of the reconstruction level distributions, to avoid theoretical uncertainties in the extrapolation to the full phase space. The fiducial volume is based on the selection of physics objects defined in Table 5.2.

At particle level, electron and muon kinematics includes the four-momentum of the photons radiated in a cone of radius 0.1, around the final state lepton direction ("dressed leptons"). The Z boson four-momentum is defined by adding the four momenta of selected dressed leptons within an invariant mass window 76 GeV$< m_{ll} <$106 GeV.

Leptons	dressed objects $p_T >$27 GeV and $|\eta| <$2.5
Z boson	2 same flavour and opposite charge leptons $m_{ll} = (91 \pm 15)$ GeV
b-jets	matching with a B-hadron $p_T >$20 GeV and $|y| <$2.5 ΔR(b-jets, l)$>$0.4

Table 5.2: Fiducial region at the particle level.

As mentioned before, jets are built from the final state hadrons and their decaying products, using the anti-k$_t$ algorithm. Ambiguities from overlapping leptons and jets are resolved by requiring ΔR(b-jets, l)$>$0.4.
Jets are identified as b-jets via the hadron cone matching, namely the same procedure used to flavour classify events at reconstruction level. Jets are matched to a weakly-decaying B-hadron with $p_T >$5 GeV in a cone of radius ΔR=0.3.
At particle level, signal regions are defined by the Z boson produced with at least one or at least two b-jets (i.e. Z+$\geq$1 b-jets and Z+$\geq$2 b-jets). These two regions are inclusive in the number of b-jets to allow for the measurement of inclusive integrated and differential cross sections. The choice of having exclusive b-jet regions at detector level is driven by the consequent reduction of the $t\bar{t}$ background, while at particle level the inclusive regions allow a reduced statistical uncertainty with respect to the exclusive ones.

5.2 Unfolding

5.2.1 The Bayesian method

The unfolding method is used to infer from the observed physics quantities the "true" ones, corrected for detector distortions. In this analysis, the unfolding corrects also for the differences between the detector level (exclusive Z+1 b-jet and Z+2 b-jets) and particle level (inclusive Z+≥1 b-jet and Z≥+2 b-jets) kinematic regions. The method commonly used to solve this problem is the so called "bin-by-bin" correction. In an integrated measurement the ratio of the number of Monte Carlo events reconstructed in a certain bin over the number of truth events in the same bin is determined. This factor represents an efficiency and it is used to count the number events at particle level, from the number of reconstructed events in that bin. Since this method does not account for correlations between adjacent bins, it can be used to perform differential cross section measurements only if the bin-to-bin migrations are small compared to the bin width of the distribution. If this condition is not valid, additional corrections must be determined to take into account the bin-to-bin migrations. These corrections are performed by a matrix which relates the number of Monte Carlo events generated in one bin of a certain distribution to the number of events reconstructed in any bin of the distribution (*migration matrix*). The matrix is then inverted and applied to the reconstructed distributions of real data (*response matrix*), as explained in Eq.4.1. The iterative Bayesian unfolding technique [108] used in the Z+b-jets differential cross section measurements exploits the Bayes' theorem to build the response matrix. In a first step, according to the Bayes' theorem, the response matrix can be expressed as:

$$M(T_i|R_j) = \frac{M(R_j|T_i) \cdot P_0(T_i)}{\sum_{l=1}^{N} M(R_j|T_l) \cdot P_0(T_l)} \qquad (5.1)$$

where $M(R_j|T_i)$ is the *migration matrix* which represents the probability that an event is generated at truth level (T_i) in bin i, but it is reconstructed (R_j) in bin j. $P_0(T_i)$ is the "prior", namely an a-priori known distribution of the physics quantity under investigation. In the first step of the iteration,

the truth distribution from Monte Carlo generator is used as prior. The denominator of Eq.5.1 represents the overall normalisation, which accounts for all the different processes affecting the reconstruction in the bin j. In principle, the choice of the prior function can strongly influence the final result, introducing a bias in the cross section measurement. This is solved by applying the Bayes's theorem in an iterative sequence, where the result of each iteration is used as a prior for the following step. The more iterations are performed, the less the dependence on the initial prior choice is achieved. On the other hand, while the initial bias is reduced, the statistical fluctuations in the migration matrix increase with the number of iterations and can be misinterpreted as real migrations due to detector effects. The number of iterations must therefore be a balance between the two effects.

5.2.2 Response matrices and fake corrections

Two ingredients are needed to measure the differential cross section: the response matrix and the fake correction, all described below.

For each observable, the migration matrix is filled on the x axis with the "matched" events, namely the events at reconstruction-level passing the selection of Section 4.4, with a correspondence at truth level. On the y axis the matrix is filled with truth events falling in the fiducial volume (see Section 5.1). The response matrices are obtained by inverting the migration ones and the entries in each column are normalised to the total number of truth events that fulfil the event selection in that bin. In this way, the sum of the values in a given column represents the reconstruction efficiency for the corresponding truth bin.

The response matrices are shown for the leading and second leading b-jet kinematics variables p_T and y in Fig.5.1 for the Z+$\geq$1 b-jet and Z+$\geq$2 b-jets samples. While for the rapidity observables migrations are all below 2%, the p_T distributions suffer for higher migrations in low p_T bins and are not symmetric, indicating that jets tend to be reconstructed as less energetic than at particle level. In Fig.5.2 response matrices are shown for the Z rapidity and for its angular separation from the leading b-jet in the Z+$\geq$1 b-jet processes.
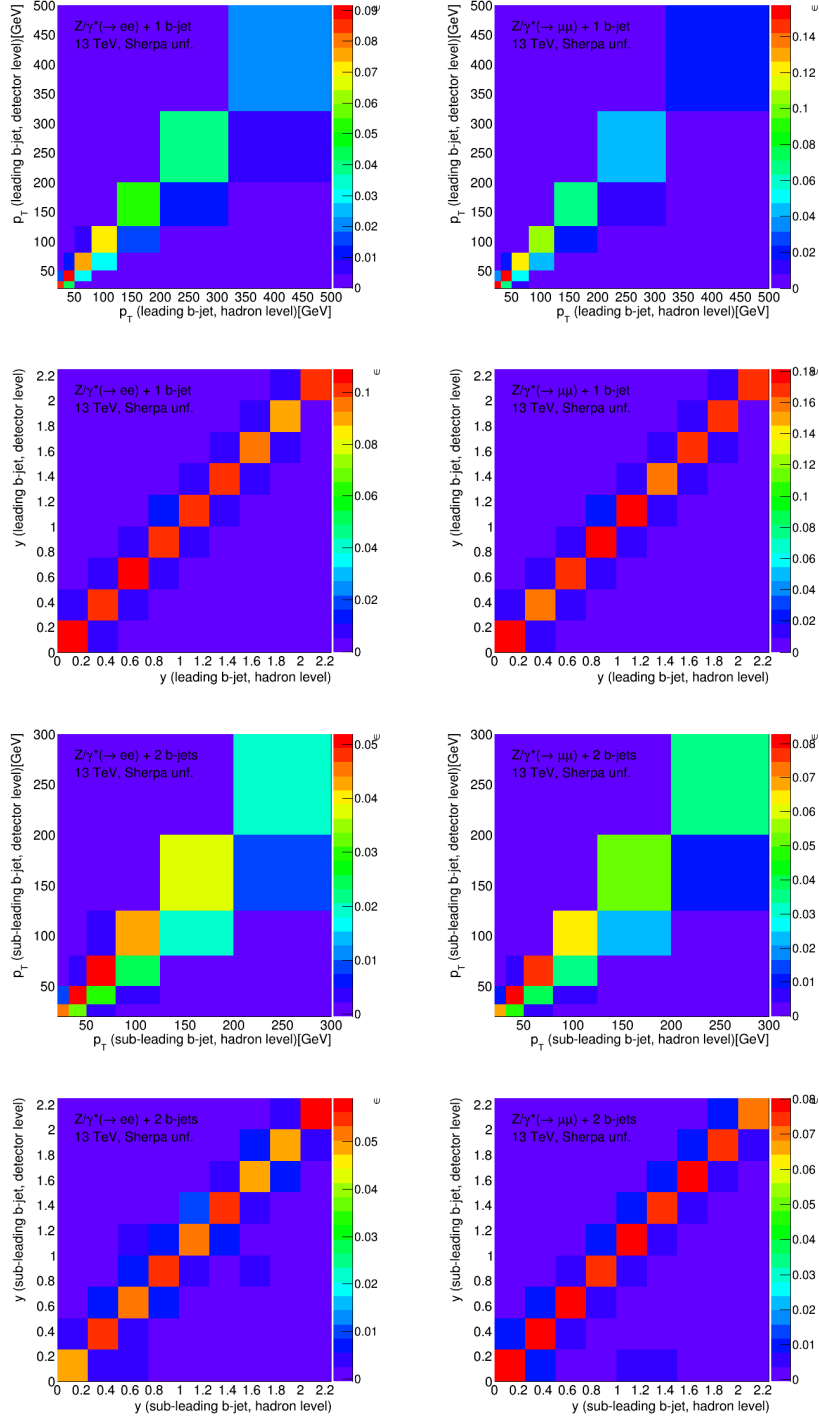
Figure 5.1: Response matrices for the kinematic distributions of the leading b-jet p_T (top row) and y (second row) for the Z+≥1 b-jet processes in the electron (left) and muon (right) channels by using SHERPA. The response matrices for the sub-leading b-jet p_T (third row) and y (bottom row) in the Z+2 b-jets mode are also shown.
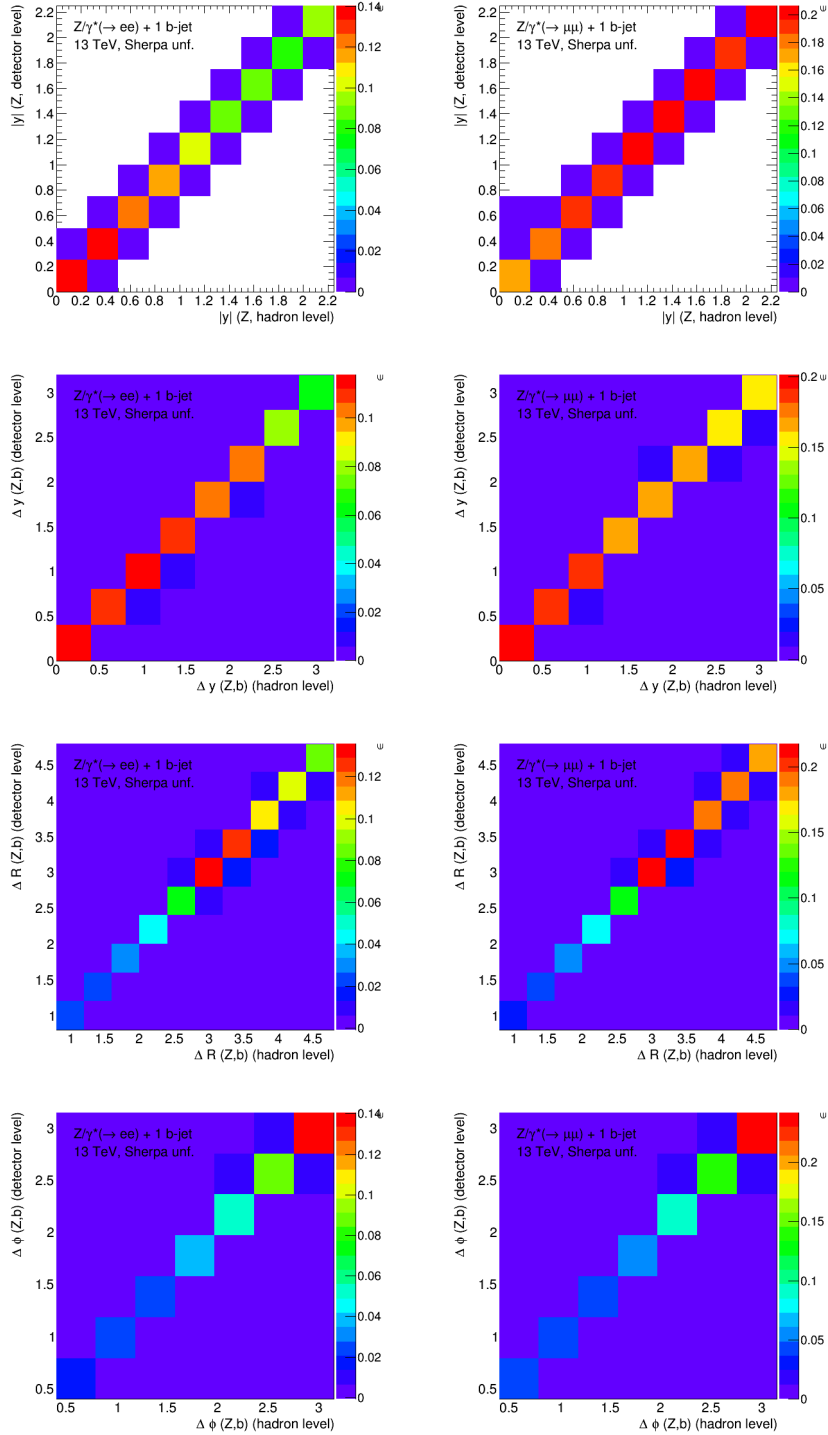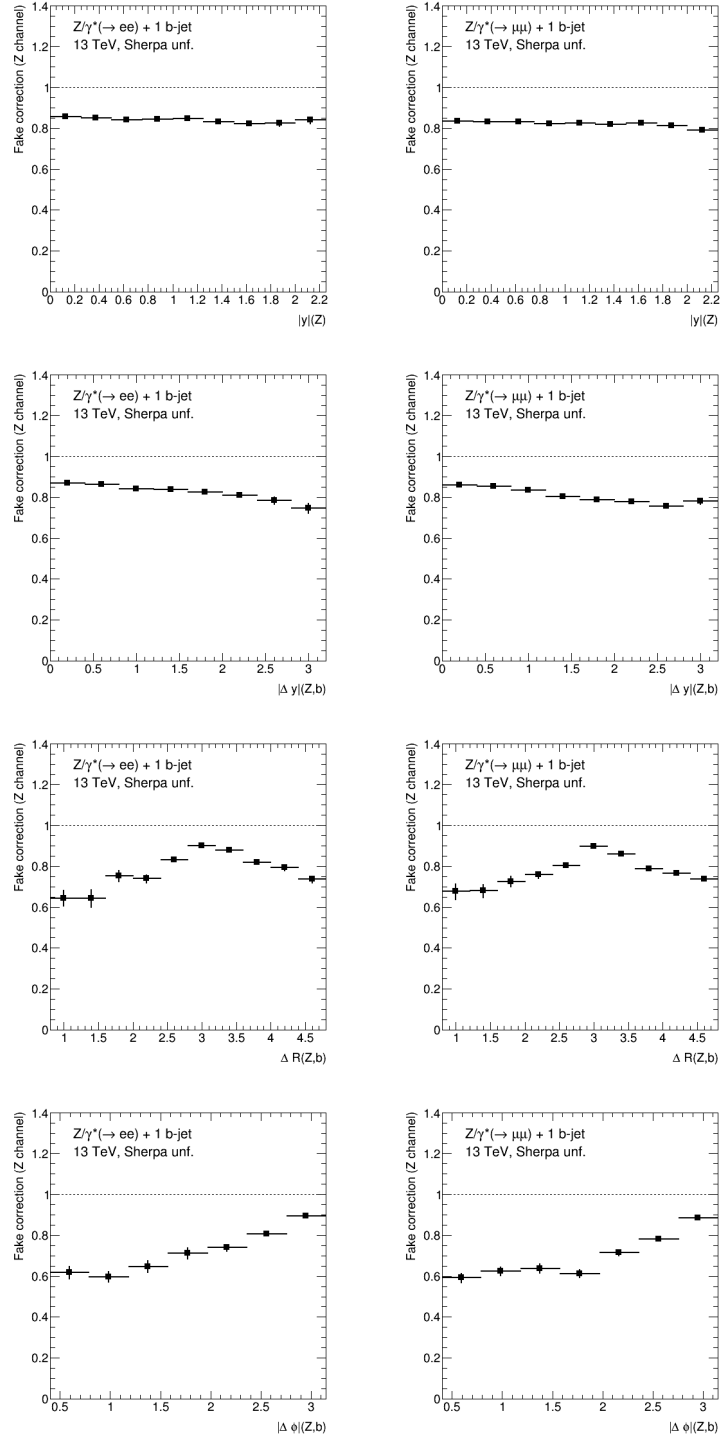
Figure 5.2: Response matrices for the kinematic distributions of y_Z (top row), $\Delta y(Z, b)$ (second row), $\Delta R(Z, b)$ (third row) and $\Delta \Phi(Z, b)$ (bottom row) for the Z+≥1 b-jet processes in the electron (left) and muon (right) channels by using SHERPA.
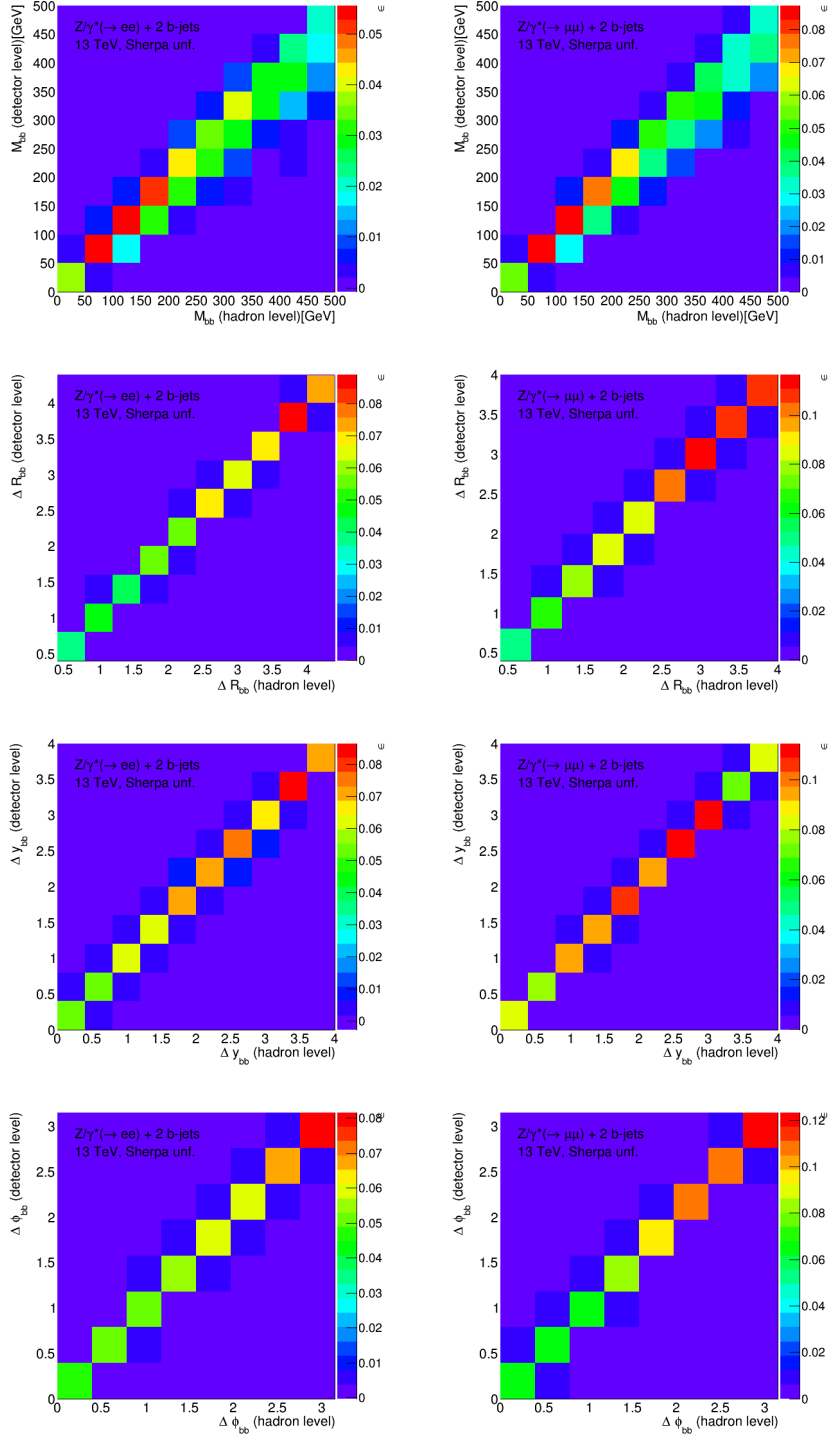
Figure 5.3: Response matrices for the kinematic distributions of M_{bb} (top row), ΔR_{bb} (second row), Δy_{bb} (third row) and $\Delta \Phi_{bb}$ (bottom row) for the Z+$\geq$2 b-jets processes in the electron (left) and muon (right) channels by using SHERPA.

Fig.5.3 presents the matrices for various di-b-jets observables, all showing migrations below 5%. In general, the diagonal terms in the muon channel are typically higher than the ones in the electron channel, meaning that the muon reconstruction is in general better than the electron one.

Distributions at reconstruction-level can contain additional contributions due to events generated outside but reconstructed inside the fiducial volume (examples are pileup jets or jets reconstructed above the p_T threshold, but corresponding to softer objects at particle level). A fake correction term (f_j) is evaluated to account for such effects and it is applied as a multiplicative factor, before the unfolding procedure (see Eq.4.1), both on the background-subtracted data and to the reconstructed Monte Carlo events. Fig.5.4 shows the fake correction distribution for the leading and second leading b-jets kinematics. The largest corrections are usually located at the edge of the kinematic distributions, as for example at low p_T, where the effect of pileup have a larger impact. Fake corrections for Z and Z-bjet observables are shown in Fig.5.5 for the electron and muon channel and for di-b-jets quantities in Fig.5.6.

5.3 Unfolding tests

In order to very the stability and consistence of the unfolding procedure, detailed studies have been performed.

5.3.1 Closure test

A basic sanity check consists in using the same sample to emulate both data and Monte Carlo distributions. Since SHERPA is used as nominal generator for the signal modelling, it has been chosen as input for this study. Fig.5.7 shows, as example, the results for the leading b-jet p_T and y in the Z+$\geq$1 b-jets analysis in both electron and muon channels. Similar results are presented for the Z+$\geq$2 b-jets signal as a function of the invariant mass of the two b-jets (M_{bb}) and their azimuthal separation ($\Delta\Phi_{bb}$) in Fig.5.8. As one would expect, the unfolded reconstructed distribution perfectly repro-
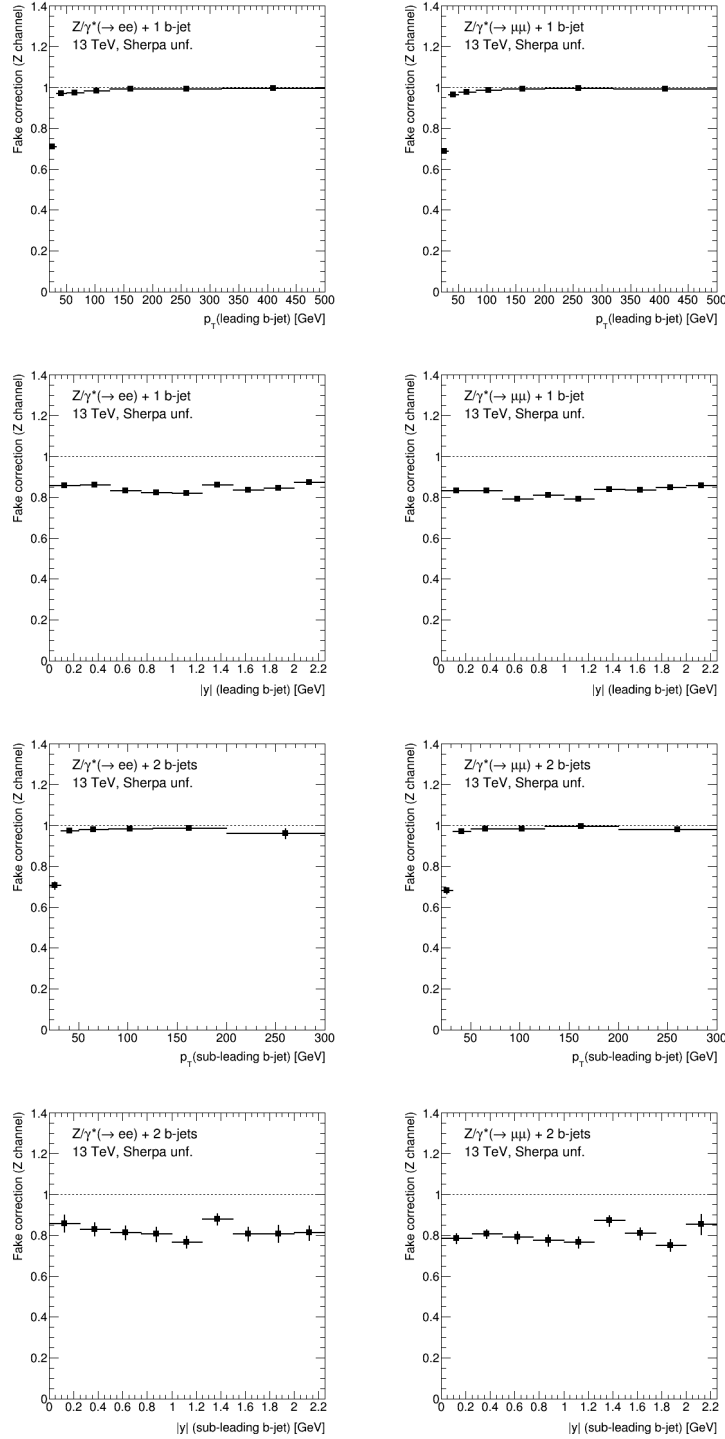
Figure 5.4: Fake correction for the kinematic distributions of the leading b-jet p_T (top row) and y (second row) for the Z+1 b-jet analysis and of the sub-leading b-jet p_T (third row) and y (bottom row) in the Z+≥2 b-jets mode in the electron (left) and muon (right) channels by using SHERPA.

Figure 5.5: Fake correction for the kinematic distributions of y_Z (top row), $\Delta y(Z, b)$ (second row), $\Delta R(Z, b)$ (third row) and $\Delta\Phi(Z, b)$ (bottom row) for the Z+$\geq$1 b-jet processes in the electron (left) and muon (right) channels, by using SHERPA.
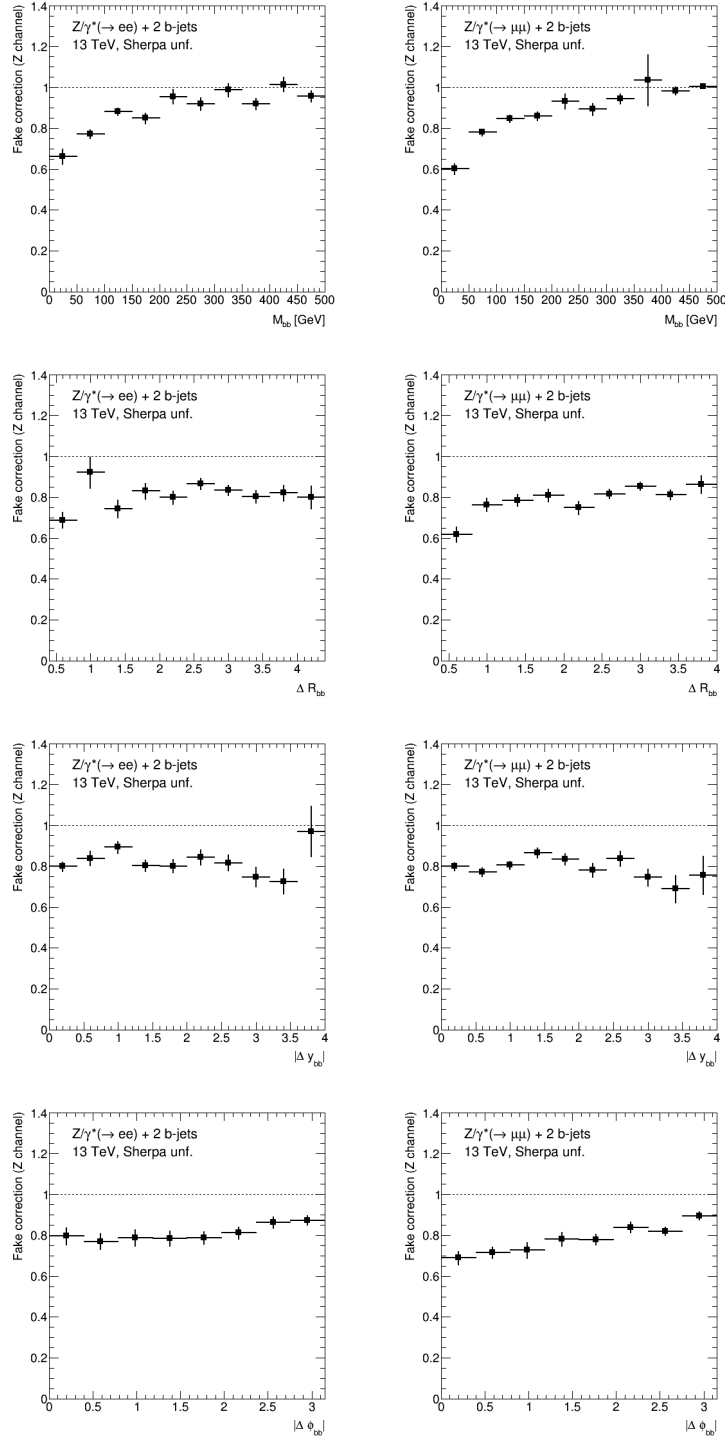
Figure 5.6: Fake correction for the kinematic distributions of of M_{bb} (top row), ΔR_{bb} (second row), Δy_{bb} (third row) and $\Delta \Phi_{bb}$ (bottom row) for the Z+2 b-jets processes in the electron (left) and muon (right) channels, by using SHERPA.

duce the Monte Carlo truth distributions, showing a good closure of the full methodology.

A second test follows directly from the previous. The events generated by SHERPA are divided in two categories according to their event number; the even events are used to emulate data and the odd ones to correct for detector effects. The results for the Z+≥1 b-jet and Z+≥2 b-jets analyses are shown in Fig.5.9 and Fig.5.10 for the same variables as before, as an example. Again the closure is ensured with the only difference that larger statistical uncertainties are observed with respect to the previous test, due to the smaller sample used.

