Simulation of the Measurement of the Inclusive Jet Cross Sections in Z($\rightarrow e^+e^-/ \rightarrow \mu^+\mu^-$)+jets Events in pp Collisions at 14 TeV with the ATLAS experiment[1]

Ester Segura i Solé

Institut de Física d'Altes Energies
Universitat Autònoma de Barcelona
Departament de Física
Edifici Cn E-08193 Bellaterra (Barcelona)

June 2009

supervised by
Martine Bosman
IFAE
Edifici Cn E-08193 Bellaterra (Barcelona)

[1] Ph.D.Dissertation

Als meus pares

This PhD thesis is based on the work realized over the last two years in the ATLAS collaboration. It has not been possible to include analysis based on data. Therefore, the analysis shown here it is an exercise previous to real data analysis. It shows first results comparing different Monte Carlo predictions. It includes hints and ideas on how to look at signal and the corresponding background events, and how they can be treated. In summary, there is a first complete example on how to analyze Z+jets events, including $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$, for testing pQCD predictions and as a background of other new physics channels. Part of this analysis has been included in an ATLAS CSC note, and it has also been presented in many conferences during summer 2008. Now it is time to perform the analysis on real ATLAS data, and I hope this thesis will be useful for those performing the new-coming analysis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

After 20 years of preparation, the Large Hadron Collider is going to be switched on in late 2009, smashing protons at an energy of 14 (10) TeV, to recreate the first moments after the Big Bang. Particles will whizz around a circular tunnel of 27 km in circumference at near light speed. The tunnel, built near Geneva, and its experiments constitute one of the largest coordinated efforts ever made to study the fundamental structure of nature. It is expected that at the energies reached in the proton-proton collisions at the LHC, unknown physical phenomena will have to occur and will be observable. Among the particle debris may lie evidence for extra dimensions, mysterious dark matter that pervades the universe or the Higgs boson, which gives mass to elementary particles.

ATLAS is one of the LHC experiments. Besides the new phenomena goals in its physics program, there is the understanding of the already known physics. The better understanding of perturbative Quantum Chromodynamics theory is one of the aims of ATLAS. Quantum Chromodynamics is the field theory which describes the strong interaction between quarks and gluons. It remains an "unsolved" theory, since no single approximation method can be used to all length scales. Perturbative QCD naturally describes a large set of high-energy, large-momentum-transfer cross sections and its formalism has provided an invaluable tool in the study of the strong interactions.

The most prominent signature of QCD at hadron colliders is the production of collimated jets of hadrons. The measurement of the production of such jets in association with a vector boson, W or Z, provides a stringent test of perturbative QCD (pQCD) calculations. Furthermore, some of new physics processes at hadron colliders, such as the production of Higgs bosons and supersymmetric particles, can be mimicked by the production of vector bosons in association with jets that constitute irreducible backgrounds to these searches.

This PhD. thesis presents the measurement of the inclusive jet cross section in $Z \rightarrow e^+e^-$ and in $Z \rightarrow \mu^+\mu^-$ events, comparing theory predictions with "real data", i.e. Monte Carlo fully-reconstructed events, for the first 1fb^{-1} of data at the ATLAS detector. Reconstructed corrected data is compared to next-to-LO (NLO) and LO pQCD predictions. Perturbative predictions are corrected for the contributions of the non-perturbative processes, like the underlying event and the fragmentation of the partons into jets of hadrons. These processes are not described by perturbation theory and must be estimated using phenomenological models. Two different reconstructed data are used, PYTHIA and ALPGEN Monte Carlo data. Comparisons of both Monte Carlo predictions are studied. Background processes are estimated proposing different data-driven methods to be applied to real data. ATLAS Cone 0.4 algorithm is used to look for jets in the events after identifying the presence of a Z boson through the reconstruction of its decay (electrons or muons). Reconstructed data is corrected for detector effects, using independent factors.

As this work was carried out before the "physics-data" start of the LHC, the presented studies are based on Monte Carlo simulations. During the preparation of a high-energy collider experiment, such simulations are important to develop efficient strategies for data analysis and for the reconstruction of the physics objects observed with the detectors.

The outline of this thesis is as follows. In the Chapter 2, the main features of the LHC collider and the ATLAS experiment are reviewed. Chapter 3 is devoted to the description of the theory of strong interactions and jet phenomenology at hadron colliders. Chapters 4 and 5 describes the analysis in detail. Final results are compiled and discussed in chapter 6. Chapter 7 is devoted to the conclusions.

In addition, at the end of the document five appendixes are included. Appendix A presents some extra table results. Appendix B details information about MCFM program. Appendix C contains a full study of jet algorithms for Z+jets events. Unfolding procedure and corrections factors for $Z \rightarrow \mu\mu$+jets events are contained in Appendix D. And finally, Appendix E summarizes results coming from other experiments (CDF and D0 at Tevatron).

Chapter 2

The LHC and the ATLAS Experiment

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [1] is installed in the 27 km long former LEP tunnel situated at CERN, Geneva (Switzerland). It will accelerate two counter-rotating beams of protons, delivered by the Super Proton Synchrotron (SPS). Collisions will take place at four interaction points where detectors are located. These include Point 1 (ATLAS detector), Point 2 (ALICE detector), Point 5 (CMS detector) and Point 8 (LHCb detector) as shown in figure 2.1 and figure 2.2.

(a) (b)

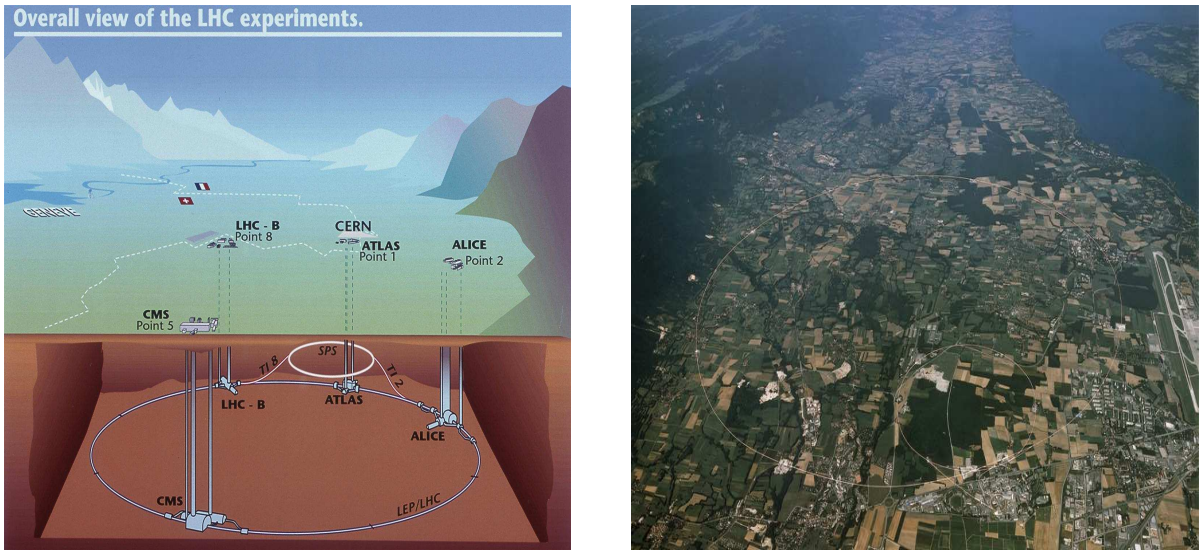


Figure 2.1: Overall view of the LHC experiments (a). Aerial view of the LHC, near Geneva (b).

The LHC will collide proton beams at energies of 7 TeV and at a peak luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$, aiming at an annual integrated luminosity of $\sim 100\,\text{fb}^{-1}$. These values are one or more orders of magnitude higher than what has been achieved by any previous experiment. The current highest-energy accelerator, the Tevatron at Fermilab, collides proton against anti-proton beams at a center-of-mass energy of 1.9 GeV and has collected $\sim 5\,\text{fb}^{-1}$ over its ten-year period of operation. The performance requirements of the LHC set significant challenges in the design and construction of the accelerator. To bend 7 TeV protons around the ring, 1,232 LHC dipoles (figure 2.3) are used, which cover $\sim$ 20 km of the

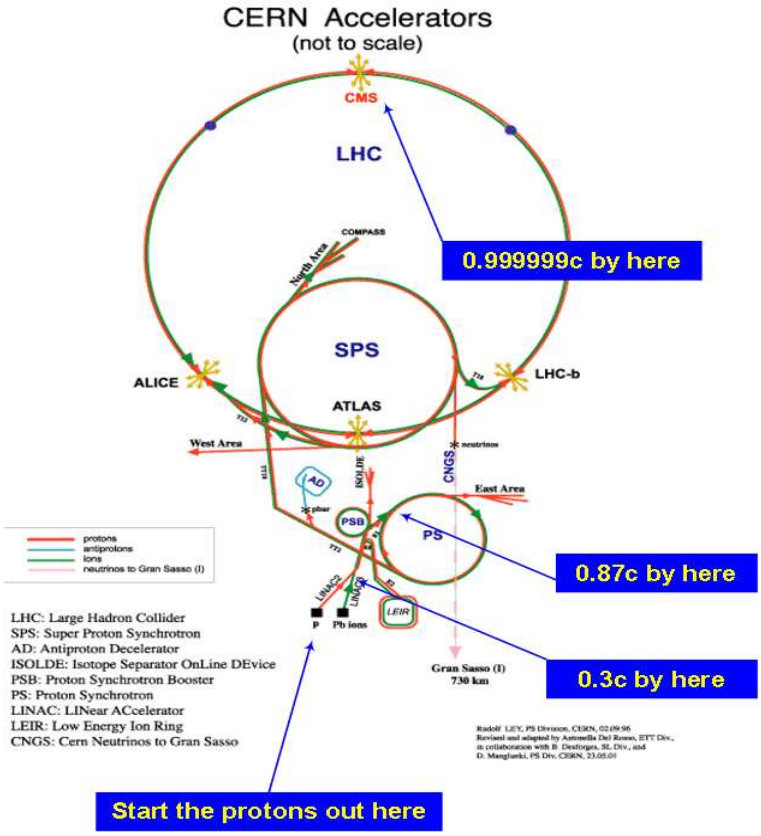


Figure 2.2: Schematic view of the LHC and the SPS accelerator ring, where the different interactions points and the corresponding detectors are shown.

ring. The beams are focused using quadrupole magnets to boost the luminosity at the collision points. 392 quadrupole magnets are used in the straight sections of the ring. The dipole magnets must produce magnetic fields of 8.36 Tesla. Such a high field is produced using niobium-titanium super-conducting magnets and super-fluid helium[1] is used for cooling to maintain the operation temperature of 1.9 K. The Tevatron accelerator reaches 4.5 Tesla at 4.2 K. HERA at DESY reaches 5.5 Tesla. Both use the Nb-Ti technology invented in the 1960s at the Rutherford-Appleton Lab.

Hadron colliders can produce high energy collisions much more efficiently than electron colliders, as synchroton radiation is much lower. The energy dissipated by the accelerated particles due to synchroton radiation in an accelerator ring of radius R is

$$\delta E = \frac{4\pi e^2}{3R}\beta^3\gamma^4$$

per revolution, where $v = \beta c$ and $E = \gamma mc^2$. If the particles are relativistic, then the γ^4 factor becomes dominant and electron colliders suffer from a large radiation loss. For example, 50 GeV electrons have a γ of 98,000 while protons would have a γ of 54 for the same energy.

Enormous hadronic activity in proton collisions generally creates "messy" events with large number of particles. It is therefore not the most appropriate environment for precision measurements of known physics features (although precision measurements will be carried on, ie. m_W, m_{top}, etc) and the focus of the physics programmes tend to be searches for signatures of new physics. Such new physics which potentially has large implications for our understanding of the universe typically relies on the availability of large amounts of energy.

[1]For LHC, 12 million liters of liquid nitrogen are vaporized during the initial cool-down of 37,000 tons of equipment for the LHC. The total inventory of liquid helium is 700,000 liters.

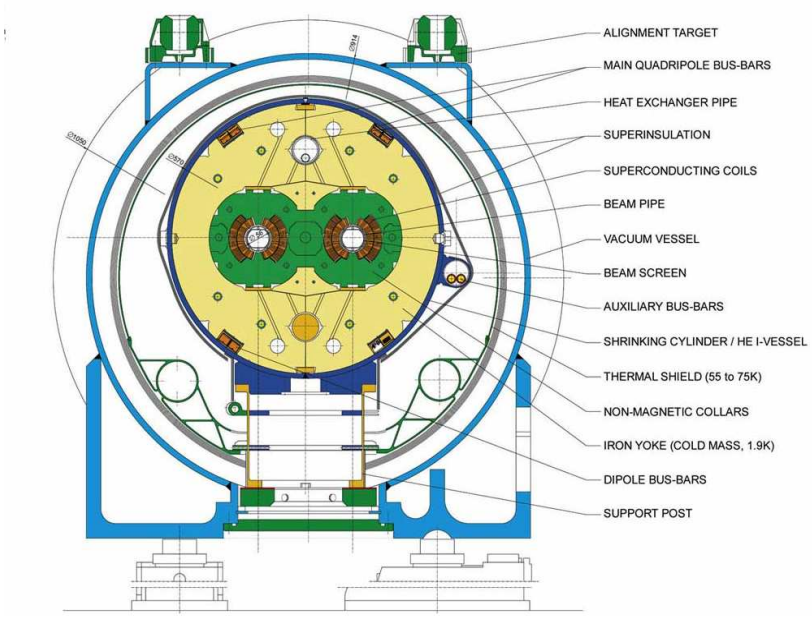


Figure 2.3: Cross section of the LHC beam-pipe with dipole magnet.

Table 2.1 summarizes some of the important parameters of the LHC proton beam. The LHC will operate partly in proton-proton mode but will also collide lead nuclei to study heavy ion collisions. The study presented in my thesis only considers proton-proton collisions. The current operational plan is to have first collisions at late 2009 at 10 TeV.

Parameter	unit	value
Ring circumference	[m]	26658.883
Number of particles per bunch		1.15×10^{11}
Number of bunches		2808
Beam energy	[GeV]	7000
Relativistic gamma		7461
Peak luminosity	$[\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	10^{34}
RMS Beam size at IP1	[μm]	16.7
Inelastic cross section	[mb]	60
Events per bunch crossing		19

Table 2.1: LHC beam parameters.

2.2 The ATLAS Detector

In case of curiosity, ATLAS is an acronym for **A** **T**oroidal **L**HC **A**pparatu**S**. It is a collaboration of about four thousand physicists from more than thirty nations, who gathered to build one of the largest experiment in the history of mankind. The ATLAS experiment was designed to exploit the full physics potential of the LHC and it is supposed to be in operation for roughly two decades. This implies basic design requirements, which satisfy a broad field of specifications. The detector and its electronics must

work reliably in a high radiation environment and provide precise measurements of various physical quantities. Theses quantities are crucial to discover and study new and interesting physics signatures and processes.

2.2.1 Overall Concept

The overall detector layout, shown in figure 2.4, is incredibly complex and described in detail in [2]. For completeness, in this chapter a brief description of the detector components is given.

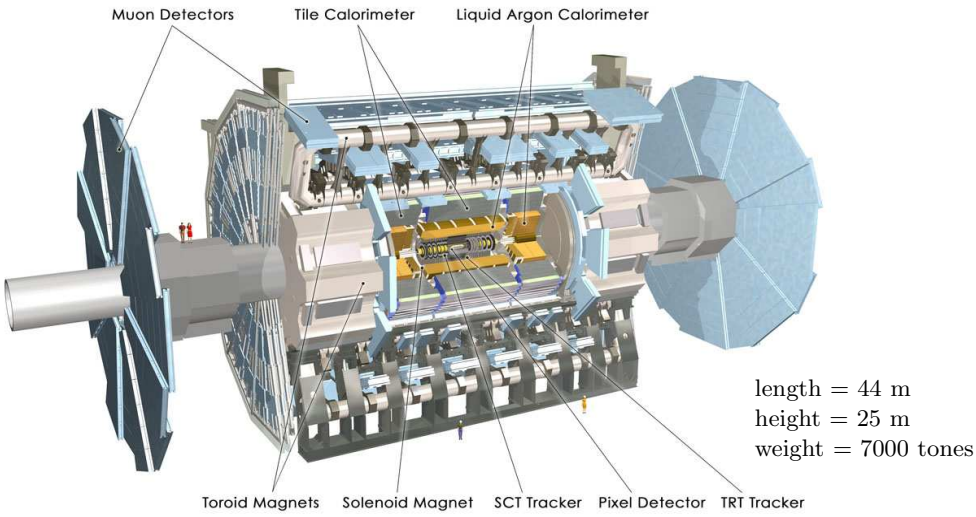


Figure 2.4: Overview of the ATLAS detector layout.

To support the physics programme described in section 2.5, a number of requirements have been set for the detector building:

- Very good electromagnetic calorimetry for electron and photon identification and measurement, complemented with full-coverage hadronic calorimetry for accurate jet and missing transverse energy (E_T^{miss}) measurements;
- High-precision muon momentum measurements, with the capacity to guarantee accurate measurements at the highest luminosity using the external muons spectrometer alone;
- Good charged particle momentum resolution and track reconstruction efficiency;
- Large acceptance in both polar angle and azimuthal angle;
- Fast and radiation hard electronics and sensor elements[2].

High accuracy and large acceptance are crucial in all parts of the detector to record the full extent of collisions. The detector must provide essential signatures of the events including electron, photon, muons, hadronic jet, vertex tagging and missing transverse energy measurements. Identification of these

[2]The effect of radiation damage is a major concern to all components, especially in the innermost tracking modules. An upgrade program to replace the inner detector is in its development phase. For example, the SCT tracker is designed to withstand a decade of radiation damage though degradation of performance is expected due to depletion of effective carrier density and increase of leakage current.

signatures needs to be optimized for a high luminosity environment where reconstruction of the objets are further complicated by the presence of pile-up[3].

To meet these requirements, the detector is a complex of state-of-art sub-detectors weighting 7000 tonnes in total. The sub-detectors systems can roughly be divided into:

- **Tracking detectors** for measurement of charged particles.
- **Calorimetry** for energy measurement of electromagnetic and hadronic particles.
- **Muon chambers** for measurement of muons.
- **Magnet system** for bending the trajectory of charged particles.

2.2.2 Nomenclature

Quantities used to describe the detector features are defined in this section.

- Coordinate system: The center of the detector defines the origin of the three axes. The beam direction defines the z-axis and the x-y plane is the plane transverse to it. The positive x-axis is pointing towards the center of the LHC ring and the positive y-axis towards the sky.
- Angles: Azimuthal angle ϕ is measured from the x-axis. The polar axis θ is measured from the positive z direction though pseudo-rapidity η is generally used instead, where $\eta = -ln(tan(\frac{\theta}{2}))$.

In hadron collisions, unlike e^+e^- colliders, the center-of-mass energy of a hard scattering is unknown and varies significantly from event to event. Rapidity (or true rapidity) of a particle is defined as $y = ln(\frac{E+p_z}{E-p_z})$ and is a useful quantity in this environment: rapidity difference of two particles is invariant under a boosting in the z direction. Pseudo-rapidity, η, approximates rapidity in the massless limit.

2.2.3 The Inner Detector

The Inner Detector (ID) [3] is the innermost part of ATLAS (see figure 2.5(a)). With a combination of high-precision, high-granularity layers in the inner part and straw tubes in the outer part, it can reconstruct the tracks of charge particles in a solenoidal magnetic field of 2T with a coverage that extends up to η =2.5.

The main requirements for the ID are:

- Tracking efficiency of 95% over the full coverage for isolated tracks with $p_T > 5$ GeV, with fake-track rates less than 1% of signal rates.
- Identification of individual particles in dense jets where the calorimeter cannot resolve the individual particles.
- Momentum measurement in a large momentum range. Below p_T= 0.5 GeV, the particles loop in the magnetic field and reconstruction is not possible. This lower limit affects the reconstruction of converted photons and J/Ψ decays.
- Distinguish between electrons and photons which create similar clusters in the Electromagnetic Calorimeter.
- Charge identification of particles with large transverse momentum for the identification of a possible Z' decay.

[3]The LHC will collide bunches of 10^{11}protons 40 million times per second. With an inelastic proton proton cross section of 60 mb, the number of inelastic scatterings per bunch crossing follows a Poisson distribution with an average of 19. This is called "pile-up".

- Decay length reconstruction used for CP-violation studies in the B-system and for a B_s^0 mixing angle.
- Tagging of jets originating from high energy b-quarks. The tagging is done by secondary vertex identification and through the identification of leptons from semi-leptonic B-meson decays.
- Momentum measurement of low energy muons which have large multiplicity scattering in the hadronic calorimeter.
- Electron/jet separation in addition to the separation already provided by the calorimeter.
- Identification of the primary vertex in the presence of many vertices from overlying minimum bias events.

The momentum and vertex resolution from physics call for high-precision measurements to be made with fine-granularity detectors. Semiconductor tracking detectors, using silicon microstrip and pixel technologies are used. The highest granularity is achieved around the vertex region using semi-conductor pixel detectors. The total number of precision layers must be limited because of the material they introduce, and because of their high cost. Photon conversions, bremsstrahlung from electrons and nuclear interactions with pions all cause a degraded calorimeter performance. Typically, three pixel layers and eight strip layers are crossed by each track. A large number of tracking points (typically 36 per track) is provided by the straw tube tracker (TRT) in the outer part, which provides continuous track-following with much less material per point and a lower cost. The combination of the two techniques gives very robust pattern recognition and high precision in both η and ϕ coordinates.

In the barrel part of the ID where $|\eta| \leq 1$ all of the detecting elements are ordered in cylindrical structures while the two end-caps have them placed in wheels. This assures that the particles pass all the detecting elements with large incident angles.

The Silicon Detectors

Two technologies are used for the ATLAS semiconductor tracker (SCT): pixel detectors placed close to the beam pipe and silicon strip detectors placed further away (as shown in figure 2.5). The SCT system is designed to provide eight precision measurements per track in the intermediate radial range, contributing to the measurement of momentum, impact parameter and vertex position, as well as providing good pattern recognition by the use of high granularity.

The silicon strip detectors have a n-type bulk with a single sided readout of n^+ strips. This choice is believed to be the most radiation resistent. In each of the layers of the detector two single sided detectors are placed back to back. In one layer the strips are parallel to the beam pipe thus measuring the ϕ coordinate directly. The strips on the back side reconstruct the z(r) coordinate in the barrel (end-cap). Each strip in the detecting element has a length of 12 cm and a width 80 μm.

The pixel detector system is designed to provide a very high-granularity, and high-precision set of measurements as close to the interaction point as possible. The system provides three precision measurements over the full acceptance , and mostly determines the impact parameter resolution. The system has approximately 80.4 million square detecting elements, each 50 μm in the Rϕ and 400 μm in z, thus giving a 2-dimensional coordinate with just one layer. The best resolution is obtained in the ϕ direction. The pixel device is placed in the layers closest to the primary vertex because of the high spatial resolution.

The most important requirements for the pixel detectors is the determination of secondary vertexes for the identification of B-meson decays, for b-tagging in top physics. It is important for pattern recognition since it has very low occupancy in spite of the close placement to the primary vertex. To fulfill this

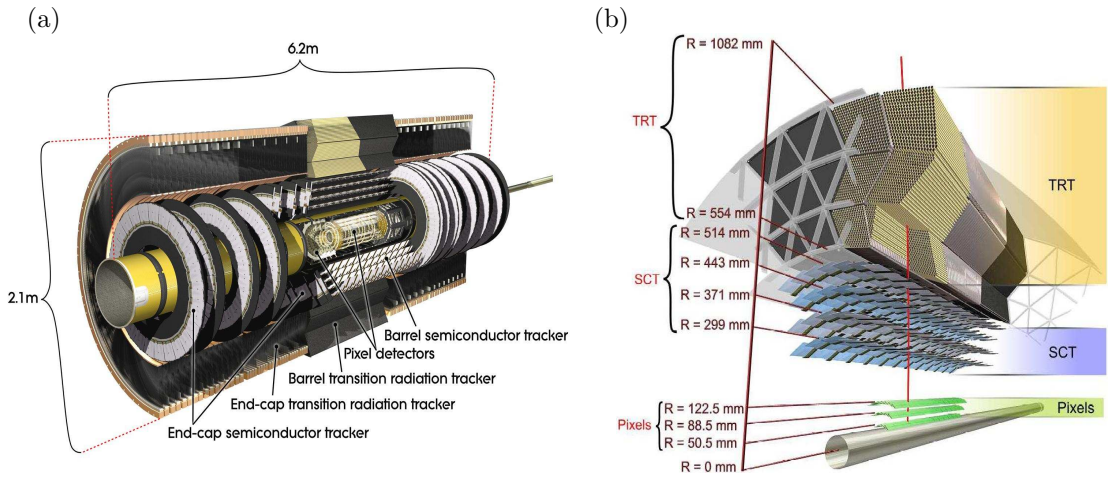


Figure 2.5: Overview of the ATLAS Inner Detector (a). Drawing showing the sensors and structural elements traversed by a charged track of 10 GeV p_T in the barrel inner detector ($\eta = 0.3$) (b). The track traverses successively the beryllium beam-pipe, the three cylindrical silicon-pixel layers with individual sensor elements of 50×400 μm^2, the four cylindrical double layers (one axial and one with a stereo angle of 40 mrad) of barrel silicon-microstrip sensors (SCT) of pitch 80 μm, and approximately 36 axial straws of 4 mm diameter contained in the barrel transition-radiation tracker modules within their support structure. .

requirement the first pixel layer is placed as close as possible to the beam pipe (the radius of the beam pipe is $29 < R < 36$ mm).

The Transition Radiation Tracker

The Transition Radiation Tracker (TRT) (shown in figure 2.5) is based on the use of straw detectors, which can operate at the very high rates expected at the LHC by virtue of their small diameter and the isolation of wires within individual gas volumes.

Electron identification capability is added by employing xenon gas to detect transition-radiation photons created in a radiator between the straws. This technique is radiation hard, and allows a large number of measurements, typically 36, to be made on every track at a modes cost. However, the detector must cope with a high occupancy and high counting rates at the LHC design luminosity. The large number of hits of the tracks is powerful in the pattern recognition state where tracks are to be found in the detector. The TRT provides additional discrimination between electrons and hadrons.

In total there are around 370,000 straws of 4 mm diameter in the TRT, which are placed radially in the end-cap and along the beam axis in the barrel region; these orientations are chosen to maximize the number of straws passed in all directions pointing away from the interaction point. The straws are filled with a xenon gas mixture for the absorption of transition radiation which also enables a faster drift-time for electrons, providing higher spatial resolution and reducing the influence from neighboring bunch crossings at the LHC. In the center of the straw there is a 30 μm gold covered tungsten wire.

Some photos of the ID are shown in figure 2.6.

2.2.4 Calorimetry

A calorimeter is usually divided into an electromagnetic and a hadronic calorimeter. This distinction is possible because of the different interaction behavior between the calorimeter and electron/photons on

(a) (b)

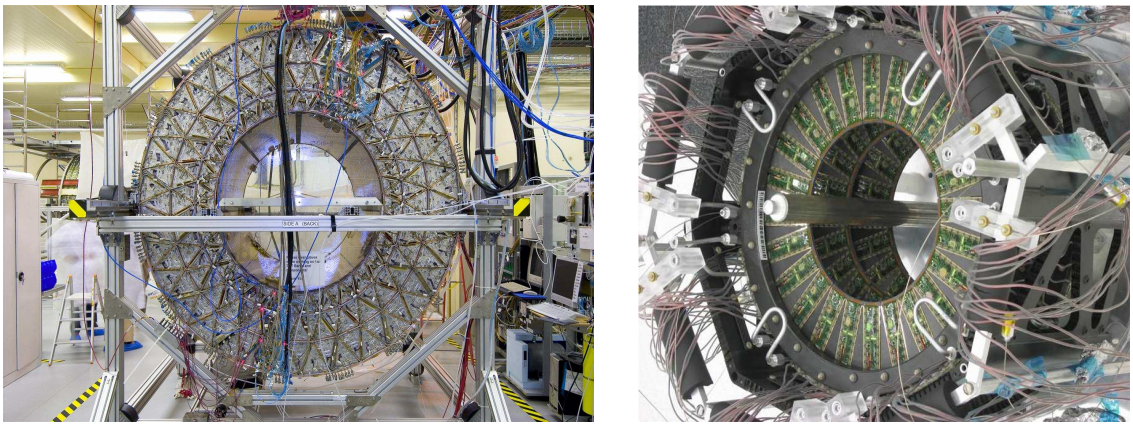


Figure 2.6: View of the Inner Detector ATLAS TRT (September 2005) (a) and picture of a pixel detector end-cap (b).

one side and hadrons on the other side. An extensive overview of calorimeters in particle physics is given in [4].

The ATLAS calorimeter [5], shown in figure 2.7, consists of an electromagnetic calorimeter covering the pseudo-rapidity region $|\eta| < 3.2$, a hadronic barrel calorimeter covering $|\eta| < 1.7$, hadronic end-cap calorimeters covering $1.5 < |\eta| < 3.2$, and forward calorimeters covering $3.1 < |\eta| < 4.9$.

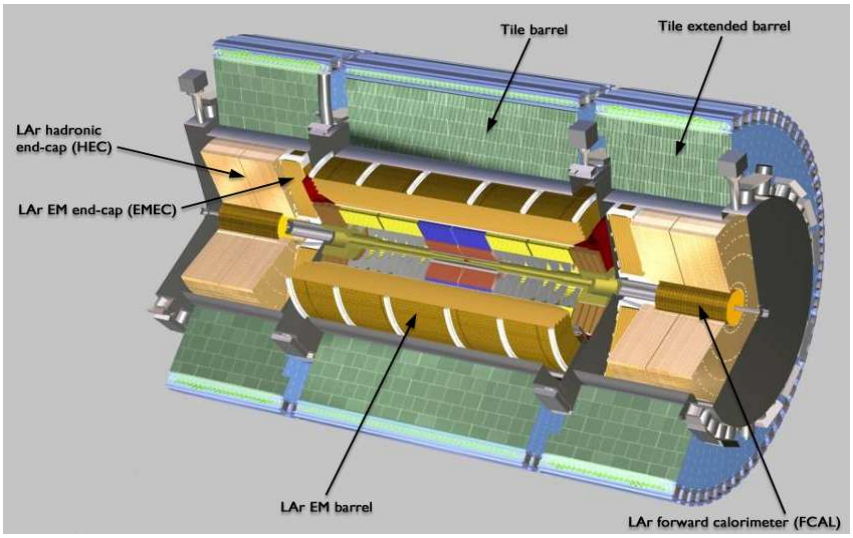


Figure 2.7: General view of the ATLAS calorimeter system.

The Liquid Argon Calorimeter

The electromagnetic (EM) calorimeter is a lead-liquid argon sampling calorimeter with accordion-shape absorbers and electrodes. The lead gives the shower development with its short radiation length and the secondary electrons create ionization in the narrow gaps of liquid argon. An inductive signal from the ionization electrons drifting under the action of an electric field across the gap-gap is registered by cooper electrodes. The accordion geometry provides a full coverage in ϕ without any cracks, and a fast

extraction of the signal (figure 2.8(a)). In the region devoted to precision physics ($0 < |\eta| < 2.5$) the EM calorimeter is segmented into four samplings:

- *Presampler:* A single thin layer of argon but no lead absorber in front. The purpose is to correct for the energy lost in the solenoid and cryostat wall.
- *1st Sampling:* The first sampling has a depth of $4.3X_0$[4]. The readout is as seen in figure 2.8(a) in thin η strips which provides an excellent resolution in the η coordinate for γ/π^0 separation. The ϕ coordinate is not suited for this since converted photons will open up in the magnetic field and produce clusters with widths similar to π^0 clusters.
- *2nd Sampling:* The majority of the energy is deposited in the $16X_0$ of the second sampling. Clusters with energy below 50 GeV are fully contained. For the position measurement of the cluster the two coordinates are equally important resulting in square cells of size $\Delta\eta$ x $\Delta\phi = 0.02454$ x 0.02454.
- *3rd Sampling:* Only the highest energy electrons will reach this deep in the detector. The clusters are at this point wide and the cell size can be doubled in the η direction without loss of resolution.

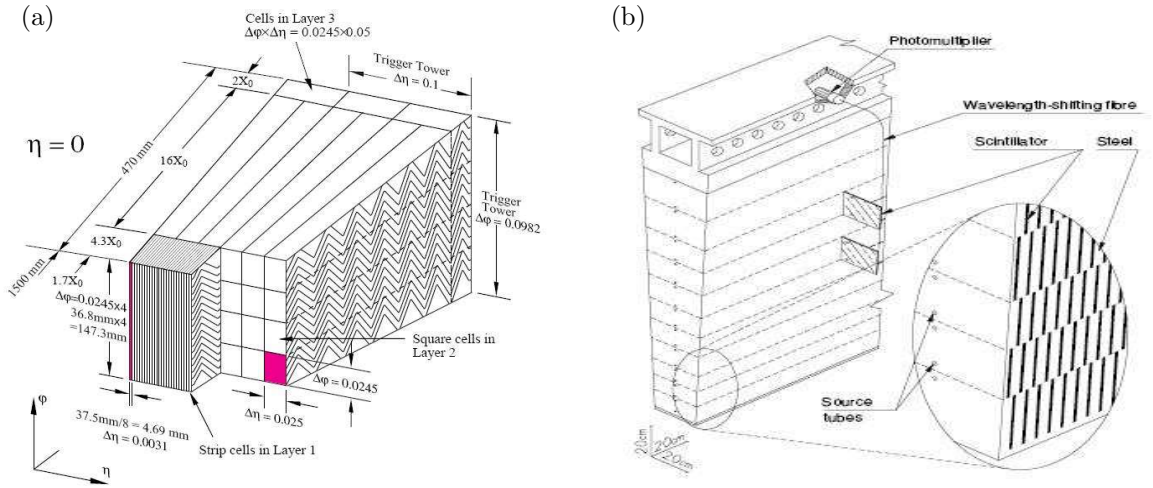


Figure 2.8: Sketch of a barrel module of the EM calorimeter, where the different layers are clearly visible (a). Schematic view of the tile calorimeter, the optical readout, the tiles, the fibres and the photomultipliers are shown (b).

For the high-η region ($2.5 < |\eta| < 3.2$), i.e. for the end-cap inner wheel, the calorimeter is segmented in two longitudinal sections and has a coarser larger granularity than for the rest of the acceptance, which is sufficient to satisfy the requirements (reconstruction of jets and E_T^{miss}). The end-cap EM calorimeters start at $|\eta| = 1.5$ and continue down to $|\eta| = 3.2$ but with an increased cell size above the $|\eta| = 2.5$. There is a crack with bad energy resolution where the end-cap and barrel calorimeters meet.

In the forward region calorimeter is also liquid argon technology to withstand the high radiation levels. The design is simpler than the EM calorimeter and has parallel copper plates as absorbers placed perpendicular to the beam.

[4]The radiation length, X_0, is the longitude after which an electron radiates a bremsstrahlung photon or the longitude after which a photon creates an e^+e^- pair.

The very forward hadronic calorimeter with a coverage down to $|\eta| = 4.9$ is made of copper/tungsten. The choice of these materials is necessary to limit the width and depth of the showers from high energy jets close to the beam pipe, and to keep the background level low in the surrounding calorimeters from particles spraying out from the forward region.

The Tile Calorimeter

The tile calorimeter is a sampling calorimeter using steel as absorber and scintillator as the active medium. It is located in the central rapidity region, reaching out to $|\eta| = 1.7$, where the liquid argon electromagnetic calorimeter takes over. It is subdivided in a central barrel and two extended barrels. The radial depth of the tile calorimeter is approximately 7.4λ (interaction lengths). The resolution of the hadronic calorimeter is $0.5/\sqrt{E} \oplus 3\%$ for $|\eta| < 3$ and $1/\sqrt{E} \oplus 10\%$ for $3 < |\eta| < 5$, as detailed in table 2.2.

Each barrel consists of 64 modules of size $\delta\phi \sim 0.1$, made of steel plates and scintillating tiles. The orientation of the scintillating tiles radial and normally to the beam line allows for almost full ϕ calorimeter coverage. The light created in the scintillators is read out with wavelength shifting fibres to photomultipliers placed on the outside of the calorimeter. The fibres absorb the light produced in the scintillators and re-emit it at longer wavelengths where it reaches the photomultipliers through total reflection inside the fibres (figure 2.8(b)).

To improve hermiticy of the hadronic calorimeter, special modules are placed in the gap region between the barrel and the extended barrel, as shown in figure 2.9(a). These modules are made of steel-scintillator sandwiches with the same sampling fraction as the rest of the tile calorimeter and with thin scintillator counters in the sectors where the available space in the gaps is even more limited (e.g. the region where the cables and cooling of the EM calorimeter and the ID pass through). The cumulative amount of material in front of the EM calorimeter, in the EM calorimeter, in each hadronic calorimeter layer and before the first layer of the muon spectrometer is shown in figure 2.9(b).

(a) (b)

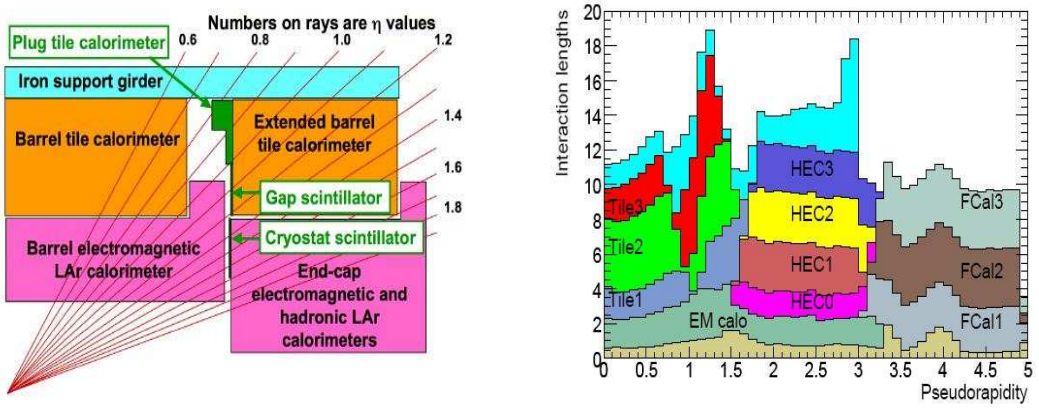


Figure 2.9: Schematic of the transition region between the barrel and end-cap cryostats (a). Cumulative amount of material, in units of interaction length, as a function of $|\eta|$, in front of the EM calorimeter, in the EM calorimeter itself, in each hadronic layer, and the total amount of material in front of the first active layer of the muon spectrometer (up to $|\eta| < 3.0$) (b).

2.2.5 The Muon System

The muon spectrometer [6] forms the outer part of the ATLAS detector. It is designed to detect charged particles exiting the calorimeters and to measure their momentum in the pseudo-rapidity range up to $|\eta| < 2.7$. It is also designed to trigger on these particles in the region $|\eta| < 2.4$. The general layout of the muon system is shown in figure 2.10.

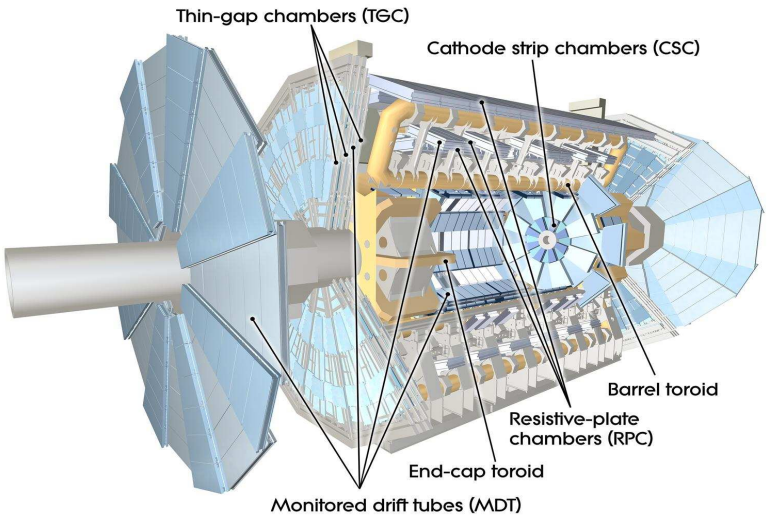


Figure 2.10: The ATLAS Muon Detector.

The muon system has to fulfill the following requirements:

- A good transverse momentum resolution in the low p_T region. The limit is defined by the ability to detect the $H \rightarrow ZZ^*$ decay in the muon channel with a high suppression of the background. A momentum resolution of around 1% is required.
- Sufficient resolution at high p_T for good charge identification to identify $Z' \rightarrow \mu^+\mu^-$ process.
- A rapidity coverage of $|\eta| < 3.0$ and a hermetic system to prevent particles to escape through holes.
- Measurement of spatial coordinates in two dimensions to provide good mass resolution.
- A low rate of both punch-through hadrons and fake tracks.

Precision-tracking chambers in the barrel region are placed between and on the eight coils of the superconducting barrel toroid magnet (see chapter 2.2.6), while the end-cap chambers are in front of and behind the two end-cap toroid magnets. The ϕ symmetry of the toroids is reflected in the symmetric structure of the muon chamber system, consisting of eight octants. The chambers in the barrel are arranged in three concentric cylindrical shells around the beam axis. In the end-cap regions, muon chambers form large wheels, perpendicular to the z-axis. Figure 2.11 gives cross-sections in the planes transverse to, and containing, the beam axis. In the center of the detector ($|\eta| \approx 0$), a gap in chamber coverage has been left open for services to the solenoid magnet, calorimeters and the ID. In the barrel/end-cap transition region with $1.1 < |\eta| < 1.3$ (combined reconstruction) or < 1.7 (stand-alone reconstruction), the middle stations are missing for the initial data-taking (EES and EEL chambers, shown in figure 2.11(b))), which results in a large degradation of the resolution and a drop of the reconstruction efficiency (using combined (barrel region) or stand-alone (end-cap region) muon reconstruction).

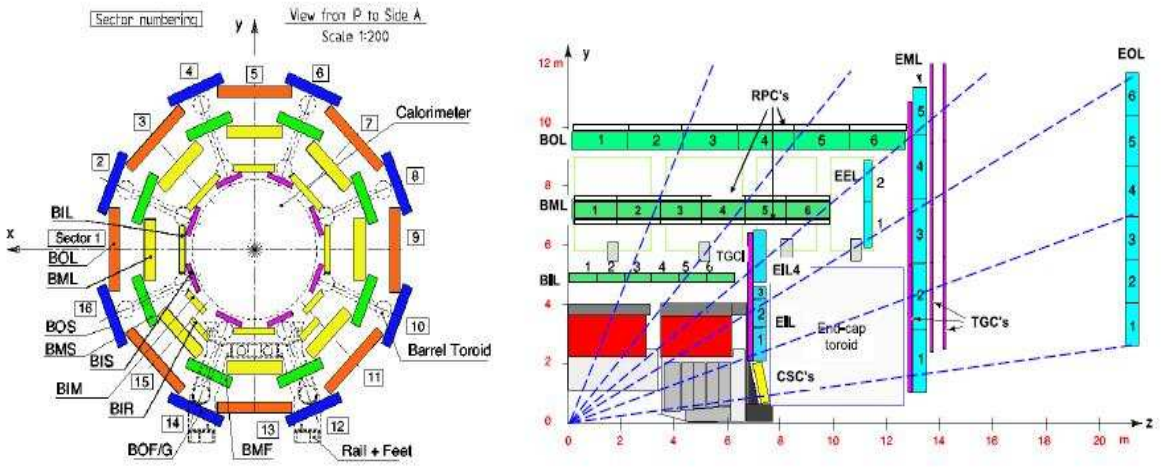


Figure 2.11: Cross-section of the barrel muon system perpendicular to the beam-axis (non-bending plane), where the three concentric cylindrical layers are shown (a). Cross-section of the muons system in a plane containing the beam axis (bending plane). Infinite-momentum muons are illustrated by the dashed lines (b).

The precision momentum measurement is performed in the Monitored Drift Tube chambers (MDTs). They cover the pseudo-rapidity range $|\eta| < 2.7$. These chambers consist of three to eight layers of drift tubes made out of aluminium tubes of 30mm diameter and lengths from 70 cm to 630 cm. To measure the coordinate in the bending plane of the magnet the tubes are placed transverse to the beam axis. In order to reduce the level of fake tracks reconstructed from random associations of background hits each set of MDTs are required 2 superlayers each with 3 or 4 layers of tubes. Each MDT has a resolution of 80 μm, which results in a momentum resolution $\Delta p_T/p_T < 10^{-4}$ for tracks with $p_T > 300$ GeV.

In the forward region ($2 < |\eta| < 2.7$), where track density is higher, Cathode-Strip Chambers (CSC) are used in the innermost tracking layer due to their higher rate capability and time resolution. The CSCs are multiwire proportional chambers segmented into strips in orthogonal directions. This allows both coordinates to be measured. The resolution of a chamber is 40 μm in the bending plane and 5 mm in the transverse plane.

To obtain a good resolution in the momentum measurement, a high-precision optical alignment system is needed. It monitors the positions and internal deformations of the MDT chambers. It is complemented by track-based alignment algorithms.

An essential design requirement of the muon system was the capability to trigger on muon tracks. The precision-tracking chambers have been complemented by a system of fast trigger chambers. In the barrel region ($|\eta| < 1.05$), Resistive Plate Chambers (RPC) were selected, while in the end-cap region ($1.05 < |\eta| < 2.4$), Thin Gap Chambers (TGC) are used. The trigger chambers measure both coordinates of the track, one in the bending (η) plane and one in the non-bending (ϕ) plane, while MDT chambers only give the η coordinate.

The general performance goals of the ATLAS detector are summarized in table 2.2.

Detector component	Required resolution	η coverage
tracking	$\sigma_{p_T}/p_T = 0.05\% \, p_T \oplus 1\%$	± 2.5
EM calorimeter	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2
hadronic calorimeter		
barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2
forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 \leq \lvert\eta\rvert \leq 4.9$
muon spectrometer	$\sigma_{p_T}/p_T = 10\%$ (at $p_T = 1$ TeV)	± 2.7

Table 2.2: Required resolution and coverage for the different detector components.

2.2.6 The Magnet System

The ATLAS superconducting magnet system is an arrangement of a central solenoid providing the ID with a magnetic field, surrounded by a system of three large air-core toroids generating the magnetic field for the muon spectrometer.

The solenoid magnet is placed inside the EM calorimeter. This is different from most other detector designs where the magnet is placed outside the EM calorimeter. A small magnetic field also reduces the transverse spread of showers. The major problem is the increased amount of material in front of the calorimeter which causes many particles to start showering before they reach the active part of the calorimeter. The solenoid superconducting magnet is cooled indirectly by helium at 4.5 K. To reduce the material, the magnet shares the cryostat with the liquid argon calorimeter. The length is considerably shorter than the inner tracking system. This is the result of a compromise: a short coil reduces the material in front of the calorimeter and a long coil makes the magnetic field more uniform in the ID. The magnetic field along the z-direction is 2 T at the interaction point.

The toroid magnet is divided into a barrel part and two forward regions (figure 2.12). Each of the three toroids consists of eight coils assembled radially and symmetrically around the beam axis as shown in figure 2.13. The toroid coils are of a flat racetrack type with two double-pancake windings made of 20.5 kA aluminium-stabilized NbTi superconductor and are in separate cryostats for the barrel. In the forward region, the toroid field is also formed by eight superconducting coils but they are placed in a common cryostat. With a toroid field, particles will cross the complete pseudo-rapidity range, almost perpendicular to the field. This means that the field integral, which is the important factor for momentum resolution, can be kept high even in the forward region. The low number of coils to form the toroid field results in a field strength that varies strongly with the ϕ coordinate. The field in the barrel is 2 T and in the end-caps from 4 T to 8 T (at $\phi = 0$) as shown in figure 2.12(b).

(a) (b)

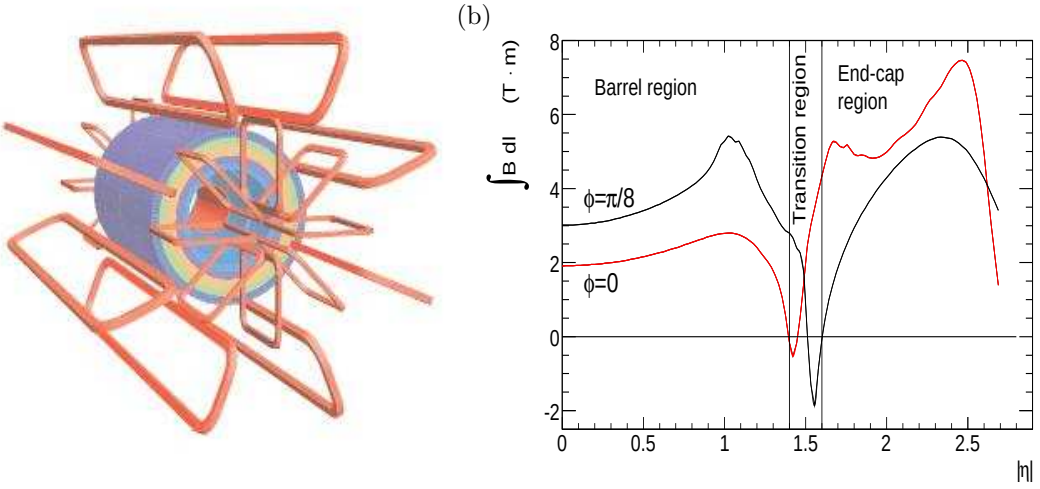


Figure 2.12: Geometry of magnet windings and tile calorimeter steel. The eight barrel toroid coils, with the end-cap coils interleaved are visible. The solenoid winding lies inside the calorimeter volume (a). And predicted field integral as a function of $|\eta|$ from the innermost to the outermost MDT layer in one toroid octant, for infinite momentum muons. The curves correspond to the azimuthal angles $\phi = 0$ (red) and $\phi = \pi/8$ (black) (b).

(a) (b)

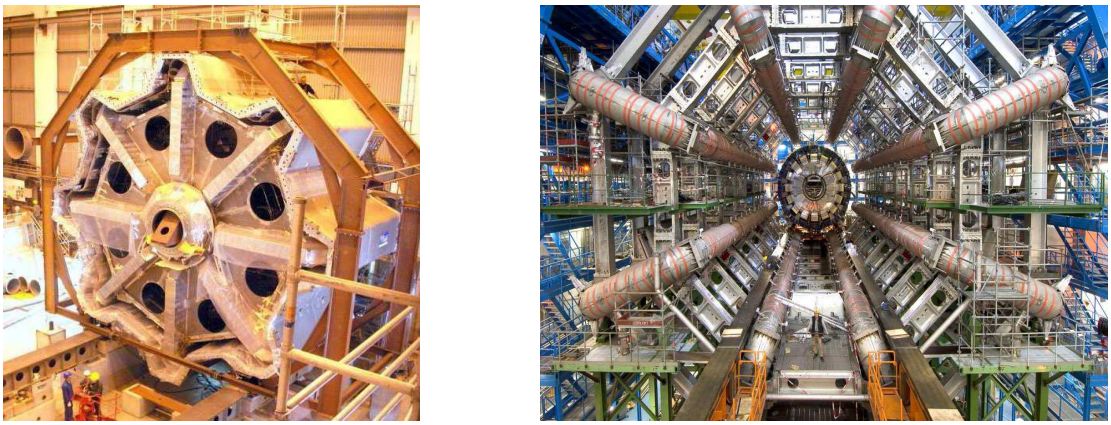


Figure 2.13: End-cap toroid cold mass. The eight flat, square coil units and eight keystone wedges (with the circular holes) are visible (a). Barrel toroid as installed in the underground cavern; note the symmetry of the supporting structure. (b).

2.3 Trigger and Data Acquisition System in ATLAS

The data-size of one recorded collision is of the order of 1MB. Since bunch crossings occur every 25 ns (i.e. at a rate of 40 MHz), this would result in a data volume which cannot be stored with today technologies. To be handled by the ATLAS computing system a reduction to 300 MB/s is needed. The goal of the ATLAS Trigger and Data Acquisition system (TDAQ) is to reduce the rate of candidate collisions from 40 MHz to ∼200 Hz without loss of interesting physics events. The TDAQ system has three levels as shown in figure 2.14.

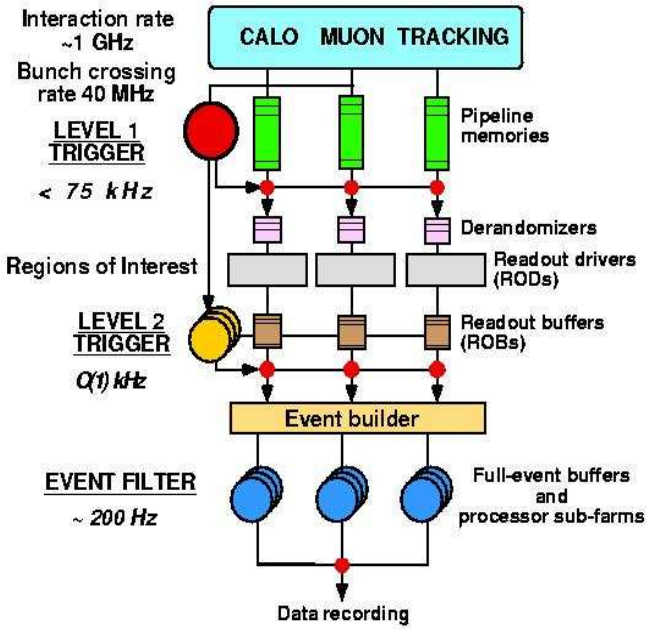


Figure 2.14: Diagram of the Trigger/Data Acquisition system.

The first level (LVL1) is hardware based. It uses the information coming from the calorimeters with reduced granularity and from the muon trigger chambers (RPCs and TGCs stations). The latency, time in which data of all sub-detectors can be stored in a pipeline, of the level one trigger is 2μs, which leads to a target rate of 75kHz. An important task of LVL1 is to define the so-called region of interest (RoI) for each event. The RoIs are regions in the detector, where possibly interesting objects might be present. The LVL1 trigger passes the event information within the RoIs from the read-out buffers (ROBs) to the second level trigger (LVL2).

The LVL2 trigger is software based and uses the full granularity in the RoIs of the detector and also the ID. The target rate is 1 kHz, with a latency of 1ms to 10 ms, depending on the complexity of the event. The access of the LVL2 trigger to the full event would exceed the required maximal latency and hence the concept of RoIs is needed. The disadvantage of this approach is that an interesting object, which has failed LVL1 trigger, cannot be found by LVL2. If an event passes the LVL2 trigger requirements, all information of the event is collected from the ROBs by the so-called Event Builder and passed to the third trigger level, which is called Event Filter (EF).

The EF makes the final decision whether the event is recorded for further analysis or not. Its target rate is ∼200 Hz. The Event Filter is also software based and runs on a computer farm near the ATLAS pit. This allows for a relatively long decision time of the order of one second. As a consequence, the EF has access to full event information. More sophisticated reconstruction algorithms can be used. Events

which are accepted by the EF are written in mass storage devices and available for further offline analysis.

Even though the TDAQ system highly reduces the output rate, the total space needed by the ATLAS experiment is of the order of 1 PetaByte (10^{15} bytes) per year. The event size is of 1.5 MB. This makes a powerful computing environment necessary, which is introduced in section 2.4.

The ATLAS trigger menu defines the operation of the trigger system and its conditions. A condition is a combination of an object, e.g. an electron, and a certain threshold, e.g., $p_T \geq 20$ GeV. Table 2.3 shows the trigger menu for the low luminosity phase of LHC.

LVL1 signature	Rate (kHz)
EM25I	12.0
2EM15I	4.0
MU20	0.8
2MU10	0.2
J200	0.2
3J90	0.2
4J65	0.2
J50+XE60	0.4
TAU25+XE30	2.0
MU10+EM15I	0.1
Others	5.0

HLT signature	Rate (Hz)
e25i	40
2e15i	<1
2γ20i	2
μ20i	40
2μ10	10
j400	10
3j165	10
4j110	10
j70+xE70	20
τ35+xE45	5
μ10+e15i	1
2μ6 for B-physics	10

Table 2.3: Subset of items from two illustrative trigger menus at LVL1 (left) and at the HLT (right) for a luminosity of $2\text{x}10^{33}\text{cm}^{-2}\text{s}^{-1}$.

2.4 ATLAS Computing

2.4.1 The ATLAS Software Framework

The software for the experiment is developed by a working group, comparable in manpower to a sub-detector working group in hardware. The complexity of the detector also implies a complex underlying software and hence the development of a common software framework for the experiment which is called Athena [7].

Athena allows an integrated communication between various software applications, called algorithms, within the framework. All algorithms have access via the StoreGate to data of other algorithms, e.g. the general event information or the detector description. Moreover, a common framework ensures a common approach of software developing, e.g. messaging or access on disk and re-use of already written code-segments, and a common underlying design of the software packages. The Athena framework allows a dynamic loading of libraries and it is organized in form of plug-in modules allowing flexible configuration of various algorithms to be executed. The configuration is done via the so-called JobOptions file which allows user specification of the algorithms via Python scripting.

The main purpose of the ATLAS software is to generate, simulate, digitize and reconstruct proton-proton collisions, i.e., events, in the LHC environment. This Athena software chain is illustrated in figure 2.15.

The advantage of the modular approach becomes obvious since the simulation, digitization and reconstruction algorithms for the sub-detector systems can be implemented independently. The generation process includes the proton-proton collision itself, calculating the position and momentum four-vectors of all the particles which are produced in the collision. The generation is based on various Monte Carlo

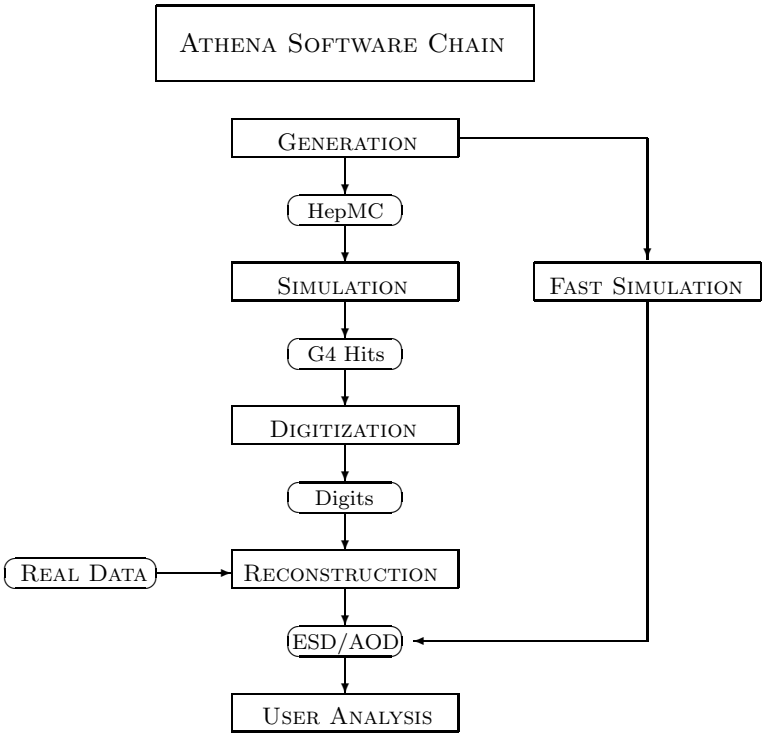


Figure 2.15: Illustration of the Athena software chain including generation, simulation, digitization and reconstruction.

generator programs, which are explained in section 3.5.

