THE UNIVERSITY OF CHICAGO

MEASUREMENT OF JET PRODUCTION IN ASSOCIATION WITH A W BOSON AT ATLAS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

CERN-THESIS-2011-330
08/11/2011

BY
IMAI JEN–LA PLANTE

CHICAGO, ILLINOIS
DECEMBER 2011

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The completion of this thesis is thanks to my advisor, Jim Pilcher. He guided me to work on interesting topics and made sure the studies converged. He kept me focused and motivated, and I am grateful to have his support.

Thanks to all the members of my thesis committee. Young-Kee Kim made herself available, from convincing me to choose the University of Chicago to scheduling the defense. Jonathan Rosner provided helpful feedback, and Cheng Chin was kind enough to participate late at night from across the ocean.

I received tremendous help from the University of Chicago ATLAS group. The first version of this analysis was published thanks to Monica Dunford. Maria Fiascaris' dedication and enthusiasm have inspired me to keep going at the toughest moments. Sasha Paramonov is a one man physics army, and I am lucky he has been on my side.

Thanks to my ATLAS colleagues, particularly those I worked with on TileCal and on W+jets. Martina Hurwitz is a great trailblazer. Kelby Anderson and the staff of the Electronics Design Group were generous with their time and experience. I learned a lot from collaborating with Hugo Beauchemin, Erik Devetak, Brian Martin, and Nicola Venturi. I would gladly share another office with Eric Feng, Hyeon Jin Kim, Peter Onyisi, Rishi Pravahan, and Jae Yu.

I had excellent academic support on my way to graduate school. The teachers at Bellevue College fueled my excitement about science, especially Robert Hobbs and Larry Susanka. Lupine Miller and the NASA Space Grant staff made me feel welcome at the

University of Washington even before the first day. I am proud to be a Dovie and thank Pam Samuelson and Bob Glushko for their encouragement.

I learned as much from undergraduate research at the University of Washington as from classes. Thanks to David Cobden, Henry Lubatti, Mark Tuttle, and Gordon Watts for supervising my projects. Henry introduced me to the beauty of statistics, the machine shop, and high energy physics. I am grateful for his guidance.

Studying abroad changed the way I see the world. I thank Martin and Kathleen Richardson at the University of Central Florida for that opportunity. Anis Jouini and Janelle Shane were patient with my non-existent French vocabulary and supportive of my research at the Laboratoire de Physico-Chimie des Matériaux Luminescents in Lyon.

Many friendships have sustained me during this journey. Thanks to Wojtek Fedorko and the University of Chicago sailing club for making the first years fly. With the camaraderie of my fellow graduate students, I had fun even at the most difficult times. During the long wait for data, Esteban Fullana, Belen Salvachua, Giulio Usai, and Matteo Volpi kept things interesting. Ceri Perkins gave valuable insight on life at CERN. When the data arrived and work got crazy, my friends from Achilles International provided an outlet and an inspiration. I am fortunate to have such good friends.

The love of my family always puts things in perspective. Brian and my sister Ilan, thank you for reminding me that science is fun. I am glad to be in almost the same time zone now, and I know we will stay connected at any distance. Miguel, thank you for improving the state of my world and challenging me to think differently. The joy of being together is the greatest discovery of my LHC years.

Most of all I am grateful to my parents, David and Rita, my first teachers and best supporters. They have given me countless opportunities and the freedom to explore. All of my successes are thanks to them.

ABSTRACT

This study examines the associated production of jets with a W boson, identified by the decay $W \rightarrow e\nu$. The measurement uses the full data sample collected by the ATLAS experiment during the inaugural year of proton-proton collisions with center-of-mass energy $\sqrt{s} = 7$ TeV at the Large Hadron Collider. Cross sections are presented for $W \rightarrow e\nu$ events with each jet multiplicity from zero through four or more jets and as a function of jet transverse momenta, the scalar sum of jet and lepton transverse momenta, the invariant mass of jets, and angular distributions. The results represent the first experimental test of perturbative QCD predictions with next-to-leading order precision for $W + 4$ jets.

CHAPTER 1

INTRODUCTION

The study of particle physics made significant progress in the late 20th century with the development and experimental confirmation of the Standard Model. This powerful framework describes with remarkable precision the observations that have been made to date, yet it is limited and incomplete [2]. The agreement between theory and experiment is both encouraging, as it proves the usefulness of the model, and frustrating, as it gives few clues to the deeper structure of nature.

The Large Hadron Collider (LHC) was designed and constructed over a period of more than 10 years to unravel the mystery of what more fundamental physics could exist. To do this, it collides protons at energies almost an order of magnitude higher than those previously produced in a laboratory. The resulting interactions are carefully examined and compared to theoretical prediction to look for any differences.

A key part of the program is the measurement of the production of quarks and gluons in association with a $W^{\pm}$ boson, commonly known as W+jets, which is the subject of this thesis. The measurement tests the performance of the experimental apparatus and the understanding of physics modeling at the new collision energies. Once verified, this information can be used to study even more complex and unusual processes within the Standard Model and to search for new ones.

Together with the Z boson, the $W^{\pm}$ bosons are the weak force carriers of the Standard Model. While much less powerful than the strong force that holds together atomic nuclei and the electromagnetic force between electrically charged particles, the weak force is no less important. For example, radioactive decays converting atoms of one element or isotope to another occur through the emission of a virtual W boson.

Figure 1.1 [3] shows the masses, charges, and spins of the $W^{\pm}$ and Z bosons and the other elementary particles of the Standard Model. The remaining bosons, or spin one particles, carry the strong (gluon) and electromagnetic (photon) forces. The fermions, or spin 1/2 particles, are the quarks, leptons, and their antiparticles, which can all interact with W bosons. The electron is perhaps the most familiar, as the carrier of electric charge in circuits. Once produced, the W boson quickly decays to either a quark and antiquark (*e.g.* $W \rightarrow \bar{u}d$) or a pair of leptons, including a neutrino or antineutrino (*e.g.* $W \rightarrow e\bar{\nu}$).

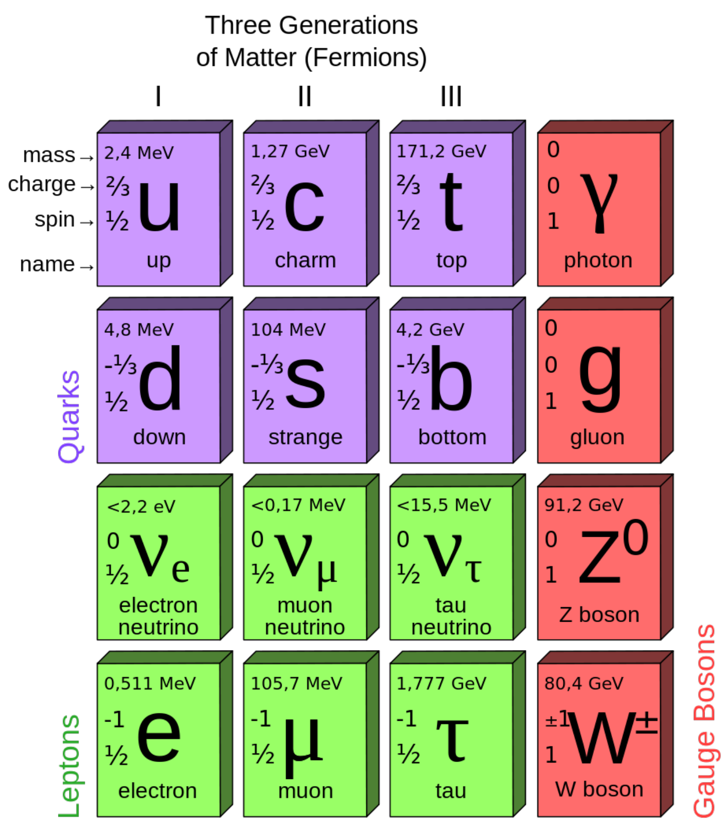

Figure 1.1: Elementary particles of the Standard Model.

In this analysis, the W boson is identified by its decay to an electron and antineutrino, or positron (antielectron) and neutrino, which occurs for about one in nine W bosons that are produced.[1] This gives a clear event signature, as the neutrino is identified by an imbalance of momentum in the transverse plane, perpendicular to the direction of the colliding protons. This quantity, conventionally called missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, is calculated by taking the negative vector sum of all measured momenta in the transverse plane. To satisfy momentum conservation, any remainder is attributed to the neutrino, the only Standard Model particle that can pass through the detector without depositing measurable energy.

Events containing a W boson are then classified by the number of accompanying quarks and gluons, collectively called partons. The partons cannot be observed directly, but produce collimated sprays of energetic particles called jets. Measuring the production rates and properties of the jets in these events gives a precise test of Standard Model predictions that rely on sophisticated theoretical and numerical techniques.

Perturbative QCD (pQCD) predictions for W+jets have recently been extended to next-to-leading order (NLO) precision in the strong coupling constant, α_s, for W+4 jets [1], reducing the uncertainty by about half. Figure 1.2 shows an example from that publication of one of the virtual diagrams evaluated in the calculation. The five final-state objects make this a groundbreaking achievement. Testing the validity of such tools at the LHC is essential to understanding new observations there.

The measurement described here uses all of the proton-proton collision data collected during the inaugural year of LHC operation, at an unprecedented center-of-mass energy of $\sqrt{s} = 7$ TeV. An earlier version of the analysis using about 4% of the data has already

1. As charge is not considered in this analysis, "electron" is used for either an electron or positron, "neutrino" for a neutrino or antineutrino, and "W" for both the $W^{\pm}$ bosons.

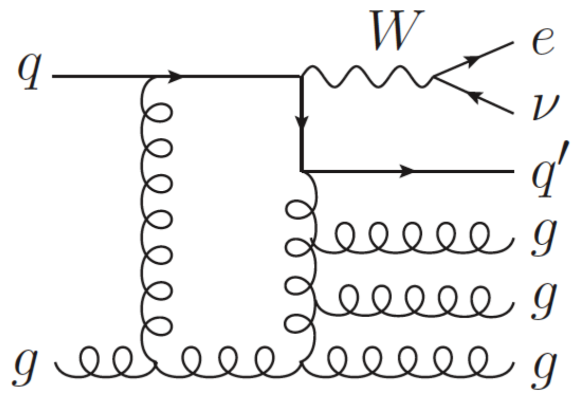

Figure 1.2: Virtual diagram for the production of a W boson plus four partons, with the W decaying to an electron and neutrino.

been published [4]. Another measurement of W+jets at the LHC is in progress [5].

W+jets measurements at lower collision energies have been reported by the CDF [6, 7] and D0 [8] collaborations and compared to NLO predictions up to W+3 jets [9]. The agreement observed has built confidence in the predictions. However, the challenges grow with each additional parton, and this measurement provides the first comparison of experimental results with the NLO prediction for W+4 jets.

This thesis includes several different types of results. Inclusive cross sections are the rates at which different numbers of jets are produced in association with a W boson. These results are first directly calculated from the numbers of events observed. They are then corrected for experimental effects such as the resolution of jets in the detector, to obtain the rates at which the particles are produced, known as particle-level cross sections. Detector independence makes these useful to compare with theoretical predictions and results from other experiments. The production rates as a function of different observables such as jet direction and momentum, collectively called differential cross sections, are also presented both before and after correction.

The inclusive and differential cross sections are compared to predictions from Monte Carlo event generators and show exceptional agreement. The particle-level cross sections are also compared to NLO predictions up to W+4 jets. As noted above, this represents the first experimental test of those predictions.

CHAPTER 2

EXPERIMENT

Physics at the energy frontier justifies pioneering experimental apparatus. This chapter describes that equipment. The main features of the LHC and its remarkable performance during early operation are presented first, followed by an overview of the ATLAS detector, highlighting the parts essential to this analysis.

2.1 The Large Hadron Collider

The world's largest machine is what makes this measurement possible. The LHC [10] combines recycled infrastructure at CERN with new technologies to deliver proton-proton collisions at energies over three times higher than at previous hadron colliders. As its own prototype, it has exceeded performance expectations during the first years of data taking, albeit slightly behind schedule.

The design of the LHC takes advantage of a 27-km-long tunnel used previously by the Large Electron Positron collider, at an average depth of about 100 m underground. While often approximated as a ring, the tunnel consists of sixteen alternately curved and straight sections. Each straight section includes an interaction point where the beams can be made to collide. Four of the eight possible interaction points are used for experiments.

ATLAS [11], the detector used for this measurement, is one of those four main experiments. It is described in the next section. CMS [12], the experiment located on the opposite side of the LHC, shares its physics goals with ATLAS. They are both sometimes called general-purpose detectors as they are designed to study a wide variety of particles and interactions [13]. Each has its own strengths, and the two experimental teams challenge each other to reach the physics goals first [14]. For collision data they rely on the smooth operation of the LHC and its supporting infrastructure.

The injection complex, which is shown schematically in Fig. 2.1 [15], includes the Super Proton Synchrotron (SPS), which provided collisions for the original discovery of the W and Z bosons [16]. Earlier stages in the chain, the Linac, Booster, and Proton Synchrotron, sequentially accelerate protons to 25 GeV. Finally, the SPS accelerates them to 450 GeV, the minimum energy at which they circulate in the LHC. The protons are injected and circulated in 0.2 ns bunches, which are spaced apart by 25 ns.[1]

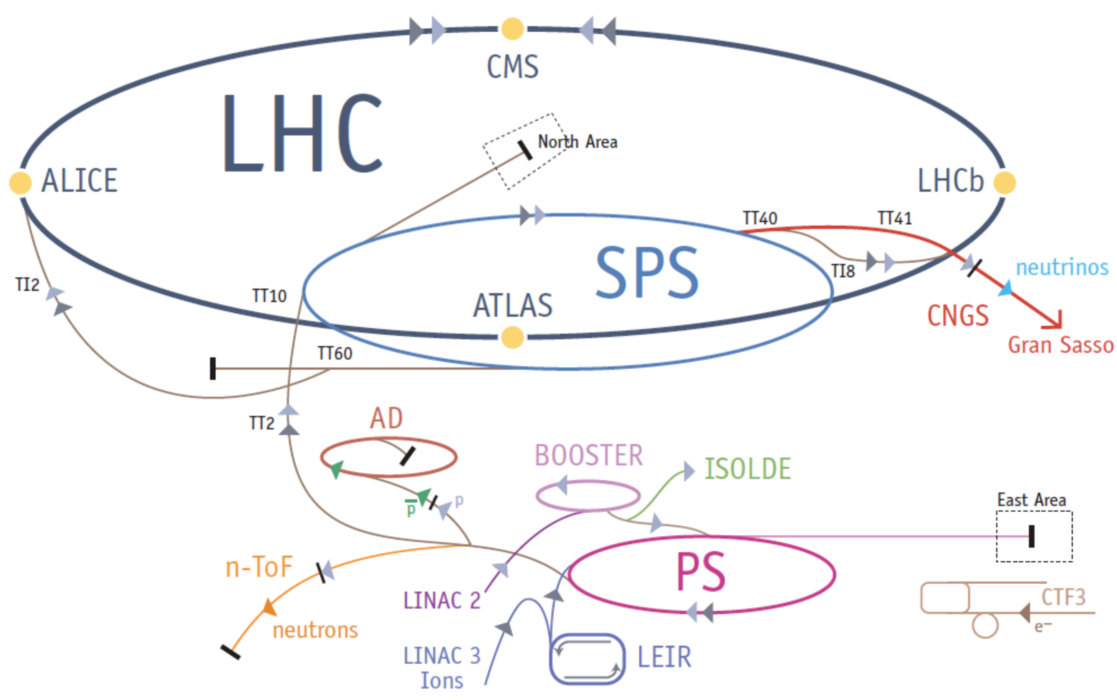

Figure 2.1: Schematic of the LHC injection complex.

1. Not all bunches were filled during early running. The nearest bunch spacing in 2010 was 150 ns. An approximate bunch length from 2010 is given for comparison.

The two counter-circulating proton beams are then ramped to their full collision energy. For the early data used in this analysis, the beams are each accelerated to 3.5 TeV, half the design energy, for collisions at center-of-mass energy $\sqrt{s} = 7$ TeV. This milestone was first achieved on March 30, 2010. Figure 2.2 shows one of the first collision events from that day, recorded with the ATLAS detector.

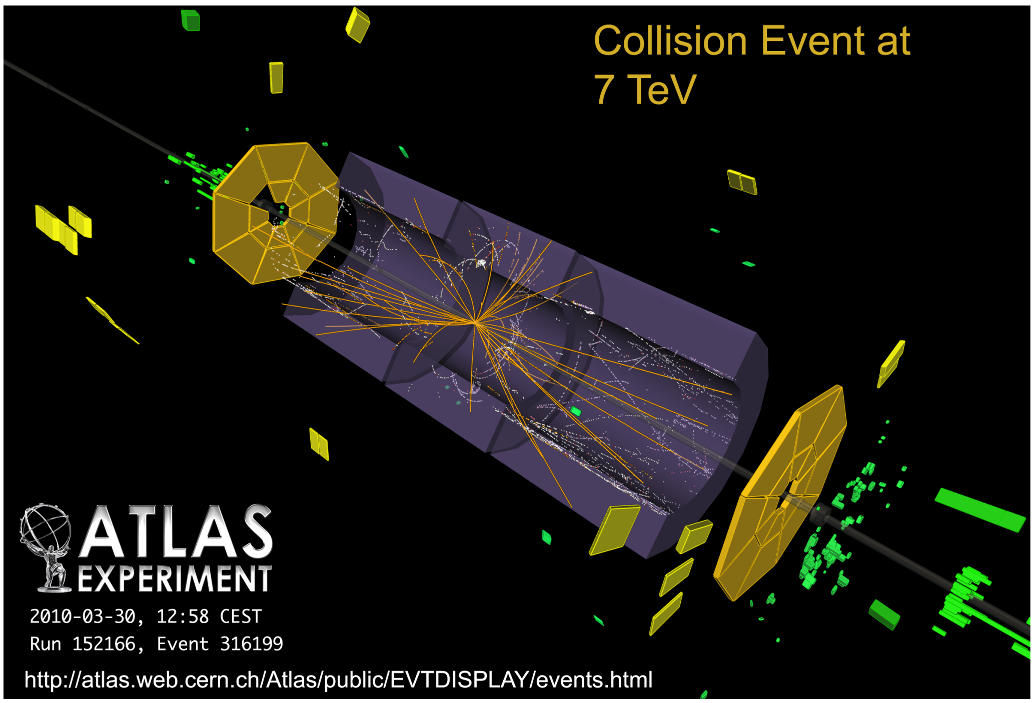

Figure 2.2: One of the first 7 TeV proton-proton collisions recorded at the LHC.

The accumulation of data by ATLAS during 2010 is shown in Fig. 2.3. As the LHC performance, in terms of both the collision rate and fraction of the time when it delivered collisions, grew rapidly during the year, the vertical axis is shown on a logarithmic scale. A measure of this performance is the rate of events produced per unit cross section, called the instantaneous luminosity. The total amount of data is characterized by the integrated luminosity, which is obtained by integrating the instantaneous luminosity over time and

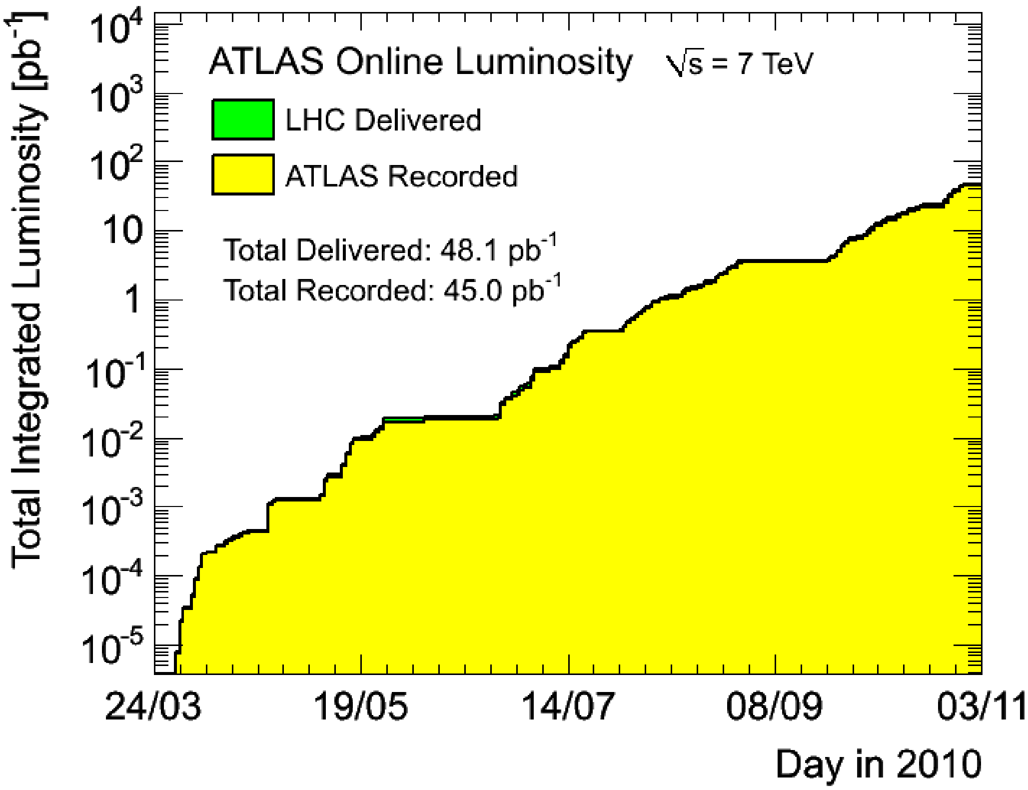

Figure 2.3: Timeline of the total integrated luminosity collected by ATLAS during 2010.

is expressed in units of inverse cross section.[2]

The steep slope of the integrated luminosity over time shows the growing confidence in machine operation during the year. At the design operating parameters, the energy stored in each beam of the LHC will reach 350 MJ. Notably, this is enough to melt 500 kg of copper [15]. During normal operations, the stored energy in the beam is released into a dedicated absorber called a beam dump at the end of data-taking periods. However, if control of the beam is lost, the amount of stored energy could determine the level

2. Although not necessarily related to area, particle interaction probabilities are called cross sections and are measured in barns (1 barn = 10^{-28} m^2). Since this unit is large for the processes observed at the LHC, picobarns (1 pb = 10^{-12} barns) are typically used.

of damage to the machine. Balancing the risk of such an occurrence with the goal of producing as many collisions as possible, the stored energy in the beam, and thus the instantaneous luminosity, was increased gradually during 2010, by relaxing limits on the numbers of protons per bunch and occupied bunches.

Also satisfying in Fig. 2.3 is the difference between the yellow histogram showing data recorded by ATLAS and the green histogram showing data delivered by the LHC, which is barely distinguishable from it. Over 93% of the data delivered by the LHC was recorded by ATLAS. A small fraction of the delivered data was not recorded, mainly due to runs discarded due to problems with individual parts of the detector and time spent starting and stopping data-taking. Some additional data is removed by the application of quality criteria (see Section 3.1).

The luminosities shown in Fig. 2.3 are from an initial estimate based on the LHC operating parameters. This had an uncertainty of $\pm 11\%$ [17]. It was later improved to a more precise measurement with an uncertainty of $\pm 3.4\%$, accompanied by a 3.6% decrease of the central value, using Van der Meer scans in which the beams were displaced in the transverse plane to determine their width and height [18].

The rapid acceleration of data-taking during 2010 boosted the size of the data sample. It extended the kinematic reach of the cross section measurement and enhanced the statistics for measurements of the differential distributions. However, the improvements were accompanied by changes in the operating conditions, producing a phenomenon called pileup that needed to be taken into account in the analysis.

Pileup describes the superposition of additional proton-proton interactions on the collision being studied. It is classified into in-time and out-of-time pileup. In-time pileup describes the interaction of additional protons in the same bunch crossing, while out-of-time pileup refers to the effects of interactions in nearby bunch crossings. Pileup is an

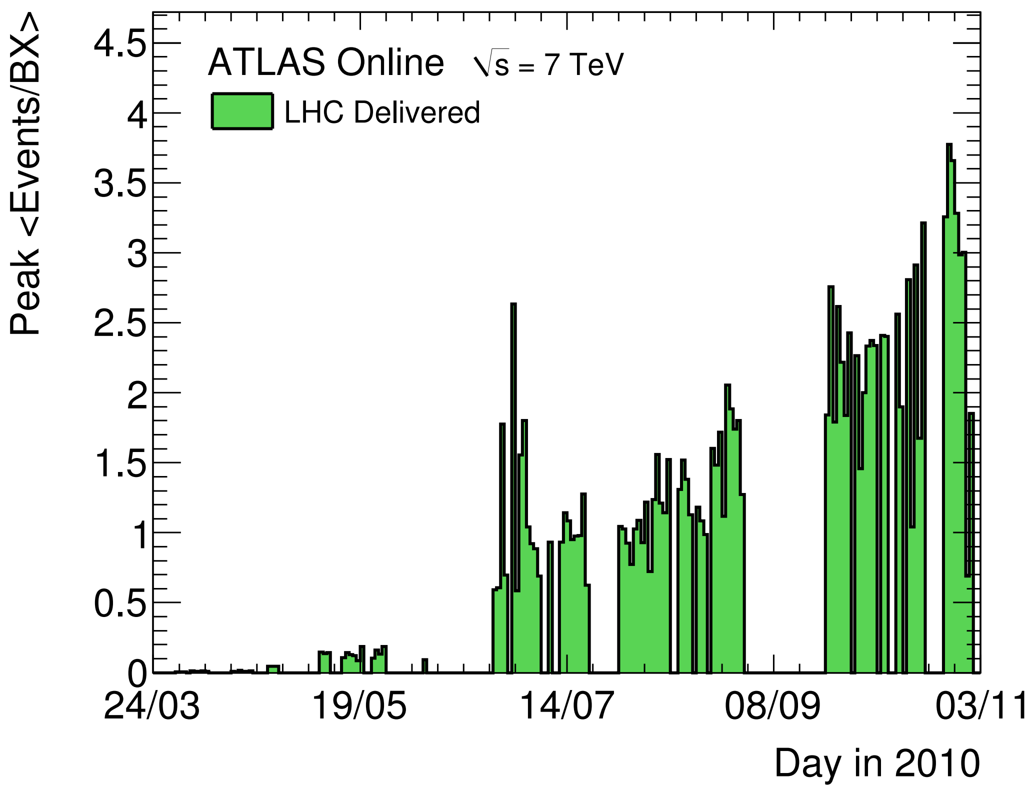

Figure 2.4: Timeline of the peak number of interactions per bunch crossing during 2010.

expected side effect of the improving LHC performance.

In-time pileup increased because the number of protons per bunch grew from about 10^{10} to 10^{11} during the year. Raising the number of protons per bunch significantly increased the final data sample because the instantaneous luminosity scales quadratically with this parameter. However, it also increased the probability for multiple protons to interact in the same bunch crossing.

Figure 2.4 shows the peak number of interactions per bunch crossing as a function of time. Comparing with Fig. 2.3, most of the data was recorded when the peak number of interactions was high. For example, the total integrated luminosity approximately doubled in the last two weeks of data taking. The average for the whole data sample is about two interactions per bunch crossing.

Another improvement came from circulating additional pairs of colliding bunches, which linearly increases the instantaneous luminosity. To protect the LHC as noted above, initially only a few bunches were circulated at any given time, with the occupancy increasing during the year to a minimum bunch spacing of 150 ns. The proximity of colliding bunches can cause out-of-time pileup due to residual signal in the detector electronics during the time window captured for the next event. However, with the bunch spacing well above the design spacing of 25 ns, there was little out-of-time pileup in 2010.