5.3.2 Tests on the number of iterations

As mentioned in Section 5.2.1, the bayesian unfolding is an iterative procedure which in principle depends on the choice of the prior. This dependence weakens increasing the number of iterations, causing on the other hand the increase of statistical uncertainties. A detailed study has been performed on these two effects, by building two functions describing them separately.
To estimate the dependence on the choice of the initial prior, the following variable is introduced:

$$
f = \sum_{n=1}^{N_{bins}} \frac{|U(i)_n - U(i-1)_n|}{U(i)_n} \quad ,
$$
(5.2)

where the relative difference between the result of iteration i $(U(i)_n)$ and the preceding $(U(i-1)_n)$ is evaluated for all the iterations and summed over all the bins (n) of the analysed physics quantity. For each iteration, f is calculated, normalised to the total number of bins (N_{bins}) and shown as a function of the number of iterations in Fig.5.11. Since in the bayesian unfolding, the result of step $i-1$ is used as prior for the i-th iteration, this discrete function gives an idea of the dependence on the first prior, and it is expected to decrease as a function of the number of iterations and to converge to a value which remains stable with further iterations.
To study the behaviour of the statistical uncertainty, the relative statistical
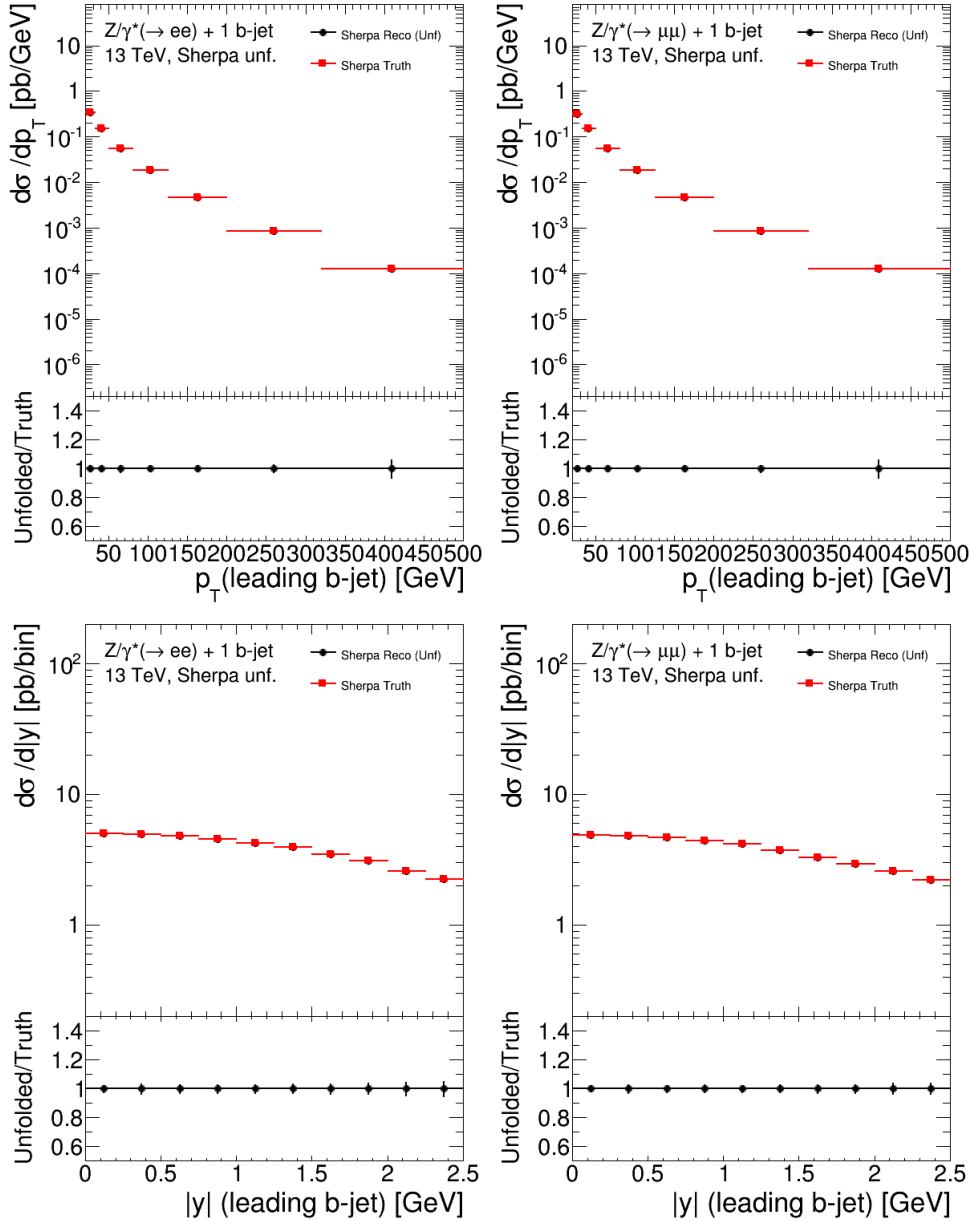
Figure 5.7: Closure test for the unfolding procedure, by using SHERPA both to emulate signal and to correct for detector effects. Differential cross sections as a function on the leading b-jet p_T (top) and y (bottom) for the production of the Z+≥1 b-jet process are shown in the electron (left) and muon (right) channels.
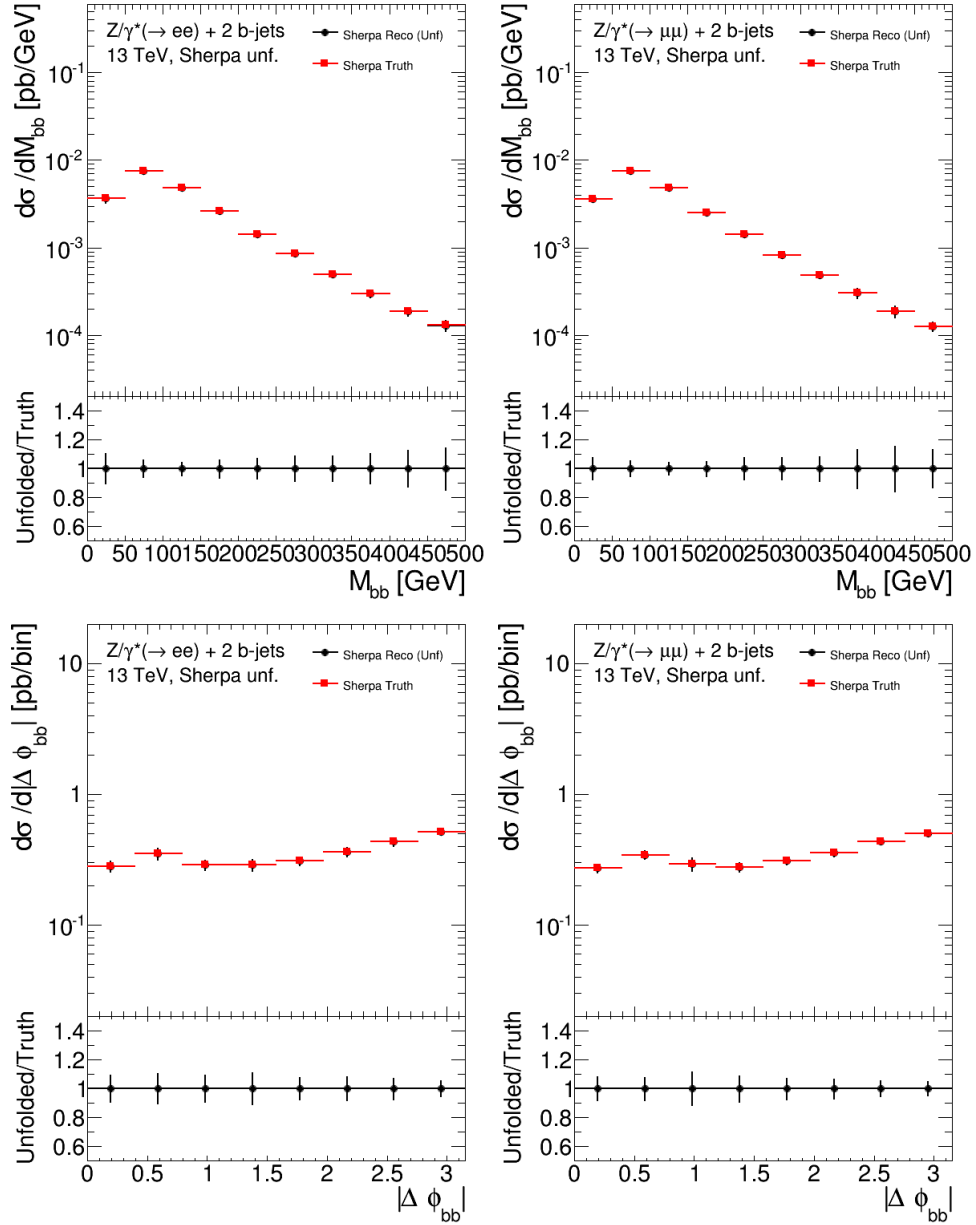
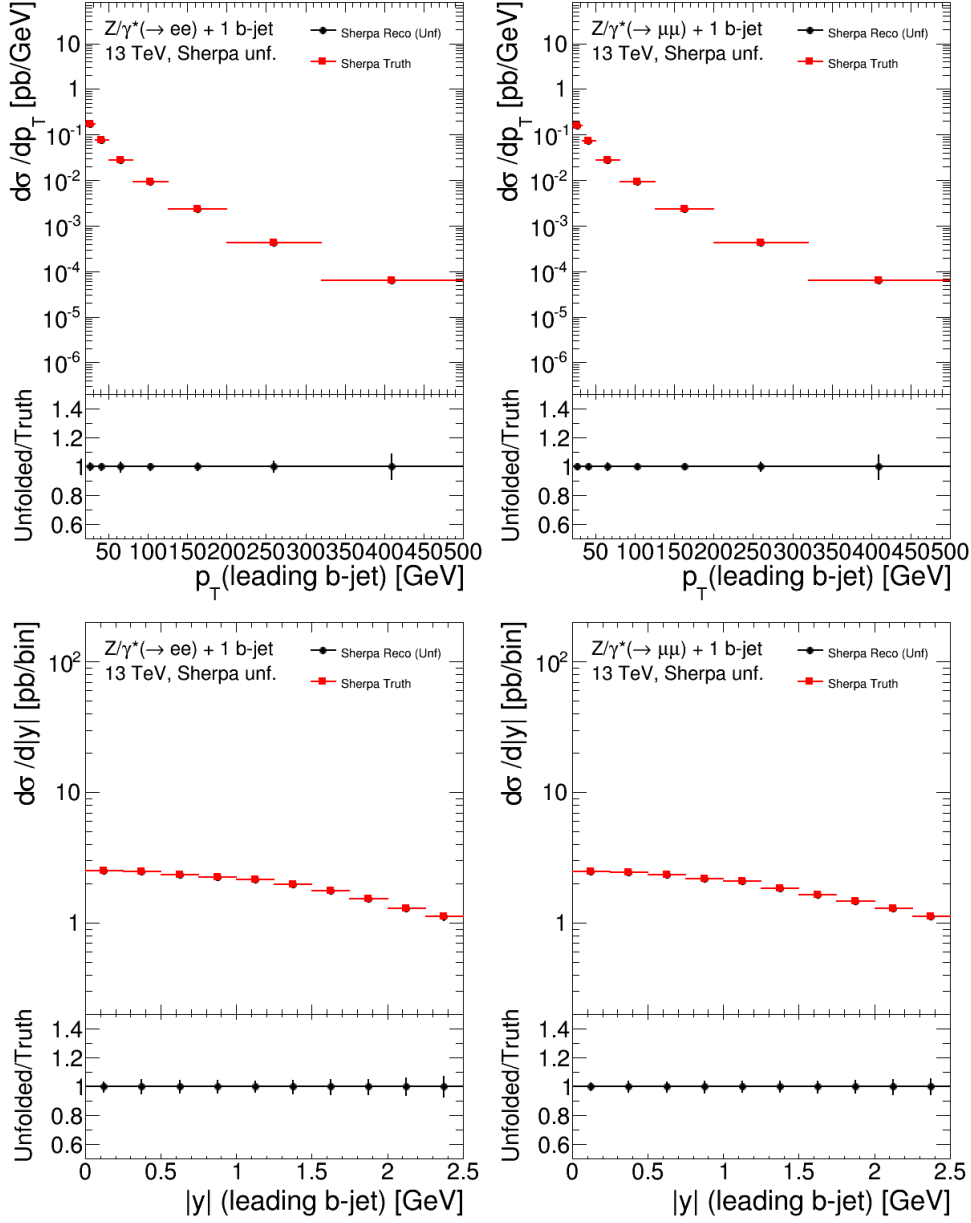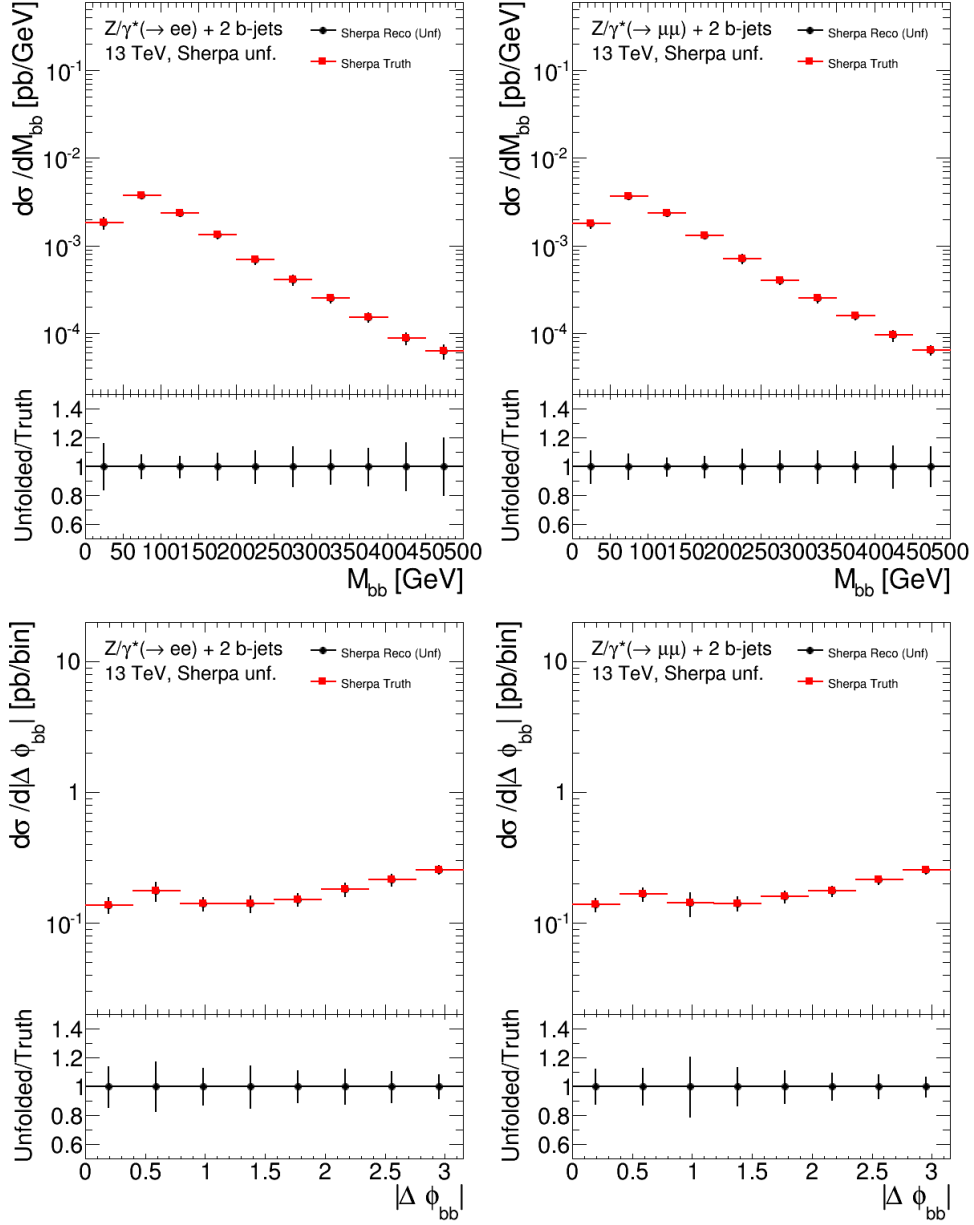Figure 5.8: Closure test for the unfolding procedure, by using SHERPA both to emulate signal and to correct for detector effects. Differential cross sections as a function on the M_{bb} (top) and $\Delta\Phi_{bb}$ (bottom) for the production of the Z+$\geq$2 b-jets process are shown for the electron (left) and muon (right) channels.

Figure 5.9: Closure test for the unfolding procedure, by using half SHERPA sample to emulate data and the other for the unfolding. Differential cross sections as a function on the leading b-jet p_T (top) and y (bottom) for the production of the Z+1 b-jet process. Cross sections are shown for the electron (left) and muon (right) channels.

Figure 5.10: Closure test for the unfolding procedure, by using half SHERPA sample to emulate data and the other for the unfolding. Differential cross sections as a function on the M_{bb} (top) and $\Delta\Phi_{bb}$ (bottom) for the production of the Z+2 b-jets process. Cross sections are shown for the electron (left) and muon (right) channels.

error is introduced:

$$g = \sum_{n=1}^{N_{bins}} \frac{\Delta(U(i)_n)}{U(i)_n} \quad , \tag{5.3}$$

where $\Delta(U(i)_n)$ is the statistical error associated to the i-th iteration and $U(i)_n$ is the corresponding unfolded result. The g variable is evaluated for each step of the unfolding algorithm, normalised for the total number of bins and plotted as a function of the number of iterations.
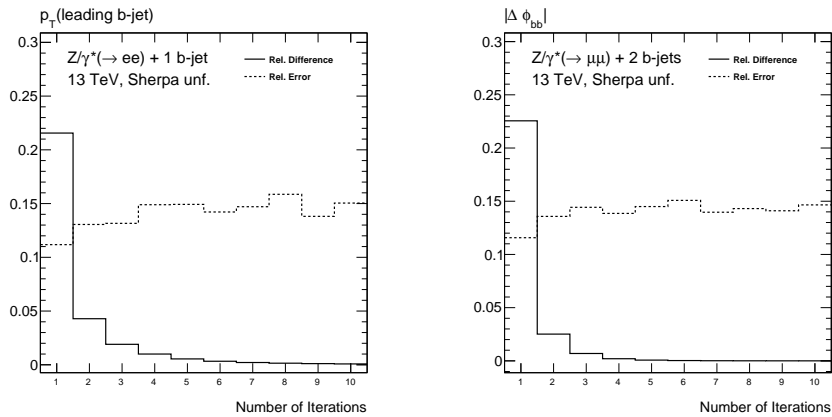
Figure 5.11: Test on the number of iterations by using SHERPA to unfold data. The relative difference (black line) and the relative error (dashed line) functions are shown for the distributions of the leading b-jet p_T in the electron channel (left) and the $\Delta\Phi_{bb}$ in the muon channel (right), for Z+$\geq$1 b-jet and Z+$\geq$2 b-jets processes, respectively.

Fig.5.11 shows the relative difference (f) and the relative error (g) distribution as a function of the number of iterations for the leading b-jet p_T and the $\Delta\Phi_{bb}$. The dependence from the initial prior is $\leq$25% in the first iteration, and decreases steeply below 5% after from second step, while modest reductions are achieved in the following iterations. As the statistical uncertainty presents a slow increase after the second iteration, a number of iterations equal to two is chosen, representing the best compromise between the two competitive effects.

5.4 Uncertainties on the unfolding

Uncertainties on the unfolding procedure are mainly due to two sources: the modelling of the b-jet observable implemented in the generator employed to correct for detector effects and the statistical fluctuations of the Monte Carlo sample. A brief description is given in the following.

5.4.1 Modelling and Template of flavour fit

The flavour fit procedure used to extract the normalisations of Z+b, Z+c and Z+l processes is sensitive to the modelling of the properties of the hadrons decaying into jets, defined by the used generator. Although the normalisations are consistent among SHERPA, ALPGEN and MADGRAPH, some differences remain and should be considered as a systematic source. The choice of the generator have additional impact on the migration matrices, fake corrections and, obviously, on the initial priors. In this context, the effects due to the choice of the nominal generator are considered as modelling systematics.

The modelling uncertainty is evaluated by repeating the full analysis chain by using ALPGEN instead of SHERPA, from the signal and background estimation, to the flavour fit and the unfolding procedure. The difference in the final cross section measurements obtained by using the two generators is taken as modelling uncertainty. Finally, its impact on the integrated and differential cross section is symmetrised. This preliminary approach gives an overestimation of the uncertainty on both the flavour fit and on the unfolding uncertainties. In view of the publication, the two sources will be disentangled: this procedure will be used for the flavour fit systematic, while for the modelling uncertainty, impacting the unfolding inputs (such as matrices and priors), a data-MC reweighting (based on the difference between SHERPA and ALPGEN) will be applied.

5.4.2 Statistical fluctuations on Monte Carlo samples

The procedure used to unfold the detector level measurements can be affected by the finite size of the Monte Carlo sample. This can cause distortions

in the migration matrices and have impact on the statistical error affecting the fake correction distributions. These effects are studied with a toy-Monte Carlo, where various inputs to the unfolding are allowed to fluctuate independently according to Gaussian distributions.

The particle level yield is modified by extracting the population in each bin according to a Poisson distribution with mean equal to the original Monte Carlo prediction. Correspondingly, the population of each bin in the unfolding matrix, is defined by extracting it from a binomial distribution with number of trials equal to the population of each bin and success probability according to the originally predicted probability of reconstructing in bin j the object belonging to bin i at truth level. Finally, the number of fake signal events bin is extracted from a binominal distribution with a number of trials equal to the expected number of reconstructed objects in that bin and success probability equal to the fake signal rate predicted in the same bin. The unfolding is then carried out using the various migration matrices and fake correction factors obtained. For each bin of the unfolded distribution the ratio to the nominal results of the analysis is estimated and distributions of these ratios for all bins are produced. The distributions over 100 pseudo experiments have a Gaussian behaviour with mean equal to one and the sigma of the distribution is used as systematics uncertainty on the Monte Carlo statistics.

The impact of the unfolding systematic uncertainties is reported in Section 6.2 together with all other systematic contributions.

Chapter 6

Results

In Chapters 4 and 5 all the steps concerning the measurements of the cross sections for the production of the Z+1 b-jet and Z+2 b-jets processes have been presented. In particular the definition of the particle level observables is detailed in Section 5.1.

In Section 6.1 a brief summary of the Monte Carlo calculations that are compared to the final cross section results is given. In Section 6.2 the impact of each systematic contribution in the cross section measurements is detailed for each variable studied in this thesis. The results of the integrated cross sections for the Z+$\geq$1 b-jet and Z+$\geq$2 b-jets processes are presented in Section 6.3. Finally the differential cross sections of the Z+$\geq$1 b-jet and Z+$\geq$2 b-jets as a function of several b-jet observables are described in Section 6.4 and Section 6.5, respectively.

6.1 Theoretical description of the Z+b-jets processes

The theoretical predictions for the production of the Z boson in association with b-jets is challenging. As anticipated in Chapter 1, two main processes are investigated, which are referred to as the 4*Flavour Number Scheme* (4FNS) and the 5*Favour Number Scheme*. In the 4FNS only the four lightest quark (i.e. u, d, c and s) can be present in the initial stage and generate a final state with at least two b-quarks through the gluon splitting. On the

other hand, in the 5FNS the b-quark is assumed to contribute to the hard
process and its PDF is calculated from the gluon density evolution above the
b-mass threshold. Therefore another interesting feature to investigate is the
treatment of the b-quark mass ($m_{b-quark}$) in the PDFs set of the previous
schemes.

Physical observables involving jet production are sensitive to non perturba-
tive effects:

- hard double parton interactions (DPI), characterised by additional b-
  jets produced in a secondary interaction in the same p-p collision where
  the Z boson is produced;

- underlying events, representing the additional hadronic activity from
  the proton remnants, which increase the overall production of hadrons;

- fragmentation and hadronisation, mapping the particle level predic-
  tions into final state hadrons.

The Monte Carlo calculations used in this thesis for the prediction of the
Z+b-jets processes have been described in detail in Section 4.3. Since the
cross section results are compared directly to these calculations, the main
features of SHERPA, MADGRAPH and ALPGEN calculations are summarised
in Table 6.1.

Monte Carlo calculations				
Generator	ME precision	scheme	$m_{b-quark}$	PDFs
SHERPA	NLO	5FNS	✓	NNPDF3.0nnlo
MADGRAPH	LO	5FNS	✗	NNPDF3.0nnlo
ALPGEN	LO	4FNS	✓	CTEQ6L1

Table 6.1: Summary of the Monte Carlo calculations used for comparison with
the Z+b-jets cross section measurements. MADGRAPH 5_AMC@NLO is referred
to as MADGRAPH in the table and in the text. $m_{b-quark}$ indicates the mass of the
b-quark.

6.2 Systematic uncertainties on the final results

The uncertainty contributions on the final results come from the reconstruction level systematics and the systematics related to the unfolding procedure. In the flavour fit procedure, all the detector level systematics explained in detail in Section 4.5 are considered as nuisance parameters, being their contribution free to float in the fit. The errors associated to the obtained scale factors are highly correlated to the variations applied to the Z+jets background processes, when estimating the impact of each detector level uncertainty. Since the unfolding procedure is performed after applying the variation corresponding to each detector level systematics, for both signal and background, adding the error on the scale factor in the unfolding routine would imply double counting of detector level systematics. For this reason the errors on the scale factors are not propagated further.

From Fig.6.1 to Fig.6.6, the relative contribution of each systematic source is shown as a function of several observables. Since the uncertainties are in agreement in the two lepton modes, the breakdown of the systematic components is show for one channel only. The various uncertainties are represented with different colours, while the black line represents the total systematic band, given by the sum in quadrature of all the components. Fig.6.1 and Fig.6.2 show the uncertainties for the kinematic of the leading and sub-leading b-jets, respectively. The dominant uncertainty is given by the modelling, explained by the large difference between SHERPA and ALPGEN description of b-jet quantities. It is important to recall that the modelling involves both the difference between the two generators and the difference in the template shape used in flavour fit. The second largest systematics is given by the b-tagging systematics (on average $\sim$20%). The Monte Carlo statistics dominates only in the bins where the number of events is particularly small. In few bins of the leading b-jet p_T and sub-leading b-jet y, the total systematics is very large: this is not surprising, considering the fluctuations of Monte Carlo generators in these regions, where the statistics is very low.

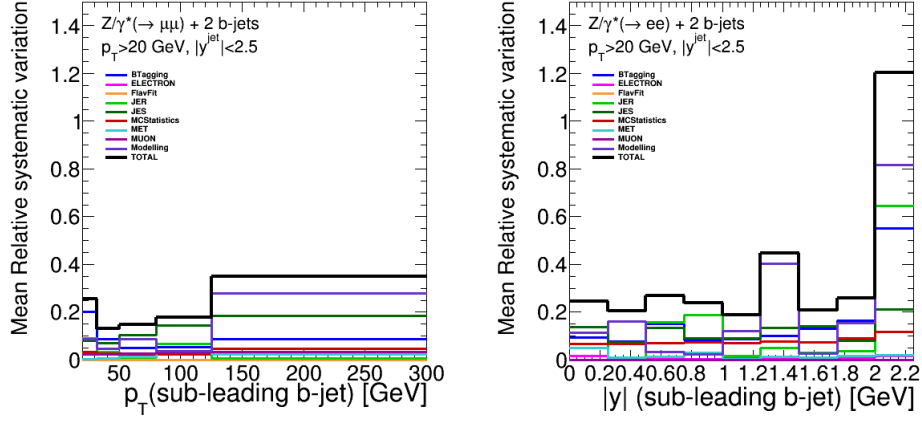The Z boson rapidity systematics is on average 25% (see Fig.6.3), with the
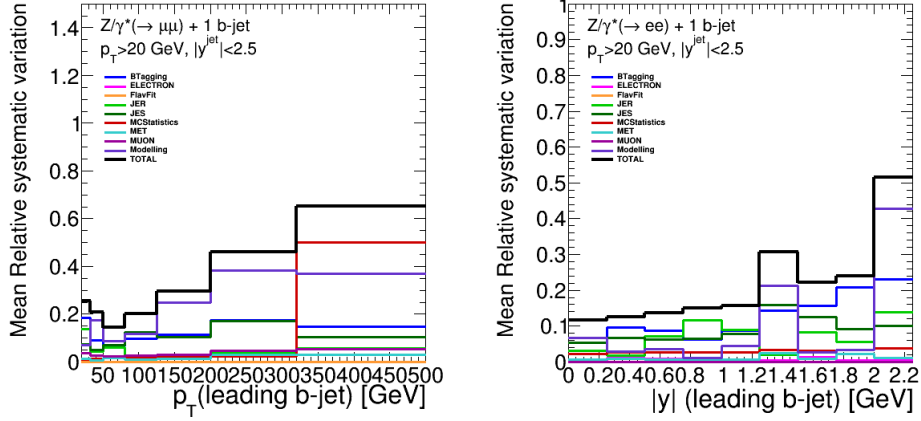
Figure 6.1: Breakdown of systematics uncertainty for the Z+≥1 b-jet cross section measurement as a function of the leading b-jet p_T in the muon channel (left) and y in the electron channel (right) using the SHERPA generator.

exception of the forward region, where, mostly in the muon channel, the systematics reaches 55%. The Δy(Z,b) (see Fig.6.3), ΔR(Z,b) and $\Delta \Phi$(Z,b) (see Fig.6.4) observables suffer for larger systematic contributions with respect to the Z rapidity, due the Monte Carlo modelling.

Fig.6.5 and Fig.6.6 show the systematic uncertainties for the Z+≥2 b-jets processes. In general, the uncertainties on di-b-jets quantities are dominated by the modelling and the b-tagging systematics. In some bins, the JES and JER contributions are relevant ($\sim$35%), while in general, the lepton uncertainties are below 5%.

The luminosity uncertainty is not shown in the plots, while the error associated to the flavour fit is reported (orange) for completeness, but it is not accounted for the total systematic band.