The second part is the tracking of the produced particles through the detector using the Geant4 toolkit [8]. Geant4 simulates the impact of the magnetic field and the interactions with the material (multiple scattering, energy loss, photon conversions) and further decays of unstable particles. Each interaction of a particle with an active, i.e. sensitive, detector element is stored in a so called hit-object, which represents the position and type of the interaction.

During the digitization step, the response of the detector and its electronics on the various hit-objects is simulated. The final information is stored in so-called digit-objects.

The reconstruction part is based on the digitized information. The reconstruction includes various algorithms, for pattern recognition, track fitting, vertex determination and energy measurements. It should be noted that the digitized information and the real data are equivalent from a data representation point of view. The output of the reconstruction part are the so-called event summary data (ESD) and analysis object data (AOD) files. While the first includes a more detailed description of an event, the latter one only includes information which is of primal interest for physics analysis.

The last step in a standard Athena chain is the actual physics analysis, where several techniques can be applied to study a particular physics process.

2.4.2 Grid Computing

One year of data taking at the LHC results in 15 Petabytes (15 million of Gigabytes) of data, produced by the four experiments. The data have to be analyzed by physicists worldwide to discover new physics processes. In addition, billions of complex theoretical simulations of the proton-proton collision must be calculated.

In the previous LEP experiments, the computer processing was done at a computer farm, near the experiment. For LHC, storage and processing requirements exceed by far the capacities available at any

single site and hence a new approach has been chosen, which is commonly known under LHC Computing Grid (LCG) project [9]. This computing grid provides the infrastructure for the storage of the data and the necessary computing power for the physics analysis and simulations.

The data distribution follows a so-called Tier-structure (figure 2.16). The LHC raw data is recorded in a first step on tape at the so called Tier-0 center at CERN, where calibration and first-pass reconstruction takes place and a copy of the reconstructed data is stored. Tier-0 distributes a second copy of the raw data to the Tier-1 centers. Additional copies of the reconstructed data are also distributed to the Tier-1 centers. The Tier-1 centers have the responsibility for managing the permanent data storage (raw, simulated and processed data) and providing computational capacity for re-processing and for analysis processes that require access to large amounts of data. Tier-2 centers provide computational capacity and appropriate storage services for Monte Carlo event simulations and perform end-user analysis. Tier-2 centers obtain data as required from Tier-1s, and the data generated at Tier-2s is sent to Tier-1s for permanent storage. Finally, Tier-3 centers, which represent facilities at universities and laboratories, take part in the processing and analysis of the LHC data. Any group of scientists associated to a Tier-2 center can be a Tier-3. They have access to the data and analysis facilities. In Spain, there is one Tier-1 facility at Barcelona, PIC. And one distributed Tier-2 center for ATLAS, involving the IFAE (Barcelona), the IFIC (Valencia) and the UAM (Madrid).

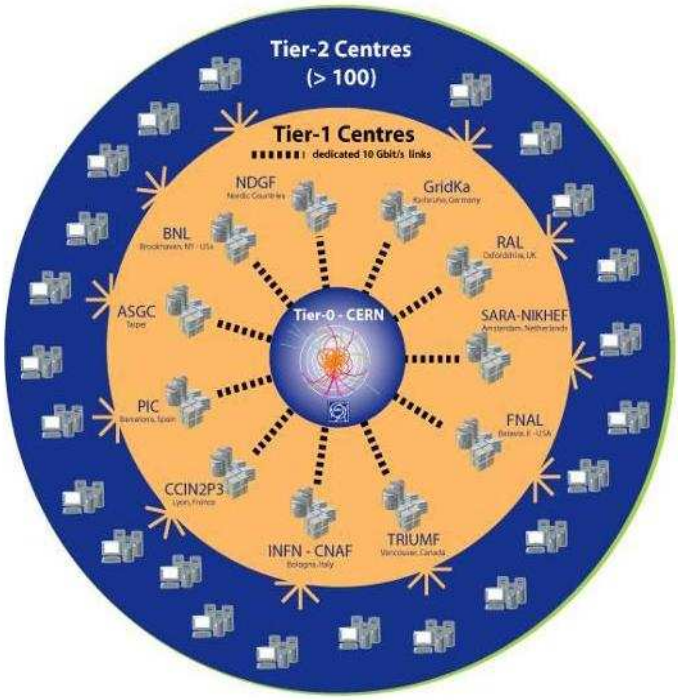


Figure 2.16: LCG tier structure.

The LCG project involves dedicated hardware and software developments. Obviously, an adequate bandwidth is needed for the data distribution within the grid. The grid-software must be compatible with heterogeneous hardware and must also ensure coherent software at all connected computers. Managing and protecting the data so they are not lost or corrupted over the lifetime of the LHC, and providing accounting mechanisms so that different groups have fair access, are some of the challenges that the LCG is addressing. Several tests of the grid infrastructure (data challenges) have been preformed during the last years to ensure the full functionality of the grid with the start of the data taking.

2.5 Physics Programme

The benchmark test of the ATLAS detector design is its discovery potential for the Higgs boson, however the high luminosity and the large center of mass energy of the LHC proton-proton collisions allow also a test of various theoretical models, e.g. supersymmetry. In each second the production of one $t\bar{t}$-pair, five Z bosons, which decay into lepton pairs, 50 W bosons, 100 QCD jets with a tranverse momentum larger than 200 GeV and half a million $b\bar{b}$-pairs are expected within the ATLAS detector during the low luminosity phase of LHC, which corresponds to an instantaneous luminosity of 10^{33} cm^{2}s^{-1}. These large rates of physics processes provide not only a good opportunity for high precision tests of the Standard Model, but are also a huge background to many hypothetic new physics channels. Some of these physics channels are shortly discussed in the following [10].

2.5.1 The Higgs Boson

The production cross-section of a Higgs boson with a mass of 1 TeV is expected to be more than 100 fb, which corresponds to roughly 1,000 events in one year during the low luminosity phase. The width of the Higgs boson Γ_H depends on its mass m_H, i.e.

$$\Gamma_H \propto (m_H)^3$$

and is expected to be in the order of a few MeV for $m_H \approx 100$ GeV and rises up to 100 GeV for $m_H \approx 600$ GeV. Therefore, Higgs bosons with large masses cannot be identified as a clear peak. The couplings of the Higgs boson to fermions are proportional to the fermion masses. This implies different dominating decay modes for different m_H-scenarios (figure 2.17(a)). For the search of the Higgs boson, three mass ranges are distinguished:

- **Low mass region:** $m_H < 130$ GeV: The decay into two b-quarks will dominate in this region since these quarks are the most massive fermions which are kinematically accessible. Due to the overwhelmingly large QCD-background, it is extremely difficult to discover the Higgs-boson in this channel. The decay $H \rightarrow \gamma\gamma$ and $H \rightarrow \tau\tau$ are the most promising searches for discovery, but even here, the irreducible background has a cross section which is 60 times larger than the signal one.

- **Intermediate mass region:** 130 GeV $< m_H <$ 180 GeV: The decays into the gauge bosons $H \rightarrow ZZ^* \rightarrow 4l$ and $H \rightarrow WW^* \rightarrow l\nu l\nu$ becomes dominating in this regime[5]. For $m_H \approx 170$ GeV the four lepton decay gets suppressed, once the decay mode into two real W bosons opens up. In this case, it is crucial to understand the background. The signal significance exceeds 5σ, assuming a relative uncertainty on the background contribution of 5%. This is not the most promising mass regime for the discovery of the Higgs-Boson.

- **High mass region:** $m_H > 180$ GeV: In this mass regime the decay in two real Z bosons is dominating, which is the most clear channel for the Higgs-boson search. The dominating background is the continuum production of Z boson pairs.

Figure 2.17(b) illustrates the overall sensitivity for a Standard Model Higgs-boson for an integrated luminosity of 100 fb^{-1}, which corresponds to a few years of running at low luminosity. It can be seen that a signal significance of 5 standard deviations can be achieved. It is expected that higher integrated luminosities in the range of 300 fb^{-1}allow for a determination of the Higgs boson mass to a precision of 0.1% for a 120 GeV $< m_H <$ 400 GeV and its cross section to a precision of roughly 10%. It should be also noted that ATLAS provides good possibilities to discover a Higgs sector within the Minimal Supersymmetric Standard Model (MSSM) [11].

[5]The letter l indicates the decay into a lepton, either an electron or a muon.

(a) (b)

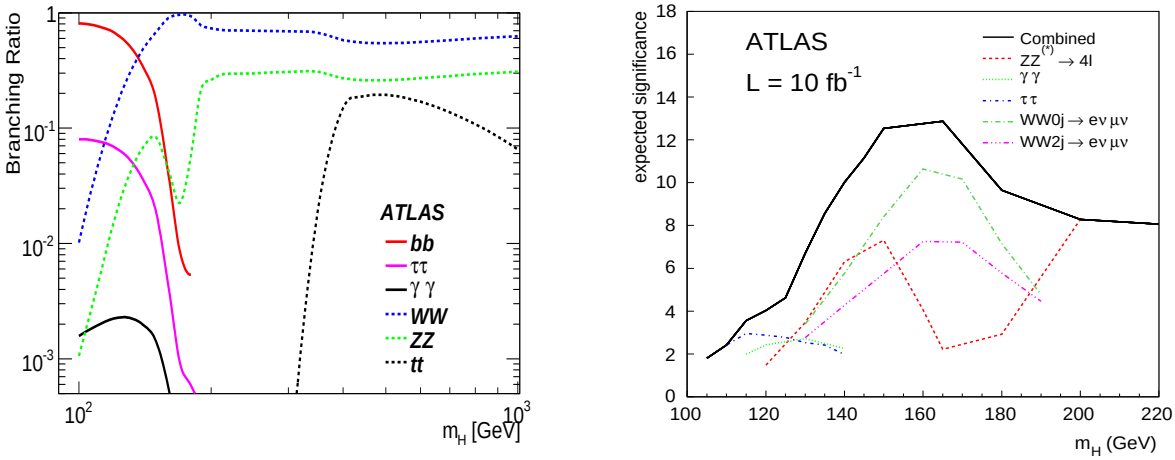


Figure 2.17: Branching ratios for the relevant decay modes of the SM Higgs boson as a function of its mass (a). The ATLAS sensitivity for the discovery of the SM Higgs boson. The statistical significance is plotted for various decay channels and the combination with an integrated luminosity of 10fb^{-1} for the lower mass range (b).

Last news, from March of 2009, excludes the Higgs mass region 160-170 GeV with 95 % confidence level. This exclusion comes from new results of combined data of Tevatron experiments. Figure 2.18 shows the different excluded regions of the Higgs mass, including this last result. Earlier experiments at the Large Electron-Positron Collider at CERN excluded a Higgs boson with a mass of less than 114 GeV at 95% confidence level. And quantum effects calculations involving the Higgs boson require its mass to be less than 185 GeV.

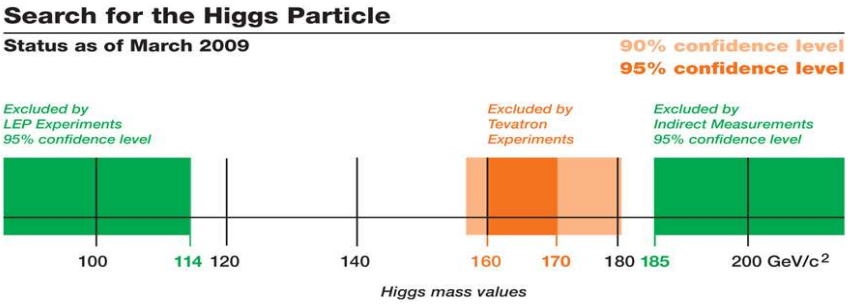


Figure 2.18: Excluded Higgs regions. This is the last update, from March 2009, where the region 160-170 GeV of the Higgs mass is excluded at 95 % confidence level, from Tevatron experiments.

2.5.2 Supersymmetry

The discovery of supersymmetric particles [12], which exists at the electro-weak scale, should be relatively easy compared to the discovery of the Higgs boson. The cross section of the production of gluinos and squarks are comparable to the relevant Standard Model background processes at the same Q^2. 10,000 events containing supersymmetric particles with a mass around 1 TeV are expected during one year of running at low luminosity.

The actual decay modes of supersymmetric particles are model dependent. Usually, final state objects with high transverse momenta are expected, since they stem from heavy particles. A very significant feature of the decay of supersymmetric particles is a relative large amount of missing transverse energy, which is due to the lightest supersymmetric particle, the so-called neutralino. The neutralino cannot decay further, since it is already the lightest particle with a supersymmetric quantum number. Moreover it should be a weakly interacting particle and hence escapes the detector without any further interaction. These signatures are therefore useful to discriminate from the Standard Model background processes.

2.5.3 Beyond the Standard Model

The supersymmetric extension has some attractive theoretical features, but is not the only possible theoretical scenario, which could be discovered by the ATLAS detector. The discovery potential of ATLAS covers also:

- New vector bosons, which are usually labeled as W' and Z' up to a mass scale of 5 TeV.
- Leptoquarks up to a mass scale of 1.5 TeV.
- Excited quarks up to masses of 6 TeV.
- Technicolor resonances up to a mass limit of 1 TeV.
- Signature of extra-dimension models.
- New particles, predicted by Little-Higgs models.

2.5.4 Precision Measurements

The statistical uncertainty on measurements scales as $\frac{1}{\sqrt{N}}$, where N is the number of signal events. Hence, the error on most measurements is dominated by systematic uncertainties, since the statistical contribution can be usually neglected due to the enormous luminosity of LHC. Some systematic uncertainties can also be reduced with high statistic calibration samples.

Moreover, the high statistic in some channels allows a clean cut-based selection of signal processes. An example is the determination of the top quark mass. While the experiments at the Tevatron collider have a few hundred recorded top quark events [13] , which require sophisticated methods for the signal selection and mass determination, the ATLAS detector is expected to record thousands of top quark events during a day in the low luminosity phase. This large statistics allow simple hard cuts, which ensure a very clean and well understood event sample.

The precise measurement of the top-quark mass is interesting not only because it is the only quark with mass at the electroweak scale, but also because it constrains the Higgs boson mass through loop corrections. A Higgs-boson discovered with a mass which is predicted by precision top-quark measurements would be a very remarkable result of the Standard Model. The mass of the top quark will be measured in the semi-leptonic decay mode of a top-quark pair. The lepton, stemming from W boson decay, is used to tag the event, while the jets resulting from the decay of the second top quark, are used for the mass determination. This process can also be used for the determination of the jet energy scale, i.e. the precision to which the average energy of the particle jets can be reconstructed in the detector. Assuming an uncertainty of 1% on the jet energy scale in ATLAS, a precision of $\Delta m_{top} \approx 1 - 1.5$ GeV can be reached.

In order to predict the mass of a Standard Model Higgs boson, also the mass of the W boson must be measured to high precision. Figure 2.19 illustrates the top quark and W boson mass relation [14]. The mass of the W boson is known to ~ 25 MeV. The precision measurement of the W boson mass in hadron colliders is achieved by using templates based W mass fits. Templates of the p_T^l and M_T^W distributions

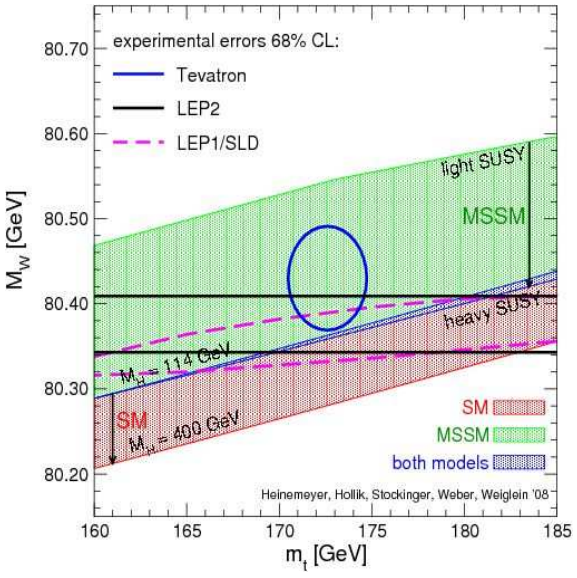


Figure 2.19: Dependence of the Higgs boson mass on the masses of the W boson and the top quark. The latest measurements of these W and top mass suggest a light Higgs mass, i.e. below 300 GeV.

are produced with varying m_W values, and compared to the corresponding distribution observed in data. The expected dominant systematic contributions of this measurement are the absolute calibration of the lepton energy scale and various theoretical uncertainties. It is expected that this measurement is a complex measurement in ATLAS and will take several years to reach the precision of today experiments.

The last aspect of precision measurements at ATLAS mentioned here is B-physics, which includes:

- Precision measurements of CP-violation in B-mesons.
- Measurements of rare decays which are strongly suppressed in the Standard Model.
- Overconstraining the CKM matrix by a precise measurement of flavor oscillations in B_s^0 and B_d^0.

It is expected that ATLAS delivers several more precise results on B-physics than the present lower energy colliders. Some measurements can be even comparable to those from LHCb experiment. For these measurements an optimal performance of the ID is necessary, since it is responsible for the reconstruction of a second vertex and for the identification of the b quarks.

2.6 Start-up ATLAS News

"CERN announces the start-up date for LHC" *Geneva, August 7th 2008. CERN has today announced that the first attempt to circulate a beam in the Large Hadron Collider (LHC) will be made on September 10th. This news comes as the cool down phase commissioning CERN's new particle accelerator reaches a successful conclusion. Television coverage of the start-up will be made available through Eurovision.*

"First beam in the LHC" *Geneva, September 10th 2008. The first beam in the Large Hadron Collider at CERN was successfully steered around the full 27 kilometers of the world's most powerful particle accelerator at 10h28 this morning. This historic event marks a key moment in the transition from over two decades of preparation to a new era of scientific discovery. "It's a fantastic moment", said LHC project leader Lyn Evans, "we can now look forward to a new era of understanding about the*

origins and evolution of the universe." Figure 2.21 shows the first beam (a),(c) and (d) and the second beam (b) events seen at ATLAS.

"Incident in LHC sector 3-4" *Geneva, September 20th 2008. During commissioning (without beam) of the final LHC sector (sector 3-4) at high current for operation at 5 TeV, an incident occurred at mid-day on Friday 19 September resulting in a large helium leak into the tunnel.*

"LHC re-start scheduled for 2009" *Geneva, September 23rd 2008. Investigations at CERN following a large helium leak into sector 3-4 of the Large Hadron Collider (LHC) tunnel have indicated that the most likely cause of the incident was a faulty electrical connection between two of the accelerators magnets. Before a full understanding of the incident can be established, however, the sector has to be brought to room temperature and the magnets involved opened up for inspection. This will take three to four weeks. Full details of this investigation will be made available once it is complete.*

"CERN releases analysis of LHC incident" *Geneva, October 16th 2008. Investigations at CERN following a large helium leak into sector 3-4 of the LHC tunnel have confirmed that the cause of the incident was a faulty electrical connection between two of the accelerator's magnets. This resulted in mechanical damage and release of helium from the magnet cold mass into the tunnel. Proper safety procedures were in force, the safety systems performed as expected, and no one was put at risk. Sufficient spare components are in hand to ensure that the LHC is able to restart in 2009, and measures to prevent a similar incident in the future are being put in place. "This incident was unforeseen", said CERN Director General Robert Aymar, "but I am now confident that we can make the necessary repairs, ensure that a similar incident can not happen in the future and move forward to achieving our research objectives".* The summary of the analysis of the 19 September 2008 incident at the LHC can be found at [15].

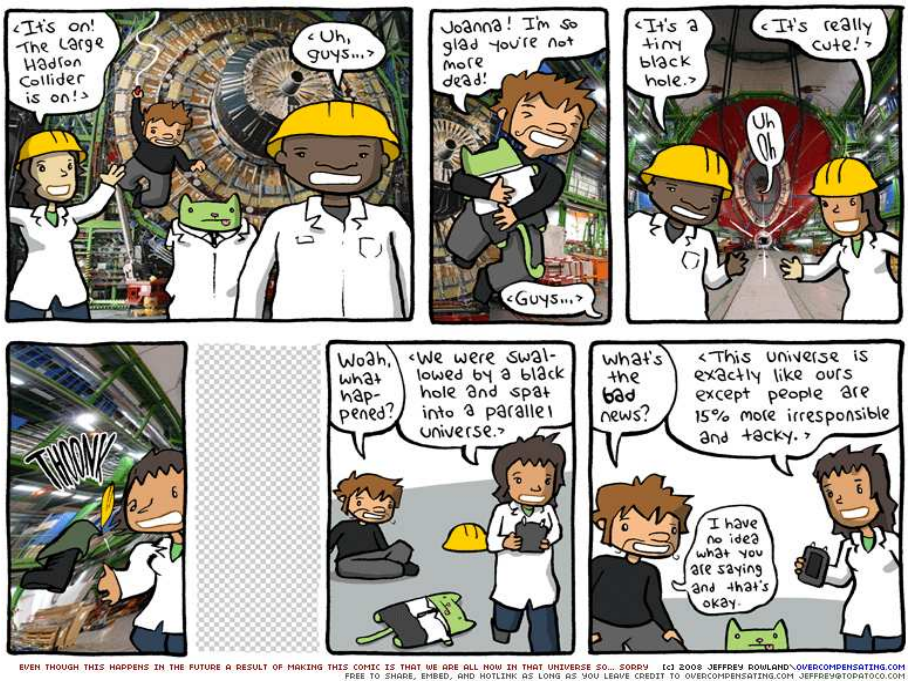


Figure 2.20: The Large Hadron Collider is on!

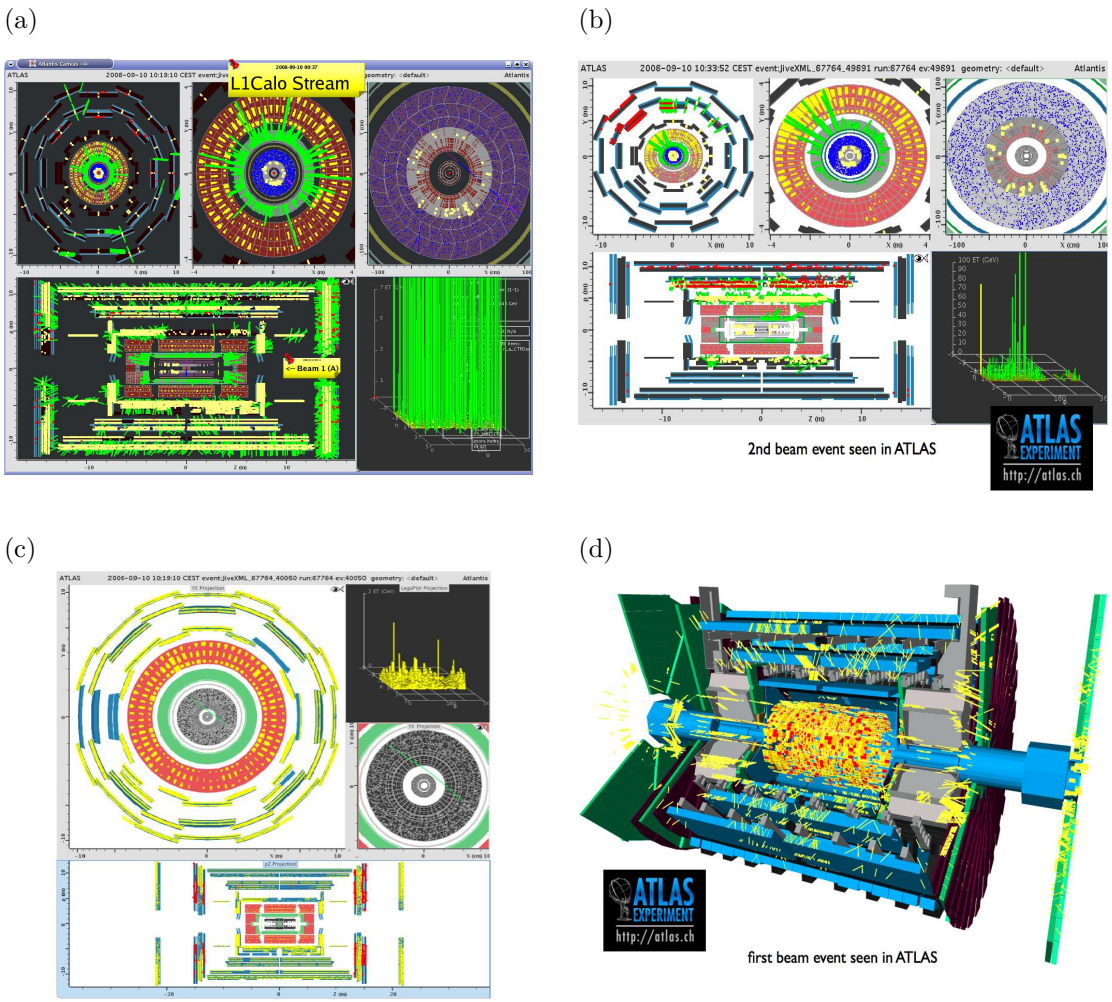


Figure 2.21: First single beam events seen by the ATLAS experiment, on 10th September 2008. Different views show the interaction of the beam with different parts of the detector ATLAS. To avoid risks, LHC decided to open the collimators progressively, each shot resulted in proton collisions with the collimators, visible in the detectors ("splash event"). ATLAS was running with Pixel and SCT barrel off and SCT endcaps, forward calorimeters and forward muon chambers at reduced HV.

"Follow up of the incident of 19 September 2008 at the LHC" *Geneva, December 5th 2008. CERN today confirmed that the Large Hadron Collider (LHC) will restart in 2009. This news forms part of an updated report, published today, on the status of the LHC following a malfunction on 19 September. "The top priority for CERN today is to provide collision data for the experiments as soon as reasonably possible," said CERN Director General Robert Aymar. "This will be in the summer of 2009." The initial malfunction was caused by a faulty electrical connection between two of the accelerator's magnets. This resulted in mechanical damage and release of helium from the magnet cold mass into the tunnel. Proper safety procedures were in force, the safety systems performed as expected, and no one was put at risk. Detailed studies of the malfunction have allowed the LHC's engineers to identify means of preventing a similar incident from reoccurring in the future, and to design new protection systems for the machine. A total of 53 magnet units have to be removed from the tunnel for cleaning or repair, of these, 28 have already been brought to the surface and the first two replacement units have been installed in the tunnel. The current schedule foresees the final magnet being reinstalled by the end of March 2009, with the LHC being cold and ready for powering tests by the end of June 2009. Full details of the timetable to restart are available in the report published today.* [http://press.web.cern.ch/press/PressReleases/Releases2008/PR17.08E.html]

"New schedule for the Large Hadron Collider" *Geneva, February 2009. The CERN Management today confirmed the restart schedule for the Large Hadron Collider resulting from the recommendations from the Chamonix workshop. The new schedule foresees first beams in the LHC at the end of September this year, with collisions following in late October. A short technical stop has also been foreseen over the Christmas period. The LHC will then run through to autumn next year, ensuring that the experiments have adequate data to carry out their first new physics analyses and have results to announce in 2010. The new schedule also permits the possible collisions of lead ions in 2010. This new schedule represents a delay of 6 weeks with respect to the previous schedule which foresaw LHC "cold at the beginning of July". The cause of this delay is due to several factors such as implementation of a new enhanced protection system for the busbar and magnet splices, installation of new pressure relief valves to reduce the collateral damage in case of a repeated incident, application of more stringent safety constraints, and scheduling constraints associated with helium transfer and storage. In Chamonix there was consensus among all the technical specialists that the new schedule is tight but realistic. The enhanced protection system measures the electrical resistance in the cable joints (splices) and is much more sensitive than the system existing on 19 September. The new pressure relief system has been designed in two phases. The first phase involves installation of relief valves on existing vacuum ports in the whole ring. Calculations have shown that in an incident similar to that of 19 September, the collateral damage (to the interconnects and super-insulation) would be minor with this first phase. The second phase involves adding additional relief valves on all the dipole magnets and would guarantee minor collateral damage (to the interconnects and super-insulation) in all worst cases over the life of the LHC. One of the questions discussed in Chamonix was whether to warm up the whole LHC machine in 2009 so as to complete the installation of these new pressure relief valves or to perform these modifications on sectors that were warmed up for other reasons. The Management has decided for 2009 to install relief valves on the four sectors that were already foreseen to be warmed up. The dipoles in the remaining four sectors will be equipped in 2010.*

Chapter 3

QCD and Boson+jets Production

3.1 Introduction

The ancient Greeks recognized just four elements, earth, wind, fire and water. These elements formed the building blocks of matter, of all they could see and of all they could touch. This concept laid dormant for almost two thousand years, until the middle ages and the allure of alchemy brought people to the systematic study of matter. In the nineteenth century, Mendeleev categorized about 60 elements before his death and this number has continued to grow to over 110 accepted today. These elements are, however, not considered now quite so elemental, although in some sense of course they remain so. Rutherford opened Pandora's box and what is now considered elemental, is smaller than the Greeks, or indeed anyone else, could have possibly imagined.

3.2 The Standard Model

The Standard Model (SM) [16, 17, 18] is an incredibly successful description of the interactions of elementary particles that best accommodate all experimental observations to date. It is a quantum field theory which describes the interactions of spin-$\frac{1}{2}$ point-like fermions, whose interactions are mediated by bosons. The spin-1 gauge bosons are a consequence of the $SU(3) \times SU(2) \times U(1)$ local symmetry, called gauge invariance.

The fermions can be grouped into two classes: leptons and quarks. The left-handed states are doublets under the $SU(2)$ group, while the right-handed states are singlets. There are three generations of fermions, each generation identical except for the increasing mass, as depicted in figure 3.1. The origin of this structure remains a mystery.

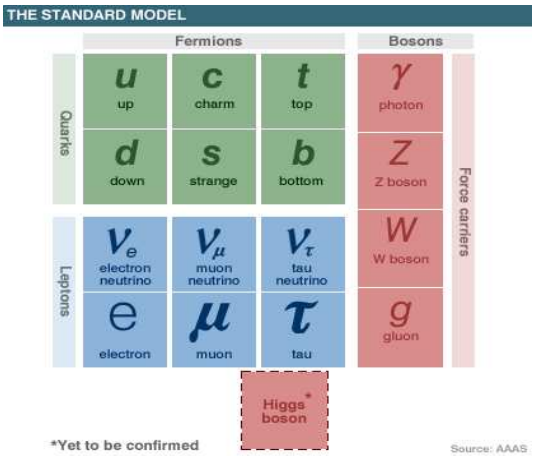


Figure 3.1: Standard Model picture.

There are three leptons with electric charge -1, the electron (e), the muon (μ) and the tau (τ) and three electrically neutral leptons, the neutrinos ν_e, ν_μ and ν_τ. Similarly, there are three quarks with electric charge $\frac{2}{3}$, up (u), charm (c) and top (t), and three with electric charge $-\frac{1}{3}$, down (d), strange (s) and bottom (b), shown in table 3.1 [19]. The quarks are triplets under the $SU(3)$ group and thus carry an additional charge, referred to as color. There is mixing between the three generation of quarks, which is parametrised by the Cabibbo-Kobayashi-Maskawa (CKM) matrix.

Fermion	particle	charge [e]	mass
"light" quarks	down (d)	-1/3	1.5 to 3.3 MeV
	up (u)	+2/3	3.5 to 6.0 MeV
	strange (s)	-1/3	$104 \pm^{26}_{34}$ MeV
"heavy" quarks	charm (c)	+2/3	$1.27 \pm^{0.07}_{0.11}$ GeV
	bottom (b)	-1/3	$4.2 \pm^{0.17}_{0.07}$ GeV
	top (t)	+2/3	171.2 ± 2.1 GeV
leptons	electron (e)	-1	0.511 MeV
	e-neutrino (ν_e)	0	$\ell^+\ell^-1$ MeV
	muon (μ)	-1	105.66 MeV
	μ-neutrino (ν_μ)	0	$\ell^+\ell^-1$ MeV
	tau (τ)	-1	1.777 GeV
	τ-neutrino (ν_τ)	0	$\ell^+\ell^-1$ MeV

Table 3.1: Fermions and their respective properties.

The $SU(2) \times U(1)$ symmetry group describes the so-called electroweak interaction. It can be spontaneously broken by the existence of a Higgs field [20, 21]. This leads to the emergence of massive vector bosons, the W and the Z, which are the mediators of the weak force, while the photon, massless, is the mediator of electromagnetism. The electroweak symmetry breaking is currently a topic of intense study, and indeed one of the primary goals of some of the most recent particle accelerators. The $SU(3)$ group describes the strong interaction, quantum chromodynamics (QCD). Eight vector gluons mediate this interaction. They carry color charges themselves, and are thus self-interacting. Properties of bosons are shown in table 3.2.

Boson	Force	charge [e]	mass
gluon (g)	strong force	0	0
photon (γ)	electromagnetic	0	0
$W^\pm$	weak force	±1	80.4 GeV
Z^0	weak force	0	91.2 GeV

Table 3.2: Vector bosons and the forces they mediate.

3.2.1 Beyond the Standard Model

It is noted that the SM does not contain an explanation for the particular pattern of quantum numbers we obtain, and neither does it explain the family structure. It also contains 19 free parameters, an unacceptable feature for a fundamental particle theory (three charged-lepton masses, six quark masses and four parameters to describe their mixing in weak interactions, three independent interaction strengths and a CP-violating parameter for the strong interaction, the $W^\pm$ and Higgs boson masses). The observation of non-zero neutrino mass is also difficult to incorporate into the SM.

Despite its successes, new theories that attempt to go beyond the Standard Model [22] are being developed, to account for some of these issues. One of the more extended new theories is the Supersymmetric (SUSY) model (detailed in section 3.7).

3.3 Quantum Chromodynamics (QCD)

Quantum Chromodynamics (QCD) is the gauge field theory that describes the strong interaction in the Standard Model. It was developed in 1973 [23, 24] in the context of Quantum Field Theory based in SU(3) symmetry group. It is a non-abelian theory and the corresponding Lagrangian, describing quarks q_k and gluon $\mathcal{A}^A_\alpha$ fields, is given by:

$$\mathcal{L}_{QCD} = \sum_{flavor} \bar{q}_a (i\gamma^\mu D_\mu - m_q)_{ab} q_b - \frac{1}{4} F^A_{\alpha\beta} F^{\alpha\beta}_A \tag{3.1}$$

where the sum runs over the six different flavors of the quarks, and $F^A_{\alpha\beta}$ is the field strength tensor for the spin-1 gluon field $\mathcal{A}^A_\alpha$,

$$F^A_{\alpha\beta} = \partial_\alpha \mathcal{A}^A_\beta - \partial_\beta \mathcal{A}^A_\alpha - g f^{ABC} \mathcal{A}^B_\alpha \mathcal{A}^C_\beta \tag{3.2}$$

where $D_\mu = \partial_\mu + ig\mathcal{A}^\alpha_\mu t^\alpha$ is the covariant derivative, t^α are the Gell-Mann matrices and f^{ABC} are the structure constants of the SU(3) color group. Indices A,B,C run over the eight color degrees of freedom of the gluon field. The charge associated to the strong interaction is the *color*. The third term of equation 3.2 shows the non-abelian nature of QCD. This terms describes the interaction between gluons. This self-coupling leads to the two main characteristics of the strong interaction: *asymptotic freedom* and *confinement*.

The strong coupling constant, $\alpha_s = \frac{g^2}{4\pi}$, depends on the scale of the interaction. At high energies (small distances), the strong interaction proceeds via color fields of reduced strength and the quarks and the gluons behave as essentially free, non-interacting, particles. This is what is called asymptotic freedom. It is important to remark that asymptotic freedom is the basis of the perturbative approach to QCD (pQCD) calculations. On the other hand, at low energies (large distances) the coupling strength asymptotically diverges, making impossible the production of free quarks. If the energy of the field between a pair of quarks $q\bar{q}$ increases (distance between them increases), at some point a new pair of $q\bar{q}$ is created. Therefore, only colorless bound states, hadrons, can be created. This situation is called confinement, and it is on the base of jet formation. The value of α_s as a function of the energy scale Q is shown in figure 3.2.

3.3.1 The Quark Parton Model and the Deep Inelastic Scattering

Deep inelastic scattering (DIS) experiments were performed in the 1960s in order to understand the proton and the neutron internal structure, giving place to the first approach to QCD theory, the Quark Parton Model, which first established the physical reality of quarks.

At lowest order, in deep inelastic electron-proton scattering a single photon, Z^0 (neutral current scattering) or $W^\pm$ (charged current scattering) is exchanged between the electron and the proton. The boson interacts with an object inside the proton called *parton*. In the context of QCD, partons are associated with quarks and gluons. The time scale of the hard scattering process is very short compared to that of inter-parton interactions, hence the other partons on the proton can be regarded as spectators in the scattering process. Figure 3.3 shows a graphical representation of a deep inelastic scattering of electron-photon.

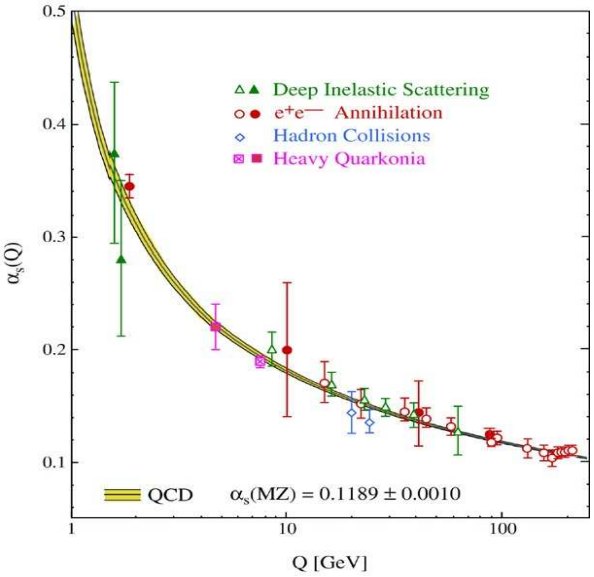


Figure 3.2: Value of the running strong coupling α_s as a function of the energy scale Q. Confinement region at low energies Q, and asymptotic freedom region at high energies Q can be distinguished in the figure.

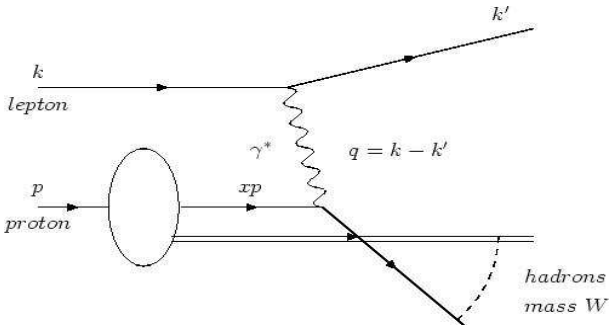


Figure 3.3: Schematic view of a neutral current deep inelastic scattering process: the photon interacts with a quark inside the proton, resolving the proton internal structure.

The four-momentum transfer squared from the electron to the photon, Q, can be used to describe the kinematics of the inclusive reaction $ep \rightarrow eX$ at a given center-of-mass energy, $\sqrt{s}$:

$$Q^2 = -q^2 = (k - k')^2 \tag{3.3}$$

where k and k' are the four-momenta of the ingoing and outgoing electron, respectively. Further we can define:

$$x = \frac{Q^2}{2\,(p \cdot q)} \tag{3.4}$$

where p is the 4-momentum of the incoming proton. The variable x can be interpreted as the fraction of the proton 4-momentum carried by the struck quark, and it is called the Bjorken variable. The kinematics of a DIS collision is completely determined by the two above variables (3.3) and (3.4) (Q and x). Figure 3.4 shows the kinematic range of DIS events accessible from different experiments.

In the Quark Parton Model, the DIS cross section can be considered as the incoherent sum of the scattering of point-like spin-$\frac{1}{2}$ particles, called quarks. The differential cross section for neutral current

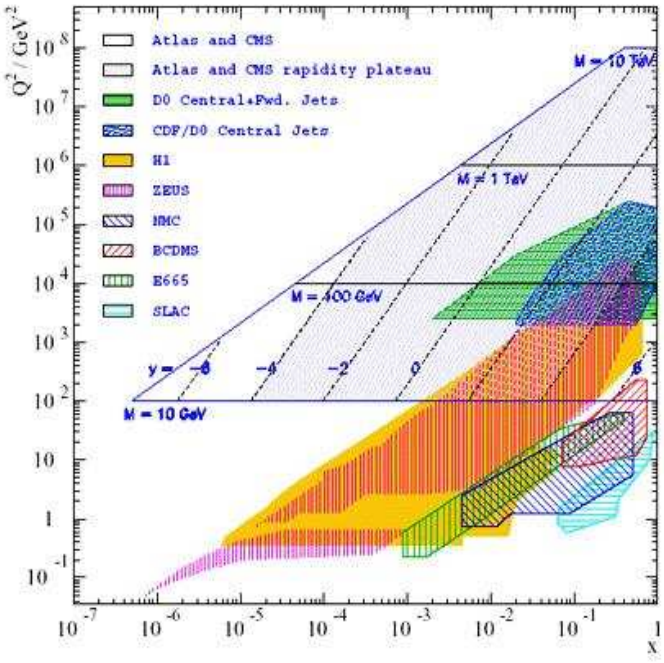


Figure 3.4: Kinematic range of Deep Inelastic Scattering events accessible from different experiments, including ATLAS and CMS.

$ep \to eX$ can be written as:

$$\frac{d^2\sigma_{NC}}{dx\,dQ^2} = \frac{4\pi\alpha^2}{xQ^4}\left(y^2x \cdot F_1(x,Q^2) + (1-y)\cdot F_2(x,Q^2)\right) \tag{3.5}$$

where y is the ratio $\frac{Q^2}{sx}$. F_1 and F_2 are the so-called structure functions of the proton, and are a convenient way to express the incalculable part of the hadronic vertex. They are defined as follows:

$$F_1(x) = \sum_i e_i^2\, f_i(x) \qquad \textit{and} \qquad F_2(x) = \sum_i e_i^2\, x\, f_i(x) \tag{3.6}$$

The term $f_i(x)\,dx$ gives the probability of finding a parton of type i inside the proton, carrying a fraction of the proton momentum between x and $x + dx$. And e_i is the charge of the parton in units of electron charge. The functions $f_i(x)$ are known as the *Parton Distribution Functions* (PDFs) (see section 3.3.3).

The second identity of (3.6) predicts that the cross section only depends on one variable, x. This property is called Bjorken scaling. Approximate scaling is observed in the data $x \approx 0.1$, but violation of scaling is observed for lower and higher x. Figure 3.5 shows F_2 versus Q^2 measured by different experiments.

A better description of the proton is given by Quantum Chromodynamics. QCD extends the naive Quark Parton Model by allowing interactions between the partons via the exchange of gluons.

3.3.2 Hadron-hadron Scattering

Scattering processes at high energy hadron colliders can be classified as either *hard*, where processes involve large amounts of transverse momenta, or *soft* (figure 3.6). QCD is the underlying theory for all such processes, but the approach and level of understanding is very different for the two cases. For hard processes, like high p_T jet production, the rates and event properties can be predicted with good precision using perturbation theory. For soft processes, like the underlying event, rates and properties

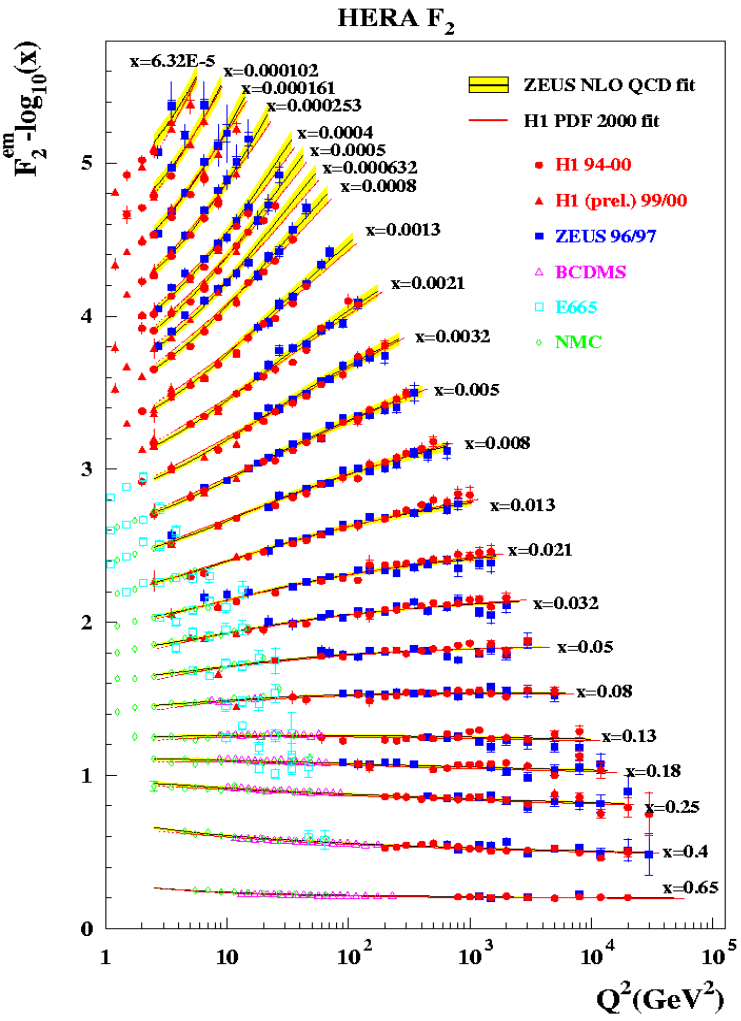


Figure 3.5: Structure function F_2 as a function of Q^2 and x measured in DIS and fix target experiments.

are dominated by non-perturbative QCD effects, which are less understood. For many hard processes, soft interactions are occurring along with the hard interactions and their effects must be understood for comparisons to be made to perturbative predictions.

It was first pointed out by Drell and Yan that parton model ideas developed for deep inelastic scattering could be extended to certain processes in hadron-hadron collisions. The paradigm process was the production of a massive lepton pair by quark-antiquark annihilation, the Drell-Yan process, and it was postulated that the hadronic cross section $\sigma(AB \to \mu^+\mu^- + X)$ could be obtained by weighting the subprocess cross section $\hat{\sigma}$ for $q\bar{q} \to \mu^+\mu^-$ with the PDFs $f_{q/A}(x)$ extracted from the deep inelastic scattering:

$$\sigma_{AB} = \int dx_a dx_b \; f_{a/A}(x_a) f_{b/B}(x_b) \; \hat{\sigma}_{ab \to X} \tag{3.7}$$

where, for the Drell-Yan process, $X = l^+l^-$ and $ab = q\bar{q}$, $\bar{q}q$. The good agreement between theoretical predictions and the measured cross sections provided confirmation of the parton model formalism and

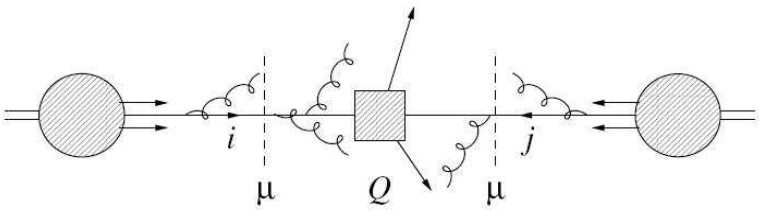


Figure 3.6: In hard hadron-hadron collision, constituent partons for each incoming hadron interact at short distance (large momentum transfer Q^2).

allowed for the first time rigorous, quantitative treatment of certain hadronic cross sections.

Problems appeared when perturbative corrections from real and virtual gluon emission were calculated. Large logarithms from gluons emitted collinear with the incoming quarks appeared to spoil the convergence of the perturbative expansion. These logarithms could be absorbed in the definition of the parton distributions, since they were the same as those that arise in deep inelastic scattering structure function calculations, giving rise to logarithmic violations of scaling. The key point was that all logarithms in the Drell-Yan corrections could be factorized into renormalized parton distributions and *factorization theorems* were derived, showing that this was a general feature of hard scattering processes. Taking into account the logarithm corrections, (3.7) becomes:

$$\sigma_{AB} = \int dx_a dx_b \; f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \; \hat{\sigma}_{ab\to X} \tag{3.8}$$

The Q^2 that appears in the PDFs is a large momentum scale that characterizes the hard scattering. Changes to the Q^2 scale of the order of $\mathcal{O}(1)$ are equivalent in this leading logarithm approximation.

The final step in the theoretical development was the recognition that the *finite* corrections left behind after the logarithms had been factored were not universal and had to be calculated separately for each process, giving rise to perturbative correction $\mathcal{O}(\alpha_s^n)$ to the leading logarithm cross section of (3.8). Schematically

$$\sigma_{AB} = \int dx_a dx_b \; f_{a/A}(x_a, \mu_F^2) \, f_{b/B}(x_b, \mu_F^2) \; \times \; [\; \hat{\sigma}_0 + \alpha_s(\mu_R^2) \; \hat{\sigma}_1 + \dots \;]_{ab\to X} \tag{3.9}$$

where μ_F is the *factorization scale*, which can be thought of as the scale that separates the long- and short-distance physics, and μ_R is the *renormalization scale* for the QCD running coupling. Formally, the cross section calculated at all orders in perturbation theory is invariant under changes in these parameters, the μ_F^2 and μ_R^2 dependence of the coefficients ($\hat{\sigma}_1$,...) exactly compensates the explicit scale dependence of the parton distributions $f_{a/A}(x_a, \mu_F^2)$, $f_{b/B}(x_b, \mu_F^2)$ and the coupling constant $\alpha_s(\mu_R^2)$. As more terms are included in the perturbative calculation, more exact becomes this compensation. In the absence of a complete set of higher order corrections, a specific choice for the two scales must be done. This choice can be translated to different numerical results. Typically it is often assumed $\mu_F = \mu_R$ and of the order of the typical momentum scales of the hard scattering process.

In a hadronic interaction, as shown in figure 3.7, partons that participate in the short distance scattering (hard scattering) can emit radiation (quarks and gluons), prior to the scattering yielding the possibility of *initial state radiation (ISR)*. The remnants of the original hadron are no longer color single states and will interact, presumably softly, with each other generating an underlying distribution of soft partons, the beginning of the *underlying event (UE)*. After large momentum transfer scattering processes, new gluons and quarks may be produced by another radiation step (*final state radiation, FSR*). Mainly these partons will have low energy and/or be nearly collinear with the hard scattered partons. The FSR and ISR are described in Monte Carlo programs.

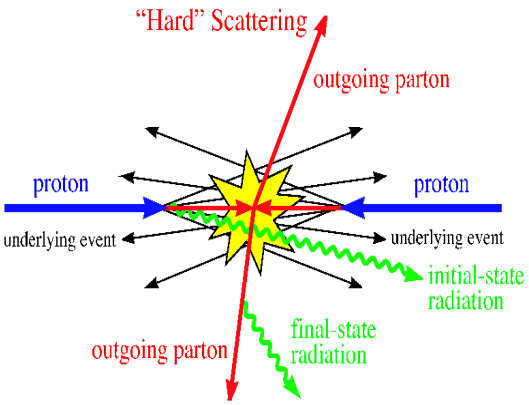


Figure 3.7: Pictorial representation of a hard scattering event.

The final step in the evolution of long distance states involves a non-perturbative hadronization process, that organizes the colored degrees of freedom from the showering and from the softer interactions of other initial state partons into color-singlet hadrons with physical masses. The hadronization step is accomplished in a model dependent fashion in different Monte Carlos (see section 3.4.2).

In a hadronic event containing jets, the underlying event (UE) consists of all event activity except the high-p_T jets from the hard scattering process. The underlying structure of hadronic interactions has not been fully understood yet and it is not clear how it should be modeled. The underlying event contains components associated with beam-beam remnant interactions, particles coming from the initial and final state radiation, from color strings stretching between the underlying event and the highest-p_T jet and from secondary parton interactions. To measure the underlying event using di-jet events, for instance, the leading jet (the jet with highest transverse momentum) is used to select the event (jet event illustrated in figure 3.8(a)). Then, three different regions are defined in terms of the azimuthal angle between charged particles and the leading charged jet. This angular difference is given by $\Delta\phi = \phi_{particle} - \phi_{leadjet}$. The region $|\Delta\phi| < 60°$ is referred to as toward the leading charged jet and the region $|\Delta\phi| > 120°$ is called away from the leading jet. The region transverse to the leading jet is defined by $60° < |\Delta\phi| < 120°$, and is used to study the underlying event. The event regions defined by $\Delta\phi$ are illustrated in figure 3.8(b).

3.3.3 Parton Distribution Functions

Parton Distribution Functions (PDFs) are currently determined from experiment. The most comprehensive analysis are being done by the Coordinated Theoretical-Experimental Project on QCD (CTEQ) [25], used in this analysis, and by Martin, Robers, Stirling and Thorne team (MRST) [26] groups. These groups perform a global fit to data experiments of several different types. To perform such fit one chooses a parametrization for the parton distributions at some standard factorization scale μ_0. Given some set of values for the parameters describing $f_{a/A}(x, \mu_0)$, one can determine $f_{a/A}(x, \mu)$ for higher values of μ by using evolution equations. Then the QCD cross section formulae give predictions for all the experiments that are being used. To obtain the best fit to all the experiments, one varies the parameters in $f_{a/A}(x, \mu_0)$. There are roughly 1400 data points in the CTEQ fit and only 25 parameters available to fit these data. Figure 3.9 shows the PDFs using different parametrizations, using ZEUS data [27].

(a) (b)

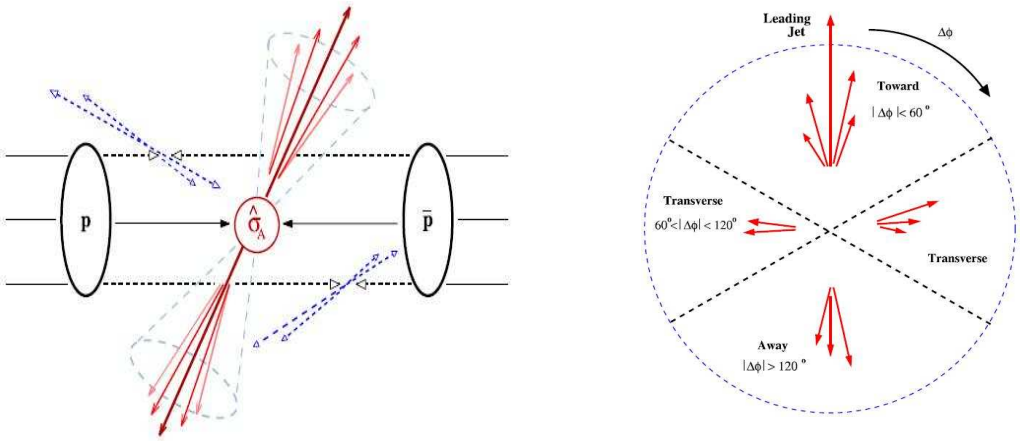


Figure 3.8: Illustration of a jet produced by a hard parton-parton scattering in a pp collision (a). Event regions defined in terms of the azimuthal angle between charged particles and the leading charged jet, $\Delta\phi = \phi_{particle} - \phi_{leadjet}$ (b).

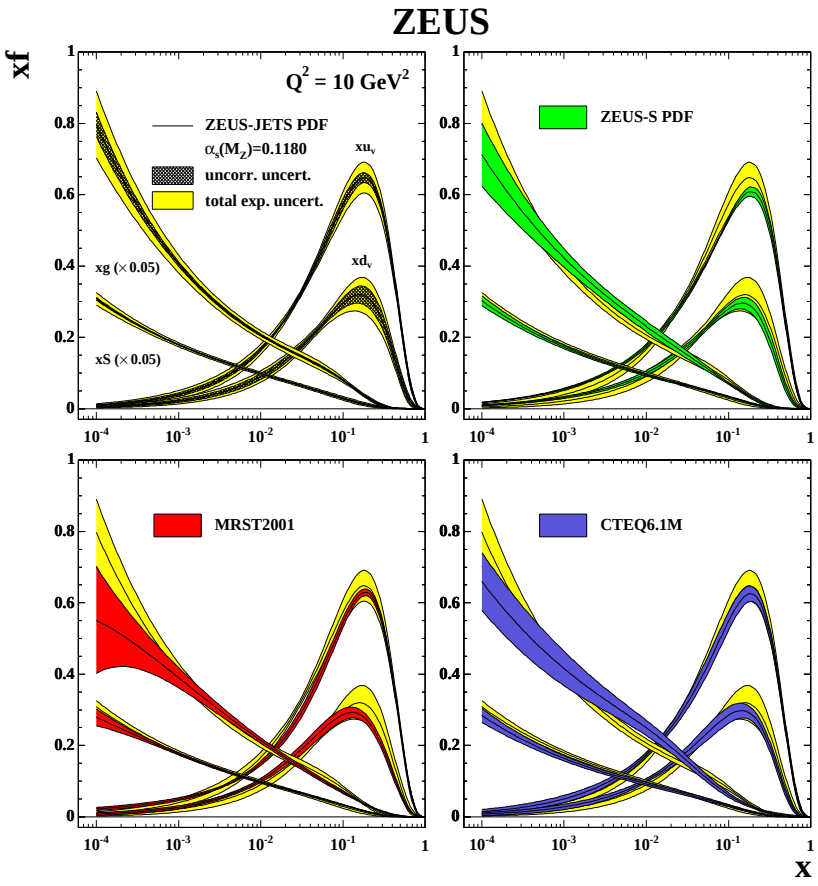


Figure 3.9: Gluon, sea, up and down PDFs extracted from the ZEUS-JETS fit and the uncorrelated and total error bands (a). PDFs extracted form the ZEUS-JETS fit compared to ZEUS-S PDFs (b). PDFs extracted from the ZEUS-JETS fit compared to MRST2001 PDFs (c). PDFs extracted from the ZEUS-JETS fit compared to CTEQ6.1 PDFs. The total experimental uncertainty bands are shown for each PDF set.

3.3.4 Boson + jets pQCD Predictions

Processes involving the production of $W^{\pm}/Z^0$ bosons in association with jets are of great importance at hadron collider experiments. The presence of a hard scatter means that these interactions occur at a scale that should make perturbative QCD applicable, and thus are excellent channel to test such predictions. In addition, due to the presence of a vector boson, the process scale is at energies of $Q^2 \approx M_Z^2$.