The pileup corrections applied in this analysis are described in Sections 3.2.3 and 3.6.5. As the LHC approaches its design operating parameters, both in-time and out-of-time pileup will continue to increase. Eventually the peak instantaneous luminosity is expected to reach 10^{34} cm^{-2} s^{-1}, compared to the maximum of 2 x 10^{32} cm^{-2} s^{-1} achieved in 2010. With pileup becoming more significant, this measurement using only data with little pileup will provide an important benchmark for future studies.

2.2 The ATLAS detector

The smooth operation of the LHC in 2010 set the stage for data-taking. No less important was the readiness of the ATLAS (A Toroidal LHC ApparatuS) detector. After years of design, construction, and tests, the expected performance and understanding of the detector [11] were established with the first proton-proton collisions at 0.9 and 2.36 TeV [19]. As the measurement reported here demonstrates, ATLAS was well prepared for the arrival of 7 TeV collisions.

The detector consists of several complementary subsystems shown in Fig. 2.5 [11]. The key parts for this measurement are described in the following sections. Together

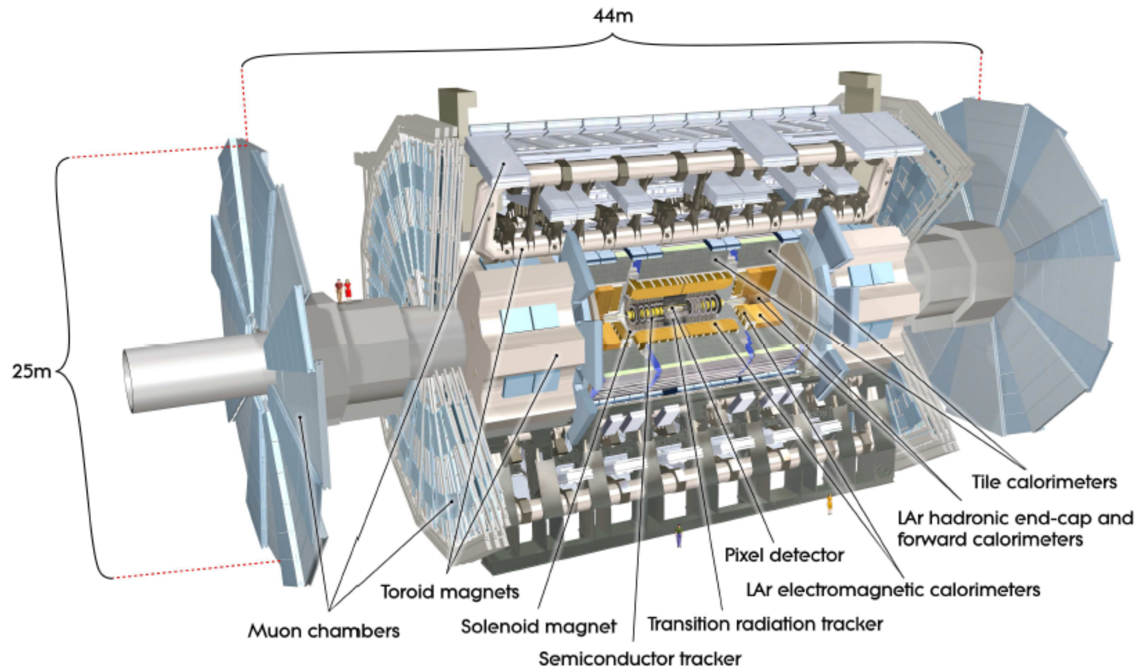

Figure 2.5: Schematic of the ATLAS detector.

they provide a comprehensive view of each proton-proton collision event, including the identification of different particles and measurements of their kinematic properties.

Measurements are performed with respect to a right-handed coordinate system, with the origin defined as the nominal *pp* interaction point at the center of the detector. The positive x-axis is defined by the horizontal direction from the interaction point to the centre of the LHC ring, with the positive y-axis pointing upwards, while the counterclockwise beam direction as seen from above defines the z-axis. The azimuthal angle ϕ is measured around the beam axis and the polar angle θ is the angle from the z-axis. The Lorentz invariant rapidity and pseudorapidity are defined as $y = 0.5 \times \ln[(E + p_z)/(E - p_z)]$ and $\eta = -\ln\tan(\theta/2)$, respectively.

2.2.1 Inner detector

The inner detector (ID) is particularly important to this analysis for the identification of electrons, as discussed in Chapter 3. It provides the first position and momentum measurements for charged particles coming from the interaction point. With the innermost layer just 50.5 mm from the center of the beam pipe, the ID must withstand high radiation levels and contribute as little material as possible. These requirements are reflected in the design of the electronics and choice of low density materials including carbon fiber.

Figure 2.6 [11] shows the different technologies that make up the ID. Working outward from the beam pipe, the system consists of the pixel detector, the semiconductor tracker (SCT), and the transition radiation tracker (TRT). These subsystems are all immersed in a uniform magnetic field parallel to the beam axis, with a strength of 2T. They provide transverse momentum resolution $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 0.05\%\ p_{\mathrm{T}} \oplus 1\%$ for charged tracks with $p_{\mathrm{T}} > 0.5$ GeV and pseudorapidity $|\eta| < 2.5$.

The pixel detector [20] consists of three concentric cylinders of solid-state sensors (barrel) and three identical disks of sensors on each side (endcap), perpendicular to the beam. There are 1744 pixel sensors, each with 46080 individual readout channels, for a total of about 80 million pixels. The fine granularity of the pixel detector is crucial to the vertex reconstruction and impact parameter resolution achieved at ATLAS.

Each sensor is machined to a thickness of 250 μm and has an active area of $50 \times 400\ \mu\mathrm{m}^2$. Charged particles ionize when passing through the silicon bulk material of a sensor. The signal is picked up and transferred to the front-end electronics via a low-resistance bump bond. Position information is given by the location of the pixel in the sensor and of the sensor itself. Transverse momentum is determined by reconstructing curved tracks from individual pixel hits, in combination with the other ID subsystems.

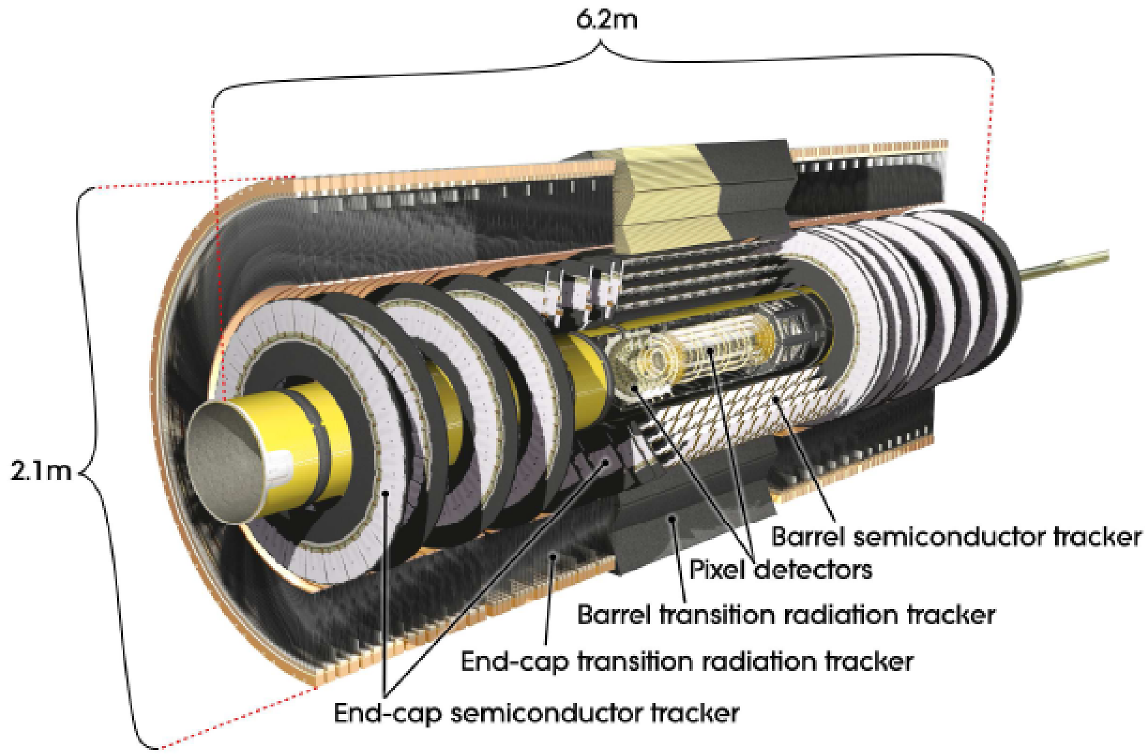

Figure 2.6: Layout of the inner detector.

For the cylindrical barrel sensors the position resolution is about 10 μm in $R - \phi$ space, where R is the radial distance from the center of the detector, and 115 μm longitudinally.

Surrounding the pixel detector are the silicon sensors of the SCT, covering four concentric cylinders in the barrel region and nine disks in each endcap, again perpendicular to the beam. The sensors are arranged in double layers of strips, with a crossing angle of 40 mrad between them to form stereo pairs. The crossing points form the SCT hits used in track reconstruction. There are typically at least four SCT hits per track, each with a resolution of about 17 μm in $R - \phi$ space and 580 μm longitudinally.

The TRT is a unique feature of the ATLAS tracking system compared to the other general-purpose LHC experiment, CMS. This outermost layer of the ID extends to pseudorapidity $|\eta| < 2.0$. It uses the transition radiation produced by relativistic charged particles passing between materials with different dielectric constants to distinguish particles of different mass, such as electrons and pions. Figure 2.7 shows an example of this discrimination using 2009 data and simulation as described in [19]. The probability of hits on tracks that are above a high voltage threshold depends on the Lorentz factor, $\gamma = \mathrm{E/mc}^2$, where E is the energy of a charged particle and m is its mass. Thus the fraction of high-threshold hits is greater for electrons (mass 0.5 MeV), even at low momenta, than for pions (mass 140 MeV).

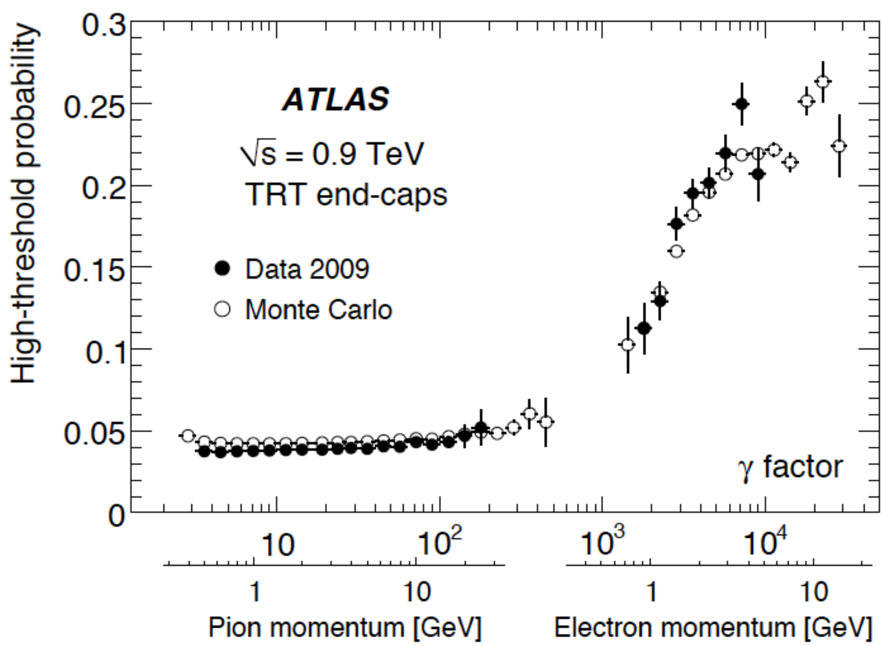

Figure 2.7: Discrimination of electrons and pions in the transition radiation tracker.

Thin layers of transition material alternate with gas-filled drift tubes called straws in the TRT. The straws have a diameter of just 4 mm. The position of the wire inside each drift tube is known to within 50 μm, giving a hit resolution of 130 μm in $R-\phi$ space. There are over 7 million straws, arranged such that charged particles typically cross 36 straws, giving tracking information as well as particle identification based on the size of the transition radiation signal.

2.2.2 Calorimeters

The next layers of the detector are the calorimeters. Figure 2.8 [11] shows their relative positions. Together they allow precise energy measurement over a large range of energies and contribute to particle identification. For this analysis, the identification and energy measurement of electrons (see Section 3.4) and the reconstruction of jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ (see Sections 3.5 and 3.6) using the calorimeters are particularly important.

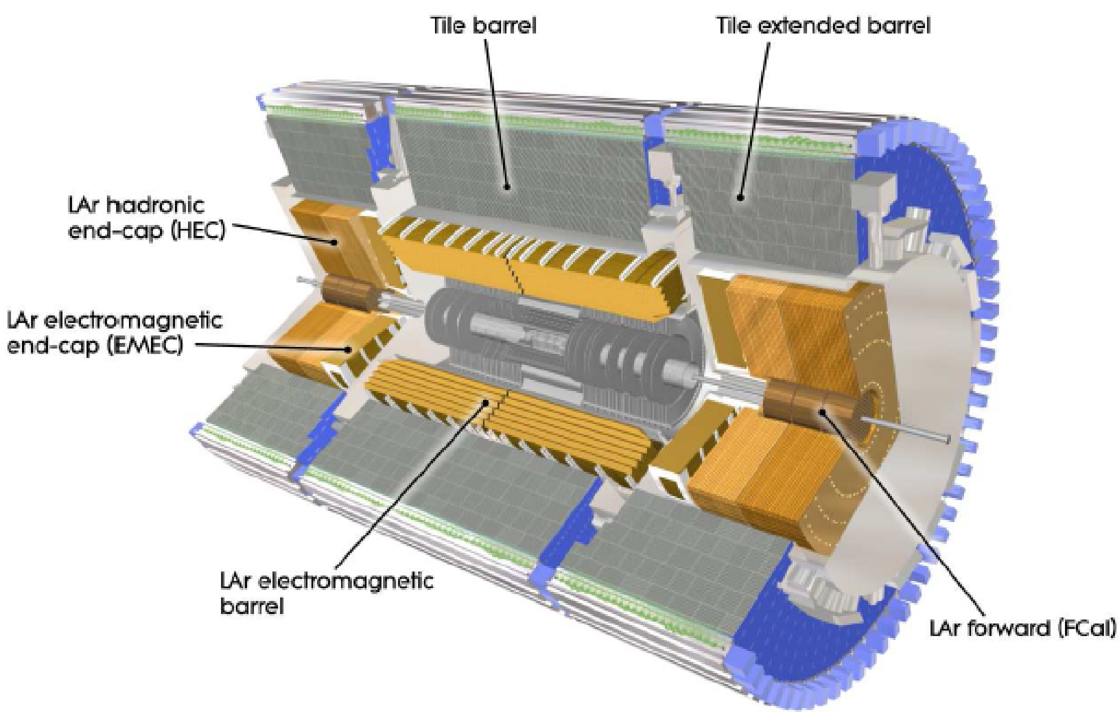

Figure 2.8: Layout of the calorimeter system.

ATLAS uses sampling calorimeters, with alternating layers of passive material absorbing the energy of particle showers and active material where the local number of charged particles is measured. Thus only a fraction of the deposited energy is sampled. The calorimeters are also non-compensating, with a varying response to EM or hadronic energy deposition. These effects are taken into account through calibration, *e.g.*, using measurements from portions of the detector exposed to beams of known energy [21], and by correcting the measured jet energies as discussed in Section 3.6.2.

The thickness of the calorimeters, measured in units of interaction length, is key to contain the showers of even highly energetic particles and measure their energy. Figure 2.9 from Ref. [11] shows this thickness as a function of pseudorapidity, $|\eta|$, with the different layers indicated. They are symmetric around the beam axis in ϕ and about the origin in η. The thickness of the electromagnetic (EM) calorimeter, better expressed in radiation lengths for electromagnetic processes, X_0, is at least 22 to 24 X_0.

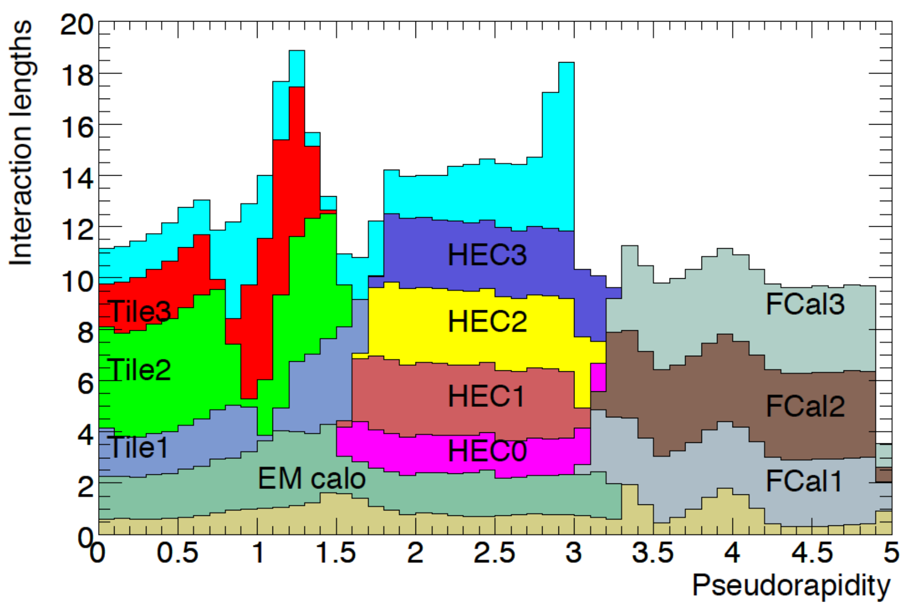

Figure 2.9: Cumulative amount of material in the calorimeters, in units of interaction length. Colors indicate the different layers, which are described in the text.

Liquid argon [22] is used as the active material in the EM calorimeter and forward calorimeter (FCAL), as well as the hadronic endcap calorimeter (HEC). The EM calorimeter covers pseudorapidity $|\eta| < 3.2$, with a typical energy resolution of $\sigma_E/E = 10\% \ / \ \sqrt{E} \oplus 0.7\%$. It includes the EM barrel and two endcaps shown in Fig. 2.8. In the region with ID tracking, $|\eta| < 2.5$, the EM calorimeter has a granularity as fine as 0.025×0.025 in $\Delta\eta$ x $\Delta\phi$ and three depth layers. This allows precise shower shape measurements for electron identification. Beyond this region, the granularity increases to 0.1×0.1.

The hadronic calorimeters are designed to provide a jet resolution of $\sigma_E/E = 50\% \ / \ \sqrt{E} \oplus 3\%$. The central region is covered by the Tile Calorimeter (TileCal), using lead and scintillating plastic tiles. As illustrated in Fig. 2.8, the TileCal consists of four cylinders, placed symmetrically with two on each side of the interaction point. The inner sections (barrel) cover $|\eta| < 0.8$ and the outer sections (extended barrel) cover $0.8 < |\eta| < 1.7$. Each section is segmented into three depth layers, with a $\Delta\eta$ x $\Delta\phi$ granularity of 0.1×0.1 in the inner layers and 0.2×0.1 in the last layer. The scintillating tiles are oriented perpendicular to the beam and interleaved with steel providing the lead absorber. Scintillation light is delivered to photomultiplier tubes via wavelength shifting fibers.

Dedicated hadronic energy measurement is extended to higher pseudorapidities by the HEC, which covers $1.5 < |\eta| < 2.5$ with a granularity of 0.1×0.1 and $2.5 < |\eta| < 3.2$ with a granularity of 0.2×0.2 in $\Delta\eta$ x $\Delta\phi$. It also uses liquid argon as the active material, with copper as the absorber. The thickness of the four HEC layers is shown in Fig. 2.9.

The forward calorimeter (FCAL) provides lower granularity but coverage up to $|\eta| < 4.9$, which is particularly important for measuring the imbalance of momentum from a neutrino (see Section 3.5). Due to space limitations, it has a high density, with the first depth layer using copper absorber and the other two using tungsten. It measures the energy of both EM and hadronic interactions, with a design resolution of $\sigma_E/E = 100\% \ / \ \sqrt{E} \oplus 10\%$.

2.2.3 Trigger

At the LHC design luminosity, ATLAS could see about one billion collisions per second. Not all events are equally interesting and only a small fraction can be saved. The trigger system is designed to reduce the event rate from 40 MHz to about 200 Hz, while quickly sorting out the most promising events. Only events passing all three stages of the trigger are fully reconstructed and recorded for offline analysis.

The first stage, the Level-1 (L1) trigger, must cope with the highest event rate. To speed up event processing, it is a built-in part of the detector hardware and uses only some information directly from the calorimeters and the muon trigger chambers, rather than fully reconstructed objects. The calorimeter information is particularly important for this analysis, which primarily uses an electron trigger, with events selected by a photon or $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger as a cross-check. The likely presence of an electron, photon, or large $E_{\mathrm{T}}^{\mathrm{miss}}$ is determined from regions of the calorimeters passing energy thresholds. These Regions of Interest (RoIs) provide a starting point for the next trigger selection.

Events passing the L1 trigger reach the high-level trigger, which consists of the Level-2 (L2) trigger and the event filter. With the event rate reduced to less than 100 kHz by the L1 trigger, the L2 trigger has access to more detailed information. It reconstructs interesting objects using combined information from different subdetectors and can exploit more refined thresholds than L1. The event filter then applies additional reconstruction algorithms and calibration, for trigger selection based on objects that are very similar to those used in offline analysis. This provides a further reduction from about 3 kHz to the required 200 Hz rate at which events with an average size of about 1.3 Mbyte can be recorded.

CHAPTER 3

DATA & SELECTION

This chapter describes the events used for the analysis. The data consist of events collected during the 2010 run of the LHC and simulated events. Candidate W boson events decaying to an electron and neutrino are identified within these samples. The later sections of this chapter describe the selection criteria in detail, as well as the classification of the selected events by jet multiplicity.

3.1 Experimental data

This analysis uses all of the proton-proton collision data at center-of-mass energy $\sqrt{s} = 7$ TeV collected by ATLAS during 2010. Only runs where the LHC delivered stable beams with energies each greater than 3.4 TeV are considered. A standard set of requirements are placed on the status of different parts of the detector to ensure good quality data [23].

The ATLAS detector, solenoid and toroid magnets, inner detector vertexing, luminosity monitoring, and triggers are all required to be operational. Although muons are not considered in the event selection, the muon detection systems are required to be fully functioning. This is useful for comparing the measured cross sections with a parallel analysis of the $W \to \mu\nu$ channel. Good quality status is required for electron and jet reconstruction, in both central and forward regions, and for the reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$. These in turn require the relevant detector subsystems, particularly the calorimeters, to be fully operational. After all of these requirements, the dataset corresponds to an integrated luminosity of 36 pb^{-1}.

3.2 Simulated data

Signal and background events are simulated with the same center-of-mass energy as the data, $\sqrt{s}$ =7 TeV, using Monte Carlo methods [24]. The stable particles generated in each simulated sample are passed through the GEANT4 detector simulation [25, 26] and reconstructed using the same version of the ATHENA software framework [27] as for the data. This gives a prediction of how each process would be measured in the detector.

The parton interaction producing the W boson or background process is called hard scatter due to the relatively high momentum transfer. The remaining partons in the colliding protons can also interact. Such interactions and other QCD radiation accompanying the hard scatter are collectively called underlying event. These effects are included in the simulations, typically with modeling parameters adjusted using ATLAS measurements [28, 29].

Additional simulated collisions are overlaid with the signal and background samples to model in-time and out-of-time pileup. The out-of-time pileup corresponds to a bunch spacing of 150 ns, as for the majority of data. The in-time pileup corresponds to an average of 2.2 interactions per event. Each simulated event sample was then reweighted to match the distribution of the number of reconstructed primary vertices, a measure of pileup, to that obtained from data. The details of the reweighting are presented in Section 3.2.3.

3.2.1 Signal samples

The Monte Carlo event generators used to model the W+jets signal for this analysis are listed in Table 3.1. The samples are normalized 10.46 nb, the inclusive W cross section times $W \to e\nu$ branching ratio predicted with next-to-next-to-leading-order (NNLO) precision [30]. Chapter 7 discusses other theoretical predictions of the W+jets cross sections, which are compared with the measured results but not used in the analysis, and the uncertainty on the predictions.

Physics process	Generator
$W \rightarrow e\nu$ + jets ($0 \leq N_{\text{parton}} \leq 5$)	ALPGEN + HERWIG
$W \rightarrow e\nu$ + jets ($0 \leq N_{\text{parton}} \leq 5$)	SHERPA + HERWIG
$W \rightarrow e\nu$ inclusive	PYTHIA

Table 3.1: Samples of simulated signal events used in this analysis.

Various Monte Carlo programs are available for hadron collider physics, with different advantages as compared for example in [31]. The primary W+jets signal sample for this analysis is generated with ALPGEN v2.13 [32], which is particularly useful for final states with high jet multiplicity. It performs perturbative matrix element calculations to leading-order (LO) in the strong coupling constant, α_s, for $W + n$ partons ($n = 0 - 5$). The LO diagrams for $W + n$ partons include some that are NLO for $W + (n-1)$ partons [7]. Thus ALPGEN is sometimes called an enhanced leading-order Monte Carlo generator.

ALPGEN and other Monte Carlo event generators rely on parton distribution functions (PDFs) to model the structure of the colliding protons. This is discussed in more detail in Chapter 7. The CTEQ6L1 PDF set is used to generate the Alpgen samples [33] for this analysis.

Samples for W plus up to five additional partons are produced from matrix element calculations in ALPGEN and combined by their relative cross sections. The partons from the matrix element calculation are evolved to stable particles through fragmentation and hadronization, collectively called a parton shower, using another Monte Carlo simulation, HERWIG [34, 35]. Overlap between the matrix element calculation and the parton shower is removed according to the MLM matching scheme [36] using a p_T threshold of 20 GeV.

The primary W+jets sample generated with Alpgen is used throughout this analysis in different ways. In particular, it is used to correct the measured cross sections for detector effects such as resolution and efficiency, as described in Chapter 5. For comparison, a W+jets sample produced with a different Monte Carlo event generator, Sherpa [37], is also used. Like Alpgen, it combines matrix element calculations with parton showering, but using the CKKW matching scheme [36] instead of MLM and a higher p_{T} threshold of 30 GeV. Sherpa has its own parton showering and default underlying event tune. Another difference is the use of an NLO PDF set, CTEQ6.6M [38]. Comparing results using different simulations shows the impact of these choices.

The jet multiplicity spectrum of a third $W \to e\nu$ sample simulated with Pythia [39] is also considered. This sample includes only leading-order W production with no partons plus possible jets from parton showering, so it is not expected to model the higher jet multiplicities as well as Alpgen and Sherpa. However, it is convenient to study the impact of selection criteria on the inclusive sample since the events are unweighted. MRST 2007 LO* [40], a PDF set optimized for use with leading-order Monte Carlo event generators, is used.

3.2.2 Background samples

All background processes except QCD multijet production were modeled using Monte Carlo simulated events. The QCD background, from events where only jets are produced, presents special challenges and is estimated from data, as described in Chapter 4. The simulated background samples and their cross sections including branching ratios are listed in Table 3.2.