6.3   Integrated cross sections

Another goal of this thesis is the measurement of the integrated cross sections for the production of the Z+≥1 b-jet and Z+≥2 b-jets processes. For this purpose a bin-by-bin method has been applied to the distribution of the exclusive b-jets multiplicity ($n_{b-jets,ex}$, see Fig.4.23), and the measured

Figure 6.2: Breakdown of systematics uncertainty for the Z+≥2 b-jets cross section measurement as a function of the second leading b-jet p_T in the muon channel (left) and y in the electron channel (right) using the SHERPA generator.
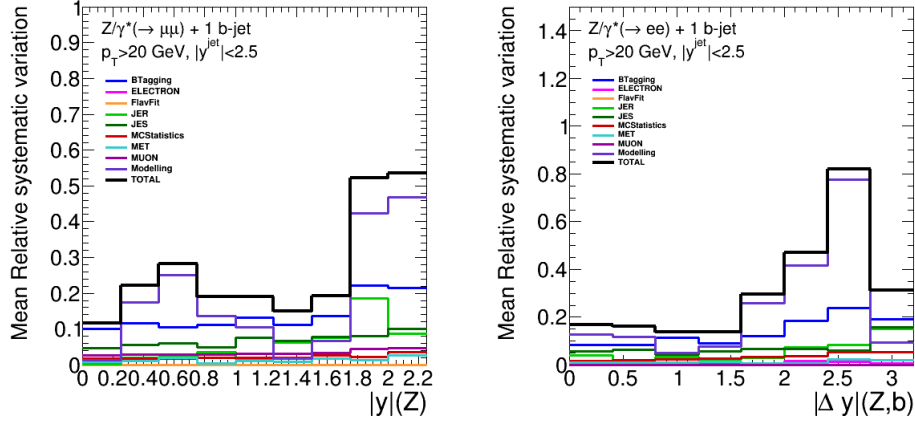
Figure 6.3: Breakdown of systematics uncertainty for the Z+≥1 b-jet cross section measurement as a function of the Z boson y in the muon channel (left) and $\Delta y(Z,b)$ in the electron channel (right) using the SHERPA generator.
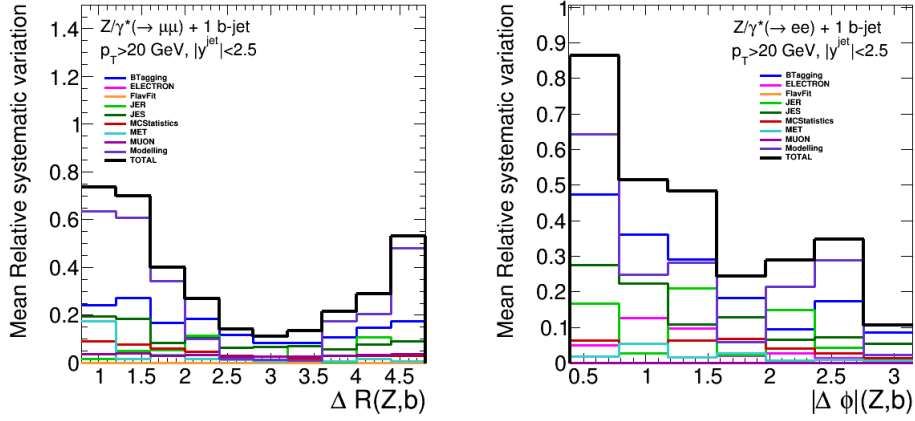
Figure 6.4: Breakdown of systematics uncertainty for the Z+≥1 b-jet cross section measurement as a function of ΔR(Z,b) in the muon channel (left) and $\Delta\Phi$(Z,b) in the electron channel (right) using the SHERPA generator.
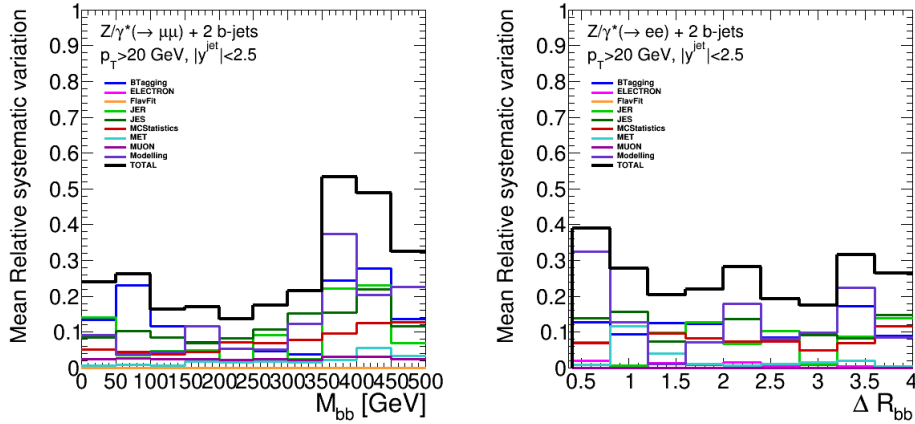
Figure 6.5: Breakdown of systematics uncertainty for the Z+≥2 b-jets cross section measurement as a function of the di-bjets observables M_{bb} in the muon channel (left) and ΔR_{bb} in the electron channel (right) using the SHERPA generator.
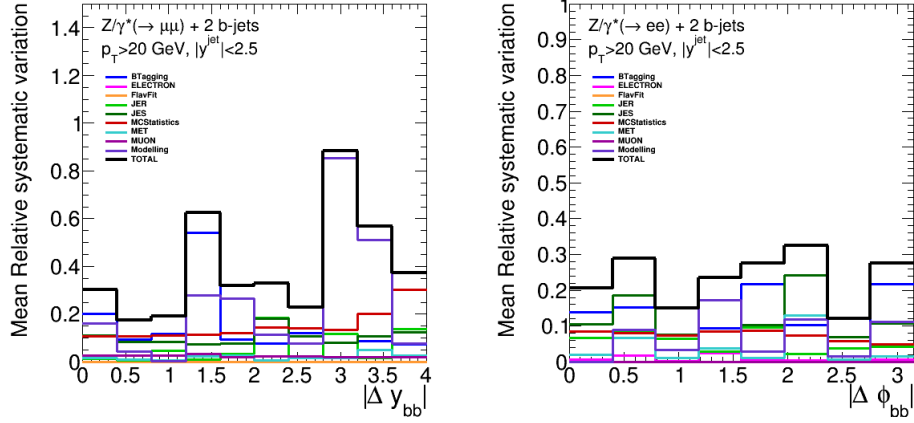
Figure 6.6: Breakdown of systematics uncertainty for the Z+$\geq$2 b-jets cross section measurement as a function of the di-bjets observables Δy_{bb} in the muon channel (left) and $\Delta\Phi_{bb}$ in the electron channel (right) using the SHERPA generator.

cross section can be expressed as:

$$\frac{d\sigma_i}{dX_i} = \frac{(N_i^{obs} - B_i)}{L \cdot \Delta X_i} C_i \tag{6.1}$$

where N^{obs} (B) is the number of observed (background) events in bin i, passing the detector level selection. The Z+jets background normalisations are properly corrected by using the scale factors resulting from the flavour fit in the combination of the two lepton channels. The number of signal events is normalised to the luminosity (L) and the bin width (ΔX) to obtain a cross section measurement. A correction term (C) is applied as a multiplicative factor, defined as the number of Monte Carlo events passing the particle level selection with respect to the number of reconstructed events. The factor C represents the correction for detector effects and the extrapolation of the measurement from the detector to the particle level. Moreover, due to the particle level selection, which is inclusive in the number of b-jets, the C term also describes the extrapolation from the exclusive b-jet selection at detector level (N_{jets}=1, 2), to the inclusive b-jet selection at particle level ($N_{jets} \geq$1, 2). Assuming bin-to-bin migrations are small compared to the bin width, the unfolding procedure is not applied for this measurement. In Fig.6.7 the

correction factors C are shown. Large corrections are visible in both electron and muon channels and are due to detector level requirements, not applied at particle level:

- exclusive b-jet region definition at detector level;

- veto on additional non b-jets;
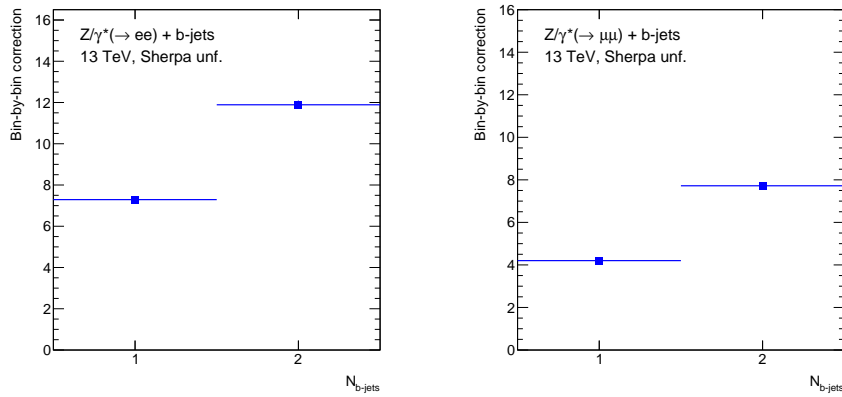
- flavour labelling on reconstructed jets.

Figure 6.7: Distributions of efficiency correction factors employed in the bin-by-bin method, in the electron (left) and muon (right) channels using SHERPA generator.

The integrated cross sections for Z+$\geq$1 b-jet and Z+$\geq$2 b-jets processes are evaluate as the content of the first and second bin, respectively, of the differential cross section as a function of the inclusive b-jet multiplicity. The differential cross sections as a function of the inclusive b-jet multiplicity are presented in Fig.6.8. Data are corrected with SHERPA and compared to the predictions from MADGRAPH and ALPGEN. The particle level measurements have a similar trend in the electron and muon channels.

The integrated cross section results are summarised in Table 6.2 and Table 6.3 for the Z+$\geq$1 b-jet and Z+$\geq$2 b-jets processes, respectively, and compared with the predictions from SHERPA, ALPGEN and MADGRAPH. Measured data are presented with both the statistical, systematic and luminosity uncertainties, being the systematic component evaluated as the sum
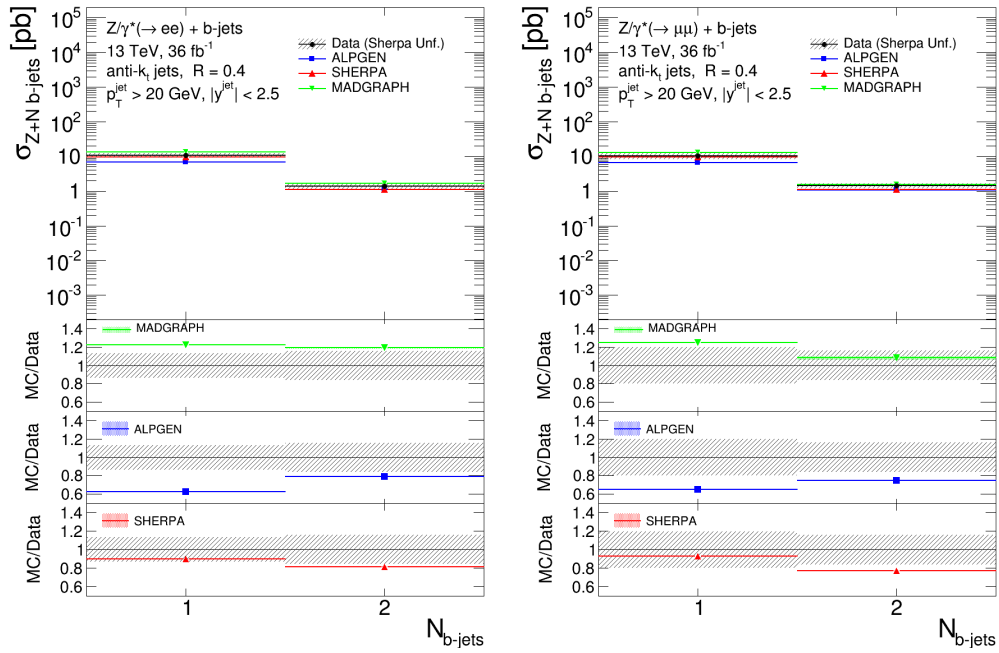
Figure 6.8: Differential cross section as a function of the b-jets inclusive multiplicity. Data are unfolded with SHERPA (red) and compared to MADGRAPH (green) and ALPGEN (blue) predictions. Results are shown in the electron (left) and muon (right) channels. The uncertainty (shaded band) includes all the reconstruction level systematics, with the exception of the luminosity uncertainty.

in quadrature of the components shown in Section 6.2. Monte Carlo calculations are shown with the statistical uncertainty only.

Z+≥ 1 b-jet cross sections		
Sample	Electron channel σ [pb]	Muon channel σ [pb]
	value ± stat ± syst ± lumi	value ± stat ± syst ± lumi
Data	10.8 ± 0.1 ± 1.5 ± 0.2	10.22 ± 0.08 ± 2.01 ± 0.21
Monte Carlo	value ± stat	value ± stat
SHERPA	9.72 ± 0.03	9.46 ± 0.03
MADGRAPH	13.27 ± 0.05	12.76 ± 0.07
ALPGEN	6.78 ± 0.03	6.64 ± 0.02

Table 6.2: Integrated cross section for Z boson production in association with at least 1 b-jet for the electron and muon channels. The measured value is provided with statistical, systematic and luminosity uncertainties. Monte Carlo predictions are reported with the statistical error only.

Z+≥ 2 b-jets cross sections		
Sample	Electron channel σ [pb]	Muon channel σ [pb]
	value ± stat ± syst ± lumi	value ± stat ± syst ± lumi
Data	1.37 ± 0.04 ± 0.21 ± 0.03	1.44 ± 0.03 ± 0.23 ± 0.03
Monte Carlo	value ± stat	value ± stat
SHERPA	1.117 ± 0.007	1.105 ± 0.007
MADGRAPH	1.635 ± 0.015	1.56 ± 0.03
ALPGEN	1.087 ± 0.011	1.076 ± 0.008

Table 6.3: Integrated cross section for Z boson production in association with at least 2 b-jets for the electron and muon channels. The measured value is provided with statistical, systematic and luminosity uncertainties. Monte Carlo predictions are reported with the statistical error only.

The measurement Z+≥1 b-jet cross section reaches a precision of 14%(20%) in the electron (muon) channel. The uncertainty is dominated by the exper-

imental systematics, whose main contributions are given in the breakdown plot (see Fig.6.9) and in Table 6.4. The results on the Z+≥1 b-jet cross section show the NLO SHERPA calculations (5FNS) are well comparable to the measured values. The LO MADGRAPH generator overestimates the measured value by ∼20%, but it is still in reasonable agreement with them within the errors. On the other hand the LO 4FNS ALPGEN calculation underestimates the measured cross section in data up to 40%.

The measurement of the cross section of the Z production in association with 2 b-jets has a comparable precision, with a total uncertainty of 16% in both the electron and muon channels. The result favours the the NLO SHERPA and LO MADGRAPH predictions derived with the 5FNS. The 4FNS ALPGEN calculation underestimates the cross section measurement by about 20%.

The results are in agreement with Run 1 analysis [35] performed at a centre of mass energy of 7 TeV for both Z+≥1 b-jet and Z+≥2 b-jets processes. Nevertheless, while in Run 1 data favoured the 4FNS and the 5FNS for the description of Z+≥1 b-jet and Z+≥2 b-jets, respectively, this does not seem to be confirmed from Run 2 results, where both processes are better modelled by the 5FNS. Differences among the various predictions appear mostly when comparing Matrix Element calculations performed at different orders in the perturbative expansion. As can be concluded looking at Tables 6.2-6.3 and Fig.6.8, the NLO SHERPA generator shows the best overall agreement with measurements.

6.4   Differential cross sections for Z+≥1 b-jet

All the measured particle level differential cross sections for the Z+≥1 b-jet process are shown from Fig.6.10 to Fig.6.15 for both electron and muon channels. The measured cross sections are shown as black points with statistical error only, while the overall uncertainties, defined as the sum in quadrature of all the components listed in Section 6.2, are reported as dashed bands. For comparison, the predictions by SHERPA, MADGRAPH and ALPGEN Monte Carlo generators are also shown with the statistical uncertainty only.

The leading b-jet p_T distributions are presented in Fig.6.10 for the electron and muon channels. All the Monte Carlo generators predict softer jets com-
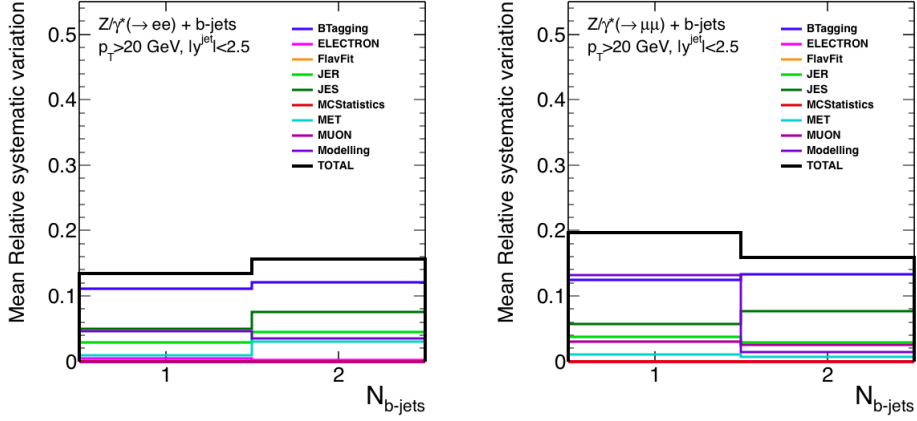
Figure 6.9: Breakdown of systematics uncertainty for the Z+≥1 b-jet and Z+≥2 b-jets cross section measurements for the electron (left) and muon (right) channels.

Summary of relative systematics uncertainties				
	Electron channel		Muon channel	
Systematics	Z+≥1 b-jet (%)	Z+≥2 b-jets (%)	Z+≥1 b-jet (%)	Z+≥2 b-jets (%)
BTagging	11.1	12.1	12.1	13.6
JES	5.0	7.5	5.5	7.7
JER	2.9	4.5	3.9	2.9
Modelling	6.5	4.5	15.3	6.8
Muon	-	-	3.2	2.5
Electron	0.4	0.2	-	-
MET	1.0	3.0	1.2	0.6

Table 6.4: Relative systematic uncertainties on the integrated Z+≥1 b-jet and Z+≥2 b-jets cross section measurements for the electron and muon channels.
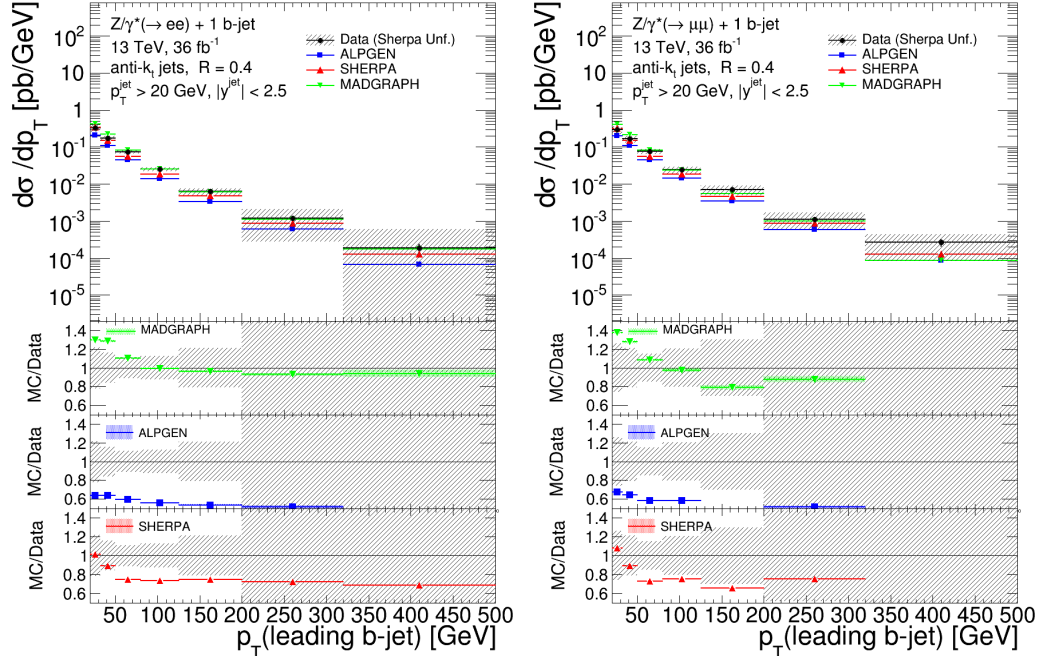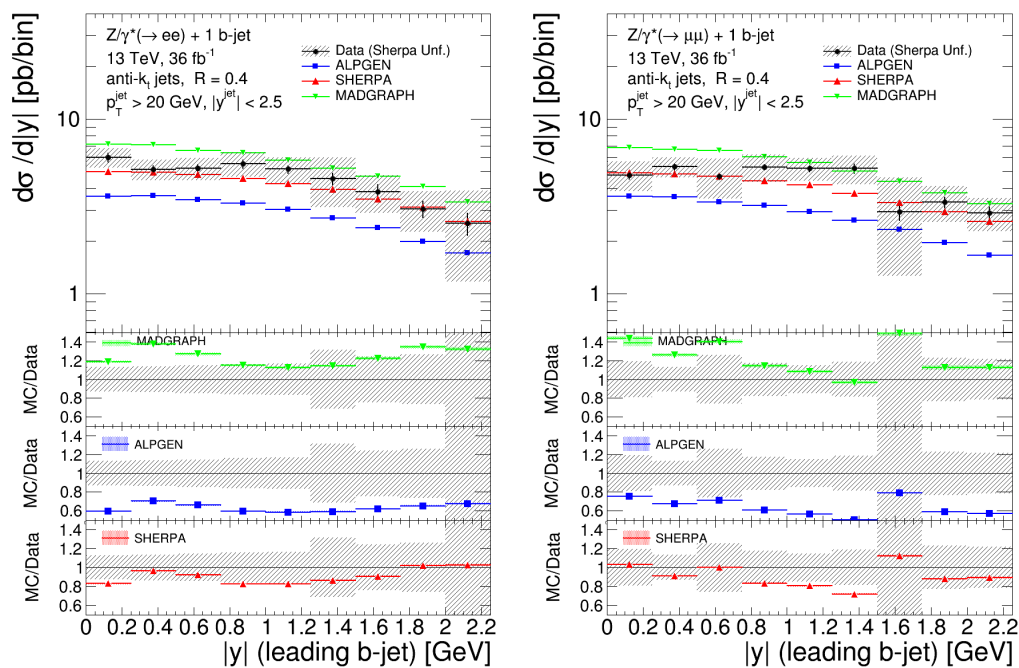
Figure 6.10: Differential cross sections for the production of a Z boson in associaton with at least 1 b-jet as a function of the leading b-jet p_T for the electron (left) and muon (right) channels. The data points (black) are presented with the total uncertianty (gray shaded band) and compared to predictions from SHERPA (red), ALPGEN (blue) and (MADGRAPH), which are represented with the statistical uncertainty only.

pared to data, but for p_T >50 GeV the 5FNS (SHERPA and MADGRAPH) seem to be more in agreement with data. The leading b-jet rapidity (see Fig.6.11) is well modelled by all the predictions, as one can notice from the MC/Data comparison, where, apart from an overall normalisation already discussed in Section 6.3, the agreement in the shape is similar.

In Fig.6.12 the differential cross section as a function of the Z boson y is presented. Even if this observable is well described by all Monte Carlos, the best agreement is achieved by the 5FNS SHERPA predictions.

The Z-b-jets observables are shown in Fig.6.13-6.15. Apart from an overall normalisation factor, the shapes of the $\Delta y(Z,b)$, $\Delta R(Z,b)$ and $\Delta\Phi(Z,b)$ are well modelled in all cases.

Figure 6.11: Differential cross sections for the production of a Z boson in asso-
ciaton with at least 1 b-jet as a function of the leading b-jet y for the electron
(left) and muon (right) channels. The data points (black) are presented with the
total uncertianty (gray shaded band) and compared to predictions from SHERPA
(red), ALPGEN (blue) and (MADGRAPH), which are represented with the statistical
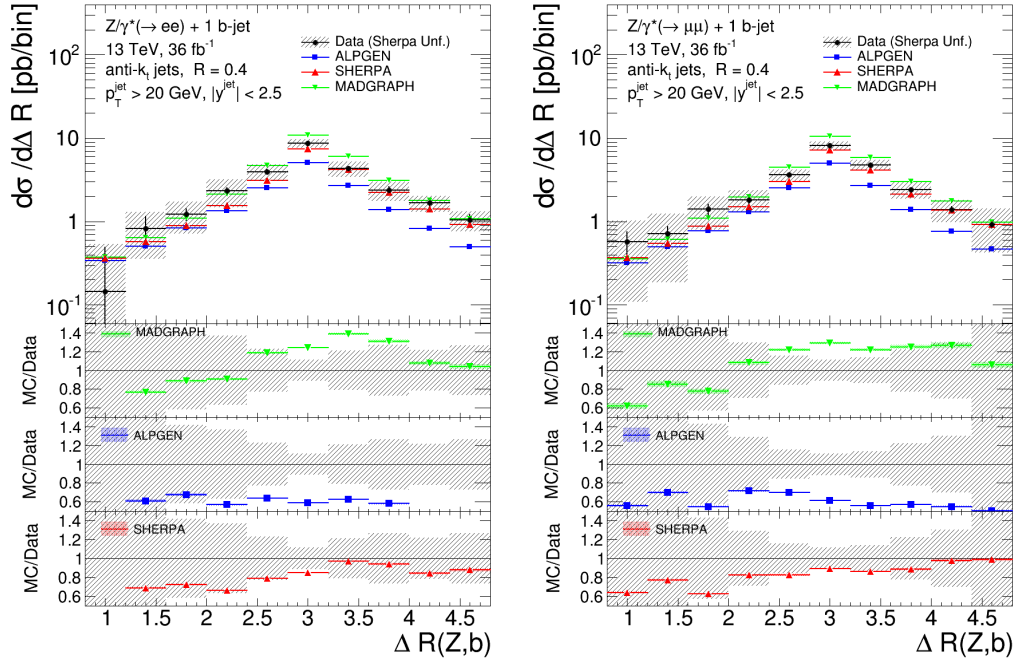uncertainty only.
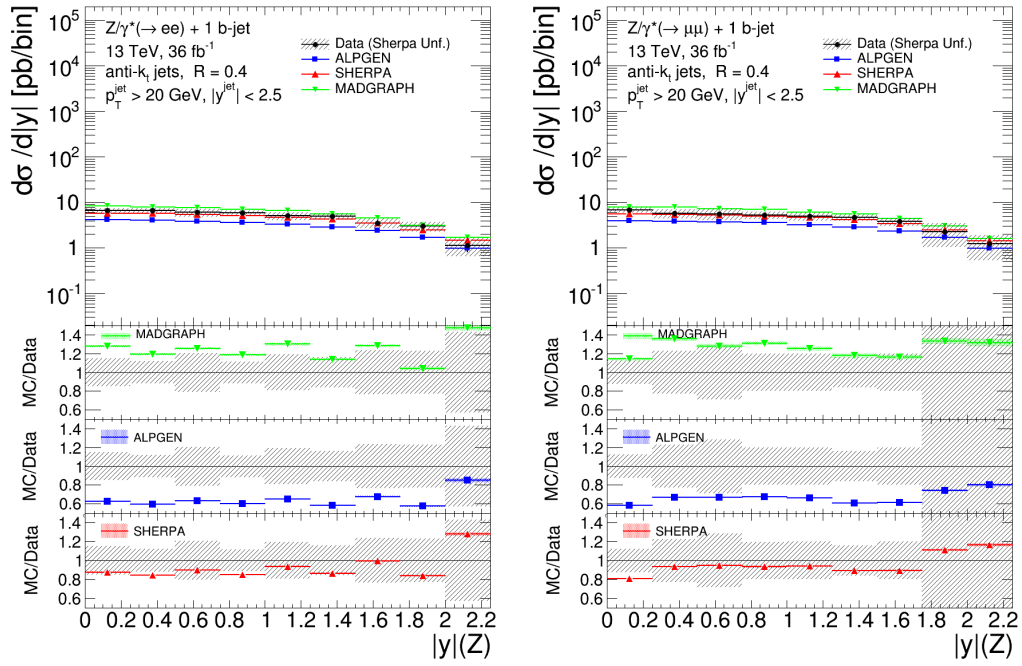
Figure 6.12: Differential cross sections for the production of a Z boson in association with at least 1 b-jet as a function of the Z y for the electron (left) and muon (right) channels. The data points (black) are presented with the total uncertianty (gray shaded band) and compared to predictions from SHERPA (red), ALPGEN (blue) and (MADGRAPH), which are represented with the statistical uncertianty only.
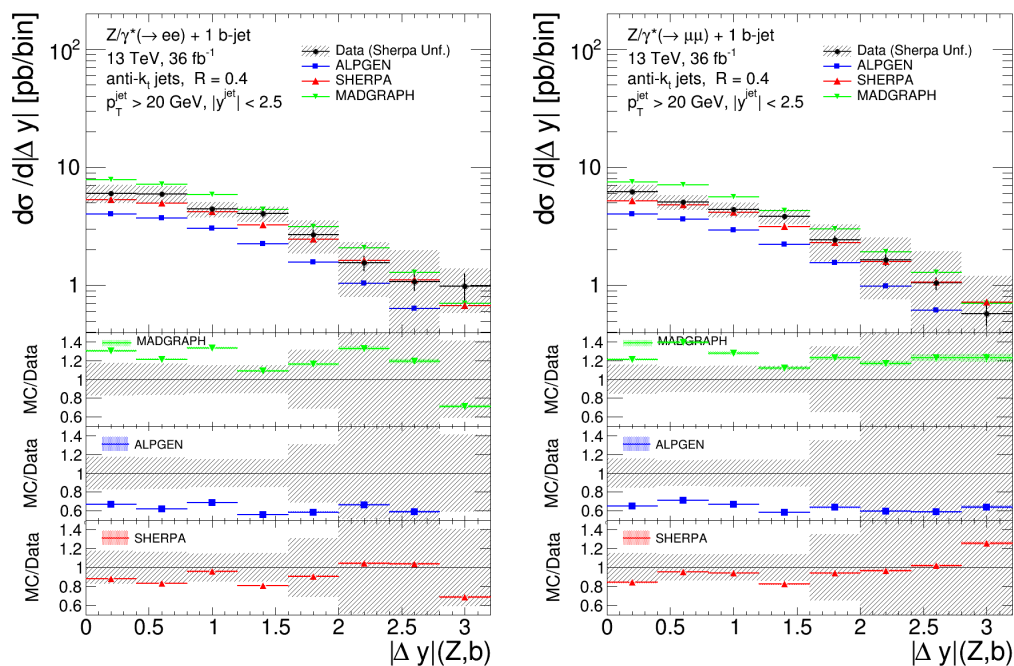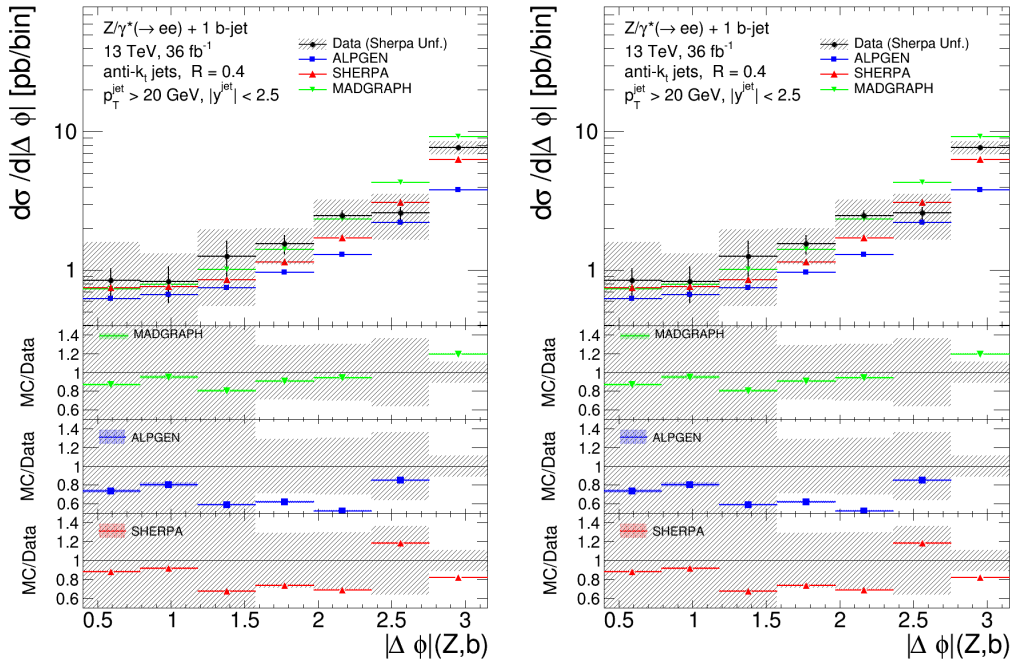
Figure 6.13: Differential cross sections for the production of a Z boson in association with at least 1 b-jet as a function of the Δy(Z,b) for the electron (left) and muon (right) channels. The data points (black) are presented with the total uncertianty (gray shaded band) and compared to predictions from SHERPA (red), ALPGEN (blue) and (MADGRAPH), which are represented with the statistical uncertainty only.

Figure 6.14: Differential cross sections for the production of a Z boson in association with at least 1 b-jet as a function of the $\Delta R(Z,b)$ for the electron (left) and muon (right) channels. The data points (black) are presented with the total uncertianty (gray shaded band) and compared to predictions from SHERPA (red), ALPGEN (blue) and (MADGRAPH), which are represented with the statistical uncertainty only.