The Z +jets production cross section can be decomposed into its multi-jet components:

$$\sigma_{Z+jets} = \sigma_{Z+1jet} + \sigma_{Z+2jets} + \sigma_{Z+3jets} + \dots \tag{3.10}$$

where every jet multiplicity cross section can be developed in perturbative terms on α_s as follows:

$$\sigma_{Z+1jet} = a_1\alpha_s + a_2\alpha_s^2 + a_3\alpha_s^3 + a_4\alpha_s^4 + \dots \tag{3.11}$$

$$\sigma_{Z+2jets} = b_2\alpha_s^2 + b_3\alpha_s^3 + b_4\alpha_s^4 + \dots \tag{3.12}$$

$$\sigma_{Z+3jets} = c_3\alpha_s^3 + c_4\alpha_s^4 + \dots \tag{3.13}$$

$$\dots$$

The a_i, b_i, c_i coefficients are in general function of the jet definition, in particular the cone size used to cluster the partons into jets, and the transverse momentum, rapidity and separation cuts imposed on the jets. Lowest order on α_s contributions in these expansions, i.e. the leading order (LO) contributions $(a_1, b_2, c_3, ...)$, can be calculated from the matrix elements. Figure 3.10(a) shows the tree level diagrams that contribute to the LO Z +2jets and figure 3.10(b) to the LO Z +3jets. Next-to-leading order (NLO) corrections (a_2, b_3, c_4) are nowadays known only for Z +1 and Z +2 jet processes. The NLO contributions include virtual and real corrections to the tree level processes:

$$\sigma_{Z+n}^{NLO} = \sigma_{Z+n}^{virtual} + \sigma_{Z+n}^{real} \tag{3.14}$$

$$= \sigma_{Z+n}^{1\text{-}loop} + \sigma_{Z+(n+1)}^{tree\text{-}level} \tag{3.15}$$

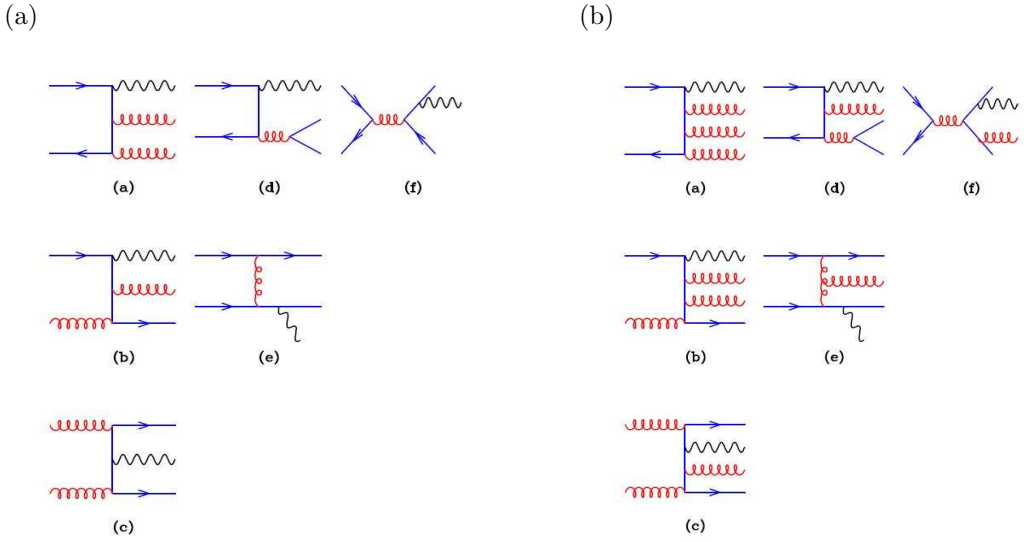


Figure 3.10: Sample tree-level diagrams for the process parton + parton → W/Z+ 2 partons (a) and for the process parton + parton → W/Z+ 3 partons (b). As usual the vector boson is denoted by a wavy line.

Virtual NLO corrections consist in one-loop corrections to the LO diagrams. Some examples of one-loop diagrams for Z +1jet are shown in figure 3.11. Real corrections include the diagrams with the radiation of an extra parton in the final state (n+1).

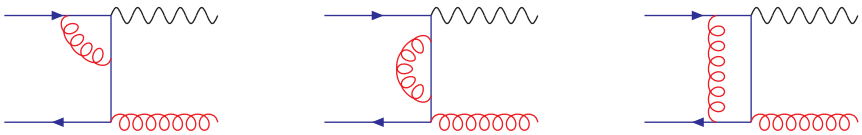


Figure 3.11: Some examples of Feynman one-loop diagrams, contributing to virtual NLO order corrections for Z +1jet processes.

Next-to-next-to-leading order contributions (NNLO) include two-loop corrections to the LO diagrams, one-loop corrections to the NLO diagrams and the radiation of an extra parton. Predictions for NNLO are not available yet for Z+jets processes, since the number of diagrams increases dramatically with the order of the calculation.

QCD corrections have been very well studied and a variety of calculations and Monte Carlo (MC) generators exist. Next-to-leading order (NLO) perturbative QCD predictions for W,Z + 1, 2 jets signatures are available (in the codes MCFM [28] and DYRAD [29]). W/Z + 3 jets predictions are only available at leading order (LO) calculation, at the moment. MCFM (*A Monte Carlo for FeMtobarn processes at Hadron Colliders*) is a parton-level Monte Carlo program which gives NLO predictions for a range of processes at hadron colliders, allowing to obtain full perturbative NLO predictions for any infra-red safe variable, like the cross section. MCFM has been used in this analysis, more information can be found in appendix B.

The principal motivation for performing a next-to-leading order calculation is to reduce the uncertainties in leading order predictions. In particular, any perturbative prediction contains an unphysical dependence on renormalization and factorization scales (often chosen to be equal, as we shall do here). The magnitude of cross-sections and the shape of differential distributions can vary greatly between two different choices of scale, which is often interpreted as an inherent theoretical uncertainty which is then ascribed to the predictions. Another strategy is to argue for a particular choice of scale, based on the physics of the process under consideration. A next-to-leading order calculation is an invaluable tool for investigating the issue of scale dependence. The logarithms that are responsible for the large variations under changes of scale at leading order are exactly canceled through to next-to-leading order. As a result, one expects that next-to-leading order predictions are more stable under such variations. In addition, the next-to-leading order result may provide further evidence to support a particular scale choice that may have been deemed appropriate at leading order. As an example, figure 3.12(a) shows the scale dependence of the exclusive Z + 1 jet differential cross-sections $d\sigma/dp_T$, integrated over the range $15 < p_T < 200$ GeV. Here, calculations for the next-to-leading order predictions have been performed within the program MCFM. The leading order prediction rises sharply as the scale is decreased, while the corrections produce a far flatter curve that exhibits a much less pronounced dependence on the scale choice. The corresponding results for the 2-jet processes are shown in figure 3.12(b). As in the 1-jet case, the renormalization and factorization scales are set equal. As anticipated, the prediction for NLO shows a considerable reduction in scale dependence.

(a) (b)

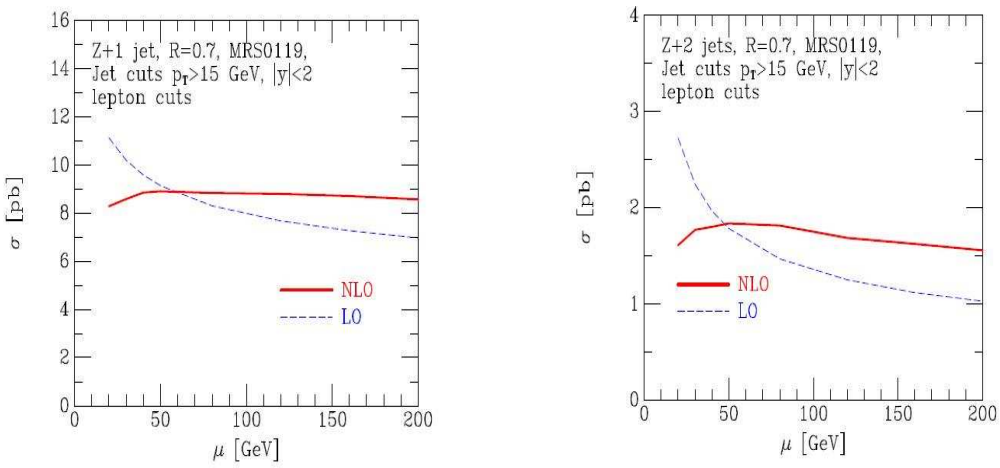


Figure 3.12: The scale dependence of the Z +1 jet predictions (a) and of the Z +2 jet predictions (b), with the factorization and renormalization scales equal and given by μ, for LO and NLO.

3.4 QCD Phenomenological Models

3.4.1 The Underlying Event

The description of the underlying event energy requires a non-perturbative phenomenological model. There are currently a number of models available, primarily inside parton shower Monte Carlo programs, to predict this process. An understanding of this soft physics is interesting in its own right but is also essential for precision measurements of hard interactions where the soft physics effects need to be subtracted.

One of the models for the underlying event is provided by JIMMY package, which should be linked to HERWIG Monte Carlo event generator. It is designed to generate multiple parton scattering events in hadron-hadron events. JIMMY implements the ideas of the eikonal model, which assumes that individual hard scatters are uncorrelated. The recent version 4.3 of JIMMY (which is linked to HERWIG version 6.5), works efficiently as a generator of underlying events in high E_T jet events and other hard processes in hadron-hadron collisions. For a given PDF set, the main adjustable parameter is the minimum transverse momentum of partonic scattering, and a variable related to the effective proton radius. Varying these one is able to get a good description of data. However, poor description of minimum bias data in which there is no hard scale is still obtained. This is probably due to the fact that the cut on transverse momentum is a hard cutoff and there is no soft component below it. JIMMY predicts twice as much underlying event activity as PYTHIA at the LHC, with tunings that give equally good descriptions of current data.

The PYTHIA model for the underlying event utilizes a multiple parton interaction framework with the total rate for parton-parton interactions assumed to be given by perturbative QCD. A cutoff, p_{Tmin}, is introduced to regularize the divergence as the transverse momentum of the scattering goes to zero. This is the main free parameter of the model and basically corresponds to an inverse color screening distance. The version of PYTHIA used in the analysis (version 6.4) has a more sophisticated treatment of color, flavor and momentum correlations in the remnants.

3.4.2 Parton Showers

Figure 3.13 depicts the basic structure of showering and hadronization processes, which are now going to be detailed.

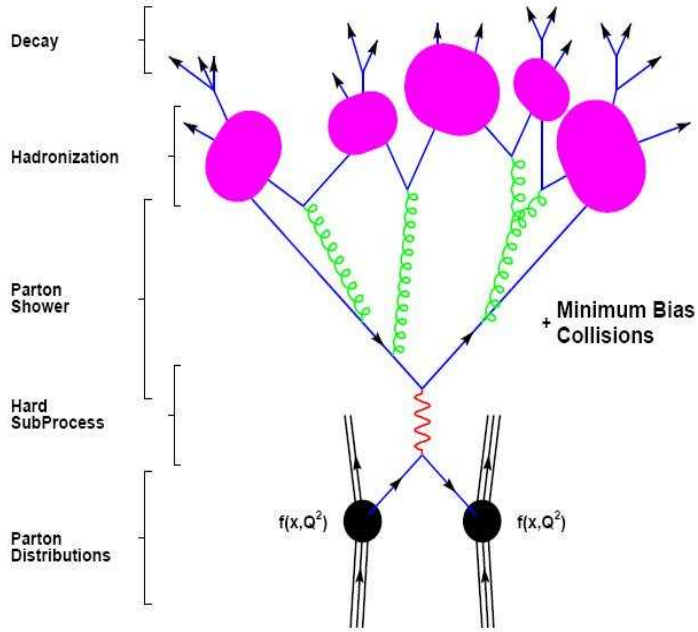


Figure 3.13: The basic structure of showering and hadronization processes are shown here.

In both initial and final-state showers, the structure is given in terms of decays $a \to bc$, specifically $q \to qg$, $q \to q\gamma$, $g \to gg$, and $g \to q\bar{q}$. Each of these processes is characterized by a splitting function $P_{a\to bc}(z)$. The branching rate is proportional to the integral $\int P_{a\to bc}(z)\, dz$. The z value picked for a decay describes the energy sharing, with daughter b taking a fraction z and daughter c the remaining $1-z$ of the mother energy. Once formed, the daughters b and c may in turn branch, and so on, as shown in figure 3.14. Each parton is characterized by some virtuality scale Q^2, which gives an approximate sense of time ordering to the cascade. In the initial-state shower, Q^2 values are gradually increasing as the hard scattering is approached, while Q^2 is decreasing in the final-state showers. Shower evolution is cut off at some lower scale Q_0, typically around 1 GeV for QCD branchings. From above the process, a maximum scale Q_{max} is introduced, where the showers are matched to the hard interaction itself. The relation between Q_{max} and the kinematics of the hard scattering is uncertain, and the choice made can strongly affect the amount of well-separated jets. Initial-state radiation is handled within the backwards evolution scheme. This is done starting from the two incoming partons at the hard interaction, tracing the showers backwards in time, back to the two shower initiators. In other words, given a parton b, one tries to find the parton a that branched into b. The evolution in the Monte Carlo is therefore performed in terms of a sequence of decreasing Q^2 and increasing momentum fractions.

3.4.3 Matrix Element

In order to compute exactly the result of a given number of emissions a possibility is to consider only those diagrams corresponding to the emission of real particles. Basically, the number of emissions coincides with the perturbative order in α_s. This option forms the core of *Tree Level Matrix Element* generators.

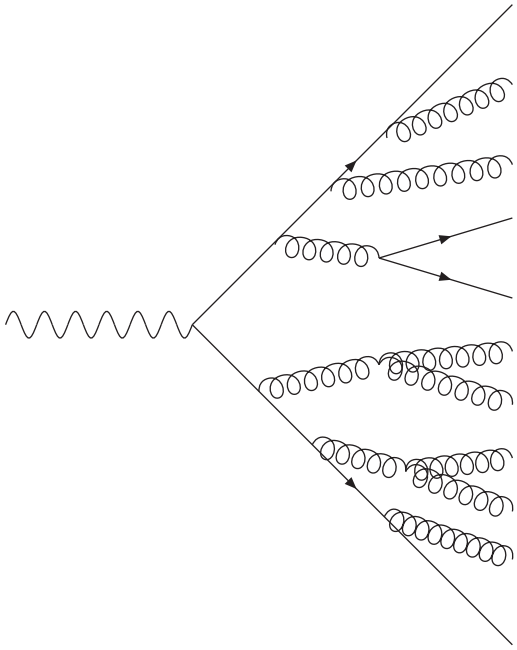


Figure 3.14: A parton cascade or shower development of successive branchings. Each outgoing line is source of a new cascade, until all outgoing lines have stopped branching (cutoff scale Q_0), then outgoing partons have to be converted into hadrons via a hadronization model.

In general, these codes are used within an event generator. Currently, the maximum number of emissions allowed is between five and ten.

The computation of tree-level matrix elements with a fixed number of legs (i.e. fixed number of partons in the final state) is performed with parton-level generators, like ALPGEN (see section 3.5.1). They describe a specific final state to lowest order in perturbation theory (virtual loops are not included in the matrix elements). These programs generally do not include any form of hadronization, thus the final states consist of bare quarks and gluons. The combination of tree-level matrix element generators and showering and hadronization programs, as HERWIG or PYTHIA, is essential, for instance, for analyses based on multi-jet configurations. The combination of matrix elements with parton showers is the origin of extensive studies. Matrix element generators are good for well separated jets and parton shower generators for the structure inside jets, so clearly the two complement each other and a marriage is desirable. To do this, without double counting[1] or gaps in the phase space coverage, is less trivial, and several alternative approaches have been developed, as merging, vetoed parton showers, etc. (more details can be found in [31, 32]).

3.4.4 Hadronization

QCD perturbation theory, formulated in terms of quarks and gluons, is valid at short distances. At long distances, QCD becomes strongly interacting and perturbation theory breaks down. In this confinement regime, the colored partons are transformed into colorless hadrons, a process called either hadronization or fragmentation. So, hadronization is described as the mechanism by which quarks and gluons produced in hard processes form the hadrons that are observed in the final state. This is an intrinsically non-perturbative process.

[1] Combining matrix element with parton shower processes involves the risk that the same parton which was already generated in the matrix element is added once again during the shower evolution. A solution for this problem, known as double counting, is the *matching* procedure to remove additional jets, which occur twice, matching each hard jet generated in the matrix element with a parton shower jet [30].

To try to describe the mechanism of hadron formation we can so far resort to different phenomenological models. The main current models are *cluster* and *string* hadronization, schematically shown in figure 3.15. These models are used, for instance, in HERWIG and PYTHIA event generators, respectively.

- *Cluster Model* [33, 34]: The model starts by splitting gluons non-perturbatively, $g \rightarrow q\bar{q}$, after the parton shower. Color-singlet $q\bar{q}$ combinations have lower masses and a universal spectrum due to the color pre-confinement [35, 36, 37] property of the shower. These color-singlet combinations are assumed to form clusters, which mostly undergo simple isotropic decays into pairs of hadrons, chosen according to the density of states with appropriate quantum numbers. This model has few parameters and a natural mechanism for generating transverse momenta and suppressing heavy particle production in hadronization. However, it has problems in dealing with the decay of very massive clusters, and inadequately suppressing baryon and heavy quark production.

- *Lund String Model* [38, 39]: This model is based on the dynamics of a relativistic string, representing the color flux stretched between the initial $q\bar{q}$. The string produces a linear confinement potential and an area law for matrix elements:

$$|M(q\bar{q} \rightarrow h_1 \cdots h_n)|^2 \propto e^{-bA} \tag{3.16}$$

  where A is the area swept out by the string in energy-momentum space. The strings breaks up into hadrons via $q\bar{q}$ pair production in its intense color field. Gluons produced in the parton shower give rise to *kinks* on the string. The model has extra parameters for the transverse momentum distribution and heavy particles suppression. It has some problems describing baryon production, but less than the cluster model.

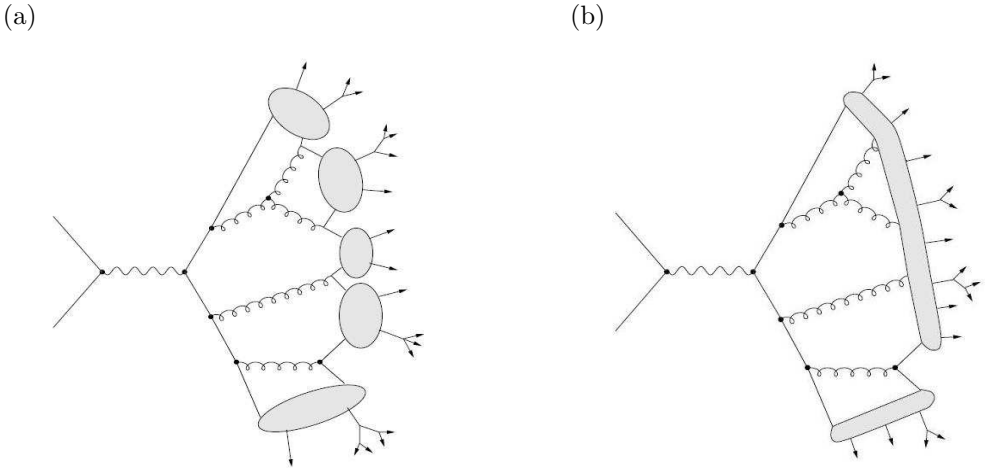


Figure 3.15: Cluster (a) and string (b) hadronization models.

3.5 Monte Carlo Generators

This section is devoted to briefly describe some of the most used Monte Carlo generator programs, linked to the previous phenomenological processes described above. Theoretical predictions form a necessary part of any particle physics experiment. A successful way to produce hypothetical events with the distribution predicted by theory is through the so-called event generators. The spectrum of event generators is very broad, from general-purpose ones to more specialized ones. Among more specialized programs, many deal with the matrix elements for some specific set of processes, a few with topics such as parton showers or particle decays, but there are no freestanding programs that handle hadronization. In the end, many specialized programs are therefore used as plugins for the general-purpose ones. In the following, we focus on the generators programs which are used in the analysis.

3.5.1 ALPGEN

ALPGEN [40] is a Monte Carlo generator for hard multi-parton processes in hadronic collisions. The program performs leading order (LO) QCD exact matrix elements calculations for a large set of parton-level processes of interest in the study of LHC and Tevatron data. Parton-level events are generated providing full information on their color and flavor structure, enabling the evolution of the partons into fully hadronized final states. So, the development of partonic cascades, with the subsequent transformation of the partons into observable hadrons are carried out by Monte Carlo programs such as HERWIG (see section 3.5.3) or PYTHIA (section 3.5.2). The consistent combination of the parton-level calculations with the partonic evolution given by the shower MC programs is the subject of extensive work. In the case of $W/Z(\rightarrow ll) + N$ jets process, N can reach for the moment, $N \leq 5$.

3.5.2 PYTHIA

The PYTHIA [41] program is a general purpose generator for hadronic events in pp, e^+e^- and e^+p high-energy colliders, comprising a coherent set of physics models for the evolution from a few-body hard processes to a complex multi-hadronic final state. PYTHIA contains around 240 different 2→n processes, all at leading order. It is mainly optimized for 2→1 processes and 2→2 processes. The physics areas covered include QCD 2→2 partonic scattering, SM processes including $\gamma, Z^0, W^\pm$ singly or in pairs or SUSY processes, among others. In the generation of the basic partonic processes listed above, initial- and final-state showers are added to provide more realistic multipartonic configurations. PYTHIA for ATLAS uses Q^2-ordered parton showers. It uses the Lund string hadronization model. Unstable particles are allowed to decay. In cases where better decay models are available elsewhere, e.g. for $\tau^\pm$ with spin information or for B hadrons, such decays can be delegated to specialized packages, as PythiaB. New PYTHIA version v6.4 [42] has introduced major changes to the description of minimum bias interactions and the underlying event. There is a new, more sophisticated scenario for multiple interactions, new p_T-ordered initial- and final-state showers (ISR and FSR) and a new treatment of beam remnants.

PYTHIA has a number of different physical parameters. The basic parameter is the lower limit of the transverse momentum, p_T-min, used in the calculation of $2 \rightarrow 2$ "hard" cross section. The minimum p_T is used as a regulating parameter. In addition, a uniform distribution of matter and a single or double Gaussian matter distribution can be used, introducing fluctuations in the multiplicity distributions, particularly at large multiplicities. The final ATLAS tuning of the PYTHIA model [43, 44, 45] for version 6.3, uses the complex scenario with a double Gaussian matter distribution (MSTP(82)=4) with a core-size (PARP(84)) of 0.5 compared to the default of 0.2, and a slightly lower p_T-min (PARP(82)) of 1.8 GeV compared to the default of 1.9 GeV.

3.5.3 HERWIG

HERWIG [46] is a Monte Carlo event generator for simulation of hadronic final states in lepton-lepton, lepton-hadron and hadron-hadron collisions. The program includes a large range of hard scattering processes together with initial- and final-state radiation using the angular-ordered parton shower. It uses the cluster hadronization model and a cluster-based simulation of the underlying event. HERWIG contains hard 2→n scattering processes for both Standard Model and its supersymmetric extension. The following types of processes are included: 2→2 QCD scattering processes, including heavy flavour production, electroweak processes, like $\gamma/\gamma^*/Z^0/W^\pm/H^0$ production, SUSY, etc. HERWIG needs to be linked to a dedicated package, JIMMY [47] to produce the underlying event activity (see section 3.4.1).

3.5.4 MCatNLO

The MCatNLO [48, 49] event generator includes the full next-to-leading order QCD corrections in the computation of hard processes. The package includes hadronic collisions, with the production of the following final states: W^+W^-, $W^\pm Z$, ZZ, $b\bar{b}$, $t\bar{t}$, H^0, etc. Incorporating the NLO matrix elements provides a better prediction of the rates while improving the description of the first hard hadron emission. As with any other parton shower based Monte Carlo, MCatNLO is capable of giving a sensible description of multiple soft/collinear emissions. For the same reason, and at variance with usual NLO programs, propagation through the shower and subsequent hadronization gives a final state description at the hadron level. One feature of MCatNLO as opposed to standard MC's is the presence of negative weights (which appear in higher order perturbative calculations, NLO in this case). Thus in unweighted event generation MCatNLO produces unit weight events with a fraction (typically ∼15%) having weight -1. MCatNLO is based on HERWIG generator, so from a technical point of view, the structure of MCatNLO is separated in a NLO former part, which produces an event file, and shower and hadronization part, which is just HERWIG, augmented by the capability of reading the event file.

3.6 Jets in Hadron Colliders

Jets are clusters of "objects" that have to be defined by a clustering algorithm, a set of mathematical rules. These clusters of particles should have kinematic properties that can be related to the corresponding properties of the energetic partons produced in the hard scattering process. There are two broad groups of jet algorithms, those based in one form or another on cones, *cone algorithms*, and those that involve repeated recombination of particles that are nearby in some distance measure, *sequential recombination algorithms*. From a theoretical point of view, the following features are desirable for an ideal jet algorithm:

- *Infrared safety*: The presence of additional soft particles between two particles belonging to the same jet should not affect the recombination of these two particles into a jet. In the same sense, the absence of additional particles between these two should not disturb the correct reconstruction of the jet. An example is shown in figure 3.16.
- *Collinear safety:* A jet should be reconstructed independent of the fact that a certain amount of transverse momentum is carried by one particle or by two collinear particles.
- *Invariance under boosts:* The algorithm should find the same solution independent of boosts in the longitudinal direction. This feature is important in *pp* colliders where the center of mass of the parton-parton collision can be boosted with respect to the *pp* center of mass.
- *Boundary stability:* The kinematic variables used to describe the jets should have kinematic boundaries that are insensitive to the details of the final state.

- *Order independence:* The algorithm should find the same jets at parton, hadron and detector level. This feature is desirable from theoretical and experimental point of view (figure 3.17).
- *Straightforward implementation:* The algorithm should be straightforward to implement in perturbative calculations.

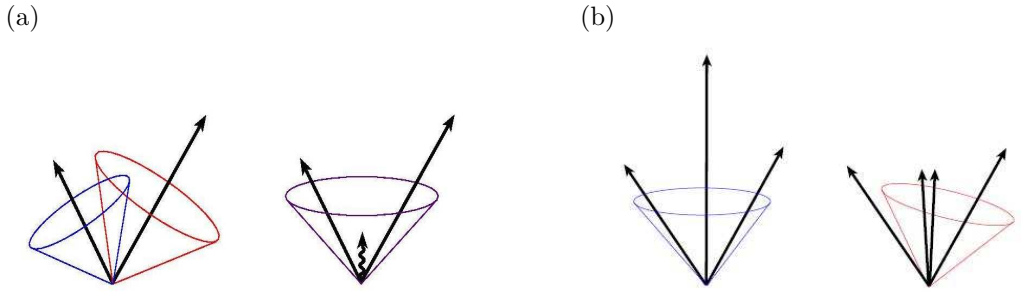


Figure 3.16: An illustration where the presence of soft radiation between two jets may cause a merging of the two jets (right) that would not occur in the absence of the soft radiation (left). An example of infrared sensitivity (a). One possible collinear problem, possible sensitivity to E_T ordering of the particles that act as seeds (b).

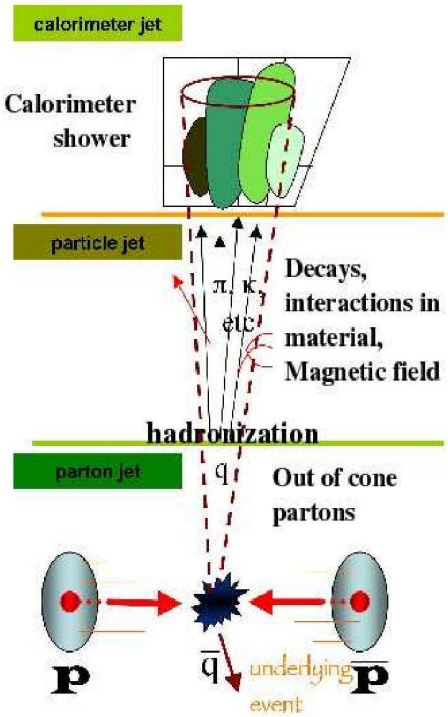


Figure 3.17: A representation of the different stages of jet production and reconstruction: parton, hadron or particle and calorimeter or detector level, where a good jet algorithm should find the same jets.

Jet algorithms used in ATLAS at present are detailed in section 5.1.3. Literature about jets in hadron-hadron collisions can be found in [50, 51, 52].

3.7 Supersymmetry

Supersymmetry (SUSY) imposes a new symmetry between the fermions and bosons [53, 54, 55, 56, 57]. The supersymmetric extensions of the Standard Model (SM) make improvements to phenomenological problems in the physics of elementary particles: it provides a natural solution for the gauge hierarchy problem[2][58], if sparticles exist at the TeV scale. Moreover, the extrapolation of LEP data within the framework of supersymmetric extension yields a precise unification of gauge couplings at a scale of $\sim 10^{16}$ GeV [59, 60]. Due to these properties, SUSY is one of the most attractive alternatives beyond the SM and has been the subject of many studies in particle physics. However, up to now, no direct evidence for SUSY has been found. It is essential to examine the properties of any new states of matter at energy scales close to the threshold for the new phenomena, or in high energy collisions at the TeV energy scale.

Sparticle production of gluinos ($\tilde{g}$) and squarks ($\tilde{q}$) occurs dominantly via strong interactions and its rate may be expected to be considerably large at the LHC. Gluino production leads to a large rate for events with multijets via series of cascade decays and the neutral lightest supersymmetric particle (LSP) in the final state which remains stable and undetectable, if R-parity is conserved. LSP's carry off apparently large missing transverse energy (E_T^{miss}). If no third generation partons in the initial state are considered and gluino and squark production rates are fixed by QCD in terms of the gluino and squark masses ($m_{\tilde{g}}$ and $m_{\tilde{q}}$), then inclusive SUSY searches with the early data rely on excesses of events in the channel of "multijets + large E_T^{miss}" [61] which is a model-independent feature. Experimental signatures, corresponding SUSY scenarios and their background processes are summarized in table 3.3 . They cover realistic supersymmetric models of minimal supergravitiy (mSUGRA) [62, 63, 64, 65], anomaly-mediated SUSY breaking (AMSB) [66, 67] and gauge-mediated SUSY breaking (GMSB) [68, 69].

jet multiplicity	additional signature	covered scenario	background
≥ 4	no lepton	mSUGRA, AMSB, split SUSY, heavy squark	QCD, $t\bar{t}$, W/Z
	single lepton (e,μ)	mSUGRA, AMSB, split SUSY, heavy squark	$t\bar{t}$, W
	dilepton (e,μ)	mSUGRA, AMSB, GMSB	$t\bar{t}$
	ditau	GMSB, large tan β	$t\bar{t}$, W
	$\gamma\gamma$	GMSB	—
~ 2	—	light squark	Z

Table 3.3: Summary of experimental signatures with E_T^{miss} and corresponding SUSY scenarios and SM background processes.

3.7.1 Z/W Boson Production Background

The processes of $Z \rightarrow \nu\overline{\nu}$ in association with multijets will give rise to final states with large E_T^{miss} and could be a dominant background for the no-lepton signature (figure 3.18). For this background contamination, the expectation is derived from the MC distribution of $Z \rightarrow \nu\overline{\nu}$ with the normalization determined by $Z \rightarrow l\bar{l}$ data, where l is e or μ.

[2]Since the gravity is not incorporated in the SM theory, this can lead to the conclusion that SM is an effective theory only valid below some energy cut off scale Λ. This cut off scale causes problems when one calculates quantum corrections to the mass of the Higgs boson. It makes also difficult to explain why the energy scale of the weak interactions, as reflected in the masses of the W and Z bosons, is so much smaller than that of unification or quantum gravity. This is what is called "hierarchy problem".

Identified leptons from Z decays can be replaced by neutrinos, and E_T^{miss} recalculated for $Z \to \nu\overline{\nu}$ process. Since the production mechanisms are the same in both processes, E_T^{miss} can be correctly estimated by $Z \to l\overline{l}$ events in data. We can better measure the $Z \to l\overline{l}$ yield thanks to small backgrounds in the final state.

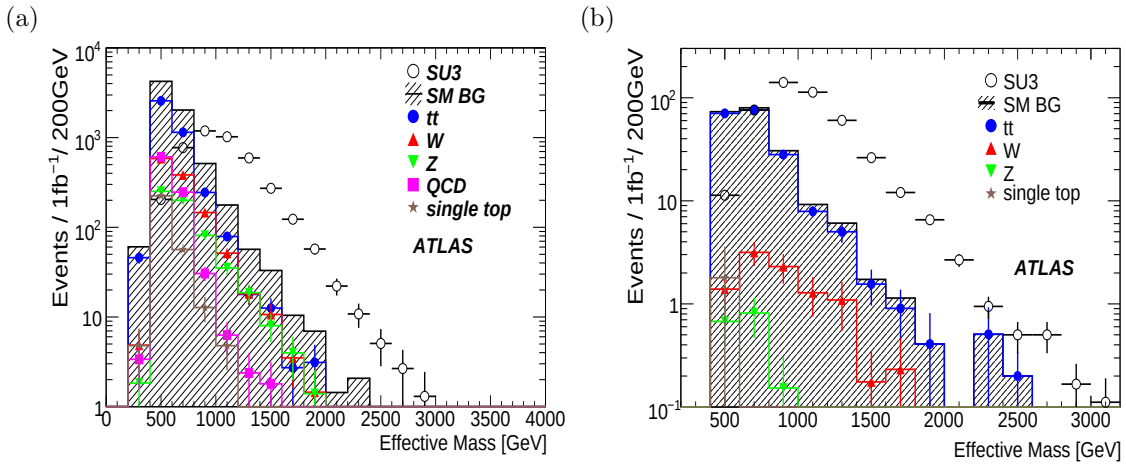


Figure 3.18: Effective mass distributions of the SUSY signal and background processes for the non-lepton mode (a) and one-lepton mode (b), for an integrated luminosity of 1 inverse femtobarn. The open circles show the SUSY signal (SU3 benchmark point). The shaded histogram (a) and black histogram (b) show the sum of all Standard Model backgrounds; different symbols show the various components.

Chapter 4

Z+jets Cross-Section Measurement: from Parton to Hadron Level

The purpose of the analysis presented in this thesis is to compare parton level predictions, i.e. theoretical calculations, to detector measurements, i.e."real" data, at hadron level (or particle jet level), as shown in figure 4.1. The whole analysis is performed mainly in two parts. The first part, which is detailed in this chapter, corresponds to the steps followed to measure the jet cross section in $Z(ll)$+jets events, from parton to hadron level. The second part, contained in the next chapter, describes the steps from detector to hadron level.

Therefore, in this chapter we describe the theoretical Next-to-Leading-Order (NLO) and Leading-Order (LO) predictions using the MCFM program [28] (see section 3.3.4). Jet non-perturbative corrections are derived from Monte Carlo events, and MCFM predictions, corrected for these non-perturbative effects, are compared to Monte Carlo generated events at hadron level (reference level in the analysis). The Monte Carlo data samples are described in tables 4.2 and 4.3. Finally, theoretical systematic uncertainties are discussed.

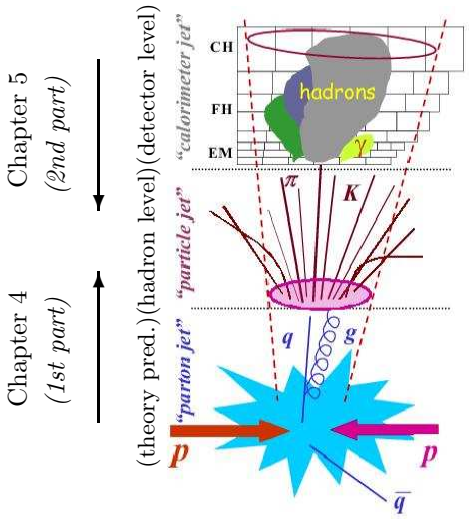


Figure 4.1: Scheme of jet production and reconstruction. Different levels are shown: parton jet level, particle jet or hadron level and calorimeter or detector jet level. In this chapter, steps followed to measure the jet cross section in $Z(ll)$+jets events, from parton to hadron level, are detailed.

4.1 Corrections from Parton to Hadron Level

The cross-section calculation and the differential distributions of the process Z+jets up to NLO order in α_s QCD predictions can be obtained with the parton-level Monte Carlo MCFM program. Nowadays, only some processes are available. Vector boson production in association with N jets for $N \leq 3$ is calculated at LO in references [70, 71]. The same process in association with N jets, up to $N \leq 2$ is calculated at NLO in reference [72]. There is ongoing work for NLO order calculations for $N \leq 3$ jets processes.

MCFM version 5.1, and processes $Z(l,l)$, $Z(l,l)$+1jet and $Z(l,l)$+2jets (up to NLO order), and $Z(l,l)$+3jets (only at LO order) are used in the analysis (more technical information can be found in appendix B). Renormalization, μ_R, and factorization, μ_F, scales are set to $M_Z^2 + p_{T_Z}^2$ for each event. MCFM is interfaced with the LHAPDF 5.2.3 [73], which provides a unified and easy way to interface to PDF sets [74]. CTEQ6.1m Parton Distribution Functions (PDFs) are used for the generation, which were the standard NLO best-fit in 2006, when MCFM v5.1 was released. The error on the inclusive cross section stemming from the PDF uncertainty is calculated using the complete set of error PDFs from the CTEQ6.1.m. MCFM predictions apply similar cuts on leptons and jets to those used in the particle identification of the analysis (see section 5.2). Only "crack region" cuts, which means a small interval in the pseudo rapidity plateau, are not used at parton generation due to technical difficulties. This should not bias the result.

Comparisons of the data cross section measurement and theoretical predictions will be made at hadron level which implies that MCFM predictions for the observables have to be corrected for non-perturbative effects resulting from jet fragmentation and from the underlying event. Both corrections are expected to decrease in relative terms as jet p_T increases.

The non-perturbative corrections are determined by comparing the multiplicity and the p_T distribution of jets with a cone of radius 0.4 clustered on the final-state particles in $Z \rightarrow \mu\mu$ Monte Carlo samples generated with PYTHIA 6.403 (a) using the standard ATLAS PYTHIA tune [45] (see section 3.5.2) and (b) with fragmentation and multiple-particle interactions switched off. To the extend to which the two partons that can comprise a jet in MCFM mimic the effects of the parton shower in PYTHIA [75], the corrections derived from the above procedure can be applied to the MCFM output. For jets with cone radius 0.4 (see section 5.1.3 for a description of jet algorithms in ATLAS) with $p_T > 40$ GeV, the effects of fragmentation and underlying event cancel up to a residual correction at the percent level, which is then applied to the MCFM predictions. These corrections, i.e. fragmentation and underlying event, will be validated with real data.

Figure 4.2 shows the correction to inclusive jet p_T distribution from parton to hadron level from the fragmentation (a) and from both fragmentation and UE (b) for cone 0.4 jets, which is the selected jet algorithm for the analysis. As expected, without fragmentation the out-of-cone energy loss for low-p_T jets is smaller, such that the number of low-p_T jets is reduced in the transfer from parton to hadron level. With underlying event the energy collected in the jet cluster increases such that the number of low-energy jets increases in the transfer from parton to hadron level. For $p_{T\,\mathrm{jet}} > 40$ GeV the non-perturbative effects are at the percent level. The corrections to inclusive Z +≥N jets cross-sections obtained for non-perturbative effects using cone 0.4 jet algorithm are detailed in table 4.1.

The study of the non-perturbative effects has also been performed on other type of jet clustering algorithms (more information about jet clustering can be found in section 5.1.3). Same process has been performed. The corrections for jet cluster cone algorithm with a radius of 0.7, and jet k_T algorithm with parameter value D=0.4 and 0.6 are shown in figure 4.3.

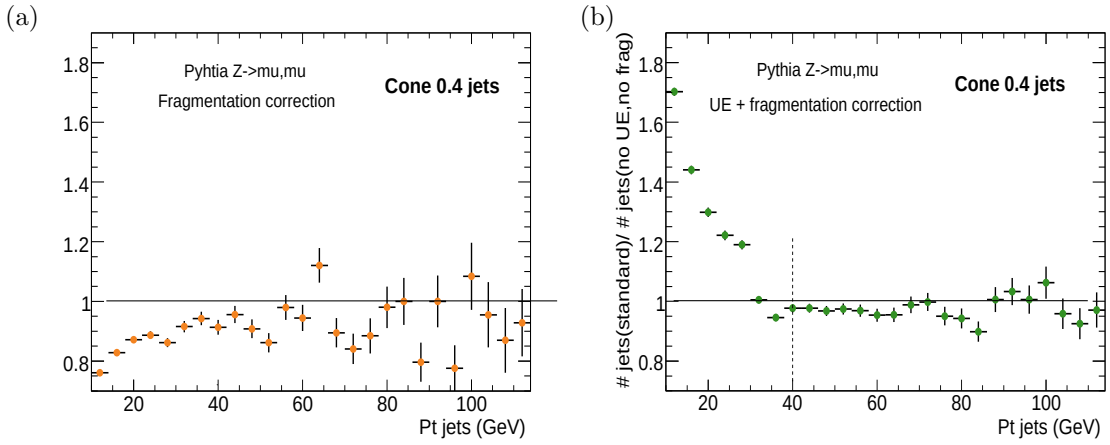


Figure 4.2: Ratio of Cone 0.4 Jet p_T distributions between standard PYTHIA and PYTHIA without fragmentation (a) and between standard PYTHIA and PYTHIA without non-perturbative corrections (b).

Process	Correction
$Z \rightarrow \mu\mu + \geq 1$jet	0.98 ± 0.01
$Z \rightarrow \mu\mu + \geq 2$jets	0.98 ± 0.01
$Z \rightarrow \mu\mu + \geq 3$jets	0.95 ± 0.06

Table 4.1: Corrections to $Z+ \geq$ N jets cross section, from parton to hadron level for non-perturbative effects (fragmentation and underlying event) using cone 0.4 jets.

The effect of fragmentation is to reduce the amount of energy inside the jet cone. Thus, from fragmentation effects alone, jets at the hadron level tend to have lower p_T than jets at the parton level. The effect of the underlying event is to add energy to the hadron level jet. In general, the underlying event tends to add more energy to the jet than the energy lost by fragmentation, but the exact ratio depends on the radius of the jet. Whereas the fragmentation corrections for Cone 0.7 jets are smaller than for Cone 0.4 jets, the underlying event corrections are larger due to the larger cone size. k_T 0.4 jets show the lowest combined corrections since fragmentation and underlying event effects basically cancel out. The performance of k_T 0.6 jets is comparable to the one of cone 0.4 jets. Except for cone 0.7 jets, the non-perturbative effects are negligible for jets with $p_T > 40$ GeV with the current PYTHIA tune.

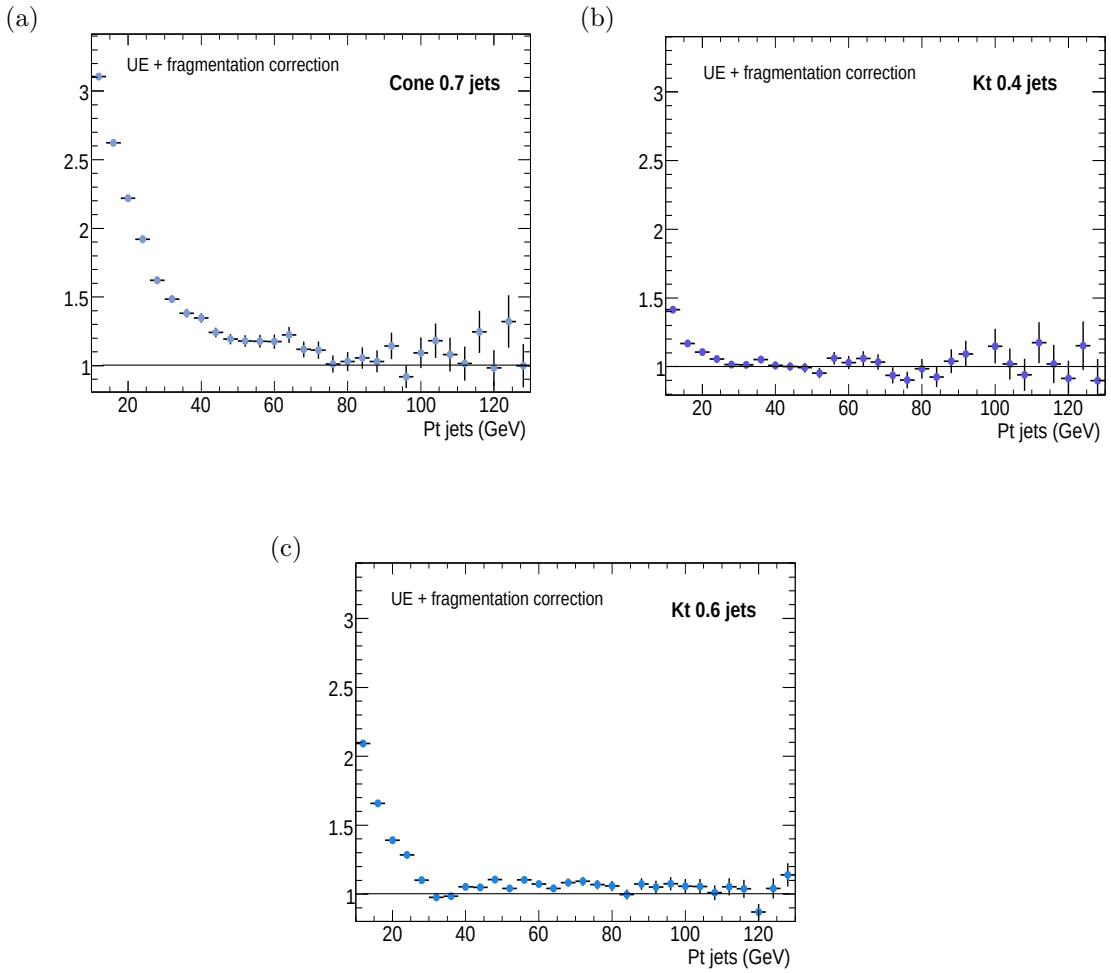


Figure 4.3: Ratio of the inclusive Jet p_T distributions between standard PYTHIA and PYTHIA without non-perturbative corrections for (a) Cone 0.7 jets, (b) k_T 0.4 jets and (c) k_T 0.6 jets.

4.2 Monte Carlo Data Sets

The Monte Carlo data sets for the signal processes $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ are generated with ALPGEN (see section 3.5), interfaced with HERWIG, using the leading order PDF set CTEQ6LL[1] [76] (Hereafter, when we refer to ALPGEN, it is understood that it is always interfaced with HERWIG). The generation is done with a renormalization and factorization scale of $m_Z^2+P_{T_Z}^2$ and a MLM (M.L.Mangano) matching technique [30, 77] (i.e. an algorithm which matches the parton shower and the matrix element calculations to avoid double counting (see section 3.4.3) with jet cuts at $p_T > 20$ GeV and $|\eta| < 6$ (*jet generator-level filter cuts*)). All but the highest jet multiplicity sample are exclusive, i.e. events are only kept if all jets with $p_T > 20$ GeV and $|\eta| < 6.0$ are matched to a matrix element parton. The highest multiplicity sample, Z+5 partons, is inclusive, events with additional jets softer than the partons from the matrix element are not discarded. Thus, there can be more than 5 jets in this sample. The di-lepton mass is required to be larger than 40 GeV and lower than 200 GeV. A generator-level filter requires one seeded-cone jet, with a radius of $R = 0.4$, with $p_T > 20$ GeV and $|\eta| < 5.0$, and two electrons/muons with $p_T > 10$ GeV and $|\eta| < 2.7$ in the event. The final Monte Carlo data sets are obtained following the standard prescription, which consists of merging the samples of Z+n partons (where n= 0 - 5), each sample weighted by the product of the respective sample cross section, the MLM matching efficiency

[1] CTEQ6LL (also known as CTEQ6L1) is a CTEQ6 PDF set with LO fit and with LO order in α_S

and the efficiency of the event-level generator filter. For the comparison with the fixed-order theoretical predictions, the merged data sets are normalized to the NLO inclusive $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ cross sections. Due to the jet filter used in their generation, the fully-simulated data sets can not be used to derive the global normalization factor. For this purpose, we use additional $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ ALPGEN data sets which are produced with the same conditions, but without the jet generator-level filter applied.

PYTHIA signal and background samples are generated with version 6.323 ($Z \rightarrow ee$, $Z \rightarrow \mu\mu$, $W \rightarrow e\nu$, $W \rightarrow \mu\nu$, $Z \rightarrow \tau\tau$) or version 6.403 ($t\bar{t}$, filtered QCD multi-jet, QCD ($b\bar{b}(\mu\mu)$) using the corresponding ATLAS underlying event tune (see section 3.5.2). PYTHIA $Z \rightarrow ee$ and $W \rightarrow e\nu$ events are preselected with a generator-level filter requiring one electron with $p_T > 10$ GeV and $|\eta| < 2.7$. The filter for the corresponding processes with muon final states, $Z \rightarrow \mu\mu$ and $W \rightarrow \mu\nu$, requires one muon with $p_T > 15$ GeV and $|\eta| < 2.8$. PYTHIA $Z \rightarrow \tau\tau$ events are generated with a filter requiring two electrons/muons with $p_T > 15$ GeV and $|\eta| < 2.8$. For each of the $Z \rightarrow \ell^+\ell^-$ samples, the di-electron/di-muon mass is required to be larger than 60 GeV.

The jet background for the electron channel is simulated with a PYTHIA QCD multi-jet sample with a minimum hard-scattering transverse momentum of 15 GeV. A generator-level filter requires a jet of $p_T \geq 17$ GeV clustered in a narrow region of $\Delta\eta \times \Delta\phi = 0.06$, a size similar to an electron cluster. The QCD multi-jet background for the muon channel is estimated with a PYTHIA $b\bar{b}(\mu\mu)$ sample (generated using an interface generator to PYTHIA, called PythiaB, specialized in B decays). Two muons with $p_T > 4$ GeV and $p_T > 6$ GeV, respectively, are required in the final state.

Top quark pair production has been simulated using the Monte Carlo generator MCatNLO, version 3.1. The hard process of $t\bar{t}$ production is calculated at NLO, so that diagrams that produce one additional parton in the final state are included at matrix element level. Fragmentation and hadronization is simulated using HERWIG. No cuts are applied at generation level other than the lepton flavor separation according to W boson decay type that allowed the subdivision of the generated events into fully-leptonic, non-fully hadronic and fully-hadronic. The top sample selected for the analysis consists of the non fully hadronic events. The top mass is set to $m_{top} = 175$ GeV. This data sample is used in $Z \rightarrow \mu\mu$ analysis part.

Tables 4.2 and 4.3 summarize the fully simulated signal and background data samples used in the analysis, for the $Z \rightarrow ee$ channel and for the $Z \rightarrow \mu\mu$ channel. These samples are all part of the ATLAS official MC data samples. The number of events used for each process and the generation and reconstruction version used are shown.

Two different detector geometries are available in the ATLAS event generation. The first one, an "ideal geometry", corresponds to the best knowledge of the material, with no misalignments (except a 4 cm shift of the electromagnetic calorimeter endcaps, already known). This geometry has been used to derive misalignment corrections and some performance studies. The second geometry is a distorted one, misalignments and extra material between Inner Detector and Calorimeter are introduced. These "misaligned" data have been used for this study.

Sample	Process	events	$\int L\,dt(\mathrm{pb}^{-1})$	generation	reconstruction
ALPGEN+HERWIG					
8130	$Z \rightarrow ee$ + 0partons	80000	537	12.0.5	12.0.6.4
8131	$Z \rightarrow ee$ + 1partons	98800	716	12.0.5	12.0.6.4
8132	$Z \rightarrow ee$ + 2partons	96200	1905	12.0.5	12.0.6.4
8133	$Z \rightarrow ee$ + 3partons	49000	3025	12.0.5	12.0.6.4
8134	$Z \rightarrow ee$ + 4partons	46250	10120	12.0.5	12.0.6.4
8135	$Z \rightarrow ee$ + 5partons	16950	10030	12.0.5	12.0.6.4
PYTHIA					
5144	$Z \rightarrow ee$	172900	121	11.0.4	12.0.6.4
5146	$Z \rightarrow \tau\tau$	170000	2208	11.0.5	12.0.6.5
5568	$t\bar{t}$	50000	100	12.0.7	12.0.6.1
5104	$W \rightarrow e\nu$	179900	16.5	11.0.4	12.0.6.1
5802	Dijets (JF17)	3012000	0.004	12.0.3	12.0.6.4

Table 4.2: Fully simulated signal and background samples used in $Z \rightarrow e^+e^-$ analysis.

Sample	Process	events	$\int L\,dt(\mathrm{pb}^{-1})$	generation	reconstruction
ALPGEN+HERWIG					
8142	$Z \rightarrow \mu\mu$ + 0partons	26000	525	12.0.5	12.0.6.4
8143	$Z \rightarrow \mu\mu$ + 1partons	50000	360	12.0.5	12.0.6.4
8144	$Z \rightarrow \mu\mu$ + 2partons	21000	400	12.0.5	12.0.6.4
8145	$Z \rightarrow \mu\mu$ + 3partons	18000	1100	12.0.5	12.0.6.4
8146	$Z \rightarrow \mu\mu$ + 4partons	7000	1500	12.0.5	12.0.6.4
8147	$Z \rightarrow \mu\mu$ + 5partons	5000	3000	12.0.5	12.0.6.4
6108	$W \rightarrow \mu\nu$ + 1partons	45000	28	12.0.5	12.0.6.5
6109	$W \rightarrow \mu\nu$ + 2partons	35000	60	12.0.5	12.0.6.5
6110	$W \rightarrow \mu\nu$ + 3partons	46000	270	12.0.5	12.0.6.5
6111	$W \rightarrow \mu\nu$ + 4partons	32000	630	12.0.5	12.0.6.5
6112	$W \rightarrow \mu\nu$ + 5partons	10000	500	12.0.5	12.0.6.5
8154	$Z \rightarrow \tau\tau$ + 0partons	10000	200	12.0.5	12.0.6.1
8155	$Z \rightarrow \tau\tau$ + 1partons	9500	170	12.0.5	12.0.6.1
8156	$Z \rightarrow \tau\tau$ + 2partons	9500	450	12.0.5	12.0.6.1
8157	$Z \rightarrow \tau\tau$ + 3partons	5000	700	12.0.5	12.0.6.1
8158	$Z \rightarrow \tau\tau$ + 4partons	5000	2500	12.0.5	12.0.6.1
8159	$Z \rightarrow \tau\tau$ + 5partons	4000	1800	12.0.5	12.0.6.1
PYTHIA					
5145	$Z \rightarrow \mu\mu$	160000	100	11.0.4	12.0.6.1
5105	$W \rightarrow \mu\nu$	45000	4	11.0.4	12.0.6.1
17506	QCD($b\bar{b}(\mu\mu)$)	123000	1.1	12.0.3	12.0.6.2
5568	$t\bar{t}$	75000	170	12.0.3	12.0.7.1
5146	$Z \rightarrow \tau\tau$	38000	490	11.0.5	12.0.6.1
MCatNLO					
5200	$t\bar{t}$	150000	320	12.0.4	12.0.6.4

Table 4.3: Fully simulated signal and background samples used in $Z \rightarrow \mu^+\mu^-$ analysis.

4.3 Comparison of Event Generators and MCFM at the Hadron Level

We consider the comparison of theory and measurement for theoretically well-defined quantities: the inclusive cross section for $Z(ll)+ \geq$1jet, $Z(ll)+\geq$2jets, $Z(ll)+\geq$3jets and $Z(ll)+\geq$4jets and the differential cross sections with respect to the p_T of the leading jet and the p_T of the second leading jet. The partonic MCFM predictions are corrected to hadron level taking into account the residual energy loss due to non-perturbative effects for jets with $p_T > 40$ GeV, by applying the factors determined previously (see table 4.1). In order to study separately reconstruction and generation effects we consider here only Monte Carlo generator information (*truth-level information*), while reconstruction effects are studied in chapter 5.

The two Monte Carlo samples used (PYTHIA and ALPGEN) are normalized to the inclusive NLO Z cross-section (0-jet bin), as determined in MCFM. The normalization factors used are shown in table 4.4. Figures 4.4(a)-(c) show the comparison of the distribution of the jet multiplicities and the p_T of the leading and second leading jets for ALPGEN and PYTHIA $Z \rightarrow \mu^+\mu^- +$ jets samples with the NLO (LO) calculations from MCFM. The corresponding numerical results are given in table 4.5 (for jet multiplicity distribution) and in appendix A (for the p_T of the leading and second leading jet distributions). The errors on the generator distributions are purely Monte Carlo statistics whereas the errors on the MCFM cross section correspond to the statistical uncertainty and to the statistical error from the folding of MCFM parton level to the hadron level (errors shown in table 4.1). PDF uncertainties (see section 4.4) will be included later.

Monte Carlo sample	Normalization factor
ALPGEN $Z \rightarrow ee$	1.33 ± 0.06
PYTHIA $Z \rightarrow ee$	1.25 ± 0.06
VBF loose filter cut $Z \rightarrow ee$	3.05 ± 0.01
ALPGEN $Z \rightarrow \mu\mu$	1.36 ± 0.01
PYTHIA $Z \rightarrow \mu\mu$	1.07 ± 0.01
VBF loose filter cut $Z \rightarrow \mu\mu$	3.05 ± 0.01

Table 4.4: Factors used to normalize Monte Carlo data to NLO MCFM predictions, at the inclusive Z cross-section, for $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ processes. The factor called "VBF loose filter cut" is the factor due to the requirement of at least 1 jet in the ALPGEN data sample. We obtain it by comparing our data sample to an inclusive $Z \rightarrow \ell\ell$ one, where no jet requirement is asked.

The NLO MCFM predictions (dark triangles) are larger than the LO predictions (open triangles), as shown in figure 4.4(a), by 20 to 30% for the $Z + \geq 1$ jet , $Z + \geq 2$ jets cross sections (i.e. where NLO predictions are available). PYTHIA (upside down red triangles) predicts only a larger $Z + \geq$1jet cross section than ALPGEN (blue squares), whereas it predicts a lower cross sections for higher jet multiplicities ($Z + \geq 2$, 3, 4 jets) because of the missing higher order contributions at the hard-scatter level. It is important to note that all selected jets are required to have $p_T > 40$ GeV. Both Monte Carlo generators predict a lower cross section than the NLO MCFM calculation for final states with more than one jet ($Z + \geq 2$ jets), and also even less than calculated by MCFM at LO with this particular choice of PDFs and renormalization/factorization scale, for $Z + \geq 2$, ≥ 3 jets. The difference between the predictions of PYTHIA and ALPGEN, and between both generators and MCFM, amounts to 10-60% depending on the jet multiplicity.