The $Z \to ee$ background process is simulated with Alpgen, just as for the primary $W \to e\nu$ sample. Background samples with tau leptons, $Z \to \tau\tau$ and $W \to \tau\nu$, are

Physics process	Generator	Cross section × BR (nb)
$Z \to ee$ ($m_Z > 40$ GeV)	ALPGEN + HERWIG	1.07
$Z \to \tau\tau$ ($m_Z > 60$ GeV)	PYTHIA	0.989
$W \to \tau\nu$	PYTHIA	10.46
Dibosons WW	HERWIG	44.9×10^{-3}
Dibosons WZ ($m_Z > 60$ GeV)	HERWIG	18.5×10^{-3}
Dibosons ZZ ($m_Z > 60$ GeV)	HERWIG	5.96×10^{-3}
$t\bar{t}$	POWHEG	0.165
Single-top $t \to \ell\nu q$ (s-channel)	MC@NLO + HERWIG	4.3×10^{-4}
Single-top $t \to \ell\nu q$ (t-channel)	MC@NLO + HERWIG	6.34×10^{-3}
Single-top (Wt)	MC@NLO + HERWIG	13.1×10^{-3}

Table 3.2: Simulated background samples used in this analysis.

simulated with PYTHIA in order to include tau decays using a dedicated program [41]. Like the signal, all of these sample are normalized to NNLO cross section predictions, with the values given in the table.

Backgrounds from diboson production, WW, WZ, and ZZ, are simulated using HERWIG. This leading-order parton shower simulation is sufficient as the diboson backgrounds are very small. The samples are normalized using cross sections obtained with the MC@NLO [42] Monte Carlo simulation, which includes next-to-leading order matrix elements. MC@NLO is also used to simulate and normalize the single-top samples. For the event simulation, the partons from the matrix element calculations are showered with HERWIG.

One drawback of MC@NLO is that it produces some events with negative weights. POWHEG [43], another Monte Carlo simulation with next-to-leading order precision that gives only positive event weights, is more convenient to model the $t\bar{t}$ background. For a precise normalization, the cross section is calculated including both next-to-leading order and next-to-next-to-leading-log corrections [44]. As $t\bar{t}$ is an important background at high jet multiplicity, special care is taken with the uncertainty (see Section 4.2). Additional

$t\bar{t}$ samples simulated with AcerMC [45] and showered with Pythia were used to assess the sensitivity to increased or decreased initial and final state radiation (ISR and FSR).

3.2.3 Pileup reweighting

Mismodeling of the pileup conditions in simulation could affect various observables including jet multiplicity. To match the average conditions, Monte Carlo simulated events are reweighted to match the multiplicity of in-time pileup interactions to that observed in data. A weight is applied to each simulated event depending on the number of primary vertices. As for the event selection, a vertex is counted if it has at least three tracks and is within 200 mm of the beam spot along the beam axis.

Figure 3.1 shows the distribution of the number of primary vertices in data (markers) and in the Monte Carlo sample simulated with Alpgen (blue). A pre-selection is applied, requiring events with at least one primary vertex, passing the quality criteria described in Section 3.1 in the case of data, and passing a single electron trigger. The data fall into two distinct periods with different triggers applied, as described in the next section.

The earlier period is shown at left in Fig. 3.1. It corresponds to just 0.9% of the data. The number of vertices peaks at one for the data, as there was very little pileup during this period. In contrast, the number of vertices peaks at two for the remainder of the data, as shown at right in Fig. 3.1. For both periods, the distributions in data and Monte Carlo are different, with closer agreement for the later period. The pileup weights correct this difference.

For each number of primary vertices, the pileup weight is the ratio of the fraction of events with that vertex multiplicity in data to the same fraction in Monte Carlo. The different weights for the two data-taking periods are shown in the second and third columns of Table 3.3. The final weights, obtained by combining those weights by their relative integrated luminosities, are shown in the fourth column of the table.

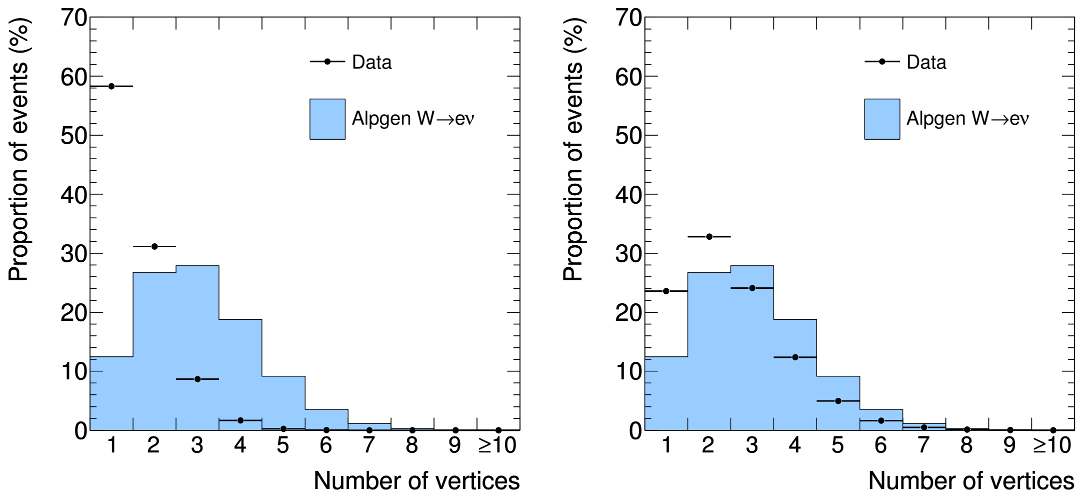

Figure 3.1: Number of reconstructed primary vertices in Monte Carlo simulated $W \to e\nu$ events before pileup reweighting compared to the first 0.9% of data from early runs (left) and to the remainder of the data collected (right).

Number of vertices	Weight A	Weight B	Weight applied
1	4.6864	1.8947	1.9477
2	1.1659	1.2282	1.2270
3	0.3107	0.8639	0.8534
4	0.0882	0.6588	0.6480
5	0.0267	0.5396	0.5299
6	0.0082	0.4613	0.4527
7	0.0028	0.4131	0.4053
8	0.0016	0.4066	0.3989
9	0.0010	0.4021	0.3945
$\geq$10	0.0009	0.3663	0.3594

Table 3.3: Weights used to match the pileup of simulated events to the data, as obtained with the first 0.9% of data from early runs (Weight A) and the remainder of the data collected (Weight B), and the resulting combined weights applied to the simulated events.

3.3 Event selection overview

This section summarizes the event selection and its impact on the data sample. The details of the electron selection follow. Section 3.5 describes the reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$, which is important both for the event selection and the estimation of the background from QCD events with one or more jets and no real W boson, as described in the next chapter. Finally, Section 3.6 is devoted to the discussion of jet counting.

Events are first selected using a single electron trigger. For early data runs, where the high-level trigger described in Section 2.2.3 was not yet activated, the L1 trigger is applied. The trigger requires a 0.1×0.1 region of the EM calorimeter in $\Delta\eta$ x $\Delta\phi$ with a transverse energy above a threshold corresponding to an offline p_{T} of approximately 15 GeV. The other 99% of data was collected with the high-level trigger, requiring a single electron with $p_{\mathrm{T}} > 15$ GeV satisfying identification criteria similar to those used offline (described as "medium" in Section 3.4). A simulation of the same trigger is applied to the Monte Carlo event samples.

Triggered events are further required to have at least one reconstructed primary vertex. To ensure that it is well-measured, the vertex must consist of at least three tracks. Its position along the beam axis must be within 200 mm of the beam spot, near the center of the detector.

Events passing the trigger and vertex criteria are identified as $W \rightarrow e\nu$ candidates by requiring exactly one electron with $p_{\mathrm{T}} > 20$ GeV (see Section 3.4) and significant $E_{\mathrm{T}}^{\mathrm{miss}}$, associated with the presence of a neutrino (see Section 3.5). The electron p_{T} and direction and $E_{\mathrm{T}}^{\mathrm{miss}}$ are then used to construct the transverse mass of the W candidate, attributing the $E_{\mathrm{T}}^{\mathrm{miss}}$ to the neutrino from the W decay, as follows:

$$M_{\mathrm{T}}(W) = \sqrt{2p_{\mathrm{T}}E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos[\phi(e) - \phi(E_{\mathrm{T}}^{\mathrm{miss}})])} \tag{3.1}$$

Estimating the W mass from momentum components in the transverse plane avoids assumptions necessary to infer the full mass, as discussed in Section 4.2. The $M_{\mathrm{T}}(W)$ distribution peaks near the W mass for events containing a W [46]. This peak is used to boost the signal-to-background ratio of the selected data, by requiring $M_{\mathrm{T}}(W) > 40$ GeV.

The peak is visible in Fig. 3.2, where the $M_{\mathrm{T}}(W)$ distribution of the data (black markers) is compared to expectations (stacked histograms) including the $W \to e\nu$ signal (white histogram). The QCD background is estimated from data (see Section 4.1, while the signal and remaining backgrounds are predicted using the Monte Carlo simulated events samples, normalized to the luminosity of the data sample. This figure also shows the increase in the background from top pair production (green histogram) as the jet multiplicity increases. Top events peak near the W mass as well since they contain a real W.

Table 3.4 shows the impact of the selection criteria on the data sample and on $W \to e\nu$ events simulated with Pythia. The efficiency of each criterion is determined from the fraction of simulated events passing the selection relative to the total number of $W \to e\nu$ events considered.

3.4 Electron selection

Electron candidates are reconstructed with information mainly from the EM calorimeter and inner detector. Reconstruction starts with 0.025 x 0.025 ($\Delta\eta$ x $\Delta\phi$) cells in the second layer of the EM calorimeter, where most of the energy of an electron is deposited. A 3×5 group of cells with at least one matching ID track forms the seed of a larger

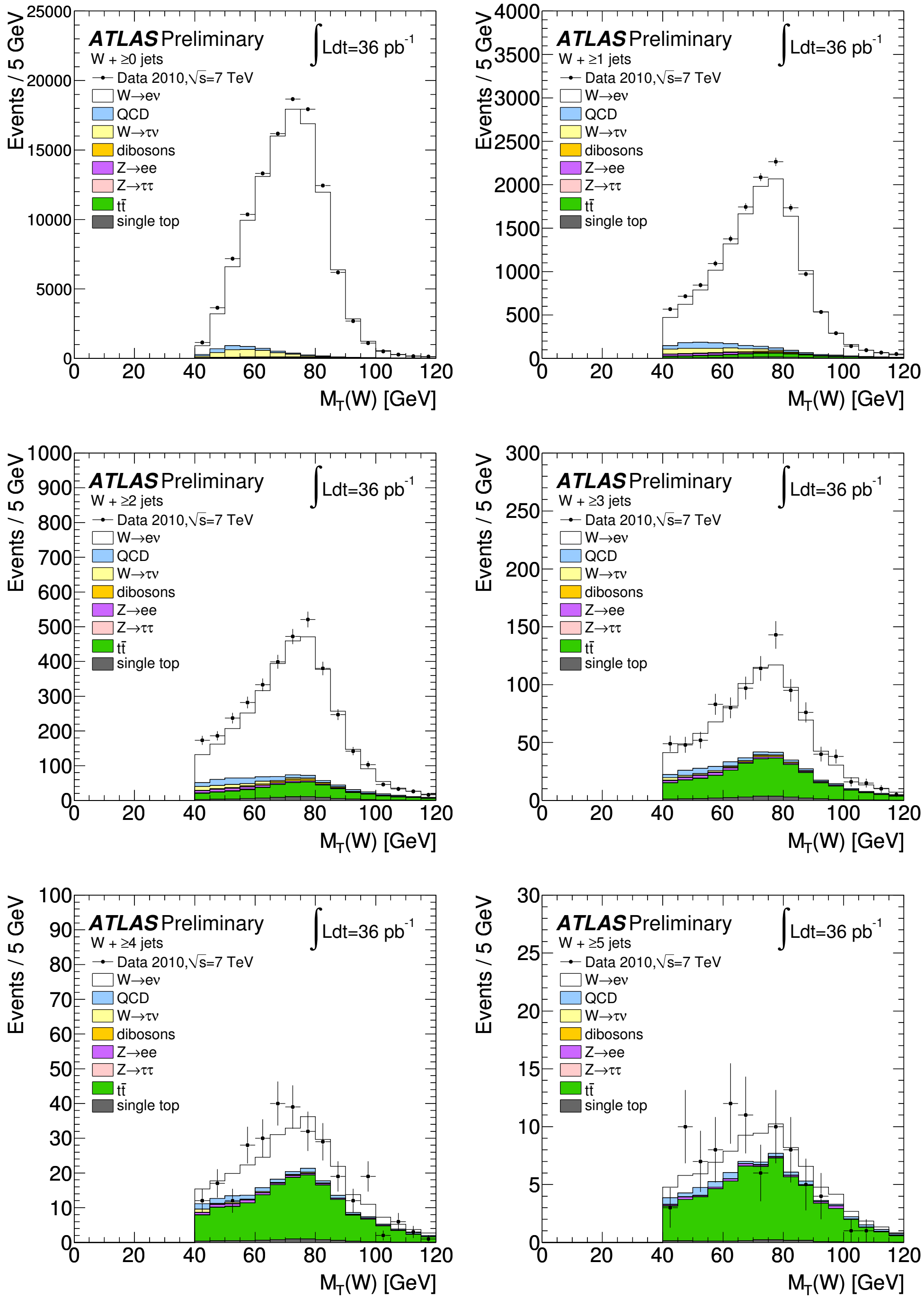

Figure 3.2: Comparison of the transverse mass distribution, $M_{\mathrm{T}}(W)$, between data and expectations, for inclusive jet multiplicities from zero (top left) through five (bottom right).

Selection criterion	Number of events	Efficiency (%)
Trigger	13096801	57.7
Elec η	12488739	55.5
OTX cleaning	12155032	53.1
Elec p_{T}	4961292	47.2
N. of elec	325914	35.8
$E_{\mathrm{T}}^{\mathrm{miss}}$cleaning	325223	35.8
$E_{\mathrm{T}}^{\mathrm{miss}}$	121923	28.3
Transverse mass	112333	27.8

Table 3.4: Total number of events in the data sample that pass the various selection criteria. The signal efficiency determined from simulated events is also shown.

cluster of calorimeter cells. The energy deposited in this cluster is corrected for lateral and longitudinal leakage of the electron energy beyond it, as described in Ref. [47]. The η and ϕ coordinates for the electron are taken from the matching track. The transverse energy calculated from those variables, $E_{\mathrm{T}} = \mathrm{E}_{\mathrm{cluster}}/\cosh(\eta_{\mathrm{track}})$, is required to be greater than 20 GeV.

To obtain an event sample in which the electron is measured by the ID tracking and is incident on the region of the electromagnetic calorimeter with the highest granularity, the electron pseudorapidity is required to be within $|\eta| < 2.47$. Any electron candidate with pseudorapidity between $1.37 < |\eta| < 1.52$ is excluded. Additional material in this transition region between the EM barrel and endcap calorimeters reduces the measurement precision for such candidates.

A few percent of cells in the EM calorimeter were not available for data-taking, mainly because of problems with the front-end readout electronics [48]. Electrons with clusters affected by these problems are removed, in a procedure described as OTX cleaning in Table 3.4. The same procedure is applied for simulated events, removing electrons in regions removed for data.

Electrons passing the kinematic cuts are also required to pass the identification and isolation criteria described in the following sections. An event must contain exactly one isolated electron passing the "tight" criteria and no second isolated candidate passing the more efficient "medium" criteria. Rejecting events with two electron candidates helps reduce backgrounds such as $Z \to ee$.

3.4.1 Identification

The electron identification criteria balance high efficiency with good rejection of jets faking electrons. Three categories of increasingly strict identification criteria are defined, called loose, medium, and tight. The medium definition includes all loose criteria and the tight definition includes medium. These criteria are listed in Table 3.5. Constant thresholds are indicated in the table, while others were optimized in bins of η and E_{T} as described in [47].

The loose criteria use only the middle layer of the EM calorimeter and additional information from the hadronic calorimeter. They have an expected jet rejection of about 500. The jet contamination in events selected by requiring a loose electron candidate is useful to study the properties of these jets for estimation of the QCD background (see Section 4.1.1).

No tracking information is included until the medium criteria, which also include requirements on the shower shape using the first layer of the EM calorimeter. A track matching the pseudorapidity of the EM cluster is required, separating electrons from photons. The track is required to have a minimum number of hits in the inner detector to give a reliable reconstruction.

The tight criteria are used to select the final electron candidate. They ensure a well-measured electron, and have an expected jet rejection about 100 times more powerful than the loose criteria. Strict requirements are placed on the agreement between measurements in the EM calorimeter and inner detector, and particle identification using transition radiation (see Section 2.2.1) is used. This lowers the typical efficiency for electrons from $W \rightarrow e\nu$ from above 90% with the medium criteria alone to just under 80%.

3.4.2 Isolation

An important background in this analysis arises from the misidentification of hadronic jets as electrons. Although only a small fraction of hadronic jets pass the electron selection, the QCD background cannot be neglected because the inclusive jet cross section is many orders of magnitude large than the W production cross section. The estimation of this background comes with large systematic uncertainties, as discussed in Section 4.1, so it is necessary to reduce it as much as possible.

The identification criteria described in the previous section limit the QCD background. To further reduce it, an additional calorimeter-based electron isolation is applied. It exploits the narrower width of EM showers compared to hadronic ones.

The electron isolation significantly reduces the QCD background, while retaining over 96% of signal events. It is defined by requiring the transverse energy in a cone of radius $R = 0.2$ around the electron, excluding the energy of the electron itself, to be less than 4 GeV. The impact of this requirement is illustrated in the next chapter.

3.4.3 Calibration & efficiency

The electron energy scale is calibrated using $Z \rightarrow ee$ events from data, with crosschecks using additional data, as described in [47]. $Z \rightarrow ee$ events are fit using the Z mass to extract the scale as a function of the measured electron energy in different η bins. The correction varies with η, ranging from about –2% to 4%. The width of the Z peak is also

Type	Description	Reversed
Loose		
Hadronic leakage	Ratio of E_{T} in the hadronic calorimeter (first layer only if $\lvert\eta\rvert < 0.8$ or $\lvert\eta\rvert > 1.37$) to E_{T} of the EM cluster	
Middle layer of EM calorimeter	Ratio of the energy in 3×7 cells over the energy in 7×7 cells centered at the electron cluster position	
	Lateral width of the shower	
Medium		
First layer of EM calorimeter	Total shower width	X
	Ratio of the energy difference between the largest and second largest energy deposits in the cluster over the sum of these energies	X
Track quality	Number of hits in the pixel detector (≥ 1)	
	Number of total hits in the pixel and SCT detectors (≥ 7)	
	Transverse impact parameter (< 5 mm)	X
Track - cluster matching	$\Delta\eta$ between the cluster position in the first EM layer and the extrapolated track (< 0.01)	
Tight		
Track - cluster matching	$\Delta\phi$ between the cluster position in the middle EM layer and the extrapolated track (< 0.02)	
	Ratio of the cluster energy to the track momentum	X
	Tighter $\Delta\eta$ requirement (< 0.005)	
Track quality	Tighter transverse impact parameter requirement (< 1 mm)	X
TRT	Total number of hits in the TRT	X
	Ratio of the number of high-threshold hits to the total number of hits in the TRT	X
Conversions	Number of hits in the innermost pixel layer (≥ 1)	
	Veto electron candidates matched to reconstructed photons	

Table 3.5: Electron identification criteria. The indicated subset of the criteria are reversed for selection of the QCD background control sample (see Section 4.1.1).

used to adjust the resolution in Monte Carlo simulated events to match the slightly worse resolution observed in data. Uncertainty on the electron energy scale and resolution and the resulting impact on $E_{\mathrm{T}}^{\mathrm{miss}}$ is included among the systematics uncertainties as discussed in Chapter 6.

The efficiency of the electron trigger, reconstruction, and identification, with isolation, is included when calculating the W+jets production cross sections. Any difference in electron observables between data and Monte Carlo could change the efficiency and bias the cross section measurements. This has been checked both by comparing the observables and by measuring the efficiency with data using tag-and-probe methods.

For trigger efficiency, the measured difference between data and Monte Carlo is just 0.5 $\pm$ 0.5%, which is applied as a correction to the data. The reconstruction efficiency is well modeled, and only the systematic uncertainty of 1.5% is derived from data. However, the identification efficiency differs between data and Monte Carlo due to imperfect modeling of the electron shower shape in the calorimeters and of the fraction of high-threshold hits in the TRT. The Monte Carlo predictions for these observables do not match the measured distributions. Future improvement of the modeling could give better agreement on the efficiency. In the present analysis, the efficiency is corrected using the measured difference between data and Monte Carlo, as a function of electron η and E_{T}, and uncertainty on the efficiency determination is included among the systematic uncertainties. The efficiency difference reaches 15% at low E_{T}.

Requiring an angular separation of $\Delta R > 0.5$ between any jet and the electron candidate, as described in Section 3.6, ensures that the electron identification efficiency is flat as a function of jet multiplicity. The efficiency has also been compared between data and Monte Carlo as a function of jet p_{T}, with no significant differences found. Thus the electron efficiency corrections are applied independently from the jet measurements.

3.5 Missing transverse energy reconstruction

Neutrinos are the main source of $E_{\mathrm{T}}^{\mathrm{miss}}$ since they pass through all the layers of ATLAS undetected. However, fake $E_{\mathrm{T}}^{\mathrm{miss}}$ can also be produced by noise and non-collision backgrounds, as well as the limited detector resolution, coverage, and uniformity. The experiment is designed to minimize these effects, and its excellent performance [49] allows the reliable use of this observable.

$E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed from three-dimensional groups of calorimeter cells passing noise-suppression criteria [11]. These are referred to as topological clusters, shortened to topoclusters. They are designed to follow the topology of energy showers in the calorimeters.

To construct topoclusters, noise thresholds are determined for each cell, including both electronic noise and the impact of pileup, based on calibration data. In each event, any cell with an energy significance of more than four standard deviations forms the seed of a topocluster. Neighboring cells with an energy significance of more than two standard deviations and any of their three-dimensional neighbors passing the same criterion are included. A last layer of all neighboring cells is then added, regardless of their energy.

For $E_{\mathrm{T}}^{\mathrm{miss}}$ calibration, cells are associated with different objects. First, cells within electrons and photos are included with the corresponding electromagnetic calibration. Then remaining cells associated with τ-jets or soft jets with $p_{\mathrm{T}} < 20$ GeV are included with local hadronic or EM calibration based on cluster topology (LCW) [50]. Cells from higher p_{T} jets are calibrated accordingly, as described in the next section. Finally, any remaining cells in topoclusters are included with LCW calibration and an additional correction based on associated tracks. The details of this and the object definitions are given in Ref. [49].

3.6 Jet counting

As jets are key to this measurement, the jet selection criteria are described in detail. The jet definition consists not only of the jet algorithm, but also the inputs to the jet finding. Once calibrated, jets with significant p_{T} are cleaned of any low-quality jets and required to pass kinematic criteria. Finally, jets associated with additional vertices due to pileup interactions are removed. The jet multiplicity of the data sample is shown at the end of this section.

3.6.1 Definition

Jets are reconstructed using the anti-kt algorithm [51] with distance parameter $R = 0.4$. The choice of the anti-kt algorithm is motivated by dedicated studies showing its robustness in the ATLAS environment [52]. Its use is facilitated by being implemented in the FastJet software package [53], which has become widely used in high energy physics.

The inputs to the jet algorithm in reconstructed events from data or simulation are the topoclusters described above in Section 3.5. In simulated events the jet algorithm can also be applied to particles before detector simulation and reconstruction, forming truth jets. All stable particles with a proper lifetime of at least 10 ps are considered, excluding electrons from W decay in $W \rightarrow e\nu$ events. Truth jets are used for jet calibration as described below and for the cross section correction and comparisons of Chapters 5 and 8, respectively.

Anti-kt is a sequential recombination algorithm: it successively groups together the topoclusters or truth particles into larger energy clusters, which eventually form the jets. At each stage, the distance from the beam of each cluster i, with transverse momentum k_{Ti}:

$$d_i = k_{\mathrm{Ti}}^{-2} \tag{3.2}$$

and the distance from any other cluster j:

$$d_{ij} = min(k_{\mathrm{Ti}}^{-2}, k_{\mathrm{Tj}}^{-2}) \frac{\Delta R_{ij}^2}{R^2} \tag{3.3}$$

are found for all clusters. The two clusters with minimum distance d_{ij} are merged, as long as this exceeds d_i, and the procedure continues until no further clusters can be merged. The angular separation of the clusters is:

$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \tag{3.4}$$

where y is the rapidity and ϕ is the azimuthal angle of each cluster. Thus the distance parameter R sets a characteristic size for the jets.

The anti-kt algorithm has several desirable features. It is infrared and collinear safe, so the results are stable to the presence of additional soft or nearby particles. It also produces jets with a roughly conical shape and radius approximating the distance parameter, which is useful for track-matching and the removal of overlapping physics objects. Using distance parameter $R = 0.4$ allows for higher jet multiplicities compared to other common settings, such as $R = 0.6$.

3.6.2 Calibration

During jet finding, all calorimeter cells in topoclusters are included, with their energies calibrated at electromagnetic scale. An additional calibration is then applied to correct the jet energy scale to the hadronic scale [54]. This is necessary due to the hadronic nature of jets and non-compensation of the calorimeters, as discussed in Chapter 2.2.2. The calibration also includes corrections for energy losses in inactive regions of the detector or beyond the depth of the calorimeters and increased energy in the presence of pileup interactions.

The calibration is derived from simulated events, as a function of jet p_{T} and η, and supported by studies with data [55]. Reconstructed jets are matched to truth jets within an angular separation $\Delta R < 0.3$. The calibration is obtained from the mean of a Gaussian fit to the distribution of measured p_{T} divided by true p_{T} for matched jets in each p_{T} and η bin.

Anti-kt jets with distance parameter $R = 0.4$ are used to derive the calibration. However, the jets considered for the calibration differ slightly from those used to measure the W+jets cross sections. Differences due to quark vs. gluon jets and non-isolated jets are included in the systematic uncertainty on the jet energy scale, as discussed in Chapter 6.

The first difference arises because the dijet calibration sample consists mainly of gluon jets. Meanwhile, quark jets are more likely than gluon jets to be produced in association with a W boson. They form narrower showers and thus tend to have a higher measured p_{T} than gluon jets with the same p_{T}. This difference has been studied using simuated events and tested with events from data containing a photon and jet, where the jet is more likely to come from a quark [56]. As the effect is only up to a few percent at low jet p_{T}, no additional correction is applied and the full difference is taken as a systematic uncertainty.

Another difference is that the jets for calibration are required to be isolated, with no second jet within angular separation $\Delta R < 1.0$, while no such criterion is applied in the W+jets analysis. This can result in a lower measured p_{T} for the jets in the W+jets sample than for the jets used in calibration. This effect has been studied by comparing the ratio of measured to true p_T between isolated and non-isolated jets in simulated events and found to be less than 3% for the jets used in this analysis. Again the effect is included in the JES uncertainty.

3.6.3 Jet quality selection

After applying the jet energy scale calibration, quality checks on jets with calibrated $p_T > 20$ GeV are applied to the data to reject problematic events and jets. The criteria applied to the jets are categorized as either loose or medium, with the medium criteria including all of the loose criteria. The requirements are summarized in Table 3.6, with details given in Reference [56].