Figure 6.15: Differential cross sections for the production of a Z boson in associaton with at least 1 b-jet as a function of the $\Delta\Phi$(Z,b) for the electron (left) and muon (right) channels. The data points (black) are presented with the total uncertianty (gray shaded band) and compared to predictions from SHERPA (red), ALPGEN (blue) and (MADGRAPH), which are represented with the statistical uncertainty only.

6.5 Differential cross sections for Z+≥2 b-jets

Particle level differential cross sections for the Z+≥2 b-jets analysis are shown from Fig.6.16 to Fig.6.21 for both electron and muon channels. The measured cross sections are shown as black points with statistical error only, while the overall uncertainties are reported as dashed bands (sum in quadrature of all the systematic components listed in Section 6.2). For comparison, the predictions by SHERPA, MADGRAPH and ALPGEN Monte Carlo generators are also shown with the statistical uncertainty only.

As for the leading b-jet, the p_T of the sub-leading b-jet is predicted softer by all Monte Carlo generators (see Fig.6.16), with similar shapes, and differences increase for p_T >200 GeV. In Fig.6.17 the y of the sub-leading is reasonably well described by all calculations, taking into account the systematic uncertainties.

The b-jets invariant mass M_{bb} is presented in Fig.6.18. Both SHERPA and ALPGEN present a slightly softer spectrum than data, even if the level of agreement is within 20% additional to the overall scale. The ΔR_{bb} (see Fig.6.19) distribution shows an improvement with respect to Run 1 results in the low range. This is explained by the better modelling of gluon split in $b\bar{b}$ pairs in the low ΔR_{bb} range in theoretical predictions.

Fig.6.20 shows the cross sections as a function of the Δy_{bb}, where a difference in shape with respect to data is clearly visible for all generators. The $\Delta\Phi_{bb}$, presented in Fig.6.21, is modelled consistently by all Monte Carlo generators and in agreement with data.

6.6 Final remarks

The Run 1 measurements, performed at a centre of mass energy $\sqrt{s}$ =7 TeV, show clearly that data favours the 4FNS predictions in the description of Z+≥2 b-jets processes and the 5FNS calculations for the Z+≥1 b-jet analysis. This seems to be not confirmed at the centre of mass energy of $\sqrt{s}$ =13 TeV: the Z+≥1 b-jet integrated cross section in both electron and muon channels, is better modelled by the 5FNS, but in the differential measurements a clear preference of any of the two production modes cannot be solidly affirmed.
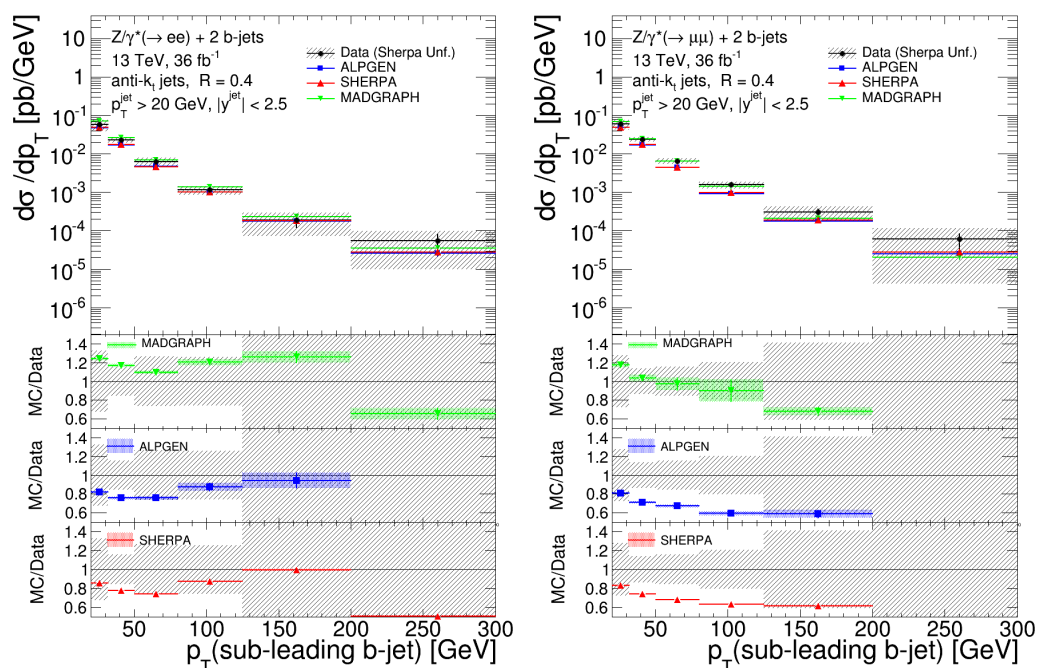
Figure 6.16: Differential cross sections for the production of a Z boson in associaton with at least 2 b-jets as a function of the sub-leading b-jet p_T for the electron (left) and muon (right) channels. The data points (black) are presented with the total uncertianty (gray shaded band) and compared to predictions from SHERPA (red), ALPGEN (blue) and (MADGRAPH), which are represented with the statistical uncertainty only.
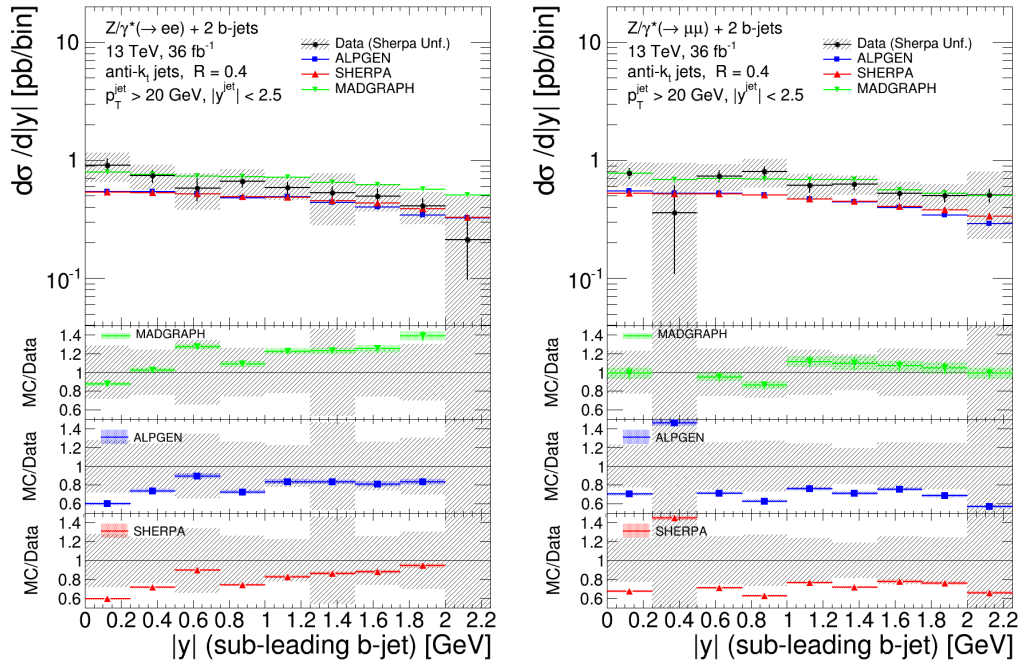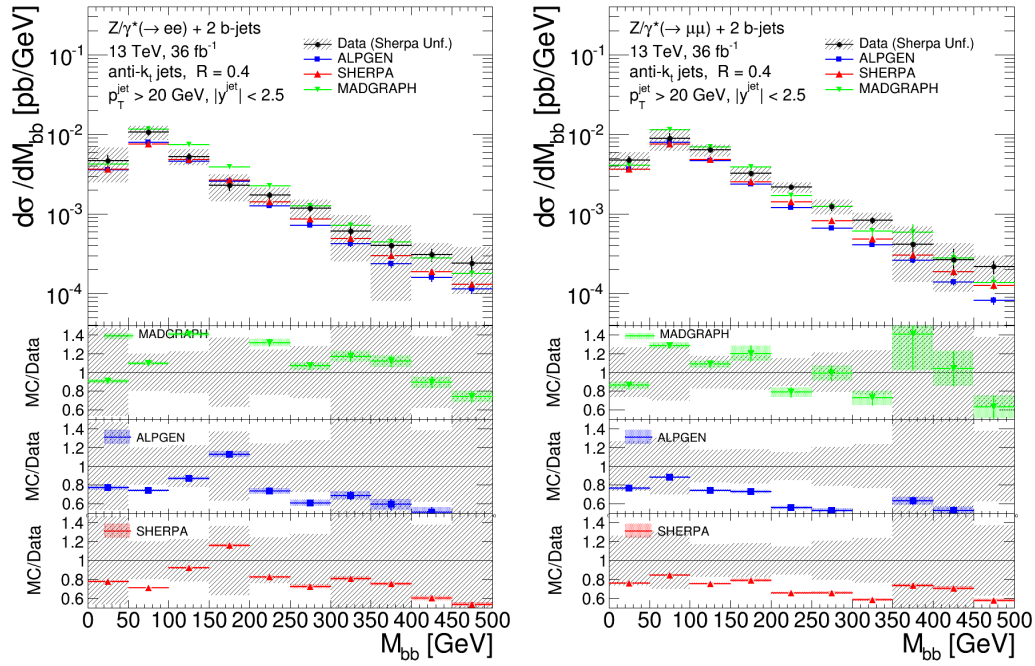
Figure 6.17: Differential cross sections for the production of a Z boson in assoiaton with at least 2 b-jets as a function of the second leading b-jet y for the electron (left) and muon (right) channels. The data points (black) are presented with the total uncertianty (gray shaded band) and compared to predictions from SHERPA (red), ALPGEN (blue) and (MADGRAPH), which are represented with the statistical uncertainty only.
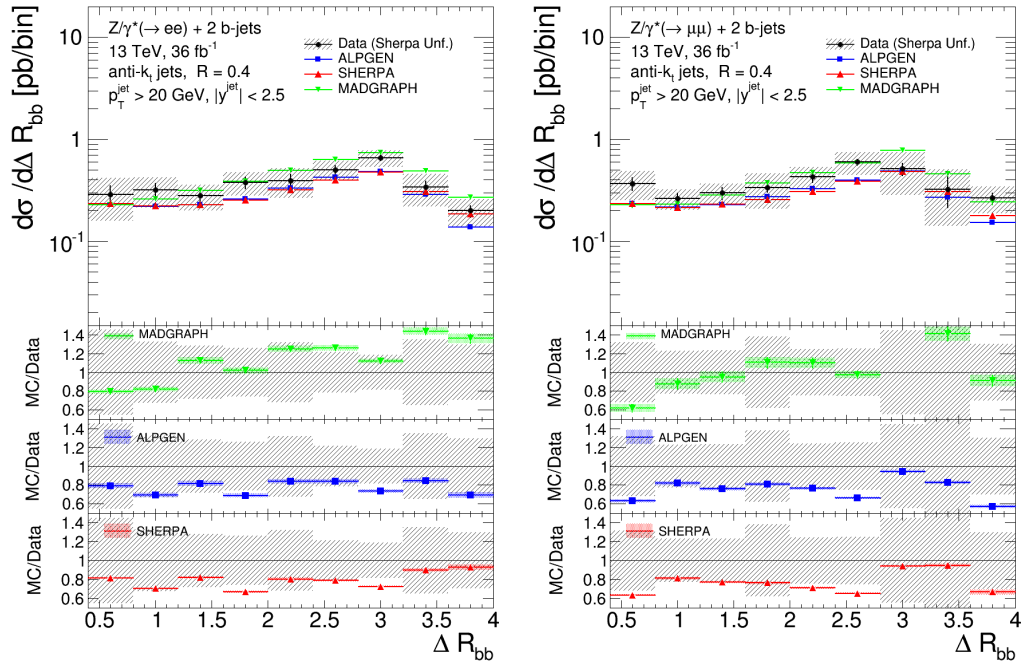
Figure 6.18: Differential cross sections for the production of a Z boson in associaton with at least 2 b-jets as a function of M_{bb} for the electron (left) and muon (right) channels. The data points (black) are presented with the total uncertianty (gray shaded band) and compared to predictions from SHERPA (red), ALPGEN (blue) and (MADGRAPH), which are represented with the statistical uncertainty only.
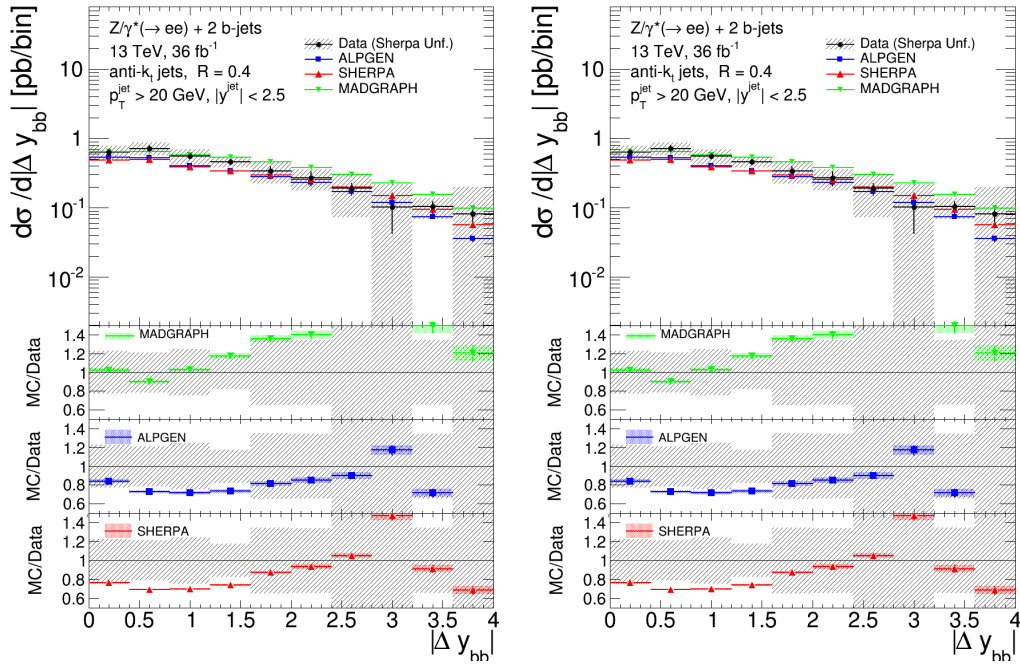
Figure 6.19: Differential cross sections for the production of a Z boson in associaton with at least 2 b-jets as a function of ΔR_{bb} for the electron (left) and muon (right) channels. The data points (black) are presented with the total uncertianty (gray shaded band) and compared to predictions from SHERPA (red), ALPGEN (blue) and (MADGRAPH), which are represented with the statistical uncertainty only.

Figure 6.20: Differential cross sections for the production of a Z boson in associaton with at least 1 b-jet as a function of the Δy_{bb} for the electron (left) and muon (right) channels. The data points (black) are presented with the total uncertianty (gray shaded band) and compared to predictions from SHERPA (red), ALPGEN (blue) and (MADGRAPH), which are represented with the statistical uncertainty only.
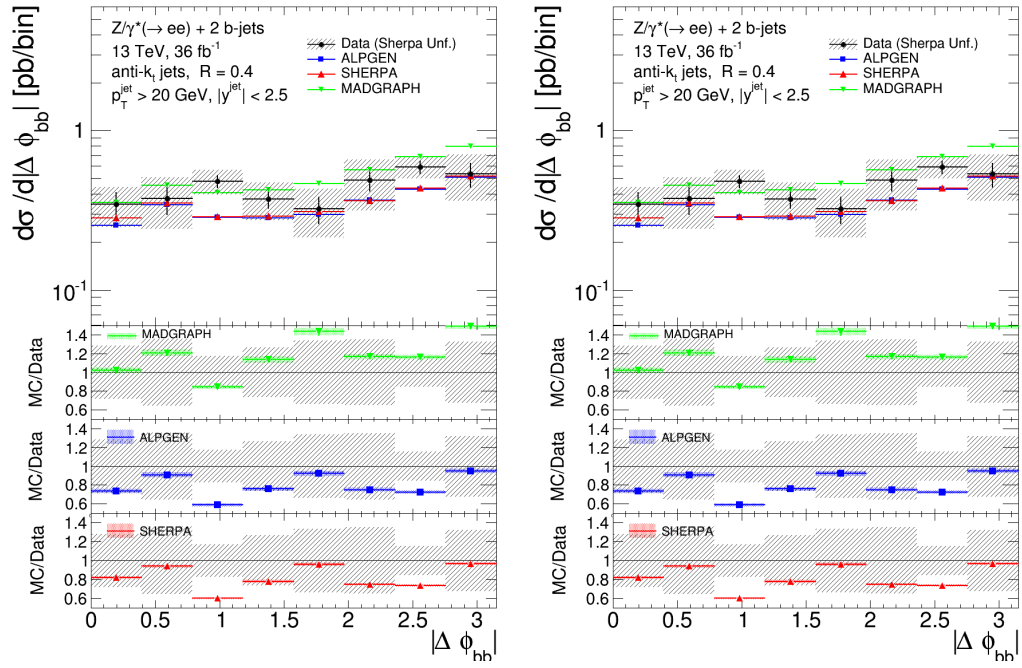
Figure 6.21: Differential cross sections for the production of a Z boson in associa-ton with at least 1 b-jet as a function of the $\Delta\Phi_{bb}$ for the electron (left) and muon (right) channels. The data points (black) are presented with the total uncertianty (gray shaded band) and compared to predictions from Sherpa (red), Alpgen (blue) and (Madgraph), which are represented with the statistical uncertainty only.

In general, the NLO SHERPA calculations improved the agreement on the overall scale in the Z+$\geq$1 b-jet cross section with respect to Run 1, while a more modest impact can be noticed for the Z+$\geq$2 b-jets measurements.
The differential cross sections are in general in good agreement with data, within the systematics uncertainties. The only exception is represented by the leading and sub-leading b-jet p_T, where independently of the flavour number scheme and the precision order of the matrix element computation, SHERPA, ALPGEN and MADGRAPH predict a softer spectrum with respect to data.

6.7   Future perspectives

The measurement of the Z+b-jets production is performed with data collected by ATLAS at the centre of mass energy of $\sqrt{s}$ =13 TeV and corresponding to a luminosity of 36 fb^{-1}. The full analysis is presented in this thesis, starting in Chapter 4 from the event selection and the background rejection criteria, with a detailed study of the Z+c-jets and Z+light-jets components. From the detector level, the measurements are extrapolated to particle level through the unfolding procedure in Chapter 5. The results of the integrated and differential cross sections of the Z+$\geq$1 b-jet and Z+2 b-jets processes are presented in this Chapter and compared to several Monte Carlo predictions, using both the 4FNS and the 5FNS.
In view of the publication of the measurement, few further steps are to be done:

- **Lepton channel combination**
  Measurements are performed separately in the electron and muon channels, obtaining a good compatibility not only in the integrated results, but also in the shape of the differential cross sections. The combination will benefit from the higher statistics and will positively impact the uncertainty at the edges of some experimental distributions, dominated by statistical uncertainty. For the combination a detailed study of correleated and uncorrelated uncertainties must be done, as in the case of the inclusive cross section measurement (see Chapter 3).

- **b-tagging uncertainty**
  A new calibration method is foreseen from January 2018, which will provide higher precision in the b-tagging measurements, needed to better identify and reject the c- and light-jets. The new calibration will help in the discrimination of the Z+c-jets and Z+light jets through the flavour fit, since the discriminating variable used to extract their normalisation is the b-tagging weight (MV2c10).

- **Modelling uncertainty**
  The modelling has the largest impact on the total systematics of the final results. In this thesis, it corresponds to the difference in the final results obtained with the nominal generator (SHERPA) or repeating the full analysis chain with another one with similar statistics (ALPGEN). The modelling in this case involves uncertainties on both the flavour fit and the unfolding procedure. This is a conservative estimate and this strategy will be used in the future only for the uncertainty associated to the template shapes in the flavour fit, together with a treatment of correlations of reconstruction level uncertainties. To account for the different descriptions of Monte Carlo generators, a reweighting procedure will be applied to SHERPA distributions according to differences with ALPGEN.

- **Fixed order theoretical predictions**
  In this thesis the results are compared to predictions from the NLO 5FNS SHERPA, the LO 5FNS MADGRAPH and the LO 4FNS ALPGEN. In the next future, fixed NLO theoretical calculations will be available at both 4FNS and 5FNS, in order to deeply investigate differences, to be attributed to the presence of the b-quark in the initial state.

Conclusions

In this thesis the measurement of the Z boson cross section performed with the early Run 2 data collected by the ATLAS detector at a centre of mass energy of $\sqrt{s}$=13 TeV is presented. The cross section is performed in the electron and muon decay channels, where the measurement reaches a precision comparable with several NNLO QCD calculations. The measured fiducial and total cross sections are combined taking into account the correlated and uncorrelated uncertainties:

$$\sigma^Z_{fid} = 778.6 \pm 2.8(\text{stat}) \pm 5.6(\text{syst}) \pm 16.4(\text{lumi})\text{pb}$$
$$\sigma^Z_{tot} = 1981.2 \pm 7.0(\text{stat}) \pm 38.1(\text{syst}) \pm 41.6(\text{lumi})\text{pb}$$

The ratio of the W to Z cross sections benefits from the partial (lepton related) or total (luminosity) cancellation of several systematic uncertainties and reaches a precision of <1% (excluding a 5% of luminosity uncertainty):

$$\sigma_{W^\pm}/\sigma_Z = 10.31 \pm 0.04(\text{stat}) \pm 0.20(\text{syst})$$

In the future, theoretical calculations will benefit from the high precision reached in the ratio measurement, to constrain NNLO PDFs and Monte Carlo modelling. In this context, I personally focused on the measurement in muon channel, including the event selection, the background estimation, the efficiency determination and the evaluation of the systematic uncertainties.

In this thesis, the first measurement of the production of the Z boson produced in association to b-jets at the centre of mass energy of $\sqrt{s}$ =13 TeV is also presented. A dataset corresponding to 36 fb^{-1} collected by the ATLAS

detector is used to measure the integrated and differential cross sections of the Z+≥1 b-jet and Z+≥2 b-jets processes in the electron and muon decay channels.

The Z+≥1 b-jet production is studied as a function of the leading b-jet kinematics, by measuring the transverse momentum and the rapidity, of primary importance to test QCD predictions and the modelling of Monte Carlo generators. The Z boson rapidity and the difference in rapidity to the leading b-jet are studied in detail, given the crucial sensitivity of these distributions to PDFs. The difference in the azimuthal and radial coordinates between the Z boson and the leading b-jet are also measured.

The Z+≥2 b-jets differential cross sections are measured as a function of the kinematics of the b-jet pairs, like the invariant mass of the $b\bar{b}$ system (M_{bb}) and the difference in the rapidity (Δy_{bb}), azimuthal ($\Delta\Phi_{bb}$) and radial (ΔR_{bb}) coordinates. The modelling of these physics observables is crucial for Higgs studies, where, for example, in the $b\bar{b}$ decay channel, it represents one of the main uncertainties contributing to the measurement. Ancillary measurements of the sub-leading kinematics are also presented in the same phase space, as a cross check of the overall consistency of the analysis.

The differential cross section measurements reach an average precision of 30%, with the exception of the edge of some distributions, characterised by a reduced statistics. The dominant systematic uncertainty is due to the modelling of the Monte Carlo generators, which has an impact in the Z+c-jets and Z+light-jets background determination and in the extrapolation of the cross sections from detector to particle level.

The measured values of the integrated cross sections of the Z+≥1 b-jet and Z+≥2 b-jets processes in the electron and muon channels are

$$\sigma_{Z(\to ee)+\geq 1\mathrm{b-jet}} = 10.8 \pm 0.1(\mathrm{stat}) \pm 1.5(\mathrm{syst}) \pm 0.2(\mathrm{lumi})\mathrm{pb}$$

$$\sigma_{Z(\to ee)+\geq 2\mathrm{b-jets}} = 1.37 \pm 0.04(\mathrm{stat}) \pm 0.21(\mathrm{syst}) \pm 0.03(\mathrm{lumi})\mathrm{pb}$$

$$\sigma_{Z(\to \mu\mu)+\geq 1\mathrm{b-jet}} = 10.22 \pm 0.08(\mathrm{stat}) \pm 2.01(\mathrm{syst}) \pm 0.21(\mathrm{lumi})\mathrm{pb}$$

$$\sigma_{Z(\to \mu\mu)+\geq 2\mathrm{b-jets}} = 1.44 \pm 0.03(\mathrm{stat}) \pm 0.23(\mathrm{syst}) \pm 0.03(\mathrm{lumi})\mathrm{pb}$$

and reach an average precision of 16%, dominated by the uncertainties on the b-tagging and the Monte Carlo modelling. Cross section results are in

good agreement in the two lepton decay channels. Integrated and differential cross sections are compared to several NLO and LO Monte Carlo generators, employing two different generation mechanisms (the 4FNS or the 5FNS).

In the past, measurements of the Z+b-jets processes were performed by AT-LAS at a centre of mass energy of $\sqrt{s} =7$ TeV. The results of these previous studies enhanced that the data favour the 5FNS in the description of Z+$\geq$2 b-jets production and the 4FNS in the modelling of the Z+$\geq$1 b-jet processes. This seem not to be confirmed by the results at the centre of mass energy of $\sqrt{s} =13$ TeV, where in general the combination of NLO+5FNS provides the best agreement with data for the description of both Z+$\geq$1 b-jet and Z+$\geq$2 b-jets processes. An improvement with respect to Run 1 results is clearly visible in the low range of the ΔR_{bb} observable, where gluon splitting starts to dominate: the underestimation observed in all predictions studied in Run 1 seem to be cured at $\sqrt{s} =13$ TeV.

The Z+b-jets production offers a clean experimental probe for many aspects of the strong interaction phenomenology at LHC. An improvement in the accuracy of the theoretical predictions is worthwhile for further investigation. In particular, it is worth noticing that the long debate on the differences between the 4FNS and the 5FNS implementation of perturbative QCD predictions for processes with b-quark in the initial state can find in accurate Z+b-jets measurements a useful ground for testing.

In perspective, the full data sample collected by the ATLAS experiment during Run 2 is expected to bring an improvement of the experimental statistical uncertainties by more than a factor two compared to current precision. Moreover in view of the publication of the measurement the treatment of some experimental uncertainties will improve, due mostly to the Monte Carlo modelling and the new b-tagging calibration precisions.

I have been deeply involved in the Z+$\geq$1 b-jet and Z+$\geq$2 b-jets cross section measurements, following all the aspects of the study, both in the muon and electron channels. In particular, I focused on the event selection, the background estimation through the flavour fit, the unfolding procedure and the evaluation of the systematic uncertainties.

During the long shut-down between the Run 1 and Run 2 of LHC, the sensors of LUCID, the ATLAS official luminosity monitor, have been replaced and

equipped with a radioactive Bismuth-207 source. As a technical activity, I directly participated to the choice and the equalisation of LUCID photomultipliers.

Finally I participated to studies of the ATLAS muon reconstruction performances aimed to the evaluation of the data-driven correction factors ("Scale Factors") needed to tune Monte Carlo simulations.

Appendix A

LUCID: the ATLAS Luminosity Detector

A.1 Introduction

Luminosity is a key quantity for any physical measurement, since it relates the rate of a particular process to its cross section. An overview of luminosity concept is provided in Section A.2, followed in Section A.3 by a description of the methods used in ATLAS to measure the luminosity. The general description of ATLAS luminometers is presented in Section A.4. The details of the dedicated luminometer LUCID (Luminosity measurement Using Cherenkov Integrating Detector) are presented in Section A.5. The final 2016 performances will be presented in Section A.6.

A.2 Luminosity overview

The *instantaneous luminosity* $\mathcal{L}$ is defined as the ratio between the rate of any process (R) and its cross section (σ). It is expressed in units of cm^{-2}s^{-1} and it is independent of the process itself.

$$\mathcal{L} = \frac{R}{\sigma} \tag{A.1}$$

The instantaneous luminosity can be inferred from the machine parameters: if the two beams are made of identical bunches, these are Gaussian in shape and perfectly overlapping without crossing angle, then the luminosity is given by:

$$\mathcal{L} = f_r n_b \frac{N_1 N_2}{4\pi \sigma_x \sigma_y} \tag{A.2}$$

where f_r is the beam revolution frequency, n_b is the number of bunches in each beam, $N_{1,2}$ are the number of protons in each beam, $\sigma_{x,y}$ are the gaussian transverse profile of the beams. At LHC the instantaneous luminosity is expected to decrease with the exponential law $\mathcal{L} = \mathcal{L}_0 e^{\frac{-t}{\tau}}$, with $\tau \cong 14$ h.

The *integrated luminosity* L is obtained by integrating the instantaneous luminosity over a certain time interval t and is expressed in units of cm^{-2}:

$$L = \int_0^t \mathcal{L}(t') dt' \tag{A.3}$$

Due to the degradation in time, the integrated luminosity is evaluated in short periods, called Luminosity Blocks (LB), in which it can be considered constant.

A.3 Luminosity measurements in ATLAS

The luminosity of a p-p collider can be expressed as

$$\mathcal{L} = \frac{R_{inel}}{\sigma_{inel}} \tag{A.4}$$

where R_{inel} is the rate of inelastic collisions and σ_{inel} is the p-p inelastic cross section. For a storage ring, operating at a revolution frequency f_r and with n_b bunch pairs colliding per revolution, this expression can be rewritten as

$$\mathcal{L} = \frac{\mu f_r n_b}{\sigma_{inel}} \tag{A.5}$$

where μ is the average number of inelastic interactions per bunch crossing. The luminosity detectors can provide only an experimental quantity which is proportional to the luminosity. In ATLAS all the luminometers can only

measure the average number of visible interactions per bunch crossing (μ_{vis}). The μ_{vis} is related to the real μ value by the acceptance and efficiency of the detector: $\mu_{vis} = \mu\epsilon$. The luminosity can thus be rewritten as:

$$\mathcal{L} = \frac{\mu f_r n_b}{\sigma_{inel}} = \frac{\mu_{vis} f_r n_b}{\epsilon \sigma_{inel}} = \frac{\mu_{vis} f_r n_b}{\sigma_{vis}} \tag{A.6}$$

where σ_{vis} represents the calibration constant necessary to pass from a relative to an absolute luminosity value.

At LHC the absolute luminosity scale of each luminosity detector is determined by means of dedicated runs, called *Van der Meer* (vdM) scans, which allow to infer the delivered luminosity at one point in time from the measurable parameters of the colliding bunches. The method consists in moving the beams transversely with respect to each other while recording the counting rate of at least one luminosity monitor. Separation scans are performed in both the vertical and horizontal directions. The observed rate is recorded while scanning and from this measurement two bell-shaped curves are obtained, with the maximum rate at zero separation (see Fig.A.1). From the area under the curves the values $\sigma_{x,y}$ are obtained and the luminosity at zero separation is inferred from Eq.A.2. The values of σ_{vis} is finally extracted from Eq.A.3 using the measured values of the luminosity and μ_{vis}. To achieve the desired accuracy on the absolute luminosity, these scans are not performed during normal physics operations, but rather under carefully controlled conditions with a limited number of colliding bunches and a modest peak interaction rate ($\mu \leq 2$).