Process	$Z(ll)$	$Z(ll)+\geq$ 1jet	$Z(ll)+\geq$ 2jets	$Z(ll)+\geq$ 3jets	$Z(ll)+\geq$ 4jets
MCFM NLO	781940.7± 365.1	109307.2±107.3	24828.57±107.9	– ± –	– ± –
MCFM LO	700968.7± 278.1	77208.1± 31.2	21429.6± 32.6	5109.9± 17.9	– ± –
ALPGEN	781956.0±1384.9	93114.6±803.5	19078.4±269.8	3694.9± 94.1	691.2± 33.7
PYTHIA	781941.3±2660.8	104464.1±972.5	15047.6±369.1	2363.1± 146.3	470.8± 65.3

Table 4.5: Cross-sections (in fb) and the corresponding statistical errors for $Z \rightarrow \mu\mu$ processes comparing MCFM NLO and LO predictions, at hadron level, with ALPGEN and PYTHIA truth-level data normalized to inclusive NLO MCFM cross-section. These results are shown in figure 4.4(a).

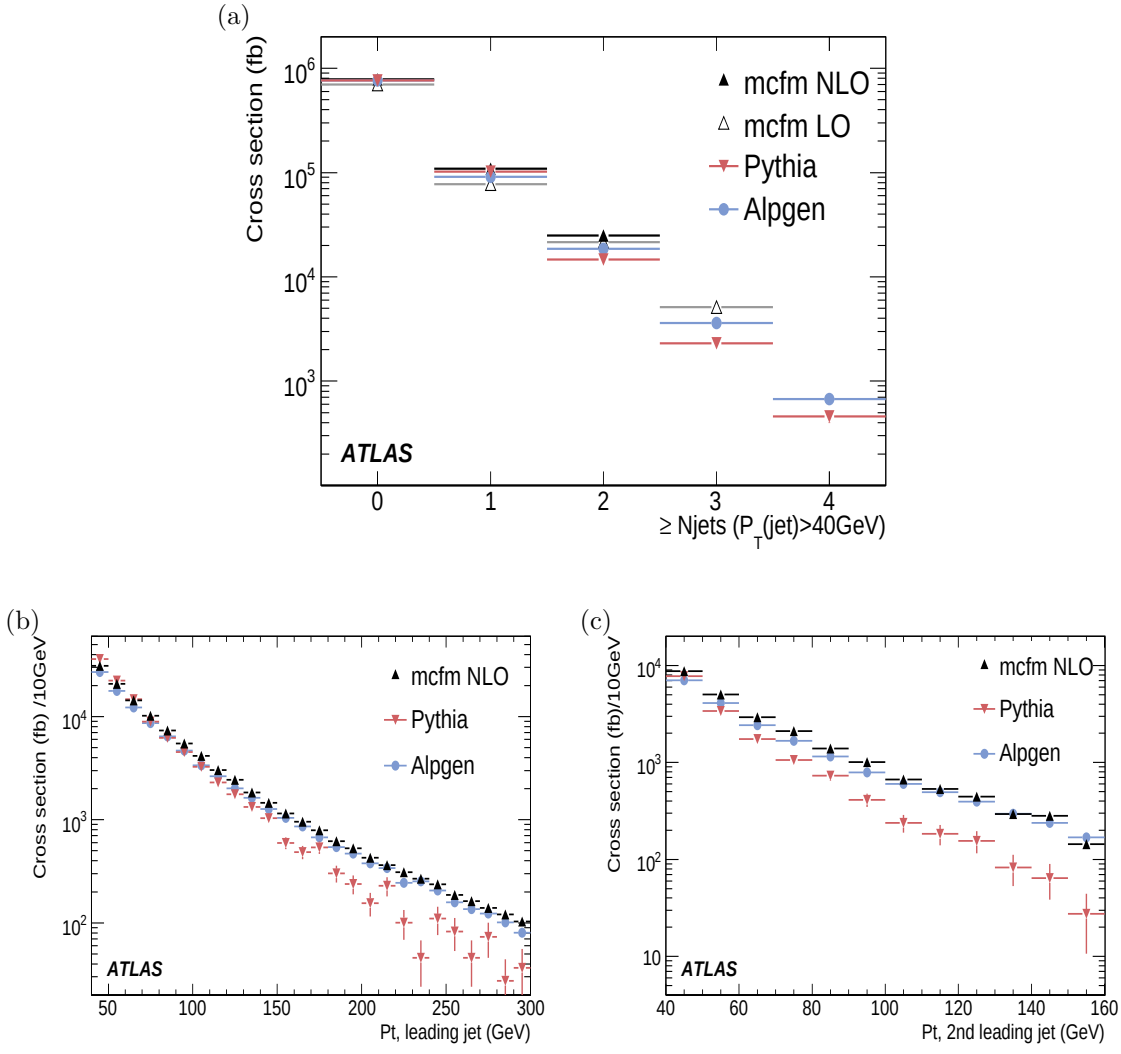


Figure 4.4: Comparison of the inclusive jet cross-section (a), the p_T distribution of the leading jet (b), and the p_T distribution of the second leading jet (c) for the $Z \rightarrow \mu^+\mu^-$+jets process from PYTHIA and ALPGEN Monte Carlo with NLO (LO) MCFM predictions for 1fb^{-1} of data. All results are normalized to inclusive Z cross-section predicted by MCFM at NLO. Only statistical errors are shown (corresponding to the generated sample statistics) for 1fb^{-1}.

A comparison of the differential cross sections as a function of the jet p_T (figure 4.4(b),(c)), indicates that the inclusive cross sections shown in figure 4.4(a) are very sensitive to the minimum jet p_T required by the selection. PYTHIA predicts larger cross sections than even NLO MCFM for low jet p_T ($p_T < 50 - 70$ GeV). But, while the shape of the jet p_T distribution predicted by ALPGEN agrees very well with the NLO MCFM prediction, PYTHIA generates a clearly softer p_T spectrum. So, we might be able to discriminate between ALPGEN and PYTHIA predictions for the leading jet spectrum with the first fb^{-1} of data.

4.4 Systematic Uncertainties

The uncertainties due to the PDFs on the MCFM predictions are calculated using the complete set of CTEQ6.1m PDF errors. The positive (+) and negative (-) deviations along the 20 eigenvectors of CTEQ6.1m PDFs are taken from MCFM results. Then, asymmetric uncertainties are obtained by summing in quadrature the maximal deviation in each direction associated to each of the 20 eigenvectors, as follows:

$$\delta\sigma^+ = \sqrt{\sum_{i=1}^{20} \max\{ \left(\sigma(a_i^+) - \sigma(a_0)\right), \left(\sigma(a_i^-) - \sigma(a_0)\right) \}^2} \tag{4.1}$$

$$\delta\sigma^- = \sqrt{\sum_{i=1}^{20} \min\{ \left(\sigma(a_i^-) - \sigma(a_0)\right), \left(\sigma(a_i^+) - \sigma(a_0)\right) \}^2} \tag{4.2}$$

where $\sigma(a)$ is the prediction of the cross section determined using the PDFs with the parameters in vector a. The results obtained are shown in table 4.6.

Process	$Z(ll)$	$Z(ll)+ \geq$ 1jet	$Z(ll)+ \geq$ 2jets	$Z(ll)+ \geq$ 3jets (LO)
σ (fb)	781940.7	109307.2	24828.6	5109.8
stat. error	$\pm$365.1	$\pm$107.3	$\pm$ 107.9	$\pm$ 17.6
$\delta\sigma^+$	-	5647.4	1366.4	255.3
$\delta\sigma^-$	-	2216.5	579.5	149.0
pdf uncertainty	-	$\pm$ 5647.4 (5.2%)	$\pm$ 1366.4 (5.5%)	$\pm$ 255.3 (5.0%)

Table 4.6: NLO (and LO for $\geq$ 3 jets) MCFM cross-sections predictions (in fb) and their associated PDF uncertainties. Last row is the largest uncertainty between the two previous ones, which is used in the final results. Statistical errors are also shown and are much smaller in all processes.

The total PDF uncertainty on the inclusive cross section varies from 2 to 5% for $Z+\geq$ 1 jet, from 2 to 6% for $Z+\geq$ 2 jets, and from 3 to 5% for $Z+\geq$ 3 jets. At low jet p_T PDF uncertainty is $\leq$ 5% and at high jet p_T decreases to $\leq$ 3% (numerical results are summarized in section A.1).

Another source of systematic uncertainty on the theoretical predictions is the dependence on the choice of renormalization/factorization scale (μ_R, μ_F). MCFM predictions could also be obtained using the MRST PDFs set, and compared with the theoretical predictions obtained here. Nevertheless, these are out of reach of this study, since it requires too much computing time. Other experiments have studied this source of systematic uncertainty. In CDF studies, changing the renormalization/factorization scales $\mu_0 = \sqrt{M_Z^2 + p_{T,Z}^2}$ by a factor of 2 from its nominal value, $\mu_0 \times 2$ or $\mu_0/2$ results in a prediction of the cross section varying between 10% and 15%. Similar variations have been found by the D0 experiment [78], when changing the scale by a factor of 2. Therefore, this source of systematic uncertainty must be considered in future comparison of Z+jets cross-section measurements with the theoretical predictions, since it can be larger than the uncertainty coming from the PDFs.

Chapter 5

Z+jets Cross-Section Measurement: from Detector to Hadron Level

This chapter is dedicated to explain in detail the steps followed in the second step of the measurement of the Z(ll)+jets cross section, going from detector to hadron level. Two separate analyses have been performed, one for the electron channel, $Z \rightarrow e^+e^-$, and another one for the muon channel, $Z \rightarrow \mu^+\mu^-$. These two analyses have been performed in parallel, using similar, but not identical, tools. This has provided us an efficient and fast way to cross check results on both analyses. This chapter describes the reconstruction of the physical objects used in the analyses, the event selection performed, the efficiencies of the related reconstructed objects, the estimation of the background contributions, the detector effects contributions to jet reconstruction (i.e, unfolding procedure from detector to hadron level), and finally, a discussion on the related systematic uncertainties.

5.1 Electron, Muon and Jet Reconstruction in ATLAS

5.1.1 Electron Reconstruction

The "sliding window" algorithm [79] is used to find and reconstruct electromagnetic clusters in the calorimeter. Clusters are positioned in order to maximize the corresponding amount of energy. These clusters have a fixed size depending on the reconstructed object and on the region of the calorimeter. For electrons, a larger size compared with photons is needed due to their larger probability of interaction in the upstream material and also due to their bending inside the magnetic field. For instance, the energy in the barrel electromagnetic calorimeter for electrons is collected over an area corresponding to 3×7 cells in the middle layer, equivalent to $\Delta\eta \times \Delta\phi = 0.075 \times 0.175$.

For each reconstructed cluster, a track within a $\Delta\eta \times \Delta\phi$ window of 0.05×0.10 with a momentum p compatible with the cluster energy E ($E/p < 10$) is searched for. If one is found, the reconstruction algorithm checks for the presence of an associated photon conversion. Then, if a matched track is found and a "no conversion" is flagged, an electron candidate is created. Otherwise, the candidate is classified as a photon.

The next step is the calibration of the electrons candidates. First of all, the raw signal extracted from each cell (in ADC counts) is converted to deposited energy (*electronics calibration*). Then, the energies deposited in the different cells of each layer of a cluster are summed, and an energy-weighted cluster position is calculated for each layer, taking into account different important effects:

- ϕ modulation of the reconstructed energy due to the accordion geometry (different amount of absorber material in ϕ).

- No full shower containment in the η window and finite granularity of the cells. Modulation on energy and bias in the measured position depending on the particle impact point within a cell.

- Not perfectly projective particles because a particle from a vertex away from the origin intersects the calorimeter at slightly different η positions in each layer.

These corrections have been calculated using both simulation and test beam data. Nevertheless, some local or "medium range" inhomogeneities in the calorimeter response can be found. It is planned to use $Z \rightarrow e^+e^-$ data to correct for these inhomogeneities and also to contribute to the knowledge of the absolute energy scale. More information about these corrections can be found in [80].

Standard identification of high-p_T electrons is based on many cuts that can be applied independently. Three reference sets of cuts have been defined to optimize different physics analyses: loose, medium and tight,

- **Loose cuts:** Simple electron identification based only on limited calorimeter information (cuts are applied on the hadronic leakage and on shower-shape variables from the middle layer of the EM calorimeter only). It provides an excellent identification efficiency ($\sim$ 88%), but a low jet background rejection ($\sim$570).

- **Medium cuts:** Improvement by adding cuts on the strips in the first layer of the EM calorimeter (with a finer granularity of $\Delta\eta \times \Delta\phi = 0.0031 \times 0.0245 \times 4$, as shown in figure 2.8(a) than the middle layer, $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$) and on the tracking variables. Strip-based cuts are effective in the rejection of $\pi^0 \rightarrow \gamma\gamma$ decays. The energy deposit pattern from π^0's is often found to have two maxima, which can be used to distinguish from the energy deposit pattern of an electron. Some of the strip-based variables are: difference between the energy associated with the second largest energy deposit and energy associated with the minimal value, $\Delta E_s = E_{max2} - E_{min}$, second largest energy deposit, R_{max2}, total shower width, ω_{stot}, shower width for three strips around the maximum strip, w_{s3}, and fraction of energy outside core of three central strips, within seven strips, F_{side}. Track quality cuts are also added. These variables include the number of hits in the pixel detector (at least one), the number of hits in the pixel and the SCT (at least nine) and the transverse impact parameter ($<$ 1mm) (see section 2.2.3 for Inner Detector information). The medium cuts increase the jet rejection by a factor of 3-4 with respect to loose cuts set ($\sim$ 2200), while reducing the efficiency by $\sim$ 10% ($\sim$ 77%).

- **Tight cuts:** Use of all the particle-identification tools available for electrons. In addition to previous cuts, cuts are applied on the number of vertexing-layer hits (rejection of electrons from conversions), on the number of hits in the Transition Radiation Tracker (TRT) (section 2.2.3), on the ratio of high-threshold hits to the number of hits in the TRT (rejection of the dominant background from charged hadrons, which radiate much less than electrons), on the difference between the cluster and the extrapolated track positions in η and ϕ, and on the ratio of cluster energy to track momentum (large energy depositions by electrons due to the transition radiation[1]). It increases the jet rejection up to $\sim 10^5$ and decreases the efficiency down to $\sim$62%.

[1] Transition radiation is produced by relativistic charged particles, when crossing the interface of two media of different dielectric constant. The energy loss of the charged particle depends on its Lorentz factor, γ, which depends on the mass of the particle. So light charged particles, like electrons, will emit much more radiation than charged hadrons, making this feature suitable for particle discrimination.

5.1.2 Muon Reconstruction

ATLAS uses a variety of strategies for identifying and reconstructing muons. *Standalone* muons are reconstructed by finding tracks in the muon spectrometer ($|\eta|_{max} \leq 2.7$), and extrapolating them to the beam line. *Combined* muons are found by matching standalone muons with nearby tracks from the Inner Detector, and then combining the measurements from both systems. *Tagged* muons are found by extrapolating inner detector tracks to the spectrometer and searching for nearby hits. Calorimeter tagging algorithms are also being investigated, to tag inner detector tracks using the presence of minimum ionizing signal in calorimeter cells.

The tracking system of ATLAS, the Inner Detector, can detect muons and other charged particles with a hermetic coverage for $|\eta| < 2.5$, providing important confirmation of muons found by the spectrometer over that η range. The 2 Tesla solenoidal magnet enables the inner detector to provide an independent measurement of the momentum of the muon (and other charged particles). Over most of the acceptance for p_T in the range between 30 and 200 GeV, the momentum measurements may be combined from inner detector and muon spectrometer to give a precision measurement better than either alone. The inner detector dominates below this range, while the muon spectrometer above it.

The output data to be used in physics analyses have two different muon collections for each processed event. Each muon collection corresponds to a given combined algorithm used: Staco and Muid. Currently, Staco muon collections are used in physics analysis. Staco and Muid algorithms are combined muon-spectrometer tracks with inner detector tracks to identify muons. Staco (STAtistical COmbination) is an algorithm which statistically merges the two independent track measurements derived from the inner detector track and from the spectrometer track. The merging is done by combining the 5×5 covariance matrices of the two tracks to obtain the benefit of two independent measurements. The combination uses the following: for two tracks at a reference location defined by their parameters vectors, P_1 and P_2, and their covariance matrices, C_1 and C_2, the parameter vector of the combined track, P, is the solution of the equation:

$$(C_1^{-1} + C_2^{-1}) \times P = C_1^{-1} \times P_1^{-1} + C_2^{-1} \times P_2$$

The combined track covariance matrix, C, is given by:

$$C = (C_1^{-1} + C_2^{-1})^{-1}$$

where the associated chi-squared is:

$$\chi^2 = (P - P_1)^T \times C_1^{-1} \times (P - P_1) + (P - P_2)^T \times C_2^{-1} \times (P - P_2)$$

Muid algorithm, unlike Staco, performs a global refit of all hits associated with the previous tracks, taking into account the material (multiple scattering and energy loss) and magnetic field in the calorimeter and muon spectrometer. In Muid, tracks are first matched by forming a chi-square with five degrees of freedom from the difference between the five track parameters and their summed covariance from the Inner Detector and standalone fits (This is the chi-square described for Staco above). Combined fits are then performed on matches with chi-square probability $\chi^2 > 0.001$. When no match satisfies this condition, a combined fit is attempted for the best match within a road about the standalone track. All measurements from the Inner Detector and the Muon Spectrometer systems are then incorporated into the global combined fit.

5.1.3 Jet Reconstruction

The ATLAS calorimeter system has around 270,000 cells. To find jets it is necessary to combine these cells into larger signal objects with physically meaningful four-momenta. Two different approaches are available: *calorimeter signal towers* and *calorimeter topological cell clusters*.

Calorimeter towers are built projecting the cells in a fixed grid in pseudorapidity and azimuth angle. The tower bin size is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the whole acceptance region of the calorimeter ($|\eta| < 5.0$). Projective calorimeter cells which completely fit inside the tower contribute their total energy to the tower signal, while non-projective cells and projective cells larger than the tower size contribute a fraction of their signal to different towers, depending on the fraction of the cell area within the towers. Then, tower signals are a sum of weighted cell signals, on the basic electromagnetic energy scale[2]. Finally, jet reconstruction from calorimeter towers starts with a re-summation step. Some towers may have negative tower signal, due to signal fluctuations from noise (electronics and physics from pile-up), giving unphysical four-momentum. Combining negative signal towers with nearby positive signal towers, or eliminating those without nearby positive signal towers, leads to cancelation of some of the noise fluctuations. The resulting "protojets" represent one or a few towers and have physically valid four-momenta. They are the input of the jet finder algorithm.

Topological cell clusters are based on the idea of reconstruction on three dimensions of energy "blobs" representing the shower development of each particle entering the calorimeter. The clustering starts with seed cells with a signal-to-noise ratio, or signal significance above a certain threshold, S, $\Gamma = E_{cell}/\sigma_{noise,cell} > S = 4$. Direct neighbor cells, in three dimensions, to the seed cells are collected into the cluster. Then, neighbors of neighbors cells are considered for those added cells with a signal significance over a secondary threshold, N, $\Gamma > N = 2$. Finally, a ring of guard cells with signal significance above a basic threshold, P, $\Gamma > P = 0$, is added to the cluster. After the initial clusters are built, a splitting algorithm looks for local signal maximums, and if they are found, clusters are split. As towers, clusters are initially built using the electromagnetic energy scale cell signals. In addition, clusters can be calibrated to a local hadronic energy scale [81]. Here, the treatment of negative signals is different from calorimeter towers. The noise suppression applied by the cell clustering reduces already any noise contribution. A schematic diagram of the different reconstruction sequences for jets is shown in figure 5.1.

The most commonly used jet finders in ATLAS are the seeded fixed-cone finder with split and merge, and the k_T algorithm. Other jet finder implementations will be available in the future: FASTJET package [82] (K_T, anti-K_T, Aachen/Cambridge flavor K_T [83]), as well as the seedless infrared-safe cone algorithm SISCONE [84].

The implementation of the iterative seeded fixed-cone jet finder is described in [50]. Firstly, all input "protojets" are ordered in decreasing transverse momentum p_T. If the object with higher p_T is above the threshold (typically $p_T > 1$ GeV) , all the objects within a cone with $R_{cone} < \Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, where R_{cone} is the value of the fixed radius of the cone, are combined. A new direction is then calculated from the four-momenta inside this cone, and a new cone is centered around this new direction. Objects inside this new cone are collected, and again the direction is updated. This process is repeated until the final cone is stable, which means that its direction does not change in the recombination. This final stable cone is called jet. Afterwards, the next seed is taken from the initial list, and the process is repeated until a new final stable cone is formed, finding a new jet. This continues until no more seeds are available. Jets formed using this process can share constituents and signal objects contributing to a

[2]This is the raw signal for ATLAS calorimeters. The nomenclature indicates that this scale has been derived from electron signals.

cone at some iteration can be lost at a later iteration.

This algorithm is not infrared safe, but can be partly recovered by introducing a split and merge step after the jet formation (no overlapping cones are left), but still extra soft particles can lead to new stable cones between previous built cones [85]. Jets which share constituents with more than a certain fraction of the p_T of the less energetic jet are merged, while they are split if this fraction is below a certain value (in ATLAS >0.5).

A narrow ($R_{cone} = 0.4$) and a wide ($R_{cone} = 0.7$) cone jet finder algorithms are run in ATLAS.

The default sequential recombination algorithm in ATLAS is the k_T algorithm. This jet finder algorithm introduces the distance d_{ij} between pairs ij of initial objects (partons, particles, reconstructed detectors objects) and the d_{iB} between the object i and the beam.

These distances are defined as:

$$d_{ij} = min(p_{T,i}^2, p_{T,j}^2)\frac{\Delta R_{ij}^2}{R^2}\,, \quad where\ \Delta R_{ij}^2 = \Delta\eta_{ij}^2 + \Delta\phi_{ij}^2$$

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

The minimum d_{min} of all d_{ij} and d_{iB} is found. If d_{min} is a d_{ij}, the corresponding pair of objects i and j are combined into a new object k using four-momentum recombination. Then, both objects i and j are removed from the initial list, and the new object k is added to it.

If d_{min} is a d_{iB}, the object i is considered to be a jet and it is removed from the list. This procedure is repeated with the new objects in the list (and the corresponding new d_{ij} and d_{iB}) until no objects are left in the list. All initial objects in the list end up to be either a jet by themselves or to be part of a jet. Unlike cone algorithms no objects are shared between jets, every particle in an event is assigned to a unique jet. k_T jet algorithm is, by design, infrared and collinear safe to all orders of calculation. The distance parameter D (the only free parameter in the implementation of the k_T algorithm) allows some control on the jet size. The default configurations in ATLAS are $D = 0.4$ for narrow and $D = 0.6$ for wide jets.

Historically, the calibration scheme for calorimeter jets in ATLAS is based on cell signal weighting and can be used in both towers and topological clusters. The basic idea, which was developed in the CDHS experiment and refined in the H1 experiment, is that low signal densities in calorimeter cells indicate a hadronic signal in a non-compensating calorimeter, as ATLAS, and thus they need a signal weight for compensation of the order of the e/π signal ratio, while high signal densities are more likely generated by electromagnetic showers, and therefore do not need extra signal weighting.

In this calibration approach, all calorimeter signal cells contributing to a jet are retrieved. The signal in each cell i is weighted by a function depending on the cell position $\vec{X}_i$ and the cell signal density $\rho_i = E_i/V_i$, where E_i is the electromagnetic energy signal of the cell and V_i is the volume of the cell. The weighting functions, $w(\rho_i, \vec{X}_i)$, are universal, they do not depend on any jet feature or variable. Then, the calibrated jet has a four-momentum $(E_{jet,calo}, \vec{p}_{reco})$, recalculated from the weighted cell signals:

$$\left(E_{jet,calo},\, \vec{p}_{reco}\right) = \left(\sum_i^{N_{cells}} w(\rho_i, \vec{X}_i)E_i\,,\ \sum_i^{N_{cells}} w(\rho_i, \vec{X}_i)\vec{p}_i\right)$$

These signal weighting functions have been determined using seeded fixed size cone jets, $R_{cone} = 0.7$, using fully simulated QCD dijet events with the ideal detector geometry. Additional calibration functions parametrized in p_T and η are used to correct for residual non-linearities. These corrections have also been calculated for other standard jet finding configurations.

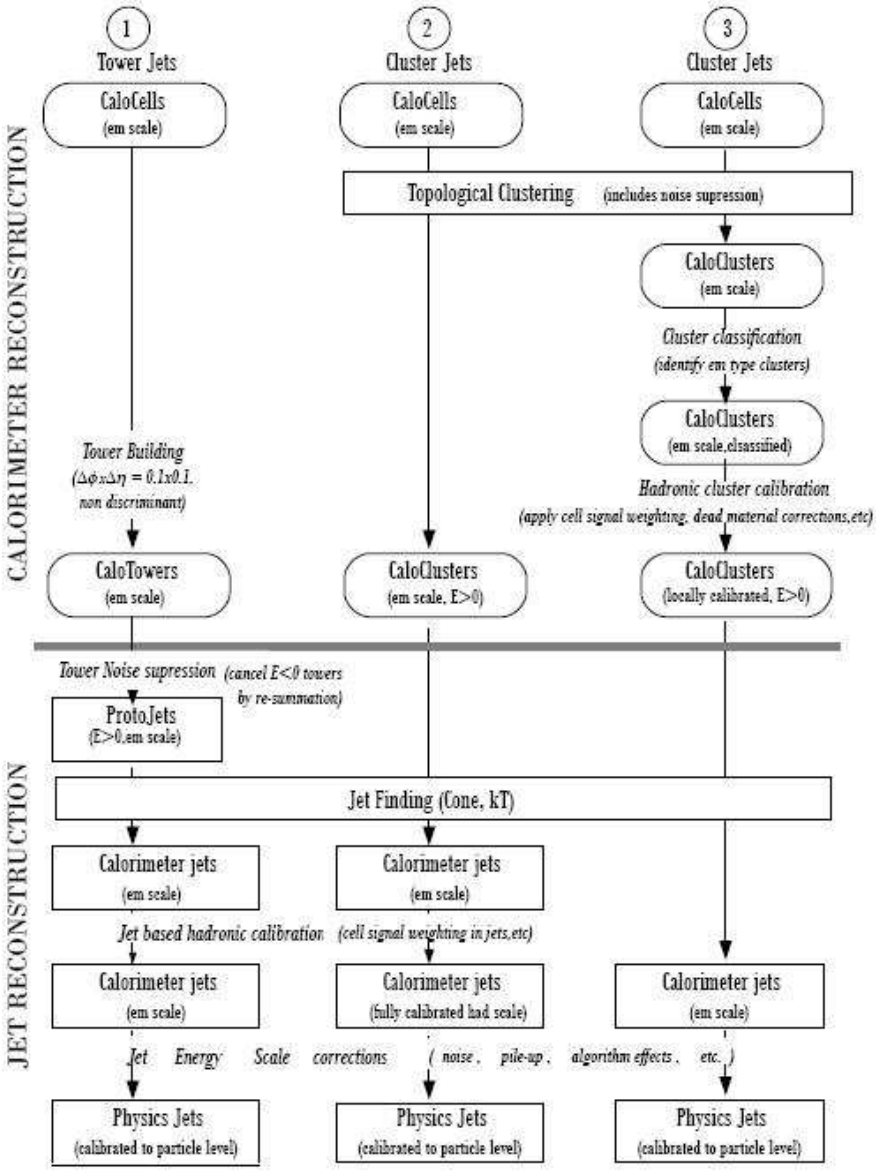


Figure 5.1: Schematic view of the reconstruction sequences for jets from calorimeter towers (left), uncalibrated (center) and calibrated (right) topological calorimeter cell clusters in ATLAS.

5.2 Particle Identification and Trigger

5.2.1 Particle ID

Electron ID

An electron candidate is required to have $p_T > 25$ GeV and to lie in the range $|\eta| < 2.4$ (central region of the EM calorimeter, corresponding to the precision measurement range) , excluding the barrel-to-endcap calorimeter crack region ($1.37 < |\eta| < 1.52$). The electrons are required to fulfill the *medium electron* identification,which consists of requirements on the calorimeter shower-shape and the matched track. No calorimeter isolation cuts are applied, only the implicit isolation cut present in the trigger is used (see section 5.2.2). For the Z selection, a distance $\Delta R > 0.2$ between the two electrons candidates is required. Electron reconstruction efficiency[3] as a function of $|\eta|$ is shown in figure 5.2(a), where a drop in efficiency at $|\eta| > 1.5$ is shown corresponding to the barrel region.

Muon ID

A muon candidate requires the combined reconstruction of an inner detector track and a muon spectrometer track, using the Staco algorithm. Muons are required to have $p_T > 15$ GeV and $|\eta| < 2.4$ (coverage for detecting muons in the Inner Detector), excluding the range $1.2 < |\eta| < 1.3$ (barrel/end-cap transition region) and the central point $|\eta| < 0.1$. Isolation is applied by requiring the energy deposition in the calorimeter to be less than 15 GeV in a cone of $\Delta R = 0.2$ around the extrapolation of the muon track. Muon reconstruction efficiency as a function of η is shown in figure 5.2(b).

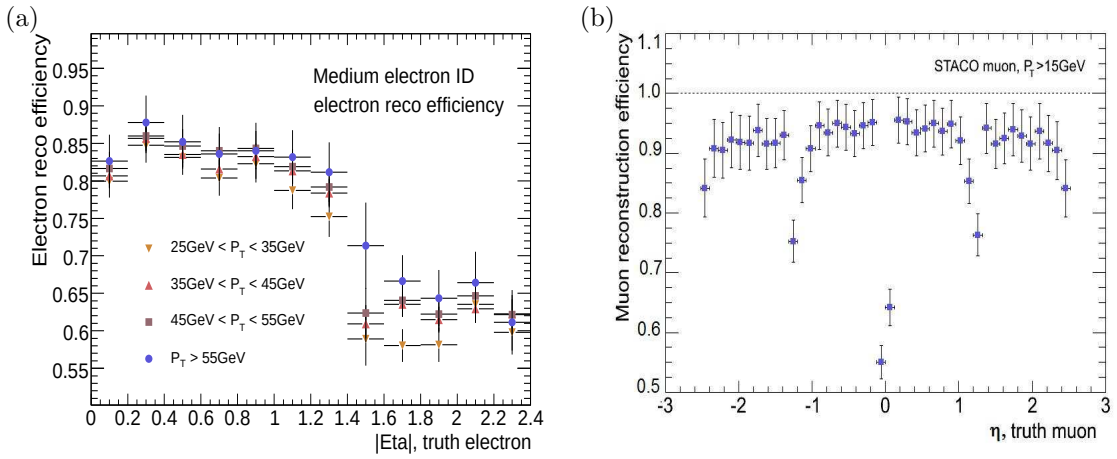


Figure 5.2: Electron reconstruction efficiency for different p_T values as a function of η (a). Electron reconstruction efficiency is clearly lower in the end-cap regions than in the barrel part (transition region at $|\eta| \sim 1.35$). Combined Staco muon reconstruction efficiency for $p_T > 15$ GeV as a function of η, where a drop of efficiency in the barrel/end-cap region transition is observed, $|\eta| \approx 1.2$-1.3 (b).

[3]Efficiency is defined as *Efficiency = (Number of Truth objects matched by a reconstructed object) divided by (Total number of Truth objects)*

Jet Reconstruction

We use jets clustered with the standard ATLAS seeded-cone algorithm with a radius of $R = 0.4$, built from either calorimeter towers ($Z \rightarrow e^+e^-$ analysis) or topological clusters ($Z \rightarrow \mu^+\mu^-$ analysis), and calibrated to hadron level, using global hadron calibration (H1-style [86]). Historically, calorimeter tower jets have been used in ATLAS, but nowadays topological cell clusters are also being used. They perform differently and, in order to enrich the analysis, we use both types of clustering with the same jet finder algorithm.[4] Appendix C details the study performed on different jet algorithms for Z+jets, before coming to this choice. Figure 5.3 shows reconstruction efficiency for calorimeter tower jets and for topological cluster jets as a function of its transverse momentum p_T[5]. The lepton and jet candidates must be separated by $\Delta R > 0.4$ (jets which are closer than this value to a selected lepton are removed). This removal cut is also applied in MCFM. Finally, it is required that the jet transverse momentum be larger than 40 GeV and that the jet be in the range $|\eta| < 3.0$.

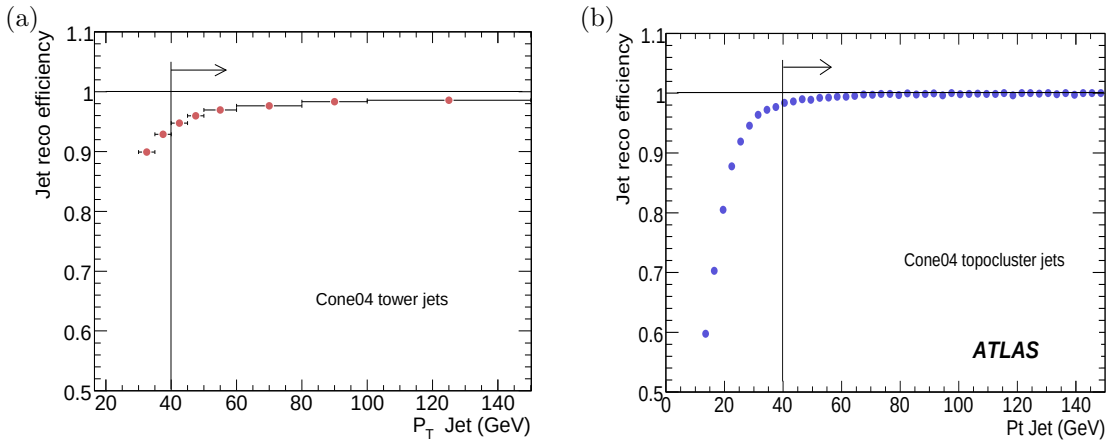


Figure 5.3: Jet reconstruction efficiency using Calorimeter Tower jets (a) in $Z(e,e)$ events and Topological Cell Cluster jets (b) in $Z(\mu,\mu)$ events, using seeded-cone $R = 0.4$ jet algorithm.

Z Selection

Two leptons (e or μ) candidates are required within $81 < M_{l,l} < 101$ GeV. If more than two leptons are reconstructed (it represents less than 0.5% of the events), the pair with closest invariant mass to M_Z is selected. The analysis is performed with events with a selected Z. All distributions are shown for events with a reconstructed Z.

5.2.2 Trigger Selection

In the electron channel, $Z \rightarrow e^+e^-$+jets are required to pass the isolated di-electron trigger (*2e15i*) or the isolated single-electron trigger (*e25i*)[6]. In the muon channel, $Z \rightarrow \mu^+\mu^-$+jets events are required to pass the di-muon trigger (*2mu10*) or the single-muon trigger(*mu20*)[7]. No matching to the trigger object is required.

[4] As explained, two different analysis have been performed in parallel, using similar but not identical tools. This has provided us an efficient way to check results. In this case, different input objects to the jet algorithm have been used, providing us with similar and compatible results.

[5] Jet reconstruction efficiency is defined with respect to truth-level jet.

[6] Isolation requirement means to ask for energy deposited inside a ring between 2×2 and 4×4 trigger towers to be below than 4 GeV (trigger towers have a size of 0.1×0.1 each one).

[7] Both muon trigger elements are non-isolated, since LVL2 muIsol algorithm was not yet implemented.

For the Z inclusive sample, Monte Carlo truth based information and data-driven tag-and-prove method [87] have been implemented. Good agreement between both methods has been found. Trigger efficiency is 99.63 ± 0.11% ($Z \rightarrow e^+e^-$) and 99.15 ± 0.25 % ($Z \rightarrow \mu^+\mu^-$) for the Z inclusive sample. Trigger efficiencies have been evaluated also as a function of jet multiplicity, using Monte Carlo truth based information. As shown in figure 5.6, trigger efficiency i stable as a function of jet multiplicity (up to 4 jets). Trigger efficiency for $Z+ \geq 1, 2, 3, 4$ jets events is ≈ 1.5-2.0% ($Z \rightarrow e^+e^-$) and ≈ 0.5-1.0% ($Z \rightarrow \mu^+\mu^-$) lower than the corresponding efficiency for the inclusive sample. Trigger efficiencies have been also studied as a function of the p_T of the leading jet, showing no bias for high p_T values, as shown in figure 5.6(b).

5.3 Lepton Reconstruction in a Multi-Jet Environment

The production of the Z boson with increasing number of jets in the event has an impact on the kinematics of both leptons and jets. Figure 5.4 shows the distributions of truth-level electron p_T and jet p_T. Figure 5.5 shows ΔR between electrons and between muons, and the minimum ΔR between each lepton (electron or muon) and the closest jet (for different jet multiplicities). As expected, in high-jet multiplicity events, the electrons (or muons) are more boosted (larger p_T and lower ΔR between them) and the distance between electrons (muons) and the closest jet becomes smaller. The average jet transverse momentum decreases with the number of jets.

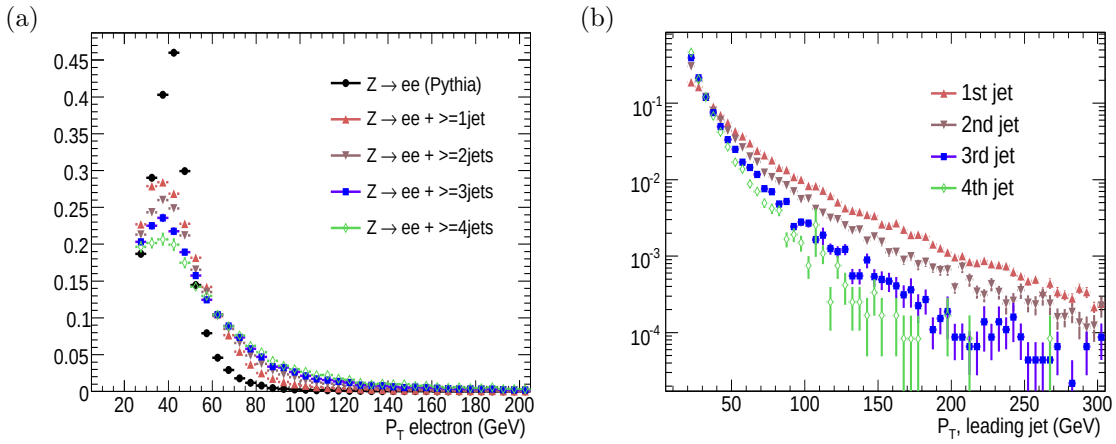


Figure 5.4: Distribution of electron p_T(a) and jet p_T(b) for samples with a different minimum number of reconstructed jets with $p_T > 20$ GeV. The events are generated with PYTHIA for the inclusive sample (black dots (a)) and with ALPGEN for Z+jets (colored dots). All histograms are normalized to unity.

Due to the OR of the single electron and dielectron trigger channels used in the analysis, the impact of an efficiency loss of the isolated single electron trigger for large jet multiplicities is reduced to the percent level. No isolation has been applied at reconstructed level. The total Z reconstruction efficiency (offline+trigger) for $Z \rightarrow e^+e^-$ events is stable with respect to both the jet multiplicity and the transverse momentum of the leading jet, as shown in figure 5.6.

In the case of the muon channel, reconstruction efficiencies and rejection for QCD multi-jet background are investigated for different muon isolation criteria. The isolation requirement for this analysis ($E_T < 15$ GeV inside a cone of ΔR=0.20) is chosen such that it presents no significant bias for events with large jet multiplicities and large jet p_T while at the same time providing a sufficiently large rejection for the QCD multi-jet background, as shown in figure 5.7 and table 5.1.

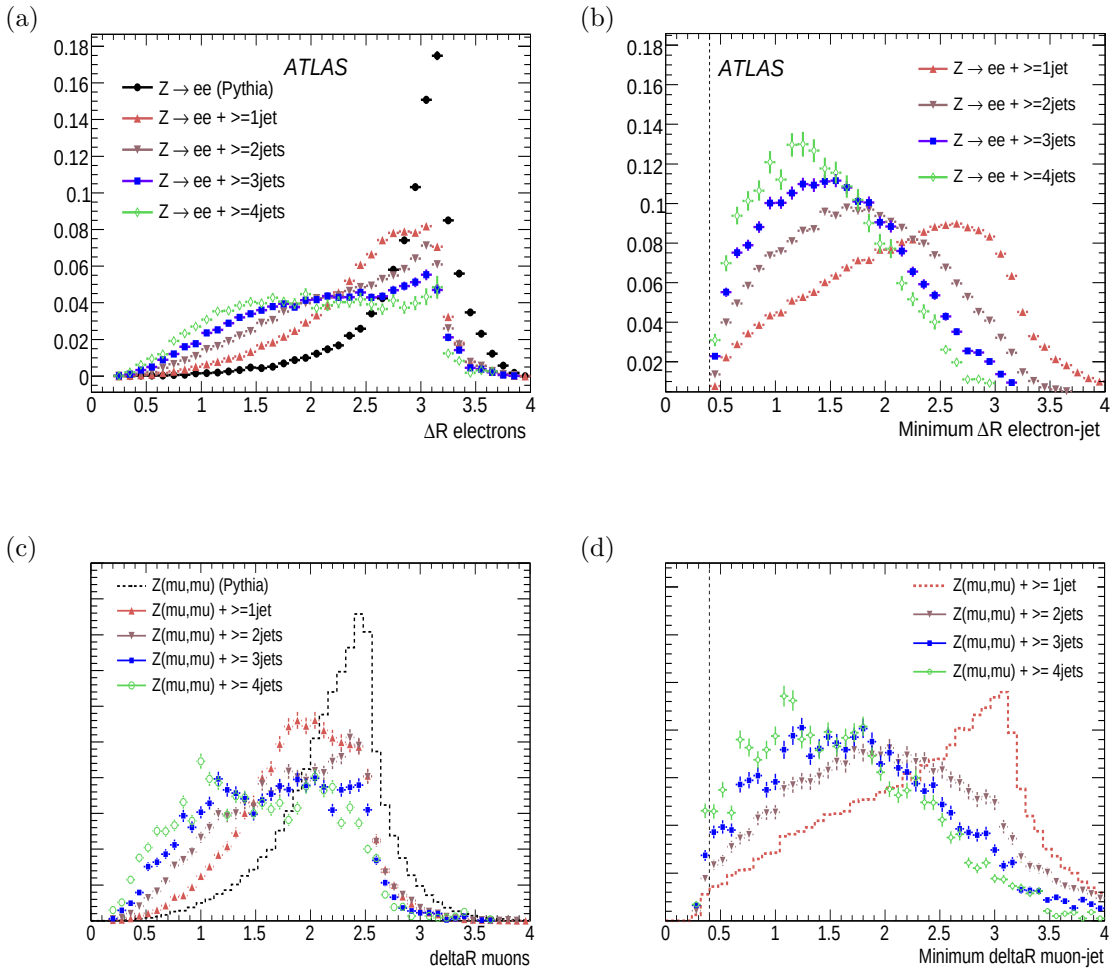


Figure 5.5: Distribution of ΔR between electrons (a), the minimum ΔR between each electron and the jets (b), distribution of ΔR between muons (c), and the minimum ΔR between muon and the jets (d) for samples with a different minimum number of reconstructed jets with $p_T > 20$ GeV. The events are generated with PYTHIA for the inclusive sample (black dots (a) and (c)) and with ALPGEN for Z+jets (colored dots). All histograms are normalized to unity.

	no isolation		$E_T < 15$ GeV in cone 0.20		$E_T < 15$ GeV in cone 0.40	
Process	events	fraction (%)	events	fraction (%)	events	fraction (%)
ALPGEN(signal) & PYTHIA(backgrounds), MCatNLO (top bgd)						
$Z \rightarrow \mu^+\mu^-$	57220±668	78.8±0.9	56619±664	95.2±1.1	54182±649	97.9±1.2
QCD($b\bar{b}$)	12112±1298	16.7±1.8	1600±171	2.7±0.3	0±898	0.0±1.6
$t\bar{t}$	1946±75	2.7±0.1	1084±63	1.8±0.1	953±53	1.7±0.1
$W \rightarrow \mu\nu$	1327±11	1.83±0.01	175±5	0.29±0.01	175±5	0.32±0.01
$Z \rightarrow \tau\tau$	22±10	0.03±0.01	12±5	0.02±0.01	12±5	0.02±0.01

Table 5.1: Number of events expected from signal and background in $Z \rightarrow \mu^+\mu^-$+jets, with an invariant mass $81 < M_{\mu,\mu} < 101$ GeV and requiring at least 1 jet, using different isolation criteria for muons, for $\int L\,dt = 1\text{fb}^{-1}$. The fraction in % is relative to all events selected (signal + backgrounds).

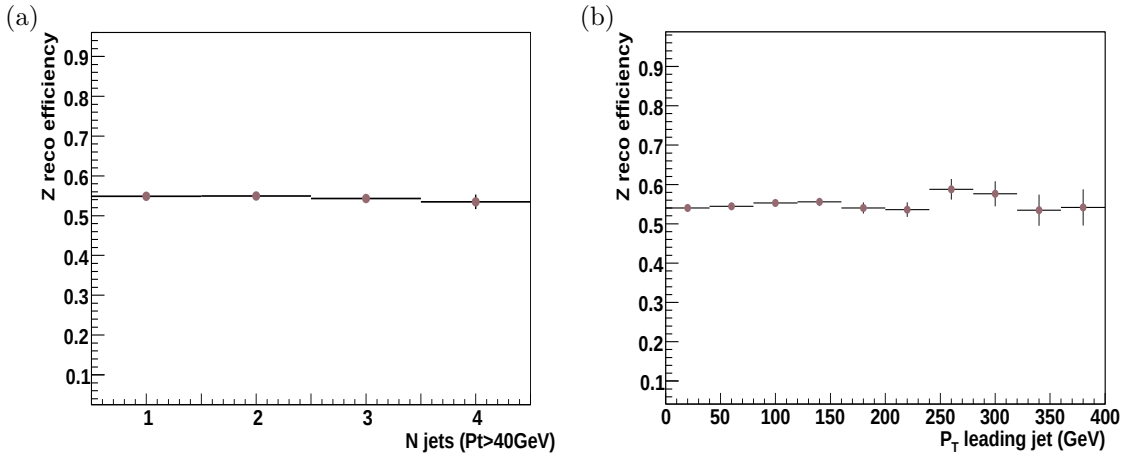


Figure 5.6: Event reconstruction efficiency as a function of the jet multiplicity (a) and the p_T of the leading jet (b), for $Z \rightarrow e^+e^-$ events, including trigger selection.

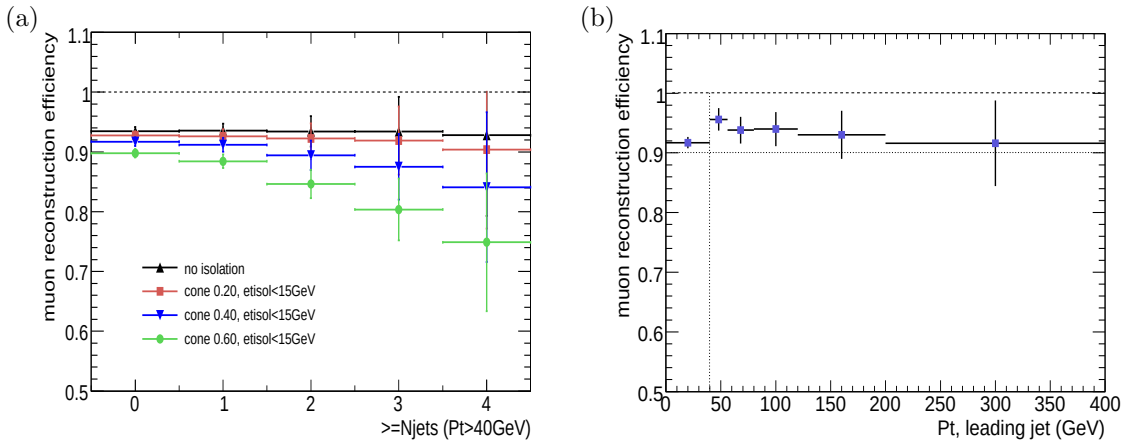


Figure 5.7: Muon reconstruction efficiency for different muon isolation cuts (a), as a function of the jet multiplicity, and vs the p_T of the leading jet (b).

5.4 Background Estimation

For the evaluation of background processes to the $Z \rightarrow l^+l^-$+jets signal process, we have considered processes with real leptons, such as top events, $W \rightarrow l\nu$, $Z \rightarrow \tau\tau$ and QCD multi-jet production. Background processes are estimated from fully-simulated Monte Carlo samples, generated with PYTHIA (backgrounds for $Z \rightarrow e^+e^-$+jets and backgrounds for $Z \rightarrow \mu^+\mu^-$+jets) and MCatNLO (top background for $Z \rightarrow \mu^+\mu^-$+jets), as shown previously in tables 4.2, 4.3. They are compared with the signal distributions derived from the respective ALPGEN Z+jets data sets.

5.4.1 QCD Jets Background Estimation

QCD dijet background is going to be a major background in many physics channels. Due to its huge cross section, orders of magnitudes larger than Z+jets processes, and its topology it is going to be difficult to quantify precisely.

$Z \rightarrow e^+e^-$+jets Channel

In the case of the electron channel, QCD background is derived from a filtered PYTHIA dijet sample. In this sample, the summed transverse energy of all stable particles (excluding muons and neutrinos) in a region $\Delta\eta \times \Delta\phi = 0.12 \times 0.12$ is required to be greater than 17 GeV. The goal of this filter at the generator level is to increase, in an unbiased way, the probability that the selected jets pass the electron identification cuts after the event simulations, i.e. to enrich the electromagnetic component of the filtered dijets at the generator level in order to get a larger jet to electron fake rate.

Statistics for this sample is increased by applying a very loose electron selection. An electromagnetic tower with transverse energy above 3 GeV is taken and a track from among all reconstructed tracks which do not belong to a photon conversion pair reconstructed in the inner detector is required to match the cluster within a broad $\Delta\eta \times \Delta\phi$ window of 0.05×0.10. The ratio, E/p, of the energy of the selected cluster to the momentum of the track is required to be lower than 10. After the electron selection, a re-weighting of the events with the additional rejection factor for the final electron ID is performed.

$Z \rightarrow \mu^+\mu^-$+jets Channel

QCD multi-jet background for highly-energetic muons results mainly from decays of $B\bar{B}$ mesons. We thus use a $b\bar{b}(\rightarrow \mu\mu)$ Monte Carlo sample to evaluate this background. This is an optimal solution for the lack of statistics of QCD dijet events, due to its large cross section and very small muon efficiency production (need of a huge Monte Carlo production). This solution has also been used in the measurement of the inclusive Z cross-section [88]. If we compare a $b\bar{b}$-sample and an inclusive lepton QCD sample, the origin of high energetic muons agree to a large extend in both samples and hence the above assumption is justified (see results on table 5.2). The QCD-sample is therefore labeled as $b\bar{b} \rightarrow \mu\mu$ in the following.

Sample	Fraction of mesons, with $\geq$ one b-quark	Fraction of mesons, with $\geq$ one c-quark	Fraction of other origin
Incl. lepton sample	0.73 ± 0.15	0.23 ± 0.08	0.04 ± 0.04
$b\bar{b}$-sample	0.734 ±0.001	0.243 ±0.001	0.023±0.001

Table 5.2: Comparison of the origin of high energetic muons in $b\bar{b}$ sample and inclusive lepton QCD sample. Muon $p_T \geq 10$ GeV is required.

In order to get a good estimation of this background, even for events with large jet multiplicity (≥3,4 jets) where statistics is small, the event selection has been modified, leaving as the last cut to be applied, the isolation cut on muons, since it represents the most restrictive cut, and adding a medium one, called soft isolation cut, which allows a higher energy deposition in the calorimeter (less than 25 GeV) in a cone size of $\Delta R = 0.2$. QCD($b\bar{b} \rightarrow \mu\mu$) events are selected by requiring initially non-isolated truth-level muons. Then, these events are reweighted by applying the following factors (figure 5.8):

- factor (Truth muons non isolated to soft isolated) 1.82 ± 0.33 (5.1)
- factor (Truth muons soft isolated to isolated) 4.17 ± 0.80 (5.2)
- factor (Truth muons to reconstructed muons (non isolated)) 1.51 ± 0.33 (5.3)

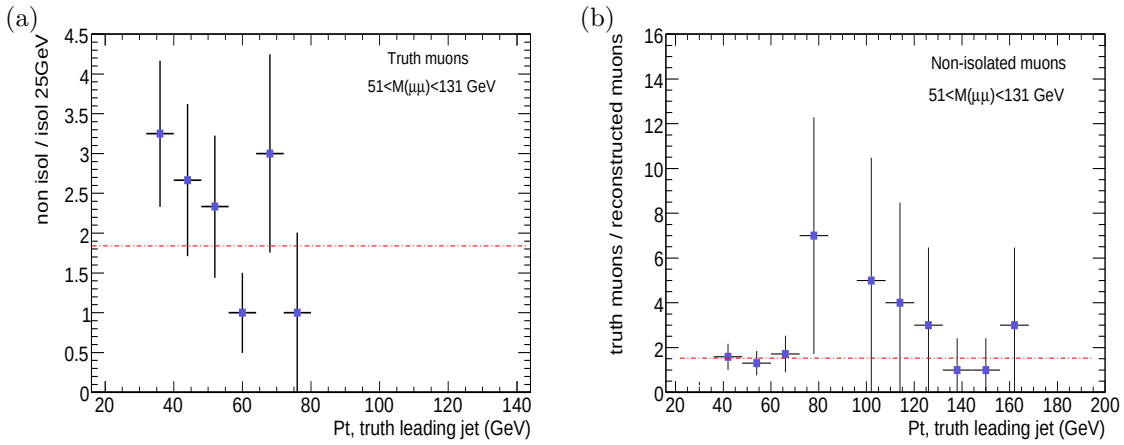


Figure 5.8: Factors applied on the QCD dijet background estimation. Factor (5.1) relating non isolated truth muons to soft isolated muons (a) and efficiency factor (5.3) relating reconstructed to truth muons of QCD dijets (b), as a function of the p_T of the jet. The invariant mass window used to compute these factors is $51 < M_{\mu,\mu} < 131$ GeV.

A scaling factor of 0.13 ± 0.02 applied to the di-muons candidates with an invariant mass of $51 < M_{\mu,\mu} < 131$ GeV , to get an estimation for the muons inside the invariant mass of $81 < M_{\mu,\mu} < 101$ GeV, used in the analysis. This scaling factor is calculated selecting events with truth muons inside $51 < M_{\mu,\mu} < 131$ GeV and comparing with the number of events obtained requiring pairs of truth muons in $81 < M_{\mu,\mu} < 101$ GeV. These muons are required to pass the default cuts for the analysis (detailed in section 5.2), with no isolation cut. Figures 5.9 show the p_T distribution of the QCD leading jet and second leading jet, selecting non-isolated truth-level muons, within the different invariant mass regions, in order to show that no obvious bias is induced on the p_T distribution using the scaling factor on di-muons selected inside the invariant mass window 51-131 GeV and scaling them down to the window of 81-101 GeV .

The second factor (5.2), where final isolation cut is performed, depends on the jet multiplicity, due to the topology of QCD events. This effect has been taken into account on the final QCD jet estimation, applying an increasing factor on each jet multiplicity (factor ≈ 4.2 ± 0.4 for ≥1 jet, factor ≈ 8.34 ± 0.2 for ≥2 jets, factor ≈ 16.7 ± 0.1 for ≥3 jets and ≈ 33.2 ± 0.1 for ≥4 jets).

In summary, two non-isolated truth muons are selected within an invariant mass of $51 < M_{\mu,\mu} < 131$ GeV , where better statistics is available. Then, the factors (5.1) and (5.2) are applied to these muons. Jet multiplicity dependence on the second factor (5.2) is taken into account. This estimation of truth isolated muons is translated to reconstructed muons using the factor (5.3). Finally, the scaling factor of the invariant mass is also applied. At the end, we get an estimation on reconstructed isolated muons within $81 < M_{\mu,\mu} < 101$ GeV.

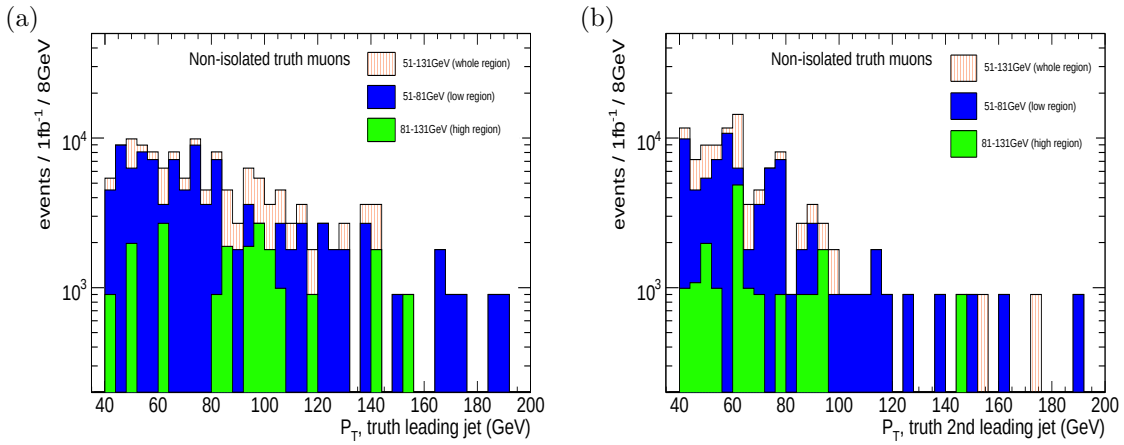


Figure 5.9: p_T jet distribution for different di-muon invariant mass regions for QCD di-jet background, where no bias in the distributions is shown.

QCD di-jet background final estimation results, for the muon channel, are shown in table 5.3. Estimation uncertainties are of the order of 50% of the QCD di-jet cross-section. Note that only 1.1pb^{-1} of data has been used (a factor 10^3 lower in statistics), which implies a worse precision in the factors used in the estimation method. Ultimately, this type of estimation methods will be applied to real data.

QCD jets	$Z+\geq$1jet	$Z+\geq$2jets	$Z+\geq$3jets	$Z+\geq$4jets	$Z+\geq$5jets	$Z+\geq$6jets
selected initial events	137505.7	97062.8	39544.1	11683.5	3594.9	898.7
statistical uncertainty	± 3515.3	± 9339.9	± 5961.5	± 3234.2	± 1797.5	± 898.7
final estim. events	1568.0	553.4	112.7	16.7	2.6	0.2
statistical uncert.	±126.7	±53.2	±17.0	±4.6	±1.3	±0.3
estimation uncert.	±940.7	±332.0	±67.6	±9.9	±1.5	±0.2

Table 5.3: Final number of events estimated for QCD background and its corresponding errors, using $b\bar{b}(\mu\mu)$ data sample and changing cut selection order as described, for $\int L\,dt = 1\text{fb}^{-1}$. Statistical uncertainty comes from the generated data sample statistics and estimation uncertainty comes from the uncertainty on the different factors used in the method. Both errors are related.

5.4.2 Top Background Estimation

One of the main backgrounds for Z+jets analysis is the top pair production, together with QCD di-jet processes (previously described). Top events have a similar signal topology as Z+jets processes, including leptons and multi-jets production. The main difference is the presence of neutrinos and b-jets in top events. Both characteristics are not going to be easily measured within the first days of data taking. For many physics analysis, top pairs need to be understood in a very precise way. For the purposes of our analysis, two different top estimation approaches have been studied for the $Z(\mu,\mu)$ channel.