Events are rejected if a jet fails the loose criteria. For jets failing the medium criteria, the jet is rejected and the event is kept. The efficiency of these criteria has been tested in data using a tag-and-probe method on events with two jets.

	Loose	Medium
HEC spikes	($f_{\mathrm{HEC}} > 0.5$ and $\lvert f_{\mathrm{HEC\ quality}}\rvert > 0.5$) or $\lvert \mathrm{neg_E}\rvert > 60$ GeV	$f_{\mathrm{HEC}} > 1 - \lvert f_{\mathrm{HEC\ quality}}\rvert$
Coherent EM noise	$f_{\mathrm{EM}} > 0.95$ and $f_{\mathrm{EM\ quality}} > 0.8$ and $\lvert\eta\rvert < 2.8$	$f_{\mathrm{EM}} > 0.9$ and $f_{\mathrm{EM\ quality}} > 0.8$ and $\lvert\eta\rvert < 2.8$
Non-collision background	$\lvert \mathrm{t_{jet}}\rvert > 25$ ns or ($f_{\mathrm{EM}} < 0.05$ and $f_{\mathrm{ch}} < 0.05$ and $\lvert\eta\rvert < 2$) or ($f_{\mathrm{EM}} < 0.05$ and $\lvert\eta\rvert \geq 2$) or ($f_{\mathrm{max}} > 0.99$ and $\lvert\eta\rvert < 2$)	$\lvert \mathrm{t_{jet}}\rvert > 10$ ns or ($f_{\mathrm{EM}} < 0.05$ and $f_{\mathrm{ch}} < 0.1$ and $\lvert\eta\rvert < 2$) or ($f_{\mathrm{EM}} > 0.95$ and $f_{\mathrm{ch}} < 0.05$ and $\lvert\eta\rvert < 2$)

Table 3.6: Selection criteria used to reject fake jets from non-collision background.

3.6.4 Kinematic selection

Kinematic criteria described above are then applied to the remaining jets. They are required to have calibrated $p_{\mathrm{T}} > 30$ GeV and rapidity $|y| < 4.4$. Only jets with angular separation $\Delta R > 0.5$ from the selected electron are counted. This avoids double counting the electron as a jet or considering jets that may be affected by a nearby electron.

3.6.5 Jet-vertex fraction

To study only those jets produced in association with a W boson, jets likely to have come from additional pileup interactions are rejected. This is done by calculating a property called the jet-vertex fraction (JVF) [57] for each jet in an event. The JVF is the fraction of the total p_{T} of tracks matched to a jet that is carried by those tracks that originate from the primary vertex. Matching tracks are those within a cone of radius $R = 0.4$ around the jet.

Figure 3.3 shows the JVF distribution for all jets passing the quality and kinematic criteria described above, in $W \to e\nu$ candidate events. Jets with insufficient tracking information to calculate the JVF, for example those outside the tracking coverage, are assigned a value of -1, and automatically accepted. Jets with $|JVF| < 0.75$ are rejected, as they are unlikely to correspond to the primary vertex. While effective for pileup jets, the overall the impact of this requirement is small, with an efficiency of 99.1% for data and 98.8% for Monte Carlo.

Table 3.7 shows the jet multiplicity of the data sample and the impact on the jet multiplicity from removing pileup jets using the JVF. The JVF has an overall effect of just 1%, indicating that there are few jets from pileup in the data sample. As a conservative estimate of the impact of pileup jets on the measured cross sections, the full effect of applying the JVF criterion to the data is included among the systematic uncertainties, as a function of each measured observable. This uncertainty is small compared to others, as shown in Chapter 6.

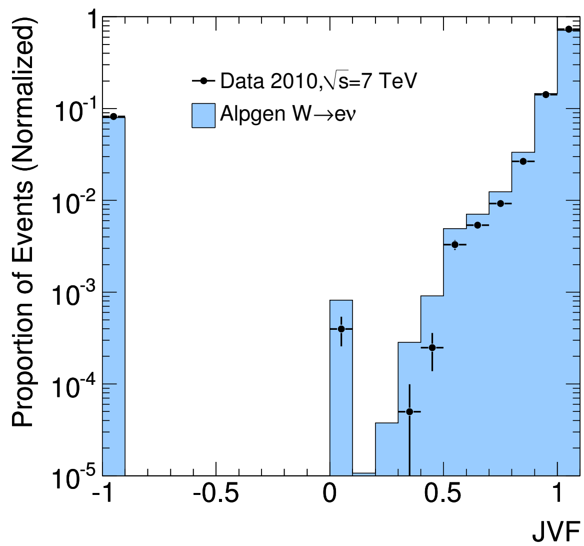

Figure 3.3: Jet-vertex fraction distribution (see text) for all jets in the data sample and in Monte Carlo simulated $W \to e\nu$ events.

Jet multiplicity	Number of events	
	Before JVF criterion	Final selection
Njet = 0	97439	97599
Njet = 1	11160	11054
Njet = 2	2716	2675
Njet = 3	703	694
Njet = 4	222	223
Njet $\geq$ 5	93	88

Table 3.7: Jet multiplicity of the data sample before and after removing jets from pileup.

CHAPTER 4

BACKGROUNDS

Most backgrounds to $W \to e\nu$ can be described as leptonic, containing an electron or a tau lepton that decays to one. They consist of $Z \to ee$, $Z \to \tau\tau$, $W \to \tau\nu$, dibosons (WW, WZ, and ZZ), single-top, and $t\bar{t}$. These backgrounds are estimated from Monte Carlo studies, using the samples and normalization described in Section 3.2.2.

The QCD background from events containing only jets is more challenging. A small fraction of these events can pass the electron selection, either through a heavy quark decay or a jet that mimics the characteristics of an electron. As noted previously, this small fraction still presents an important background since the inclusive jet cross section is much larger than the W production cross section.

The very low probability of such events passing the selection makes generation of sufficient statistics of simulated events to estimate the background prohibitively difficult. In addition, the rate at which QCD events pass the selection cannot be modeled with sufficient precision. Thus data-driven methods are critical for estimating the QCD background. The primary method used in this analysis and an alternative used as a crosscheck are presented in Section 4.1.

At low jet multiplicity, the expected $W \to e\nu$ signal far exceeds the backgrounds. However, for three or more jets the $t\bar{t}$ background becomes significant. Section 4.2 describes the in-situ validation of the $t\bar{t}$ background estimates. Contributions from the other backgrounds are small across all jet multiplicities, as shown in the final section of this chapter.

4.1 QCD multijet events

Two data-driven methods of estimating the QCD background have been developed for this analysis. Each method uses a background-enriched control sample selected from data to model the shape of the QCD background and a fit of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution to determine the normalization. Two templates are fit to the data: a QCD template from the control sample and a combined leptonic template from Monte Carlo simulated events, consisting of $W \to e\nu$, $Z \to ee$, $Z \to \tau\tau$, $W \to \tau\nu$, dibosons, single-top, and $t\bar{t}$. For the leptonic template, the samples of simulated events are normalized relative to each other using (N)NLO cross section predictions and combined. The overall normalization of the leptonic template and the normalization of the QCD template are the free parameters in the fit. The fits are done independently for different jet multiplicities, using the binned maximum likelihood fit implemented in ROOT TFractionFitter [58] and assuming Poisson statistics for the templates [59].

4.1.1 Background estimation

The electron identification criteria are designed to reduce background contamination from jets. The tight criteria do this so well that it is difficult to study the small remaining QCD background. To obtain a large control sample of QCD multijet events that is statistically independent from the primary data sample, certain criteria are reversed.

The control sample is selected from events with an electron candidate passing the loose identification criteria and the usual kinematic requirements. The normal trigger selection is not imposed as it includes medium identification. Instead events passing any enabled electron or photon trigger are considered. The efficiency of this trigger combination is not needed for the QCD background estimation since the normalization is determined by fitting.

At least one of the criteria indicated in Table 3.5 is required to fail, and the remaining medium criteria on the track quality and matching to the EM cluster are required to pass. The reversed criteria use the first layer of the EM calorimeter, the transverse impact parameter, the matching between the cluster energy and track momentum and the TRT. These criteria have been chosen to select a control sample with the same shape as the QCD background passing the primary selection, and they have been validated with Monte Carlo simulated events [60, 61]. Possible sensitivity to the choice of reversed criteria is included among the systematic uncertainties by varying the selection of the control sample as described below.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is used to fit for the normalization as it gives clear shape discrimination between the QCD and leptonic samples. As noted previously, a defining characteristic of $W \rightarrow e\nu$ events is large $E_{\mathrm{T}}^{\mathrm{miss}}$ due to the presence of a neutrino. Also important at high jet multiplicity is the $t\bar{t}$ background, which has large $E_{\mathrm{T}}^{\mathrm{miss}}$ again due to a neutrino from W decay. This distinguishes the leptonic template from QCD, where any residual $E_{\mathrm{T}}^{\mathrm{miss}}$ comes mainly from detector effects. To exploit this, the QCD control sample and leptonic template are selected with no requirement on $E_{\mathrm{T}}^{\mathrm{miss}}$, and the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution between 0 and 100 GeV is used for the fit. Once the QCD template is normalized, the final criteria of $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV is applied to determine the number of QCD events passing the full selection.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ shape discrimination and results of the fit are shown in Fig. 4.1. The fits are performed independently on the samples of events with one through four jets, to avoid double counting, and for events with five or more jets as a single sample, so that all events are considered. The agreement between the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of the data sample and the fitted templates is shown for all jet multiplicities. This figure also illustrates the background reduction, particular for QCD, from requiring $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV.

The systematic uncertainty on the QCD background includes uncertainty on the leptonic template from the Monte Carlo simulated samples and on the estimation method. All of the systematic uncertainties on the leptonic background that are discussed in Chapter 6 are taken into account in the uncertainty on the QCD background estimate. For each uncertainty, the leptonic template was varied before fitting and a new QCD normalization was determined.

The QCD background uncertainties and an indication of their size based on events with one or more jets are shown in Table 4.1. The uncertainties are taken symmetrically and summed in quadrature. Assumption in the estimation method including the fit range and control sample selection are considered.

Source	Description	Effect (%)
Leptonic template	All background systematic uncertainties described in Chapter 6	+25.3, –23.1
Fit range	Lower range 0 - 35 GeV	±20.8
	Upper range 15 - 100 GeV	±9.5
Reversed selection criteria	Fail one remaining medium and one remaining tight requirement	±13.4
	No medium requirements	±32.0
Trigger	Difference from alternative QCD method described in Section 4.1.2	±15.9

Table 4.1: Systematic uncertainties on the QCD background. The impact on the QCD estimate for events with one or more jets is shown.

The estimated QCD background could depend on the fit range if the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of the data is not well matched by the templates. Figure 4.1 shows that this is not the case, and it has been checked explicitly by varying the fit range. A lower range of 0 to 35 GeV and an upper range of 15 to 100 GeV produced only small variations.

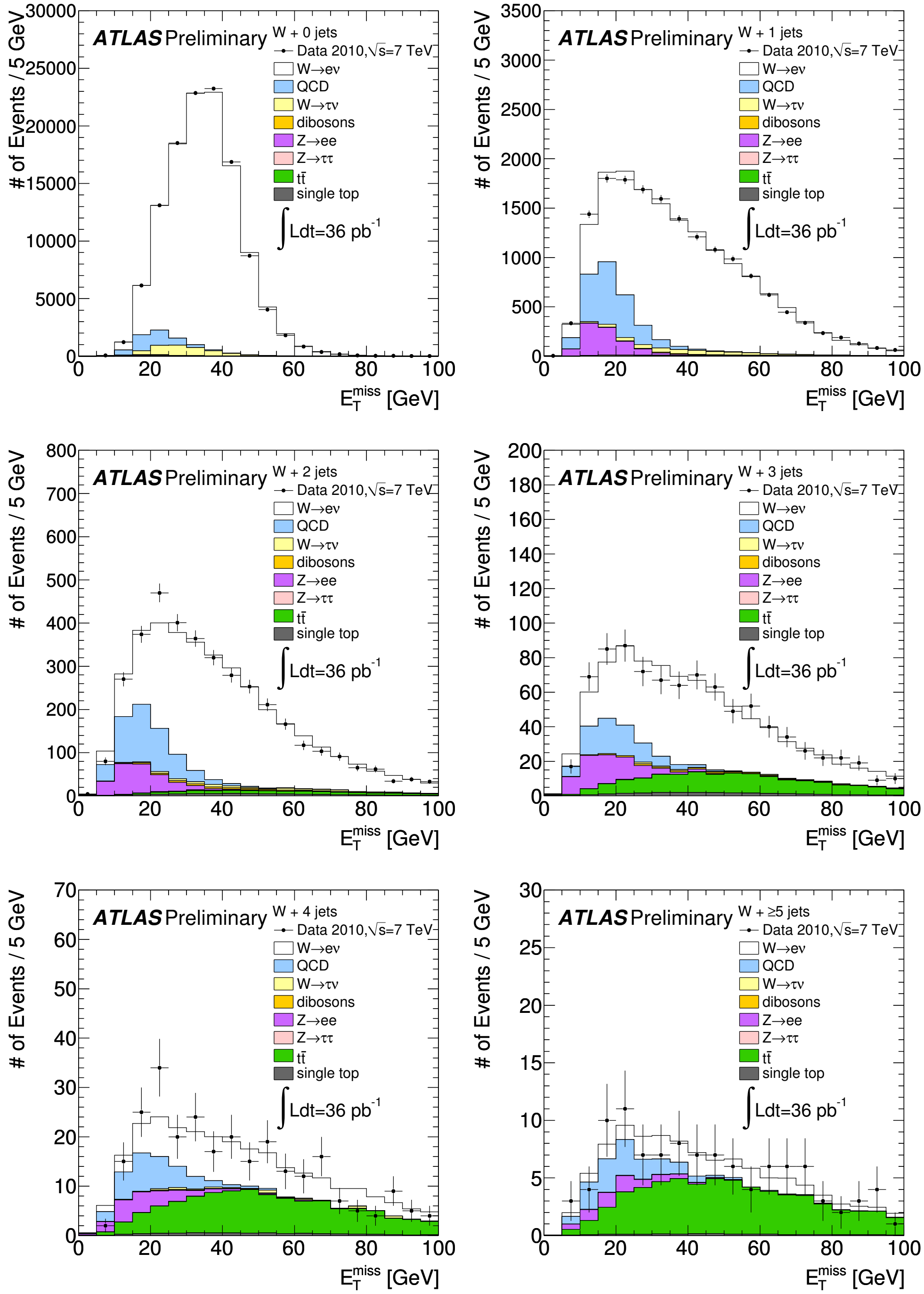

Figure 4.1: Results of fitting the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution to obtain the normalization of the QCD background template from data, for jet multiplicities from zero (top left) through five (bottom right). The final event selection requires $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV.

The choice of selection criteria for the control template has been tested by varying the requirements. First the sensitivity to the reversed criteria is evaluated by explicitly requiring an additional medium requirement (the number of pixel hits, combined pixel and SCT hits, or the matching between track and cluster) and an additional tight requirement (refined matching between track and cluster or rejection of photon conversions) to fail. Alternatively, the medium tracking requirements are removed from the selection. This has the largest impact on the estimated QCD background for events with one or more jets, contributing an uncertainty of 32%.

Finally, the difference in estimated QCD background from an alternative data-driven method described in the next section is included as a systematic uncertainty. The second method uses an explicit trigger selection. Taking the difference between the methods gives an estimate of any bias from selecting the control sample using a combination of triggers.

4.1.2 Crosscheck

A second QCD estimation method using data, which was developed when first measuring the W+jets cross sections with a smaller data sample [4], is used as a crosscheck. The QCD control sample in this alternative method is selected by substituting photon identification requirements for those of the electron. These events have shower shapes in the electromagnetic calorimeter which closely resemble electrons but have no associated charged track. The control sample is dominated by QCD multijet events that fail the tracking requirements but pass the photon EM cluster criteria and thus have similar event distributions to QCD multijet events passing the electron selection. A reversed isolation cut is applied to improve the $W \to e\nu$ signal rejection in the control sample. The transverse energy in the EM calorimeter in a cone of radius $R = 0.2$ around the photon, excluding the energy of the photon itself, is required to be greater than 15% of the transverse energy of the photon.

For early data-taking where the primary selection uses the L1 trigger, which has no tracking requirements, the same trigger is used for the control sample. For the remainder of the data, a photon trigger is used. Events are weighted to correct for the random rejection of a selected fraction of events, a standard procedure called trigger prescaling, as the event rate increased. This correction restores the relative contributions from periods with different pileup conditions.

As in the primary QCD estimation method, the normalization of the control sample is determined by fitting the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution independently for different jet multiplicities. The three samples with zero, one, and two jets are fit individually, while events with three or more jets are fit as a single sample to have sufficient statistics. To exploit the shape discrimination at low $E_{\mathrm{T}}^{\mathrm{miss}}$ the fit is performed in the range $E_{\mathrm{T}}^{\mathrm{miss}} > 10$ GeV, and the final cut of $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV is then applied to the normalized control sample to obtain the estimated QCD background in the signal region.

A comparison of the QCD estimates obtained with the two different methods is shown in Table 4.2. For events with five or more jets, the statistics of the data are so low that the uncertainty of the QCD background estimate is 100%. The cross section is not measured for this sample, due to the even larger effect of uncertainty on the $t\bar{t}$ background, which is discussed in Section 4.2.

4.1.3 Impact of electron isolation

The effects of the electron isolation requirement on the QCD background, as measured from data, and on the signal, as estimated from the sample of Alpgen Monte Carlo events, are shown in Table 4.3. Relative isolation efficiencies are calculated by taking

Jet multiplicity	Primary method $N_{\text{bkg}} \pm$ stat $\pm$ sys	Crosscheck $N_{\text{bkg}} \pm$ stat $\pm$ sys
Njet ≥ 0	$1488.8 \pm 115.9\ ^{+250.4}_{-230.6}$	$1612.5 \pm 132.7\ ^{+568.2}_{-561.6}$
Njet ≥ 1	$481.4 \pm 56.9\ ^{+128.9}_{-118.9}$	$557.9 \pm 68.7\ ^{+176.3}_{-172.8}$
Njet ≥ 2	$145.3 \pm 34.4\ ^{+54.7}_{-53.4}$	$211.3 \pm 39.3\ ^{+53.0}_{-52.3}$
Njet ≥ 3	$38.4 \pm 19.9\ ^{+16.8}_{-16.5}$	$24.9 \pm 13.7\ ^{+22.0}_{-21.5}$
Njet ≥ 4	$15.4 \pm 10.3\ ^{+11.2}_{-11.0}$	$5.0 \pm 2.8\ ^{+4.4}_{-4.7}$
Njet ≥ 5	$5.4 \pm 4.4\ ^{+2.8}_{-2.6}$	$0.0 \pm 0.0\ ^{+0.0}_{-0.0}$

Table 4.2: Difference between QCD methods.

the ratio of the number of events passing the final selection to those passing the selection without requiring isolation. The QCD background is significantly reduced by the isolation requirement, especially at high jet multiplicities where more than 50% of the background is rejected. The signal efficiency is above 97% for all jet multiplicities.

Jet multiplicity	QCD background			$W \to e\nu$		
	Before Iso.	After Iso.	Eff. (%)	Before Iso.	After Iso.	Eff. (%)
Njet ≥ 0	2928.5	1488.8	50.8	104740.3	103365.4	98.7
Njet ≥ 1	1261.4	481.4	38.2	12497.0	12284.0	98.3
Njet ≥ 2	426.5	145.3	34.1	2755.2	2703.3	98.1
Njet ≥ 3	131.5	38.4	29.2	562.2	550.8	98.0
Njet ≥ 4	44.8	15.4	34.4	113.2	110.7	97.8
Njet ≥ 5	12.8	5.4	42.2	22.2	21.6	97.3

Table 4.3: Impact of requiring electron isolation. The number of estimated QCD background events and the expected number of signal events before and after requiring isolation and the efficiencies determined by the fraction of events surviving the requirement are shown.

4.2 Top pair production

As illustrated in the next section, top pair production is the most important background to W+jets at high multiplicity. It closely matches the W+jets event signature because each top quark[1] decays to a W boson and bottom quark, which likely produces a jet. While one W boson decaying via $W \to e\nu$ will give an ideal W candidate, over 50% of the time the other W boson will decay to a quark-antiquark pair. This likely produces two more jets, for a total of four.

The $t\bar{t}$ background is approximately three times larger than the signal for events with four or more jets, and the uncertainty on this background must be treated carefully. Two kinds of uncertainty are possible: uncertainty on the shape of the $t\bar{t}$ background as a function of each measured observable, *e.g.*, jet multiplicity, and uncertainty on the overall normalization. The former is evaluated using the simulated $t\bar{t}$ event samples with modified QCD radiation described in Section 3.2.2. The latter is taken from uncertainty on the predicted cross section and validated with data.

Uncertainty on the $t\bar{t}$ background shape due to initial state radiation (ISR) and final state radiation (FSR) is included among the systematic uncertainties for each measured cross section. Variations of ISR and FSR alone and simultaneously were considered. ISR alone produced the most variation in the predicted number of events with four or more jets and is taken as a conservative estimate of the uncertainty. The impact of this variation on the $t\bar{t}$ shape is shown in Figs. 4.2 and 4.3.

Figure 4.2 compares the shape of the electron pseudorapidity and jet multiplicity distributions for the Monte Carlo simulated $t\bar{t}$ and W+jets event samples. The electron pseudorapidity is shown for the full sample, regardless of jet multiplicity, as the results are similar for different jet multiplicities. While the distribution is more central for $t\bar{t}$ events than for $W \to e\nu$, there is little sensitivity to ISR variation. The jet multiplicity is more sensitive, with differences visible even on the log scale. Events with five or more

1. Since charge information is not used here, both top and anti-top quarks are called top.

jets have the largest $t\bar{t}$ background and the most sensitivity and are not included in the final results.

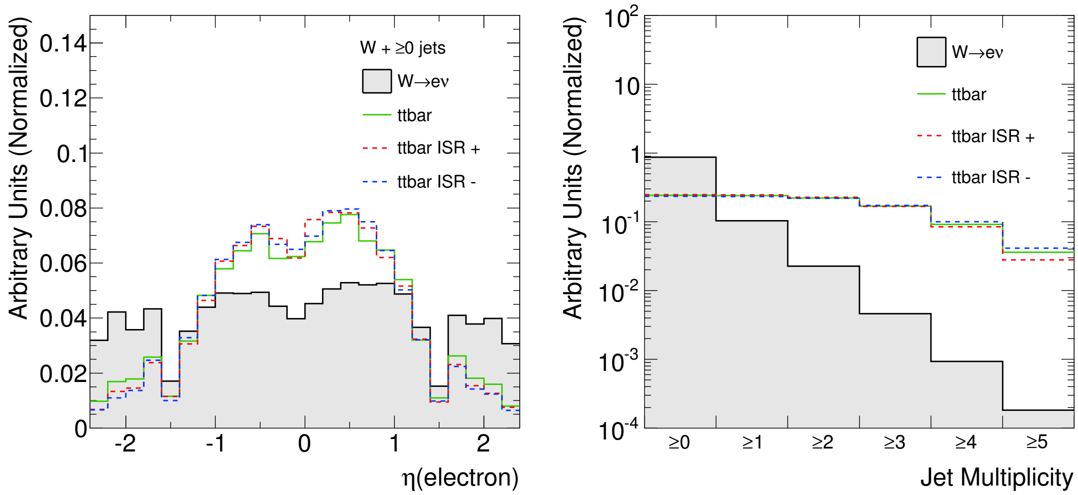

Figure 4.2: Shape difference between Monte Carlo simulated $W \to e\nu$ events and $t\bar{t}$ events passing the selection for the electron pseudorapidity (top) and inclusive jet multiplicity (bottom). Dashed lines show the shapes for $t\bar{t}$ events with modified QCD radiation. The electron pseudorapidity is hardly affected, while high values of the jet multiplicity are more sensitive to modified QCD radiation.

Figure 4.3 presents similar comparisons for the invariant mass of the combined system formed by the W candidate and a jet. Combinations of the W with each jet up to the fourth sub-leading jet are included. The transverse momentum of the neutrino from the W decay is determined from $E_{\mathrm{T}}^{\mathrm{miss}}$ and the W mass, but there are two solutions for the longitudinal component. The solution with smaller absolute value is used, as in [62].

For events containing a W boson and jet from top quark decay, the distribution of the reconstructed W+jet mass is more likely to peak near the top mass. This difference between $t\bar{t}$ and W+jets events and its evolution with jet multiplicity is visible in Fig. 4.3. Meanwhile, as for electron pseudorapidity, the W+jet mass is relatively insensitive to ISR variation.

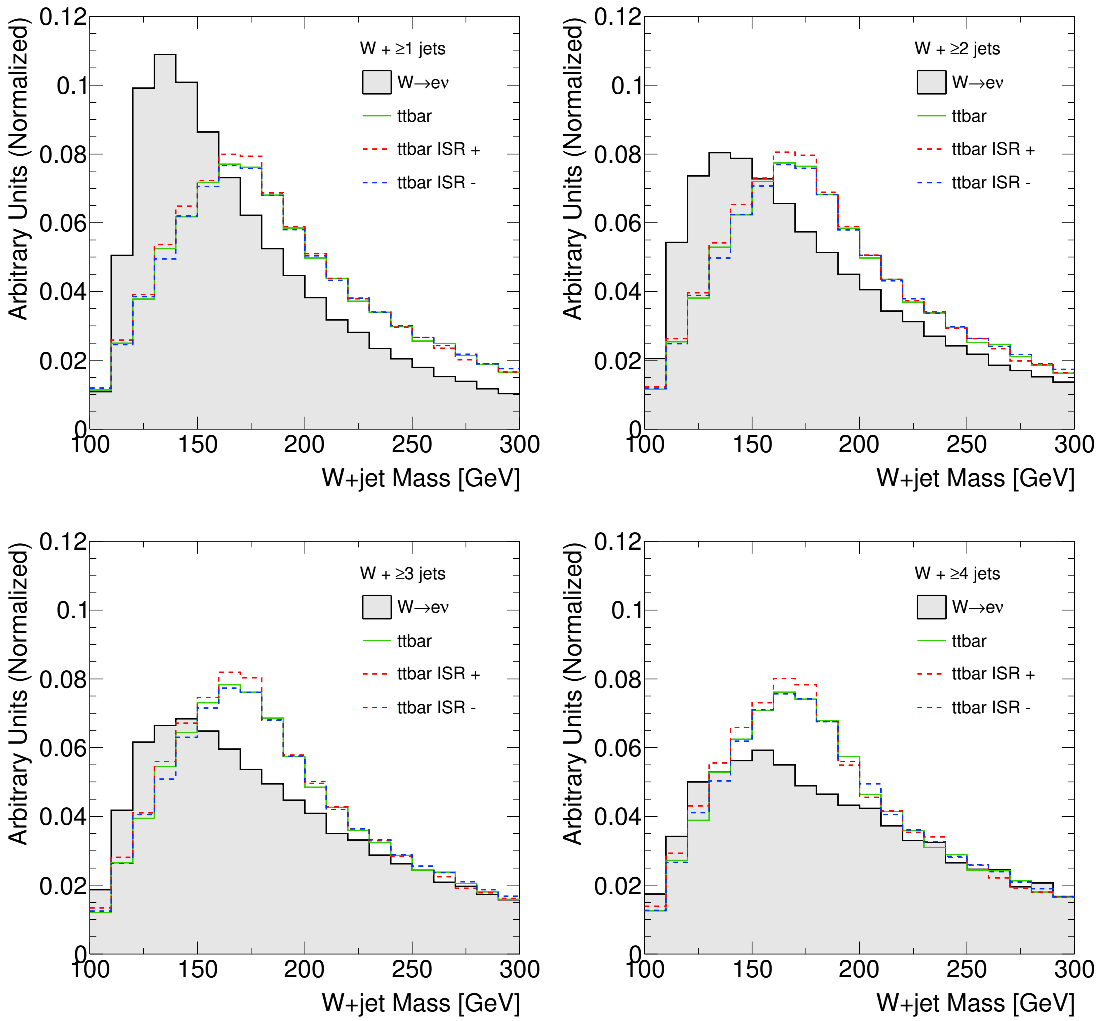

Figure 4.3: Shape difference between Monte Carlo simulated $W \to e\nu$ events and $t\bar{t}$ events passing the selection for the invariant mass reconstructed from the W boson candidate and a jet, in events with one or more jets (top left) through four or more jets (bottom right). Distributions for $t\bar{t}$ events with altered ISR settings are also shown. The shapes of the distributions are almost independent of the ISR settings.