A.4   The ATLAS luminosity detectors

The luminosity detectors in ATLAS are designed to reach three goals:

- Providing final absolute integrated luminosity values for offline analyses, for the full data sample as well as for selected periods, meaning that the luminosity must be provided for each LB. In physics analyses, in fact, data are used only if some quality criteria, provided LB by LB, are satisfied. To avoid discarding too many data, short LB are needed.
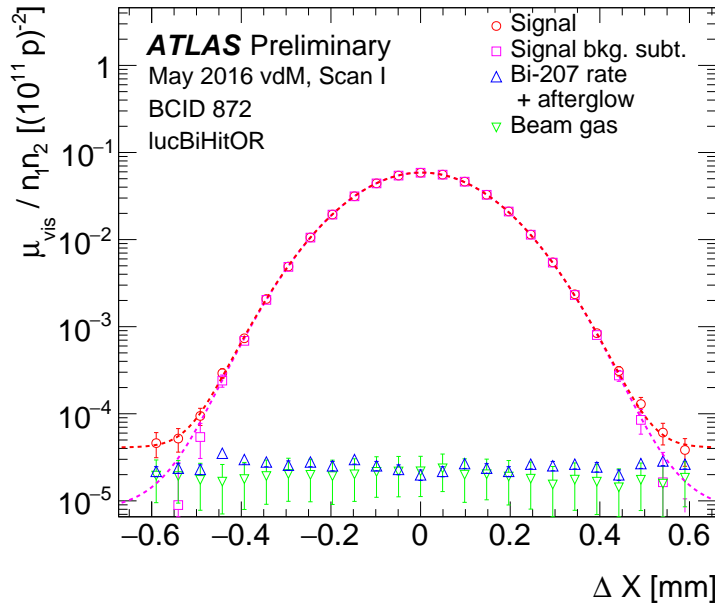
Figure A.1: Visible interaction rate per bunch crossing and per unit bunch-population product, for LUCID Hit OR algorithm versus nominal beam separation during horizontal scan 1 in the May 2016 luminosity-calibration session. The total rate measured for a single colliding-bunch pair at position 872 in the fill pattern is shown as red circles, and the background-subtracted rate as magenta squares. The background (blue triangles) is estimated from the rate measured in the preceding unfilled bunch slot. The beam-gas background (green triangles) measured using non-colliding bunches is also shown. The background-subtracted rate is fitted by a Gaussian multiplied by a sixth-order polynomial (dashed curve). The error bars are statistical only.

Typical values of LB length are of the order of 1-2 minutes. Each LB is identify by a number which uniquely tags it within a run.

- Providing fast online luminosity monitoring (1-2 seconds) to LHC, as required for efficient beam steering and machine optimization, as for example beam centering through mini-scans. The fast luminosity measurement is also used to efficiently tune the trigger. The fraction of recorded data, called *prescale*, in fact, can be changed according to the beam degradation, in order to optimize at each time the data acquisition band width.

- Fast checking of running conditions such as monitoring the structure of the beam and beam-related backgrounds.

Since there is no single experimental technique fulfilling all the above requirements, a number of complementary measurements (algorithms) and detectors have to be considered: parallel measurements of absolute and relative luminosity are mandatory. Each detector is characterised by different acceptance, response to pileup, systematic uncertainties and background sources. The redundancy of measurements guarantees the control over the systematic uncertainties and the possibility to measure the luminosity in every LHC running condition. In Fig. A.2 the main luminosity detectors in ATLAS are shown.
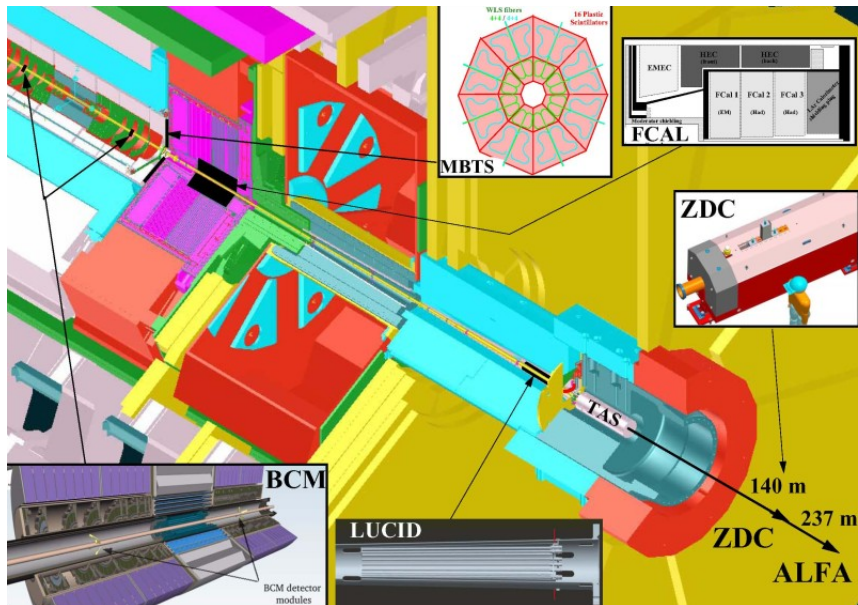
Figure A.2: ATLAS main luminometers.

In the following the subsystems used for luminosity determination are described in order of increasing pseudorapidity coverage.

Inner Detector

The Inner Detector (ID) is used to measure the momentum of charged particles over a pseudorapidity interval of $|\eta| < 2.5$, as explained in detail in

Chapter 2. Given the high efficiency in tacking particles (from 10% for p_T at 100 MeV to around 86% for p_T above a few GeV), the main application of the ID for luminosity purpose is to count the primary vertices and the tracks produced in inelastic p-p collisions during a LB.

MBTS

For the initial running period at low instantaneous luminosity ($<10^{33}$cm^{-2}s^{-1}), ATLAS has been equipped with segmented scintillator counters, the Minimum Bias Trigger Scintillators (MBTS), located at z = ±365 cm from the interaction point (IP). It covers the range 2.09< $|\eta|$ <3.84. The main purpose of MBTS detector is to provide a trigger on the minimum collision activity during a p-p bunch crossing (LB basis). Light emitted by scintillators is collected by wave length-shifting optical fibres and guided to a photomultiplier tube (PMT). The MBTS signal, after being shaped and amplified, are fed into leading-edge discriminators and send to the Central Trigger Processor (CTP) for trigger purposes.

BCM

The primary purpose of the Beam Condition Monitor (BCM) is to monitor beam losses and provide fast feedback to the accelerator operations team. It is an essential ingredient of the detector protection system, providing a fast accelerator abort signal in case of large beam loss. BCM consists of four small diamond sensors arranged around the beam-pipe on each side of the IP, at a distance of z = ±184 cm, covering an area of $|\eta|$=4.2. The horizontal and vertical pairs of the detector are read out separately, leading to two independent luminosity measurements, on LB basis.

LUCID

LUCID is a Cherenkov detector specifically designed to measure luminosity in the pseudorapidity range 5.6< $|\eta|$ <6.0. A more complete description of LUCID is presented in Section A.5. It is the only detector that can provide luminosity measurement both on bunch-by-bunch basis and on LB basis.

Calorimeters

In addition to the detectors listed above, further luminosity-sensitive methods have been developed using components of the ATLAS calorimeter system. These techniques do not identify particular events, but rather measure average particle rates over longer time scales. The Tile Calorimeter (TileCal) is the central hadronic calorimeter of ATLAS. It is a sampling calorimeter constructed from iron plates (absorber) and plastic tile scintillators (active material) covering the pseudorapidity range $|\eta| <1.7$. The complete description of the calorimeter is presented in Chapter 2. The current drawn by the readout PMTs is proportional to the total number of particles interacting in a given TileCal cell and provides a signal proportional to the total luminosity summed over all the colliding bunches present at a given time.

The Forward Calorimeter (FCal) is a sampling calorimeter that covers the pseudorapidity range $3.2< |\eta| <4.9$. Each of the two end-cap modules is divided into three longitudinal absorber matrices, one made of copper (FCal-1) and the other two of tungsten (FCal-2/3). Each matrix contains tubes arranged parallel to the beam axis filled with liquid argon as the active medium. Each FCal-1 matrix is divided into 16 ϕ-sectors, each of them fed by four independent high-voltage lines. The high voltage on each sector is regulated to provide a stable electric field across the liquid argon gaps and, similar to the TileCal PMT ones, the currents provided by the FCal-1 high-voltage system are directly proportional to the average rate of particles interacting in a given FCal sector.

Both calorimeter luminosity measurements are made on LB basis.

A.4.1 ATLAS luminosity algorithms

ATLAS primarily uses *Event Counting* algorithms to measure luminosity, where a bunch crossing is said to contain an "event" if the criteria for a given algorithm to observe one or more interactions are satisfied. The two main algorithm types being used are EventOR (inclusive counting) and EventAND (coincidence counting). Additional algorithms have been developed using *Hit Counting* and average *Particle Rate Counting*, which provide a cross-check of the linearity of the event counting techniques.

Interaction rate determination

Most of the primary luminosity detectors consist of two symmetric detector elements placed in the forward ("A") and backward ("C") direction from the interaction point. For *Event Counting* algorithms, a threshold is applied to the analogue signal output from each readout channel, and every channel with a response above this threshold is counted as containing a "hit". In an EventOR algorithm, a bunch crossing is counted if there is at least one hit on either the A side or the C side. Assuming that the number of interactions in a bunch crossing can be described by a Poisson distribution, the probability of observing an OR event can be computed as:

$$P_{EventOR}(\mu_{vis}^{OR}) = \frac{N_{OR}}{N_{BC}} = 1 - e^{-\mu_{vis}^{OR}} \tag{A.7}$$

where N_{OR} is the number of bunch crossings, during a given time interval, in which at least one p-p interaction satisfies the event-selection criteria of the OR algorithm under consideration, and N_{BC} is the total number of bunch crossings during the same interval. Solving for μ_{vis} in terms of the event counting rate yields:

$$\mu_{vis}^{OR} = -ln\left(1 - \frac{N_{OR}}{N_{BC}}\right) \tag{A.8}$$

In case of an EventAND algorithm, a bunch crossing is counted if there is at least one hit on both sides of the detector. This coincidence condition can be satisfied either from a single p-p interaction or from individual hits on either side of the detector from different p-p interactions in the same bunch crossing. Assuming equal acceptance for sides A and C, the probability of recording an AND event can be expressed as:

$$P_{EventAND}(\mu_{vis}^{AND}) = \frac{N_{AND}}{N_{BC}} = 1 - 2e^{-(1+\sigma_{vis}^{OR}/\sigma_{vis}^{AND})\mu_{vis}^{AND}/2} + e^{-(\sigma_{vis}^{OR}/\sigma_{vis}^{AND})\mu_{vis}^{AND}}$$
$$\tag{A.9}$$

This relationship cannot be inverted analytically to determine μ_{vis}^{AND} as a function of N_{AND} and N_{BC}, so a numerical inversion is performed instead. When $\mu_{vis} \gg 1$, event counting algorithms loose sensitivity as fewer and fewer events in a given time interval have bunch crossings with zero observed interactions. In the limit where $N/N_{BC} = 1$, it is no longer possible to use

event counting to determine the interaction rate μ_{vis} and more sophisticated techniques must be used. One example is a *Hit Counting* algorithm, where the number of hits in a given detector is counted rather than just the total number of events. This provides more information about the interaction rate per event and increases the luminosity at which the algorithm saturates.

Under the assumption that the number of hits in one p-p interaction follows a Binomial distribution and that the number of interactions per bunch crossing follows a Poisson distribution, the average probability to have a hit in one of the detector channels per bunch crossing can be calculated as:

$$P_{HITOR}(\mu_{vis}^{HIT}) = \frac{N_{HIT}}{N_{BC}N_{CH}} = 1 - e^{\mu_{vis}^{HIT}} \tag{A.10}$$

where N_{HIT} and N_{BC} are the total numbers of hits and bunch crossings, respectively, during a time interval, and N_{CH} is the number of detector channels. The expression above enables μ_{vis}^{HIT} to be calculated from the number of hits as:

$$\mu_{vis}^{HIT} = -ln\left(1 - \frac{N_{HIT}}{N_{BC}N_{CH}}\right) \tag{A.11}$$

An additional type of algorithm that can be used is a *Particle Counting* algorithm, where some observable is directly proportional to the number of particles interacting in the detector. These should be the most linear of all of the algorithm types, and in principle the interaction rate is directly proportional to the particle rate. The TileCal and FCal current measurements are not exactly particle counting algorithms, as individual particles are not counted, but the measured currents should be directly proportional to luminosity. Similarly, the number of primary vertices is directly proportional to the luminosity, although the vertex reconstruction efficiency is significantly affected by pileup.

Online algorithms

The online luminosity monitoring can be used to provide luminosity information for machine tuning independently on the detector status. The determination and publication of instantaneous luminosity is performed by an application suite called the Online Luminosity Calculator (OLC). OLC

analyses raw information such as event and hit counts published by lumi-
nometers every 2 seconds to determines μ and then the luminosity. Due to
the small time allowed for online measurement, no background subtraction is
performed at this stage and OLC outputs are the instantaneous luminosity
averaged on the number of colliding bunches.

Most ATLAS detectors provide a LB-averaged luminosity, except LUCID
which is able to provide also bunch-by-bunch luminosity information for each
LB. For this subsystem OLC calculates the bunch integrated luminosity using
the following sum over all colliding BCIDs:

$$\mathcal{L} = \sum_{i \in BCID} \frac{\mu_i^{vis} f_r}{\sigma_{vis}} \quad . \tag{A.12}$$

A.5 LUCID description

LUCID is the only ATLAS detector designed entirely to luminosity mea-
surements. After the Run 1 shut-down, LUCID has been redesigned (called
LUCID 2 in the following) in order to cope with the new data taking condi-
tions foreseen for Run 2:

- the LHC machine peak instantaneous luminosity increase of a factor
  about two passing from Run 1 ($0.77 \cdot 10^{34}$ cm^{-2}s^{-1}) to Run 2 ($1.7 \cdot 10^{34}$
  cm^{-2} s^{-1});

- the change of the LHC beam pipe material in the old LUCID zone from
  stainless steel to aluminium;

- the decreased bunch spacing in the LHC machine (from 50 ns to 25
  ns).

The last point had as a consequence the redesign of the readout electronics.
The first two points had as consequence a significative increase of the detec-
tor occupancy which affects:
1) luminosity algorithm saturation (see for instance comment following Eq.A.4.1);
2) photomultipliers (PMT) saturation (drawn current);
3) PMT lifetime;
4) PMT radiation hardness.

The common solution to all these problems is to decrease the detector granularity and dimensions. This imposed the choice of PMTs having a small photo-cathode diameter (from 14 mm for the old LUCID to 10 mm). A complete description of choice and characterization of the installed PMTs is presented in [109], while in the following I describe in more detail my contribution in equalizing the response of all sensors in order to find the optimal working point.
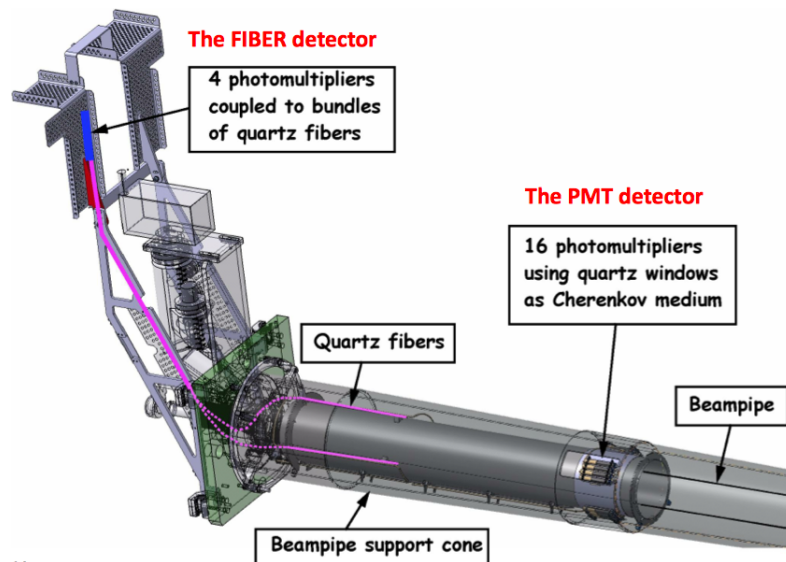
A.5.1 The LUCID design

Figure A.3: Sketch of one module of the new LUCID and its main components.

LUCID 2 consists of two modules symmetrically positioned around the beam pipe at ± 17 m from the interaction point (IP). Each module is installed on a carbon fibre cylindrical support tube surrounding the LHC beam-pipe, at a radial distance of about 10 cm from the beam-pipe. The active detectors in each module are:

- 16 photomultipliers (PMTs), attached to the carbon fibre support, whose quartz window acts as a Cherenkov radiator. Charged particles crossing the window produce light which is converted into an electri-
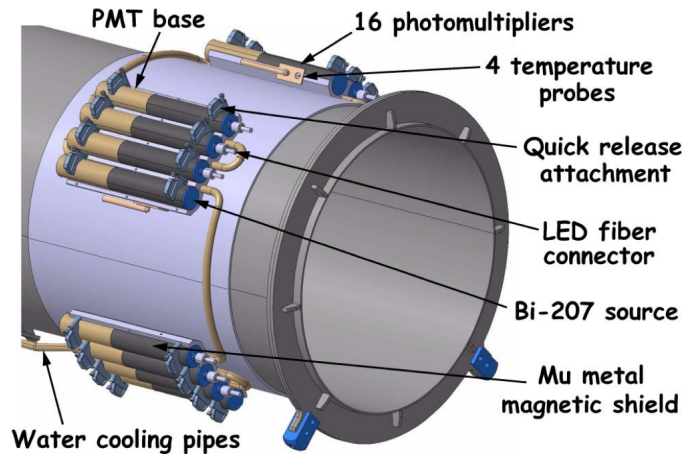
Figure A.4: Details of the PMT region, including the services such as temperature probes, PMT calibration fibres, cooling pipes, and mu-metal shields for PMT protection against the stray magnetic field.

cal current in the PMT cathode and amplified in the dynode chain, producing a measurable signal.

- 4 bundles of optical quartz fibers, acting as Cherenkov radiator for charged particles. The light produced is read-out by four PMTs placed at about one meter from the detector, outside the high radiation region. These PMTs are calibrated using a LED system, monitored by a pin diode.

A sketch of a LUCID 2 module is shown in Fig.A.3, while the details of the PMT region are shown in Fig.A.4, including the services such as temperature probes, PMT calibration fibres, cooling pipes and mu-metal shields for PMT protection against the stray magnetic field.

The 16 PMTs are arranged in four groups with different features, each acting as an independent detector:

- *Bi detector*: four Hamamatsu R760 PMTs with diameter of 10 mm equipped with liquid ^{207}Bi radiation source on the window for calibrating purposes;

- *Bi2 detector*: four Hamamatsu R760 PMTs with diameter of 10 mm

equipped with liquid [207]Bi radiation source on the window for calibrating purposes;

- *Modified Bi detector*: four Hamamatsu R760 PMTs with diameter of 7 mm. The acceptance has been further reduced with an aluminium layer deposited on the window. Also modified PMTs are equipped with [207]Bi radiation source for calibrating purposes;

- *Spares*: four Hamamatsu R760 PMTs of 10 mm diameter equipped with [207]Bi radiation source turned off and used as spares.

A.5.2 The LUCID electronics

The LUCID read-out system is based on custom-made VME boards called LUCROD (LUCid ReadOut Driver), two per sides. The LUCROD boards are placed directly in the experimental cavern, 15 m from the detector, in order to avoid dispersion or attenuation effects on signals. The electronic signals from PMTs are guided through low loss transmission coaxial cables to the board, in order to preserve the original shape and to avoid pole-zero compensation circuity. The signal duration is thus guaranteed to be below the 25 ns of LHC bunch spacing, as visible in Fig.A.5.
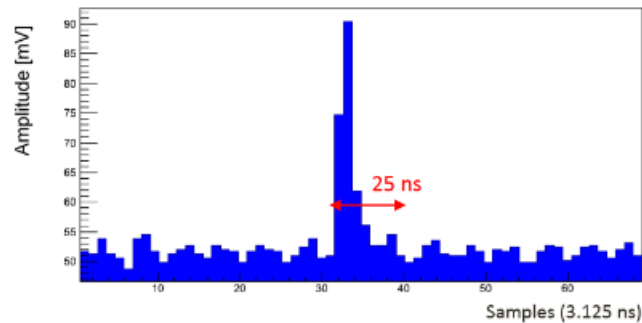
Figure A.5: Typical [207]Bi signal as recorded by the LUCROD card. The signal duration is within the 25 ns LHC BC duration.

In the LUCROD boards, signals are amplified and then digitalized by a Flash Analogue to Digital Converter (FADC). An FPGA integrates the

pulses and measures their amplitudes. The FPGA is also used to finely adjust the delays, the threshold and the signal alignment with the ATLAS data acquisition system (DAQ). LUCROD main block scheme is reported in Fig.A.6. The LUCROD board is implemented in a VME 9U card with 16 input channels and therefore each card contains eight of the above described unit. The output from the first eight FPGAs are fed into a main FPGA which sums up the charge from different PMTs for individual bunch crossing and calculates single sided hit pattern. A hit is defined as an electric signal above a fixed amplitude threshold.
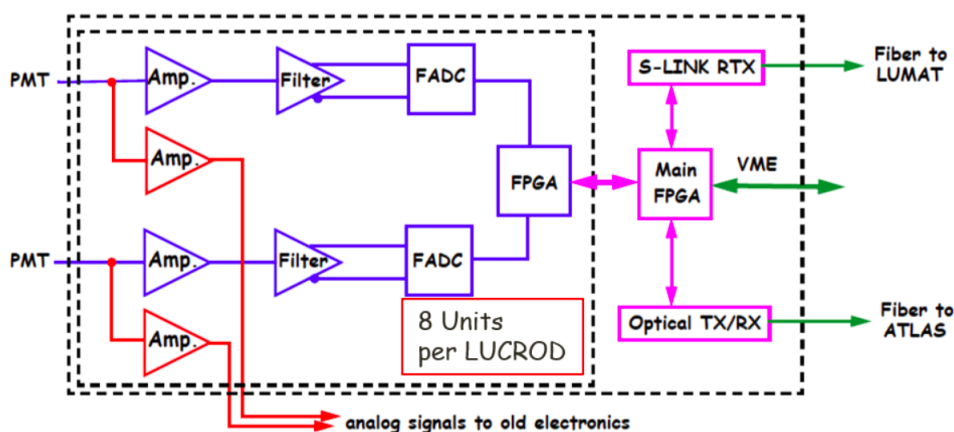
Figure A.6: LUCROD scheme for two input channels.

LUCRODs transmit the hit pattern to another readout custom made board named LUMAT through optical links at 1.3 Gbit/s. LUMAT correlates hit patterns coming from the two sides of the detector and provides coincidence algorithms (AND algorithms for example).

LUCID Luminosity Algorithms

The new LUCID 2 is designed to measure the luminosity using two different approaches which are in many aspects complementary with respect to systematic effects and sensitivity to instrumental issues and background contamination. The two different methods are:

- *Event/Hit counting*: this type of algorithms counts either the number of event with at least one hit detected in LUCID or the number of hits in the event. Each algorithm can be in single side or coincidence mode.

- *Particle flow*: in this case no threshold is applied on the signal amplitude and a characteristic of the signal (charge or amplitude), proportional to the passage of one or more particles, is recorded.

While the counting algorithms depend on the geometrical acceptance and efficiency of the detector and are based on a statistical assumption (the average number of interaction per bunch crossing is Poissonian distributed), the particle flow method provides a quantity directly proportional to luminosity. This method is thus free from saturation effect but it depends on the linearity of the read out chain and on the stability of PMT gain. A reliable and robust calibration system is thus mandatory.

A.5.3 Calibration System
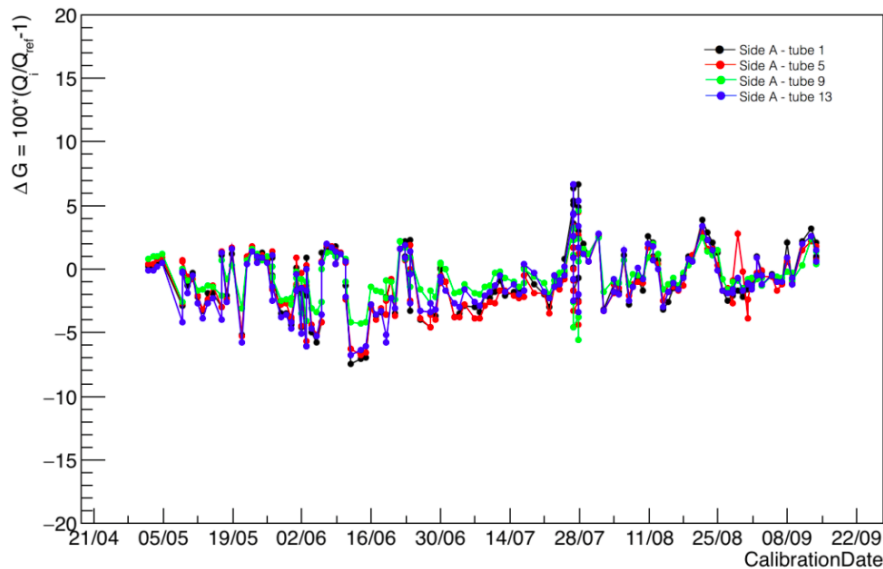
Figure A.7: Trending plot of gain changes as a function of time for ^{207}Bi calibrations in 2016. The requirement of a maximum fluctuation of 5% has been achieved.

In order to correct for gain loss and ageing during data taking, a day by day monitoring of the PMT response is crucial. This is performed in dedicated calibration runs, usually taken at the end of every LHC fill.

As presented in the description of the LUCID 2 detector, all sets of PMTs are equipped with liquid ^{207}Bi radioactive source on the top of quartz window for calibration purposes. This method of calibration is completely new and used for the first time in LUCID. The electrons from the ^{207}Bi are produced by internal conversion with a maximum energy of about 1 MeV. The PMT quartz window, with a thickness of 1.2 mm, has a Cherenkov kinetic energy threshold of 175 keV and therefore the ^{207}Bi electrons have enough energy to cross the thickness of the window and produce an amount of light comparable to a high energy minimum ionizing particle traversing the window with an energy above the Cherenkov threshold.

The calibration procedure consists of monitoring the mean charge and amplitude distribution and adjust the PMT HV to keep the fluctuation of gain at 5% maximum with respect to a fixed nominal value. A change in gain of 5% corresponds, in fact, in a change in the luminosity value of 1% for the counting algorithms and of 5% for the charge algorithm.

The final stability of the results obtained using this procedure is clearly visible in Fig.A.7: the requirement of a maximum gain fluctuation of 5% has been achieved with very few exceptions corresponding to very long LHC fills when a large gain loss is experienced. Due to the great success on stability and reliability of the method, this type of calibration is under deeper analysis for future development (LUCID for High-Luminosity LHC).

A.5.4   Equalization of charged particle response

The equalization of the response of all PMTs to charged particles is important for LUCID to properly asses the luminosity measurement.

I personally took the responsibility of equalize the PMTs before the installation in LUCID. The procedure consisted in measuring the charge output from each PMT using the ^{207}Bi source and the LUCID readout card LUCROD. The experimental setup was composed by:

- a **Black Box** where PMTs were placed, totally isolated from the outside

environment in order for PMTs to be sensitive only to the source;

- **2 Hamamatsu R760** PMTs where calibrated simultaneously. They were located in the same basement, one in the vertical position and the other in the horizontal position;

- **2 disk-shaped ^{207}Bi sources** were placed in close contact to PMTs. In order to avoid photo-electron looses, the source dimensions were larger than the PMT diameter;

- **2 Power Supplies** in a NIM crate to feed PMTs with high voltage;

- a **LUCROD** board to sample and store signals. Each PMT was connected to a different electronic channel. The trigger was required to be the logic OR of the two input channels.

A Gaussian fit on the charge distribution obtained from the LUCROD board has been performed and the error from fit has been considered as the final systematic uncertainty on the measurement. The high voltage of each PMT has then been adjusted to give the same charge output for each tube. A common value of the charge Q_{ref}=300 mV·ns that corresponds to a gain of 10^5 was chosen. In practice the charge from the ^{207}Bi source was measured at five different values of HV in steps of 10 volts in the region around the nominal charge. For each point, the HV (V in Eq.A.13) corresponding to exactly the nominal charge was calculated using the formula:

$$\frac{Q_{ref}}{Q_i} = \left(\frac{V}{V_i}\right)^{\beta} \tag{A.13}$$

where the measured β parameter defined in Eq.A.13 characterizes the gain of the PMT. The HV working point was set as the average of the five HV values obtained in this procedure. This procedure was applied to all the PMTs and thus the response to the ^{207}Bi source was equalized for all the PMTs under study.

The reproducibility as well as the time stability of the measurement were studied by fixing the PMTs at the nominal HV and by evaluating the mean charge obtained 30 minutes and 10 hours after the scans. The charge distributions with the relative Gaussian fit in red are presented in Fig.A.8 for two

arbitrary chosen PMTs (named for convenience 1 and 2) of 10 mm diameter; the mean values of the charge are reported in Table A.1. The values obtained at different times are compatible within errors.

PMT 1	Q_{30min}	=	287.9 ± 1.0 mV·ns
	Q_{10h}	=	287.9 ± 1.1 mV·ns
PMT 2	Q_{30min}	=	312.7 ± 1.2 mV·ns
	Q_{10h}	=	310.3 ± 1.3 mV·ns

Table A.1: Mean charge at nominal HV for two arbitrary chosen PMTs obtained performing the measurement after 30 minutes and 10 hours from the equalization scans. The two measurements are compatible between errors.

Uncertainties on PMT gain equalization are dominated by the limited reproducibility of the coupling between the ^{207}Bi source and the PMT. The overall precision was evaluated to be 5 V which corresponds to about 5% uncertainty on the absolute gain of the PMT.
Unfortunately the method described above was not applicable to the modified PMTs with further reduced diameter (7 mm): because of the smaller acceptance, the number of electrons crossing the quartz window was not enough to produce a visible peak in the charge distributions. A different study based on the measurement of the anodic current was performed. From a linear fit of the double logarithmic dependence of the anodic current to the HV (see for example Fig. A.9), the gain can be inferred from the formula:

$$\frac{I_{ref}}{I_i} = \left(\frac{V_i}{V_{ref}}\right)^{\beta} \tag{A.14}$$

where I_{ref} is the current at a nominal value of HV (set for convenience to V_{ref}=650 V), I_i is the current at a certain HV (V_i) and β is related to the PMT gain.