$Z \rightarrow \mu^+\mu^-$+jets Channel

The first approach uses the presence of neutrinos in top events, coming from the decay of W. This feature can be used to distinguish both processes, our signal, with no real missing E_T, and top background, with missing E_T. Figure 5.10 shows the M_{μ^+,μ^-} invariant mass as a function of missing E_T for $Z \rightarrow \mu^+\mu^-$ +jets events and for top pair events. Signal events are centered at $81 < M_{\mu,\mu} < 101$ GeV and have small missing E_T (< 30 GeV), whereas top events present a more spread topology, with higher values of missing E_T, as expected.

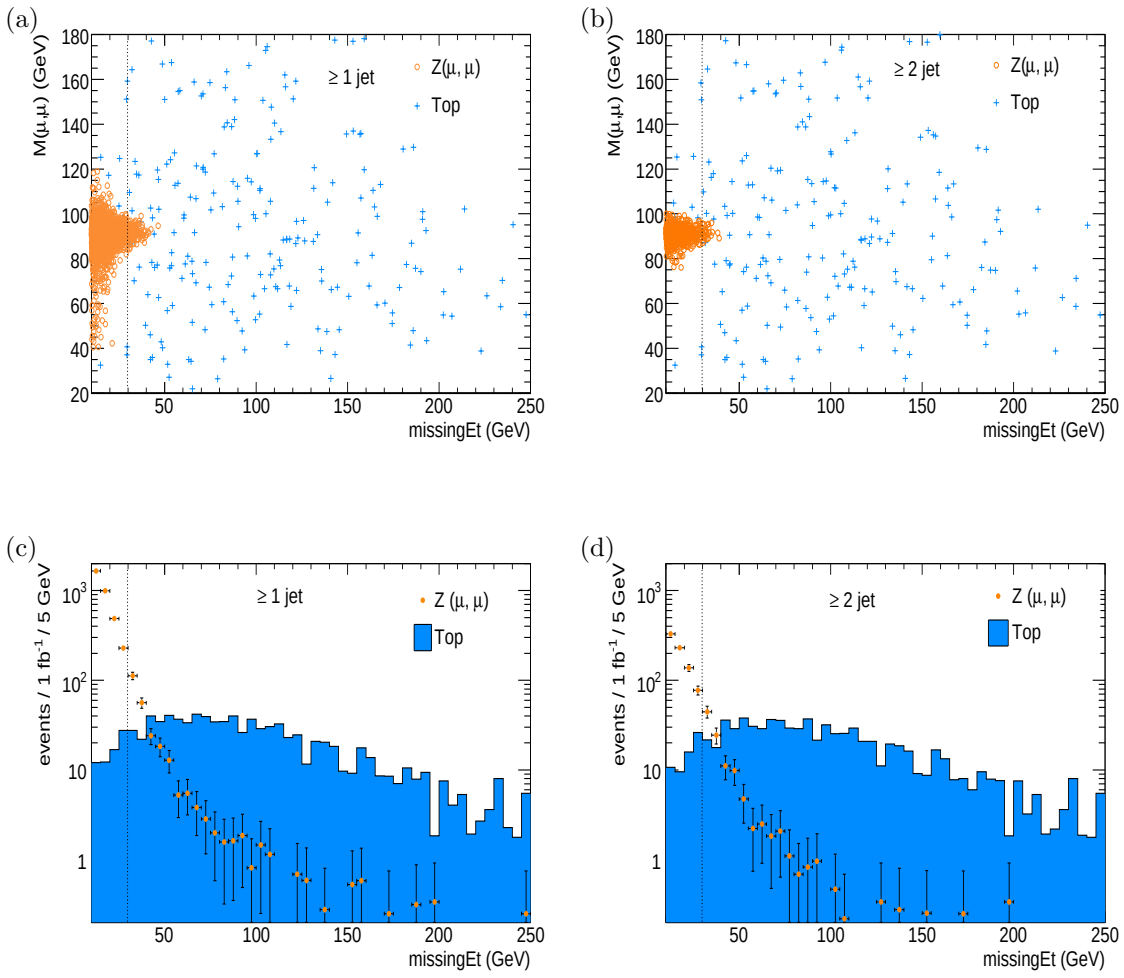


Figure 5.10: Top background and signal M_{μ^+,μ^-} invariant mass as a function of missing E_T distributions for different jet multiplicities (a), (b). Signal and top background events present clearly different topologies. Missing E_T distributions for signal and top background events (c), (d)

Consequently, a cut on missing E_T at 30 GeV can be used to distinguish signal from background. For missing $E_T < 30$ GeV, $Z \rightarrow \mu^+\mu^-$ +jets survives ($\sim 96\%$ for $Z \rightarrow \ell\ell+ \geq 1$jet), while top background decreases (to a level of 10-15%), as shown in figure 5.11(a). If the opposite cut is performed, missing $E_T > 30$ GeV, signal is reduced (only 4-20% survives depending on the jet multiplicity, mainly inside $81 < M_{\mu,\mu} < 101$ GeV), and eliminated at the region of $M_{\mu,\mu} > 110$ GeV, where only top events survive, as shown in figure 5.11(b). In the case of QCD di-jet events, if we apply a cut missing $E_T < 30$ GeV, its contribution is negligible for $M_{\mu,\mu} > 110$ GeV.

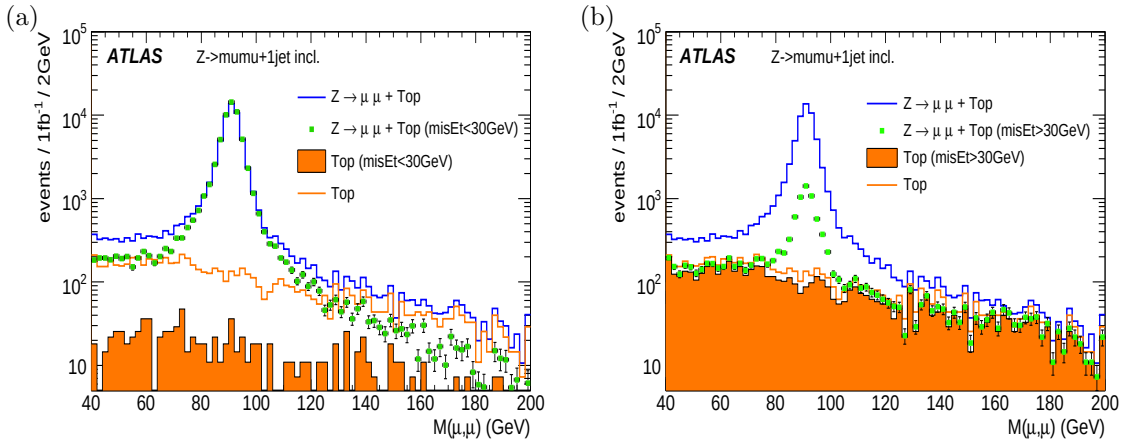


Figure 5.11: Top background and signal invariant mass for different missing E_T cuts, missing E_T <30 GeV (a) and missing E_T ≥30 GeV (b).

Although a cut on missing E_T would be effective to control the background, it can not be applied in our analysis as explained hereafter. In $Z \rightarrow \mu^+\mu^-$ +jets events, missing E_T comes from misidentified jets, so a cut on this variable would bias our final p_T jet distributions, specially at high p_T values, as shown in figure 5.12. On the other hand, it can be used to understand top background in the upper band of the Z peak (ie. $M_{\mu,\mu} \geq 101$ GeV), where no signal is present, and apply this knowledge to the central invariant mass region ($81 \leq M_{\mu,\mu} \leq 101$ GeV), where the signal is found.

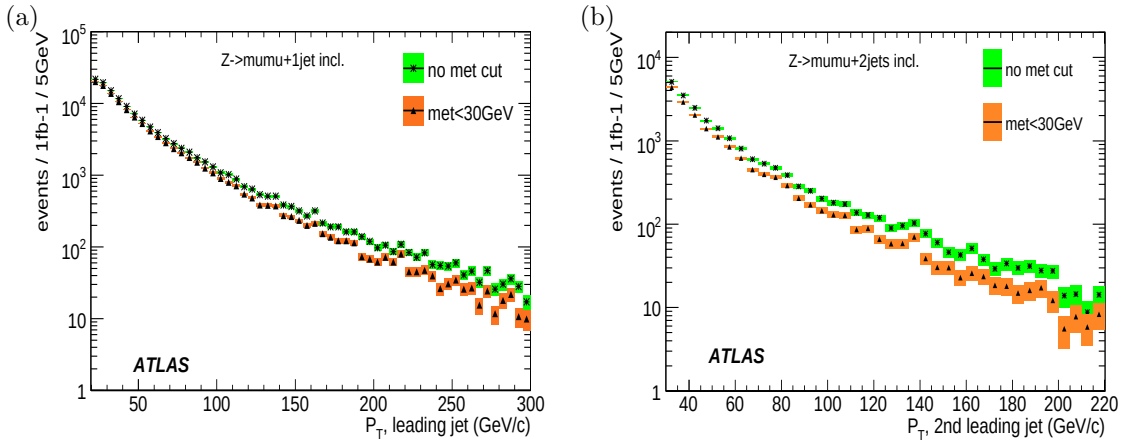


Figure 5.12: Jet p_T signal distributions applying a missing E_T cut of 30 GeV, for the leading jet (a) and for the second leading jet (b) in $Z \rightarrow \mu^+\mu^-$ events.

We select events, including top background and $Z \rightarrow \mu^+\mu^-$ +jets signal events, with an invariant mass of $M_{\mu,\mu} > 110$ GeV, and we apply a cut on missing $E_T > 30$ GeV, obtaining the distributions shown in figure 5.13, called "top estimation" (orange dots). Then, we compare it with top distribution (green area) with no missing E_T cut on the region $M_{\mu,\mu} > 110$ GeV, and we obtain a weighting factor from Monte Carlo data of 1.3 ± 0.1, obtained from the ratio of the number of selected events (to be applied to top background to obtain "top estimation").

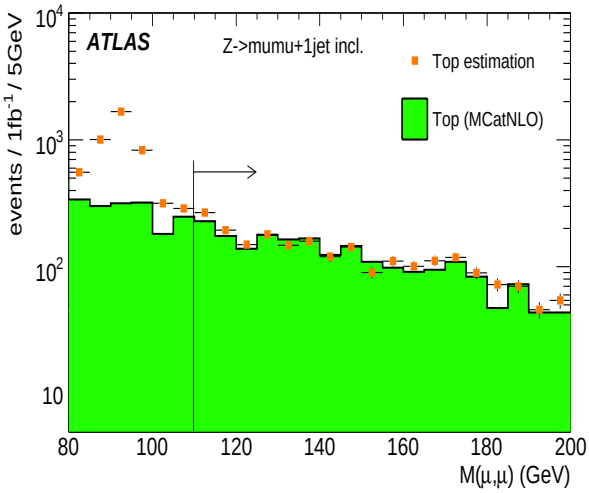


Figure 5.13: Top background estimation. Signal plus top MCatNLO background estimation (obtained from signal+top background with missing E_T cut > 30 GeV, and applying a weighting factor of 1.3) compared with top MCatNLO background with no met cut (a).

Finally, as expected, top jet p_T distributions are not biased due to the missing E_T cut since here it is true missing E_T (figure 5.14), therefore jet distributions from top estimation will not be biased, and consequently, they will contribute equally to our final $Z \rightarrow \mu^+\mu^-$ +jets cross section result.

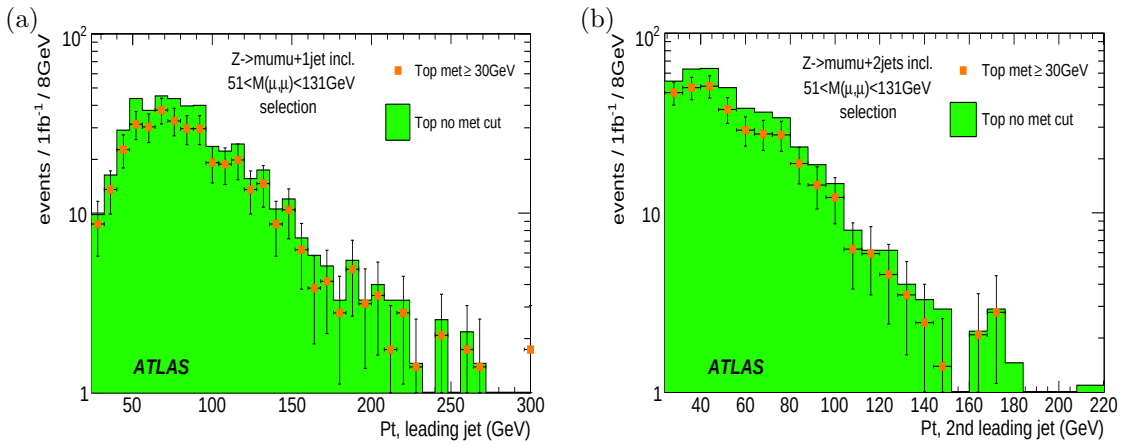


Figure 5.14: Jet p_T distributions applying a missing E_T cut of 30 GeV and compared with no missing E_T cut, for the leading jet (a) and for the second leading jet (b) in top events, with no factor applied.

The second approach to estimate top background is determining the relation of signal and background when using the usual analysis lepton selection (i.e, two muons, for the $Z(\mu\mu)$ channel) and when using a different selection: electron[8] + muon, where top events survive but Z+jets events are clearly eliminated. Figure 5.15 shows top background di-lepton invariant mass, selecting two muons and selecting an electron plus a muon. Both distributions have similar shape for different jet multiplicities, as expected.

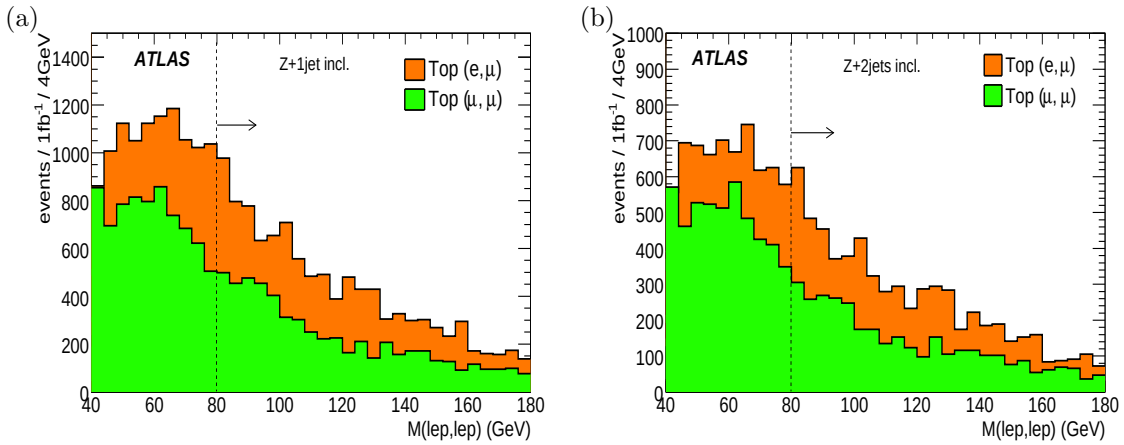


Figure 5.15: Top background invariant mass, for Electron+Muon and Muon+Muon selection.

Figure 5.16 shows the ratio between the invariant mass from top events decaying to electron+muon (e,μ) and top events decaying to muon+muon (μ, μ). If we assume this ratio to be constant for $M_{lep,lep} \geq$ 80 GeV and common to all jet multiplicities, its value is 1.64 ± 0.06. This factor can be applied on top(e,μ) events (where no signal is selected) to estimate top(μ, μ) background events in the region $M_{lep,lep} \geq 80$ GeV, and use this estimation to extrapolate and to subtract top events under the signal $Z \rightarrow \mu^+\mu^-$peak region. Other possible background contributions like diboson production, WW, ZZ, WZ, in electron+muon selection in the region $80 \leq M_{lep,lep} \leq 200$ GeV are not taken into account.

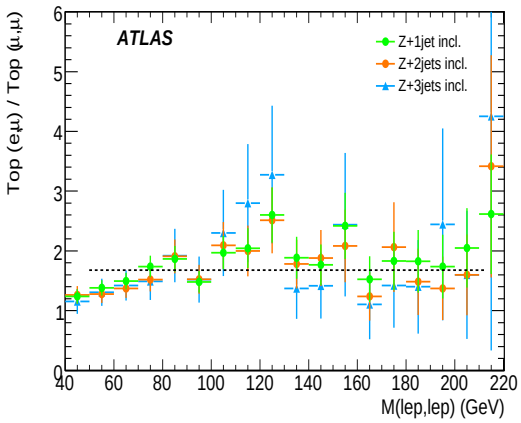


Figure 5.16: Ratio between top background invariant mass selecting an electron and a muon, and selecting two muons. For $M_{\mu,\mu} \geq 80$ GeV, a constant value of 1.64 ± 0.06 can be assumed.

[8]The electron selection follows the electron ID defined in section 5.2.1

Finally, a combination of the two previous estimation methods (electron+muon selection and a missingE_T cut) can be very useful to determine top background from real data and compare it with Monte Carlo predictions. The goal of the method is to compare top background events with a missing E_T cut (where nearly no signal events in $M_{\mu,\mu} > 110$ GeV region are), and selecting electron+muon in this region, and extrapolate our understanding to the central region ($81 \leq M_{\mu,\mu} \leq 101$ GeV), where we need to evaluate top background events. Figure 5.17 shows top background events applying this combined selection for different jet multiplicities: electron+muon selection and missing $E_T > 30$ GeV in (a) and muon+muon selection and missing $E_T > 30$ GeV in (b).

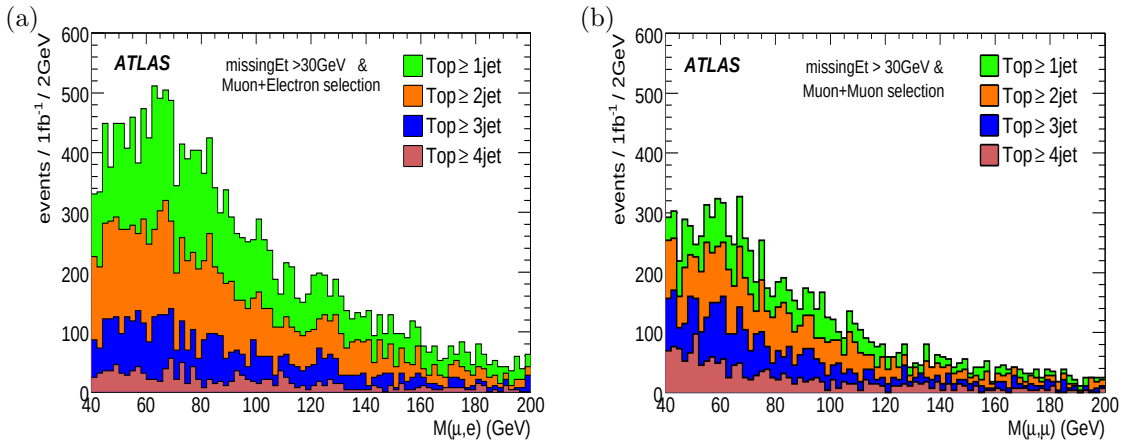


Figure 5.17: Top background (MCatNLO data sample) for different jet multiplicities, where Missing $E_T > 30$ GeV cut and one muon plus one electron selection, top(e,μ), have been applied (a) and two muons selection (b). These combined selections are used to estimate top background events, with no signal events. QCD background can be considered negligible (less than percentage level).

Factors to estimate top background have been derived, their values are shown in table 5.4. The initial selection is one truth-level electron plus one truth-level muon and a missing E_T cut of 30GeV. Then, as detailed in table 5.4, first the invariant mass window is moved from 110 GeV$\leq M_{l,l} \leq$ 200 GeV to 81 GeV$\leq M_{l,l} \leq$ 101 GeV (step 1) and then electron-muon selection is transformed to muon-muon selection (step 2). Finally, the factor to obtain top events with no missing E_T cut from top events with missing E_T cut is shown (step 3).

Table 5.5 shows the number of top events estimated from Muon+Electron and a missing $E_T > 30$GeV cut selection, in the analysis invariant mass window $81 \leq M_{l,l} \leq 101$GeV , where we need to subtract top background events. The table also contains the number of top events counted directly inside the same invariant mass window $81 \leq M_{l,l} \leq 101$GeVİn the last row, a comparison of both results in percentage is detailed. Differences increase with jet multiplicity, from 4% ($\geq$ 1jet), 10% ($\geq$ 2jets) up to 20% ($\geq$ 3jets) and 50% ($\geq$ 4jets). These differences can be taken into account as systematic uncertainties for top background estimation.

Factors used in Top estimation				
	≥1jet	≥2jets	≥3jets	≥4jets
	Top(e,μ) and missing E_T >30 GeV **110≤ M$_{e,\mu}$ ≤ 200 → 81≤ M$_{e,\mu}$ ≤ 101** GeV			
step 1	0.677 ± 0.016	0.686 ± 0.021	0.675 ± 0.032	0.629 ± 0.058
	missing E_T > 30 GeV and 81 ≤ M$_{e,\mu}$ ≤ 101 GeV **Top(e,μ) → Top(μ,μ)**			
step 2	0.436 ± 0.012	0.531 ± 0.017	0.672 ± 0.032	1.034 ± 0.084
	Top(μ,μ) and 81≤M$_{\mu,\mu}$ ≤ 101 GeV **missing E_T >30 GeV→ no missing E_T cut**			
step 3	1.307 ± 0.138	0.875 ± 0.092	0.554 ± 0.058	0.409 ± 0.042

Table 5.4: Factors used in top estimation.

Top estimation				
	≥1jet	≥2jets	≥3jets	≥4jets
initial selected events	Top(e,μ) and missing E_T >30 GeV and 110≤ M$_{e,\mu}$ ≤ 200			
	4681.5 ± 69.5	2781.0 ± 53.6	1155.6 ± 34.5	309.8 ± 17.9
estimated events	Top(μ,μ) and 81≤M$_{\mu,\mu}$ ≤ 101 GeV and no missingE_T cut			
	1123.1 ± 23.1	551.5 ± 10.6	180.4 ± 5.4	50.9 ± 3.0
counted events	Top(μ,μ) and 81≤M$_{\mu,\mu}$ ≤ 101 GeV			
	1083.7 ± 62.8	610.9 ± 47.1	225.5 ± 28.6	32.7 ± 10.9
difference estimated vs counted				
(%)	4.0	10	20	55

Table 5.5: Top estimation using factors shown in table 5.4 and applying muon reconstruction efficiency (see section 5.5) and top events counting directly from Monte Carlo, using the default analysis selection. Last row shows the difference between top estimated events and top counted events in %, which can be used as a systematic uncertainty estimation for top background events.

5.4.3 Other Background Processes

Other background processes are taken into account: $W \rightarrow e\nu$, $W \rightarrow \mu\nu$ and $Z \rightarrow \tau^+\tau^-$. We use PYTHIA data samples to count their contributions. In some cases, where available statistics is small, extrapolation from $Z \rightarrow \ell\ell+ \geq 1$jet is made to $Z \rightarrow \ell\ell+ \geq 2$jets, $Z \rightarrow \ell\ell+ \geq 3$jets and $Z \rightarrow \ell\ell+ \geq 4$jets. The ALPGEN corresponding data samples for all these background processes are also taken into account. Compatible number of background events are obtained. Final values for PYTHIA $W \rightarrow e\nu$, $W \rightarrow \mu\nu$ and $Z \rightarrow \tau^+\tau^-$ backgrounds are shown in tables 5.6 and 5.7. In both channels, their contribution represents less than 1% of the total accepted events, for the different jet multiplicities.

5.4.4 Background Estimation Results

Finally, distributions for the invariant mass $M_{lep,lep}$, number of events with $\geq$ Njets, p_T of the leading jet and p_T of the second leading jet for signal and backgrounds, including results from section 5.4.1 for QCD di-jet estimation, and from section 5.4.2 for top estimation, are shown in figure 5.18 for $Z \rightarrow e^+e^-$+jets analysis and in figure 5.19 for $Z \rightarrow \mu^+\mu^-$+jets analysis.

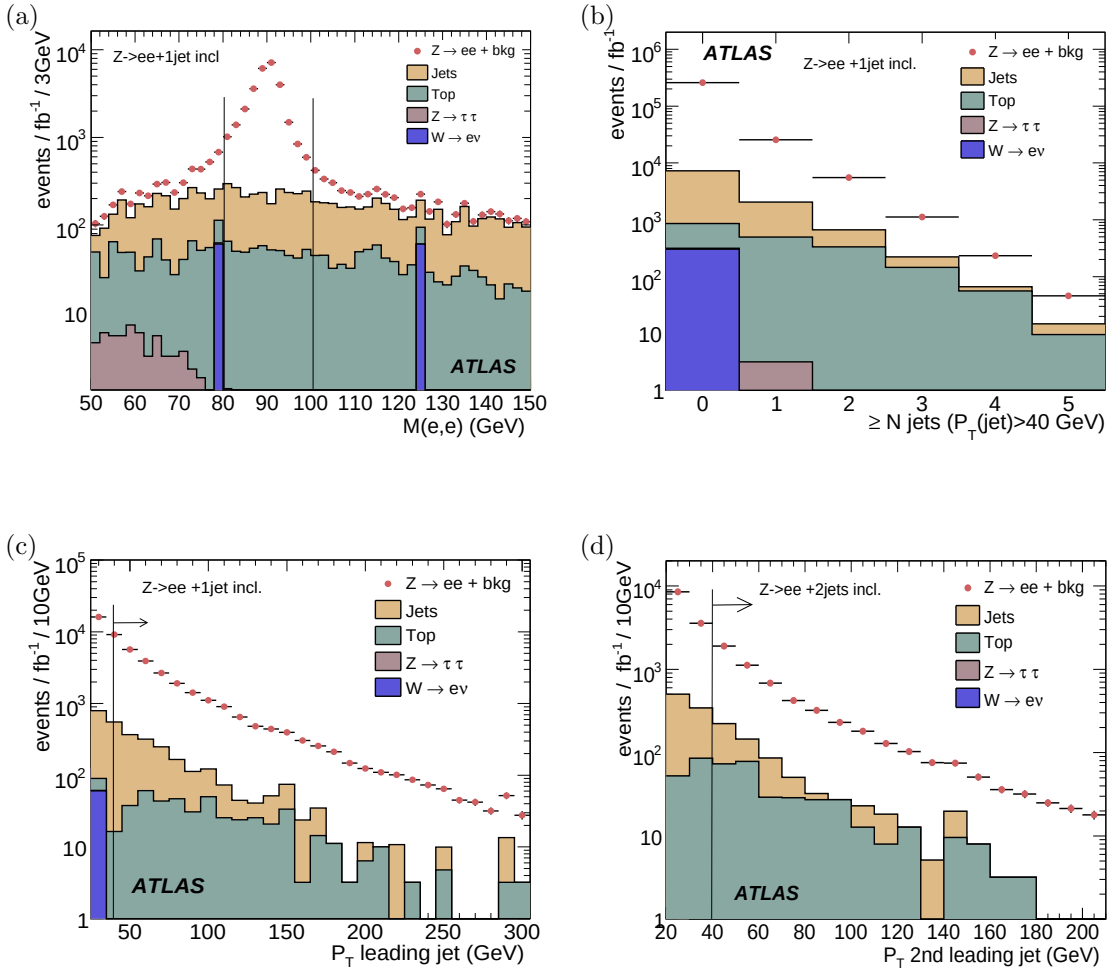


Figure 5.18: Signal and background distributions for $Z \rightarrow e^+e^-$ + jets analysis.

Tables 5.6 and 5.7 show final signal and background estimations, for $Z \rightarrow e^+e^-$+jets and $Z \rightarrow \mu^+\mu^-$+jets. For both channels, the main backgrounds are top decay and QCD di-jet events. Top decays represent between 2% ($Z \rightarrow \ell\ell+ \geq 1$jet) and 10-12% ($Z \rightarrow \ell\ell+ \geq 4$jets). In the case of QCD

	$Z \rightarrow \ell\ell+ \geq$ 1jet	$Z \rightarrow \ell\ell+ \geq$ 2jets	$Z \rightarrow \ell\ell+ \geq$ 3jets	$Z \rightarrow \ell\ell+ \geq$ 4jets
$Z \rightarrow e^+e^- +$ jets analysis				
Process	σ (fb)	σ (fb)	σ (fb)	σ (fb)
$Z \rightarrow e^+e^-$	23520±145	4894±45	900±15	168±5
QCD jets	1545±89	336±42	78±20	10±7
$t\bar{t}$	496±28	333±23	146±15	56±10
$W \rightarrow e\nu$	(28±13)	(5.9±2.6)	(1.1±0.5)	(0.2±0.1)
$Z \rightarrow \tau^+\tau^-$	3.2±1.2	(0.67±0.25)	(0.1±0.05)	(0.02±0.01)
total background	2072.2±94.2	675.6±48.0	225.2±25.0	66.2±12.2
total #events	25592±173	5570±66	1125±29	234±13
	fraction (%)	fraction (%)	fraction (%)	fraction (%)
$Z \rightarrow e^+e^-$	91.9±0.8	87.9±1.3	80.0±2.4	71.8±2.1
QCD jets	6.0±0.4	6.0±0.8	6.9±1.8	4.3±3.0
$t\bar{t}$	1.9±0.1	6.0±0.4	13.0±1.4	24.9±4.3
$W \rightarrow e\nu$	(0.1±0.05)	(0.1±0.05)	(0.1±0.05)	(0.1±0.05)
$Z \rightarrow \tau^+\tau^-$	0.01±0.01	(0.01±0.01)	(0.01±0.01)	(0.01±0.01)
total bgd/all events	8.0±0.6%	12.1±0.9%	20.0±2.3%	29.3±5.2%

Table 5.6: The accepted cross-sections (σ, in fb) and the corresponding fraction of the total sample (in %) for signal and for the background channels in the $Z \rightarrow e^+e^- +$ jets analysis, after applying the cuts outlined in section 5.2. The numbers in brackets are extrapolated from results obtained for a lower jet multiplicity.

	$Z \rightarrow \ell\ell+ \geq$ 1jet	$Z \rightarrow \ell\ell+ \geq$ 2jets	$Z \rightarrow \ell\ell+ \geq$ 3jets	$Z \rightarrow \ell\ell+ \geq$ 4jets
$Z \rightarrow \mu^+\mu^-$+jets analysis				
Process	σ (fb)	σ (fb)	σ (fb)	σ (fb)
$Z \rightarrow \mu^+\mu^-$	56619.4±664.0	12461.2±273.6	2458.1±99.3	428.1±33.4
QCD jets ($b\bar{b}$)	1568.0±126.7	553.4±53.2	112.7±17.0	16.7±4.6
$t\bar{t}$	1123.1±23.1	551.5±10.6	180.4±5.4	50.9±3.0
$W \rightarrow \mu\nu$	175.2±4.8	37.3±1.0	7.2±0.2	1.4±0.1
$Z \rightarrow \tau^+\tau^-$	12.2±5.0	2.0±2.0	1.0±1.1	0.0±0.0
total background	2878.5±129.0	1144.2±54.3	301.3±17.9	69.0±5.5
total #events	59497.9±676.4	13605.4±278.9	2759.4±100.9	497.1±33.9
	fraction (%)	fraction (%)	fraction (%)	fraction (%)
$Z \rightarrow \mu^+\mu^-$	95.2±1.1	91.6±2.0	89.2±3.6	86.1±6.7
QCD jets ($b\bar{b}$)	2.6±0.2	4.1±0.4	4.0±0.6	3.4±0.9
$t\bar{t}$	1.9±0.1	4.0±0.1	6.5±0.2	10.2±0.6
$W \rightarrow \mu\nu$	0.29±0.01	0.27±0.01	0.26±0.01	0.28±0.02
$Z \rightarrow \tau^+\tau^-$	0.02±0.01	0.02±0.02	0.04±0.04	0.0±0.0
total bgd/all events	4.8±0.2%	8.4±0.4%	10.8±0.6%	13.9±1.1%

Table 5.7: The accepted cross-sections (σ, in fb) and the corresponding fraction of the total sample (in %) for signal and for background channels in the $Z \rightarrow \mu^+\mu^- +$ jets analysis, after applying the cuts outlined in section 5.2 for $\int L\,dt = 1\text{fb}^{-1}$.

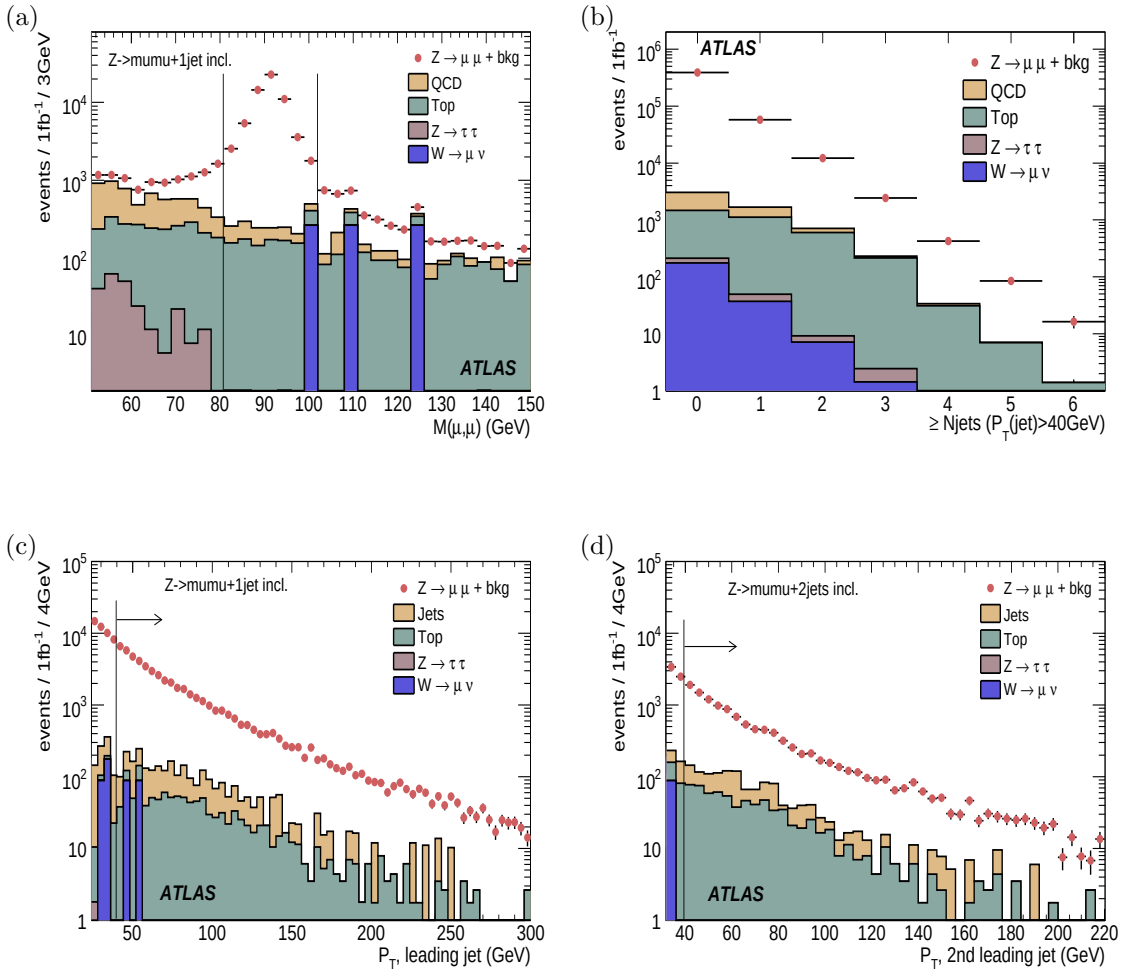


Figure 5.19: Signal and background distributions for $Z \rightarrow \mu^+\mu^- +$ jets analysis using the different background estimations.

di-jet events, these represent from 3% up to 7%. Small statistics for large jet multiplicities for these two backgrounds translate to large uncertainty on these backgrounds. So, top decays and QCD jets must be understood precisely from real data to evaluate them correctly. These percentage values must be understood as an approximated prediction.

5.5 Unfolding of the Detector Effects

The reconstructed data have to be unfolded from the detector level to the hadron level, correcting for efficiency, resolution and non-linearities in electron/muon and jet reconstruction.

In this study, the individual corrections are assumed to factorize in first approximation, and the individual contributions are investigated and corrected for separately. The corrections are detailed in the following for the case of the $Z \rightarrow e^+e^-$ +jets channel. Unfolding of the $Z \rightarrow \mu^+\mu^-$ final state is done in the same way, but using a different cone jet algorithm. It is extensively detailed in appendix D. All corrections are derived with fully-simulated Monte Carlo samples.

The dominant correction on the inclusive cross section for the $Z \rightarrow e^+e^-$ channel stems from the electron reconstruction. For each of the two electrons, the cross-section is corrected for the electron reconstruction efficiency, given as a function of the electron pseudo-rapidity, η, and its transverse momentum, p_T. These corrections are shown in figure 5.2(a) and summarized in table 5.8. The cross section is also corrected for the electron trigger efficiency with respect to the offline selection (factor 0.9963 $\pm$ 0.0011).

Electron reconstruction efficiency				
$\|\eta\|$ range	25$< p_T <$ 35 GeV	35$< p_T <$ 45 GeV	45$< p_T <$ 55 GeV	$p_T >$ 55 GeV
0.0 - 0.2	0.797 $\pm$0.019	0.807 $\pm$0.017	0.817 $\pm$0.022	0.827 $\pm$ 0.035
0.2 - 0.4	0.848 $\pm$0.022	0.857 $\pm$0.009	0.860 $\pm$0.023	0.878 $\pm$ 0.036
0.4 - 0.6	0.831 $\pm$0.023	0.835 $\pm$0.017	0.847 $\pm$0.023	0.852 $\pm$ 0.036
0.6 - 0.8	0.804 $\pm$0.024	0.816 $\pm$0.017	0.840 $\pm$0.023	0.836 $\pm$ 0.036
0.8 - 1.0	0.823 $\pm$0.042	0.833 $\pm$0.017	0.843 $\pm$0.023	0.840 $\pm$ 0.037
1.0 - 1.2	0.787 $\pm$0.025	0.813 $\pm$0.017	0.819 $\pm$0.023	0.831 $\pm$ 0.037
1.2 - 1.4	0.752 $\pm$0.027	0.784 $\pm$0.019	0.792 $\pm$0.025	0.811 $\pm$ 0.039
1.4 - 1.6	0.589 $\pm$0.036	0.609 $\pm$0.024	0.624 $\pm$0.034	0.724 $\pm$ 0.048
1.6 - 1.8	0.580 $\pm$0.025	0.635 $\pm$0.017	0.641 $\pm$0.022	0.666 $\pm$ 0.034
1.8 - 2.0	0.581 $\pm$0.026	0.615 $\pm$0.018	0.622 $\pm$0.023	0.644 $\pm$ 0.039
2.0 - 2.2	0.635 $\pm$0.025	0.629 $\pm$0.019	0.647 $\pm$0.025	0.664 $\pm$ 0.041
2.2 - 2.4	0.598 $\pm$0.024	0.623 $\pm$0.020	0.622 $\pm$0.026	0.611 $\pm$ 0.043

Table 5.8: Electron reconstruction efficiencies as a function of $|\eta|$ for four different p_T bins.

For the $Z \rightarrow \mu^+\mu^-$ analysis, muon reconstruction and trigger efficiency is 0.93 $\pm$ 0.01, for any jet multiplicity (N$\leq$4jets). This factor has been used for the whole muon p_T and η analysis range, since it has no dependence (constant value), as shown in figure 5.2(b). The inclusive cross section is also corrected for Z reconstruction efficiency (ie. select two reconstructed muons within the selected dimuon invariant mass to reconstruct the Z boson). The total Z reconstruction efficiency is 0.747 $\pm$ 0.009. This is the final correction applied to the $Z \rightarrow \mu^+\mu^-$ cross section, related to muon reconstruction and trigger efficiency.

Corrections from jet reconstruction have a comparably small impact on the overall cross section but bias the jet p_T spectrum since, in general, the detector effects are larger for low p_T jets. The reconstructed jet p_T is corrected for the non-linearity of the jet energy scale (figure 5.20(a)). For each jet required in the selection, the cross section is then corrected for the reconstruction efficiency (figure 5.20(c)) and for the effect of the jet energy resolution (figure 5.20(b)). The jet reconstruction efficiency is

determined as the fraction of truth jets which are matched to reconstructed jets. The jet energy scale and resolution are determined using a matching window of $\Delta R_{(truth-reco\ jet)} < 0.2$ and $0.5 < p_T^{reco}/p_T^{truth} < 1$ for $Z \rightarrow e^+e^-$ events and a matching of $\Delta R < 0.05$ for $Z \rightarrow \mu^+\mu^-$ events. The impact of the resolution is derived by comparing the p_T distribution of truth jets before and after a Gaussian smearing with the resolution as determined above. The errors on deriving these corrections, stemming from the limited Monte Carlo statistics, are taken into account as systematic errors on the cross section measurement.

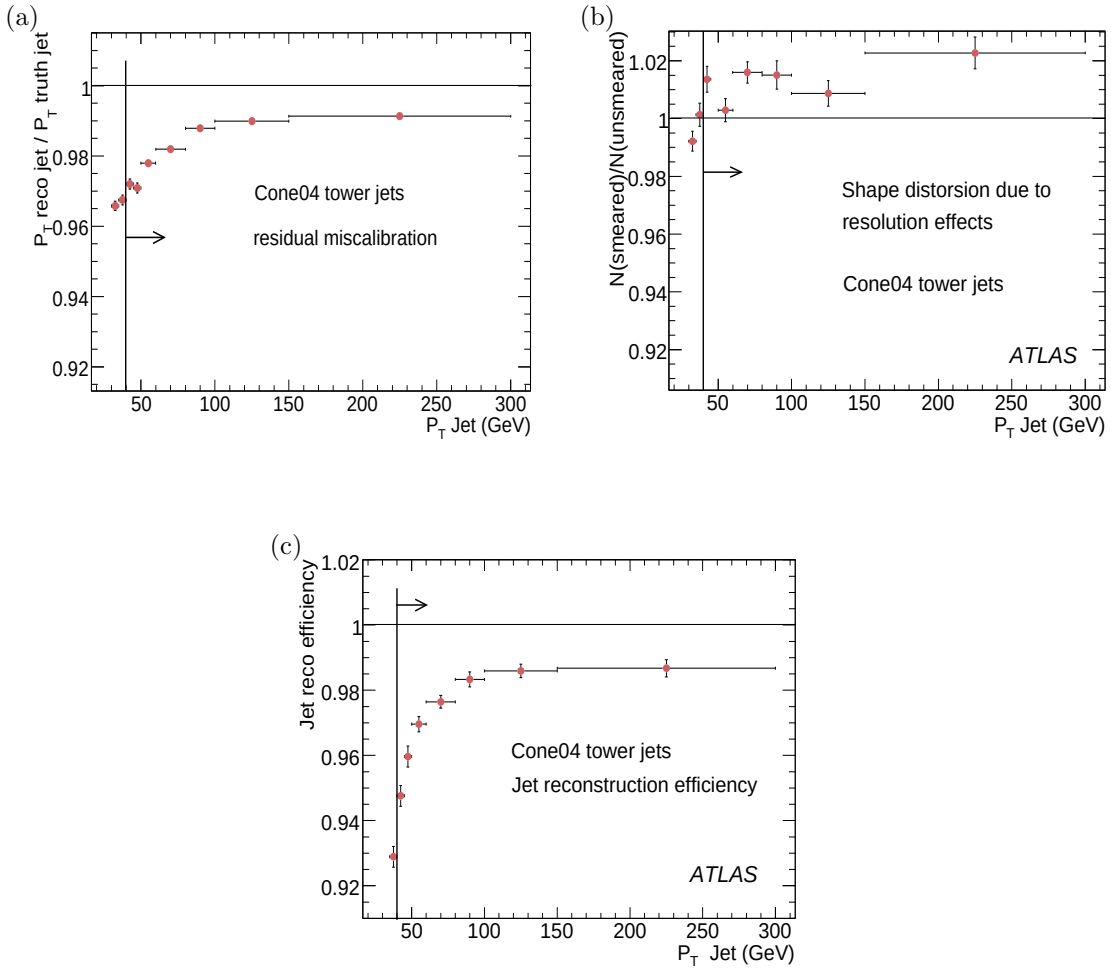


Figure 5.20: Non-linearity of the jet energy scale corrections (a), effect on the jet resolution (b) and jet reconstruction efficiency (c).

Figure 5.21(a),(b) compares the distributions of the p_T of the leading jet and the second leading jet, in different unfolding stages, with the p_T distribution of the original Monte Carlo jets (hadron level). Within the statistical and systematic errors, the p_T distributions of the Monte Carlo jets and the corrected reconstructed jets are in agreement, thus providing a consistency check for the unfolding corrections. Results corresponding to $Z \rightarrow \mu^+\mu^-$ channel, are detailed in appendix D. Figure 5.22 (a),(b) shows the ratio of the distribution of the p_T of the leading jet and the second leading jet, comparing reconstructed corrected values with values from truth-level Monte Carlo jets (hadron level). Unfolding factors have been calculated using ALPGEN data samples. For this reason a better agreement is obtained with ALPGEN (rather than PYTHIA) Monte Carlo data.

In summary, events have been selected following the recipe given in section 5.1. Crack regions are

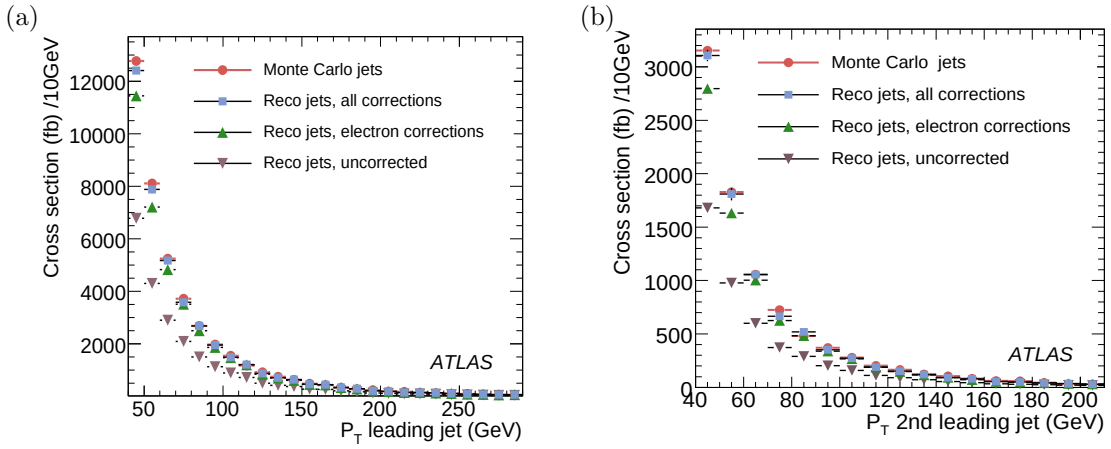


Figure 5.21: Comparison of the distributions of the p_T of the leading jet and second leading jet for the generated Monte Carlo and for the reconstructed quantities, without any correction, after electron reconstruction and triggering corrections, and final corrected quantities, including jet-related corrections. Within the statistical and systematic errors, the p_T distributions of truth jets and corrected reconstructed jets are in agreement.

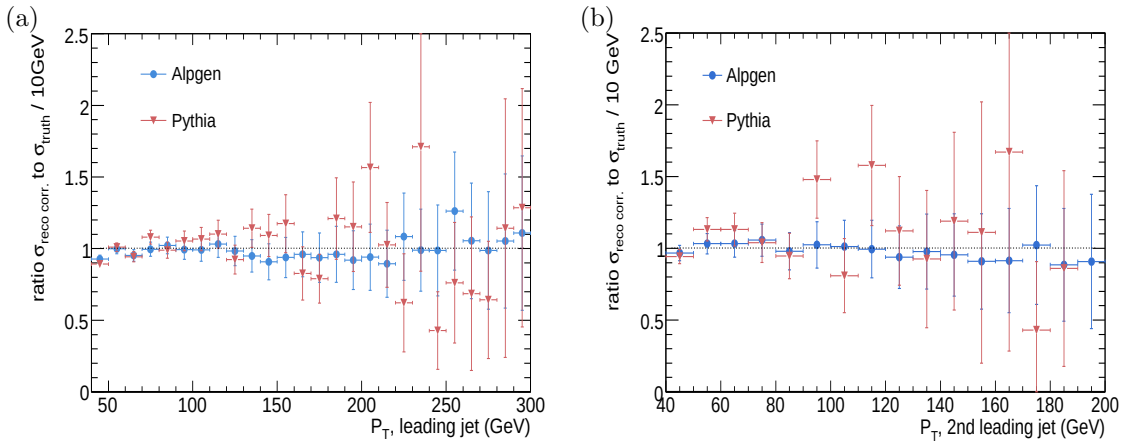


Figure 5.22: Ratio of reconstructed corrected to truth-level cross section of the p_T of the leading jet and the second leading jet, comparing ALPGEN and PYTHIA Monte Carlo predictions for $Z \rightarrow \mu^+\mu^-$ processes. Statistical errors are shown.

discarded. Background events are subtracted. Then lepton reconstruction and trigger efficiency factors are applied. Finally, unfolding jet factors are applied to jet distributions (jet reconstruction efficiency, and resolution and linearity factors) to obtain distributions at hadron level.

5.6 Systematic Uncertainties

In order to determine the expected precision of the analysis, the cross section measurement is performed on the fully-simulated ALPGEN Z + jets data sets, which are corrected to the hadron level, following the prescription of section 5.5. Systematic errors from the corrections are included.

An additional systematic uncertainty is introduced on the unfolding correction for the jet resolution due to the uncertainty on the jet resolution measurement and to the uncertainty on the shape of the p_T distribution which we use to derive the corrections. Using corrections from different event generators and varying the jet resolution within its uncertainty results in a systematic error on the cross section at the percent level.

Backgrounds can be estimated with data-driven methods, introducing additional systematic uncertainties. As seen in section 5.4, we account for that in first approximation by adding an error of 20% on the fraction of the multi-jet background for each jet multiplicity (in $Z \rightarrow e^+e^-$ channel) and an error of 20% for $Z \rightarrow \mu^+\mu^-$ process. For top background, an uncertainty of 15% is assumed. The statistical uncertainties in the samples are scaled to the number of events expected to be selected for an integrated luminosity of 1 fb^{-1}.

In the next step we evaluate the impact of the uncertainties expected for real ATLAS data taking. ATLAS expects a limited precision of the jet energy scale (JES) in the first years, starting from uncertainties at the level of 10% and converging eventually towards 1%. We obtain two benchmark scenarios, in the final results, by propagating jet energy scale uncertainties of 5% and 10% into the measured cross sections.

Figure 5.23 presents JES errors (1%, 5% and 10%) uncertainty for the inclusive cross-section (a) and for the p_T of the leading jet (b), and as an uncertainty on a ratio of 1 for the inclusive cross-section (c) and for the p_T of the leading jet (d), in the $Z \rightarrow \mu^+\mu^-$+jets channel. The systematic uncertainty on the inclusive cross-section for $\geq$4jets from a jet energy scale uncertainty of 5% is $\approx$ 15% (for the first fb^{-1}of data). This value decreases to less than 5% for a jet energy scale uncertainty of 1%.

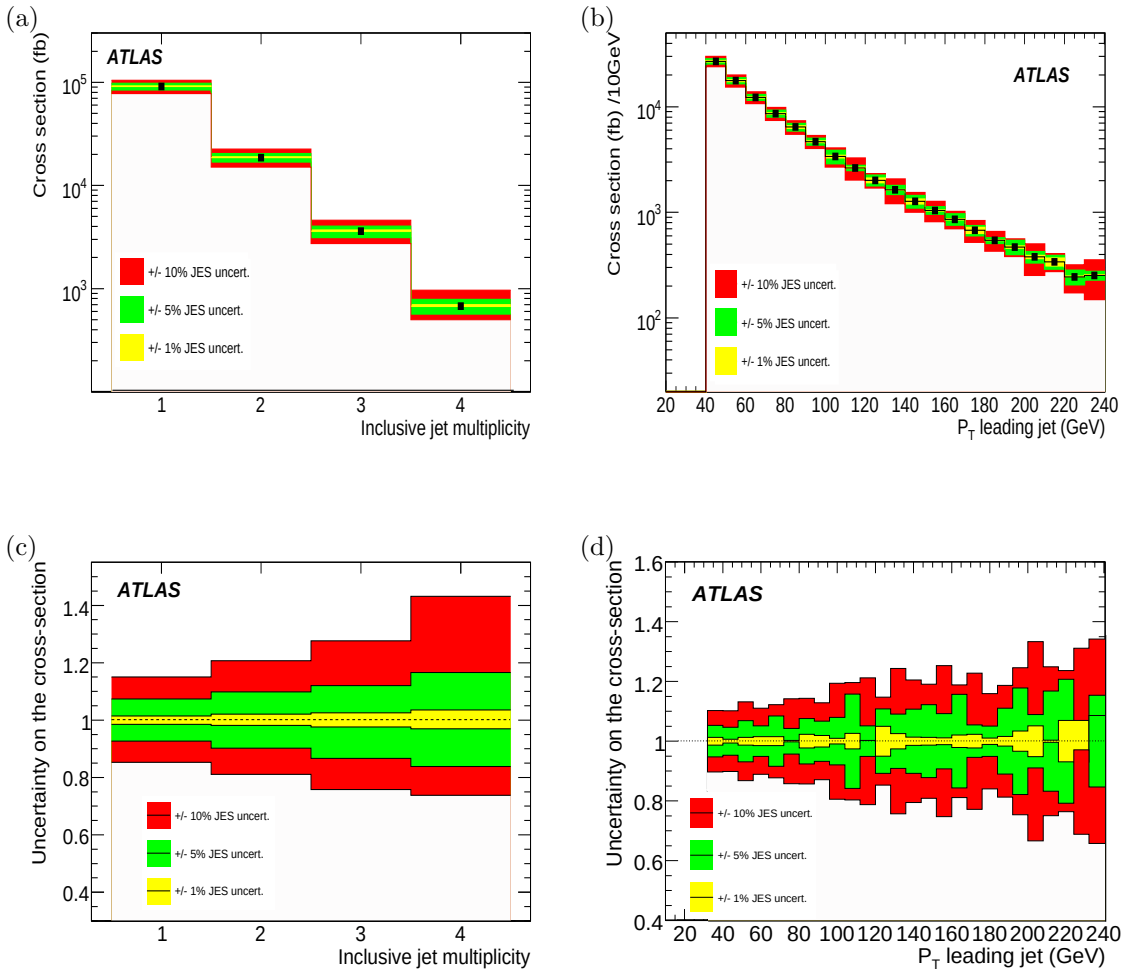


Figure 5.23: Uncertainties on $Z \rightarrow \mu^+\mu^-$ cross-sections from jet energy scale uncertainties of 1%, 5% and 10%.

Chapter 6

Z(ll)+jets Cross-Section Measurement: Results

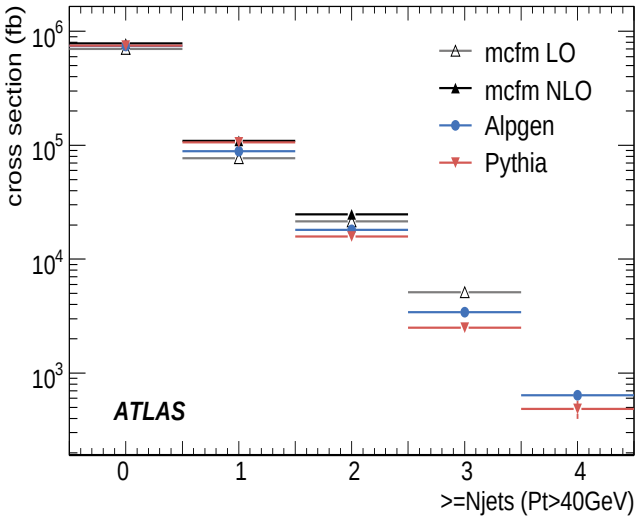
Finally, in this section we compare the results obtained in previous chapters 4 and 5. Theoretical evaluations on the cross sections from parton level corrected to hadron level, and measured cross sections corrected back to hadron level are now compared. We include systematics uncertainties coming from both sides of the analysis. Results for the first fb^{-1} of data are discussed.

6.1 Cross-Section Measurements Results

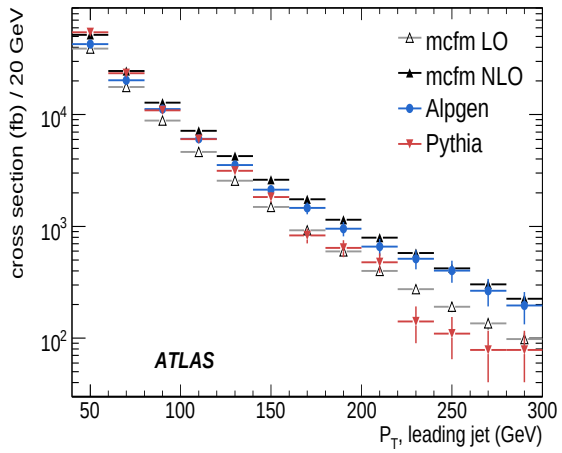
In this analysis, cross section measurements for theoretical well-defined quantities are considered. The cross section for Z+jets production is quoted at hadron level, corrected for all detector measurements effects. An unfolding technique, from the detector to the hadron level, which can be used with real data as well as with fully simulated Monte Carlo data has been developed and applied. The corrections from parton to hadron level have been also determined, taking into account non-perturbative effects (underlying event and jet fragmentation), resulting in corrections in opposite directions that partially cancel out. The resulting Z+jets cross section comparisons between theoretical NLO predictions with those obtained from reconstructed Monte Carlo data are shown in figure 6.1 for $Z \rightarrow \mu^+\mu^-$ and in figure 6.2 for $Z \rightarrow e^+e^-$. The two Monte Carlo samples are normalized to the inclusive NLO cross section, as determined by MCFM.

Figure 6.1(a) shows the comparison of the inclusive jet cross section for $Z \rightarrow \mu^+\mu^-$ events for ALPGEN and PYTHIA Monte Carlos with LO and NLO MCFM predictions for 1 fb^{-1} of data. Figures 6.1(b), (c) show the differential cross section for the p_T of the leading and second leading jet at hadron level for ALPGEN and PYTHIA Monte Carlos. Inclusive jet cross section from ALPGEN Monte Carlo data agrees well with NLO MCFM predictions for the leading jet, and the shape of the p_T distributions of the leading jet and the second leading jet predicted by ALPGEN agrees well with NLO MCFM predictions. Due to the tuning of the leading soft radiation in the parton shower, PYTHIA predicts a larger inclusive cross-section for $Z \rightarrow \mu^+\mu^- + \geq 1$ jet but a clearly softer p_T spectrum for high-p_T values ($p_T > 150$ GeV), which lines up with LO MCFM predictions for the leading jet.