The shape differences between W+jets and $t\bar{t}$ shown in Figs. 4.2 and 4.3 allow a test of the $t\bar{t}$ background normalization. Figure 4.4 shows that despite the $t\bar{t}$ shape difference, data and predictions agree within uncertainties for all jet multiplicities. This has also been checked for the electron pseudorapidity.

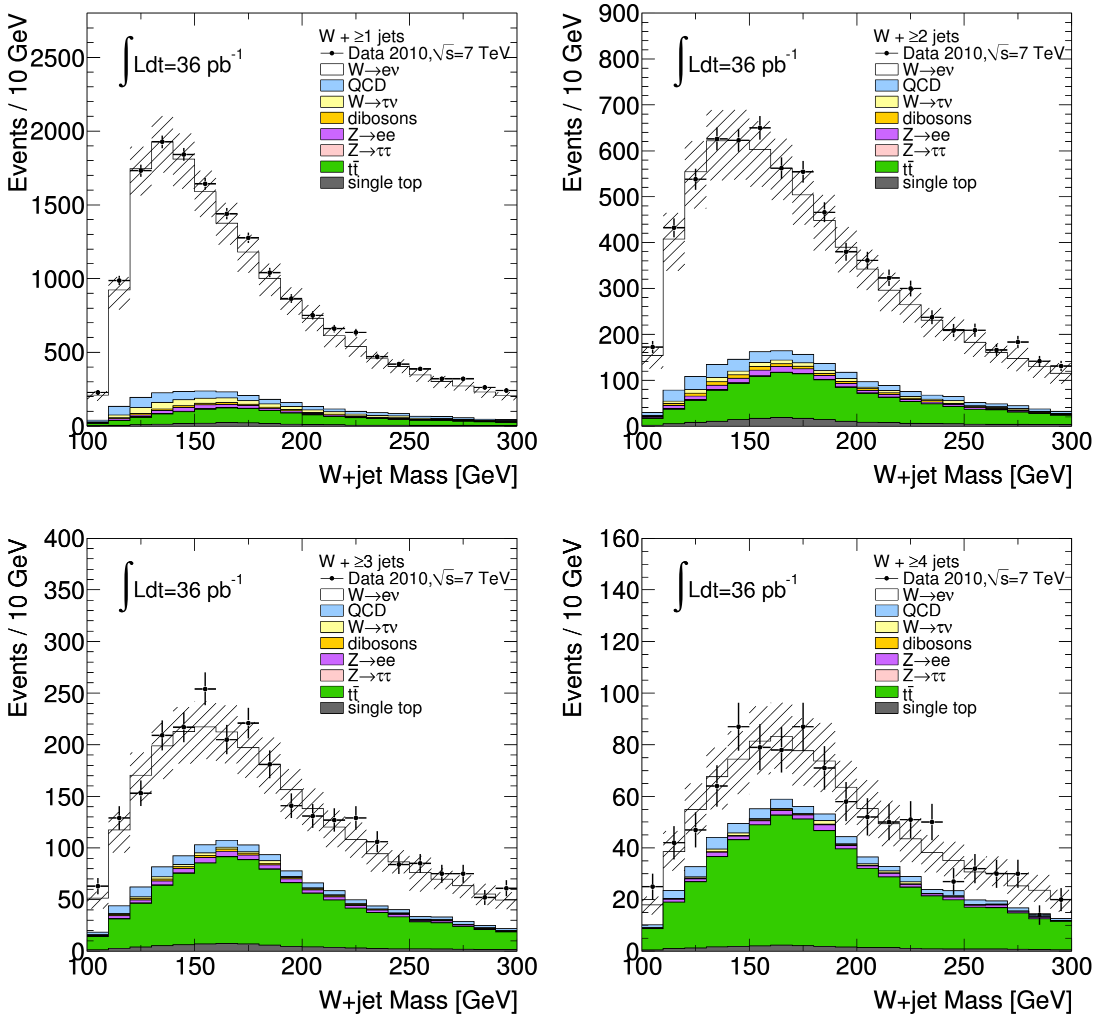

Figure 4.4: Reconstruction of the invariant mass of the W boson candidate plus one jet, showing the agreement between data and predictions for all jet multiplicities from one (top left) through four (bottom right). The shaded band shows the combined statistical and systematic uncertainty on the predictions.

4.3 Summary of backgrounds

Table 4.4 shows the estimated backgrounds for each inclusive jet multiplicity from zero to five jets. The expected numbers of signal events per jet multiplicity is also shown for comparison. Uncertainty on the QCD background is determined as discussed in Section 4.1 and the remaining uncertainties are detailed in Chapter 6.

The total background is just 5% for the inclusive sample and grows relative to the signal with increasing jet multiplicity. For events with four or more jets, the highest jet multiplicity measured, the uncertainty of less than 20% on the $t\bar{t}$ background results in about 10% uncertainty on the number of events after background subtraction. At the next higher jet multiplicity, events with five or more jets, the relative size of the backgrounds and their uncertainty are too large to reliably measure the W+jets cross sections.

Background	Njet ≥ 0 $N_{\text{bkg}} \pm$ stat $\pm$ sys	Njet ≥ 1 $N_{\text{bkg}} \pm$ stat $\pm$ sys	Njet ≥ 2 $N_{\text{bkg}} \pm$ stat $\pm$ sys
QCD	$1488.8 \pm 115.9\ ^{+250.4}_{-230.6}$	$481.4 \pm 56.9\ ^{+128.9}_{-118.9}$	$145.3 \pm 34.4\ ^{+54.7}_{-53.4}$
$W \to \tau\nu$	$2712.8 \pm 24.8\ ^{+197.1}_{-201.2}$	$353.0 \pm 8.9\ ^{+38.5}_{-49.9}$	$54.9 \pm 3.5\ ^{+6.5}_{-13.5}$
Dibosons	$131.3 \pm 0.6\ ^{+8.7}_{-8.7}$	$95.6 \pm 0.5\ ^{+6.8}_{-7.0}$	$40.9 \pm 0.3\ ^{+4.5}_{-5.5}$
$Z \to ee$	$189.2 \pm 3.5\ ^{+40.5}_{-41.8}$	$152.9 \pm 3.1\ ^{+32.5}_{-35.1}$	$55.5 \pm 1.7\ ^{+16.2}_{-17.6}$
$Z \to \tau\tau$	$132.0 \pm 1.7\ ^{+9.8}_{-10.0}$	$31.3 \pm 0.8\ ^{+3.6}_{-4.7}$	$8.8 \pm 0.4\ ^{+1.0}_{-1.4}$
$t\bar{t}$	$435.3 \pm 2.9\ ^{+37.8}_{-33.9}$	$432.2 \pm 2.9\ ^{+37.7}_{-33.7}$	$400.6 \pm 2.8\ ^{+37.4}_{-31.9}$
single top	$119.7 \pm 0.6\ ^{+9.3}_{-9.3}$	$112.9 \pm 0.6\ ^{+8.9}_{-8.9}$	$79.5 \pm 0.5\ ^{+7.1}_{-7.4}$
$W \to e\nu$	103365.4	12284.0	2703.3
Background	Njet ≥ 3 $N_{\text{bkg}} \pm$ stat $\pm$ sys	Njet ≥ 4 $N_{\text{bkg}} \pm$ stat $\pm$ sys	Njet ≥ 5 $N_{\text{bkg}} \pm$ stat $\pm$ sys
QCD	$38.4 \pm 19.9\ ^{+16.8}_{-16.5}$	$15.4 \pm 10.3\ ^{+11.2}_{-11.0}$	$5.4 \pm 4.4\ ^{+2.8}_{-2.6}$
$W \to \tau\nu$	$8.2 \pm 1.3\ ^{+2.0}_{-1.5}$	$2.5 \pm 0.7\ ^{+1.6}_{-1.1}$	$0.0 \pm 0.0\ ^{+0.0}_{-0.0}$
Dibosons	$9.1 \pm 0.2\ ^{+1.4}_{-1.9}$	$1.6 \pm 0.1\ ^{+0.3}_{-0.5}$	$0.2 \pm 0.0\ ^{+0.1}_{-0.1}$
$Z \to ee$	$19.8 \pm 1.1\ ^{+7.0}_{-7.8}$	$6.9 \pm 0.6\ ^{+1.8}_{-1.9}$	$2.1 \pm 0.2\ ^{+0.6}_{-0.7}$
$Z \to \tau\tau$	$1.7 \pm 0.2\ ^{+0.2}_{-0.4}$	$0.3 \pm 0.1\ ^{+0.1}_{-0.1}$	$0.1 \pm 0.1\ ^{+0.0}_{-0.0}$
$t\bar{t}$	$302.6 \pm 2.5\ ^{+36.7}_{-27.9}$	$165.3 \pm 1.8\ ^{+30.8}_{-24.7}$	$64.7 \pm 1.1\ ^{+16.6}_{-18.6}$
single top	$28.4 \pm 0.3\ ^{+3.3}_{-3.8}$	$8.0 \pm 0.2\ ^{+1.2}_{-1.5}$	$1.9 \pm 0.1\ ^{+0.4}_{-0.5}$
$W \to e\nu$	550.8	110.7	21.6

Table 4.4: Background estimates for the electron channel. The number of signal events predicted from the ALPGEN Monte Carlo simulated $W \to e\nu$ sample is also shown for comparison. The luminosity uncertainty of ±3.4% is not included in the systematic uncertainty.

CHAPTER 5

CROSS SECTION CALCULATION

The inclusive and differential cross sections measured in this analysis are presented for a limited phase space, which is explicitly defined in Section 5.1. This phase space is motivated by the fiducial acceptance of the detector and kinematic thresholds that ensure a high-precision measurement. No correction is made to a total cross section in order to avoid assumptions on the cross section variation in unmeasured regions of phase space.

To produce the final cross sections, the results are corrected for all known detector effects and divided by the integrated luminosity of the data sample. They can then conveniently be compared to theoretical predictions, as is done in Chapter 8, or the results of other experiments. The correction procedure, known as unfolding, is described in Section 5.2. It corrects for inefficiencies within the measurement phase space, detector resolution, and other experimental effects such as nonlinearity of the detector response.

5.1 Phase space

The limited phase space for the cross section measurement is defined in Table 5.1. The phase space criteria closely follow the criteria for the selection of reconstructed events, which were described in Chapter 3.

The phase space criteria are applied to particle kinematics after Monte Carlo generation and before detector simulation. Electrons are considered after QED radiation, with the p_{T} of any photons within a cone of radius of 0.1 around the electron direction included. The jet selection is applied to truth jets, as defined in Section 3.6.

Kinematic Property	Criterion Applied
Electron p_{T}	$p_{\mathrm{T}} > 20$ GeV
Electron η	$\lvert\eta\rvert < 2.5$
Neutrino p_{T}	$p_T > 25$ GeV
W transverse mass	$M_T > 40$ GeV
Jet algorithm	Anti-Kt, $R = 0.4$
Jet p_{T}	$p_{\mathrm{T}} > 30$ GeV
Jet rapidity	$\lvert y\rvert < 4.4$
Jet isolation	$\Delta R(e, \mathrm{jet}) > 0.5$

Table 5.1: Phase space criteria for the measured cross sections.

5.2 Unfolding

Unfolding detector-level results to particle level involves the correction of each measured quantity, *e.g.*, jet multiplicity, for experimental effects, such as smearing due to detector resolution, with which they have been convolved or "folded." This challenge is not unique to particle physics, and many statistical techniques have been developed to address it, as reviewed for example in [63].

First, it is possible to apply a simple correction factor, by taking the ratio of Monte Carlo predictions before and after applying the detector simulation. The correction factor can be calculated for each bin of a distribution, *e.g.*, each number of jets for the jet multiplicity distribution, and this technique is known as bin-by-bin unfolding. This procedure provides a basic starting point, and the results of applying bin-by-bin unfolding are compared below to the more robust method chosen for the final analysis.

The bin-by-bin unfolding method relies completely on Monte Carlo modeling of experimental effects. Since the Monte Carlo simulation cannot perfectly model the detector response, a technique with less dependence on the detector simulation is desirable. One such technique is to apply an iterative method based on Bayes' theorem [64]. This analysis uses a modified version of the method implemented in the RooUnfold package [65].

The number of iterations is handled differently, as discussed below.

Mathematically, the method can be expressed as in [65], where the unfolded distribution:

$$\hat{n}(C_i) = \sum_{j=1}^{n_E} M_{ij} n(E_j) \tag{5.1}$$

is obtained by applying the unfolding matrix:

$$M_{ij} = \frac{P(E_j|C_i) n_0(C_i)}{\epsilon_i \sum_{l=1}^{n_C} P(E_j|C_l) n_0(C_l)} \tag{5.2}$$

to the measurements, $n(E_j)$. The unfolding matrix is based on the joint probability, $P(E_j|C_i)$, that a result, C_i, comes from a measurement, E_j, given the efficiencies, $\epsilon_i \equiv \sum_{j=1}^{n_E} P(E_j|C_i)$, and an initial assumption, $n_0(C_l)$, of the unfolded distribution. In this analysis, the number of bins for each unfolded distribution, n_C, is the same as the number of bins for the measured distribution, n_E.

From Equation 5.1, each bin of the unfolded distribution can be seen as a sum of possible migrations from all bins of the measured distribution. When the off-diagonal elements of the unfolding matrix are non-trivial, indicating significant migrations between bins, the bin-by-bin unfolding method underestimates the statistical uncertainty. The uncertainty determined with the Bayesian method is more accurate as it takes into account migrations from bins with lower statistics. In the bin-by-bin method only the net migration is considered and the statistical uncertainty on the bin itself is scaled accordingly.

The Bayesian method described by Equations 5.1 and 5.2 is applied iteratively. On each iteration, the initial distribution, $n_0(C_l)$, is replaced with the unfolded distribution obtained from the previous iteration. This reduces sensitivity to the first assumed distribution.

In addition to the starting distribution, the initial response matrix is obtained from the Monte Carlo simulated event sample. As the results of this analysis are all one-dimensional distributions, each response matrix is a two-dimensional histogram of the distribution before detector simulation and reconstruction versus the corresponding distribution after these effects. The binning for both axes is the same, and it is used for the results shown in Chapter 8. The fraction of events passing the selection criteria before the detector simulation but failing after reconstruction is used to apply an overall efficiency correction.

The method is sensitive to the number of iterations applied. Additional iterations decrease the dependence on the original Monte Carlo spectrum. However, in the limit of infinite iterations, the results will follow any statistical fluctuations of the data. To avoid this, the number of iterations used for the final result in a given bin of the unfolded distribution is limited by requiring the last iteration to produce a significant change relative to the statistical uncertainty. Different threshold values for this significance have been studied [66], leading to the requirement that the change induced by the final iteration is at least 20% of the statistical uncertainty on the bin.

The number of iterations is allowed to vary for each bin of the unfolded distributions up to this threshold or a maximum of five iterations. In bins with low significance after background subtraction, where the number of events is less than three times the uncertainty, only one iteration is performed. In this case the procedure corresponds to bin-by-bin unfolding.

The unfolding corrections obtained with this method are illustrated for the jet multiplicity distribution in Fig. 5.1. As the corrections include the selection efficiency, the overall offset from one is expected. The difference between the correction for events with zero or more jets and the correction for events with one or more jets indicates the total size of jet effects such as jet reconstruction efficiency and jet energy scale and resolution. In comparison, the differences between the corrections obtained using events simulated with ALPGEN and events simulated with SHERPA are small. Those differences are used to estimate the systematic uncertainty, as discussed below.

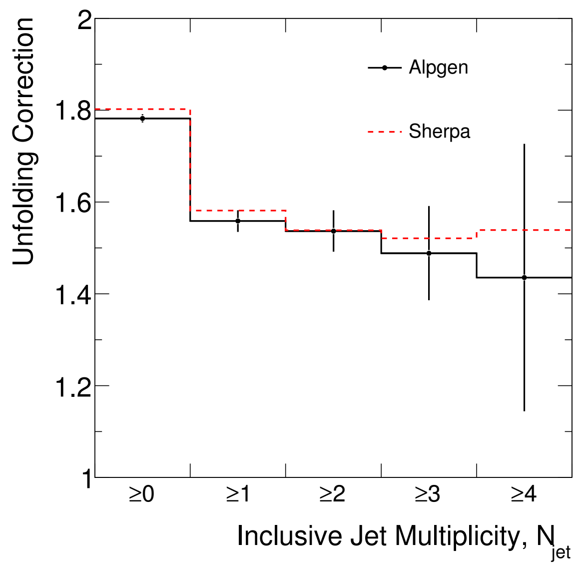

Figure 5.1: Unfolding corrections for the jet multiplicity distribution. Error bars show the statistical uncertainty of the data. The difference from the corrections obtained with a different simulated event sample, shown by the dashed line, indicates the systematic uncertainty on the unfolding.

To evaluate all of the systematic uncertainties described in Chapter 6, the unfolding is performed separately for each systematic uncertainty and the differences of the unfolded results with each uncertainty from the unfolded results with no uncertainty are combined in quadrature. The systematic uncertainty arising from the unfolding method itself is also taken into account. Several possible sources of this uncertainty were considered.

Figure 5.2 shows the impact of the unfolding method as a function of jet multiplicity. Results obtained using the bin-by-bin unfolding method or the Bayesian method with a fixed number of iterations are compared to the final results. Residual sensitivity to the assumed initial distribution was also evaluated by modifying the usual distribution taken from simulated events by the differences from the uncorrected data. The differences is each case are small.

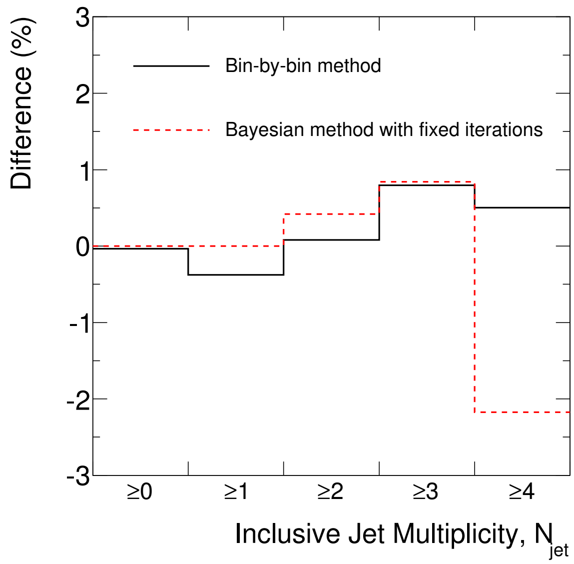

Figure 5.2: Comparison of unfolding methods. The difference between the final results and bin-by-bin unfolding (solid) or Bayesian unfolding with a fixed number of iterations (dashed) is shown as a percentage of the measured cross section for events with zero through four or more jets.

A more conservative estimate of the systematic uncertainty from unfolding was obtained using different simulated event samples to apply the method. Substituting $W \to e\nu$ events simulated with Sherpa instead of Alpgen, for both the initial distribution and migration matrix, gives a systematic uncertainty ranging from about 1% for the total sample to 6% for events with four or more jets. Estimating the uncertainty in this way takes into account uncertainty on the underlying event and PDFs, as different models are used.

CHAPTER 6

SYSTEMATIC UNCERTAINTIES

Figure 6.1 shows the different contributions to the systematic uncertainty of the cross section measurement. Since the relative size of the contributions varies greatly, larger uncertainties are shown in the top plot and smaller ones are shown on a more refined scale in the bottom plot. The dominant contribution for events with jets is from uncertainty of the jet energy scale (JES), which is discussed in the next section. Uncertainties associated with the electron are mostly small and are discussed along with the other remaining uncertainties in Section 6.2. The final section of this chapter compares the uncertainties quantitatively.

6.1 Jet uncertainties

Jets are complex objects to measure as they include many particles and QCD interactions. This measurement is particularly sensitive to uncertainty of the energies of jets because the jet p_{T} spectra decrease steeply and small changes in energy can shift many jets above or below the 30 GeVminimum jet p_{T} threshold. Uncertainties of the jet energy scale and jet energy resolution are included among the systematic uncertainties, as well as the impact of jets from additional pileup interactions.

Since jets are used in the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$, the uncertainties on jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ are correlated and are treated simultaneously. Uncertainties of the jet energy scale and resolution are each propagated to $E_{\mathrm{T}}^{\mathrm{miss}}$. When the energy of the jets in each event

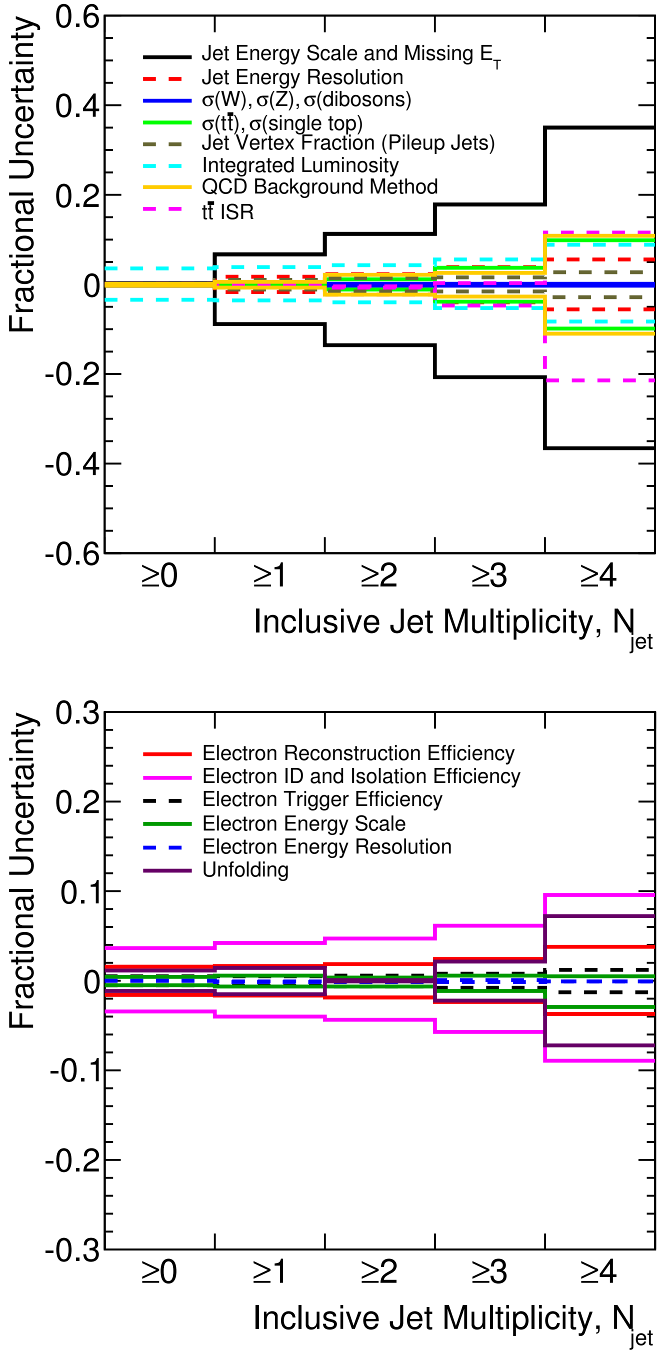

Figure 6.1: Systematic uncertainties as a function of jet multiplicity. Different contributions are shown in the two plots, with the bottom plot zoomed to show smaller contributions.

is modified to evaluate the uncertainty, the net change in jet transverse momentum is subtracted from the $E_{\mathrm{T}}^{\mathrm{miss}}$. Uncertainty of the energy of topoclusters contributing to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation but not within the selected jets is included at the same time. Then the selection requirements of $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV and $M_T(W) > 40$ GeV are reevaluated. Other $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainties are negligible in comparison [67].

Uncertainty of the jet energy scale is by far the dominant uncertainty on this measurement. As noted above, it can change the number of jets passing the minimum jet p_{T} threshold and whether an event passes the $E_{\mathrm{T}}^{\mathrm{miss}}$ and transverse mass requirements for identifying a $W \to e\nu$ event. It also has a direct effect on differential cross section measurements for observables depending on the jet energy, such as the invariant mass of jets.

Figure 6.2 illustrates the jet energy scale uncertainty, which varies with jet transverse momentum. It is determined for different pseudorapidity regions, which are covered by different parts of the calorimeters. Two representative jet pseudorapidity ranges are shown, $|\eta| < 0.3$ in the upper plot and $3.6 < |\eta| < 4.5$ in the lower plot, indicating the range of uncertainties. The total uncertainty, indicated by the shaded region, is largest at low jet transverse momentum. For jets with 20 GeV $< p_{\mathrm{T}} <$ 30 GeV, it increases from 4.1% for jets with $|\eta| < 0.3$ to 13.8% for jets with $3.6 < |\eta| < 4.5$.