A.6   Luminosity Performances in Run 2

During Run 2 data taking period (2015-2017), the instantaneous luminosity peak has increased from $5 \cdot 10^{33}$cm^{-2}s^{-1} in 2015 to $20 \cdot 10^{33}$cm^{-2}s^{-1} in 2017
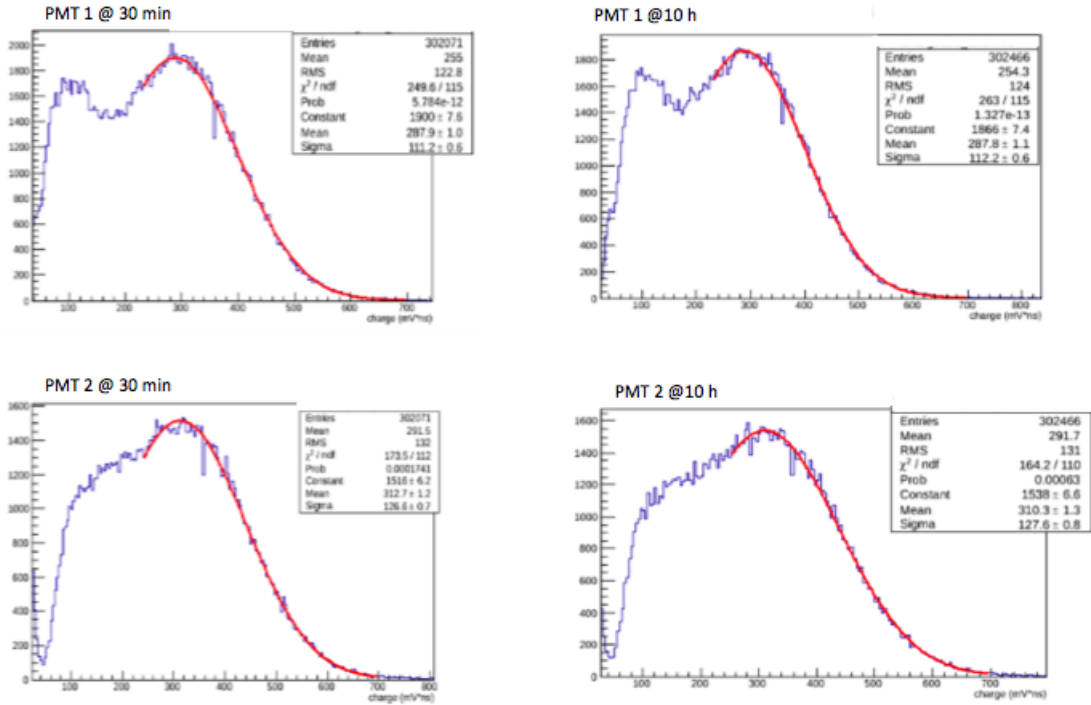
Figure A.8: Charge distribution at nominal HV for two arbitrary chosen PMTs obtained performing the measurement after 30 minutes and 10 hours from the equalization scans. The Gaussian fit on the distributions is visible in red.
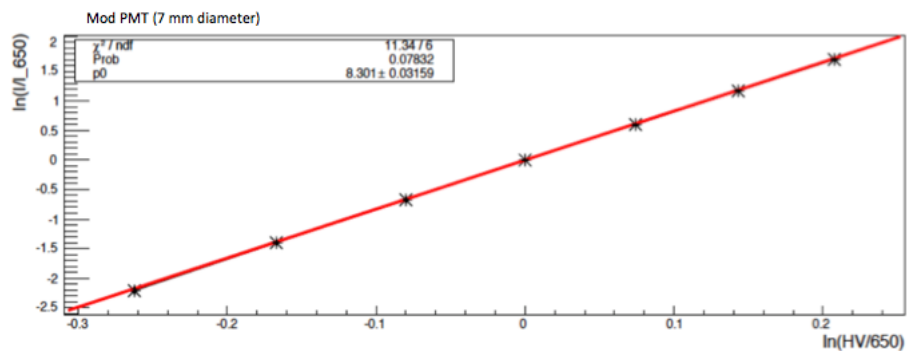

Figure A.9: Linear fit on the double logarithmic anodic current as a function of the increasing HV for a modified PMT (7 mm diameter quartz window).
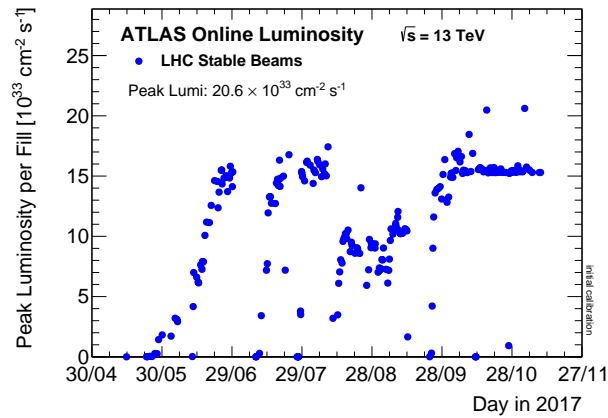
Figure A.10: Peak instantaneous luminosity delivered to ATLAS during stable beams for p-p collisions at 13 TeV centre-of-mass energy for each LHC fill as a function of time in 2017.
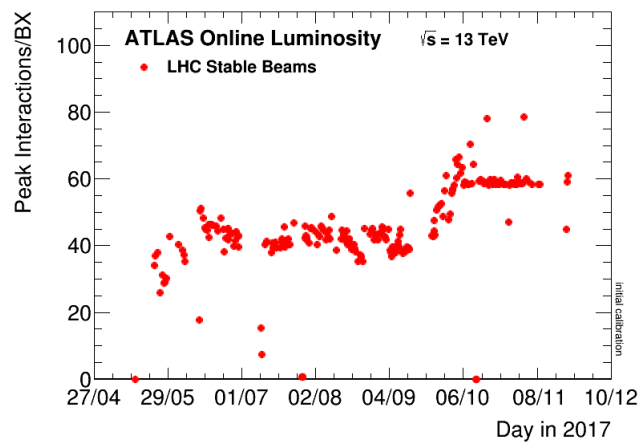
Figure A.11: Maximum number of inelastic collisions per bunch crossing during stable beams for p-p collisions at 13 TeV centre-of-mass energy for each fill in 2017.

(Fig.A.10) while the maximum number of interactions per bunch crossing increased from 28 in 2015 up to 80 in 2017 (see Fig. A.11).

For the whole period, LUCID was the so called *ATLAS-Preferred* luminosity detector, able to provide luminosity on LB basis and integrated over two seconds for beam monitoring operations. Moreover, as visible for example in Fig.A.12 where the measured pileup parameter is plotted as a function of the bunch-crossing number (BCID) averaged over the duration of the run, more than three orders of magnitude are visible between the μ measured for colliding bunches and the background in the non-colliding ones. This feature guarantees to LUCID only in the whole ATLAS experiment the possibility to provide also bunch-by-bunch luminosity measurements in all LHC data taking conditions.
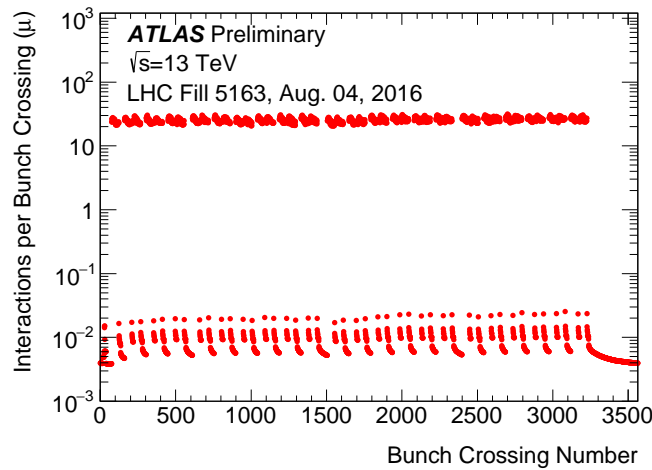

Figure A.12: Measured μ value as a function of the bunch-crossing number (BCID) averaged over the duration of the run, in a physics fill in 2016. More than three order of magnitude are visible between the values obtained in the colliding BCID and the non-colliding ones.

Due to the redundant strategy of luminosity measurement in ATLAS, the long term stability of the measurement can be inferred from the ratio between the values provided by the different luminometers. In Fig.A.13 the fractional difference in run-integrated luminosity between the LUCID Hit OR algorithm and the TILE, EMEC, FCal and track-counting algorithms is shown. Each point corresponds to an ATLAS run recorded during 25 ns
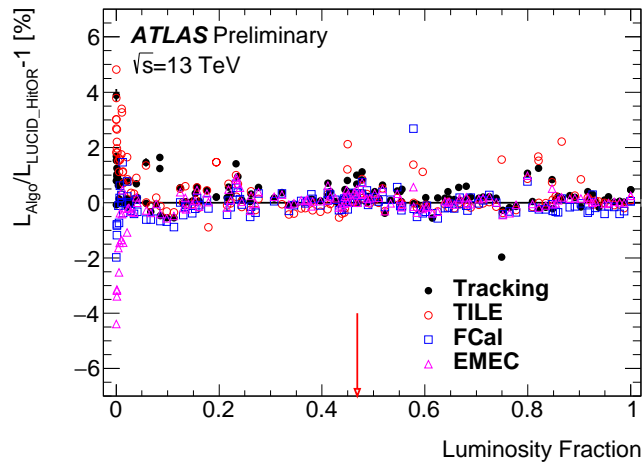
Figure A.13: Fractional difference in run integrated luminosity between the LU-CID Hit OR algorithm, and the TILE, EMEC, FCal and track-counting algorithms. Each point corresponds to an ATLAS run recorded during 25 ns bunch train running in 2016 at 13 TeV centre-of-mass energy. The luminosity measurements by TILE, EMEC, FCal and Tracking have been normalized to LUCID in the physics run indicated by the red arrow.

bunch-train running in 2016 at 13 TeV centre-of-mass energy. The luminosity measurements by TILE, EMEC, FCal and Tracking have been normalized to LUCID in a physics run recorded on August 4th, which is indicated by the red arrow. As clearly visible, a 0.7% long term stability has been achieved, except for the very first period, that actually contributes only with the 3% to the total integrated luminosity. Preliminary results indicates that the long term stability in 2017 is of the order of 2%, but the offline analysis is still ongoing to improve the result.

The long term stability is one of the components of the overall systematic uncertainty in the luminosity measurement. The other main systematics arise from the calibration using the Van der Meer method ($\sim 2\%$), the calibration transfer ($\sim 1\%$), due to the transition from low μ (VdM) to high μ (physics) runs. The final uncertainty of 2.1% and 2.2% has been achieved in 2015 and 2016 respectively. The analysis for 2017 is still ongoing.

A.7 Conclusions

The LUCID detector was redesigned to cope with the new data taking condition of LHC Run 2. Thanks to its stability and reliability, due also to a completely new and very successful calibration method, LUCID provided during the whole period the official measurement of ATLAS luminosity both online for the LHC instantaneous luminosity, used for beam monitoring and operational tuning, and offline for physics analysis, on LB and bunch-by-bunch basis. The redundant strategy used in ATLAS to evaluate the luminosity lead to a final uncertainty on the measurement of 2.1% and 2.2% in 2015 and 2016, respectively, while the analysis for 2017 is still ongoing. In this contest, I performed the measurements needed to equalize all the read-out PMTs by means of a ^{207}Bi radioactive source. This procedure proved to be essential to have an homogeneous detector in terms of gain of the single sensor and thus response to particle from p-p collisions.

Appendix B

Muon Reconstruction Performances

Muons are key to some of the most important physics results published by the ATLAS experiment at the LHC. These results include the discovery of the Higgs boson and the measurement of its properties, the precise measurement of Standard Model processes, and searches for physics beyond the Standard Model.

The performance of the ATLAS muon reconstruction at a centre of mass energies $\sqrt{s}$ =7-8 TeV has been documented in recent publications [110]. During the 2013-2015 shutdown, the LHC was upgraded to increase the centre of mass energy from 8 to 13 TeV and the ATLAS detector was equipped with additional muon chambers and a new innermost Pixel layer, the Insertable B-Layer, providing measurements closer to the interaction point. Moreover, the muon reconstruction software was updated and improved.

After introducing the ATLAS muon reconstruction and identification algorithms, the performance of the muon reconstruction in the first dataset collected at $\sqrt{s}$= 13 TeV in which I directly participated is described. Measurements of the muon reconstruction efficiencies are presented, together with the description of the main systematic uncertainties. The results are based on the analysis of a large sample of $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ decays reconstructed in 3.2 fb^{-1} of $p-p$ collisions recorded in 2015.

The full description of the analysis has been published in a public paper in

May 2016 ([111]).

B.1   Muon reconstruction and identification

Muon reconstruction is first performed independently in the Inner Detector (ID) and Muon Spectrometer (MS). The information from individual subdetectors is then combined to form the muon tracks used in physics analyses. The complete description of the ID and MS detectors and of the muon reconstruction is reported in Chapter 2. The muon reconstruction used in this work evolved from the algorithms defined in Chapter 2. These algorithms were improved in several ways. The use of a Hough transform to identify the hit patterns for seeding the segment-finding algorithm makes the reconstruction faster and more robust against misidentification of hadrons, thus providing better background rejection early in the pattern recognition process. The calculation of the energy loss in the calorimeter was also improved. An analytic parametrization of the average energy loss is derived from a detailed description of the detector geometry. The final estimate of the energy loss is obtained by combining the analytic parametrization with the energy measured in the calorimeter. This method yields a precision on the mean energy loss of about 30 MeV for 50 GeV muons.

B.2   Muon identification

Muon identification is performed by applying quality requirements that suppress background, mainly from pion and kaon decays, while selecting prompt muons with high efficiency and/or guaranteeing a robust momentum measurement. Muon candidates originating from in-flight decays of charged hadrons in the ID are often characterized by the presence of a distinctive kink topology in the reconstructed track. As a consequence, it is expected that the fit quality of the resulting combined track will be poor and that the momentum measured in the ID and MS may not be compatible. Several variables offering good discrimination between prompt muons and background muon candidates are studied in simulated $t\bar{t}$ events. Muons from W decays

are categorized as signal, while muon candidates from light-hadron decays are categorized as background. To guarantee a robust momentum measurement, specific requirements on the number of hits in the ID and MS are used. For the ID, the quality cuts require at least one Pixel hit, at least five SCT hits, fewer than three Pixel or SCT holes, and that at least 10% of the TRT hits originally assigned to the track are included in the final fit; the last requirement is only employed for η between 0.1 and 1.9, in the region of full TRT acceptance. A hole is defined as an active sensor traversed by the track but containing no hits. A missing hit is considered a hole only when it falls between hits successfully assigned to a given track. If some inefficiency is expected for a given sensor, the requirements on the number of Pixel and SCT hits are reduced accordingly.

Four muon identification selections (Medium, Loose, Tight, and High-p_T) are provided to address the specific needs of different physics analyses. Loose, Medium, and Tight are inclusive categories in that muons identified with tighter requirements are also included in the looser categories (refer to Sec.2.4.3 for a complete description of categories).

B.3 Reconstruction efficiency

A precise determination of the muon reconstruction efficiency in the region $|\eta| < 2.5$ is obtained combining the informations coming from these detectors by means of the *Tag-and-Probe* method described in the following sections.

B.3.1 Efficiency measurement in the region $|\eta| < 2.5$

The Tag-and-Probe method is employed to measure the efficiency of the muon identification selections within the acceptance of the ID ($|\eta| < 2.5$). The method is based on the selection of an almost pure muon sample from $Z \to \mu\mu$ or $J/\psi \to \mu\mu$ events, requiring one leg of the decay (tag) to be identified as a Medium muon that fires the trigger and the second leg (probe) to be reconstructed by a system independent of the one being studied. A selection based on the event topology is used to reduce the background contamination.

Three kinds of probes are used to measure muon efficiencies. ID tracks and so called *Calorimeter-Tagged* (CT) muons (see Chapter 2) both allow a measurement of the efficiency in the MS, while MS tracks are used to determine the complementary efficiency of the muon reconstruction in the ID. Compared to ID tracks, CT muons offer a more powerful rejection of backgrounds, especially at low transverse momenta, and are therefore the preferred probe type for this part of the measurement. ID tracks are used as a cross-check and for measurements not directly accessible to CT muons. A direct measurement of the CT muon reconstruction efficiency is possible using MS tracks.

The efficiency measurement for Medium, Tight, and High-p_T muons consists of two stages. First, the efficiency $\epsilon(X|CT)$ (where X stands for Medium/Tight/High-p_T) of reconstructing these muons assuming a reconstructed ID track is measured using a CT muon as probe. Then, this result is corrected by the efficiency $\epsilon(ID|MS)$ of the ID track reconstruction, measured using MS probes, using the formula:

$$\epsilon(X) = \epsilon(X|ID) \cdot \epsilon(ID) = \epsilon(X|CT) \cdot \epsilon(ID|MS) \qquad (B.1)$$

where ($X = Medium/Tight/High - p_T$). A similar approach is employed when using ID probe tracks for cross-checks. This method is valid if two assumptions are satisfied:

- the ID track reconstruction efficiency is independent from the muon spectrometer track reconstruction ($\epsilon(ID) = \epsilon(ID|MS)$).

- the use of a CT muon as a probe instead of an ID track does not affect the probability for Medium, Tight, or High-p_T reconstruction ($\epsilon(X|ID) = \epsilon(X|CT)$).

Both assumptions have been tested using generator-level information from simulation and small differences are taken into account in the systematic uncertainties.

The muons selected by the Loose identification requirements are decomposed into two samples: CT muons within $|\eta| < 0.1$ and all other muons. The CT muon efficiency is measured using MS probe tracks, while the efficiency

of other muons is evaluated using CT probe muons in a trend similar to the Medium, Tight, and High-p_T categories. The level of agreement of the measured efficiency, $\epsilon^{Data}(X)$, with the efficiency measured with the same method in simulation, $\epsilon^{MC}(X)$, is expressed as the ratio of these two numbers, called the *Efficiency Scale Factor* (SF):

$$SF = \frac{\epsilon^{Data}(X)}{\epsilon^{MC}(X)} \tag{B.2}$$

This quantity describes the deviation of the simulation from the real detector behaviour. It is of particular interest for physics analyses, where it is used to correct the simulation. I directly work on this particular part of the analysis in order to provide the correct scale factors to be applied to all ATLAS analyses.

Tag-and-Probe method with $Z \to \mu\mu$

Events are selected by requiring muon pairs with an invariant mass within 10 GeV of the Z boson mass. The tag muon is required to satisfy the Loose isolation and Medium muon identification selections and to have a transverse momentum of at least 24 GeV. Requirements on the significance of the transverse impact parameter d_0 ($|d_0|/\sigma(d_0) < 3.0$) and on the longitudinal impact parameter $|z_0|$ ($|z_0| < 10$ mm) of the tag muon are imposed. Finally, the tag muon is required to have triggered the readout of the event.

The probe muon is required to have a transverse momentum of at least 10 GeV and to satisfy the Loose isolation criteria. While this is sufficient to ensure high purity in the case of MS probe tracks, further requirements are applied to both the ID track and CT muon probes. In the case of ID tracks, an isolation requirement is applied which is considerably stricter than the Loose selection in order to suppress backgrounds as much as possible. In addition, the invariant mass window is tightened to 5 GeV around the Z boson mass, rather than the 10 GeV used in the other cases. For CT muon probes, additional requirements on the compatibility of the associated calorimeter energy deposit with a muon signature are applied to further enhance the purity. The ID probe tracks and CT probe muons must also have transverse and

longitudinal impact parameters consistent with being produced in a primary pp interaction, as required for tag muons. A probe is considered successfully reconstructed if a reconstructed muon is found within a cone in the $\eta - \phi$ plane of size $\Delta R = 0.05$ around the probe track.

A small fraction (about 0.1%) of the selected tag-probe pairs originates from sources different than $Z \to \mu\mu$ events. For a precise efficiency measurement, these backgrounds must be estimated and subtracted. Contributions from $Z \to \tau\tau$ and $t\bar{t}$ decays are estimated using simulation. Additionally, multijet events and $W \to \mu\nu$ decays in association with jet activity (W+jets) can yield tag–probe pairs through secondary muons from heavy- or light-hadron decays. As these backgrounds are approximately charge-symmetric, they are estimated from the data using same-charge (SC) tag-probe pairs. This leads to the following estimate of the opposite-charge (OC) background, N^{Bkg}, for each region of the kinematic phase-space:

$$N^{Bkg} = N_{OC}^{Z,t\bar{t}MC} + T \cdot \left(N_{SC}^{Data} - N_{SC}^{Z,t\bar{t}MC} \right) \tag{B.3}$$

where $N_{OC}^{Z,t\bar{t}MC}$ is the contribution from $Z \to \tau\tau$ and $t\bar{t}$ decays, N_{SC}^{Data} is the number of SC pairs measured in data and $N_{SC}^{Z,t\bar{t}MC}$ is the estimated contribution of the $Z \to \mu\mu$, $Z \to \tau\tau$, and $t\bar{t}$ processes to the SC sample. T is a global transfer factor that takes into account the charge asymmetry of the multijet and W+jets processes, estimated in data using a control sample of events obtained by inverting the probe isolation requirement. For MS (ID) tracks, a value of T=1.7 (1.1) is obtained, while for CT muon probes the transfer factor is T=1.2. The systematic uncertainties in the transfer factor vary between 40 and 100% and are included in the systematic error in the reconstruction efficiency described in the following section.

The efficiency measured in the data is corrected for the background contributions described above, by subtracting the predicted probe yields attributed to these sources from the number of observed probes,

$$\epsilon = \frac{N_R^{Data} - N_R^{Bkg}}{N_P^{Data} - N_P^{Bkg}} \tag{B.4}$$

where N_P denotes the total number of probes and N_R the number of successfully reconstructed probes. The resulting efficiency can then be compared directly to the result of the simulation.

Tag-and-probe method with $J/\psi \to \mu\mu$

The reconstruction efficiencies of the Loose, Medium, and Tight muon selections at low p_T are measured from a sample of $J/\psi \to \mu\mu$ events selected using a combination of single-muon triggers and the dedicated "muon + track" trigger.

Tag-probe pairs are selected within the invariant mass window of 2.7-3.5 GeV and requiring a transverse momentum of at least 5 GeV for each muon. The tag muon is required to satisfy the Medium muon identification selection and to have triggered the readout of the event. In order to avoid low-momentum curved tracks sharing the same trigger region, tag and probe muons are required to be $\Delta R > 0.2$ apart when extrapolated to the MS trigger surfaces. Finally, they are selected with $\Delta z_0 = |z_0^{tag} - z_0^{probe}| < 5$ mm, to suppress background. A probe is considered successfully reconstructed if a selected muon is found within a $\Delta R = 0.05$ cone around the probe track.

The background contamination and the muon reconstruction efficiency are measured with a simultaneous maximum-likelihood fit of two statistically independent distributions of the invariant mass: events in which the probe is or is not successfully matched to the selected muon. The fits are performed in six p_T and nine η bins of the probe tracks. The signal is modelled with a Crystal Ball function with a single set of parameters for the two independent samples. Separate first-order polynomial fits are used to describe the background shape for matched and unmatched probes.

Systematic uncertainties

The main contributions to the systematic uncertainty in the measurement of the efficiency SFs with $Z \to \mu\mu$ and $J\psi \to \mu\mu$ events are shown in Figs.B.1 and B.2, as a function of η and p_T, respectively.

The uncertainty in the background estimate is evaluated in the $Z \to \mu\mu$ analysis by taking the maximum variation of the transfer factor T when esti-
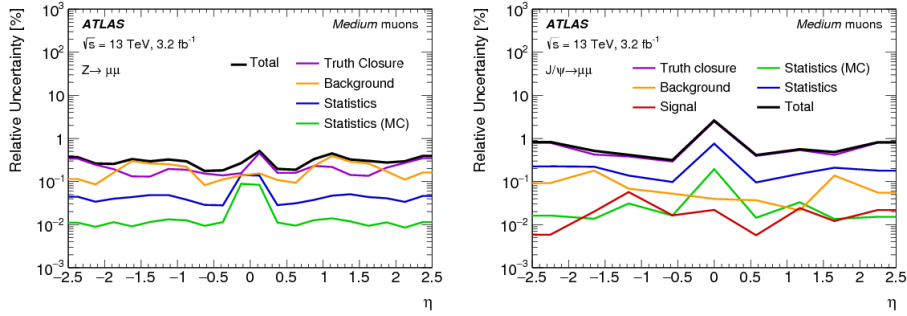
Figure B.1: Total uncertainty in the efficiency scale factor for Medium muons as a function of η as obtained from $Z \to \mu\mu$ data (left) for muons with $p_T > 10$ GeV, and from $J/\psi \to \mu\mu$ data (right) for muons with $5 < p_T < 20$ GeV. The combined uncertainty is the sum in quadrature of the individual contributions.
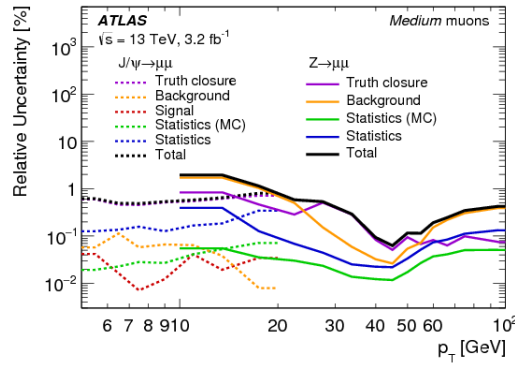

Figure B.2: Total uncertainty in the efficiency scale factor for Medium muons as a function of p_T as obtained from $Z \to \mu\mu$ (solid lines) and $J/\psi \to \mu\mu$ (dashed lines) decays. The combined uncertainty is the sum in quadrature of the individual contributions.

mated with a simulation-based approach and when assuming the background to be charge-symmetric. This results in an uncertainty of the efficiency measurement below 0.1% over a large momentum range, but reaching 1% for low muon momenta where the contribution of the background is most significant. In the $J/\psi \to \mu\mu$ analysis, the background uncertainty is estimated by changing the function used in the fit to model the background, replacing the first-order polynomial with an exponential function. An uncertainty due to the signal modelling in the fit, labelled as *Signal* in Figs.B.1 and B.2, is

also estimated using a convolution of exponential and Gaussian functions as an alternative model. Each uncertainty is about 0.1%.

The cone size used for matching selected muons to probe tracks is optimised in terms of efficiency and purity of the matching. The systematic uncertainty deriving from this choice is evaluated by varying the cone size by $\pm 50\%$. This yields an uncertainty below 0.1% in both analyses.

Possible biases in the tag-and-probe method, such as biases due to different kinematic distributions between reconstructed probes and generated muons or correlations between ID and MS efficiencies, are estimated in simulation by comparing the efficiency measured with the tag-and-probe method with the "true" efficiency given by the fraction of generator-level muons that are successfully reconstructed. This uncertainty is labelled as *Truth Closure* in Figs.B.1 and B.2. In the $Z \to \mu\mu$ analysis, agreement better than 0.1% is observed in the high momentum range. This uncertainty grows at low p_T, and differences up to 0.7% are found in the $J/\psi \to \mu\mu$ analysis. A larger effect of up to 1-2% is measured in both analyses in the region $|\eta| < 0.1$. In the extraction of the efficiency scale factors, the difference between the measured and the "true" efficiency cancels to first order. To take into account possible imperfections of the simulation, half of the observed difference is used as an additional systematic uncertainty in the SF.

No significant dependence of the measured SFs with p_T is observed in the momentum range considered in the $Z \to \mu\mu$ analysis. An upper limit on the SF variation for large muon momenta is extracted from simulation, leading to an additional uncertainty of 2-3% per TeV for muons with $p_T > 200$ GeV. The efficiency scale factor is observed to be independent of the amount of pile-up.

Results

Fig.B.3 shows the muon reconstruction efficiency as a function of η as measured from $Z \to \mu\mu$ events for the different muon selections. The efficiency as measured in data and the corresponding scale factors for the Medium selection are also shown in Fig.B.4 as a function of η and ϕ. The efficiency at low p_T is reported in Fig.B.5 as measured from $J/\phi \to \mu\mu$ events as a
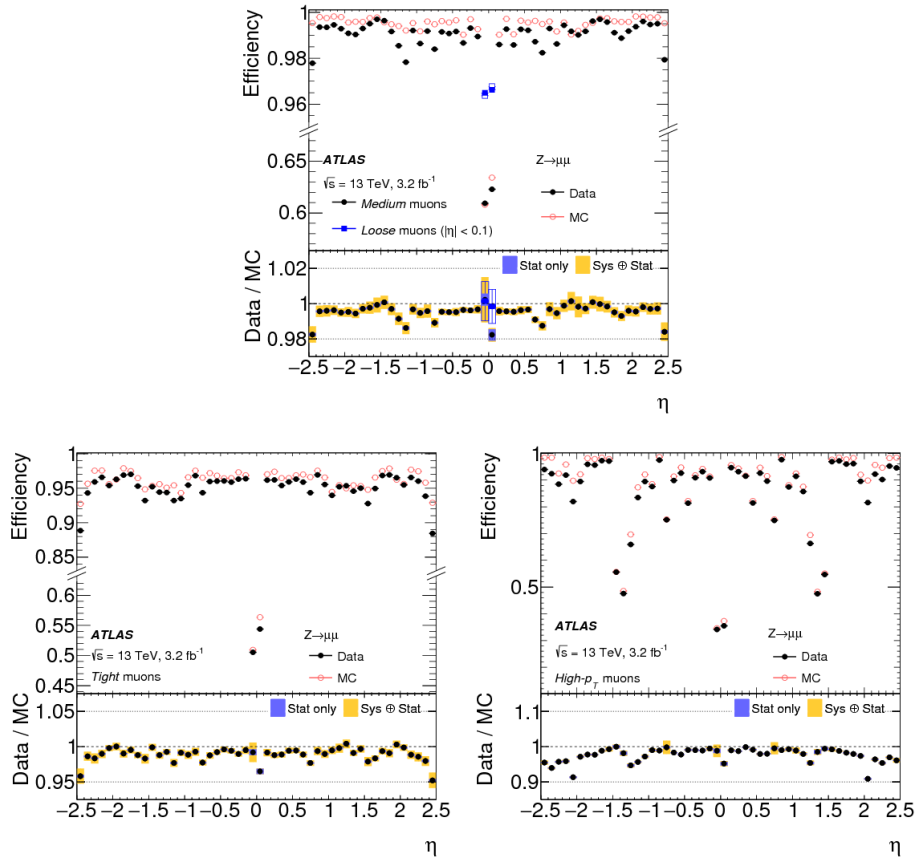
Figure B.3: Muon reconstruction efficiency as a function of η measured in $Z \to \mu\mu$ events for muons with $p_T > 10$ GeV shown for Medium (top), Tight (bottom left), and High-p_T (bottom right) muon selections. In addition, the top plot also shows the efficiency of the Loose selection (squares) in the region $|\eta| < 0.1$ where the Loose and Medium selections differ significantly. The error bars on the efficiencies indicate the statistical uncertainty. Panels at the bottom show the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties.

function of p_T in different η regions.