(a)



(b)



(c)

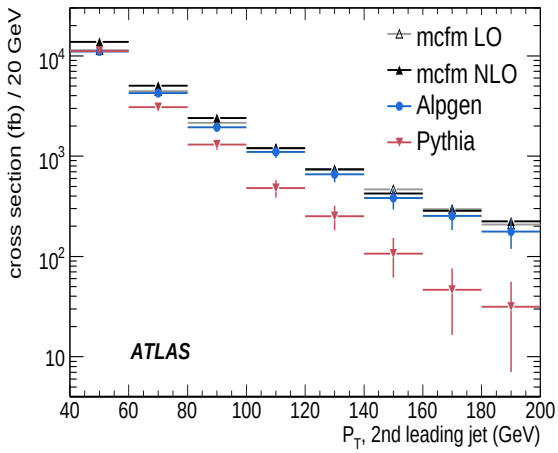


Figure 6.1: $Z \rightarrow \mu^+\mu^-$ final cross section (a) and $Z \rightarrow \mu^+\mu^-$ final cross sections as a function of the p_T of the leading jet (b) and the p_T of the second leading jet (c). In summary, NLO (LO) MCFM predictions are corrected for non-perturbative effects and ALPGEN and PYTHIA reconstructed results are unfolded to hadron level. Only statistical uncertainties are shown.

Figure 6.2 shows, for $Z \rightarrow e^+e^-$ events, the comparison of the inclusive jet cross section (a) and the differential cross section for the p_T of the leading jet (b) at the hadron level for ALPGEN Monte Carlo with LO and NLO MCFM calculations for 1 fb^{-1} of data. The error bars are calculated only from intrinsic Monte Carlo quantities as the quadratic sum of statistical errors from the Monte Carlo sample size and the systematic errors from the unfolding corrections derived from Monte Carlo. As in $Z \rightarrow \mu^+\mu^-$ events, the shape of the jet p_T distribution predicted by ALPGEN agrees well with the shape predicted by NLO MCFM. Inclusive jet cross section by ALPGEN agrees well with NLO MCFM predictions up to $Z+ \geq 2$ jets.

(a) (b)

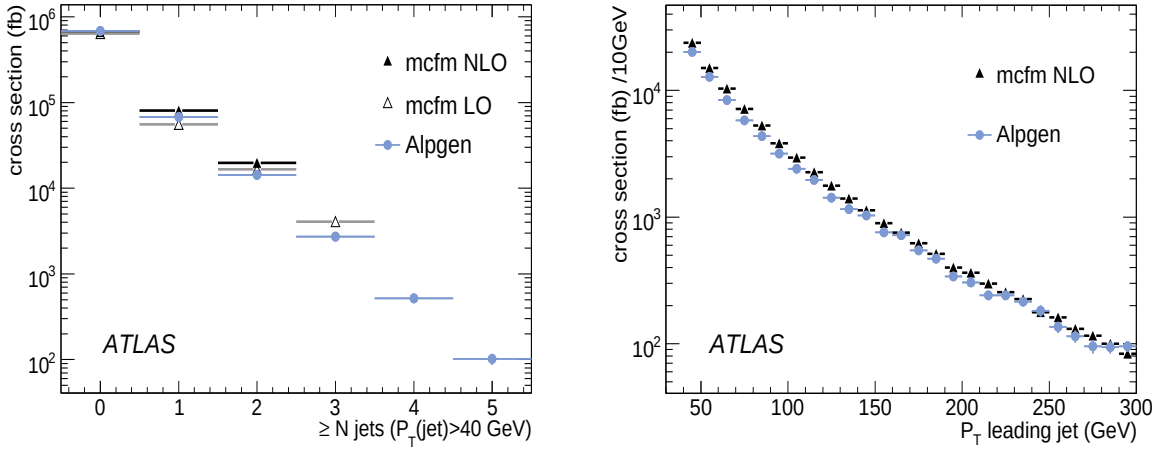


Figure 6.2: $Z \rightarrow e^+e^-$ final cross sections, comparing MCFM predictions with Monte Carlo data. Only statistical uncertainties are shown.

Looking at Z+jets cross sections, we can discriminate between LO and NLO predictions and we might be able to discriminate between ALPGEN and PYTHIA predictions for the leading jet p_T spectrum, for the first fb^{-1} of data, if statistics is the major uncertainty in results (more in section 6.4).

6.2 Statistical and Systematic Uncertainties

Figure 6.3(a) shows statistical and systematic uncertainties on $Z \rightarrow e^+e^- +$jets inclusive cross sections. With an error of 5% (10%) on the JES we expect an uncertainty of 5-15% (15-35%) for the different jet multiplicities (N$\leq$ 1-4 jets), on the cross-section measurements for the first fb^{-1}of data. Figure 6.3(b) shows statistical and systematic uncertainties on $Z \rightarrow \mu^+\mu^-$ + jets inclusive cross-section. With an error of 5% (10%) on the JES we expect an uncertainty of 10-20% (20-40%) for the different jet multiplicities (N$\leq$ 1-4 jets), on the cross-section measurements for the first fb^{-1}of data. JES error is the largest contribution to the uncertainty in the Z+jets cross section. Table 6.1 contains a summary of the different uncertainties on the $Z \rightarrow e^+e^- +\geq$ N jets and on the $Z \rightarrow \mu^+\mu^- +\geq$ N jets cross-sections.

QCD dijet background assumes 20% uncertainty ($Z \rightarrow e^+e^-$ analysis), which represents 2-3 % on the inclusive $Z \rightarrow e^+e^- +\geq$ N jet cross section. For the $Z \rightarrow \mu^+\mu^-$ analysis, a 50% uncertainty for QCD dijet background is obtained, from QCD background estimation (see section 5.4.1), which represents between $\sim$ 2.0% - 5.0% on the inclusive $Z \rightarrow \mu^+\mu^- +\geq$N jet cross section. Top background can be assumed to have a 15% uncertainty and represents 1-2 % uncertainty for $Z \rightarrow e^+e^-$ cross section. This 15% uncertainty is in agreement with ATLAS top quark pair cross-section analysis [89]. If we consider a 15% uncertainty in top background for $Z \rightarrow \mu^+\mu^- +\geq$1, 2 jet events, a $\sim$ 1 % uncertainty on the $Z \rightarrow \mu^+\mu^- + \geq$ N jets cross section is obtained. In the case of $\geq$3, 4 jets, a 30% uncertainty for top estimation is assumed, which represents 1-2% uncertainty on $Z \rightarrow \mu^+\mu^- +\geq$ N jets cross section. Therefore, background uncertainties do not represent an important part of the uncertainties of the resulting cross sections (less than 5%).

Statistical error for signal increases from 1% to 7% as jet multiplicity increases. Thus, only at high jet multiplicities statistical error becomes relevant. PDF uncertainty remains approximately constant as jet multiplicity increases and represents less $<$ 5% for the inclusive Z+jet cross section. Integrated luminosity uncertainty is not considered in this study, but it is not necessary since we normalize our measurements to the inclusive Z cross section.

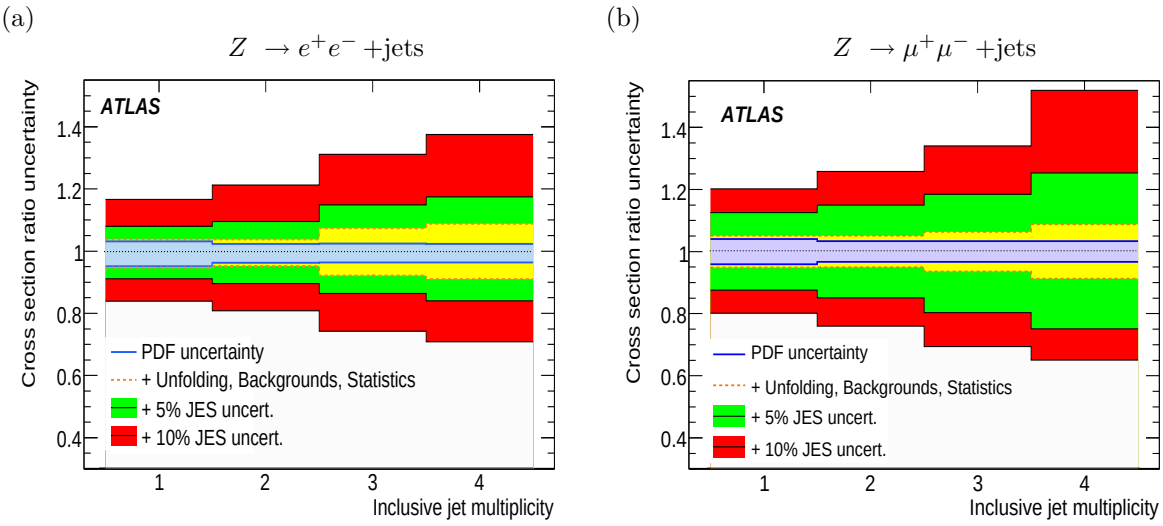


Figure 6.3: Statistical and systematic uncertainties for $Z \rightarrow e^+e^-$ channel (a) and for $Z \rightarrow \mu^+\mu^-$ channel (b) of the inclusive jet cross section.

Uncertainty	$Z \rightarrow e^+e^-$+jets	$Z \rightarrow \mu^+\mu^-$+jets
PDF	$< 5\%$	$< 5\%$
5% (10%) JES	5-15 (15-35%)	10-20 (20-40%)
QCD dijet background	2-3%	2-5%
Top background	1-2%	1-2%
Statistical	1-7%	1-7 %
Total	15-40%	20-50%

Table 6.1: Summary of the different uncertainties on the cross-sections (in fb) for $Z \rightarrow e^+e^-$ + N jets and $Z \rightarrow \mu^+\mu^-$ + N jets events in percentage level, for the first 1 fb^{-1}of data.

In summary, for the first fb^{-1} of data, Z+jet inclusive jet cross section can be determined up to 15-20% for $Z+\geq$ 1jet, 20-25% for $Z+ \geq$ 2jets, 30-35 % for $Z+\geq$ 3jets and 40-50 % for $Z+\geq$ 4jets.

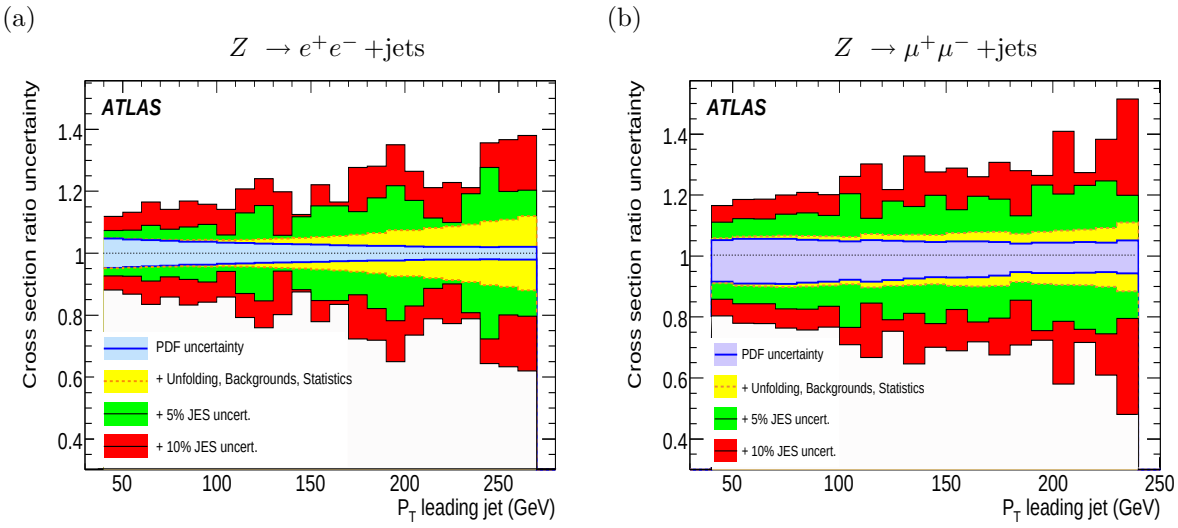


Figure 6.4: Statistical and systematic uncertainties for $Z \rightarrow e^+e^-$ channel (a) and $Z \rightarrow \mu^+\mu^-$ channel (b) for the differential cross section as a function of the p_T of the leading jet.

Figures 6.4 show the statistical and systematic uncertainties for the differential cross section as a function of p_T of the leading jet, for $Z \rightarrow e^+e^-$ (a) and for $Z \rightarrow \mu^+\mu^-$ (b) events. PDF uncertainty decreases as p_T increases, from 5% to 2-3%. On the other hand, statistical uncertainty is relevant for high-p_T jets only (10-5%). JES uncertainty also increases with p_T of the leading jet, from 5 to 10% for 5% JES uncertainty, and from 5 to 15-20% for 10% JES uncertainty. So, the uncertainty for low-p_T jets (p_T <100 GeV) is ∼4 (∼12)% for $Z \rightarrow e^+e^-$ +jets and ∼6 (∼12)% for $Z \rightarrow \mu^+\mu^-$ +jets (depending on the JES uncertainty assumed 5 or (10) %). The uncertainty for high-p_T jets (p_T > 180 GeV) is ∼10 (15-22)% for $Z \rightarrow e^+e^-$ +jets and 15-25 (30-45)% for $Z \rightarrow \mu^+\mu^-$ +jets. A summary of the different uncertainties for the differential cross section split in low and high p_T regions are shown in table 6.2.

Uncertainty	$Z \rightarrow e^+e^-$+ jets		$Z \rightarrow \mu^+\mu^-$+jets	
	low-p_T	high-p_T	low-p_T	high-p_T
PDF	5%	2-3%	5%	2-3%
Statistical	-	5-10%	-	5-10%
JES 5% (10%)	2-5%(8-14%)	∼10%(15-22%)	5-8%(10-14%)	6-16%(20-30%)
Total	<8% (∼15%)	15-20 (25-38)%	∼12 (20)%	15-25 (30-45)%

Table 6.2: Summary of the different uncertainties on the cross-sections (in fb) as a function of the p_T of the leading jet for $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ processes in percentage level, for the first 1 fb^{-1}of data. Low-p_T jet values mean $p_T < 100$ GeV and high-p_T corresponds to $p_T > 180$ GeV.

6.3 Comparison with Tevatron Results

In this section, a comparison of the main results and systematic uncertainties of the different results coming from Tevatron experiments, CDF [90] and D0 [78] is compiled. Both experiments have published analysis of $Z(e^+, e^-)$+jets cross-section measurements. More details are given in appendix E. At the Tevatron collider [91], protons and antiprotons collide at a center of mass energy of 1.96 TeV, while at the LHC protons and protons will collide at 14 TeV.

CDF analyses have studied the inclusive $Z/\gamma^*(\rightarrow e^+e^-)$+N jets cross section up to N≥ 3 jets. D0 analyses have looked at the inclusive $Z/\gamma^*(\rightarrow e^+e^-)$+N jets cross section up to N≥ 4 jets. Our ATLAS analysis has studied cross section up to N≥ 4 jets for both channels, $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$. All studies have looked at the jet cross section as a function of the p_T of the leading jet and of the second leading jet, and D0 analyses even for the third leading jet. CDF has used 2.5 fb^{-1} of data, while D0 studies for 0.4 fb^{-1} and 1.04 fb^{-1} of data. ATLAS Monte Carlo data have been scaled to 1 fb^{-1}of data.

In all analyses, parton level predictions are provided by the MCFM program, using CTEQ6.1M PDFs and setting factorization and renormalization scale to $\mu_0^2 = M_Z^2 + p_{T,Z}^2$. Uncertainties due to dependence on μ scale have been studied ($\mu = \mu_0/2$ and $\mu = \mu_0 \times 2$) in CDF and D0 analyses, leading to 10 - 15% uncertainty for the different jet multiplicities (CDF analyses), and 10 % uncertainty on the cross section for the leading jet and for the second leading jet (D0 analyses). PDF uncertainties varies from 2 % at low p_T^{jet} to 10% at high p_T^{jet} (CDF) and 5-10 % for the first and second leading jet and 5-15% for the third leading jet (D0). In our analysis, PDF uncertainty represents 5 % uncertainty on the jet cross section at low p_T and 3 % uncertainty at high p_T values.

In CDF analyses, QCD backgrounds are estimated using data-driven techniques (including QCD and W+jets processes) representing 7 % for Z +≥ 1 jet and 8% for Z+ ≥ 2 jets. Other background events are estimated using Monte Carlo PYTHIA generator ($t\bar{t}$, $Z/\gamma^*(\rightarrow \tau\tau)$+jets, WW, WZ, ZZ, $Z/\gamma^*(\rightarrow e^+e^-) + \gamma$), representing 4 % for ≥ 1 jet and 6 % for ≥ 2 jets. QCD backgrounds have an uncertainty of 15 % and leads to 1-2 % uncertainty on the jet cross section. Non-QCD backgrounds have an uncertainty of 30 %, which translates to 1 % uncertainty on the jet cross section. In the case of D0 analyses, QCD backgrounds (multi-jet and W+jets events) represents less than 1 % and non-QCD backgrounds less than 6%. In our analysis, backgrounds represent less than 10% for Z+≥ 1, 2 jets and less than 15-20% for Z+≥ 3, 4 jets. With an uncertainty of 50 % on the QCD estimation and 15% on the top background estimation, uncertainties on the cross section results 1-5 %, mostly due to QCD dijet events uncertainty. It is clear that in ATLAS future analyses, real data will be necessary to understand background processes using data-driven methods and reducing uncertainties on their estimation. On

the other hand, final uncertainties on the cross section are rather similar between ATLAS predictions and CDF and D0 results, although smaller relative errors on background are obtained by CDF/D0. In addition, in our analysis a smaller invariant mass window $M_{l,l}$ has been used in order to reduce the background level.

The measured cross sections as a function of the inclusive jet multiplicity up to $N_{jet} \geq 3$ jets are compared to MCFM predictions in all three studies. Tevatron results show that the LO pQCD predictions underestimate the measured cross sections, while good agreement is observed between data and NLO pQCD predictions.

In D0 analysis, for the leading jet, PYTHIA shows a more steeply falling spectrum than observed in data. Our analysis also shows the same behavior of PYTHIA data compared with NLO MCFM prediction. Comparisons of sub-leading jets show that PYTHIA predict more steeply falling p_T^{jet} spectra than observed in data, in agreement with expectations based on the limited validity of the soft/collinear approximations of their parton shower. In this case, the same behavior is shown in our analysis, comparing PYTHIA data with NLO and LO predictions. Finally, D0 results show comparisons with the ALPGEN+PYTHIA event generators. This generator combines tree-level matrix elements with parton showers, thereby utilizing matrix elements also for sub-leading jets. For all three p_T^{jet} spectra, ALPGEN+PYTHIA predicts lower productions rates than observed in data, but the shapes of the spectra are well described. In our analysis, ALPGEN+HERWIG generator predicts the same p_T shape for the leading and second leading jet as NLO pQCD predictions.

One of the most important source of systematic uncertainty in all three cases is the Jet Energy Scale. In CDF analyses, JES is known to 3%, but uncertainties on the jet cross section represents 2 % at low p_T^{jet} ($\approx$ 30 GeV) and 3 % at high p_T^{jet} ($\approx$ 400 GeV). In our analysis, considering an uncertainty of 5 % on the JES, uncertainties on the jet cross section represents a 3-5 % uncertainty at low jet p_T ($\approx$ 40 GeV) and 10-15 % at high jet p_T ($\approx$ 250 GeV). Reducing the JES uncertainty to less than 5 % level uncertainty will be a very important goal for ATLAS.

Finally, the total systematic uncertainty in the CDF analyses is 7-14 % for $Z/\gamma^*(\rightarrow e^+e^-)+ \geq$ 1jet and 10-15 % for $Z/\gamma^*(\rightarrow e^+e^-)+ \geq 2$ jets. In our analysis, 15 % for $Z(\rightarrow e^+e^-)+ \geq$ 1jet, 20 % for $Z(\mu^+\mu^-)+ \geq$ 1jet, 20 % for $Z(\rightarrow e^+e^-)+ \geq 2$ jets and 25% for $Z(\mu^+\mu^-)+ \geq 2$ jets. A better understanding of the ATLAS performance will help to reduce systematic uncertainties. New more accurate theoretical parton predictions and better knowledge of the PDFs will also help to reduce uncertainties but in a longer term.

6.4 Summary of Final Results

In this section, numeric results for the cross-section measurement are summarized in different tables and figures. The final cross-section measurements for $Z \rightarrow e^+e^-+$ jets events are shown in table 6.3, with the corresponding statistical and systematic uncertainties. The final cross-section measurements for $Z \rightarrow \mu^+\mu^-$ +jets are shown in table 6.4, with the corresponding statistical and systematic uncertainties. Figure 6.5 reproduces these final results, for $Z \rightarrow e^+e^-$ in (a) and for $Z \rightarrow \mu^+\mu^-$ in (b). Only systematic errors are shown, since they represent the major error (statistic errors are very small for all the cases, as seen previously in figure 6.1(a) and 6.2(a)).

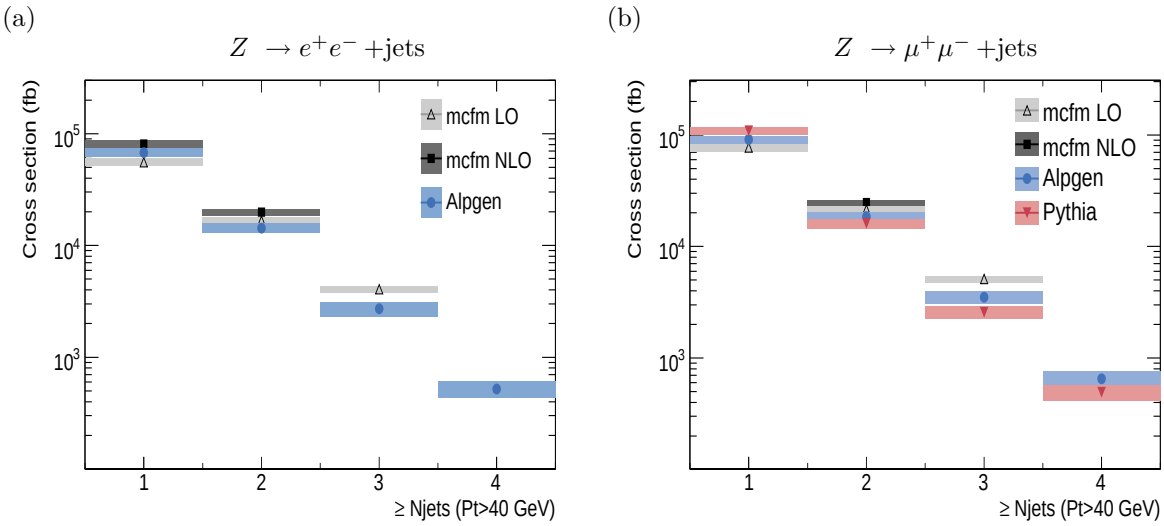


Figure 6.5: Cross sections for $Z \rightarrow e^+e^-$+jets (a) and for $Z \rightarrow \mu^+\mu^-$+jets (b), with main errors displayed (PDF uncertainty for NLO and LO predictions and 5% JES on ALPGEN and PYTHIA Monte Carlo data). For $Z \rightarrow \mu^+\mu^-$+ 1 jet, PYTHIA data and MCFM NLO corrected prediction coincides.

Process	Cross Section (fb)	± (stat.uncert.)	± (syst.uncert.)
$Z+ \geq$ 1jet			
MCFM NLO	80806.9	± 96.5	± 6428.3
MCFM LO	55563.1	± 44.4	± 4420.2
ALPGEN	67825.9	± 470.7	± 6244.4 (11106.6)
$Z+ \geq$ 2jet			
MCFM NLO	19805.8	± 254.0	± 1205.2
MCFM LO	16596.3	± 28.9	± 1009.9
ALPGEN	14325.2	± 209.0	± 1425.7 (2892.0)
$Z+ \geq$ 3jet			
MCFM NLO	–	–	–
MCFM LO	4057.5	± 18.8	± 246.9
ALPGEN	2723.2	± 135.5	± 387.8 (775.7)
$Z+ \geq$ 4jet			
MCFM NLO	–	–	–
MCFM LO	–	–	–
ALPGEN	519.3	± 33.4	± 86.7 (173.0)

Table 6.3: Cross-sections (in fb) and the corresponding uncertainties for $Z \rightarrow e^+e^-$ events comparing MCFM NLO and LO predictions, at hadron level, with ALPGEN and PYTHIA reconstructed data corrected to hadron level for 1 fb^{-1} of data. Systematic uncertainties include PDF uncertainties in the case of MCFM predictions; and unfolding, background estimation and JES (5% and 10%) uncertainties in the case of MC predictions. These results are shown in figures 6.2(a), 6.3(a) and 6.5(a).

Process	Cross Section (fb)	± (stat.uncert.)	± (syst.uncert.)
$Z+ \geq$ 1jet			
MCFM NLO	109307.2	± 107.3	± 5647.4
MCFM LO	77208.1	± 31.2	± 5647.4
ALPGEN	90700.8	± 955.1	± 6704.5 (13520.5)
PYTHIA	108614.1	± 1337.2	± 8028.6 (16190.7)
$Z+ \geq$ 2jet			
MCFM NLO	24828.57	± 107.9	± 1366.4
MCFM LO	21429.55	± 32.6	± 1366.4
ALPGEN	18474.53	± 320.4	± 1838.6 (3668.2)
PYTHIA	16115.97	± 515.1	± 1603.9 (3199.9)
$Z+ \geq$ 3jet			
MCFM NLO	–	–	–
MCFM LO	5109.9	± 17.9	± 255.3
ALPGEN	3511.0	± 107.0	± 450.9 (912.5)
PYTHIA	2568.0	± 205.6	± 329.8 (667.4)
$Z+ \geq$ 4jet			
MCFM NLO	–	–	–
MCFM LO	–	–	–
ALPGEN	652.5	± 35.3	± 107.4 (226.6)
PYTHIA	493.9	± 90.2	± 76.9 (171.5)

Table 6.4: Cross-sections (in fb) and the corresponding uncertainties for $Z \rightarrow \mu^+\mu^-$ events comparing MCFM NLO and LO predictions, at hadron level, with ALPGEN and PYTHIA reconstructed data corrected to hadron level, for 1 fb^{-1} of data. Systematic uncertainties include PDF uncertainties in the case of MCFM predictions; and unfolding, background estimation and JES (5% and 10%) uncertainties in the case of MC predictions. These results are shown in figures 6.1(a), 6.3(b) and 6.5(b).

The differential cross-section as a function of the p_T of the leading jet for $Z \rightarrow e^+e^-$+ jets events is shown in figure 6.6(a). Clearly, MCFM NLO predictions are in agreement with these from ALPGEN Monte Carlo, within the systematic and statistical uncertainty. Table 6.5 contains the different results for NLO predictions and ALPGEN Monte Carlo data for $Z \rightarrow e^+e^-$+jets events, with statistical and systematic uncertainties, for 1 fb^{-1} of data.

Differential cross-section as a function of the p_T of the leading jet for $Z \rightarrow \mu^+\mu^-$+jets events is shown in figure 6.6(b). The figure includes systematic uncertainties (not statistical uncertainties). Table 6.6 summarizes results for NLO predictions and ALPGEN Monte Carlo data, including statistical and systematic uncertainties. Table 6.7 summarizes results for LO predictions and PYTHIA Monte Carlo data. With the first fb^{-1} of data, we will be able to distinguish between NLO and ALPGEN with LO and PYTHIA from $p_T^{lead\ jet} > 150$ GeV , as shown in figure 6.6(b).

(a) (b)

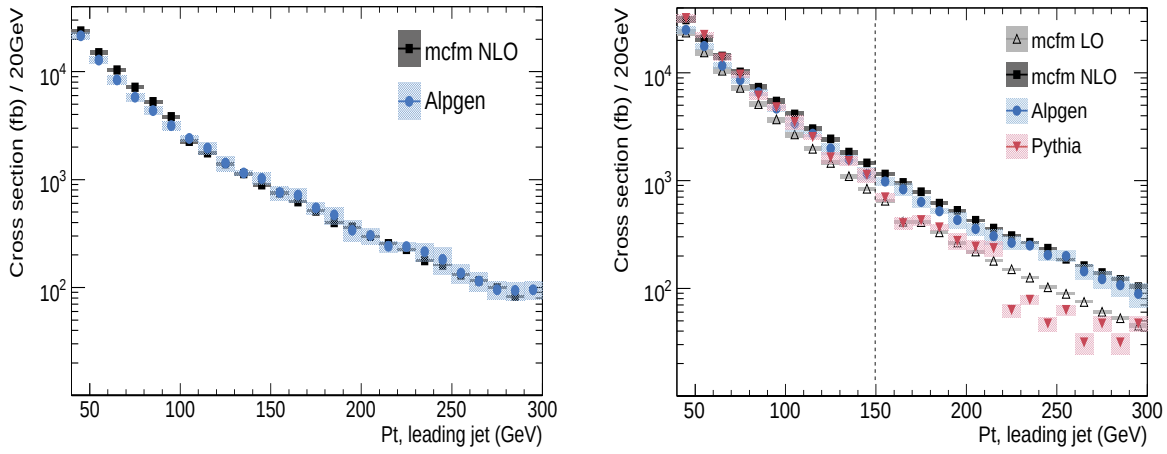


Figure 6.6: Differential cross sections as a function of the p_T of the leading jet for $Z \rightarrow e^+e^-$+jets (a) and for $Z \rightarrow \mu^+\mu^-$+jets, with statistical and systematic errors, for NLO and LO predictions and ALPGEN and PYTHIA Monte Carlo data, for the first fb^{-1} of data.

p_T leading jet, $Z \rightarrow e^+e^-$+jets									
p_T range	MCFM NLO					ALPGEN			
(GeV)	σ (fb)		± (stat.)		± (syst.)	σ (fb)		± (stat.)	± (syst.)
40 - 50	23745.0	±	7.9	±	1121.7	21445.6	±	1055.0	±1576.0 (2549.0)
50 - 60	15020.4	±	6.5	±	672.8	12789.8	±	780.2	± 947.0 (1689.1)
60 - 70	10352.1	±	5.4	±	440.6	8387.8	±	677.2	± 752.6 (1384.7)
70 - 80	7163.3	±	4.3	±	286.0	5807.4	±	392.6	± 447.1 (824.5)
80 - 90	5268.8	±	3.6	±	197.4	4348.0	±	337.0	± 368.1 (729.9)
90 - 100	3815.1	±	2.9	±	140.4	3165.4	±	274.6	± 293.7 (501.7)
100 - 110	2955.3	±	2.4	±	101.8	2411.2	±	120.8	± 142.0 (342.3)
110 - 120	2258.9	±	2.0	±	73.9	1964.6	±	250.5	± 256.6 (408.7)
120 - 130	1770.2	±	1.7	±	55.2	1421.1	±	218.9	± 219.1 (342.5)
130 - 140	1400.6	±	1.4	±	42.4	1154.5	±	57.5	± 67.3 (228.7)
140 - 150	1132.7	±	1.2	±	32.3	1027.2	±	116.8	± 121.4 (128.9)
150 - 160	894.9	±	1.0	±	24.7	756.9	±	113.1	± 116.0 (167.4)
160 - 170	754.4	±	0.9	±	19.3	719.2	±	78.8	± 109.8 (118.7)
170 - 180	622.4	±	0.7	±	15.3	545.1	±	70.9	± 73.3 (151.2)
180 - 190	512.0	±	0.6	±	12.2	469.6	±	81.4	± 84.0 (131.8)
190 - 200	398.8	±	0.5	±	9.4	339.4	±	72.6	± 74.0 (118.8)
200 - 210	362.5	±	0.5	±	8.0	304.7	±	50.3	± 51.9 (80.5)
210 - 220	298.5	±	0.4	±	6.3	240.5	±	26.0	± 27.3 (50.8)
220 - 230	255.5	±	0.4	±	5.2	240.2	±	19.8	± 23.9 (54.7)
230 - 240	224.6	±	0.3	±	4.7	214.6	±	39.5	± 41.1 (45.3)
240 - 250	177.3	±	0.3	±	3.5	182.1	±	47.9	± 50.4 (64.9)
250 - 260	161.4	±	0.3	±	3.3	135.5	±	26.0	± 27.0 (49.7)
260 - 270	130.7	±	0.2	±	2.7	114.3	±	22.2	± 23.2 (43.4)
270 - 280	115.6	±	0.2	±	2.4	95.2	±	17.8	± 18.5 (27.7)
280 - 290	99.6	±	0.2	±	1.9	93.9	±	16.6	± 18.4 (29.5)
290 - 300	83.0	±	0.2	±	1.6	95.3	±	14.3	± 17.9 (26.7)

Table 6.5: Cross section as a function of the p_T of the leading jet comparing NLO prediction (corrected to hadron level) with ALPGEN Monte Carlo data (unfolded to hadron level) for $Z \rightarrow e^+e^-$ events. These results are shown in figure 6.2(b),figure 6.4(a) and figure 6.6(a).

p_T leading jet, $Z \rightarrow \mu^+\mu^-$+jets									
p_T range	MCFM NLO					ALPGEN			
(GeV)	σ (fb)	±	(stat.)	±	(syst.)	σ (fb)	±	(stat.)	± (syst.)
40 - 50	30964.2	±	9.5	±	1814.0	25030.0	±	719.9	±1963.4 (2893.2)
50 - 60	20776.2	±	8.2	±	1172.2	17678.9	±	605.0	±1049.1 (2165.8)
60 - 70	14319.1	±	6.7	±	773.6	11692.1	±	492.0	± 703.1 (1493.2)
70 - 80	10187.0	±	5.4	±	524.1	8598.3	±	417.6	± 609.9 (1154.5)
80 - 90	7349.5	±	4.4	±	361.5	6578.3	±	365.3	± 484.2 (919.9)
90 - 100	5470.9	±	3.6	±	256.6	4651.9	±	307.8	± 319.2 (637.3)
100 - 110	4150.7	±	3.0	±	182.7	3350.9	±	261.2	± 478.0 (671.8)
110 - 120	3033.9	±	2.4	±	128.0	2725.2	±	235.6	± 136.3 (608.5)
120 - 130	2435.3	±	2.1	±	96.2	1981.3	±	200.9	± 223.3 (300.3)
130 - 140	1835.7	±	1.7	±	69.2	1552.8	±	180.9	± 150.3 (422.0)
140 - 150	1454.6	±	1.4	±	51.7	1153.4	±	155.9	± 164.3 (262.1)
150 - 160	1153.5	±	1.2	±	40.5	979.0	±	143.6	± 80.2 (221.7)
160 - 170	960.0	±	1.0	±	29.4	827.9	±	132.1	± 100.8 (156.2)
170 - 180	789.5	±	0.9	±	23.1	633.1	±	114.1	± 71.9 (153.9)
180 - 190	618.1	±	0.7	±	18.5	520.7	±	103.4	± 32.1 (112.3)
190 - 200	528.9	±	0.6	±	15.0	431.9	±	94.2	± 72.1 (86.7)
200 - 210	429.5	±	0.5	±	10.9	355.5	±	85.5	± 45.4 (122.9)
210 - 220	362.1	±	0.5	±	8.4	303.8	±	77.9	± 49.3 (63.8)
220 - 230	309.4	±	0.4	±	7.3	265.3	±	72.8	± 37.8 (71.1)
230 - 240	268.0	±	0.4	±	6.0	249.3	±	70.6	± 22.4 (101.9)
240 - 250	236.0	±	0.3	±	5.2	203.5	±	68.8	± 22.4 (101.9)
250 - 260	186.1	±	0.3	±	4.0	199.7	±	63.2	± 22.4 (101.9)
260 - 270	162.3	±	0.3	±	3.5	143.7	±	53.6	± 22.4 (101.9)
270 - 280	140.0	±	0.2	±	3.5	121.7	±	49.3	± 22.4 (101.9)
280 - 290	121.2	±	0.2	±	3.5	106.8	±	46.2	± 22.4 (101.9)
290 - 300	103.5	±	0.2	±	3.5	89.1	±	42.2	± 22.4 (101.9)

Table 6.6: Cross-section as a function of the p_T of the leading jet, comparing NLO predictions (corrected to hadron level) with ALPGEN reconstructed corrected data (unfolded to hadron level), for $Z \rightarrow \mu^+\mu^-$ events. Statistical and systematics errors are shown for all the cases. Systematic errors include unfolding, background and JES (5% and 10%) uncertainties. These results are shown in figure 6.1(b), figure 6.4(b) and figure 6.6(b).

p_T leading jet, $Z \rightarrow \mu^+\mu^-$+jets									
p_T range	MCFM LO					PYTHIA			
(GeV)	σ (fb)		± (stat.)		± (syst.)	σ (fb)		± (stat.)	± (syst.)
40 - 50	23584.1	±	31.7	±	1663.0	32187.0	±	784.2	±2524.8 (3720.5)
50 - 60	15447.5	±	34.6	±	1049.9	22501.7	±	655.7	±1335.3 (2756.6)
60 - 70	10470.2	±	35.1	±	671.2	13867.5	±	514.8	± 833.9 (1771.0)
70 - 80	7254.3	±	34.6	±	440.9	9627.2	±	424.5	± 609.9 (1292.7)
80 - 90	5125.4	±	33.5	±	297.0	6141.7	±	339.1	± 452.1 (858.8)
90 - 100	3694.5	±	31.8	±	203.7	4769.2	±	299.4	± 327.2 (653.4)
100 - 110	2679.5	±	30.5	±	138.5	3488.2	±	256.1	± 497.6 (699.5)
110 - 120	1973.4	±	28.3	±	97.8	2531.2	±	218.1	± 126.6 (565.2)
120 - 130	1459.1	±	26.7	±	67.8	1632.5	±	175.2	± 184.0 (247.4)
130 - 140	1099.6	±	23.9	±	48.8	1517.2	±	171.8	± 146.9 (412.3)
140 - 150	841.4	±	26.4	±	35.3	1130.5	±	148.3	± 161.0 (256.9)
150 - 160	648.1	±	22.5	±	27.0	699.1	±	116.6	± 57.3 (158.3)
160 - 170	413.4	±	16.3	±	15.1	401.5	±	88.4	± 48.9 (75.6)
170 - 180	412.6	±	13.7	±	14.5	426.9	±	90.0	± 48.5 (103.8)
180 - 190	331.8	±	25.3	±	12.0	365.9	±	83.3	± 22.6 (78.9)
190 - 200	265.5	±	26.1	±	9.2	274.5	±	72.1	± 45.8 (55.1)
200 - 210	220.0	±	23.8	±	6.9	243.9	±	68.0	± 31.1 (84.4)
210 - 220	179.7	±	21.1	±	5.3	235.0	±	65.9	± 38.1 (49.4)
220 - 230	150.0	±	19.3	±	4.5	62.7	±	34.0	± 8.9 (16.8)
230 - 240	125.4	±	17.3	±	3.6	78.3	±	38.0	± 7.0 (32.0)
240 - 250	102.4	±	15.3	±	3.0	47.0	±	29.5	± 7.0 (32.0)
250 - 260	89.1	±	14.1	±	2.6	62.8	±	34.0	± 7.0 (32.0)
260 - 270	74.9	±	12.7	±	2.2	31.4	±	24.1	± 7.0 (32.0)
270 - 280	60.5	±	11.4	±	2.2	47.1	±	29.5	± 7.0 (32.0)
280 - 290	52.9	±	10.4	±	2.2	31.4	±	24.1	± 7.0 (32.0)
290 - 300	45.1	±	9.4	±	2.2	47.1	±	29.5	± 7.0 (32.0)

Table 6.7: Cross-section as a function of the p_T leading jet, comparing LO predictions (corrected to hadron level) with PYTHIA reconstructed corrected data, for $Z \rightarrow \mu^+\mu^-$ events. Statistical and systematics errors are shown for all the cases. Systematic errors include unfolding, background and JES (5% and 10%) uncertainties. These results are shown in figure 6.1(b), figure 6.4(b) and figure 6.6(b).

Finally, the differential cross-section as a function of the p_T of the second leading jet is shown in figure 6.7 for $Z \rightarrow \mu^+\mu^-$+jets events. NLO and LO predictions are compared to ALPGEN and PYTHIA Monte Carlo data, at hadron level, for the first fb^{-1} of data. Only systematic error is shown in the figure. Figure 6.1(c) contains the statistical uncertainty, which increases for high-p_T jets. Tables 6.8 and 6.9 summarize the results, including statistical and systematic uncertainties.

LO and NLO predictions agree for the second lead jet p_T distribution. Also ALPGEN data agree well with theoretical predictions. PYTHIA data do not agree, and even considering their uncertainties (statistical uncertainty for high-p_T values is not negligible), PYTHIA can be distinguished from all previous predictions clearly for the first 1 fb^{-1} of data for all p_T range.

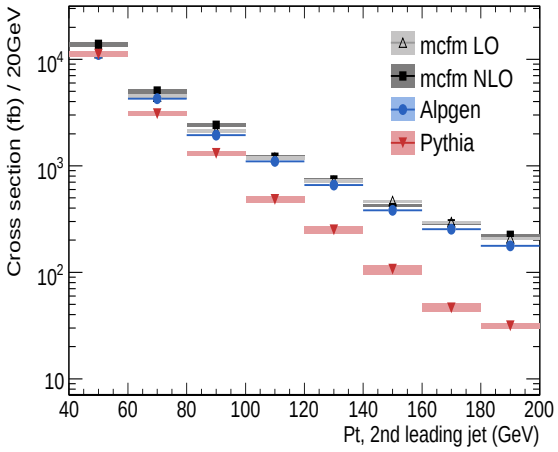


Figure 6.7: Differential cross sections as a function of the second leading jet for $Z \rightarrow \mu^+\mu^-$+jets, with systematic errors, for NLO and LO predictions and ALPGEN and PYTHIA Monte Carlo data, for the first fb^{-1} of data.

p_T 2nd leading jet, $Z \rightarrow \mu^+\mu^-$+jets										
p_T range	MCFM NLO					MCFM LO				
(GeV)	σ (fb)		±(stat.)		±(syst.)	σ (fb)		±(stat.)		±(syst.)
40 - 50	8744.6	±	1.7	±	616.6	7141.9	±	2.6	±	503.6
50 - 60	5044.4	±	1.2	±	339.9	4217.7	±	2.9	±	284.2
60 - 70	2924.6	±	0.8	±	187.5	2665.3	±	3.3	±	170.9
70 - 80	2101.1	±	0.7	±	125.6	1777.8	±	3.6	±	108.0
80 - 90	1396.4	±	0.5	±	80.9	1247.7	±	4.0	±	72.3
90 - 100	1010.1	±	0.4	±	55.7	893.9	±	4.4	±	49.3
100 - 110	667.7	±	0.3	±	34.5	678.8	±	4.8	±	35.1
110 - 120	532.9	±	0.3	±	26.4	514.3	±	5.2	±	25.5
120 - 130	445.1	±	0.3	±	20.7	403.4	±	5.8	±	18.7
130 - 140	293.2	±	0.2	±	13.0	318.9	±	6.2	±	14.1
140 - 150	281.7	±	0.2	±	11.8	257.1	±	6.5	±	10.8
150 - 160	143.6	±	0.1	±	6.0	208.7	±	7.1	±	8.7
160 - 170	176.9	±	0.1	±	6.5	158.8	±	7.0	±	5.9
170 - 180	107.7	±	0.1	±	3.8	136.9	±	7.6	±	4.8
180 - 190	180.9	±	0.2	±	6.6	109.2	±	7.8	±	4.0
190 - 200	42.2	±	0.1	±	1.5	98.9	±	8.7	±	3.4
200 - 210	58.5	±	0.1	±	1.8	93.8	±	11.6	±	2.9
210 - 220	52.2	±	0.1	±	1.5	70.4	±	9.4	±	2.1

Table 6.8: Cross-section as a function of the p_T of the second leading jet, comparing NLO predictions (corrected to hadron level) with LO predictions (corrected to hadron level) for $Z \rightarrow \mu^+\mu^-$ events, for 1 fb^{-1} of data. Statistical and systematics errors are shown for both predictions. These results are shown in figure 6.1(c) and figure 6.7.

p_T 2nd leading jet, $Z \rightarrow \mu^+\mu^-$+jets								
p_T range	ALPGEN				PYTHIA			
(GeV)	σ (fb)		±(stat.)	±(syst.)	σ (fb)		±(stat.)	±(syst.)
40 - 50	6810.7	±	362.7	±534.2 (787.2)	7304.6	±	373.6	±573.0 (844.3)
50 - 60	4245.2	±	286.4	±251.9 (520.1)	3848.8	±	271.2	±228.4 (471.5)
60 - 70	2495.7	±	219.6	±150.1 (318.7)	1968.2	±	193.9	±118.4 (251.4)
70 - 80	1763.4	±	182.7	±125.1 (236.8)	1104.6	±	143.8	± 78.4 (148.3)
80 - 90	1131.0	±	146.3	± 83.2 (158.2)	693.6	±	113.9	± 51.1 (97.0)
90 - 100	806.8	±	123.8	± 55.4 (110.5)	610.0	±	107.1	± 41.9 (83.6)
100 - 110	608.8	±	107.6	± 86.8 (122.0)	192.6	±	60.2	± 27.5 (38.6)
110 - 120	490.4	±	96.5	± 24.3 (109.5)	289.0	±	73.7	± 14.5 (64.5)
120 - 130	370.4	±	83.9	± 41.7 (56.1)	174.7	±	57.3	± 19.7 (26.5)
130 - 140	289.6	±	75.5	± 28.0 (78.7)	76.3	±	38.5	± 7.4 (20.7)
140 - 150	227.5	±	66.9	± 32.4 (51.7)	76.3	±	38.5	± 10.9 (17.3)
150 - 160	153.7	±	55.0	± 12.6 (34.8)	30.5	±	24.4	± 2.5 (6.9)
160 - 170	130.0	±	50.6	± 15.8 (24.5)	30.6	±	24.4	± 3.7 (5.8)
170 - 180	123.3	±	48.6	± 14.0 (30.0)	15.8	±	17.3	± 1.8 (3.8)
180 - 190	101.8	±	44.2	± 6.3 (22.0)	31.5	±	24.4	± 1.9 (6.8)
190 - 200	75.5	±	38.0	± 12.6 (15.1)	0.0	±	24.4	± 0.0 (0.0)
200 - 210	54.5	±	32.3	± 6.9 (18.8)	31.5	±	24.4	± 4.0 (10.9)
210 - 220	51.4	±	31.0	± 8.3 (10.8)	0.0	±	24.4	± 0.0 (0.0)

Table 6.9: Cross-section as a function of the p_T of the second leading jet, comparing ALPGEN and PYTHIA reconstructed corrected data, for $Z \rightarrow \mu^+\mu^-$ events. Statistical and systematics errors are shown for all the cases. Systematic errors include unfolding, background and JES (5% and 10%) uncertainties. These results are shown in figure 6.1(c) and figure 6.7.

Chapter 7

Summary and Conclusions

The production of jets in association with vector bosons, Z^0 or $W^{\pm}$, in hadron collision is an important process in quantum chromodynamics (QCD) and is a signature source of background for many Standard Model measurements (e.g., $t\bar{t}$ production) and in searches for new phenomena (e.g., supersymmetry). Such measurements at the Large Hadron Collider (LHC) rely on accurate descriptions of Boson+jets production by particle level event generators. These models require validation with measurements of the properties of the Boson+jets system, especially as a function of jet multiplicity.

This thesis reports on prospects for first measurements of the inclusive jet cross sections in $Z(\rightarrow e^+e^-/\ \rightarrow \mu^+\mu^-)$ + N jets (N$\leq$1, 2, 3, 4) production at the ATLAS experiment, using Monte Carlo generated data. It also presents measurements of the jet cross section as a function of the p_T of the leading jet and of the second leading jet. Results are refer to hadron level jets with $p_T > 40$ GeV and $|\eta| < 3.0$, considering defined kinematic range for the Z products: electrons with $p_T > 25$ GeV and reconstruction region $|\eta| < 2.4$, and muons with $p_T > 15$ GeV and reconstruction region $|\eta| < 2.4$, within $81 < M_{l,l} < 101$ GeV, and $\Delta R_{lepton,jet} > 0.4$.

The generated Monte Carlo data are compared to NLO pQCD predictions, computed by MCFM program using CTEQ6.1M PDFs, and the renormalization and factorization scales set to $\mu^2 = M_Z^2 + p_{T,Z}^2$, and using a seeded cone iterative algorithm with R_{cone}=0.4 to reconstruct jets at the parton level. Two different Monte Carlo generators events, PYTHIA (parton-shower based) and ALPGEN+HERWIG (matrix element based) are considered, comparing their results to pQCD predictions.

The shape of the p_T distributions of the leading jet and the second leading jet predicted by ALPGEN+HERWIG generator agrees well with pQCD NLO MCFM predictions. Due to the tuning of the leading soft radiation in the parton shower, PYTHIA predicts a larger inclusive cross section for $Z(\rightarrow l^+l^-)$ + $\geq$ 1 jet but a clearly softer p_T spectrum for high-p_T values, which lines up with LO MCFM predictions for the leading jet.

Background processes to Z+jets events are studied and estimation methods are investigated. Main background processes are QCD dijet events and top production, which can represent between 5-15% of the events inside the signal peak region. Finally, systematic uncertainties to jet cross sections in Z+jets processes are studied. Jet Energy Scale uncertainty (expected to be known to 5 % in the early stage of the experiment) will be the main limitation for the Z+jets cross section measurement. Predictions for the first fb^{-1}of data from the ATLAS detector are being presented.

The presented studies are of great importance to understand the predictive power of the various event generator models for Boson+jet processes as well as to quickly comprehend Standard Model processes from real data, and finally discover new physics at the LHC .

Appendix A

Numerical Results

In this appendix some results obtained in the analysis are summarized in different tables.

A.1 PDF Uncertainties as a Function of p_T of the Leading Jet

PDF uncertainties for the differential MCFM cross-section prediction as a function of p_T of the leading jet are shown in table A.1. As explained in section 4.4, positive (+) and negative (-) deviations are shown in second and third column. Statistical uncertainties are shown on the last column.

p_T leading jet				
p_T range (GeV)	MCFM NLO σ (fb)	PDF syst. uncert. $\delta\sigma^+$	$\delta\sigma^-$	statistical uncertainty
40 - 50	30964.2	1851.3	1073.0	± 9.5
50 - 60	20776.2	1195.1	673.8	± 8.2
60 - 70	14319.1	799.1	433.7	± 6.7
70 - 80	10187.0	533.4	289.0	± 5.4
80 - 90	7349.5	368.0	196.9	± 4.4
90 - 100	5470.9	260.5	138.2	± 3.7
100 - 110	4150.7	186.6	97.9	± 3.0
110 - 120	3033.9	130.5	68.5	± 2.4
120 - 130	2435.3	98.1	51.6	± 2.1
130 - 140	1835.7	70.6	37.2	± 1.7
140 - 150	1454.6	52.7	28.5	± 1.4
150 - 160	1153.5	41.2	22.7	± 1.2
160 - 170	960.0	30.0	16.8	± 1.0
170 - 180	789.5	23.6	13.6	± 0.9
180 - 190	618.1	18.9	11.2	± 0.7
190 - 200	528.9	15.1	9.5	± 0.6
200 - 210	429.5	11.2	7.2	± 0.5
210 - 220	362.0	8.5	6.1	± 0.5
220 - 230	309.4	7.3	5.4	± 0.4
230 - 240	268.0	6.0	4.6	± 0.4
240 - 250	236.0	5.2	4.2	± 0.3
250 - 260	186.1	3.9	3.2	± 0.3
260 - 270	162.3	3.5	3.0	± 0.2

Table A.1: NLO p_T leading jet cross-section prediction for $Z \rightarrow \mu^+\mu^-$ channel. PDF systematic uncertainties are shown. Maximum values between $\delta\sigma^+$ and $\delta\sigma^-$ systematic uncertainties are used in the final results. Statistical uncertainties are shown in the last column.

A.2 Comparison of Monte Carlo Truth-level with MCFM NLO (LO) Predictions

In this section, final results comparing truth-level cross-section as a function of the p_T of the leading jet and the p_T of the second leading jet are compared with MCFM NLO and LO predictions corrected to hadron level. These results are shown in tables A.2 and A.3.

p_T leading jet				
p_T range (GeV)	MCFM NLO σ (fb)	MCFM LO σ (fb)	ALPGEN σ (fb)	PYTHIA σ (fb)
40 - 50	30964.2± 9.5	23584.1± 31.7	26993.3± 609.3	36101.8± 626.2
50 - 60	20776.2± 8.2	15447.5± 34.6	17686.0± 493.2	22277.6± 491.9
60 - 70	14319.1± 6.7	10470.2± 35.1	12273.2± 410.8	14702.4± 399.6
70 - 80	10187.0± 5.4	7254.3± 34.6	8639.2± 344.7	8913.2± 311.2
80 - 90	7349.5± 4.4	5125.4± 33.5	6435.4± 297.5	6210.8± 259.7
90 - 100	5470.9± 3.6	3694.5± 31.8	4684.1± 253.8	4525.3± 221.7
100 - 110	4150.7± 3.0	2679.5± 30.5	3379.5± 215.6	3270.3± 188.5
110 - 120	3033.9± 2.4	1973.4± 28.3	2642.6± 190.6	2299.3± 158.0
120 - 130	2435.3± 2.1	1459.1± 26.7	2014.6± 166.4	1767.9± 138.6
130 - 140	1835.7± 1.7	1099.6± 23.9	1636.5± 150.0	1328.3± 120.1
140 - 150	1454.6± 1.4	841.4± 26.4	1270.1± 132.2	1035.3± 106.0
150 - 160	1153.5± 1.2	648.1± 22.5	1043.4± 119.8	595.4± 80.4
160 - 170	960.0± 1.0	513.4± 16.3	861.8± 108.9	485.5± 72.6
170 - 180	789.5± 0.9	412.6± 13.7	676.1± 96.4	540.5± 76.6
180 - 190	618.1± 0.7	331.8± 25.3	542.4± 86.4	302.3± 57.3
190 - 200	528.9± 0.6	265.5± 26.1	469.9± 80.4	238.2± 50.9
200 - 210	429.5± 0.5	220.1± 23.8	378.2± 72.1	155.7± 41.1
210 - 220	362.1± 0.5	179.7± 21.1	339.7± 68.3	229.0± 49.9
220 - 230	309.4± 0.4	150.0± 19.3	244.9± 58.0	100.8± 33.1
230 - 240	268.0± 0.4	125.4± 17.3	252.0± 58.9	45.8± 22.3
240 - 250	236.0± 0.3	102.4± 15.3	206.2± 53.2	109.9± 34.6
250 - 260	186.1± 0.3	89.1± 14.1	158.2± 46.6	82.4± 29.9
260 - 270	162.3± 0.3	74.9± 12.7	136.3± 43.3	45.8± 22.3
270 - 280	140.0± 0.2	60.5± 11.4	123.2± 41.2	73.3± 28.2
280 - 290	121.2± 0.2	52.9± 10.4	101.4± 37.4	27.5± 17.3
290 - 300	103.5± 0.2	45.1± 9.4	80.4± 33.3	36.6± 20.0

Table A.2: Cross section (in fb) as a function of the p_T leading jet for $Z \rightarrow \mu^+\mu^-$ channel, comparing NLO and LO prediction (corrected to hadron level) with ALPGEN and PYTHIA truth-level predictions. Statistical errors are shown for all the cases.

p_T second leading jet				
p_T range (GeV)	MCFM NLO σ (fb)	MCFM LO σ (fb)	ALPGEN σ (fb)	PYTHIA σ (fb)
40 - 50	8744.6± 1.7	7141.9± 2.6	7041.8± 245.7	7749.8± 290.1
50 - 60	5044.4± 1.2	4217.7± 2.9	4113.4± 187.8	3398.5± 192.1
60 - 70	2924.6± 0.8	2665.3± 3.3	2418.0± 144.0	1740.5± 137.5
70 - 80	2101.1± 0.7	1777.8± 3.6	1668.8± 119.6	1062.6± 107.4
80 - 90	1396.4± 0.5	1247.7± 4.0	1154.4± 99.5	732.8± 89.2
90 - 100	1010.1± 0.4	894.0± 4.4	787.7± 82.2	412.2± 66.9
100 - 110	667.7± 0.3	678.8± 4.8	601.6± 71.8	238.2± 50.8
110 - 120	533.0± 0.3	514.2± 5.2	493.3± 65.0	183.2± 44.6
120 - 130	445.1± 0.3	403.4± 5.8	394.8± 58.2	155.7± 41.1
130 - 140	293.1± 0.2	318.9± 6.2	296.1± 50.4	82.4± 29.9
140 - 150	281.7± 0.2	257.1± 6.5	238.4± 45.2	64.1± 26.4
150 - 160	143.6± 0.1	208.7± 7.1	169.2± 38.1	27.5± 17.3
160 - 170	177.0± 0.1	158.8± 7.0	142.2± 34.9	18.3± 14.1
170 - 180	107.7± 0.1	136.9± 7.6	120.6± 32.2	36.6± 19.9
180 - 190	180.9± 0.1	109.2± 7.8	115.0± 31.4	36.6± 19.9
190 - 200	42.2± 0.1	98.9± 8.7	83.1± 26.9	9.2± 10.0

Table A.3: Cross section (in fb) as a function of the p_T second leading jet, comparing NLO and LO prediction (corrected to hadron level) with ALPGEN and PYTHIA truth-level predictions. Statistical errors are shown for all the cases.

Appendix B

Running MCFM

B.1 Introduction

A complete manual for MCFM (v.5) can be found in [28]. In this analysis, the default scheme has been used to define the electroweak couplings ($ewscheme = -1$). Table B.1 summarizes the input values used in this scheme. The default scheme uses effective field theory approach, which is valid for scales below the top mass. Other approaches are available. The values for the remaining parameters in MCFM are shown in table B.2.

Parameter	Name	Input value	Output value
G_F	Gf	1.16639×10^{-5}	input
$\alpha(M_Z)$	aemmz	1/128.89	input
$\sin^2\theta_\omega$	xw	0.2312	calculated
M_W	wmass	80.419 GeV	input
M_Z	zmass	91.188 GeV	input
m_t	mt	172.5 GeV	calculated

Table B.1: Default scheme used to fix the electroweak parameters of the Standard Model and the corresponding input and calculated values.

The selected parton distribution function in our analysis is CTEQ6_m [76]. The corresponding value of $\alpha_s(M_Z)$ is 0.118. In MCFM, the value of $\alpha_s(M_Z)$ is hardwired with the parton distribution, and the parton distribution also specifies the number of loops that should be used in the running of α_s. Other parton distribution function sets are available.

MCFM can process calculations at different orders:

- *lord*: The calculation is performed at leading order (LO) only.
- *virt*: Virtual (loop) contributions to the next-to-leading (NLO) order result are calculated (+counterterms to make them finite), including also the lowest order contribution.
- *real*: In addition to the loop diagrams calculated by *virt*, the full next-to-leading order results must include contributions from diagrams involving real gluon emission (-counterterms to make them finite). Only the sum of the *real* and the *virt* contributions is physical.
- *tota*: For simplicity, this option simply runs the *virt* and *real* pieces in series before performing a sum to obtain the full next-to-leading (NLO) order result.