There are six main sources of uncertainty of the jet energy scale, as shown in Fig. 6.2: uncertainty of the simulation of the production and interaction of particles within jets (triangles), calibration procedure (crosses), calorimeter response (open squares), simulation of the detector noise levels (inverted triangles), underlying event modeling (solid circles), and the amount of material in the detector (solid squares). Uncertainty due to additional pileup interactions is also included, depending on the number of reconstructed primary vertices in each event. For typical events with two vertices, the pileup uncertainty is less

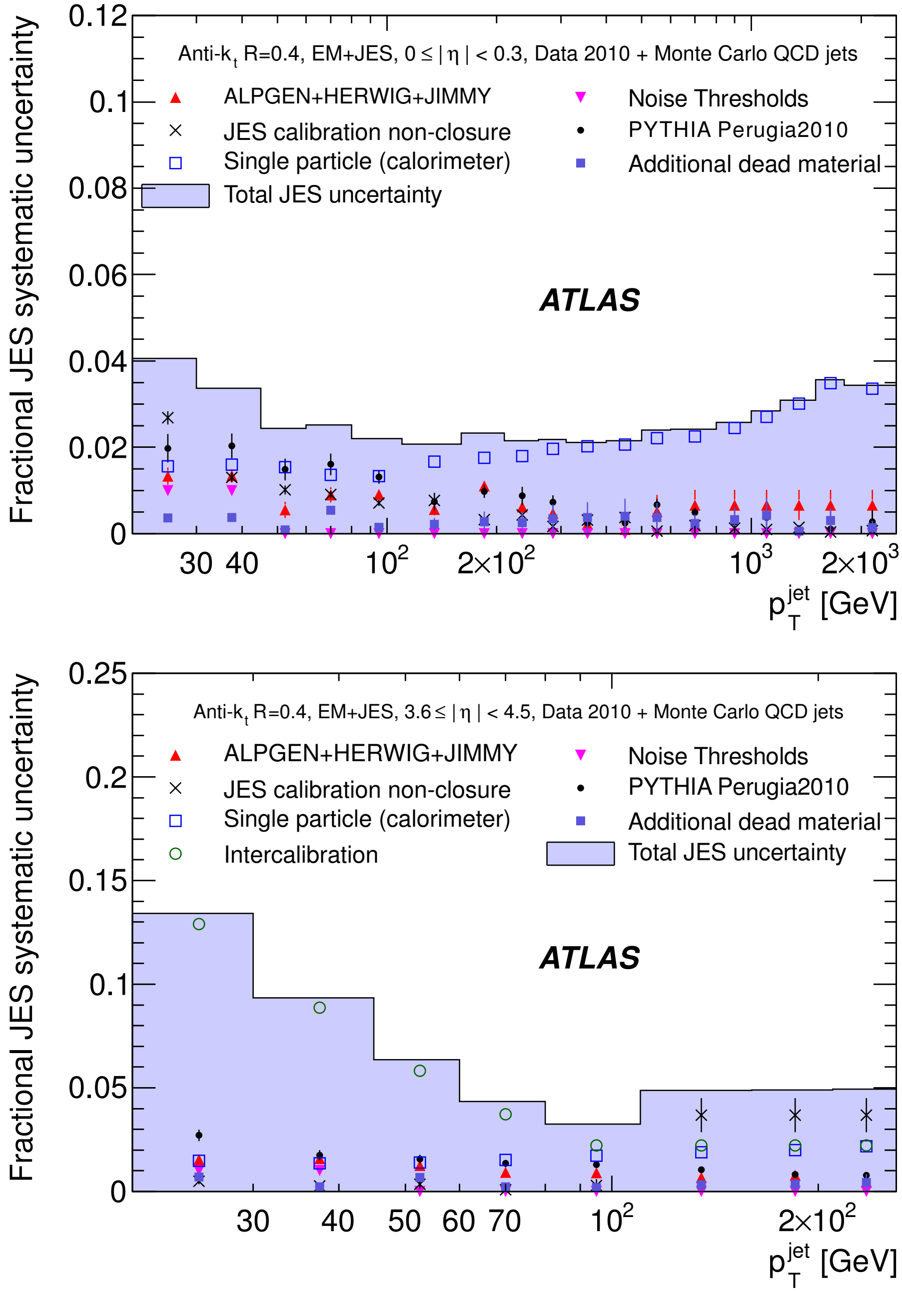

Figure 6.2: Systematic uncertainty of the jet energy scale as a function of jet transverse momentum for jets with the smallest pseudorapidities used in this analysis, $|\eta| < 0.3$ (top), and the largest, $3.6 < |\eta| < 4.5$ (bottom). The total uncertainty, obtained by summing the non-closure contribution linearly with the others in quadrature, is shown by the shaded region.

than 3%.

These uncertainties have been studied mainly using simulated events, supplemented by dedicated beam tests with portions of the calorimeter, as detailed in [56]. The jet energy scale has also been tested using data, for example by studying the transverse energy balance in events with two jets or a jet and photon, confirming the results obtained from the simulated events.

For forward pseudorapidity regions as shown in the lower plot of Fig. 6.2, there is an additional intercalibration uncertainty (open circles). This is determined by balancing the transverse momenta of jets in different pseudorapidity regions, for events with two jets. One of the jets is required to be in the region under consideration and the other in a more central region, where the jet energy scale is known more precisely.

As mentioned in Section 3.6.2, there are additional uncertainties of the jet energy scale in W+jets events due to differences from the dijet events used to obtain the jet energy scale calibration. The first comes from the higher measured energy for quark jets as are likely in W+jets than for gluon jets, which dominate the calibration sample. The difference has been parametrized as a function of the fraction of quark vs. gluon jets using simulated events and tested with data. The fraction of quark vs. gluon jets in the $W \to e\nu$ event sample simulated with ALPGEN, as a function of jet p_{T} and η, is used to determine the possible impact on the jet energy scale for each jet, which is included in the systematic uncertainty.

The second additional uncertainty of the jet energy scale comes from other jets within angular separation $\Delta R < 1.0$ from the jet of interest. These are excluded from the calibration sample. The size of this effect has been measured as function of the angular separation of each jet from its nearest neighbor. It contributes less than 3% uncertainty of the jet energy scale in all cases.

Uncertainty of the jet energy resolution has also been studied using simulated events and tested with data. The uncertainty is expressed as a fraction of the jet transverse momentum resolution, $\sigma_{p_{\mathrm{T}}}$, which is parameterized as a function of jet p_{T}:

$$\frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} = \frac{N}{p_{\mathrm{T}}} \oplus \frac{S}{\sqrt{p_{\mathrm{T}}}} \oplus C \tag{6.1}$$

N, S, and C are constants characterizing the noise, stochastic fluctuations, and energy-independent fluctuations of the energy measurements, respectively. These vary for different calorimeters and have been determined from fits to Equation 6.1 using simulated events in different pseudorapidity regions, as shown in Table 6.1. The uncertainties of the parameters correspond to an uncertainty of about 10% on the energy of each jet, ranging up to 14%. To determine the effect of this uncertainty on the measured cross sections, the energies of the jets in each event are randomly smeared to produce the appropriate increase in resolution. As the simulated resolution could also be larger than the true resolution but cannot be unsmeared, the resulting effect on the measured cross sections is taken symmetrically.

Jet pseudorapidity	S ($\sqrt{\text{GeV}}$)	N (GeV)	C
$0 < \lvert\eta\rvert \leq 0.8$	1.14 ± 0.02	2.1 ± 0.2	0.048 ± 0.005
$0.8 < \lvert\eta\rvert \leq 1.2$	1.12 ± 0.03	1.6 ± 0.2	0.045 ± 0.008
$1.2 < \lvert\eta\rvert \leq 2.1$	1.07 ± 0.03	1.9 ± 0.2	0.05 ± 0.01
$2.1 < \lvert\eta\rvert \leq 2.8$	0.93 ± 0.03	1.1 ± 0.1	0.05 ± 0.01

Table 6.1: Parameters characterizing the jet energy resolution in different pseudorapidity regions according to Equation 6.1.

A final jet uncertainty comes from jets associated with additional proton-proton interactions. As discussed in Section 3.6.5, such jets are identified using the JVF criterion. For a conservative estimate of their impact on the cross section measurement, the difference in the data with and without this criterion is included among the systematic uncertainties.

6.2 Other uncertainties

Table 6.2 lists all of the uncertainties that are considered and their sizes. The effects of each uncertainty are propagated both to the normalization of the inclusive cross sections for $W \to e\nu$ produced in association with zero or more through four or more jets and to each observable, such as jet transverse momentum, for the differential cross sections. The uncertainties are treated as uncorrelated and added in quadrature. The first three uncertainties, which are related to jets, were discussed in the previous section.

Uncertainty	Size applied
Jet energy scale	$\pm 3 - 10\%$ of jet energy (dependent on p_{T} and η)
Jet energy resolution	$\sim$10% of jet energy (varied for each jet)
Pileup jet removal	Compare data with and without JVF requirement
Trigger efficiency	$\pm 0.5\%$
Electron reconstruction eff.	$\pm 1.5\%$
Electron identification eff.	$\pm 1 - 8\%$ (dependent on E_{T} and η) $\oplus \pm 2\%$
Electron energy scale	$\pm 0.3 - 1.6\%$ of electron energy (dependent on E_{T} and η)
Electron energy resolution	$< 0.6\%$ of electron energy
W, Z background normalization	$\pm 5\%$ of background cross section
Top background normalization	$\pm 7\%$ of background cross section
$t\bar{t}$ background shape	Simulated events with modified QCD radiation
QCD background method	See Section 4.1
Unfolding method	See Section 5.2
Luminosity	3.4% of luminosity

Table 6.2: Sources of systematic uncertainty.

Uncertainties of the electron trigger, reconstruction, and identification efficiencies are derived from data as discussed in Section 3.4.3. The trigger and reconstruction efficiencies uncertainties are independent of electron E_{T} and η, while the larger identification uncertainty varies from 1% to 8% and is combined with the 2% uncertainty of the isolation efficiency. The impact of these uncertainties increases with jet multiplicity, as shown in Fig. 6.1, due to the increasing background significance.

Like the jet energy scale and resolution uncertainties, uncertainty on the electron energy scale and resolution can affect the measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$. These electron uncertainties are thus propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ and $M_T(W)$, and the selection is reevaluated as for the jet uncertainties. Uncertainty of the energy scale, measured from data as described in Section 3.4.3, varies from 0.3 to 1.6% as a function of electron E_{T} and η. The resulting impact on the cross section for events with one or more jets is about 0.7%. The resolution uncertainty measured from data, which was described in the same section, is less than 0.6% of the electron energy. This leads to an uncertainty of less than 0.1% of the cross section for events with one or more jets.

The uncertainties of the backgrounds include the cross section used to normalize the simulated background samples, the $t\bar{t}$ background shape uncertainty as discussed in Section 4.2, and uncertainty of the QCD background. The last uncertainty includes both uncertainties of the estimation method and the effect on the normalization fits from shape changes in the leptonic background template due to the jet, electron, and other background uncertainties on the shape of the leptonic background template.

As discussed in Chapter 5, uncertainty of the unfolding is estimated by taking the difference of the final results from those obtained using $W \rightarrow e\nu$ events simulated with Sherpa instead of Alpgen in the unfolding method. For example, this gives a contribution of about 1.5% to the uncertainty of the cross section for events with one or more jets.

Finally, the luminosity uncertainty is 3.4% overall. Its effect on the measured cross sections can be larger at high jet multiplicity where the backgrounds become signficant, as it enters both through the background normalization and directly in the cross section calculation. For example, a larger integrated luminosity would increase the number of background events subtracted and increase the denominator in the cross section calculation, both decreasing the measured cross section. These effects are unfolded simultaneously to account for the correlation.

6.3 Summary

Table 6.3 compares the relative systematic uncertainties as a function of jet multiplicity. At all multiplicities above zero jets, jet energy scale uncertainty is by far the dominant contribution. Uncertainties associated with the top background (top cross section uncertainty and ISR) increase with jet multiplicity, as expected when the signal-to-background ratio decreases. Meanwhile, electron uncertainties are mainly flat as a function of jet multiplicity, consistent with the independence of jet and electron selection.

Systematic uncertainties of the cross sections (in %)

Uncertainty	$N_{\text{jets}} \geq 0$	$N_{\text{jets}} \geq 1$	$N_{\text{jets}} \geq 2$	$N_{\text{jets}} \geq 3$	$N_{\text{jets}} \geq 4$
+JES:	–0.3	9.9	15.1	23.3	41.8
–JES:	0.3	–11.0	–17.9	–28.4	–55.7
JER:	±0.2	±1.6	±2.4	±4.2	±6.1
+Trigger Eff:	–0.5	–0.5	–0.6	–0.8	–1.3
–Trigger Eff:	0.5	0.6	0.6	0.8	1.3
+Ele SF Reco:	–1.5	–1.6	–1.8	–2.4	–3.7
–Ele SF Reco:	1.6	1.7	1.9	2.4	3.8
+Ele SF ID:	–3.4	–4.0	–4.4	–5.7	–8.9
–Ele SF ID:	3.6	4.3	4.7	6.1	9.6
+Ele p_T Scale:	0.5	0.6	0.4	0.6	0.5
–Ele p_T Scale:	–0.5	–0.6	–0.6	–1.1	–2.9
+Ele p_T Res.:	0.0	0.0	–0.1	–0.1	–0.4
–Ele p_T Res.:	–0.0	0.0	–0.1	–0.1	–0.4
JVF (pileup jets):	±0.0	±1.1	±1.5	±1.6	±2.8
QCD method:	±0.2	±0.7	±2.2	±2.6	±10.9
+σ(W,Z,dibosons):	∓0.1	∓0.2	∓0.2	∓0.2	∓0.2
+$\sigma(t\bar{t}$, single top):	∓0.0	∓0.3	∓1.2	∓3.8	∓9.9
+ISR:	–0.0	–0.1	–0.3	0.3	11.5
–ISR:	–0.0	–0.1	–0.7	–4.7	–21.3
+Lumi:	–3.4	–3.6	–4.0	–5.3	–8.2
–Lumi:	3.6	3.8	4.3	5.6	8.8
Unfolding:	±1.1	±1.5	±0.5	±2.0	±6.2

Table 6.3: Relative systematic uncertainties per jet multiplicity.

CHAPTER 7

THEORETICAL PREDICTIONS

As noted in the introduction, recent theoretical advances have allowed the predictions at next-to-leading order precision in the strong coupling constant, α_s, of the cross sections for up to four partons produced in association with the W boson. The additional perturbative QCD corrections between leading order and next-to-leading order reduce the uncertainty on the predictions, in particular through the dependence on certain parameter choises, which are discussed in the first section of this chapter.

All of the inclusive and differential cross sections measured in this analysis are compared with next-to-leading order predictions produced by the BlackHat-Sherpa collaboration [1]. The phase space for the predictions is the same as for the measured cross sections and was described in Section 5.1. BlackHat uses an analytic method to evaluate the virtual loop diagrams, such as the one shown in Fig. 1.2, while the remaining matrix element calculations are done by Sherpa. The uncertainties on these predictions are discussed in the following sections. Corrections of the predictions for certain effects that are not included in the perturbative QCD calculations, and the associated uncertainties, are discussed in the final section.

7.1 Renormalization and factorization scales

The separation of QCD processes into hard scatter and other processes, as discussed in Section 3.2, is not a physical distinction but a convenience to apply perturbative techniques.

It requires the choice of two energy scales to perform the calculations: the renormalization scale, which determines the strength of the QCD coupling, and the factorization scale, which separates the hard scatter parton interactions from interactions with lower momentum transfer, which are characterized by parton distribution functions. The dependence of the predictions on the scale choices is reduced as higher order corrections are included [68].

Although not necessarily the same, these scales are often chosen to be equivalent. For the BlackHat-Sherpa predictions, both scales are set to $H_T/2$, where H_T is the scalar sum of the transverse momenta of all final-state particles. Dependence of the predictions on the scale choice is then included as a systematic uncertainty.

The scale dependence is evaluated by varying each scale by a factor of two, raising the factorization scale, lowering the renormalization scale, or raising and lowering both simultaneously. The uncertainty on each bin of the measured cross section is determined by the variation producing the largest effect on the cross section in that bin. The resulting uncertainty on the predicted cross section for events with one or more jets is less than 8%.

An example of the scale uncertainty is shown in Fig. 7.1 from Reference [1]. The figure illustrates that the next-to-leading order scale dependence is much less than at leading order. This is an important improvement, as scale dependence is still the largest uncertainty on the predictions.

7.2 Parton distribution functions

The BlackHat-Sherpa predictions use the next-to-leading order parton distribution function (PDF) set CTEQ6.6. Conventional methods are applied to evaluate the uncertainty on these PDFs [69]. An orthogonal set of variations characterizing the uncertainties

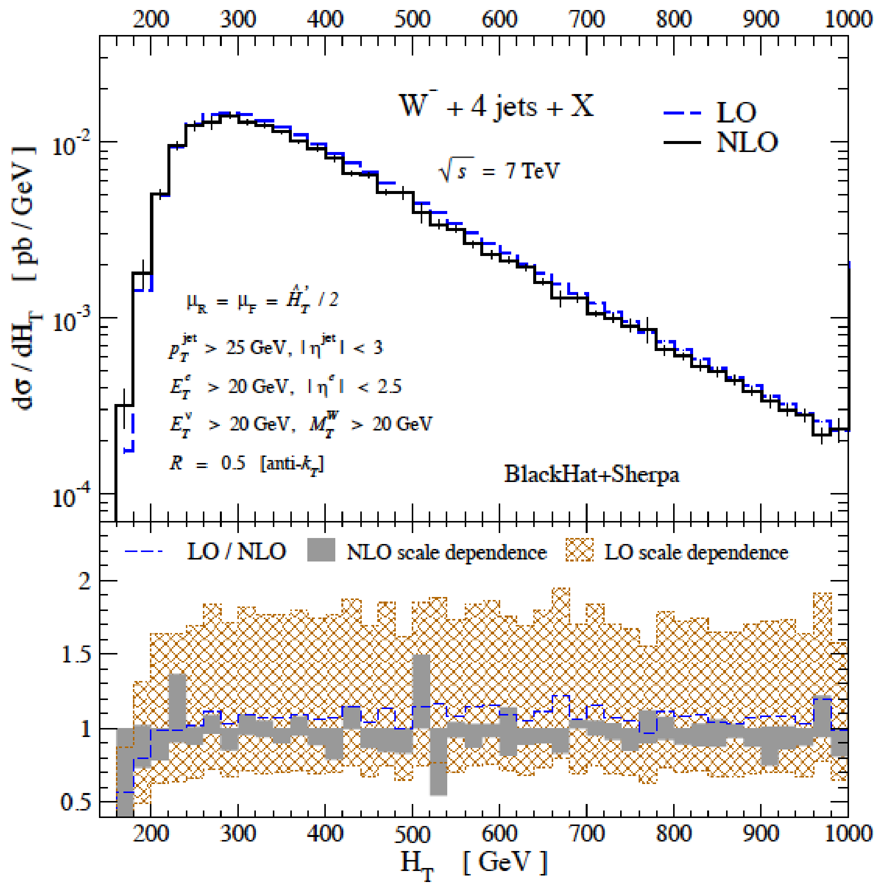

Figure 7.1: Comparison of leading and next-to-leading order BLACKHAT-SHERPA cross section predictions [1] for the H_T distribution (see text) in events containing a W boson plus four jets. The phase space differs slightly from this analysis and is indicated below the curves in the upper part of the figure. Shaded regions show the reduction in scale uncertainty between leading and next-to-leading order.

due to experimental errors on the data and the method applied to obtain the PDFs is used. Alternate PDFs with the strong coupling contant α_s varied by its uncertainty are also considered. The total uncertainty is obtained from the sum in quadrature of the effects of all variations on the predicted cross sections. The resulting uncertainty on the predicted cross section for events with one or more jets is 3.9%. The effect of using other PDF sets instead of CTEQ6.6 was shown to be typically less than the uncertainty obtained in this way [66].

7.3 Non-perturbative corrections

For comparison with the measured cross sections, the BlackHat-Sherpa predictions are corrected for the effects of parton hadronization, underlying event, and QED radiation of photons from the electron, which are not included in the perturbative QCD calculations. The corrections and their uncertainties are described in this section.

Hadronization, the evolution to stable particles of partons produced with the W boson, is not directly included in the BlackHat-Sherpa predictions. A bin-by-bin correction for this process is applied to the predictions for comparison with the measured cross sections, where it is included. The corrections are calculated by taking the ratio of results simulated with Alpgen before and after hadronization of the partons. Two different models of the hadronization, simulated with Herwig and Pythia, are considered, using the average for the corrections with the difference from either model indicating the uncertainty. The resulting correction decreases the cross section for events with one or more jets by about 6% $\pm$ 1%, with little uncertainty.

As discussed in Section 3.2, simulated events typically include modeling of the underlying event, consisting of multiple parton interactions and QCD radiation besides the hard scatter. This affects the measured cross sections and must be corrected in the

BlackHat-Sherpa predictions in order to compare. The corrections and their uncertainty are again obtained by taking the average and half the difference, respectively, of the effects of underlying event modeled with Herwig and Pythia using events simulated with Alpgen. For events with one or more jets, the underlying event increases the cross section by about 7% ± 3%.

The kinematics of the electron from W boson decay can change due to QED radiation of photons. As noted in Section 5.1, in this analysis the p_{T} of any photons within a cone of radius 0.1 around the electron direction is included in the electron p_{T}, giving a well-defined cross section. The BlackHat-Sherpa predictions do not include this effect, so a correction is applied.

The ratio of cross sections simulated using Alpgen with and without QED radiation included is used to correct the BlackHat-Sherpa predictions. The difference from corrections simulated with Sherpa, which models the QED radiation differently, is taken as an estimate of the uncertainty. The corrections are typically less than 2%, with negligible uncertainty.

CHAPTER 8

RESULTS

The first results shown in this chapter are the measured distributions of the jet multiplicity, jet transverse momenta, and other jet observables, before correction for detector effects. After background subtraction, these distributions are used to calculate the cross sections shown in Section 8.2.

8.1 Uncorrected distributions

The figures in this section compare the observed numbers of $W \to e\nu$ events selected from data with the expected numbers of events, as a function of different observables. As described in Chapter 3, expectations for the leptonic backgrounds and $W \to e\nu$ signal are modeled with Monte Carlo simulated events, while the QCD background is estimated from data.

Figure 8.1 shows the jet multiplicity, and Figs. 8.2 through 8.4 show the jet transverse momenta. The magnitudes of the jet transverse momenta are then summed with the electron transverse momentum and $E_{\mathrm{T}}^{\mathrm{miss}}$:

$$H_{\mathrm{T}} = p_{\mathrm{T}}(e) + E_{\mathrm{T}}^{\mathrm{miss}} + \sum_{j=1}^{N_{jet}} p_{\mathrm{T}}(\text{jet j}) \qquad (8.1)$$

The H_{T} distribution is shown in Fig. 8.5. The next figure displays the invariant mass of the first two jets in events with two or more jets, the first three jets in events with three

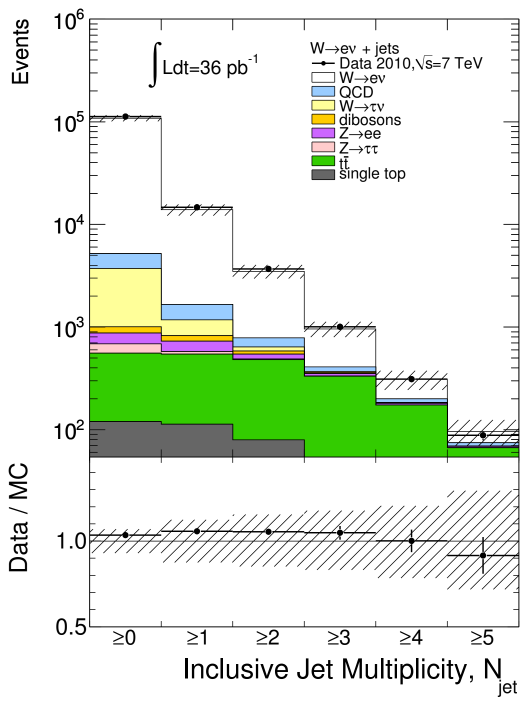

Figure 8.1: Jet multiplicity distribution in $W \to e\nu$ candidate events.

or more jets, and the first four jets in events with four or more jets. Finally, the remaining figures show angular observables for the lepton and first two jets.

In each figure, the data are shown with statistical uncertainties indicated by error bars. Uncertainty of the predicted distributions, including the statistical uncertainty of the simulated event samples, is illustrated by shaded bands. The ratios of data to expectations are shown below the plots. For all observables, the data and expectations exhibit good agreement.

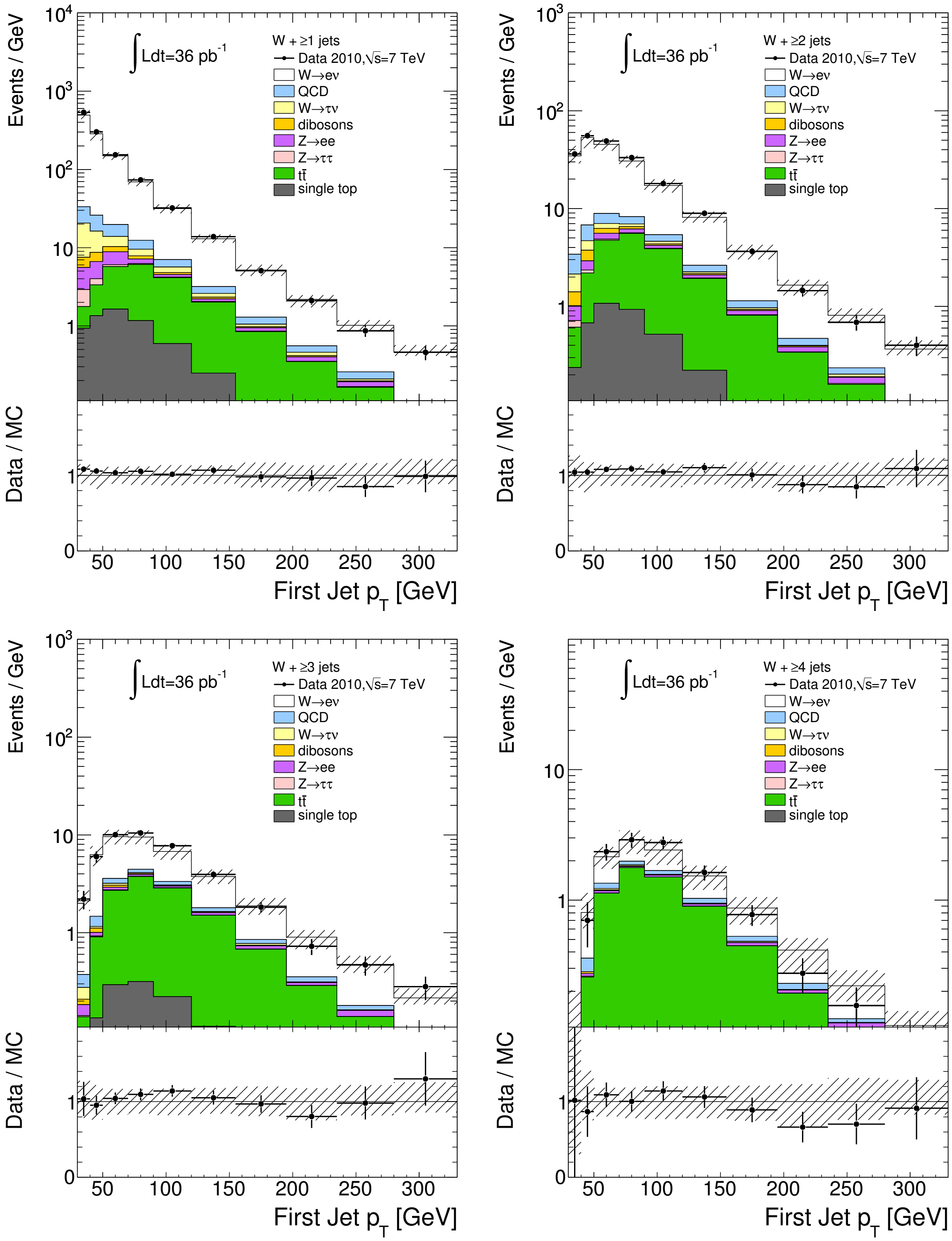

Figure 8.2: Jet p_{T} distribution for the first jet in $W \to e\nu$ candidate events with one or more, two or more, three or more, and four or more jets.