The efficiencies of the Loose and Medium selections are very similar through-out the detector with the exception of the region $|\eta| < 0.1$, where the Loose selection fills the MS acceptance gap using the calorimeter and the so called *Segment-Tagged* (ST) muons contributions. The efficiency of these selections is observed to be in excess of 98%, and between 90 and 98% for the Tight selection, with all efficiencies in very good agreement with those predicted
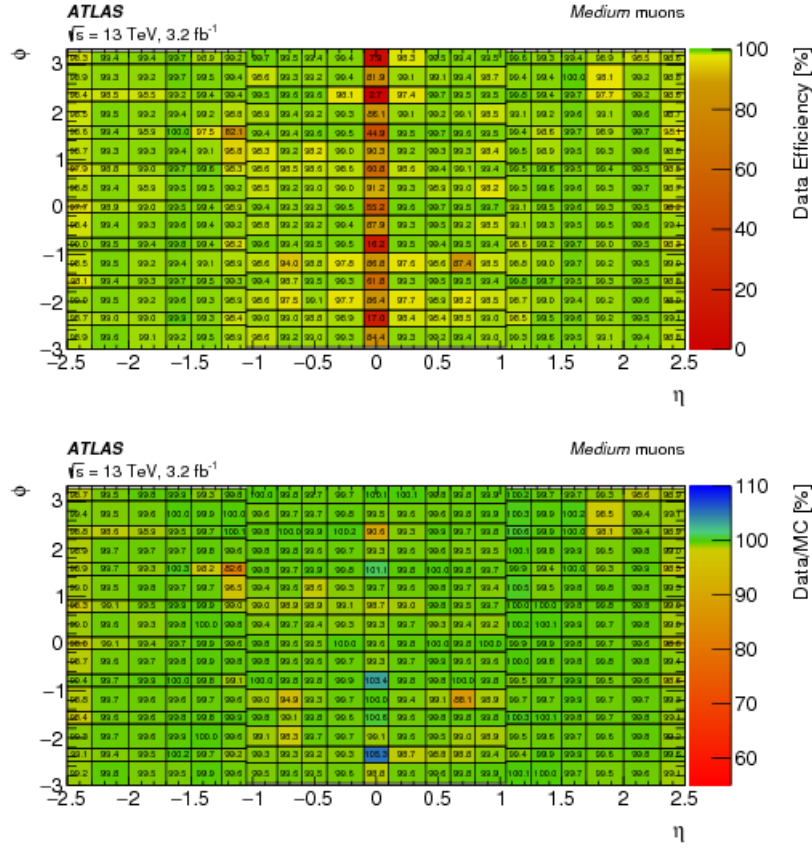
Figure B.4: Reconstruction efficiency measured in data (top), and the data/MC efficiency scale factor (bottom) for Medium muons as a function of η and ϕ for muons with $p_T > 10$ GeV in $Z \to \mu\mu$ events. The thin white bins visible in the region $|\phi| = \pi$ are due to the different bin boundaries in ϕ in the endcap and barrel regions.

by the simulation. An inefficiency due to a poorly aligned MDT chamber is clearly localised at $(\eta, \phi) = (-1.3, 1.6)$, and is the most significant feature of the comparison between collision data and simulation for these three categories. In addition, a 2%-level local inefficiency is visible in the region $(\eta, \phi) = (-1.9, 2.5)$, traced to temporary failures in the SCT readout system. Further local inefficiencies in the barrel region around $\phi = -1.1$ are also linked to temporary faults during data taking. The efficiency of the High-p_T selection is significantly lower, as a consequence of the strict requirements on momentum resolution. Local disagreements between prediction and observa-
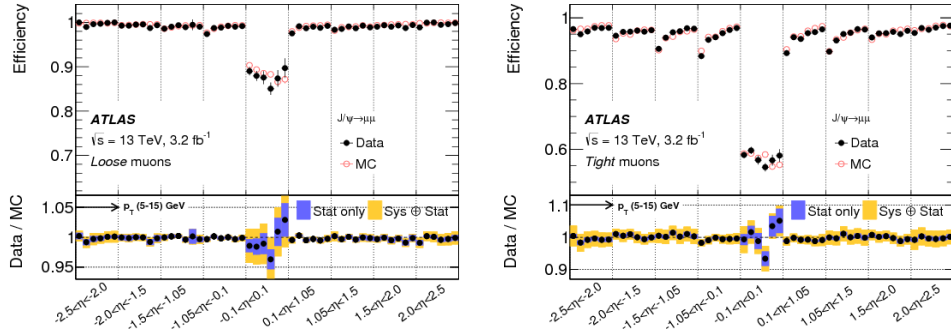
Figure B.5: Muon reconstruction efficiency in different η regions measured in $J/\psi \to \mu\mu$ events for Loose (left) and Tight (right) muon selections. Within each η region, the efficiency is measured in six p_T bins (5-6, 6-7, 7-8, 8-10, 10-12, and 12-15 GeV). The resulting values are plotted as distinct measurements in each η bin with p_T increasing from 5 to 15 GeV going from left to right. The error bars on the efficiencies indicate the statistical uncertainty. The panel at the bottom shows the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties.

tion are more severe than in the case of the other muon selections.

Apart from the poorly aligned MDT chamber, they are most prominent in the CSC region. Fig.B.6 shows the reconstruction efficiencies for the Medium muon selection as a function of transverse momentum, including results from $Z \to \mu\mu$ and $J/\psi \to \mu\mu$, for muons with $0.1 < |\eta| < 2.5$. The efficiency is stable and slightly above 99% for $p_T > 6$ GeV. Values measured from $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events are in agreement in the overlap region between 10 and 20 GeV. The efficiency scale factors are also found to be compatible.

B.3.2   Efficiency measurement in the region $|\eta| > 2.5$

As described in the previous sections, the reconstruction of combined muons is limited by the ID acceptance to the pseudorapidity region $|\eta| < 2.5$. For $|\eta| > 2.5$, the efficiency is recovered by using the so called *Extrapolated or Standalone* (ME) muons included in the Loose and Medium muon selections. A measurement of the efficiency SF for muons in the region $2.5 < |\eta| < 2.7$ (high-η region) is performed using the method described in [110]. The number of muons observed in $Z \to \mu\mu$ decays in the high-η region is normalised
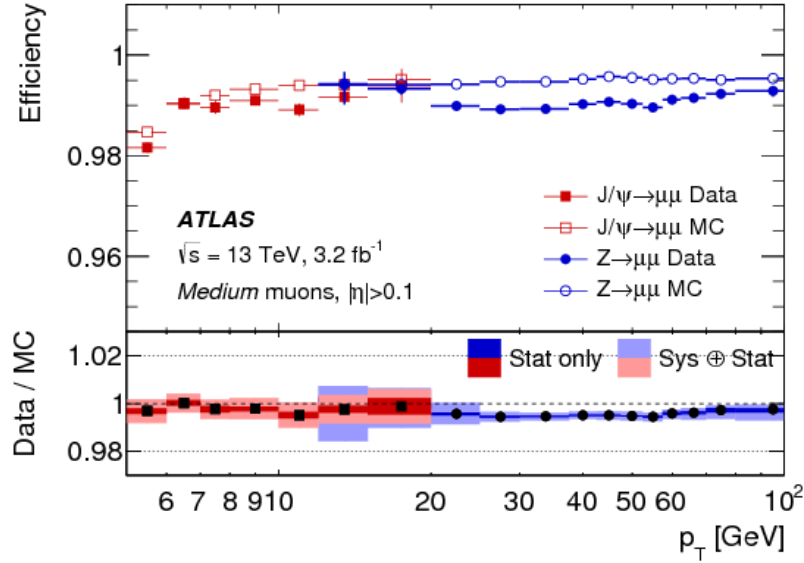
Figure B.6: Reconstruction efficiency for the Medium muon selection as a function of the p_T of the muon, in the region $0.1 < |\eta| < 2.5$ as obtained with $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events. The error bars on the efficiencies indicate the statistical uncertainty. The panel at the bottom shows the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties.

to the number of muons observed in the region $2.2 < |\eta| < 2.5$. This ratio is calculated for both data and simulation, applying all known performance corrections to the region $|\eta| < 2.5$. The SFs in the high-η region are defined as the ratio of the aforementioned ratios and are provided in four η and 16 ϕ bins. The values of the SFs measured using the 2015 dataset are close to 0.9 and are determined with a 3–5% uncertainty.

Appendix C

Flavour fit using MADGRAPH and ALPGEN

The nominal Monte Carlo generator used to simulate the Z+jets processes in the full analysis is SHERPA. In the contest on the flavour fit, since the truth labelling of jets strongly depends on the Monte Carlo used, studies are performed with MADGRAPH and ALPGEN too. This additional study is needed to verify the stability of the full fit procedure and of the results. Flavour fit results for the Z+1 b-jet analysis is presented in Section C.1 and for Z+2 b-jets in Section C.2.

C.1 Flavour fit for Z+1 b-jet analysis

The flavour fit on MADGRAPH and ALPGEN Z+jets sample is performed using the same strategy as on SHERPA. In the Z+1 b-jet the variable to fit is the b-tagging weight (MV2c10) distributed in three bins for single lepton channel. Pre-fit and post-fit plots are shown in Fig.C.1 and in Fig.C.2 for the Z+jets sample modelled with MADGRAPH and ALPGEN, respectively.

In order to understand the scale factors obtained from the fit and summarised in Section 4.6.3, a deep look to nuisance parameter pull and correlations is needed. When MADGRAPH is used as baseline, the combination presents some very strong pulls, close to 1σ: on the JES effective NP1 (the largest jet energy scale nuisance parameter), the jet flavour composition, the
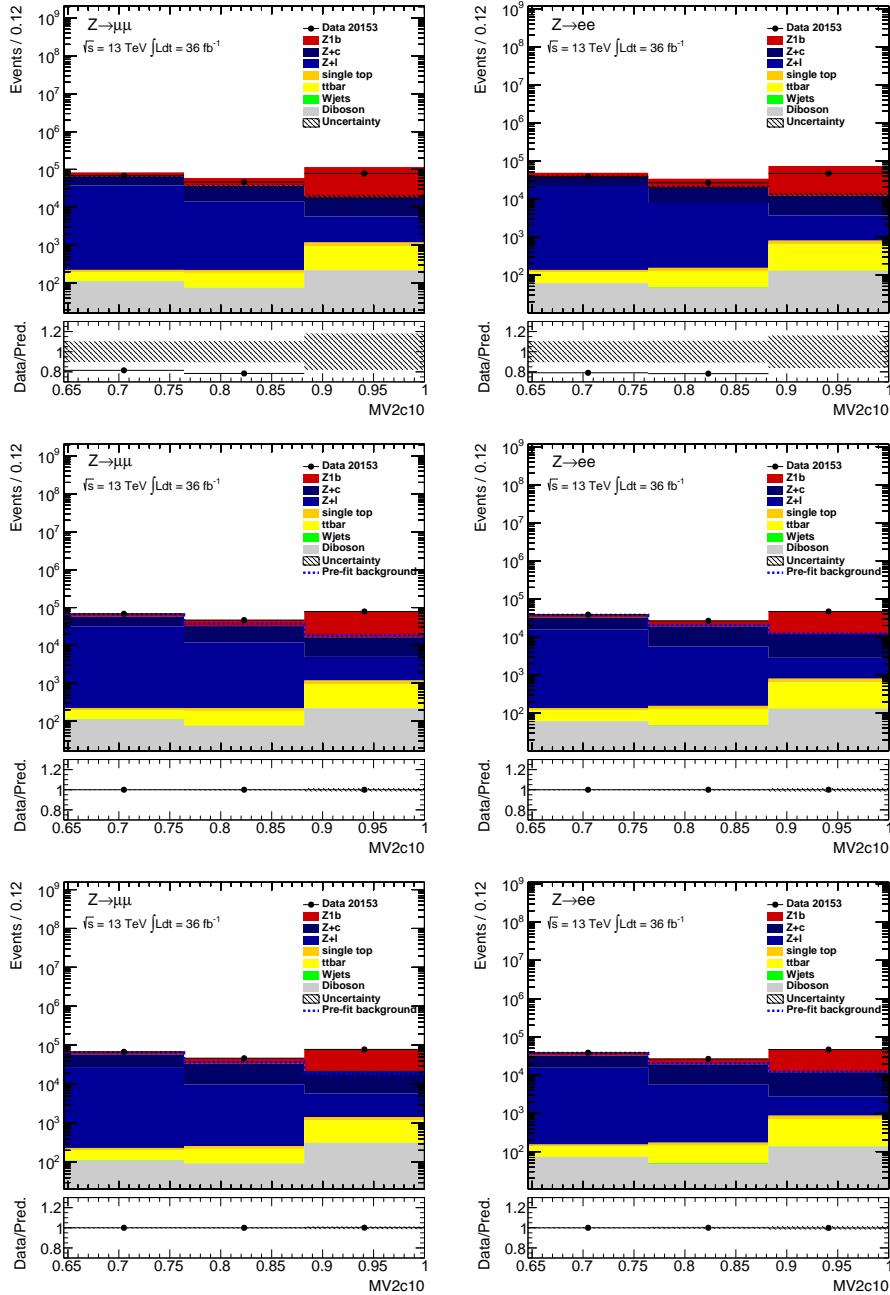
Figure C.1: Pre- (top row) and post-fit (middle and bottom row) plots for the muon (left) and electron (right) channels, using the Madgraph generator as baseline for the Z+jets samples in the Z+1 b-jet analysis. The fits shown in the middle row are performed separately in the electron and muon channels, while the ones shown in the bottom row are performed combining the two channels together.
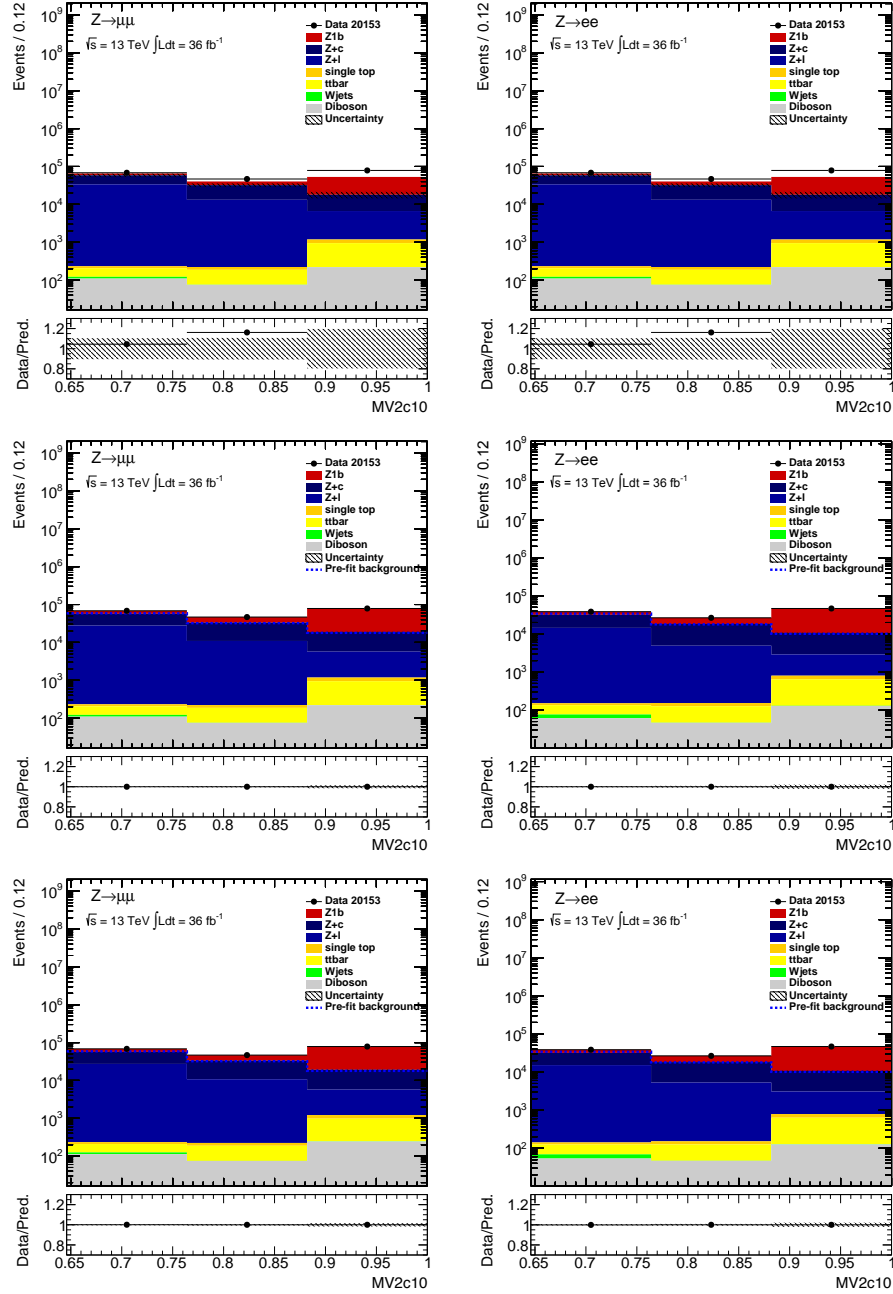
Figure C.2: Pre- (top row) and post-fit (middle and bottom row) plots for the muon (left) and electron (right) channels, using the ALPGEN generator as baseline for the Z+jets samples in the Z+1 b-jet analysis. The fits shown in the middle row are performed separately in the electron and muon channels, while the ones shown in the bottom row are performed combining the two channels together.

pileup offset and rho topology, and on the muon efficiency systematics (see Chapter 4). The pull on the jet flavour composition is expected, considering the initial standard assumption of 50% share between gluon and quark jets in all the samples. The deviation from nominal value of muon efficiency systematic can be explained by the strong variation in the fitted yields and scale factors from the muon channel alone to the combination, where the electron seems to dominate. Further checks will be performed on these parameters in the future, but it is important to remember that MADGRAPH, having the worst MC statistics, is more susceptible to fluctuations.

Looking at ALPGEN, no large pull is observed, apart from the $+1\sigma$ of the muon efficiency systematic component. In this last case, no large variation in the scale factors or yields is observed in either the muon or the electron channel, when passing from the single channel fits to the combined one.

The fit on MADGRAPH samples presents similar correlations to the SHERPA ones, and these also vary slightly from the single channel to the combination, although in this case, the combination does not reduce the number of correlations. In particular:

Zb tends to be correlated to the flavour tagging uncertainty component C0, the Z+light normalisation, and the jet flavour composition parameter and anticorrelated to the Z+c parameter;

Z+c tends to be anticorrelated with Zb, Z+l, flavour tagging uncertainty components C3 and light2, and the MC statistical uncertainty of the second bin of the distribution;

Z+l tends to be anticorrelated with Z+c, flavour tagging uncertainty component light3, the MC statistical uncertainty of the first bin of the distribution, and correlated with the MC statistical uncertainty of the second bin of the distribution, the flavour tagging uncertainty components light0 and C3, and Zb.

The correlations between the different Z+jets process normalisations is not surprising, neither with the flavour tagging uncertainties. The correlations with the MC statistical uncertainty need more investigations.
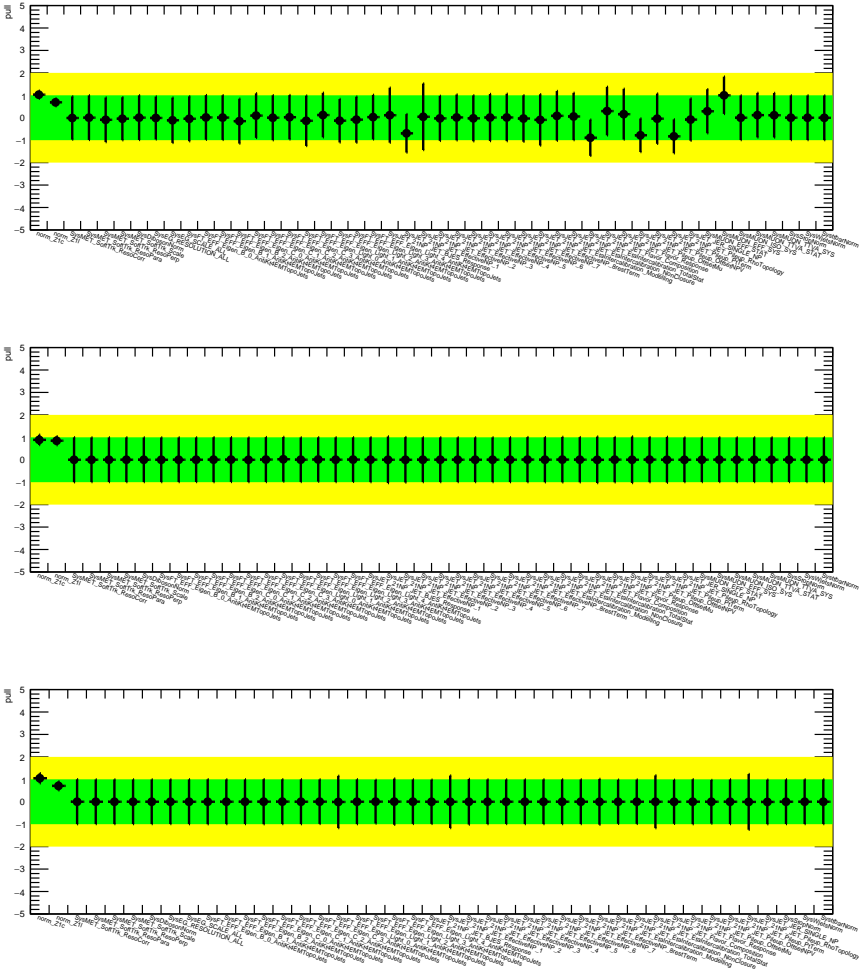
Figure C.3: Pulls of all the nuisance parameters included in the fit (apart from those representing the MC statistics) with MADGRAPH as preferred signal generator. The top plot refers to the combined muon and electron fit, the middle plot to the muon only fit, and the bottom plot to the electron only fit.
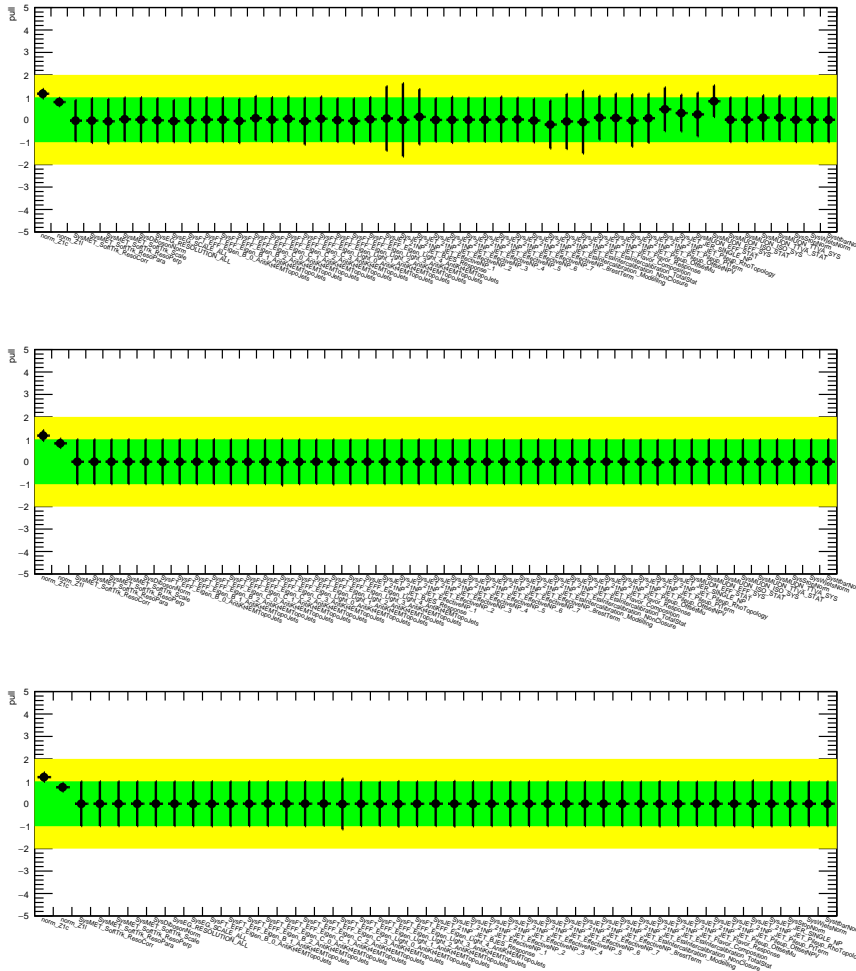
Figure C.4: Pulls of all the nuisance parameters included in the fit (apart from those representing the MC statistics) with ALPGEN as preferred signal generator. The top plot refers to the combined muon and electron fit, the middle plot to the muon only fit, and the bottom plot to the electron only fit.

On the other hand, ALPGEN presents in general less correlations than the other two generators. In particular:

Zb tends to be correlated to the flavour tagging uncertainty component C0, and the jet flavour composition parameter;

Z+c tends to be anticorrelated with Z+l, flavour tagging uncertainty components C0, C3 and light2;

Z+l tends to be anticorrelated with Z+c, and the flavour tagging uncertainty component light3, and correlated with the flavour tagging uncertainty components light0 and C3.

The correlations between the different Z+jets process normalisations is not surprising, neither correlations of those with the flavour tagging uncertainties. In this case, there is no strong correlation with any of the MC statistical uncertainties on the fitted bins.

C.2   Flavour fit for Z+2 b-jets analysis

The fit on Z+jets components is performed on MADGRAPH and ALPGEN samples, by using the technique described in Section 4.6.3. In the Z+2 b-jets analysis, the variable to fit is given by the the sum of the b-tagging weights of the two b-jets, distributed in 6 bins. Fig.C.7 and Fig.C.8 show the pre- and post-fit distributions for the MADGRAPH and ALPGEN generators respectively.

A deep understanding of the scale factors obtained from the fits (see Section 4.6.3) can be obtained by looking at the pulls and correlations of nuisance parameters. When MADGRAPH is used as baseline, single lepton channel fits are quite stable, even with a large number of constraints on the JES parameters, as can be noticed in Fig.C.9. The combined fit shows some instabilities, with JER and muon reconstruction efficiency variations close to 1σ.
In Fig.C.10 ALPGEN is used for the Z+jets modelling and presents single lepton channel fits rather stable. The combined result is preatty stable too,
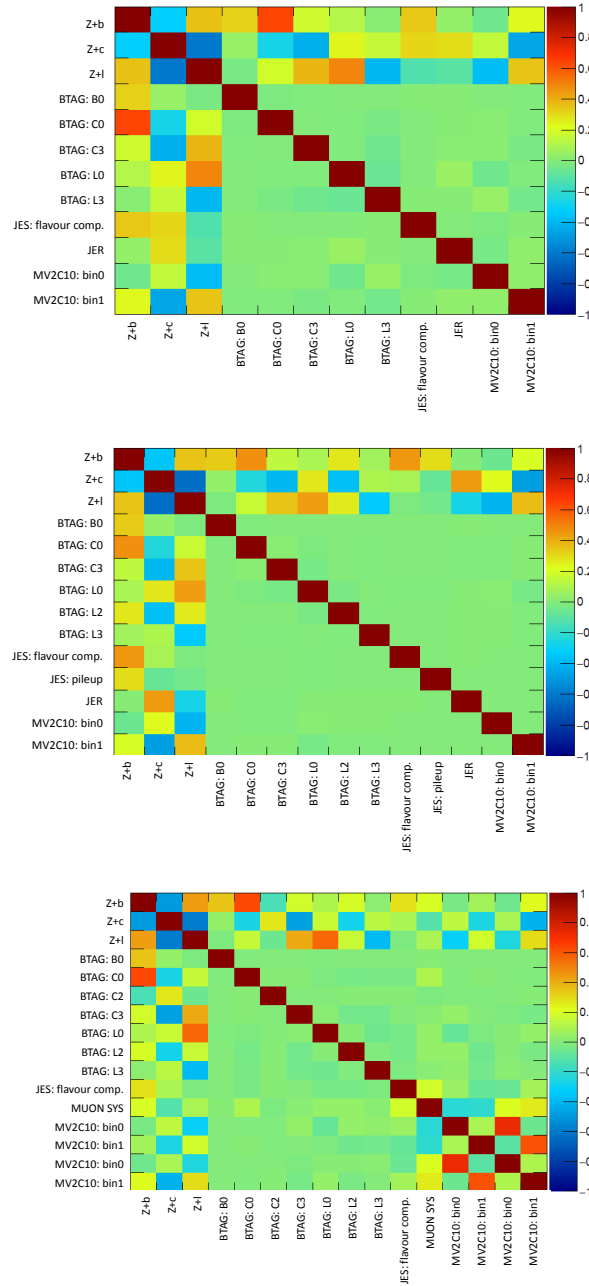
Figure C.5: Correlations of the nuisance parameters included in the fit that have at least a correlation which is in absolute value larger than 50% with any other parameter. MADGRAPH is the signal generator used in the electron only (top), muon only (middle) and combined (bottom) fits.
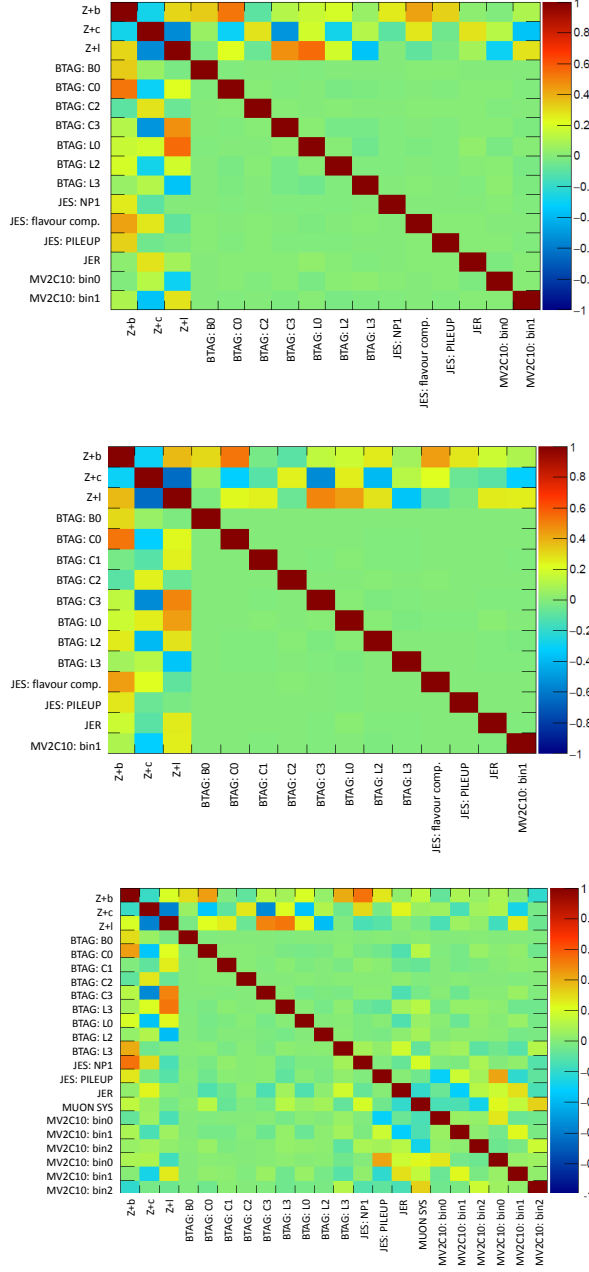
Figure C.6: Correlations of the nuisance parameters included in the fit that have at least a correlation which is in absolute value larger than 50% with any other parameter. ALPGEN is the signal generator used in the electron only (top), muon only (middle) and combined (bottom) fits.
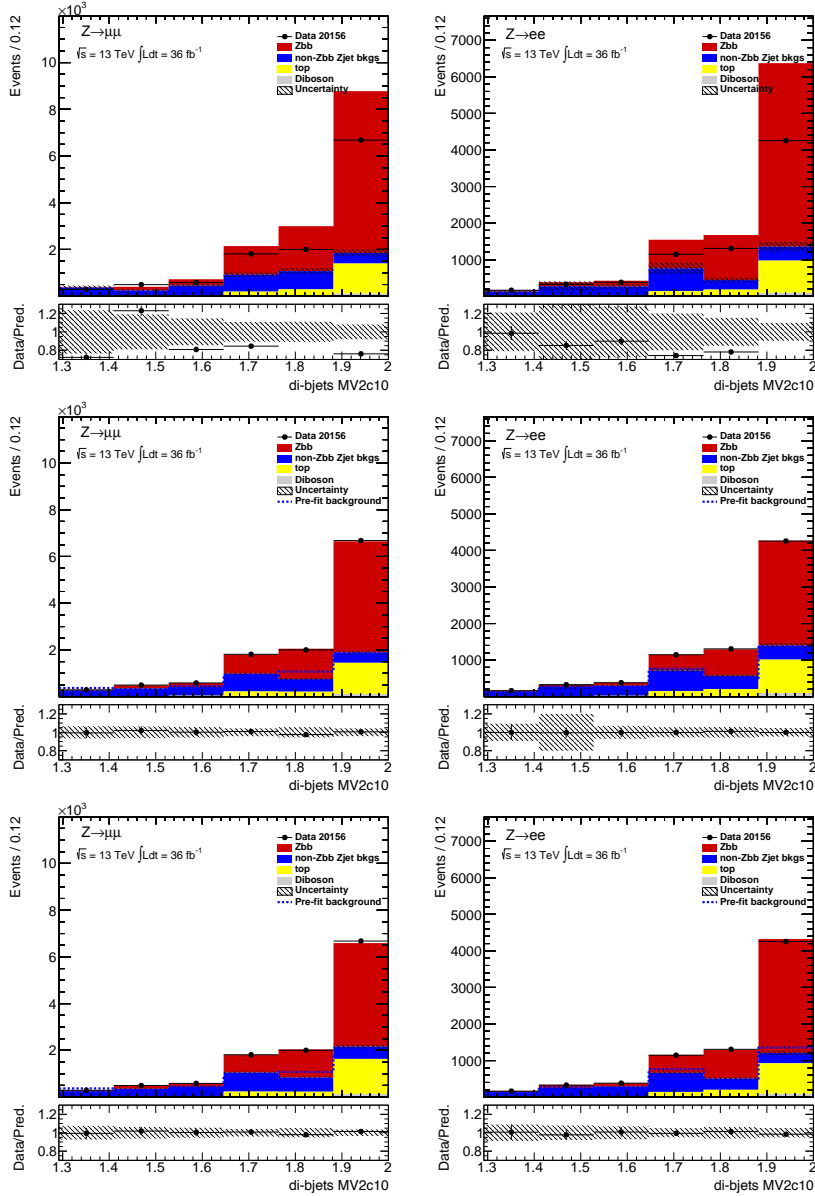
Figure C.7: Pre- (top row) and post-fit (middle and bottom row) plots for the muon (left) and electron (right) channels, using the Madgraph generator as baseline for the Z+2 b-jets analysis. The fits shown in the middle row are performed separately in the electron and muon channels, while the ones shown in the bottom row are performed combining the two channels together.
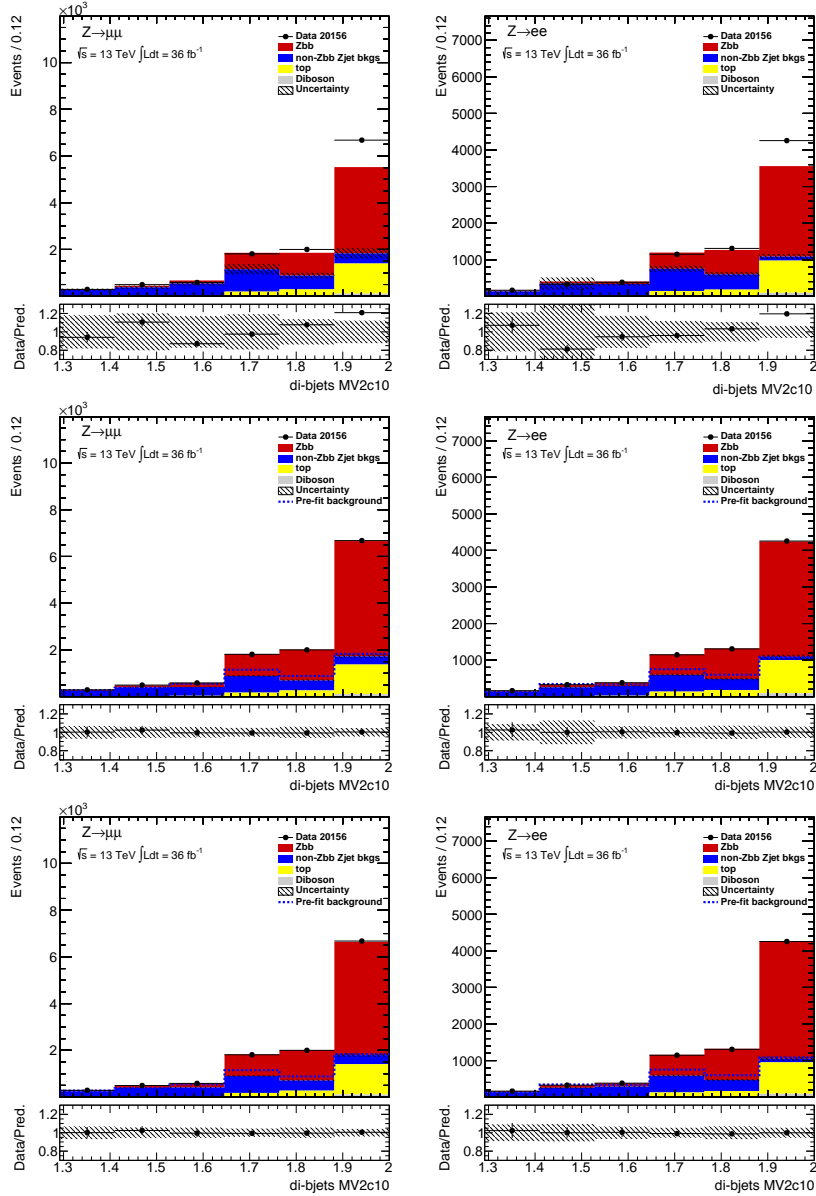
Figure C.8: Pre- (top row) and post-fit (middle and bottom row) plots for the muon (left) and electron (right) channels, using the ALPGEN generator as baseline for the Z+2 b-jets analysis. The fits shown in the middle row are performed separately in the electron and muon channels, while the ones shown in the bottom row are performed combining the two channels together.