Parameter	Fortran name	Default value
m_τ	mtau	1.777 GeV
m_τ^2	mtausq	3.1577 GeV^2
m_c^2	mcsq	2.25 GeV^2
m_b^2	mbsq	17.64 GeV^2
Γ_τ	tauwidth	2.269×10^{-12} GeV
Γ_W	wwidth	2.06 GeV
Γ_Z	zwidth	2.49 GeV
V_{ud}	Vud	0.975
V_{us}	Vus	0.222
V_{ub}	Vub	0.
V_{cd}	Vcd	0.222
V_{cs}	Vcs	0.975
V_{cb}	Vcb	0.

Table B.2: Default values for the remaining parameters in MCFM.

In our analysis, NLO order predictions have been calculated for Z+1jet and Z+2jets using the *tota* option, and *lord* for Z+3jets calculation.

B.2 Available Processes

Some available processes in MCFM are shown in table B.3.

Z-boson Production, Processes 31-33

These processes represent the production of a Z boson which subsequently decays either into electrons ($nproc = 31$), neutrinos ($nproc = 32$) or bottom quarks ($nproc = 33$). Where appropriate, the effect of a virtual photon is also included.

Z + jet Production, Processes 41-43

These processes represent the production of a Z boson and a single jet, where the Z subsequently decays into electrons ($nproc = 41$), neutrinos ($nproc = 42$) or bottom quarks ($nproc = 43$). Where appropriate, the effect of a virtual photon is also included. The calculation may be performed at NLO, although the NLO calculation of process 43 does not include radiation from the bottom quarks.

Z + 2jets Production, Process 44

This process represents the production of a Z boson and 2 jets, including also the effect of a virtual photon. The Z/γ^* decays to an e^+e^- pair. The calculation may be performed up to NLO. More details on this calculation can be found at [72, 92].

Z + 3jets Production, Process 45

This process represents the production of a Z boson and 3 jets, including also the effect of a virtual photon. The Z/γ^* decays to an e^+e^- pair. The calculation can be performed at LO only.

nproc	$f(p_1) + f(p_2) \to \cdots$	Order
31	$Z^0(\to e^-(p_3) + e^+(p_4))$	NLO
32	$Z^0(\to 3 \times (\nu(p_3) + \bar{\nu}(p_4)))$	NLO
33	$Z^0(\to b(p_3) + \bar{b}(p_4))$	NLO
41	$Z^0(\to e^-(p_3) + e^+(p_4)) + f(p_5)$	NLO
42	$Z^0(\to 3 \times (\nu(p_3) + \bar{\nu}(p_4))) - [\text{sum over } 3\nu] + f(p_5)$	NLO
43	$Z^0(\to b(p_3) + \bar{b}(p_4) + f(p_5))$	NLO
44	$Z^0(\to e^-(p_3) + e^+(p_4)) + f(p_5) + f(p_6)$	NLO
45	$Z^0(\to e^-(p_3) + e^+(p_4)) + f(p_5) + f(p_6) + f(p_7)$	LO

Table B.3: Some of the processes available at MCFM, indicated by the choice of the variable *nproc*. $f(p_i)$ denotes a generic partonic jet.

Appendix C

Jet Algorithm Analysis for Z+jets Processes

C.1 Introduction and Definitions

Among the large collection of jet algorithms available for physics analysis, Topo and Tower Cone 0.4 jet algorithms have been chosen for our analysis. Tower Cone 0.4 jet is used for $Z \rightarrow e^+e^-$ events and Topo Cone 0.4 for $Z \rightarrow \mu^+\mu^-$. In the following, a summary of the performance of different jet algorithms, using Z+jets events, is detailed. The motivation of this study was to choose an appropriate jet algorithm for the analysis.

The characteristics of the reconstructed variables depends on the choice of the calorimeter signal (Towers or Topological Clusters), the choice of the jet finder algorithm and its configuration (wide/narrow jets). All the default available jet algorithms were analyzed: cone jet and k_T jet. Tower jets and Topological cluster jets were studied. ALPGEN $Z \rightarrow \mu^+\mu^-$+jets data samples were used (see table 4.3).

Jet reconstruction performance is typically expressed in terms of expected or measured signal linearity, i.e. flatness of the detector response to particle jets over the whole kinematic range of interest, signal uniformity in pseudorapidity η and azimuth ϕ over the whole detector system coverage, and the achievable energy resolution. Additional features of jet reconstruction performance are the efficiency to find jets. This jet performance analysis is based on the jet reconstruction efficiency, the linearity in energy and uniformity in η, and the energy resolution.

Jet reconstruction efficiency is defined as follows:

$$efficiency = \frac{N_{matches\ of\ truth\ particle\ jets\ with\ reconstructed\ jets}}{N_{total\ truth\ particle\ jets}} = \frac{N_m^{jets}(R_{matching})}{N_{truth}^{jets}} \tag{C.1}$$

The matching is performed when the distance between reconstructed jet and truth jet, called matching radius, is $R_{matching} < \Delta R$, defined as follows:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{C.2}$$

Depending on the size of the jet algorithm, different values of ΔR have been chosen. Table C.1 shows the different values of ΔR. Only one match is allowed for each truth jet.

The difference of energy between the reconstructed jet and the matched truth jet is called *energy linearity* (or signal linearity). If calibration were perfect, the energy at reconstructed level should be on average equal to the truth level particle jet energy. The desirable value of linearity should be close to

Jet Algorithm	ΔR matching
Cone 0.4	0.3
Cone 0.7	0.5
k_T 0.4	0.3
k_T 0.6	0.5

Table C.1: Values of ΔR matching for the different jet algorithms. This value is used in reconstruction efficiency evaluation.

1.0, and not dependent on position or energy. Linearity is defined as follows:

$$linearity = \frac{Energy\ of\ reconstructed\ jet\ matched\ with\ a\ truth\ jet}{Energy\ of\ the\ truth\ jet} \tag{C.3}$$

In this case, the matching is at $R_{matching} < \Delta R = 0.05$ for all jet algorithms. The ΔR match cut is so small because we need to be sure we are matching exactly the same jet to obtain the difference of energies between them (reconstructed level and truth level). Signal uniformity is measured by the variation of the signal as function of the jet energy and jet direction. It is calculated as function of η, using equation C.3. And it is calculated as a function of the energy using a Gaussian fit of the distribution of the relative difference between $E^{jet}_{reconstructed}$ and E^{jet}_{truth}. Energy resolution is given by the width of the same fit distribution. Section C.6 details the different Gaussian fits performed on Topological Cone 0.4 jets.

Two different regions in η are studied: a central region where $|\eta| < 1.2$ and the whole detector region $|\eta| < 4.9$. In the case of jet p_T three different cuts are used: low cut ($p_T > 15$ GeV), medium ($p_T > 35$ GeV) and high cut ($p_T > 50$ GeV). Only efficiency results for low and medium p_T cuts are shown, since the high cut performs very similarly to the medium cut. For linearity studies low and high cut results are shown. In the $Z \rightarrow \mu^+\mu^-$ +jets analysis, jet p_T cut is set at $p_T > 40$ GeV, between the medium and the high cut. This cut was finally chosen after this jet analysis. In the jet analysis, a reconstructed $Z(\mu\mu)$ is required in each event.

C.2 Tower Jet Algorithms for Z+jets

In this first section, the jet algorithms investigated are Cone 0.4 and 0.7 jet algorithms and k_T 0.4 and 0.6 jet algorithms. They are all using CaloTowers as input.

Jet Reconstruction Efficiency

Jet reconstruction efficiency for the complete acceptance η coverage of the detector is shown in figure C.1(a). Clearly, Tower k_Ts algorithms perform better reaching already 90% efficiency at low jet p_T values ($\sim$ 20 GeV). Cone 0.4 jet shows a worse efficiency behavior, but for values of jet $p_T \sim 40$ GeV, the efficiency reaches the level of 95%. On the other hand, Tower Cone 0.7 has a low efficiency, even at high p_T ($\sim 60 - 80$ GeV), it does not reach the value $\sim 100\%$. This feature is likely related to the split-merge decision (some jet experts try to understand it [93]). Similar efficiencies shaper are shown in figure C.1(b), for the central region ($|\eta| < 1.2$). Tower Cone 0.4 jets are slightly more efficient at values around 30-40 GeV, than previously.

Figure C.2 shows the efficiencies as a function of the position in η. On the left part (a), the low p_T cut is applied. Tower k_T algorithms have efficiencies of 90% (Tower k_T 0.6 jet) and of 95% (Tower k_T 0.4 jet) and do not present any dependence on the position. Tower Cone jet efficiencies, however, do present a clear dependence on η. For high η values, i.e. forward jets ($|\eta| > 2.6$), efficiency decreases. This fact can be due to difficulties to build tower cells in the forward region for low energies because

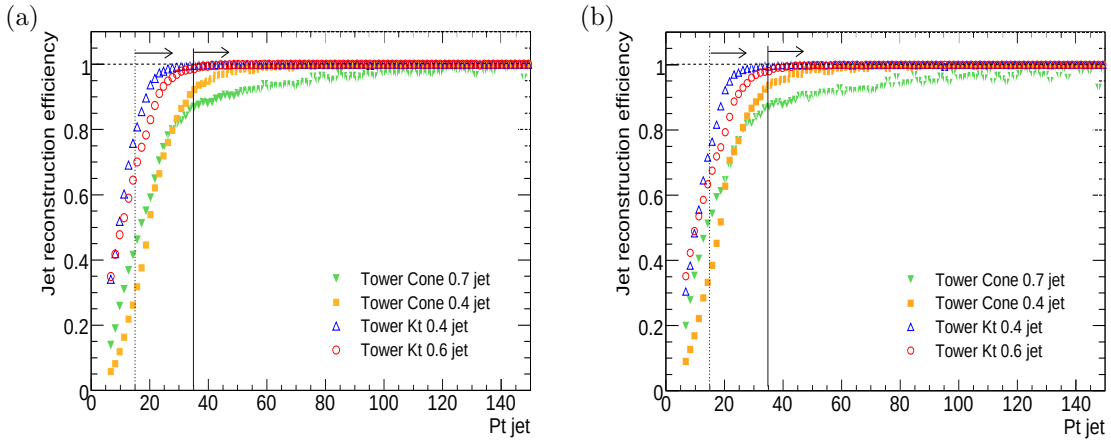


Figure C.1: Tower jet reconstruction efficiency vs p_T within the whole η region ($|\eta| < 4.9$), showing the two p_T cuts used later on (a), and for a cut $|\eta| < 1.2$ (central region) (b).

of the lack of projectivity in η. In figure C.2(b), efficiencies as a function of η for the medium cut are shown ($p_T > 35$ GeV). In this case, Tower k_T algorithms present a very good efficiency. Tower Cone jet efficiencies improve, reaching more than 90% efficiency, and the relaxing the η dependence. Tower Cone 0.4 jet performs better than Tower Cone 0.7 jet algorithm, specially for $|\eta| < 3.0$.

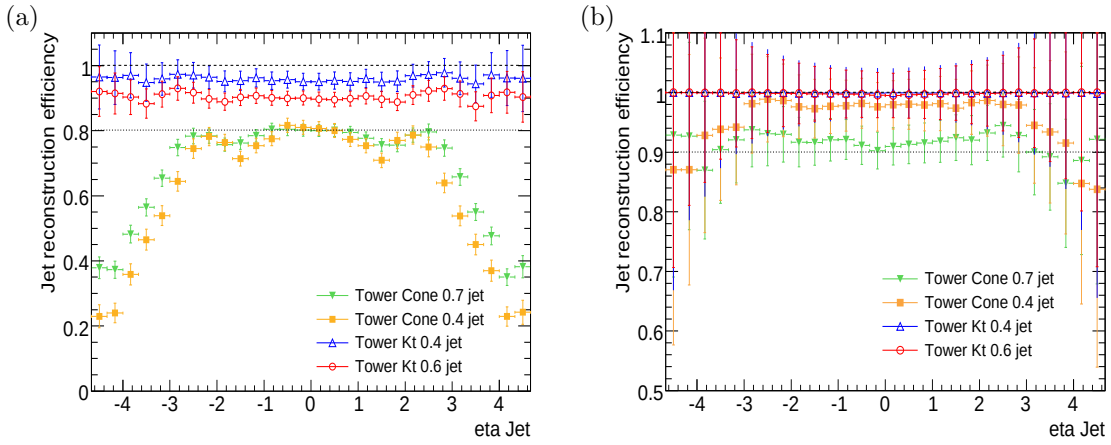


Figure C.2: Tower jet reconstruction efficiency vs η with $p_T > 15$ GeV (a) and for $p_T > 35$ GeV (b).

Finally, the efficiency performance of jet algorithms for different jet multiplicity environments is studied. Figure C.3 shows the reconstruction efficiency for (a) ≤ 6 jets and (b) ≤ 2 jets, for Tower Cone 0.7 and Cone 0.4, and for k_T 0.4 and 0.6 jet algorithms as a function of η, in the central part, where jet production from $Z + jet$ process is higher. It is clearly shown that Tower Cone 0.7 jet algorithm performs worse at central region ($|\eta| \sim 0$) than any other jet algorithm for high jet multiplicities. The other jet algorithms studied now do not present efficiency differences from low to high jet multiplicities. Therefore, for the study of high-jet multiplicity events, as in Z+jets production, an efficiency behavior like the one of Tower Cone 0.7 jet algorithm is undesirable.

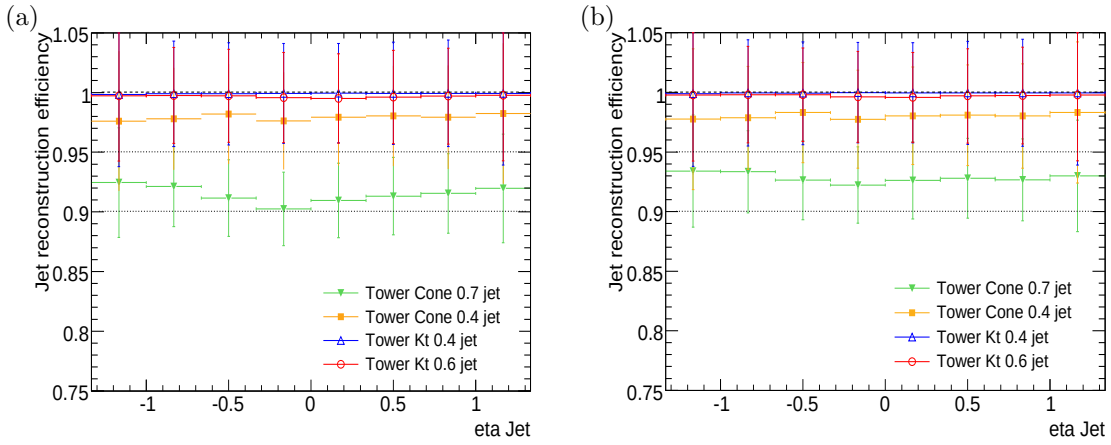


Figure C.3: Jet reconstruction efficiencies for events with ≤ 6 jets (a) and with ≤ 2 jets (b).

Jet Energy Linearity and Resolution

Figures C.4(a) and (b) show the signal uniformity for Tower algorithms. We can identify, clearly, the different parts of the detector, the central barrel, end-caps and forward. On the left, a p_T cut of 15 GeV is applied, while on the right a higher p_T cut of 50 GeV is applied. For high energies, performance is better in the forward part of the calorimeter, and uniformity gets close to 1.0. For values $|\eta| < 3.0$, signal uniformity is quite constant, except for the crack region at $|\eta| \sim 1.2 - 1.3$. From these two figures, one can see that Cone jets perform slightly better than k_T jet algorithms. The reason is that cell calibration was extracted for Cone 0.7 jets. Figures C.5(ac) and (b) show the jet energy linearity as a function of the

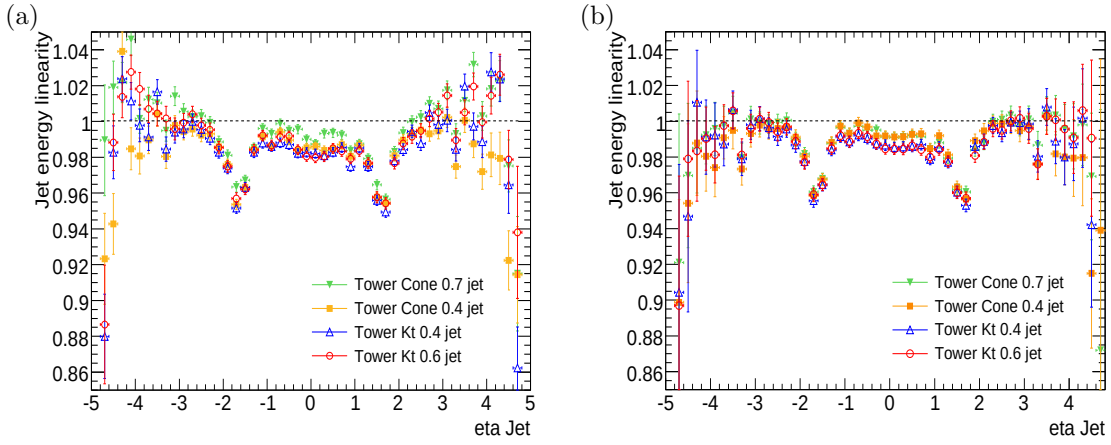


Figure C.4: Jet signal uniformity as a function of the position η for a cut $p_T > 15$ GeV (a) and for $p_T > 50$ GeV .

transverse momentum. Differences between the Cone and K_T jet algorithms are detected. Linearity for Tower Cone jet algorithms is ~ 0.99 for all p_T values. For the central region, shown on the right part of the figure (b), a small decrease of the linearity is shown, for $p_T < 50$ GeV. On the other hand, linearity of Tower k_T algorithm is much more dependent on the transverse momentum. Energy resolution is shown in figure C.5(c) and (d). At central region, resolution worsens, specially for Tower Cone 0.7 algorithm.

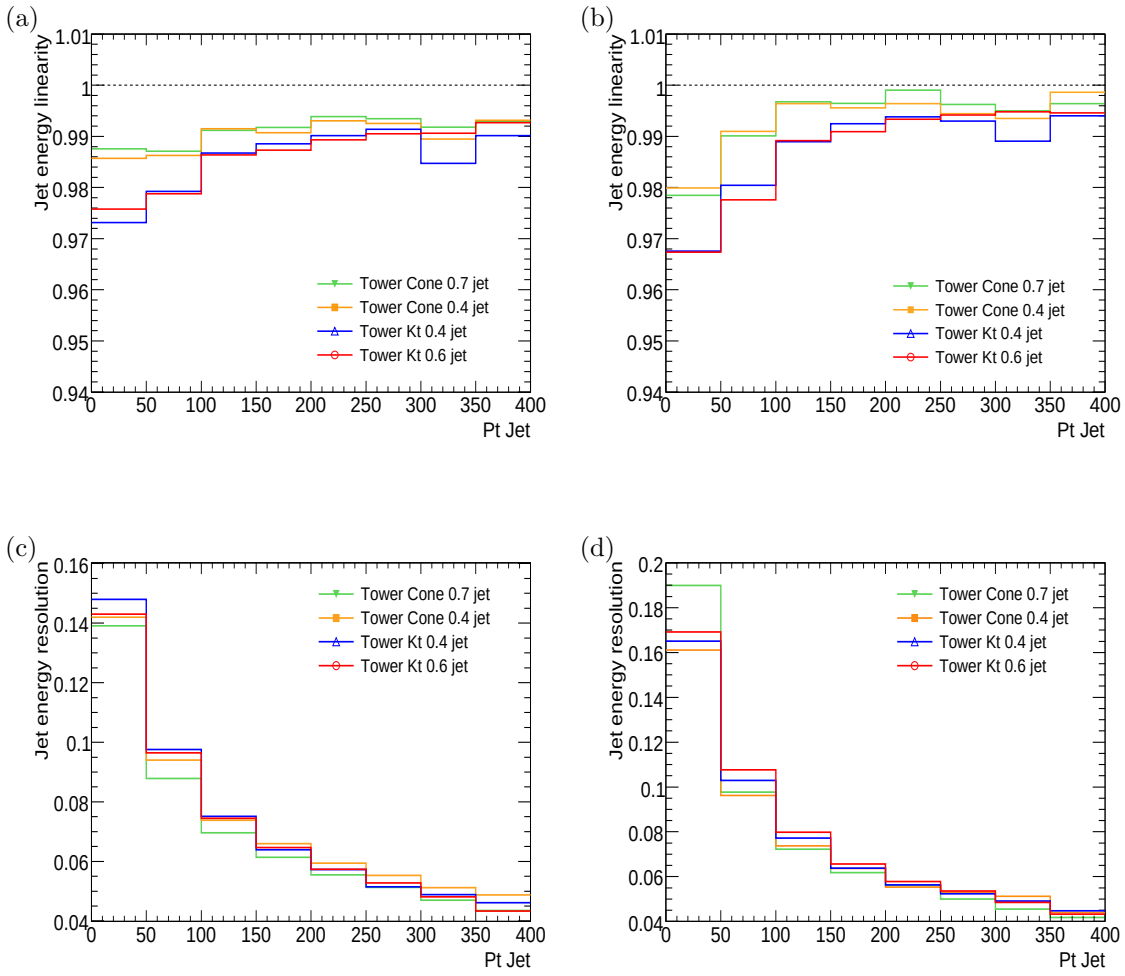


Figure C.5: Jet energy linearity as a function of p_T for all η range (a) and for central region ($|\eta| < 1.2$) (b). Jet energy resolution for all η range (c) and for central region ($|\eta| < 1.2$) (d).

In conclusion, Tower k_T jet algorithms have a better reconstruction efficiency, while Tower Cone jet algorithms present a better and more uniform linearity.

C.3 Cone Jet Algorithms for Z+jets

In this second part, the cone jet algorithms available are compared, for different inputs, Tower Cells and Topological Clusters jets, and using default cone radius values 0.4 and 0.7.

Jet Reconstruction Efficiency

Figure C.6 shows the reconstruction efficiency for Cone jet algorithms as a function of the transverse momentum of the jet. On the left (a), the whole detector region is selected. The Topological Cone 0.4 jet algorithm is the one which performs better. At energies of 35 GeV, efficiency is very good, reaching $\geq 95\%$. Tower Cone 0.4 jet algorithm has also a good efficiency, at 35 GeV, it is around 90%, and above 50 GeV it reaches 100% reconstruction efficiency. Tower and Topological Cone 0.7 jet algorithms have similar efficiency performance. The reconstruction efficiency is low, only 85% at 35 GeV, and for higher energies the efficiency maintains its low behavior. On the right (b), only the central region is selected ($|\eta| < 1.2$). The Tower Cone 0.4 jet efficiency increases, and gets to high values similar to those ones of Topological Cone 0.4 jet.

If we have a look at the efficiency versus the position η for Cone jet algorithms, the result is shown in figure C.7. For the low energy cut, on the left (a), Topo Cone 0.4 jet presents clearly a better reconstruction efficiency than the other Cone jet algorithms. Topological jets, Cone 0.4 and Cone 0.7, present a high efficiency at large η values ($|\eta| > 2.4$), while Tower jets have a decreasing efficiency. Tower Cone jet algorithms, as seen previously, do not have a good efficiency at large η values, which can be due to reconstruction difficulties at these forward positions for low energies values. On the right of figure C.7(b), reconstruction efficiency as a function of the position η is shown, for $p_T > 35$ GeV. Topo Cone jets present higher efficiencies (and nearly equal) than Tower Cone jets. For $p_T > 35$ GeV, efficiencies do not depend on the position η in any Cone jet algorithm.

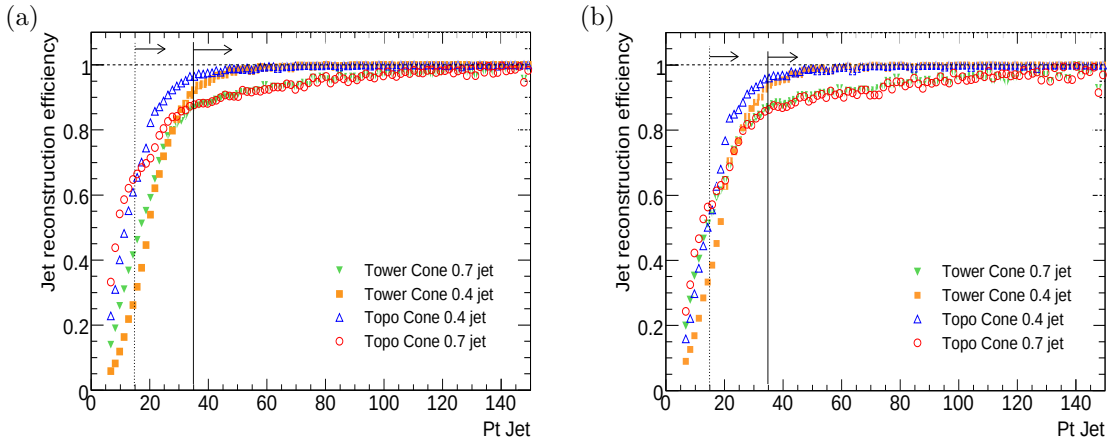


Figure C.6: Jet reconstruction efficiency vs p_T within the whole acceptance in η ($|\eta| < 4.9$) (a). The two p_T cuts used later are shown. And for $|\eta| < 1.2$ (central region) (b).

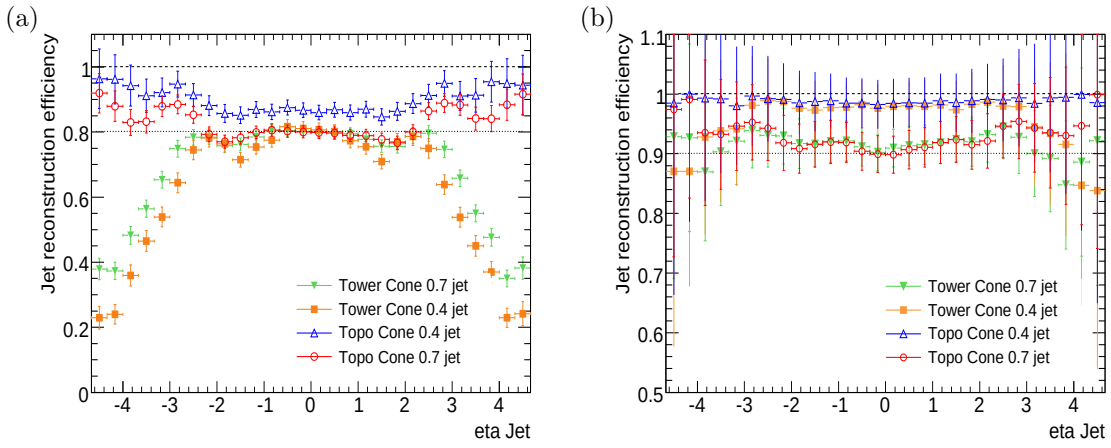


Figure C.7: Jet reconstruction efficiency vs η with $p_T > 15$ GeV (a) and for $p_T > 35$ GeV.

Finally, the efficiency performance of jet algorithms for different jet multiplicity Z+jets events is shown in figure C.8. For events with ≤ 6 jets (a), efficiency for Topo and Tower Cone 0.7 jets decreases at $|\eta| < 0.5$ in comparison with reconstruction efficiency for events with less number of jets (≤ 2 jets) (b). Topo and Tower Cone 0.4 jets do not present this efficiency behavior for high-jet multiplicity events.

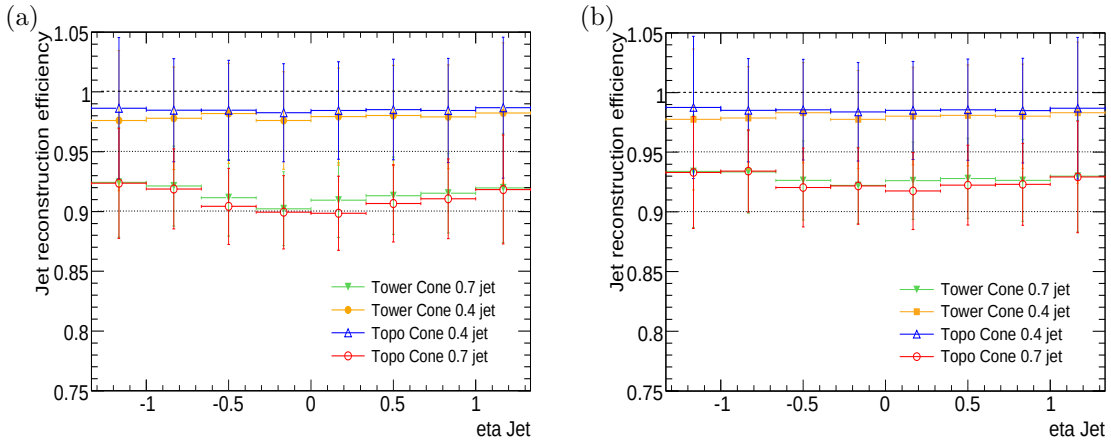


Figure C.8: Jet reconstruction efficiencies for events with ≤ 6 jets (a) and with ≤ 2 jets (b).

Jet Energy Linearity and Resolution

Figures C.9(a) and (b) show the signal uniformity for Tower and Topological Cone algorithms. We can observe a crack region at $|\eta| \sim 1.2$, and for $|\eta| > 3.0$ a non-flat signal uniformity for all Cone algorithms. For the low cut $p_T > 15$ GeV (a), Topo Cone 0.7 has the higher signal uniformity, followed by Topological Cone 0.4 and Tower Cone 0.7, with values of 0.99-1.0. Tower Cone 0.4 jet presents a lower value of signal uniformity, specially at central region. For the high cut $p_T > 50$ GeV (b), all Cone jet algorithms present a flat, high and similar signal uniformity for $|\eta| < 3.0$.

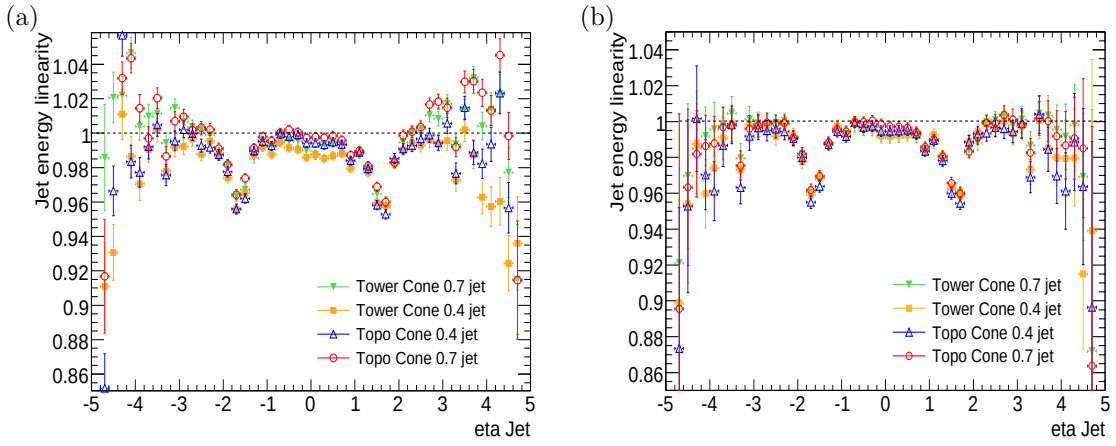


Figure C.9: Jet singal uniformity as a function of the position η for a cut $p_T > 15$ GeV (a) and for $p_T > 50$ GeV .

Figures C.10(a) and (b) show the jet energy linearity as a function of the transverse momentum. Energy linearity for Tower and Topological Cone jet algorithms is close to ~ 0.99 for all p_T range. While for the central η region, shown in figure C.10(b), a small decrease of the linearity for Tower Cone 0.7 jet is shown, for $p_T < 50$ GeV. On the other hand, linearity of Topological jets for $p_T < 50$ GeV is maintained constant and above > 0.99. Energy resolution is shown in figures C.10(c) and (d). At central region, resolution worsens, specially for Tower Cone 0.7 algorithm.

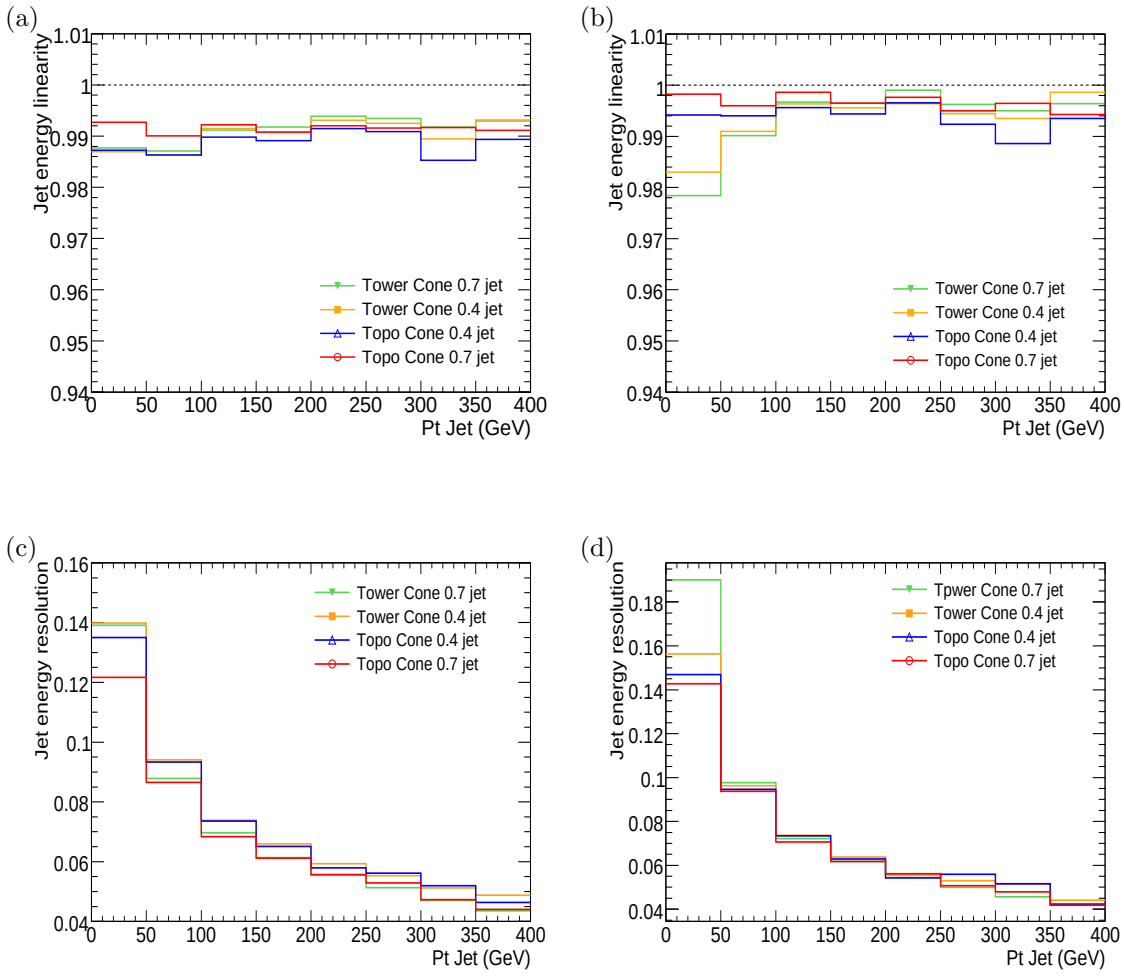


Figure C.10: Jet energy linearity as a function of p_T for all η range (a) and for central region ($|\eta| < 1.2$) (b). Jet energy resolution for all η range (c) and for central region ($|\eta| < 1.2$).

In conclusion, Tower and Topological Cone 0.4 jet have a better efficiency, and Topological Cone 0.4 and 0.7 present a better jet energy linearity. It should be added that Tower and Topological Cone 0.4 jet perform better for high-jet multiplicity events.

C.4 Mix of Jet Algorithms

Finally on this section, some selected algorithms are compared. We use Tower and Topological Cone 0.4 algorithms, since, from previous results, they all have a good performance to analyze $Z \rightarrow \mu^+\mu^-$ events, and we add Topological k_T 0.4 and k_T 0.6, not analyzed previously. From these four possible options, we arrive to our conclusion on which jet algorithm should be use in the analysis, although other criteria are also taken into account.

Jet Reconstruction Efficiency

Reconstruction efficiencies are shown in figure C.11. Topological k_T 0.4 jet algorithm has a very good efficiency, at 30 GeV efficiency reaches $\sim$ 100%. Topological k_T 0.6 jet algorithm has also a good efficiency, at 35 GeV its efficiency is close to 100%. Topological Cone 0.4 jet has a good efficiency too, at 35 GeV efficiency is close to 100%. Tower Cone 0.4 jet algorithm presents the lowest efficiency, 90% at 35 GeV. If we look for the central region, $|\eta| < 1.2$ (b), similar efficiency performances are obtained.

Figure C.12 shows the efficiency as a function of the position η. For low energies, $p_T < 15$ GeV (a), efficiencies for Topological jets are good. Topological k_T 0.4 jet has a flat efficiency around 95% for all η range. Topological k_T 0.6 jet has a good efficiency of $90 - 95\%$ in the whole region. Topological Cone 0.4 also presents a good reconstruction efficiency of around $85 - 90\%$. On the other hand, Tower Cone 0.4 jet has a low efficiency, of $\sim 80\%$ at central region, and a high decrease of efficiency for the high η values (forward region). At higher energies, $p_T > 50$ GeV (b), all reconstruction efficiencies increase. Tower Cone 0.4 jet also presents a higher efficiency, but it still has low efficiency for $|\eta| > 3.0$.

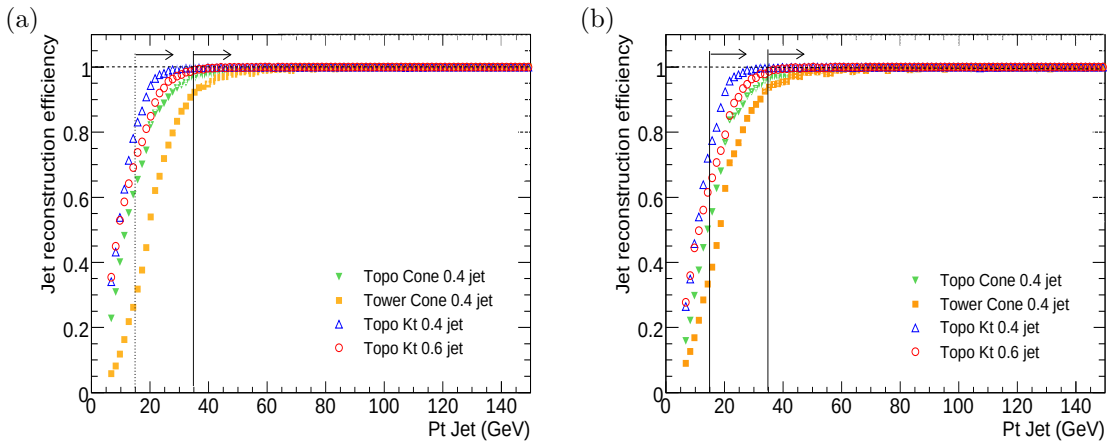


Figure C.11: Jet reconstruction efficiency vs p_T within $|\eta| < 4.9$, showing the two p_T cuts used (a), and for $|\eta| < 1.2$ (central region) (b).

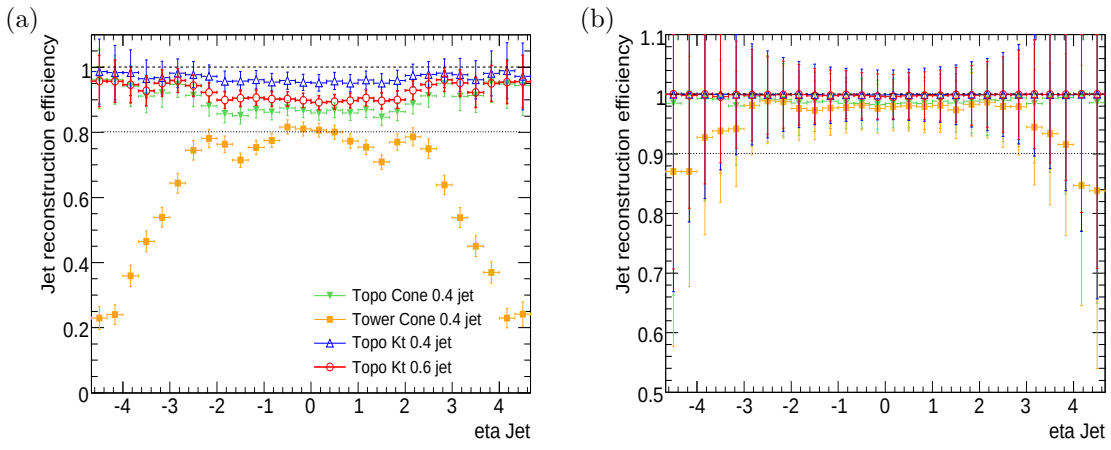


Figure C.12: Jet reconstruction efficiency vs η with a cut $p_T > 15$ GeV (a), and for $p_T > 35$ GeV (b).

Jet Energy Linearity and Resolution

Jet signal uniformity as a function of the position η is shown in figure C.13. Topological k_T 0.6, Topological k_T 0.4 and Topological Cone 0.4 jet have a high and flat signal uniformity (0.99 - 1.0), for low p_T values ($p_T > 15$ GeV) (a), except at crack region ($|\eta| \sim 1.2$) and for high η values ($|\eta| > 3.0$). This behavior is maintained for higher p_T ($p_T > 50$ GeV) (b). On the other hand, Tower Cone 0.4 jet has a

special behavior, as we have already seen. For the low cut, $p_T > 15$ GeV, signal uniformity is smaller (0.97-0.98) than any other jet algorithm. For higher p_T, this behavior disappears, and its jet signal uniformity reaches the same high value as the other jet algorithms.

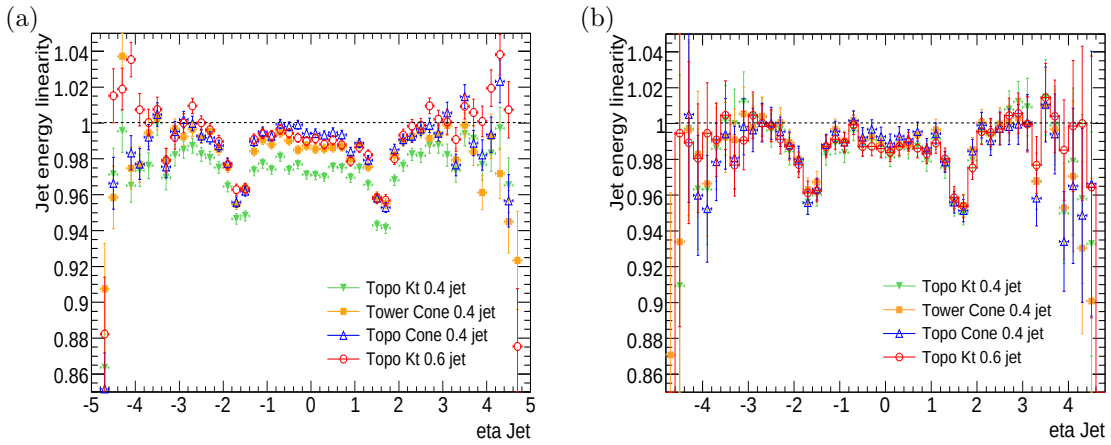


Figure C.13: Jet signal uniformity as a function of the position η for a cut $p_T > 15$ GeV (a) and $p_T > 50$ GeV (b) .

Figure C.14(a) and (b) shows the jet energy linearity for the studied jet algorithms. For the whole η range (a), all of them except Topological k_T 0.4, have a constant linearity of ~ 0.99. Topological k_T 0.4 jet has a decrease of linearity for $p_T < 100$ GeV. If we look at the central region, $|\eta| < 1.2$, Topo Cone 0.4 present the flattest and higher linearity (> 0.99). Topological k_T 0.6 and Tower Cone 0.4 jets have a small decrease of linearity for $p_T < 100$ GeV. Finally, Topological k_T 0.4 jet worsen its behavior, specially for $p_T < 50$ GeV. Figure C.14(c) and (d) show the energy resolution. At central region, resolution worsen, specially for Topological k_T 0.4 jet algorithm.

In conclusion, although Topo k_T 0.4 presents a very good efficiency, its low linearity makes this jet algorithm not optimal for Z+jets studies. Tower Cone 0.4 has a good performance, except at low energies, where reconstruction efficiency is low at large η values. For $p_T > 40$ GeV (jet cut for the analysis), the efficiency of Tower Cone 0.4 is good. Finally, Topological Cone 0.4 and Topological k_T 0.6 jets peresent a very good efficiency and a constant jet energy linearity. From this study point of view, Topological Cone 0.4 and Topological k_T 0.6 jets are the optimal ones to be used in Z+jets analysis.

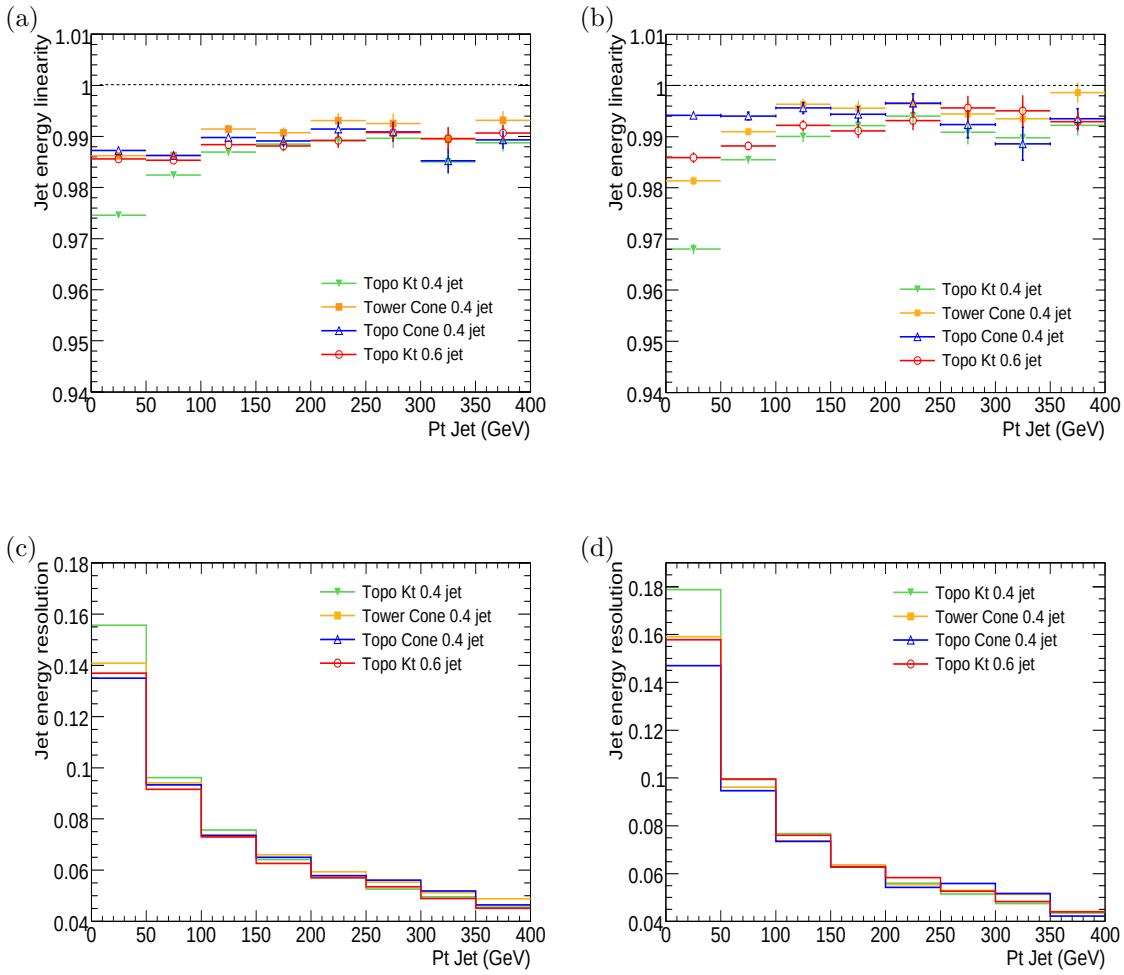


Figure C.14: Jet energy linearity as a function of p_T for all η range (a) and for central region ($|\eta| < 1.2$) (b). Jet energy resolution for all η range (c) and for central region ($|\eta| < 1.2$).

Jet Algorithm	Efficiency	Linearity & Uniformity
Tower Jets		
Cone 0.7	low eff. at 50 GeV & low eff. at high N jets	No p_T dependence, constant at $\|\eta\| < 1.2$
Cone 0.4	good eff. at $p_T > 40$ GeV, η dependence	No p_T dependence, constant at $\|\eta\| < 1.2$
k_T 0.6	good eff. at $p_T > 25$ GeV, no η dependence	p_T depend, decrease at $\|\eta\| < 1.2$, for $p_T < 100$ GeV
k_T 0.4	good eff. at $p_T > 25$ GeV, no η dependence	p_T depend, decrease at $\|\eta\| < 1.2$, for $p_T < 100$ GeV
Topological Jets		
Cone 0.7	low eff. at 50 GeV & low eff. at high N jets	No p_T dependence. Good performance at $\|\eta\| < 1.2$
Cone 0.4	good eff. at $p_T > 35$ GeV, no η dependence	No p_T dependence. Good performance at $\|\eta\| < 1.2$
k_T 0.6	good eff. at $p_T > 25$ GeV, no η dependence	No p_T dependence. Good performance at $\|\eta\| < 1.2$
k_T 0.4	good eff. at $p_T > 25$ GeV, no η dependence	p_T depend, decrease at $\|\eta\| < 1.2$, for $p_T < 100$ GeV

Table C.2: Summary of reconstruction efficiency and energy linearity for the different jet algorithms.

C.5 Summary and Conclusions

Here, we summarize all the results obtained previously for all types of jet algorithms. Table C.2 shows the summary of good and bad characteristics of the different jet algorithms.

The jet algorithms which perform better for Z+jets events, are Topo Cone 0.4 and Topo k_T 0.6 jets. Tower Cone 0.4 jet also is a good candidate, for $p_T > 40$ GeV. Finally, Tower Cone 0.4 jet algorithm has been used for $Z \rightarrow e^+e^-$ analysis and Topo Cone 0.4 for $Z \rightarrow \mu^+\mu^-$ analysis.

C.5.1 Final Results for Topo Cone 0.4 and Topo K_T 0.6 Jets

Figure C.15(a) shows the reconstruction efficiency for Topological Cone 0.4 and Topological k_T 0.6 jet algorithms. Jet energy linearity, for the whole η range is shown in figure C.15(b).

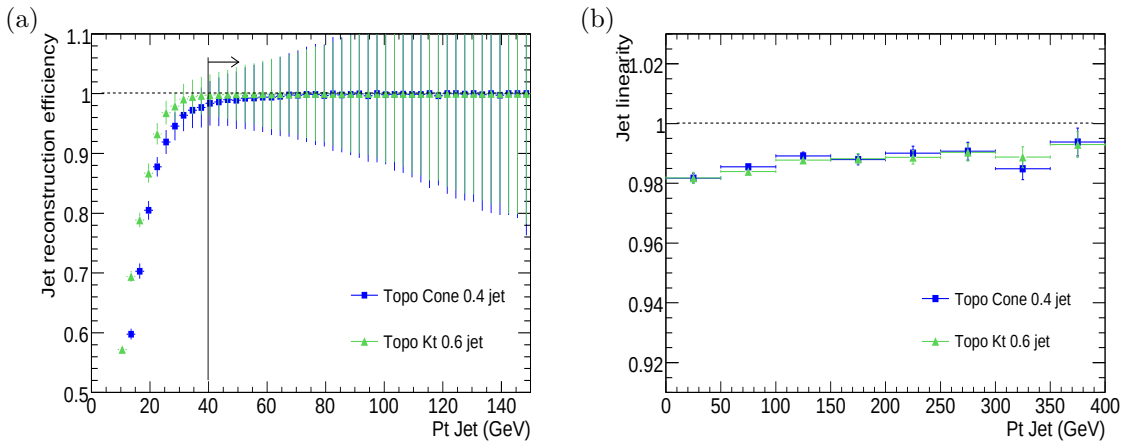


Figure C.15: Jet reconstruction efficiency (a) and jet energy linearity (b) for Topological Cone 0.4 and k_T 0.6 jets.

C.6 Linearity and Resolution Measurement

Jet energy linearity as a function of the truth jet energy is calculated using Gaussian fits. The different distributions of the E_T of the reconstructed matched jet with respect to the truth energy of the same jet, in different E_T beams, is fitted with a Gaussian function, in the range 0.8 -1.2. The mean obtain in the fit is estimated to be the energy linearity and the sigma obtained gives us the energy resolution in the E_T range studied.

Figure C.16 shows an example of the linearity and resolution measurement for Topo Cone 0.4 jets. A Gaussian fit is performed for each E_T region: 0-50 GeV, 50-100 GeV, 100-150 GeV, 150-200 GeV, 200-250 GeVand 250-300 GeV. For low E_T values, as shown in figure C.16(a) and (b), the width of the Gaussian is clearly wider than for higher E_T values, figures C.16(e) and (f). This means a worse energy resolution for low E_T, as expected. In addition, for low E_T values, distributions are asymmetric, see figure C.16(a), which means a loss of reconstructed energy in the jet. Gaussian functions have been reconstructed inside the tight window 0.8 -1.2, close to the peak value, as shown in the figures C.16(a), (b), (c), (d), (e) and (f).

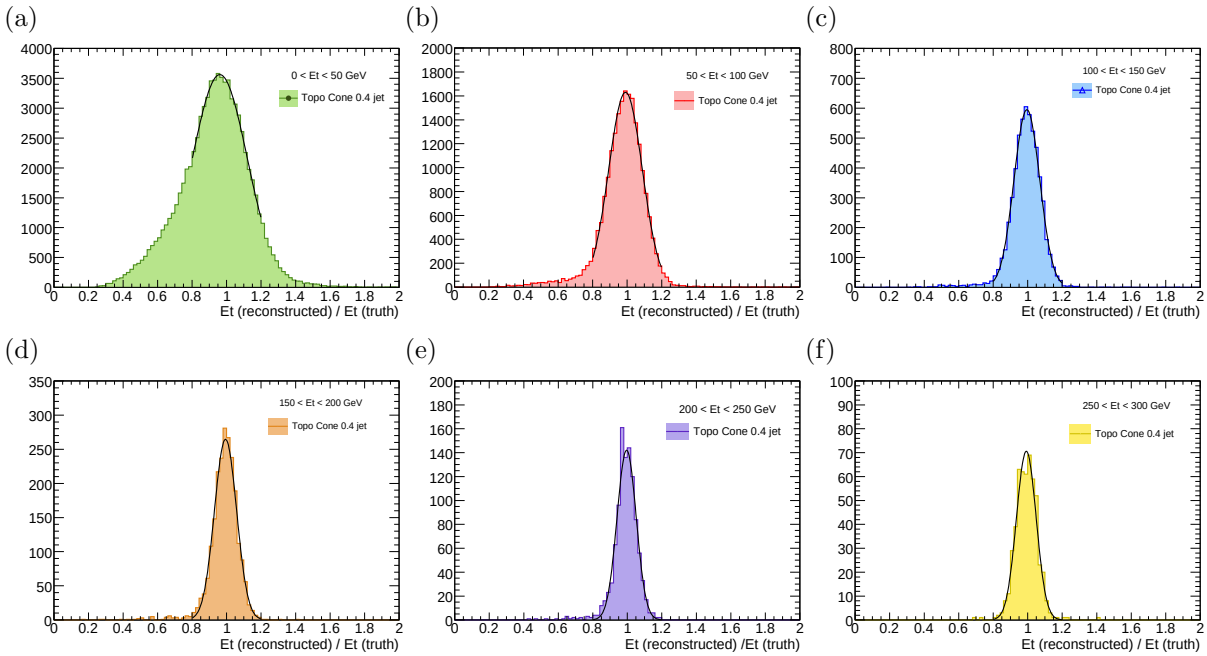


Figure C.16: Gaussian fits for the different distributions of the E_T of the reconstructed jet (matched to a truth jet) over the E_T of the truth jet. The fits have been applied for different E_T bins: 0-50 GeV(a), 50-100 GeV(b), 100-150 GeV(c), 150-200 GeV(d), 200-250 GeV(e) and 250-300 GeV(f). Gaussian fits are been performed inside de window $0.8 < x < 1.2$. The Gaussian mean has been taken as the jet energy linearity and the sigma has been used to get the energy resolution.

Appendix D

Unfolding Detector Effects for $Z \rightarrow \mu^+\mu^-$ +jets Events

D.1 Topological Cone 0.4 Jet Unfolding Process

In this section, unfolding[1] correction factors for TopoCluster Cone 0.4 jet, the jet algorithm used in $Z \rightarrow \mu^+\mu^-$ analysis, are derived. Similar results are obtained for $Z \rightarrow e^+e^-$+jets events shown in section 5.5. The unfolding procedure is meant to be the combination of detector correction factors used at detector level jets to get them to hadron level. The different detector effects are supposed to factorize, so we can treat them independently. Reconstruction efficiency for muons and jets, jet energy linearity and resolution effects are taken into account to correct reconstructed jets.

The jet reconstruction efficiency is determined as the fraction of truth jets which are matched to reconstructed jets, with a $\Delta R < 0.3$. The jet energy scale and resolution are determined using a matching window of $\Delta R_{(truth-reco\ jet)} < 0.05$ for $Z \rightarrow \mu^+\mu^-$+jets events. The impact of the resolution is derived by comparing the p_T distribution of truth jets before and after a Gaussian smearing with the resolution. Table D.1 summarizes all the unfolding jet factors for $Z \rightarrow \mu^+\mu^-$ process, using Topo Cone 0.4 jet algorithm.

unfolding jet factors			
p_T range (GeV)	reconstruction efficiency	jet energy linearity	resolution correction
40 - 60	0.984 $\pm$0.023	0.9840 $\pm$0.0009	0.982 $\pm$0.008
60 - 80	0.995 $\pm$0.036	0.9868 $\pm$0.0010	1.014 $\pm$0.009
80 - 120	0.998 $\pm$0.040	0.9892 $\pm$0.0008	1.007 $\pm$0.013
120 - 160	0.999 $\pm$0.068	0.9887 $\pm$0.0011	1.007 $\pm$0.014
160 - 240	0.999 $\pm$0.085	0.9910 $\pm$0.0011	1.003 $\pm$0.031
240 - 400	0.999 $\pm$0.137	0.9899 $\pm$0.0015	1.005 $\pm$0.052

Table D.1: Unfolding jet factors: reconstruction efficiency, jet energy linearity and resolution correction factors for Topo Cone 0.4 jet algorithms in $Z \rightarrow \mu^+\mu^-$+jets events.