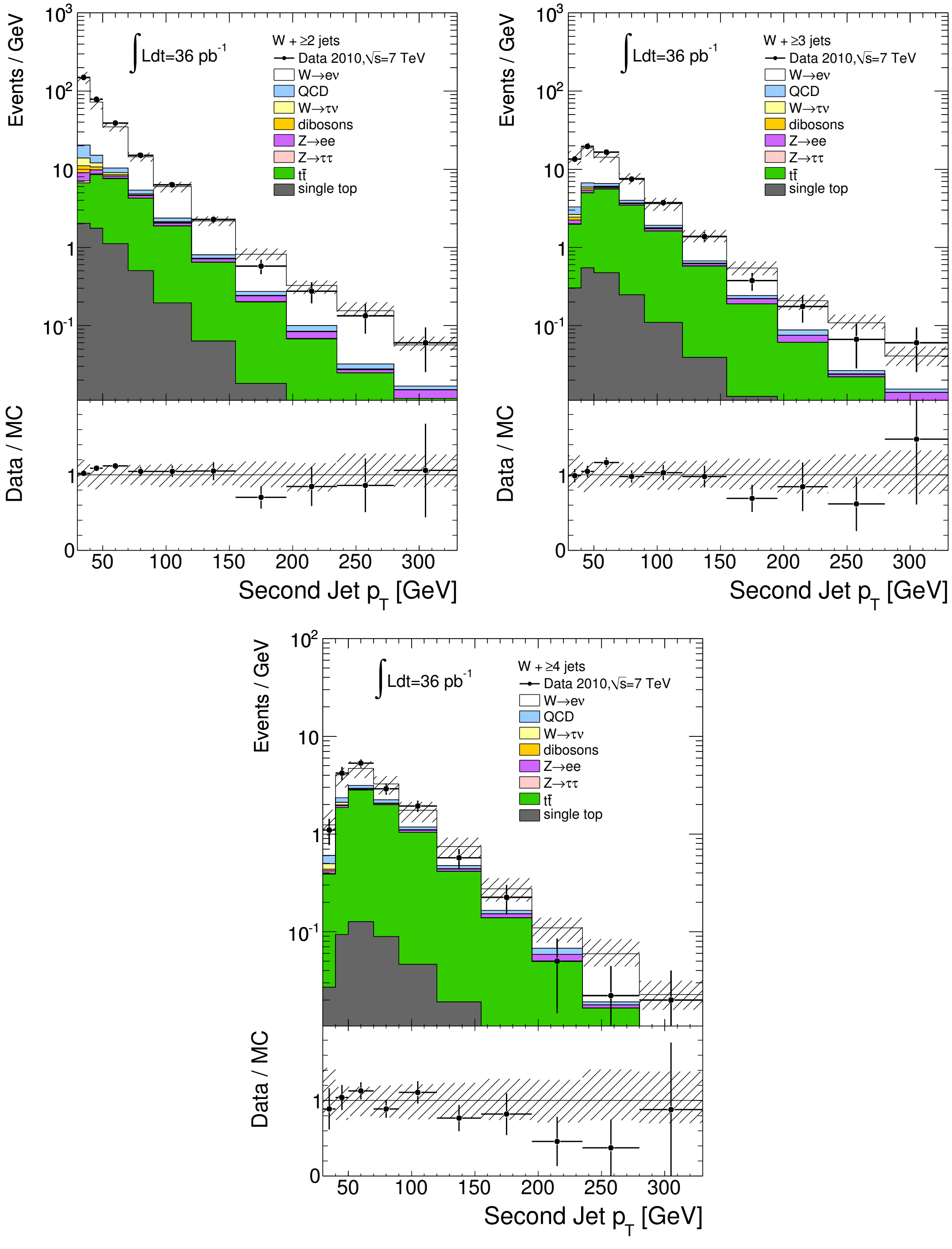

Figure 8.3: Jet p_{T} distribution for the second jet in $W \to e\nu$ candidate events with two or more, three or more, and four or more jets.

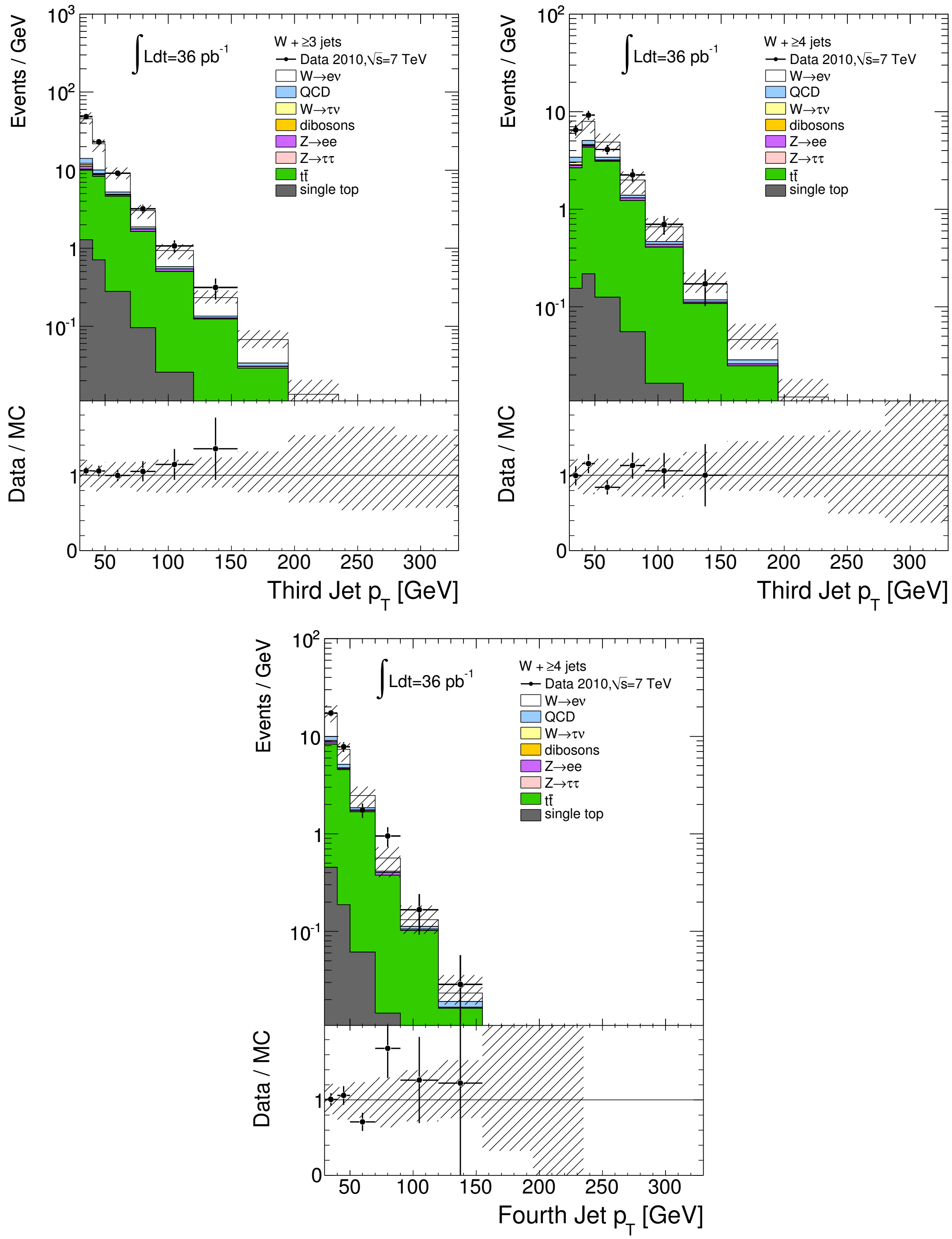

Figure 8.4: Jet p_{T} distribution for the third jet in $W \to e\nu$ candidate events with three or more (top left) and four or more jets (top right) and for the fourth jet in events with four or more jets (bottom).

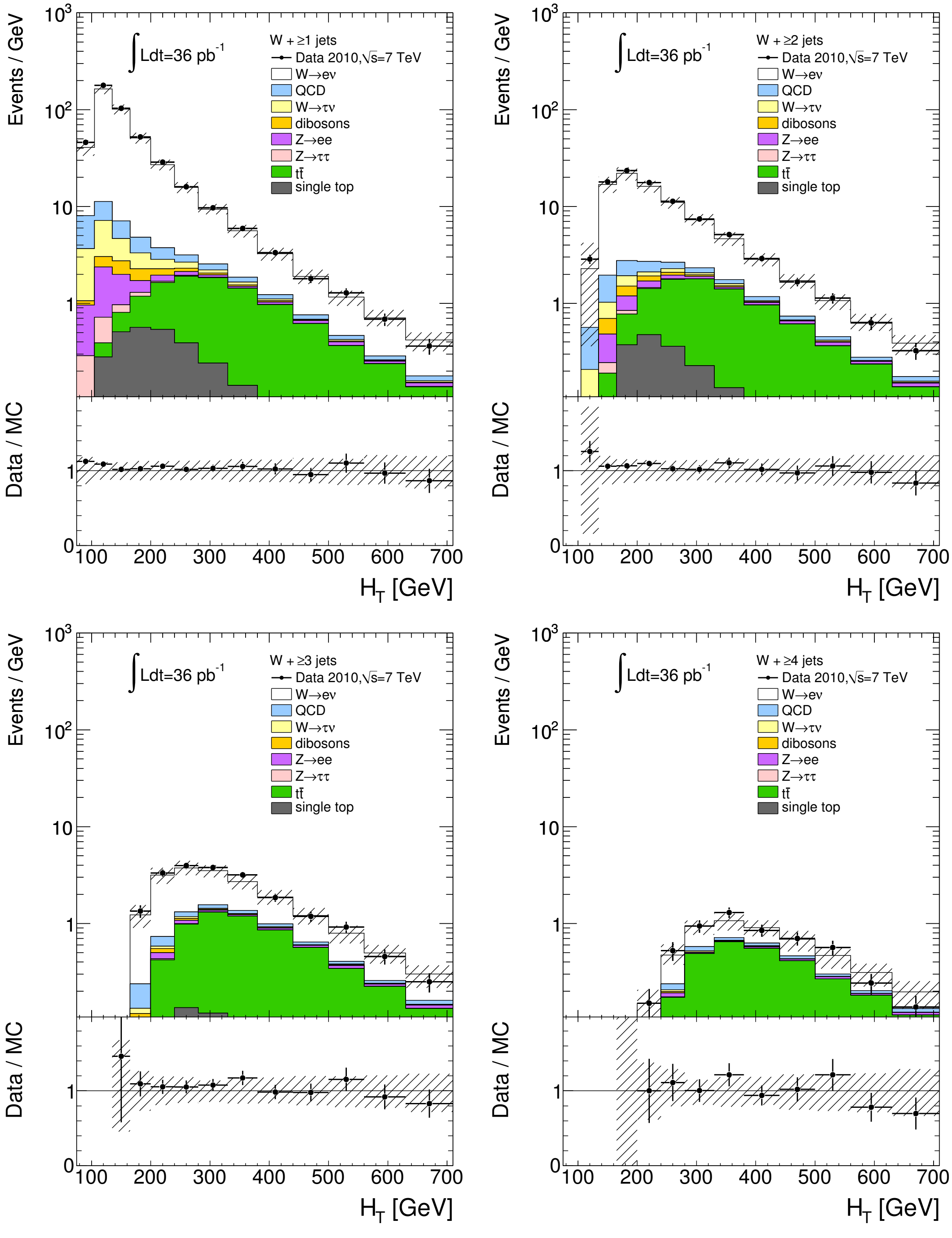

Figure 8.5: H_{T} distribution in $W \to e\nu$ candidate events with one or more, two or more, three or more, and four or more jets.

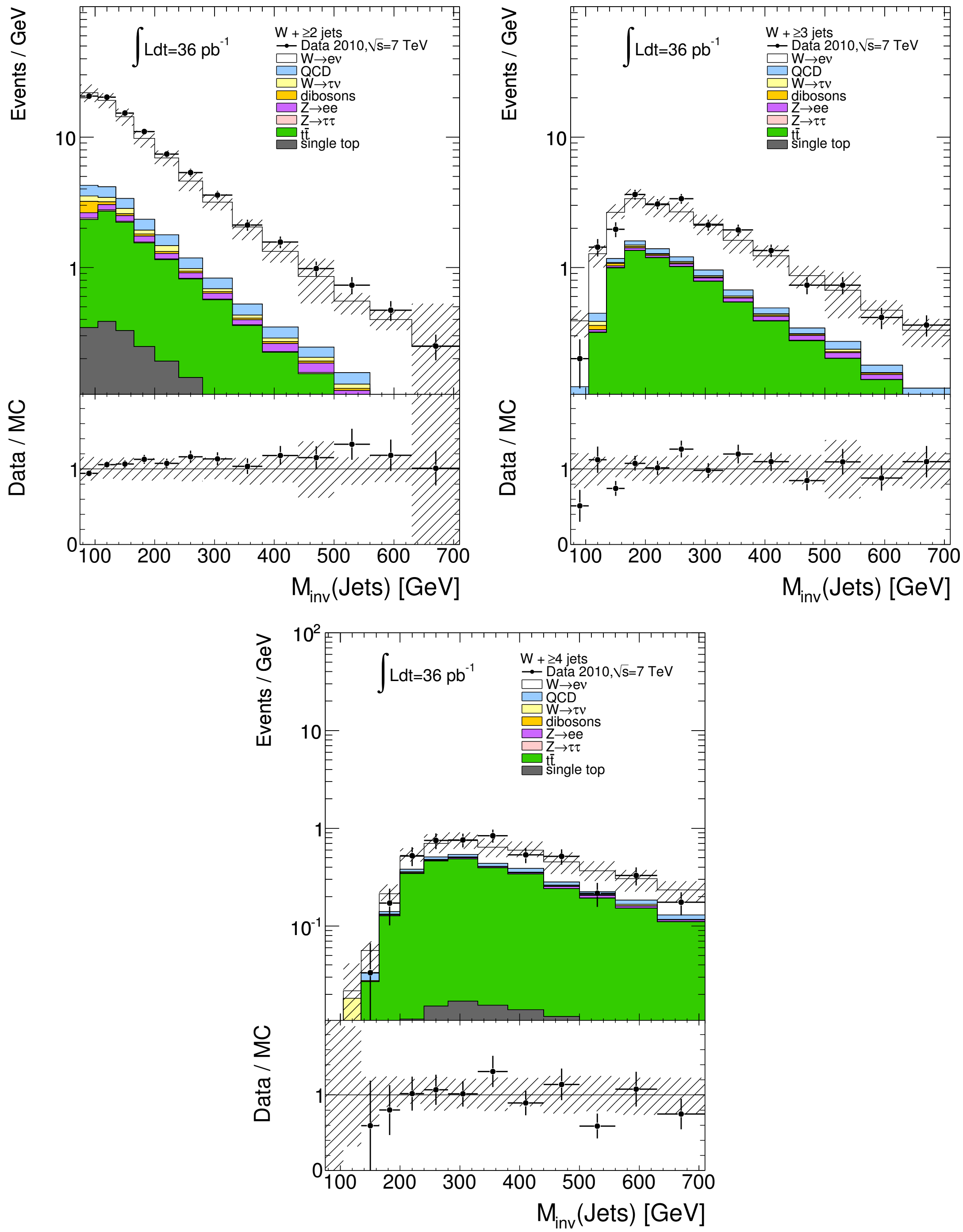

Figure 8.6: Invariant mass of jets in $W \rightarrow e\nu$ candidate events with two or more, three or more, and four or more jets.

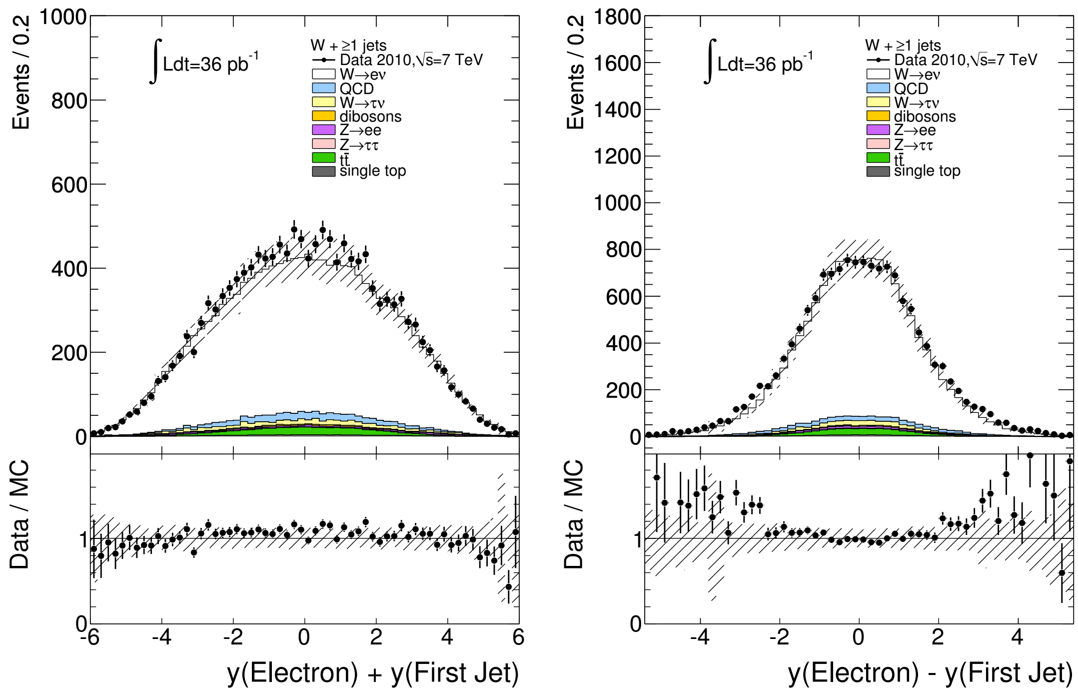

Figure 8.7: Sum (left) and difference (right) of rapidities of the electron and first jet in $W \to e\nu$ candidate events with one or more jets.

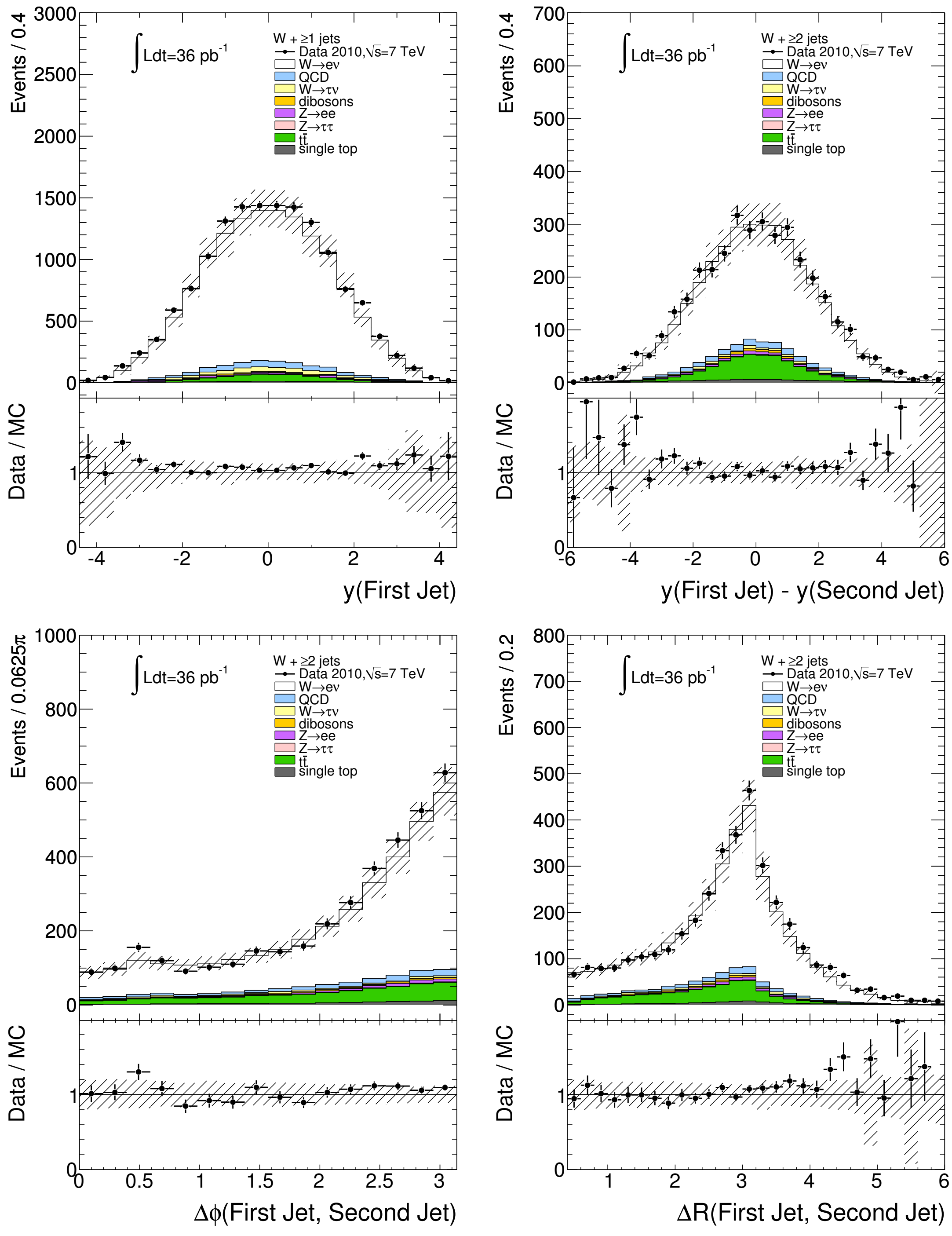

Figure 8.8: Angular distributions for the first two jets in $W \to e\nu$ candidate events with two or more jets, showing the rapidity of the first jet (top left), difference of rapidities of the first two jets (top right), difference of azimuthal angle of the first two jets (bottom left), and angular separation of the first two jets (bottom right).

8.2 Measured cross sections

This section presents the measured inclusive and differential cross sections for $W \to e\nu$ production in association with jets. The results are limited to the phase space defined in Section 5.1 and corrected for all known detector effects as described in Section 5.2. They are used to test the next-to-leading order predictions described in the previous chapter and compared to the predictions of other event simulations.

Table 8.1 shows the inclusive cross section for $W \to e\nu$ produced in association with zero or more through four or more jets. The measured results, including statistical and systematic uncertainties, are listed side-by-side with the next-to-leading order predictions obtained with BlackHat-Sherpa. The predictions are corrected for non-perturbative effects and shown with all of the uncertainties discussed in Chapter 7. For all jet multiplicities, the measured results and predictions agree.

Jet multiplicity	$\sigma \pm$ stat $\pm$ sys (pb)	NLO prediction (pb)
≥ 0	$5284.5 \pm 19.5\ ^{+294.1}_{-277.2}$	$5001.5\ ^{+629.7}_{-665.2}$
≥ 1	$564.1 \pm 6.6\ ^{+62.6}_{-52.3}$	$529.5\ ^{+45.6}_{-42.6}$
≥ 2	$123.1 \pm 2.7\ ^{+19.3}_{-16.9}$	$121.1\ ^{+6.1}_{-8.2}$
≥ 3	$24.6 \pm 1.2\ ^{+5.9}_{-5.4}$	$24.4\ ^{+2.0}_{-3.3}$
≥ 4	$4.4 \pm 0.6\ ^{+2.0}_{-2.1}$	$4.6\ ^{+0.5}_{-1.2}$

Table 8.1: The measured cross section times branching ratio for $W \to e\nu$ produced in association with zero or more through four or more jets, shown with statistical and systematic uncertainties. Next-to-leading order predictions obtained with BlackHat-Sherpa and corrected for non-perturbative effects are shown for comparison.

This is illustrated in Fig. 8.9, which also shows the ratios of cross sections with adjacent jet multiplicities:

$$\sigma(W+\geq N\,\text{jets})/\sigma(W+\geq N-1\,\text{jets}) \tag{8.2}$$

for jet multiplicities, N, from one through four. The combined statistical and systematic uncertainties on the measurements are indicated by black hashing, with the statistical uncertainty alone indicated the error bars on the data points. Predictions from events simulated with Alpgen, Sherpa, and Pythia and the next-to-leading order predictions from BlackHat-Sherpa are shown for comparison.

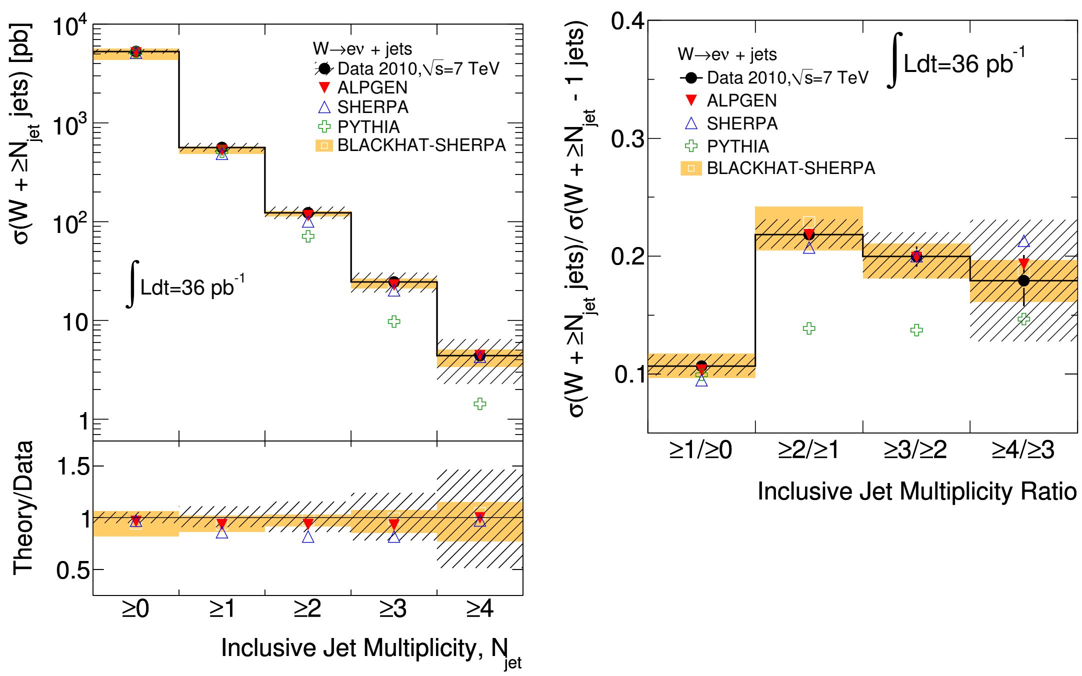

Figure 8.9: Cross section times branching ratio for $W \to e\nu$ produced in association with zero or more through four or more jets (left) and ratios of these cross sections (right).

The lower part of the left plot in Fig. 8.9 shows the ratios of the predictions to the results from data. The predictions from Pythia are excluded from the ratio due to the visibly large differences. For the jet multiplicity ratios shown in the right plot of Fig. 8.9, many uncertainties cancel, *e.g.*, uncertainty of the luminosity, resulting in a smaller total uncertainty. In both plots, there is good agreement between the measured cross sections and all of the theoretical predictions except those obtained with Pythia. This shows that the Pythia leading order parton shower simulation is insufficient to model the additional partons in W+jets production, and no further comparison with PYTHIA is made for the differential cross sections.

Figures 8.10 through 8.13 show the differential cross sections, including $W \to e\nu$ branching ratio, for all of the observables shown in the previous section. The p_{T} of the first through fourth jets in order of highest p_{T} are shown Fig. 8.10. The results are nearly indistinguishable from the predictions until the highest jet multiplicity, events with four or more jets, where the statistics of the data sample become a limitation. Figure 8.11 shows similarly good agreed between results and predictions for the differential cross sections in H_{T}, the scalar sum of jet and electron transverse momenta with $E_{\mathrm{T}}^{\mathrm{miss}}$, and the invariant mass of jets.

The final figures show the differential cross sections for six different angular observables. For the distributions of the sum and difference of rapidities between the electron and first jet, shown in Fig. 8.12, the results and predictions agree to within the uncertainties, which grow at large values of the sum and difference, where at least one of the objects has large rapidity. The agreement for angular distributions of the first two jets in Fig. 8.13 is similar, with slight departures for the SHERPA predictions, particularly in the difference of azimuthal angle between the two jets.