with only one pull due to the C0 components of the uncertainty related to b-jet calibration.
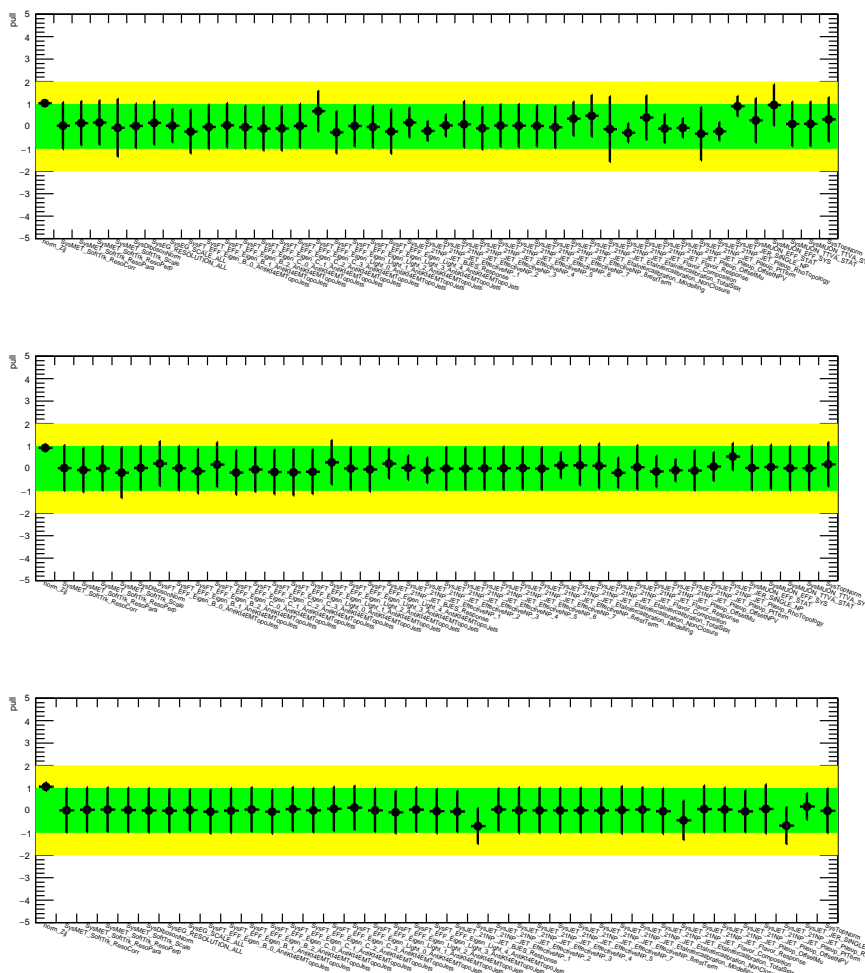

Figure C.9: Pulls of all the nuisance parameters included in the fit (MC statistics excluded) with MADGRAPH as baseline generator for Z+2 b-jets analysis. The top plot refers to the combined muon and electron fit, the middle plot to the muon only fit, and the bottom plot to the electron only fit.

Figures C.11-C.12 show the correlations among nuisance parameters for all the fits performed by using MADGRAPH and ALPGEN respectively. For both generators, the Zbb signal has strong correlation (∼70%) with the B0 component of the b-jets calibration and the Zjj background normalisation presents a evident correlation with the L0 parameter. In addition, the Zbb
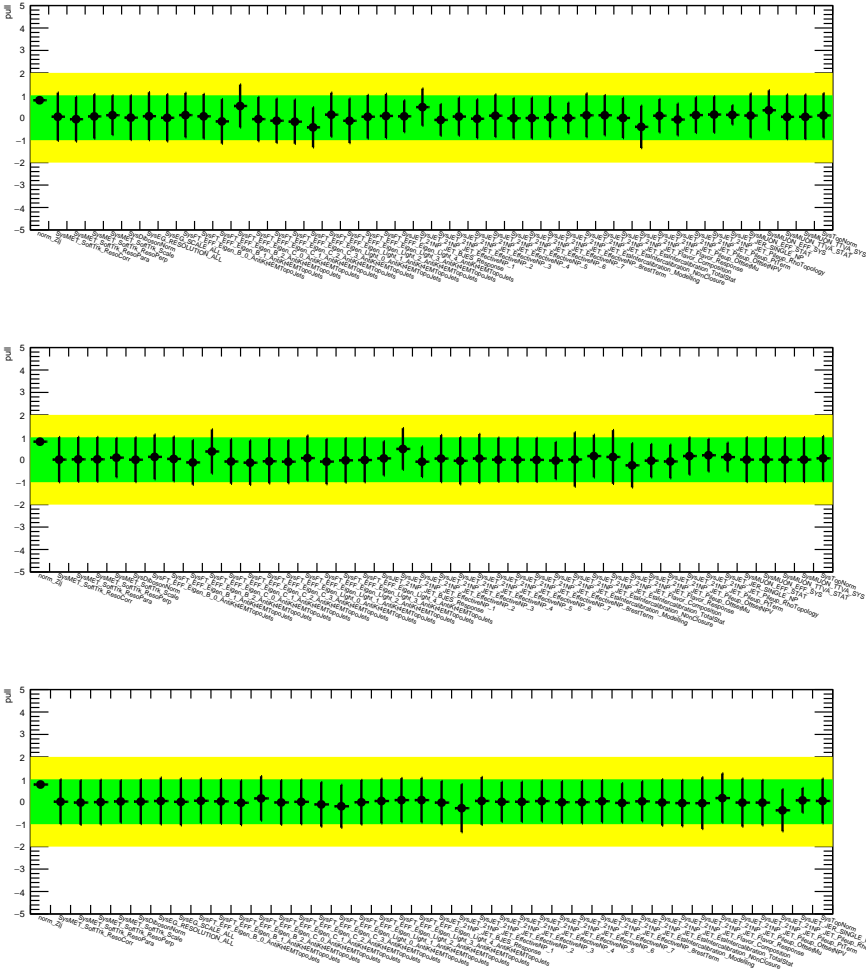
Figure C.10: Pulls of all the nuisance parameters included in the fit (MC statistics excluded) with ALPGEN as baseline generator for Z+2 b-jets analysis. The top plot refers to the combined muon and electron fit, the middle plot to the muon only fit, and the bottom plot to the electron only fit.

signal estimated with ALPGEN has a large correlation with the flavour component of JES systematics.

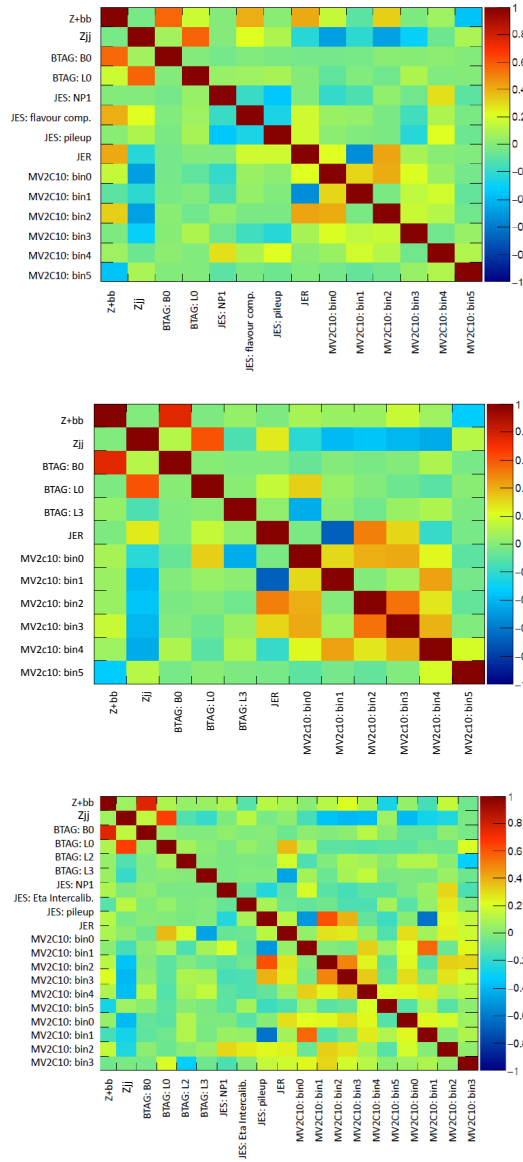Figure C.11: Correlations of the nuisance parameters in the Z+2 b-jets analysis. MADGRAPH is the signal generator used in the electron only (top), muon only (middle) and combined (bottom) fits. Only parametrs with a correlation with an absolute value larger than 50% are shown.

Figure C.12: Correlations of the nuisance parameters in the Z+2 b-jets analysis. ALPGEN is the signal generator used in the electron only (top), muon only (middle) and combined (bottom) fits. Only parametrs with a correlation with an absolute value larger than 50% are shown.
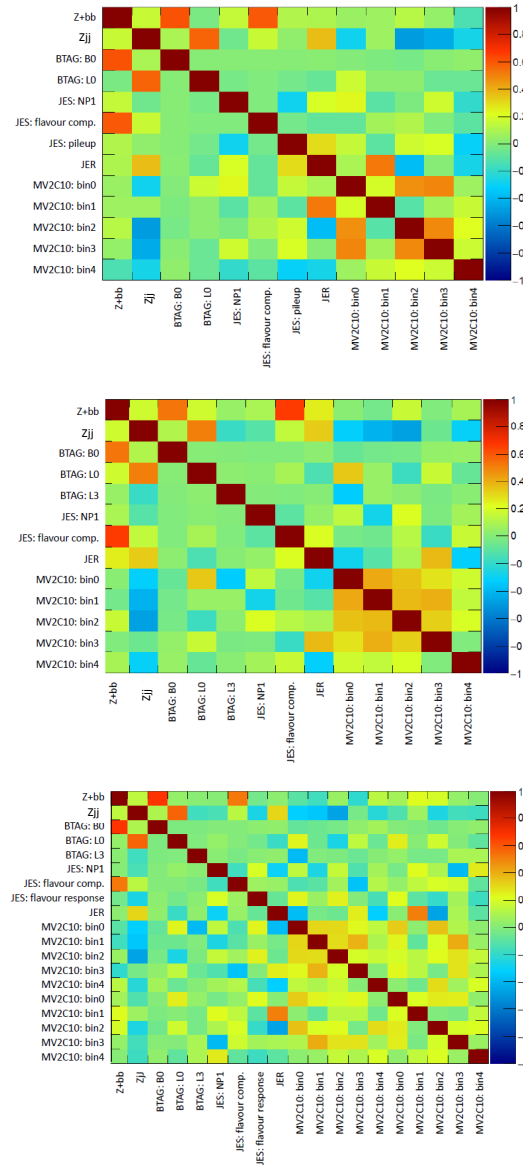
Bibliography

[1] C. Quigg, *Gauge Theories of the strong, weak and electromagnetic interactions*, Westview Press (1997).

[2] F.Halzen, A.D.Martin, *Quarks and leptons: an introduction course in modern particle physics*, John Wiley & Sons (1984).

[3] K. A. Olive et al., *Particle Data Group*, Chin. Phys. C 38 (2014) 090001 and 2015 update, http://pdg.lbl.gov/.

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

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

[6] A. Salam, *Elementary Particle Theory*, Ed. N. Svarholm (1968).

[7] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and Collider Physics*, Cambridge University Press (2003).

[8] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nuclear Physics, B44, 1 (1972) 189–213.

[9] R. Devenish and A. Cooper-Sarkar, *Deep Inelastic Scattering*, Oxford University Press (2004).

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

[11] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys., B126 (1977) 298.

[12] V. Radescu, *Review on PDFs*, QCD@LHC2016, https://indico.cern.ch/event/516210/contributions/2212433/attachments/1325679/1990171/QCDatLHC_VR.pdf.

[13] SHERPA, *Sketch of hadron hadron collisions*, https://www.opensciencegrid.org/wp-content/uploads/2014/05/event.jpg.

[14] Torbjon Sjostrand, *Monte Carlo Generators*, CERN-LCGApp-2006-06 [arXiv:0611247v1].

[15] Z. Marshall, *Simulation of Pile-up in the ATLAS Experiment*, J. Phys. Conf. Ser. 513 (2014) 022024.

[16] UA2 Collaboration (P. Bagnaia et al.), *Evidence for $Z \to e^+e^-$ at the CERN anti-p p Collider*, Phys. Lett. 129B (1983) 130-140.

[17] UA1 Collaboration (G. Arnison et al.), *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, Phys.Lett. 122B (1983) 103-116.

[18] UA1 Collaboration (G. Arnison et al.), *Observation of muonic Z^0 decay at the $\bar{p}p$ collide*, Phys. Lett. B 147 (1984) 241-248.

[19] ALEPH Collaboration (D. Decamp et al.), *Determination of the Number of Light Neutrino Species*, Phys. Lett. B231 (1989) 519-529.

[20] C. Patrignani et al., *Particle Data Group*, Chin. Phys. C, 40, 100001 (2016).

[21] CDF Collaboration (A. Abulecia et al.), *Measurements of Inclusive W and Z cross sections in $p\bar{p}$ collisions at $\sqrt{s} =1.96$ TeV*, J. Phys. G34 (2007) 2457.

[22] D0 Collaboration, *Measurement of W and Z boson production cross sections*, Phys.Rev.D60:052003,1999.

[23] S. Catani et al., *Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO*, Phys. Rev. Lett. 103:082001 (2009).

[24] ATLAS Collaboration, *Determination of the strange quark density of the proton from ATLAS measurements of the $W \to l\nu$ and $Z \to ll$ cross sections*, Phys.Rev.Lett. 109 (2012) 012001.

[25] ATLAS Collaboration, *Measurement of the inclusive $W^{\pm}$ and Z/γ cross sections in the electron and muon decay channels in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D85, 072004 (2012).

[26] ATLAS Collaboration, *Precision measurement and interpretation of inclusive W^+, W^- and $Z\gamma^*$ production cross sections with the ATLAS detector*, Eur. Phys. J. C 77 (2017) 367.

[27] ATLAS Collaboration, *Measurements of top-quark pair to Z-boson cross-section at $\sqrt{s} = 13$, 8, 7 TeV with the ATLAS detector*, JHEP 02 (2017) 117.

[28] 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. C77 (2017) 361.

[29] J. M. Campbell, R. K. Ellis, F. Maltoni and S. Willenbrock, *Production of a Z boson and two jets with one heavy-quark tag*, Phys. Rev. D 73 (2006) 054007.

[30] J. M. Campbell, R. K. Ellis, F. Maltoni and S. Willenbrock, *Associated pro- duction of a Z Boson and a single heavy quark jet*, Phys. Rev. D 69 (2004) 074021.

[31] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a $b\bar{b}$ pair in pp collisions at 13 TeV using the ATLAS detector*, ATLAS-CONF-2016-091.

[32] ATLAS Collaboration, *A measurement of the ratio of the production cross sections for W and Z bosons in association with jets with the AT-LAS detector*, Eur. Phys. J. C (2014) 74: 3168.

[33] CDF Collaboration, *Measurment of cross section for b jet production in events with a Z boson in $p\bar{p}$ collisions at $\sqrt{s}$ =1.96 TeV*, Phys. Rev. D 79 (2009) 052008.

[34] D0 Collaboration, *Measurement of the ratio of differential cross sections $\sigma(p\bar{p} \to Z + bjets)/\sigma(p\bar{p} \to Z + jets)$ in $p\bar{p}$ collisions at $\sqrt{s}$ =1.96 TeV*, Phys. Rev. D 87 (2013) 092010.

[35] ATLAS Collaboration, *Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector*, JHEP10(2014)141.

[36] CMS Collaboration, *Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s}$ =7 TeV*, JHEP 06 (2014) 120.

[37] CMS Collaboration, *Measurements of the associated production of a Z boson and b jets in pp collisions at $\sqrt{s}$ =8 TeV*, Submitted to Eur. Phys. J. C [arXiv:1611.06507].

[38] J.M. Campbell and R.K. Ellis, *MCFM for the Tevatron and the LHC*, Nucl. Phys. Proc. Suppl. 205-206 (2010) 10.

[39] Lyndon Evans and Philip Bryant, *LHC Machine*, JINST, 3:S08001, 2008.

[40] High Luminosity LHC project, http://hilumilhc.web.cern.ch/about/hl-lhc-project.

[41] Luminosity Public Results Run-2, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

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

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

[44] ATLAS Collaboration, *Insertable B-Layer Technical Design Report* , ATLAS-TDR-019.

[45] F. Hugging, *The ATLAS Pixel Insertable B-Layer (IBL)*, arXiv:1012.2742v1.

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

[47] ATLAS Collaboration, *ZeroDegreeCalorimeter*, https://twiki.cern.ch/ twiki/bin/viewauth/Atlas/ZeroDegreeCalorimeter.

[48] ATLAS Collaboration, *ATLAS Forward Protons*, http://www-hep.uta.edu/ brandta/ATLAS/AFP/AFP.html.

[49] ATLAS Collaboration, *Measurement of the total cross section from elastic scattering in p-p collisions at $\sqrt{s}$= 7 TeV with the ATLAS detector*, Nucl. Phys. B 889 (2014) 486-548.

[50] ATLAS Collaboration, *The ATLAS Data Acquisition and High Level Trigger system*, JINST 11 (2016) P06008.

[51] T. Cornelissen, et al. Concepts, *Design and Implementation of the ATLAS New Tracking (NEWT)*, ATL-SOFT-PUB-2007-007, 2007.

[52] ATLAS Collaboration, *ATLAS event at 900 GeV - 6 May 2015 - Run 264034 Event 11475271*, ATLAS-PHO-Event-2015-007.

[53] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, ATLAS-CONF-2016-024.

[54] J. Illingworth and J. Kittler, *A survey of the Hough transform*, Computer Vision, Graphics, and Image Processing, 44 (1988) 87-116.

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

[56] ATLAS Collaboration, *Jet kinematic distributions in proton–proton collisions at $\sqrt{s}$ = 900 GeV with the ATLAS detector*, ATLAS-CONF-2010-001.

[57] G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP, 0804:063, 2008.

[58] ATLAS Collaboration, *ATLAS Calorimeter Response to Single Isolated Hadrons and Estimation of the Calorimeter Jet Scale Uncertainty*, ATLAS-CONF-2011-028.

[59] ATLAS Collaboration, *Determination of the ATLAS jet energy measurement uncertainty using photon-jet events in proton-proton collisions at $\sqrt{s}$=7 TeV*, ATLAS- CONF-2011-03.

[60] ATLAS Collaboration, *Probing the measurement of jet energies with the ATLAS detector using Z+jet events from proton-proton collisions at $\sqrt{s}$=7 TeV*, ATLAS-CONF- 2012-053.

[61] ATLAS Collaboration, *Tagging ans suppression of pile-up jets with the ATLAS detector*, ATLAS-CONF-2014-018.

[62] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATL-PHYS-PUB-2015-022.

[63] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012.

[64] ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, arXiv:1512.01094v2.

[65] ATLAS Collaboration, *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s}$=13 TeVß*, ATL-PHYS-PUB-2015-023.

[66] ATLAS Collaboration, *ATLAS 2015 Z Counting Analysis*, ATL-COM-PHYS-2015-1408.

[67] ATLAS Collaboration, *Measurement of $W^{\pm}$ and Z-boson production cross sections in pp collisions at $\sqrt{s} =$13 TeV with the ATLAS detector*, Phys. Lett. B 759 (2016) 601.

[68] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP 11 (2007) 070.

[69] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP 06 (2010) 043.

[70] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D82 (2010) 074024.

[71] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. 178 (2008) 852–867.

[72] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} =$7 TeV with the ATLAS detector*, JHEP 1409 (2014) 145.

[73] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A462 (2001) 152.

[74] N. Davidson, T. Przedzinski and Z. Was, *PHOTOS Interface in C++: Technical and Physics Documentation (2010)*, arXiv: 1011.0937.

[75] T. Sjostrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP 05 (2006) 026.

[76] C. Anastasiou et al., *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, Phys. Rev. D69 (2004) 094008.

[77] A. D. Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. C63 (2009) 189–285.

[78] 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.

[79] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, ATLAS-PHYS-PUB-2012-003.

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

[81] ATLAS MC production, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AtlasProductionGroupMC15a.

[82] S. Catani and M. Grazzini, *An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC*, Phys. Rev. Lett. 98 (2007) 222002.

[83] K. Melnikov and F. Petriello, *Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$*, Phys. Rev. D 74 (2006) 114017.

[84] S. Dulat et al., *The CT14 Global Analysis of Quantum Chromodynamics*, arXiv: 1506.07443 (2015).

[85] J. Gao et al., *CT10 next-to-next-to-leading order global analysis of QCD*, Phys.Rev. D89.3 (2014) 033009.

[86] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP 04 (2015) 040.

[87] L. A. Harland-Lang et al., *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C75.5 (2015) 204.

[88] S. Alekhin, J. Bluemlein and S. Moch, *The ABM parton distributions tuned to LHC data*, Phys. Rev. D89.5 (2014) 054028.

[89] H. Abramowicz et al., *Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data*, Eur. Phys. J. C75.12 (2015) 580.

[90] ATLAS Collaboration, *Determination of the strange quark density of the proton from ATLAS measurements of the $W \to l\nu$ and $Z \to ll$ cross sections*, Phys. Rev. Lett. 109 (2012) 012001.

[91] K. A. Olive et al., *Review of Particle Physics*, Chin. Phys. C38 (2014) 090001.

[92] CMS Collaboration, *Measurement of the Inclusive W and Z Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV*, JHEP 10 (2011) 132.

[93] ATLAS Collaboration, *Measurement of the inclusive $W \pm$ and $Z/gamma$ cross sections in the electron and muon decay channels in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D85 (2012) 072004.

[94] ATLAS, Collaboration, *Standard Model Derivations*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/StandardModelASG.

[95] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP 02 (2009) 007.

[96] S. Hoche et al., *QCD matrix elements and truncated showers*, JHEP 05 (2009) 053.

[97] S. Hoche et al., *QCD matrix elements + parton showers: The NLO case*, JHEP 04 (2013) 027.

[98] M. L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP 07 (2003) 001.

[99] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. D 82 (2010) 074018.

[100] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP 07 (2002) 012.

[101] J. Alwall et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, Eur. Phys. J. C53 (2008) 473.

[102] J. Alwall et al., *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.

[103] ATLAS Collaboration, *Measurement of the differential cross-sections of prompt and non-prompt production of J/Ψ and Ψ(2S) in pp collisions at $\sqrt{s}$ =7 and 8 TeV with the ATLAS detector*, Eur. Phys. J. C 76(5), 1-47, (2016).

[104] ATLAS Collaboration, *Flavour Tagging Twiki*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/FlavourTaggingLabeling.

[105] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s}$ = 7 TeV*, Eur. Phys. J. C, 73 3 (2013) 2304.

[106] 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 3 (2013) 2306.

[107] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s}$= 8 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C 76 (2016) 653.

[108] G. D'Agostini, *A Multidimentional Unfolding Methos Based on Bayes' Theorem*, Nucl. Instrum. Meth. A362 (1995) 487-498.

[109] L. Alberghi et al., *Choice and Characterization of photomultipliers for the ATLAS LUCID detector*, JINST 11 P05014 (2016).

[110] ATLAS Collaboration, *Measurement of muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data*, Eur. Phys. J. C (2014) 74:3130.

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

Acknowledgements

Al termine di questo lungo percorso, mi sento di ringraziare tutti coloro che in qualche modo hanno lasciato un segno.

Ringrazio la mia famiglia, che mi ha sostenuta sempre, con grande fiducia. La ringrazio perchè, oltre all'affetto, ha stimolato la mia curiosità e le mie opinioni, senza impormi i suoi insegnamenti. Grazie per avermi trasmesso il valore della libertà del mio pensiero.

Ringrazio Francesco, che oltre ad essere il mio coinquilino preferito, stimola il mio percorso e la mia crescita. Grazie per aver capito che dietro un fine settimana a lavorare, c'è molto di più di un semplice senso del dovere, c'è la passione per il mio mestiere. Grazie per ironizzare sulla mia spigolosità ed essere sempre al mio fianco.

Ringrazio la nonna Maria, che mi sognava professoressa, e oggi in qualche modo lo sono diventata.

Ringrazio Grazia e Giulia, perchè sono le mie compagne. Sanno quando chiedere e quando lasciare perdere, quando sgridarmi e quando consolarmi, quando spronarmi e quando trattenermi. Sanno cosa voglio prima che lo sappia io e sono pronte ad accogliermi quando sbaglio. Vi voglio bene singolarmente, ma insieme siamo qualcosa di speciale.

Ringrazio l'Ufficio 161 bis, l'unico e imbattibile, che con Silvia e Federica raggiunge la sua migliore configurazione. Grazie ragazze per le follie delle 6, che a volte sono le 2, per il muro del pianto, per i bigliettini sul muro, per

le rassicurazioni, le risate, i canti, i balli, lo stretching, i video, le birre e il fumo passivo.

Ringrazio Viola per essere la parte metodicamente pazza della troppia. Ringrazio Serena, Chiara, Marina, Matteo e Federico per incoraggiarmi e farmi sentire speciale e Giacomo per farmi sentire quasi normale.

Ringrazio Riccardo e Renato, per le presenze e i vuoti, per i menti e i nasi, per i concerti e le briscole, i balli e i dispetti.

Ringrazio Enzo per i suoi "Stupida Camilla", sempre molto pertinenti, e il Betto per la sua ironia fantozziana. Ringrazio Shak, Fabio, Pozzi, Gabri, Zan e tutti gli amici di questi lunghi anni.

Ringrazio il gruppo ATLAS di Bologna, con cui condivido un percorso dalla laurea triennale. In particolare, Sara e Roberto, che mi hanno spinta e accompagnata fino al dottorato.

Ringrazio Laura e Benedetto per avermi fatta crescere in questi anni dal punto di vista professionale e personale. Le nostre discussioni hanno stimolato in me un pensiero via via più critico. Li ringrazio soprattutto per aver sopportato la mia testardaggine, pazientato davanti alle polemiche, gestito le mie insicurezze e frenato le mie ansie.

Ringrazio il gruppo ATLAS di analisi, quindi Evelin, Semen, Chiara, Federico, Valentina e Doug. In particolare sono grata a Evelin per avermi resa piùù consapevole sulla qualità del mio lavoro, e Semen per le nottate passate insieme a lavorare, per il sostegno, l'aiuto, le risate e la pazienza di fronte alle mie incessanti domande.