[1] Unfolding procedure at CDF is performed in a different way, where real data is used, and unfolding factor corrections are obtained comparing data with Monte Carlo data, all correction effects are contained in only one correction factor (more information in appendix E). So the method used here must not be confused with the unfolding CDF-style, although the goal is the same.

Figure D.1 shows the jet reconstruction efficiency for TopoCluster Cone 0.4 jet (a) and the Jet Energy linearity factor, due to residual miscalibrations (b), as a function of the p_T of the jet. Both correction factors are obtained for different ranges on p_T (40-60 GeV, 60-80 GeV, 80-120 GeV, 120-160 GeV, 160-240 GeVand 240-300 GeV), in order to minimize statistical uncertainty.

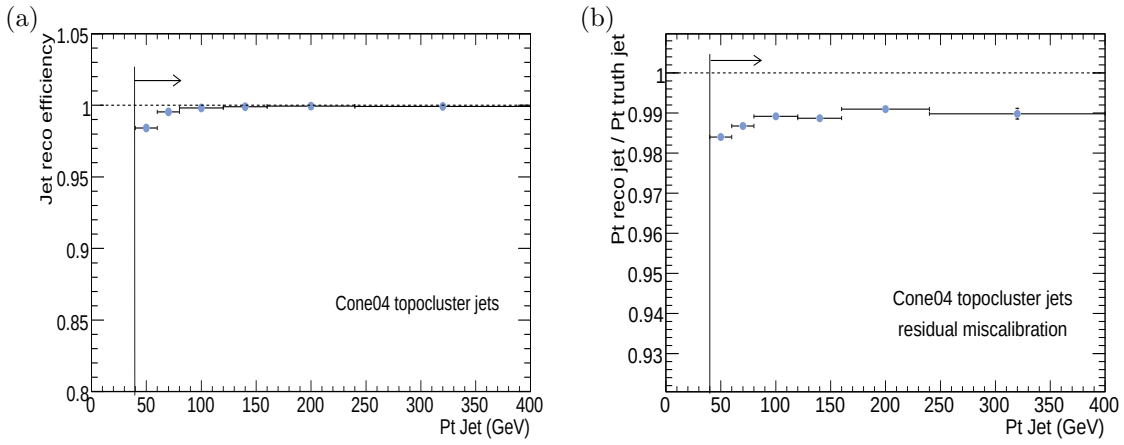


Figure D.1: Unfolding factors used from reconstructed level to hadron level, for jets in $Z \rightarrow \mu^+\mu^-$ events. Jet reconstruction efficiency distribution (a) and Jet energy linearity factors (b).

Jet energy resolution for TopoCluster Cone 0.4 jet is shown in figure D.2(a). The effects of the energy resolution are reflected in the correction factors shown in figure D.2(b).

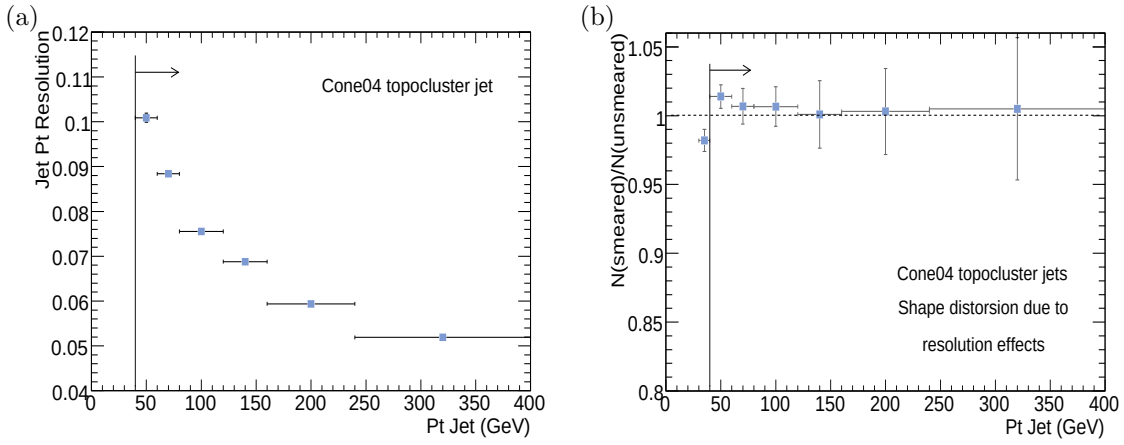


Figure D.2: Jet Resolution, for Topo Cone 0.4 jets in $Z \rightarrow \mu^+\mu^-$ events (a). Shape distortion due to resolution effects (b).

Finally, figure D.3 shows the p_T distribution of the leading jet and the second leading jet, corrected with the previous detector effects corrections and compared to the same quantities of Monte Carlo data. As we can observe, within the statistical and systematic errors, the p_T distributions of the Monte Carlo

jets and the corrected reconstructed jets are in agreement, thus providing a consistency check for these unfolding corrections.

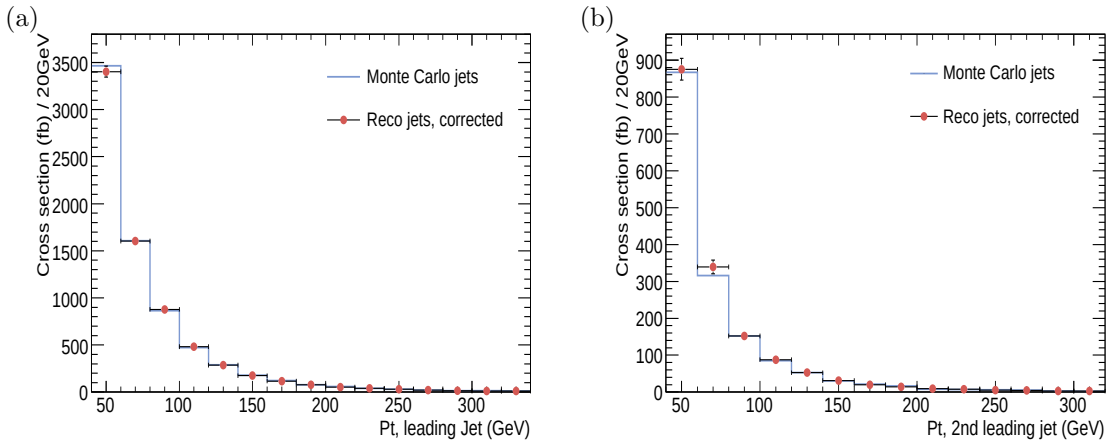


Figure D.3: Comparison of the distribution of the p_T of the leading jet (a) and the p_T of the second leading jet (b) for Monte Carlo jets and for reconstructed and corrected Topological Cluster Cone 0.4 jets.

Appendix E

Published Z+jets Results at Tevatron

The Tevatron is currently the most powerful hadron collider operational in the world. It is located at Fermilab (Batavia, IL, USA). There, protons and antiprotons are accelerated to interact with a center-of-mass energy of 1.96 TeV. The CDF [90] and D0 [78] experiments are located at the two interaction points of the Tevatron. Related results to our analysis from both experiments have been published. Here, there is a brief summary of published Z+jets results from Tevatron data.

E.1 CDF Results

Inclusive jet cross section in Z events, with Z decaying into an electron-positron pair, have been measured as a function of jet transverse momentum and jet multiplicity at $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV with the upgraded Collider Detector at Fermilab in run II, based on an integrated luminosity of 1.7 fb^{-1} and a more recent update for 2.5 fb^{-1}. The measurements cover the rapidity region $|y^{jet}| < 2.1$ and the transverse momentum range $p_T^{jet} > 30$ GeV/c. Next-to-leading order perturbative QCD predictions are in good agreement with the measured cross sections [94, 95].

The measurements are performed in the kinematic range defined as:

- Two electrons with $E_T > 25$ GeV
- At least one central electron: $|\eta^e| < 1$
- Second electron central or forward : $|\eta^e| < 1$ or $1.2 < |\eta^e| < 2.8$
- Z mass window: $66 < M_{ee} < 116$ GeV/c^2
- $\Delta R(e, jet) > 0.7$

Jets are searched for using the MidPoint [96] algorithm with R=0.7. Measurements are defined for hadron level jets with $p_T^{jet} > 30$ GeV/c and $|y^{jet}| < 2.1$.

Figure E.1(a) shows the measured M_{ee} invariant mass distribution in the selected region $46 < Mee < 136$ GeV/c^2 for data compared to the SM prediction in events with at least one jet. The plot includes 20 GeV/c^2 sidebands around the defined signal region. The measurement only includes statistical uncertainties. The main backgrounds to the $Z/\gamma^*(\rightarrow e^+e^-)$+jets sample arise from inclusive jets and W +jets events, which are estimated from the data. Other background contributions from $t\bar{t}$, $Z/\gamma^*(\rightarrow e^+e^-) + \gamma$, WW, WZ, ZZ and $Z/\gamma^*(\rightarrow \tau^+\tau^-)$ +jets final states are estimated using Monte Carlo samples. The total background for inclusive $Z/\gamma^*(\rightarrow e^+e^-)$+jets productions is about 12% for $N_{jet} \geq 1$, and increases up to about 17% for $N_{jet} \geq 3$.

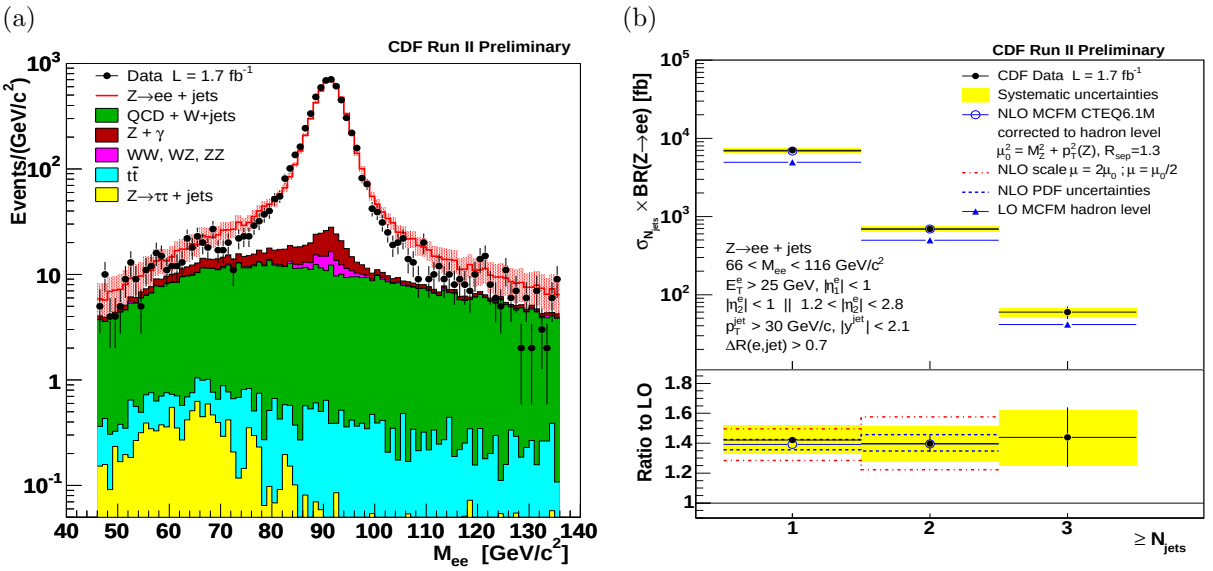


Figure E.1: Measured M_{ee} invariant mass distribution in the region $46 < Mee < 136$ GeV/c^2 for data compared to the SM prediction in events with at least one jet. The plot includes 20 GeV/c^2 sidebands around the defined signal region. The measurement only includes statistical uncertainties. (a). Total Inclusive Cross Section vs Inclusive Jet Multiplicity. Measured cross section as a function of inclusive jet multiplicity compared to NLO pQCD predictions as determined using MCFM. In the ratio plot, data and NLO theory are referred to LO predictions. Systematic uncertainties on data and theoretical predictions are included (b).

The measured cross section as a function of the inclusive jet multiplicity up to $N_{jet} \geq 3$, is shown in figure E.1(b). It is compared to the NLO pQCD predictions determined using MCFM. In the ratio plot, data and NLO theory are referred to LO predictions. Systematic uncertainties on data and theoretical predictions are included. The parton-to-hadron non-perturbative corrections vary between 1.1 and 1.4 as N_{jet} increases. The LO pQCD predictions underestimate the measured cross sections by a factor 1.4 approximately independent of N_{jet}. Good agreement is observed between data and NLO pQCD predictions.

Figure E.2 shows the 1^{st} leading jet cross section in $Z+ \geq 1$jet production as a function of p_T^{jet} compared to NLO pQCD predictions as determined using MCFM. The pQCD prediction includes an additional C_{had} factor that accounts for non-pQCD contributions. Figure E.2(b) shows the 2^{nd} leading jet cross section in $Z+ \geq 2$jets production as a function of p_T^{jet} compared to NLO pQCD predictions, including C_{had} factor. Systematic uncertainties on data and theoretical predictions are included. The cross sections decreases by more than three order of magnitude as p_T^{jet} increases from 30 GeV/c up to about 300 GeV/c. The parton-to-hadron correction, C_{had}, decreases as p_T^{jet} increases from about 1.2(1.26) at p_T^{jet} of 30 GeV/c to 1.02(1.01) for $P_T^{jet} > 200$ GeV/c for $N_{jet} \geq 1$ ($N_{jet} \geq 2$), and is dominated by the underlying event contribution. The uncertainty on C_{had} is about 10% (17%) at low p_T^{jet} and goes down to 1% at high p_T^{jet} for $N_{jet} \geq 1$ ($N_{jet} \geq 2$). Good agreement is observed between the measured cross sections and the nominal theoretical predictions.

Finally, figure E.3 shows relative contributions to the total systematic uncertainty to the 1^{st} leading jet and to the 2^{nd} leading jet cross sections as a function of p_T^{jet}. The main systematic uncertainty comes from jet energy scale uncertainty (3%), $3\% - 12\%$ for $Z+ \geq 1$jet and 5%-13% for $Z+ \geq 2$jets.

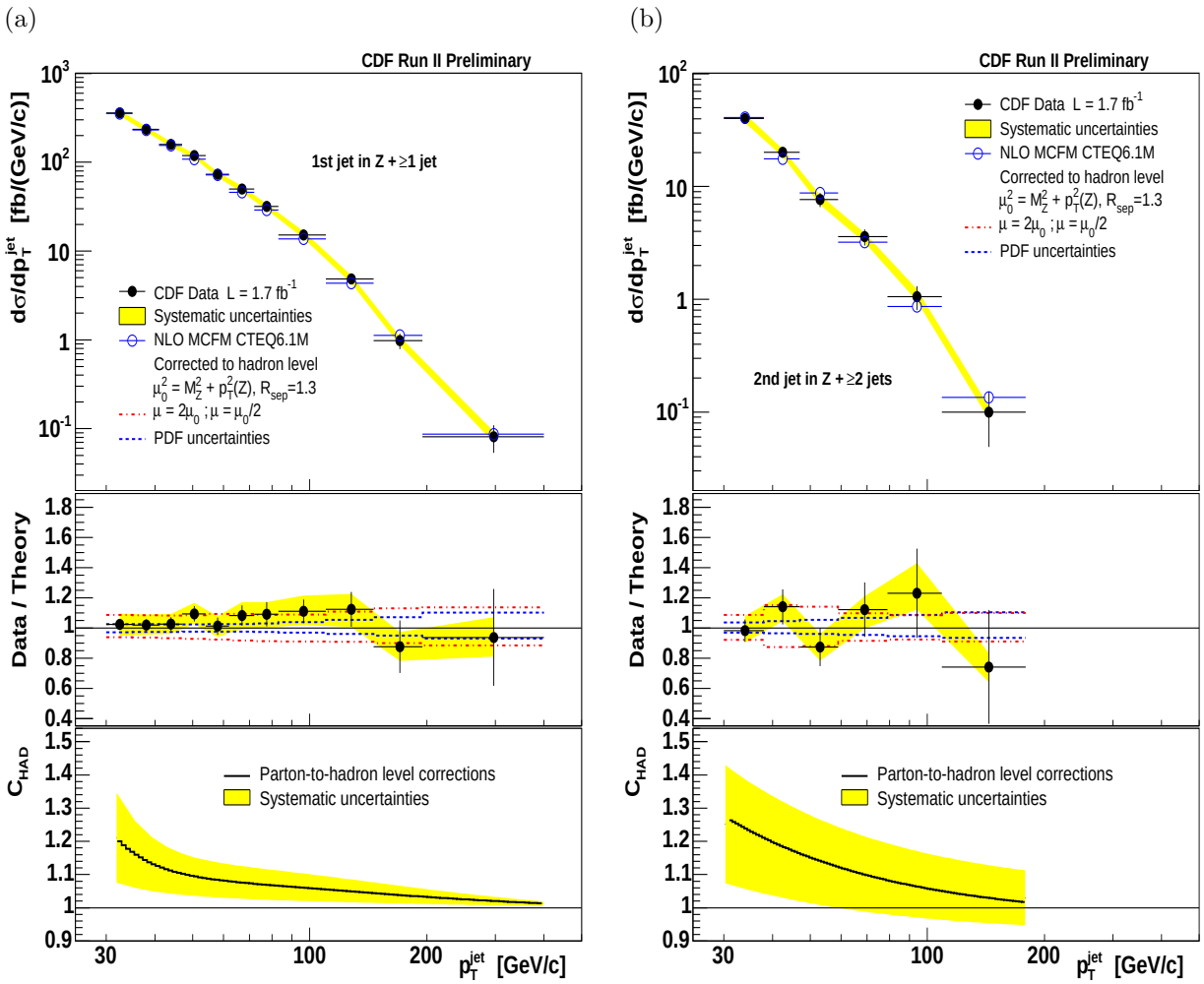


Figure E.2: 1^{st} leading jet cross section in $Z+ \geq 1$jet production (a) and 2^{nd} leading jet cross section in $Z+ \geq 2$jets production (b). Both cross sections are compared to NLO pQCD predictions from MCFM.

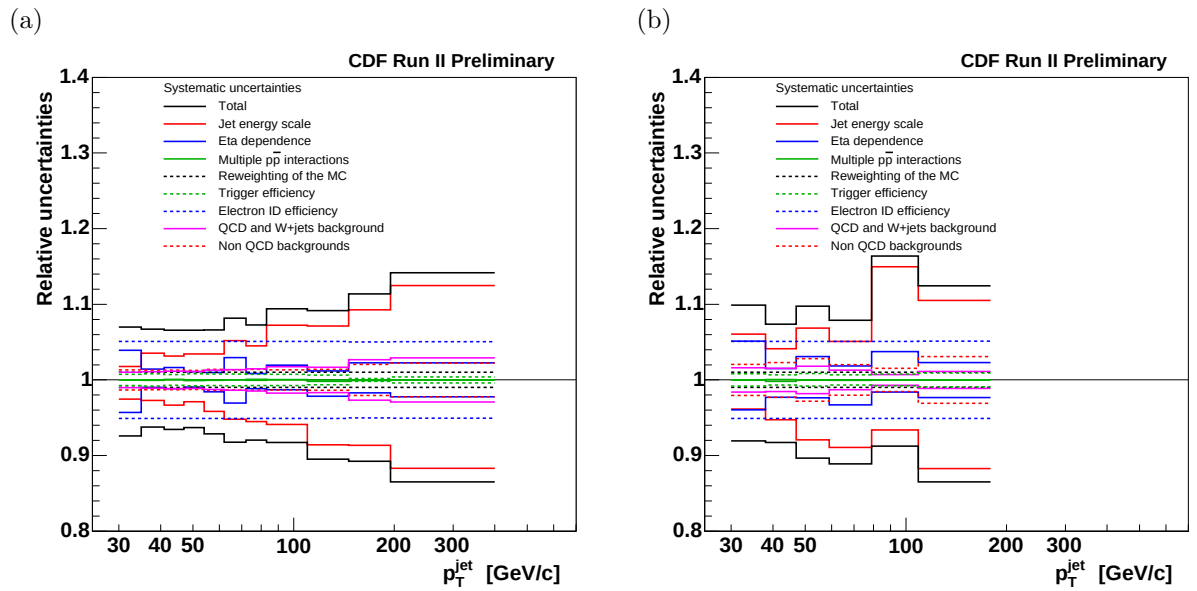


Figure E.3: 1^{st} leading jet (a) and 2^{nd} leading jet (b) cross section relative systematic uncertainties.

E.2 D0 Results

A study of events with Z bosons and associated jets produced at the Fermilab Tevatron Collider in $p\bar{p}$ collisions at a center of mass energy of 1.96 TeV has been performed. The data samples consists of nearly 14,000 $Z/\gamma^* \to e^+e^-$ candidates corresponding to an integrated luminosity of 0.4 fb^{-1} collected with the D0 detector. Ratios of the $Z/\gamma^*+ \geq n$ jet cross sections to the total inclusive Z/γ^* cross section have been measured for n = 1 to 4 jets, and found to be in good agreement with a next-to-leading order QCD calculation and with a tree-level QCD prediction with parton shower simulation and hadronization [97].

Events are required to have:

- Two electrons with $p_T > 25$ GeV
- Both central electron: $|\eta^e| < 1.1$
- At least one electron fire the trigger for the event
- Z mass window: $75 < M_{ee} < 105$ GeV/c^2
- One reconstructed primary vertex with a position along the beam direction within 60cm, spatially matched with one electron.
- $\Delta R(e, jet) > 0.4$

Jets are reconstructed using the "Run II cone algorithm" that combines particles within a cone of radius $R_{cone} = 0.5$. The transverse momentum of each jet was corrected for multiple $p\bar{p}$ interactions, calorimeter noise, out-of-cone showering effects, and energy response of the calorimeter. Jets were required to have $p_T > 20$ GeV and $|\eta| < 2.5$.

The primary background to the Z/γ^* di-electron signal is from multijet production. For the $Z/\gamma^*+ \geq 0-2$ jet samples, a convoluted Gaussian/Breit-Wigner function was used to fit the Z lineshape, and an exponential form was used to account for both the QCD background and the Drell-Yan (γ^*) component of the signal. For lower statistics $Z/\gamma^*+ \geq 3$ jet samples, the contributions from QCD and Drell-Yan components were estimated from the side bands of the Z in the dielectron invariant mass spectrum. The background contributions from QCD processes was found to be 3-5%, depending on the jet multiplicity. Other contributions are not from misidentified electrons, but correspond to other Standard Model processes (e.g., $t\bar{t}$ production, $Z \to \tau^+\tau^-$, $W \to e\nu$). These small irreducible background contributions ($< 1\%$) were also taken into account.

Two independent Monte Carlo samples were used to determine the correction factors for the cross section, both tuned to match the measured inclusive jet multiplicity and jet p_T distributions in data. The first sample was based on PYTHIA simulations. The second sample (ME-PS) was based on MADGRAPH [98] $Z/\gamma^* + n$ LO Matrix Element (ME) predictions, using PYTHIA for parton showering (PS) and hadronization.

Figure E.4(a) shows the fully corrected measured cross-section ratios for $Z/\gamma^*+ \geq n$ jets as a function of jet multiplicity, compared to three QCD predictions. The MCFM and the ME-PS predictions are generally in good agreement with data. PYTHIA predicts fewer events at high jet multiplicity. Figure E.4(b) compares the jet p_T spectra of the n^{th} jet, n= 1,2,3 in $Z/\gamma^*+ \geq n$ jet events to the ME-PS MC predictions. Good agreement can be seen over a wide range of jet transverse momenta.

(a) (b)

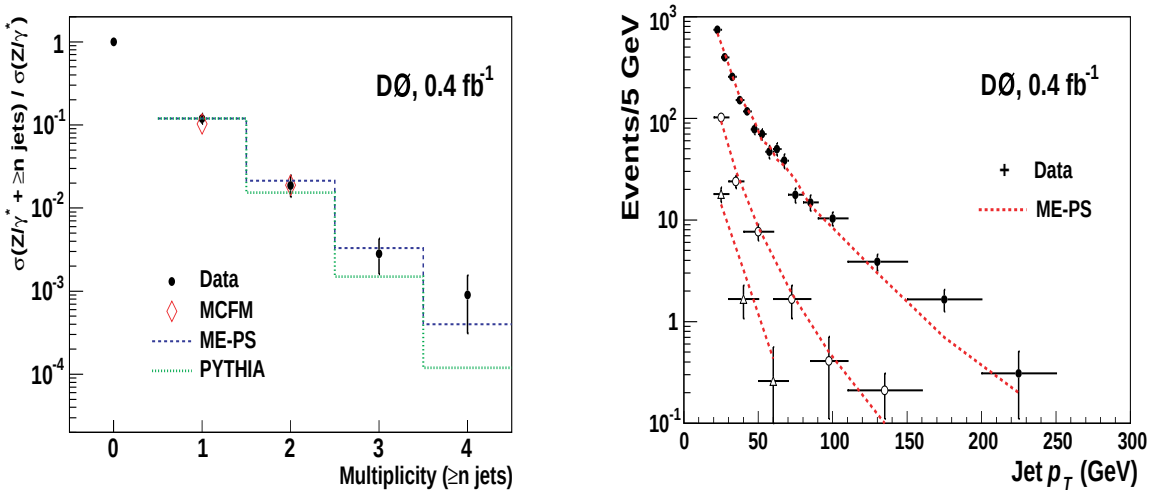


Figure E.4: Ratios of the $Z/\gamma^* \geq n$ jet cross sections to the total inclusive Z/γ^* cross section versus jet multiplicity. The uncertainties on the data (dark circles) include the combined statistical and systematic uncertainties added in quadrature. The dashed line represents predictions of LO Matrix Element (ME) calculations using PYTHIA for parton showering (PS) and hadronization, normalized to the measured $Z/\gamma^*+ \geq 1$ jet cross-section ratio. The dotted line represents the predictions of PYTHIA normalized to the measured $Z/\gamma^*+ \geq 1$ jet cross-section ratio. The two open diamonds represent predictions from MCFM. (a). Comparison between data and theory (ME-PS) for the highest p_T jet distribution in the $Z/\gamma^*+ \geq 1$jet sample (dark circles), for the second highest p_T jet distribution in the $Z/\gamma^*+ \geq 2$ jet sample (open circles), and for the third highest p_T jet distribution in the $Z/\gamma^*+ \geq 3$ jet sample (open triangles). The uncertainties on the data are only statistical. The MC distributions are normalized to the data. (b)

A new study from D0 experiment extends the previous results to the measurements of differential distributions up to the third leading jet [99]. The data corresponds to an integrated luminosity of 1.04 $\pm$ 0.06 fb^{-1}. Cross sections are presented in bins of the transverse momentum (p_T) of the N^{th} jet in events containing at least N = 1, 2, or 3 jets and are normalized to the measured inclusive Z/γ^* cross section to reduce uncertainties. The Z/γ^* is selected via its decay into an electron-positron pair. A comparison of the differential p_T distributions to leading order (LO) and next-to-leading order (NLO) perturbative QCD (pQCD) predictions from MCFM, as well as to results from several event generators: parton-shower-based event generators like PYTHIA and HERWIG, and matrix element and parton shower merging procedures adopted in event generators as ALPGEN+PYTHIA and SHERPA.

Events are required to have:

- Two electrons with $p_T > 25$ GeV, with opposite sign electric charge, $p_T > 25$ GeV and $|\eta|$ <1.1 or 1.5 < $|\eta|$ < 2.5.
- Calorimeter jets are reconstructed using iterative seed-based cone jet algorithm, with a radius R_{cone}=0.5. Jets are selected with p_T >20 GeV and $|\eta| < 2.5$.
- Events are required to pass single or dielectron trigger requirements.
- Z mass window satisfying $65 < M_{ee} < 115$ GeV

Backgrounds arising from events containing two real electrons (e.g., WW, $t\bar{t}$, and $Z/\gamma^*(\rightarrow \tau\tau)$) were estimated to be below 6% using event samples generated with PYTHIA. Backgrounds arising from events with one or more misreconstructed electrons (e.g. W+jets or multi-jet events) were found to be below 1% in the inclusive $Z/\gamma^*(\rightarrow e^+e^-)$. For the signal samples containing at least 1, 2 or 3 jets, no statistically significant contribution from W +jets or multi-jet events was observed (contribution to the signal samples at the 0.1% level using event sample generated with PYTHIA).

Corrections of the reconstructed p_T^{jet} spectra to the particle level were determined using an event sample generated with ALPGEN+PYTHIA. The simulated events were overlaid with data events from random bunch crossings to reproduce the effects of detector noise and additional $p\bar{p}$ collisions. The total uncertainty due to differences between the simulation and the data in the dependence of the electron identification efficiency on jet activity was estimated to be below 2% for all cross section measurements.

Jet energy scale (JES) corrections were derived using γ+jet and dijet events. This correction varies from $\sim$ 5% below 40 GeV to $\sim$ 0% above 80 GeV. The jet energy resolution (JER) was determined using Z/γ^* +jet events. The JER distorts the steeply falling jet p_T spectra, resulting in a net migration towards higher values of p_T^{jet}. This leads to the reconstructed p_T^{jet} spectra being (5-15)% higher than they would have been for a detector with perfect jet energy resolution.

The main systematic uncertainty of the measurements arise from the correction of the JES, contributing (50-80)% of the total systematic uncertainty of the measurements. The measurements were presented as ratios to the inclusive $Z/\gamma^*(\rightarrow e^+e^-) + X$ cross section in order to cancel the dependence on the uncertainty in the integrated luminosity of the data set.

The measurements presented above are compared with the predictions of several different theoretical models. For each model, the predicted jet p_T spectra are normalized to the predicted inclusive $Z/\gamma^*(\rightarrow e^+e^-) + X$ cross section. All predictions were generated using the CTEQ6.1M parton density functions (PDFs) and the two-loop formula for the evolution of the strong coupling constant (α_s). For the first and second jets the NLO MCFM predictions have been taken as the reference prediction; for the third jet, the leading-order (LO) MCFM prediction plays this role. The central predictions were defined using factorization and renormalization scales $\mu_F = \mu_R = \sqrt{M_Z^2 + p_{T,Z}^2}$, with M_Z and $p_{T,Z}$ denoting the mass and transverse momentum of the Z/γ^* boson. The sensitivity of the predicted cross sections to the choice of μ_F and μ_R was tested by varying their values up and down from the nominal value by a factor of two. The MCFM predictions were multiplied by correction factors accounting for multiple parton interactions (C_{MPI}) and hadronization (C_{Had}) before being compared to the measurements. The correction factors C_{MPI} and C_{Had} were estimated using inclusive $Z/\gamma^*(\rightarrow e^+e^-)$ event samples generated with PYTHIA, HERWIG+JIMMY, ALPGEN+PYTHIA and SHERPA. Both the NLO and LO MCFM predictions are in agreement with the measurements within the experimental and theoretical uncertainties (figures E.5-E.7). At NLO, varying the factorization and renormalization scales up (down) by a factor of two changes the normalized p_T^{jet} spectrum down (up) by a factor of $\approx$ 1.1 for the leading jet, compared to a factor $\approx$ 1.2 for LO. For the second leading jet, the factors are $\approx$ 1.1 (NLO) and $\approx$ 1.4 (LO), and for the third $\approx$ 1.6 (LO). The uncertainties of the MCFM predictions due to the PDFs vary from 5% at low p_T to 10% at high p_T for the two leading jets, and (5-15)% for the third leading jet.

Next, a comparison between the predictions of parton-shower based event generators PYTHIA and HERWIG+JIMMY with the measurements is made. For the leading jet, PYTHIA shows a more steeply falling spectrum than observed in data (figure E.5). The prediction of HERWIG+JIMMY shows good agreement with data at low p_T^{jet}, but resembles PYTHIA at high p_T^{jet}. Comparisons of sub-leading jets (figures E.6, E.7) show that PYTHIA and HERWIG predict more steeply falling p_T^{jet} spectra than observed in data, in agreement with expectations based on the limited validity of the soft/collinear ap-

proximation of their parton shower. Finally, comparisons with the ALPGEN+PYTHIA and SHERPA tree-level matrix element event generators are made. For all three p_T^{jet} spectra, ALPGEN+PYTHIA predicts lower production rates than observed in data, but the shapes of the spectra are well described. SHERPA predicts a less steeply falling leading p_T^{jet} spectrum than seen in data, leading to disagreements above 40 GeV. For the sub-leading p_T^{jet} spectra, SHERPA predicts higher production rates than observed in data, but the shapes are well described.

In summary, predictions of MCFM at NLO, corrected to the particle level, are found to be in good agreement with data and have a significantly smaller scale uncertainty than MCFM at LO. HERWIG and PYTHIA event generator models show significant disagreements with data which increase with p_T^{jet} and the number of jets in events. The SHERPA and ALPGEN+PYTHIA generators shown an improved description of data as compared with the parton-shower-based generators. ALPGEN+PYTHIA gives a good description of the shapes of the p_T^{jet} spectra, while predicting lower production rates than observed in data. SHERPA predicts higher production rates and a less steeply falling p_T^{jet} spectrum for the leading jet than observed in data.

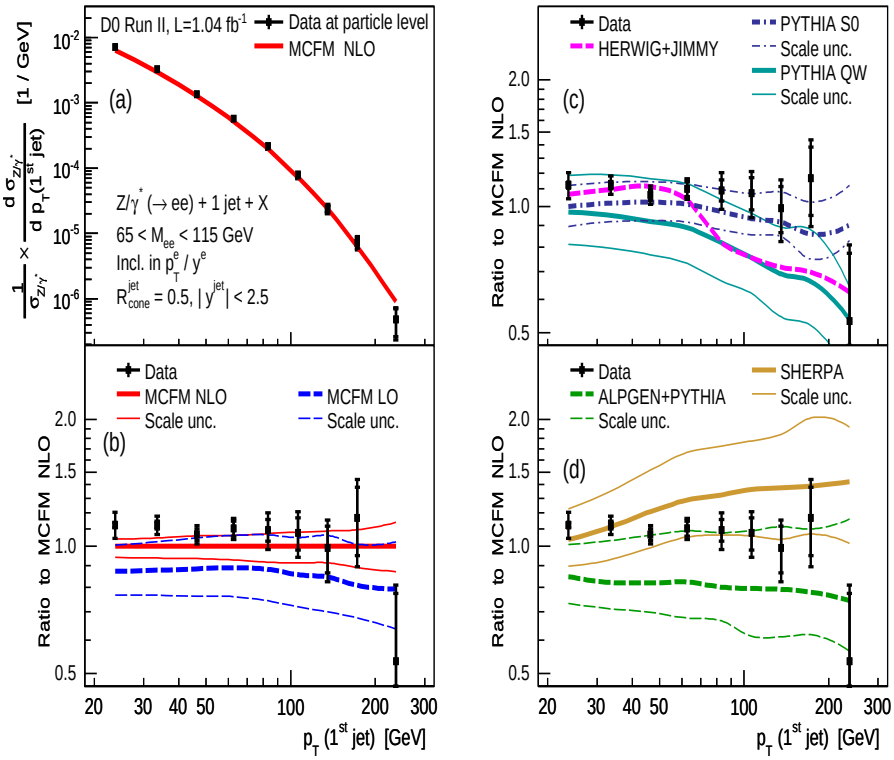


Figure E.5: The measured distribution of $\frac{1}{\sigma_{Z/\gamma^*}} \times \frac{d\sigma}{dp_T(jet)}$ for the leading jet in Z/γ^*+ jet+X events, compared to the predictions of MCFM NLO (a). The ratios of data and theory predictions to MCFM NLO are shown for pQCD predictions corrected to the particle level (b), for three parton-shower event generator models (c) and for two event generators matching matrix-elements to a parton shower (d).

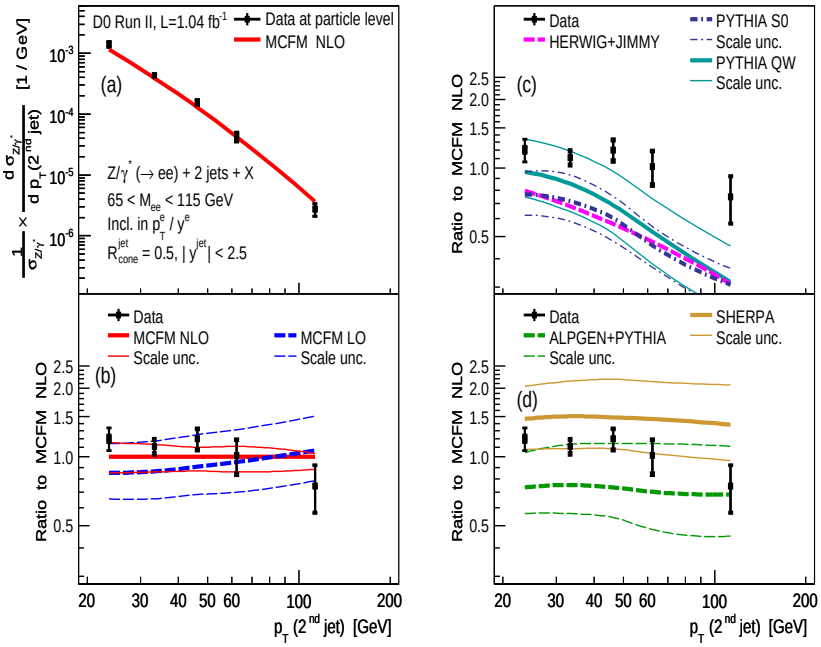


Figure E.6: The measured distribution of $\frac{1}{\sigma_{Z/\gamma^*}} \times \frac{d\sigma}{dp_T(jet)}$ for the second jet in Z/γ^*+ 2 jets+X events, compared to the predictions of MCFM NLO (a). The ratios of data and theory predictions to MCFM NLO are shown for pQCD predictions corrected to the particle level (b), for three parton-shower event generator models (c) and for two event generators matching matrix-elements to a parton shower (d).

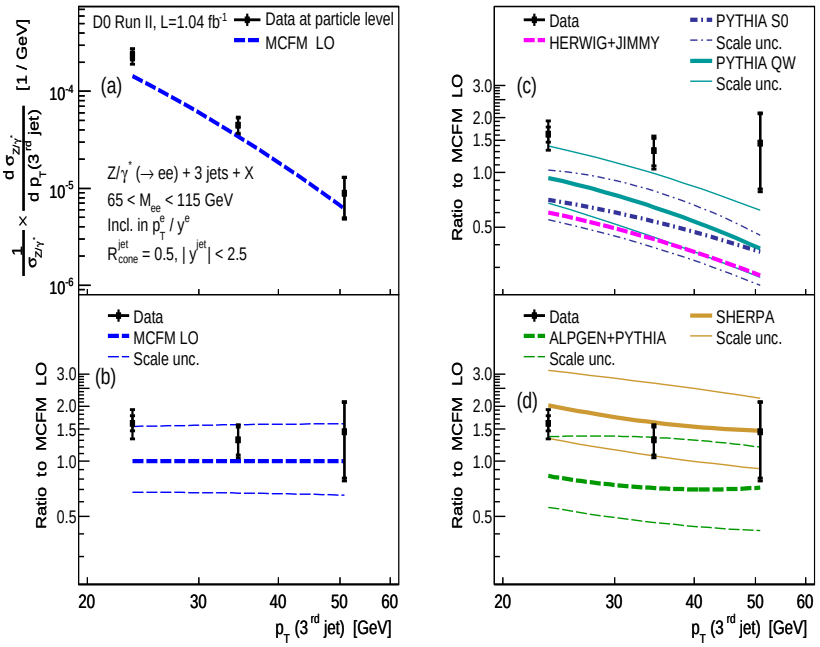


Figure E.7: The measured distribution of $\frac{1}{\sigma_{Z/\gamma^*}} \times \frac{d\sigma}{dp_T(jet)}$ for the third jet in Z/γ^*+ 3 jets+X events, compared to the predictions of MCFM NLO (a). The ratios of data and theory predictions to MCFM NLO are shown for pQCD predictions corrected to the particle level (b), for three parton-shower event generator models (c) and for two event generators matching matrix-elements to a parton shower (d).

Bibliography

[1] "LHC web page". http://lhc.web.cern.ch/lhc/.

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

[3] The ATLAS Collaboration. "The ATLAS experiment at the CERN Large Hadron Collider. Chapter 4: Inner Detector". *JINST*, 3(S08003):79–135, 2008.

[4] K.Pretzl. "Calorimeters in Astro and Particle physics". *J.Phys.*, G31(R133-R149), 2005. arXiv:physics/0502065.

[5] The ATLAS Collaboration. "The ATLAS experiment at the CERN Large Hadron Collider. Chapter 5: Calorimetry". *JINST*, 3(S08003):136–187, 2008.

[6] The ATLAS Collaboration. "The ATLAS experiment at the CERN Large Hadron Collider. Chapter 6: Muons Spectrometer". *JINST*, 3(S08003):188–229, 2008.

[7] G. Duckeck *et al.* (ATLAS Collaboration). "ATLAS computing: Technical design report". 2005. CERN-LHCC-2005-022.

[8] S. Agostinelli *et al.* (GEANT4 Collaboration). "Geant4: A simulation toolkit". *Nucl.Instrum.Meth.*, A506:250–303, 2003.

[9] I. Bird *et al.* "LHC computing Grid. Technical design report". 2005. CERN-LHCC-2005-024.

[10] ATLAS Collaboration. "ATLAS detector and physics performance. Technical design report. Vol. 2". 1999. CERN-LHCC-99-15.

[11] Ian J.R. Aitchison. "Supersymmetry and the MSSM: An elementary introduction". 2005. arXiv:hep-ph/0505105.

[12] P.Ramond. "Supersymmetry in physics: An algebraic overview". *Physica*, 15D(25), 1985.

[13] G.Brooijmans (CDF and DO Collaboration). "Top quark mass measurements at the Tevatron". 2000. arXiv:hep-ex/0005030.

[14] G. D'Agostini and G.Degrassi. "Constraining the Higgs boson mass through the combination of direct search and precision measurement results". 2000. arXiv:hep-ph/0001269.

[15] "Summary of the analysis of the 19 September 2008 incident at LHC". http://press.web.cern.ch/press/PressReleases/Releases2008/PR14.08E.html. CERN.

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

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

[18] A.Salam. *"Elementary Particle Theory"*. Stockholm, 1968.

[19] C.Amsler *et al.* (Particle Data Group collaboration). "Review of Particle Physics". *Physics Letters*, B667(1), 2008. http://pdg.lbl.gov/2008.

[20] P.W.Higgs. "Broken symmetries, massless particles and gauge fields". *Phys.Letters*, 12:132–133, July 1964. http://www.slac.stanford.edu/spires/find/hep/www?irn=3313930.

[21] P.W.Higgs. "Spontaneous symmetry breakdown without massless bosons". *Phys.Rev.*, 145:1156–1163, May 1966. http://prola.aps.org/abstract/PR/v145/i4/p1156_1.

[22] J.R.Ellis. "Beyond the standard model with the LHC". *Nature*, 448:297–301, July 2007.

[23] D.J.Gross and F.Wilczek. "Ultraviolet Behavior of Non-Abelian Gauge Theories". *Phys.Rev.Lett.*, 30:1343–1346, 1973.

[24] H.D.Politzer. "Reliable Perturbative Results for Strong Interactions?". *Phys.Rev.Lett.*, 30:1346–1349, 1973.

[25] "The Coordinated Theoretical-Experimental Project on QCD". http://www.phys.psu.edu/~cteq.

[26] A.Martin, R.Roberts, J.Stirling, and R.Thorne. "MRST group". http://durpdg.dur.ac.uk/hepdata/mrs.html.

[27] S. Chekanov and others (ZEUS Collaboration). "An NLO QCD analysis of inclusive cross-section and jet-production data from the ZEUS experiment". *Eur.Phys.J.*, C42:1–16, 2005. arXiv:hep-ph/0503274.

[28] J.M. Campbell and R.K. Ellis. "Next-to-leading order corrections to W+2jet and Z+2jet production at hadron colliders". *Phys.Rev.*, D65:113007, 2002. arXiv:hep-ph/0202176.

[29] W.T.Giele, E.W.N.Glover, and D.A.Kosower. "Higher order correction to jet cross-sections in hadron colliders". *Nucl.Phys.*, B(403):633–670, 1993. arXiv:hep-ph/9302225.

[30] S.Hoche *et al.* "Matching Parton Showers and Matrix Elements". 2006. arXiv:hep-ph/0602031.

[31] T.Sjöstrand. "Monte Carlo Generators". 2006. arXiv:hep-ph/0611247.

[32] S.Frixione and B.R.Webber. "Matching NLO QCD computations and parton shower simulation". *JHEP*, 06(029), 2002. arXiv:hep-ph/0204244.

[33] B.R.Webber. "A QCD model for jet fragmentation including soft gluon interference". *Nuclear Physics*, B(238):492–528, September 1984.

[34] G.Marchesini and B.R.Webber. "Monte Carlo Simulation of General Hard processes with coherent QCD radiation". *Nuclear Physics*, B310:461–526, 1988.

[35] D.Amati and G.Veneziano. "Preconfinement as a property of perturbative QCD". *Phys.Lett.*, B83(87), 1979.

[36] A.Bassetto, M.Ciafaloni, and G.Marchesini. "Color Singlet Distributions And Mass Damping In Perturbative QCD". *Phys.Lett.*, B83(207), 1979.

[37] G.Marchesini, L.Trentadue, and G.Veneziano. "Space-time description of colour screening via jet calculus techniques". *Nucl.Phys.*, B181(335), 1981.

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

[39] H.Bengtsson and T.Sjöstrand. "The Lund Monte Carlo for hadronic processes: PYTHIA version 4.8. *Computer Physics Communications*, 46:43–82, 1987.

[40] M.L.Mangano, M.Moretti, F.Piccinini, R.Pittau, and A.D.Polosa. "ALPGEN, a generator for hard multiparton processes in hadronic collisions". *JHEP*, 07:001, 2003. arXiv:hep-ph/0206293.

[41] T.Sjöstrand, L.Lonbland, S.Mrenna, and P.Skands. "PYTHIA 6.3 Physics and Manual". 2003. arXiv:hep-ph/0308153.

[42] T.Sjostrand, S.Mrenna, and P.Skands. "PYTHIA 6.4 Physics and Manual". *JHEP*, 05:026, 2006. arXiv:hep-ph/0603175.

[43] C.M.Buttar, D.Clements, I.Dawson, and A.Moraes. "Simulations of minimum bias events and the underlying event, MC tuning and predictions for the LHC". *Acta Phys. Polon.*, B35(1):433–441, 2004.

[44] A.Moraes, C.Buttar, and I.Dawson. "Prediction for minimum bias and the underlying event at LHC energies". *Eur.Phys.J.*, C50(2):435–466, April 2007.

[45] A.Moraes. "Modeling the underlying event: generating predictions for the LHC". April 2009. ATL-PHYS-PROC-2009-045.

[46] G.Corcella *et al.* "HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)". *JHEP*, 01:010, 2001. arXiv:hep-ph/0011363.

[47] W.M. Yao *et al* (Particle Data Group). "Review of Particle Physics". *J.Phys.*, G33:1–1232, 2006.

[48] S.Frixione and B.R.Webber. "Matching NLO QCD computations and parton shower simulations". *JHEP*, 06:029, 2002. arXiv:hep-ph/0204244.

[49] S.Frixione, P.Nason, and B.R.Webber. "Matching NLO QCD and parton showers in heavy flavour production". *JHEP*, 08:007, 2003. arXiv:hep-ph/0305252.

[50] G.C.Blazey *et al.* "Run II jet physics". 2000. arXiv:hep-ex/0005012.

[51] S.D.Ellis, J.Huston, K.Hatakeyama, P.Loch, and M.Tönnesmann. "Jets in Hadron-Hadron Collisions". *Prog. Part. Nucl. Phys.*, 60:484–551, 2008. arXiv:hep-ph/0712.2447.

[52] C.Buttar *et al.* "Standard Model Handles and Candles Working Group, Tools and Jets Summary Report: Issues in Jet Physics". 2008. arXiv:hep-ph/0803.0678.

[53] J.Wess and B.Zumino. "Supergauge Transformations in Four-Dimension". *Nucl.Phys.*, B70:39–50, 1974.

[54] P.Fayet and S.Ferrara. "Supersymmetry". *Phys.Rept.*, 32:249–334, 1977.

[55] H.P.Nilles. "Supersymmetry, Supergravity and Particle Physics". *Phys.Rept.*, 110:1–162, 1984.

[56] H.E.Haber and G.L.Kane. "The Search for Supersymmetry: Probing Physics Beyond the Standard Model". *Phys.Rept.*, 117(2-4):75–263, January 1985.

[57] R.Barbieri. "Looking beyond the Standard Model: the supersymmetric option". *Riv. Nuovo Cimento*, 11(4):1–45, 1988.

[58] G.Burdman. "New Solutions to the Hierarchy Problem". *Braz. J. Phys.*, 37(2B):506–513, June 2007. arXiv:hep-ph/0703194, `www.sbfisica.org.br/bjp/files/v37_506.pdf`.

[59] W.de Boer and C.Sander. "Global electroweak fits and gauge coupling unification". *Phys.Lett.*, B585:276–286, 2004. arXiv:hep-ph/0307049.

[60] J.R.Ellis, S.Heinemeyer, K.A.Olive, and G.Weiglein. "Indirect sensitivities to the scale of supersymmetry". *JHEP*, 02:013, 2005. arXiv:hep-ph/0411216.

[61] H.Baer, C.H.Chen, F.Paige, and X.Tera. "Signals for minimal supergravity at the CERN large hadron collider: Multi-jet plus missing energy channel". *Phys.Rev.*, D52:2746–2759, 1995. arXiv:hep-ph/9503271.

[62] A.Chamseddine, R.Arnowitt, and P.Nath. "Locally Supersymmetric Grand Unification". *Phys.Rev.Lett.*, 49:970, 1982.

[63] R.Barbieri, S.Ferrara, and C.Savoy. "Gauge Models with Spontaneously Broken Local Supersymmetry". *Phys.Lett.*, B119:343, 1982.

[64] N.Ohta. "Grand Unified Theories Based on Local Supersymmetry". *Prog.Theor.Phys.*, 70(2):542–549, 1983.

[65] L.Hall, J.Lykken, and S.Weinberg. "Supergravity as the Messenger of Supersymmetry Breaking". *Phys.Rev.*, D27:2359–2378, 1983.

[66] L.Randall and R.Sundrum. "Out of this world supersymmetry breaking". *Nucl.Phys.*, B557:79–118, 1999. arXiv:hep-th/9810155.

[67] G.F.Giudice, M.A.Luty, H.Murayama, and R.Rattazzi. "Gaugino Mass without Singlets". *JHEP*, 12:027, 1998. arXiv:hep-ph/9810442.

[68] M.Dine and A.Nelson. "Dynamical supersymmetry breaking at low-energy". *Phys.Rev.*, D48:1277–1287, 1993.

[69] M.Dine, A.Nelson, Y.Nir, and Y.Shirman. "New tools for low-energy dynamical supersymmetry breaking". *Phys.Rev.*, D53:2658–2669, 1996. arXiv:hep-ph/9507378.

[70] F.A. Berends, H. Kuijf, B. Tausk, and W.T. Giele. "On the production of a W and jets at hadron colliders". *Nucl.Phys.*, B357:32–64, 1991.

[71] F. A. Berends, W. T. Giele, H. Kuijf, R. Kleiss, and W. J. Stirling. ''multi-jet production in w,z events at p anti-p colliders". *Phys.Lett.*, B224:237–242, 1989.

[72] R. K. Ellis and J. Campbell. "Next-to-leading order corrections to W + 2 jet and Z + 2 jet production at hadron colliders". *Phys.Rev.*, D65:113007, 2002. arXiv:hep-ph/0202176.

[73] "Les Houches Accord PDF Interface". `http://hepforge.cedar.ac.uk/lhapdf/`.

[74] M.R.Whalley, D.Bourilkov, and R.C.Group. ''the les houches accord pdf interface (lhapdf) and lhaglue". `http://projects.hepforge.org/lhapdf/`, 2005. arXiv:hep-ph/0508110.

[75] J.M.Campbell, J.W.Huston, and W.J.Stirling. "Hard interactions of quarks and gluons: A primer for LHC physics". *Rept. Prog. Phys.*, 70(89), 2007. arXiv:hep-ph/0611148.

[76] J.Pumplin, D.R.Stump, J.Huston, H.L.Lai, Pavel M.Nadolsky, and W.K.Tung. "New generation for parton distributions with uncertainties from global QCD analysis". *JHEP*, 07(012):44, Jan 2002. arXiv:hep-ph/0201195.

[77] J.Alwall *et al.* "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions". *Eur.Phys.J.*, C53:473–500, 2008.

[78] V.M. Abazov *et al.* D0 Collaboration. "The Upgraded D0 Detector". *Nucl. Instrum. Meth.*, A565(463-537), 2006. arXiv:physics/0507191.

[79] W.Lampl *et al.* "Calorimeter clustering algorithms: Description and performance. CERN, 2008. ATL-LARG-PUB-2008-002.

[80] ATLAS Collaboration. "Expected Performance of the ATLAS Experiment: Detector, Trigger and Physics. Chapter 3: Electrons and Photons". page 43, Geneva, 2008. CERN-OPEN-2008-020, arXiv:0901.0512.

[81] ATLAS Collaboration. "Expected performance of the ATLAS Experiment, Detector, Trigger and Physics. Chapter 6: Jets and Missing Transverse Energy (Detector Level Jet Corrections)". *to appear*, page 261, Geneva, 2008. CERN-OPEN-2008-020.

[82] M.Cacciari and C.P.Salam. "Dispelling the N^3 myth for the k_t jet-finder". *Phys.Lett.*, B641:57–61, 2006. arXiv:hep-ph/0512210.

[83] M.Cacciari, G.P.Salam, and G.Soyez. "The anti-k_T jet clustering algorithm". *JHEP*, 04(063), 2008. arXiv:hep-ph/0802.1189.

[84] G.P.Salam and G.Soyez. "A practical seedless infrared-safe cone jet algorithm". *JHEP*, 05:086, 2007.

[85] M.H.Seymour. "Jet shapes in hadron collisions: Higher orders, resummation and hadronization". *Nucl.Phys.*, B513:269–300, 1998. arXiv:hep-ph/9707338.

[86] C.Schwanenberger (H1 Collaboration). "The jet calibration in the H1 liquid argon calorimeter". 2002. arXiv:physics/0209026.

[87] ATLAS Collaboration. "Expected Performance of the ATLAS Experiment: Detector, Trigger and Physics. Chapter 9: Standard Model (Electroweak Boson Cross-Section Measurement)". page 723, Geneva, 2008. CERN-OPEN-2008-020, arXiv:0901.0512.

[88] M.Schott. "Study of the Z Boson Production at the ATLAS Experiment with First Data". *ATLAS Thesis*, June 2007.

[89] ATLAS Collaboration. "Expected Performance of the ATLAS Experiment: Detector, Trigger and Physics. Chapter 10: Top Quark (Determination of the Top Quark Pair Production Cross-Section)". page 869, Geneva, 2008. CERN-OPEN-2008-020, arXiv:0901.0512.

[90] CDF Collaboration. "CDF Run II: Technical Design Report". 1996. FERMILAB-PUB-96-390-E.

[91] "Fermilab Beam Division". `http://www-bd.fnal.gov`.

[92] J.Campbell, R.K.Ellis, and D.L.Rainwater. "Next-to-leading order QCD predictions for W+2j and Z+2j production at the CERN LHC". *Phys.Rev.*, D68(094021), 2003. arXiv:hep-ph/0308195.

[93] A.Dotti, A.Gupta, F.Paige, C.Roda, and I.Vivarelli. "Jet Reconstruction and Calibration". *ATLAS Calorimeter Calibration Workshop*, Barcelona 2006.

[94] CDF Collaboration. "Measurement of Inclusive Jet Cross Sections in $Z/\gamma^\star(\rightarrow e^+e^-)+$ jets Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV". *Phys.Rev.Lett.*, 100(102001), March 2008. arXiv:hep-ex/0711.3717.

[95] M.D'Onofrio, M.Martinez, and O.Saltó. "Mesurement of Inclusive Jet Cross Sections in $Z/\gamma^\star(\rightarrow e^+e^-)+$ jets Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV". October 2008. arXiv:hep-ex:0810.3642.

[96] A.Abulencia *et al.* (CDF Collaboration). "Measurement of the inclusive jet cross section in $p\bar{p}$ interactions at $\sqrt{s} = 1.96$ TeV using a cone-based jet algorithm". *Phys.Rev.*, D74:071103, 2006.

[97] V.M.Abazov *et al* (D0 Collaboration). "Measurement of the ratios of the $Z/\gamma^* + \geq n$ jet production cross sections to the total inclusive Z/γ^* cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV". *Phys.Lett.*, B658(112-119), 2008. arXiv:hep-ex/0608052.

[98] F.Maltone and T.Stelzer. "MadEvent: Automatic event generation with MadGraph". *JHEP*, 02(027), 2003. arXiv:hep-ph/0208156.

[99] D0 Collaboration. "Measurements of differential cross sections of $Z/\gamma^* + jets + X$ events in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV". *(submitted to PLB)*, 2009. FERMILAB-PUB-09/066-E, arXiv:hep-ex/0903.1748.

Acknowledgments / Agraïments

This thesis has been possible thanks to many people from different countries which I have met during these last years of my life. I know I am not able to include all them here, but I would like to thank them all.

Firstly, I would like to express my gratitude to three women that have been essential to this thesis: Martine, Ulla and Lluïsa. Martine that during these years has been patient with me, teaching me physics, giving me many opportunities to learn, and contributing deeply to this thesis. Ulla that with her energy and efficiency has given an important push to this analysis and has contribute to this study. And Lluïsa that has helped me with this thesis, correcting up to the smallest detail and trying to make the whole understandable.

I also would like to give my thanks to people from ATLAS group at IFAE, who have contributed to my physics knowledge: Ilya, Luca, Matteo Cavalli-Sforza, Mario Martinez,... And also to IFAE people with who I have shared many lunches, coffees, many nice discussions, cinema sessions, mountains and excursions, SAF climbing sessions, cultural and non-cultural activities: Sigrid, Carlos, Nuria i Gabriel, Ana, José, Neus, Pilar, Carol, Francesc, Matteo, Manel, Diego, Ignasi,... and the "upstairs" people, including Pere, Juanjo and Germano.

I also want to thank the people from CERN and from McGill, where I have spent nice days and I have learned many things.

And last, but not least, people who have asked me about my thesis, many times: *"quan acabes i tornes a venir a la muntanya?"* or *"com va la tesi? ja l'has acabat?"* and gave me "ànims" to finish it: els Capdemunts, els Minis, i amics d'aquí i d'allà, m'agradaria dir a tots ells: gràcies! (ja m'estic posant les piles per tornar a fer muntanya i fer sopars!)

Finalment, els meus agraïments als meus pares i també a la meva germana que sempre m'han encoratjat a treballar en aquesta tesi, a lluitar, i a finalitzar aquesta aventura (per comenar-ne una de nova!): moltes gràcies! I també un petonet als meus dos nebodets, la Sàvila i el Sàndal, que són dos petits solets que donen sempre calor i alegria.

Ester Segura