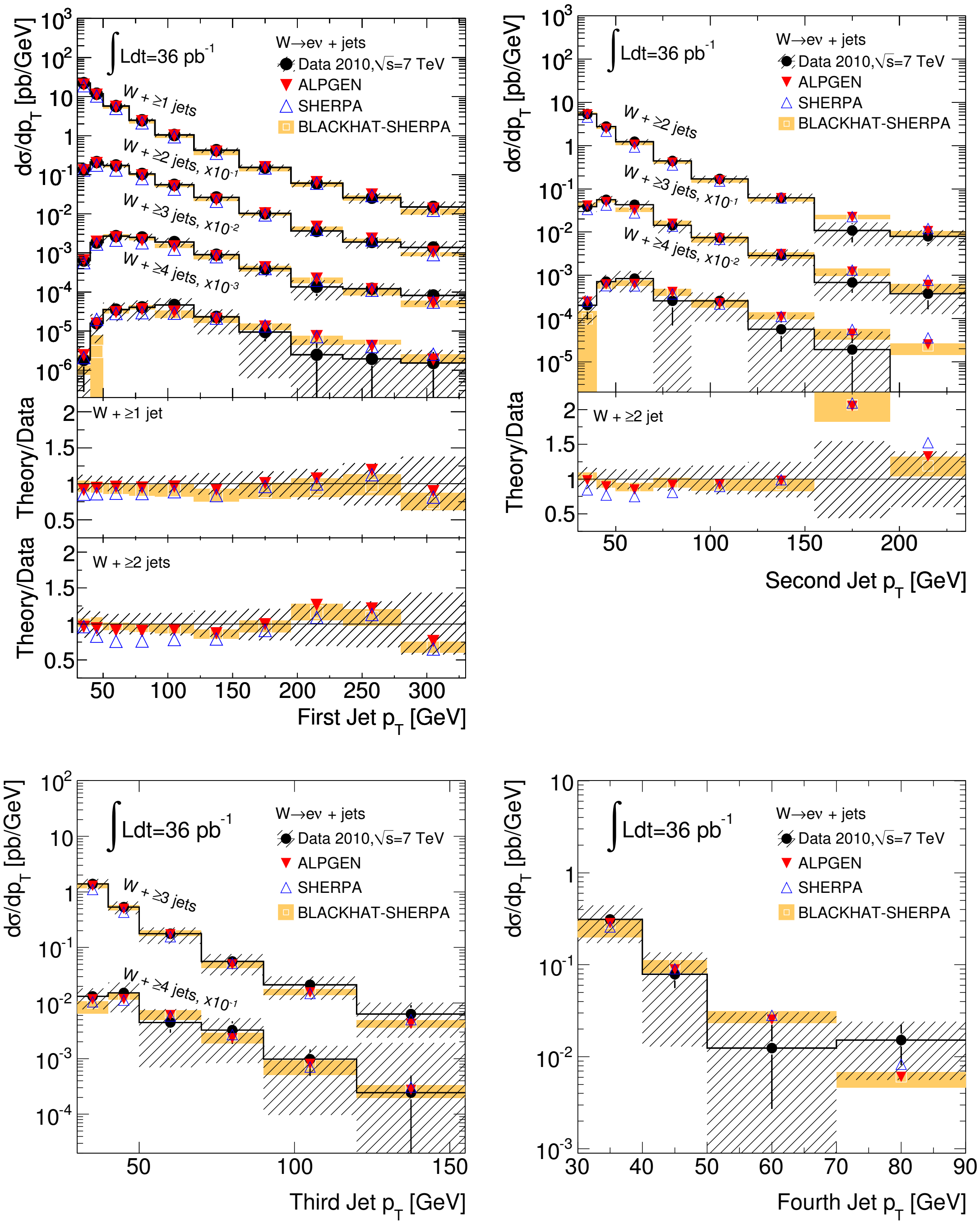

Figure 8.10: Differential $W \to e\nu$ cross section as a function of the jet p_{T}, for the first jet in events with ≥1, 2, 3, and 4 jets (top left), the second jet in events with ≥2, 3, and 4 jets (top right), the third jet in events with ≥3 and 4 jets (bottom left), and the fourth jet in events with ≥4 jets(bottom right).

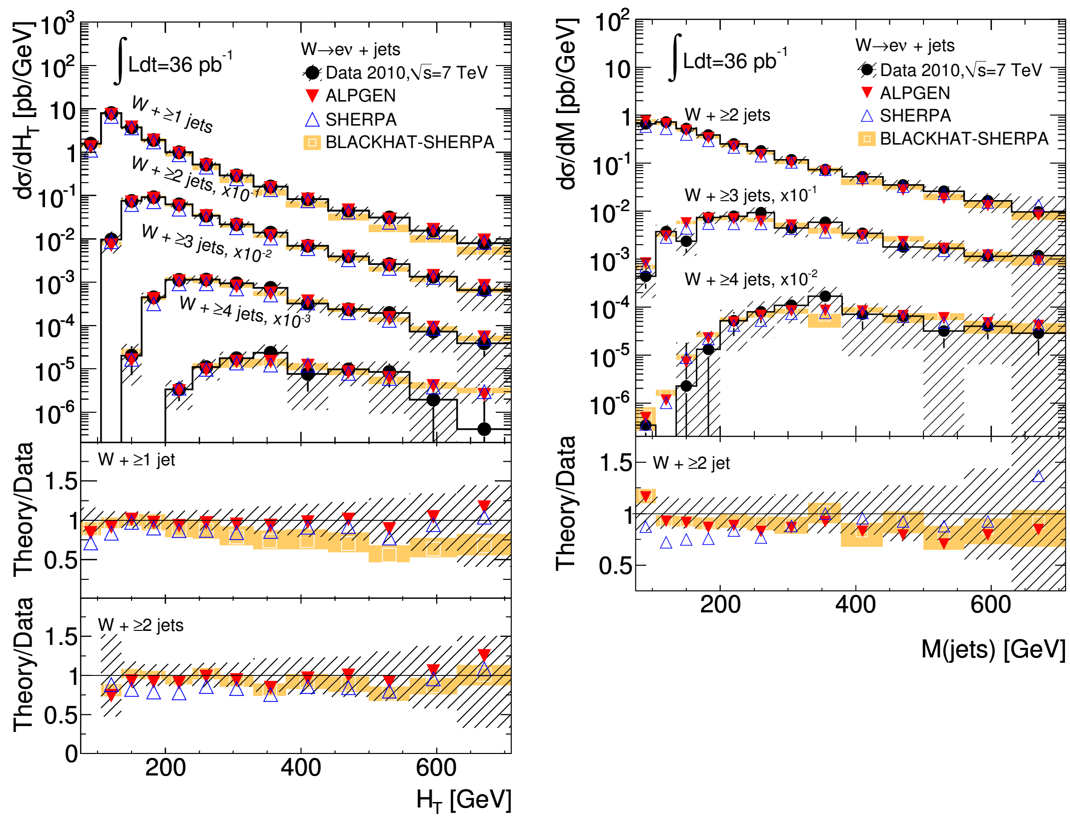

Figure 8.11: Differential $W \to e\nu$ cross section as a function of the H_{T} in events with ≥1, 2, 3, and 4 jets (left) and as a function of the invariant mass of jets in events with ≥2, 3, and 4 jets (right).

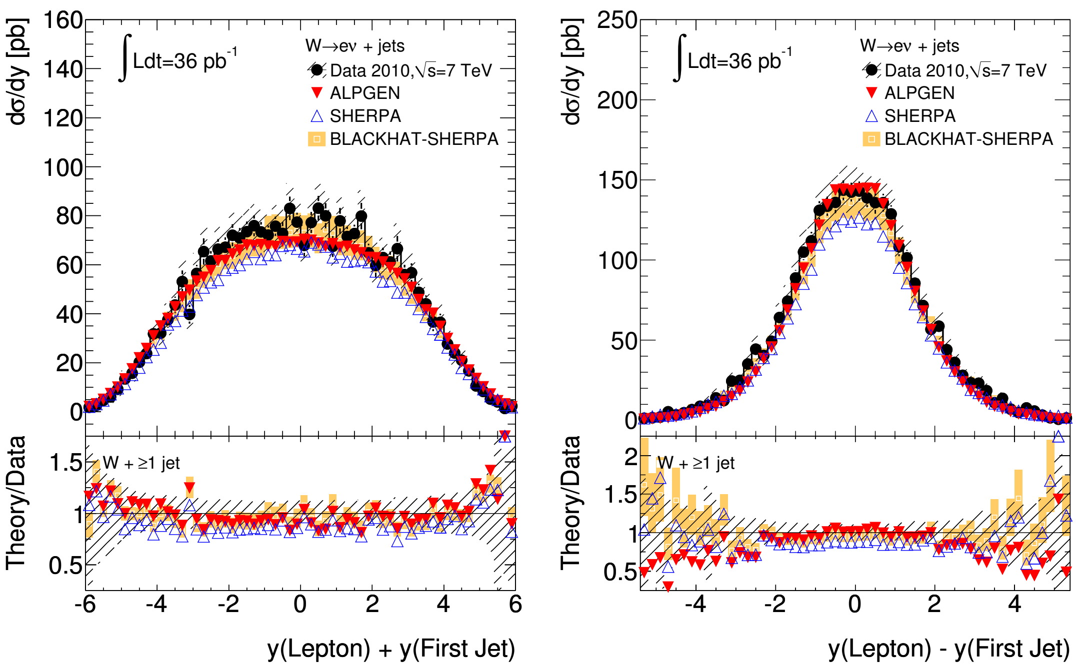

Figure 8.12: Differential cross sections for $W \to e\nu$ plus one or more jets as a function of the sum (left) and difference (right) of rapidities of the electron and first jet.

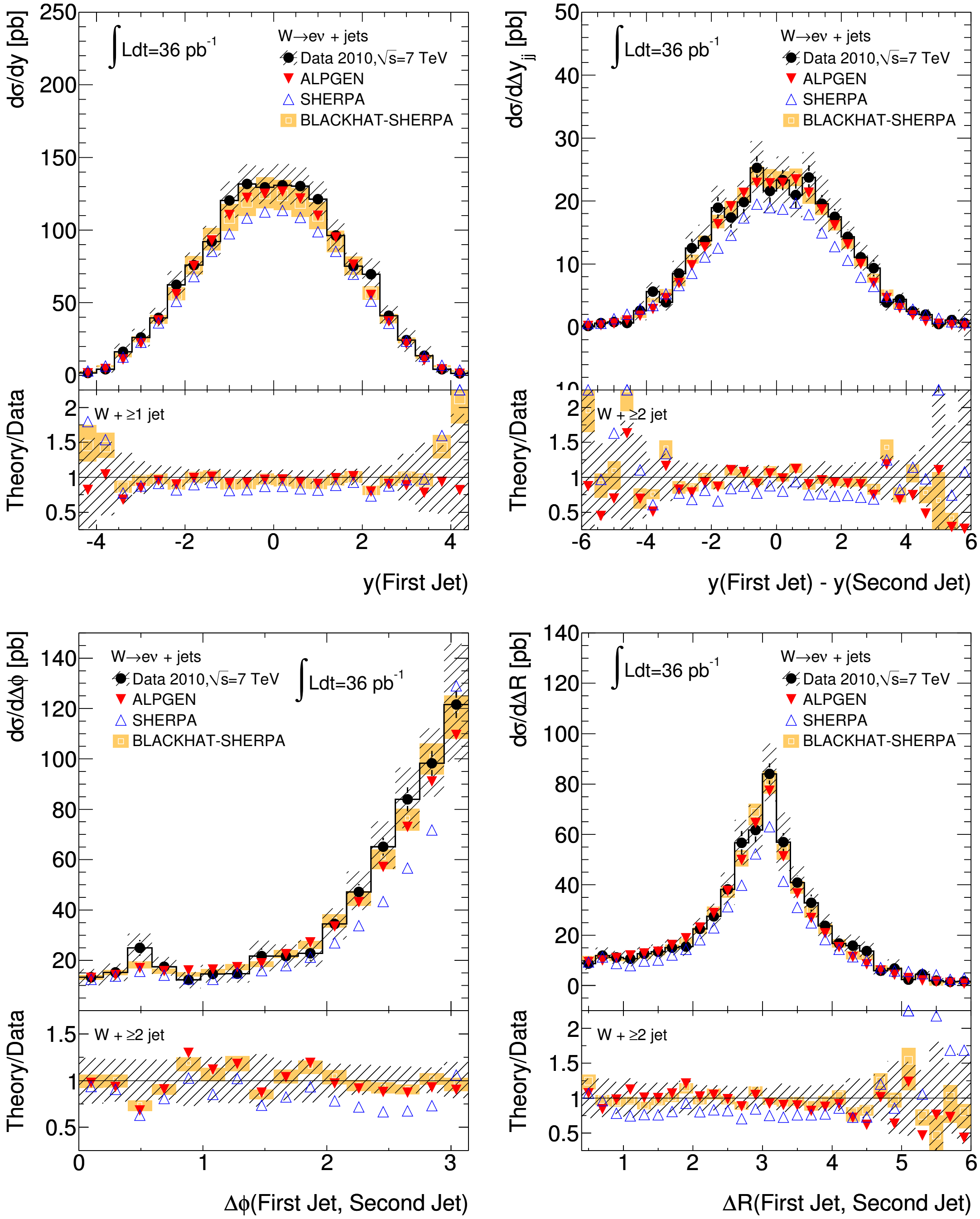

Figure 8.13: Differential cross sections for $W \to e\nu$ plus two or more jets as a function of angular observables for the first two jets, showing the rapidity of the first jet (top left), the difference of rapidities of the first two jets (top right), the difference of azimuthal angle between the first two jets (bottom left), and the angular separation of the first two jets (bottom right).

CHAPTER 9

CONCLUSIONS

The cross section measurements described in this thesis represent the first study at the Large Hadron Collider of the associated production of jets with a W boson, identified by the decay $W \rightarrow e\nu$. The method was developed using a small sample of data collected by the ATLAS detector during the initial weeks of proton-proton collisions at center-of-mass energy $\sqrt{s} = 7$ TeV. It was then extended to use the full data sample collected during 2010, allowing detailed measurements of the jet properties in events containing a W boson.

This analysis addresses the challenges of understanding the detector performance and backgrounds at an unprecedented collision energy. Comparison of the observed numbers of events with expectations demonstrates the validity of the techniques that were applied. Observables such as the jet transverse momenta, invariant mass, and angular distributions are each considered.

The results are corrected for all known detector effects and used to test the predictions of perturbative QCD calculations with next-to-leading order precision. This includes the first experimental test of next-to-leading order predictions for a W boson produced with four or more jets. The predictions describe the data well. For example, the measured cross section times branching ratio for $W \rightarrow e\nu$ plus four or more jets, 4.4 ± 0.6 (stat.) ${}^{+2.0}_{-2.1}$ (sys.) pb, agrees with the prediction of $4.6\ {}^{+0.5}_{-1.2}$ pb.

The largest systematic uncertainty in these measurements comes from uncertainty of the jet energy scale. This could be reduced by further jet calibration studies using larger data samples. However, the effects of additional interactions due to increasing pileup in data collected after 2010 will also need to be taken into account. These measurements provide a point of reference for future studies.

The level of agreement between the measured cross sections and the theoretical predictions for W+jets production exemplifies the robustness of the Standard Model in describing physics at the LHC to date. Searches for new physics will need to look further, benefitting from the experience gained here.

REFERENCES

[1] C. F. Berger *et al.* Precise predictions for $W + 4$ jet production at the Large Hadron Collider. *Phys. Rev. Lett.*, 106:092001, 2010. arXiv:1009.2338v2 [hep-ph].

[2] M. L. Mangano. Understanding the Standard Model, as a bridge to the discovery of new phenomena at the LHC. *Int. J. Mod. Phys. A*, 23, 2008. arXiv:0802.0026v2 [hep-ph].

[3] Standard Model of Elementary Particles. Wikimedia commons (Online). Available at http://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg. Accessed 17 October 2011. Values are from the 2008 Review of Particle Physics. C. Amsler *et al.* (Particle Data Group), *Phys. Lett. B*, 667:1, 2008.

[4] ATLAS Collaboration. Measurement of the production cross section for W-bosons in association with jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Lett. B*, 698:325, 2011. arXiv:1012.5382v2 [hep-ex].

[5] CMS Collaboration. Rates of jets produced in association with W and Z Bosons in pp collisions at $\sqrt{s} = 7$ TeV. *CMS-PAS-EWK-10-012*, 2011.

[6] CDF Collaboration. Measurement of the cross section for W-boson production in association with jets in $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. D*, 77:011108, 2008. arXiv:0711.4044v2 [hep-ex].

[7] B. Cooper. *A measurement of the production of jets in association with a W boson in proton-antiproton collisions at the Tevatron using data collected with the CDF Experiment.* PhD thesis, University College London, 2006. FERMILAB-THESIS-2006-61.

[8] D0 Collaboration. Measurements of inclusive W+jets production rates as a function of jet transverse momentum in $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *submitted to Phys. Lett. B*, FERMILAB-PUB-11-246-E-PPD, 2011. arXiv:1106.1475v1 [hep-ex].

[9] C. F. Berger *et al.* Precise predictions for $W + 3$ jet production at hadron colliders. *Phys. Rev. Lett.*, 102:222001, 2009. arXiv:0902.2760 [hep-ph].

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

[11] G. Aad *et al.* (ATLAS Collaboration). The ATLAS Experiment at the CERN Large Hadron Collider. *JINST 3*, S08003, 2008.

[12] R. Adolphi *et al.* (CMS Collaboration). The CMS Experiment at the CERN LHC. *JINST 3*, S08004, 2008.

[13] D. Froidevaux and P. Sphicas. General-purpose detectors for the Large Hadron Collider. *Annu. Rev. Nucl. Part.*, 56:375–440, 2006.

[14] A. Cho. Particle Physicists' New Extreme Teams. *Science*, 333:1564–1567, 2011.

[15] C. Lefevre. LHC: the guide. http://cdsweb.cern.ch/record/1165534/, 2009. CERN-Brochure-2009-003-Eng.

[16] G. Arnison *et al.* (UA1 Collaboration). Further evidence for charged intermediate vector bosons at the SPS collider. *Phys. Lett. B*, 129:273, 1983.

[17] G. Aad *et al.* (ATLAS Collaboration). Luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J.*, C71:1630, 2011.

[18] ATLAS Collaboration. Updated luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector. Technical Report ATLAS-COM-CONF-2011-011, CERN, Geneva, Mar 2011.

[19] ATLAS Collaboration. Performance of the ATLAS detector using first collision data. *J. High Energy Phys.*, 1009:056, 2010. arXiv:1005.5254v1 [hep-ex].

[20] G. Aad *et al.* (ATLAS Collaboration). ATLAS pixel detector electronics and sensors. *JINST 3*, P07007, 2008.

[21] P. Adragna *et al.* Testbeam studies of production modules of the ATLAS Tile Calorimeter. *Nucl. Instrum. Meth. A*, 606:362–394, 2009.

[22] ATLAS Collaboration. Readiness of the ATLAS liquid argon calorimeter for LHC collisions. *Eur. Phys. J. C*, 70:723–753, 2010.

[23] ATLAS Collaboration. Data-quality requirements and event cleaning for jets and missing transverse energy reconstruction with the ATLAS detector in proton-proton collisions at a center of mass energy $\sqrt{s}$ =7 TeV. Technical Report ATLAS-CONF-2010-038, CERN, Geneva, 2010.

[24] N. Metropolis. The beginning of the Monte Carlo method. *Los Alamos Science*, 15:125–130, 1987.

[25] ATLAS Collaboration. The ATLAS simulation infrastructure. *Eur. Phys. J. C*, 70:823, 2010.

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

[27] ATLAS Collaboration. ATLAS computing technical design report. Technical Report ATLAS-TDR-017, CERN-LHCC-2005-022, CERN, Geneva, 2005.

[28] ATLAS Collaboration. First tuning of HERWIG/JIMMY to ATLAS data. Technical Report ATL-PHYS-PUB-2010-014, CERN, Geneva, Oct 2010.

[29] ATLAS Collaboration. Charged particle multiplicities in pp interactions at $\sqrt{s} = 0.9$ and 7 TeV in a diffractive limited phase space measured with the ATLAS detector at the LHC and a new PYTHIA6 tune. Technical Report ATLAS-CONF-2010-031, CERN, Geneva, 2010.

[30] K. Melnikov and F. Petriello. Electroweak gauge boson production at hadron colliders through $\mathcal{O}(\alpha_s^2)$. *Phys. Rev. D*, 74:114017, 2006. arXiv:hep-ph/0609070v1.

[31] M. A. Dobbs *et al.* Les Houches guidebook to Monte Carlo generators for hadron collider physics. In *PHYSTAT2011*, Les Houches, France, Mar 2004. Working group on Quantum ChromoDynamics and the Standard Model, Physics at TeV Colliders. arXiv:hep-ph/0403045v2.

[32] M. L. Mangano *et al.* ALPGEN, a generator for hard multiparton processes in hadronic collisions. *J. High Energy Phys.*, 0307:001, 2003. arXiv:hep-ph/0206293.

[33] J. Pumplin *et al.* New generation of parton distributions with uncertainties from global QCD analysis. *J. High Energy Phys.*, 0207:012, 2002. arXiv:hep-ph/0201195v3.

[34] G. Corcella *et al.* HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *J. High Energy Phys.*, 0101:010, 2001. arXiv:hep-ph/0011363.

[35] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour. Multiparton interactions in photoproduction at HERA. *Z. Phys. C*, 72:637, 1996.

[36] J. Alwall *et al.* Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J. C*, 53:473–500, 2008. arXiv:0706.2569v2 [hep-ph].

[37] T. Gleisberg *et al.* Event generation with SHERPA 1.1. *J. High Energy Phys.*, 0902:007, 2009.

[38] P. M. Nadolsky *el al.* Implications of CTEQ global analysis for collider observables. *Phys. Rev. D*, 78:013004, 2008. arXiv:0802.0007v3 [hep-ph].

[39] S. Mrenna T. Sjöstrand and P. Skands. PYTHIA 6.4 physics and manual. *J. High Energy Phys.*, 0605:026, 2006.

[40] A. Sherstnev and R. S. Thorne. Parton distributions for LO generators. *Eur. Phys. J. C*, 55:553, 2008. arXiv:0711.2473v3 [hep-ph].

[41] N. Davidson *et al.* Universal interface of TAUOLA, technical and physics documentation. arXiv:1002.0543 [hep-ph], 2010.

[42] S. Frixione and B. R. Webber. Matching NLO QCD computations and parton shower simulations. *J. High Energy Phys.*, 0206:029, 2002.

[43] S. Frixione, P. Nason, and C. Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *J. High Energy Phys.*, 0711:070, 2007.

[44] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann. Hathor - hadronic top and heavy quarks cross section calculator. *Computer Physics Communications*, 182(4):1034 – 1046, 2011.

[45] B. P. Kersevan and E. Richter-Was. The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5. arXiv:hep-ph/0405247, 2004.

[46] J. Smith *et. al.* Transverse mass and width of the W boson. *Phys. Rev. Lett.*, 50, 1983.

[47] ATLAS Collaboration. Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data. *submitted to Eur. Phys. J. C*, CERN-PH-EP-2011-117, 2011. arXiv:1110.3174 [hep-ex].

[48] ATLAS Collaboration. Readiness of the ATLAS Liquid Argon Calorimeter for LHC Collisions. *Eur. Phys. J. C*, 70:723–753, 2010. arXiv:0912.2642v4 [physics.ins-det].

[49] G. Aad *et. al* (ATLAS Collaboration). Performance of missing transverse momentum reconstruction in proton-proton collisions at 7 TeV with ATLAS. *submitted to Eur. Phys. J. C*, CERN-PH-EP-2011-114, 2011. arXiv:1108.5602v1 [hep-ex].

[50] T. Barillari *et al.* Local Hadronic Calibration. Technical Report ATL-LARG-PUB-2009-001, CERN, Geneva, 2009.

[51] M. Cacciari, G. Salam, and G. Soyez. The anti-kt jet clustering algorithm. *J. High Energy Phys.*, 0804, 2008. arXiv:0802.1189v2 [hep-ph].

[52] L. Asquith *et. al.* Performance of jet algorithms in the ATLAS detector. Technical Report ATL-PHYS-INT-2010-129, CERN, Geneva, 2010.

[53] G. Salam, M. Cacciari, and G. Soyez. Dispelling the N^3 myth for the kt jet-finder. *Phys. Lett. B*, 641:57–61, 2006. arXiv:hep-ph/0512210v2.

[54] G. Aad *et. al* (ATLAS Collaboration). Measurement of inclusive jet and dijet cross sections in proton-proton collisions at 7 TeV centre-of-mass energy with the ATLAS detector. *Eur. Phys. J. C*, 71, 2011. arXiv:1009.5908v2 [hep-ex].

[55] ATLAS Collaboration. Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV in ATLAS 2010 data. Technical Report ATL-CONF-2011-032, CERN, Geneva, 2011.

[56] G. Aad *et. al* (ATLAS Collaboration). Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV in 2010. *to be submitted to Eur. Phys. J. C*, ATL-PERF-2011-03, 2011.

[57] D. Miller, A. Schwartzman, and D. Su. Pile-up jet energy scale correction using the jet-vertex fraction method. Technical Report ATL-COM-PHYS-2009-180, CERN, Geneva, 2009.

[58] CERN ROOT Group. http://root.cern.ch/root/htmldoc/TFractionFitter.html.

[59] R. Barlow and C. Beeston. Fitting using finite Monte Carlo samples. *Comp. Phys. Comm.*, 77:219–228, 1993.

[60] K. Lohwasser. *The W charge asymmetry: measurement of the proton structure with the ATLAS detector.* PhD thesis, University of Oxford, 2010. CERN-THESIS-2010-069.

[61] ATLAS Collaboration. Measurement of the $W \to \ell\nu$ and Z/gamma* $\to ll$ production cross sections in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *J. High Energy Phys.*, 1012, 2010. arXiv:1010.2130v1 [hep-ex].

[62] ATLAS Collaboration. Measurement of the cross section for the production of a W boson in association with b-jets in pp collisions at $\sqrt{s} = 7$ TeVwith the ATLAS detector. *submitted to Phys. Lett. B*, CERN-PH-EP-2011-132, 2011. arXiv:1109.1470v1 [hep-ex].

[63] G. Cowan. *Statistical Data Analysis*, chapter 11. Oxford University Press, USA, 1998.

[64] G. D'Agostini. A Multidimensional unfolding method based on Bayes' theorem. *Nucl. Instrum. Meth. A*, 362:487–498, 1995.

[65] T. Adye. Unfolding algorithms and tests using RooUnfold. In *PHYSTAT2011*, Geneva, May 2011. CERN, to be published. arXiv:1105.1160v1 [physics.data-an].

[66] G. Arabidze *et. al.* Measurement of the cross-section for jets produced in association with a W-boson in pp collisions at $\sqrt{s} = 7$ TeV. Technical Report ATL-COM-PHYS-2011-769, CERN, Geneva, 2011.

[67] A. Ahmad *et. al.* Measurement of the cross-section for jets produced in association with a W-boson in pp collisions at $\sqrt{s}$ =7 TeV. Technical Report ATL-PHYS-INT-2011-020, CERN, Geneva, 2011.

[68] 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.

[69] S. Alekhin *et. al.* The PDF4LHC Working Group Interim Report. arXiv:1101.0536v1 [hep-ph], 2011.