A MEASUREMENT OF W BOSON PRODUCTION IN ASSOCIATION WITH JETS AT THE ATLAS DETECTOR IN PP COLLISIONS AT $\sqrt{S} = 8$TEV

By

Matthew Mondragon

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Physics - Doctor of Philosophy

2018

ABSTRACT

A MEASUREMENT OF W BOSON PRODUCTION IN ASSOCIATION WITH JETS AT THE ATLAS DETECTOR IN PP COLLISIONS AT $\sqrt{S} = 8$TEV

By

Matthew Mondragon

This thesis presents a measurement of W boson production in association with jets using 20.2 fb^{-1} of data of proton-proton collisions at the ATLAS detector. The collisions were at a center of mass energy of $\sqrt{s} = 8$ TeV, and all the data were collected in 2012. Differential cross sections and the ratio of W^+ to W^- cross sections are shown for $W \to e\nu$ decays in the presence of jets with transverse momentum $p_{\mathrm{T}} > 30$ and jet rapidity $|\eta| < 4.4$. The observables shown have proven difficult to model in the past, including the transverse momentum of the W boson, and the scalar sum of the transverse momenta of all hard objects H_{T}. These observables test the theoretical understanding of perturbative QCD. The data are compared to leading order, next-to leading order, and next-to-next-to leading order predictions in QCD. In the W^+/W^- ratio many systematic uncertainties cancel out, improving the precision of the measurement by up to a factor of nine.

To my family, and my friends

ACKNOWLEDGMENTS

If I was asked what was the most impactful part of my graduate experience, I would certainly have to say it was the many wonderful friendships I've forged in the many years of living here. Many have come and gone, but without my friends I could never have accomplished this great task. Thank you to all of you for making my twenties fun and this time worthwhile. I'd also like to extend my love to my family, who I know I must've caused to worry constantly, and thank them for their understanding and self-restraint for not asking me "when are you going to finish"... at least not every time I called them.

And lastly, a big thanks to my advisor Joey Huston, who always provided guidance when I asked, a cool head when I was stressed, and an infinite amount of patience on this long journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

This thesis presents a measurement of the production of W bosons in association with jets (W + jets) for the data collected from the ATLAS experiment in 2012. During this run, ATLAS collected $20.2\,\text{fb}^{-1}$ worth of data from proton-proton collisions at a center of mass energy of $\sqrt{s} = 8\,\text{TeV}$. In particular, this analysis focuses on leptonic W boson decay to an electron and a neutrino with the addition of at least one or two jets.

There are several reasons to study W + jets at the Large Hadron Collider (LHC). Because of the large cross section, these events allow for a rigorous test of perturbative Quantum Chromodynamics (pQCD), and so, a lot of work has been put into the theoretical predictions of W+jets events in recent years. W production has been calculated at next-to-leading order (NLO) for events with up to as many as 5 jets [13], [14], at next-to-next-to-leading order for events with 1 jet [15], and there have been improvements to the inclusive leading order (LO) predictions as well. Comparing these predictions to data demonstrates in what regions pQCD is well understood, with what precision, and where the predictions break down. Often it is the case that some predictions do well in a certain region of phase space, and some are superior elsewhere.

Furthermore, W + jets events are a major background to many other Standard Model and beyond the Standard Model signatures. For example, a single top quark decay can also contain a neutrino, a lepton, and jets, which is the same final state products of a W + jets

event. Having a strong understanding of W + jets production allows other analyses to suppress this background and perform a more precise measurement.

In the past, W + jets has been studied at ATLAS using data collected in 2011 with center of mass energy of $\sqrt{s} = 7\,\text{TeV}$. It was shown that differential cross sections can be modeled over six orders of magnitude (from nb to fb)[1] at energies that extend into the TeV range. The best models at the time were fixed-order predictions at next-to-leading order in the strong coupling constant; α_s. These predictions however had difficulty modeling certain regions such as large globally summed transverse momentum (H_T) [16].

With the increased center of mass energy, and increased integrated luminosity of the 2012 data, this analysis can probe even greater energies than those in the past, and with finer binning. This new data can be compared with improved and higher order theoretical predictions. Moreover, observables such as the transverse momentum (p_T) of the W boson, which have not been published since 2010, are shown in this analysis with much further kinematic reach [17].

This analysis also studies the ratio of the W^+ to W^- cross sections. In the ratio, many experimental and theoretical uncertainties cancel out, allowing for high precision tests of the theoretical predictions. Also, the ratio is especially sensitive to the parton distribution functions of the valence quarks, and so these high precision measurements are useful in studying parton distribution functions (PDFs).

This dissertation is organized as follows:

1. **Introduction**

2. **The Standard Model** - This chapter provides the reader with a brief theoretical

[1] nanobarns and femtobarns are units of area used to measure cross section.

background of the Standard Model, necessary for understanding the analysis and its motivation.

3. **The LHC and ATLAS** - An overview of the LHC is given and the ATLAS detector and its components are described in detail.

4. **Object Reconstruction** - A description of how the particles measured in the detector are reconstructed into their parent particles is given.

5. **Data and backgrounds** - This chapter discusses how events are selected from the data as well which backgrounds are present, and how they are estimated.

6. **Detector Level Results** - This chapter shows the results of event selection and the estimated backgrounds at the detector level.

7. **Unfolding** - The technique for determining the experimental cross section unmodified by detector effects.

8. **Systematic Uncertainties** - A description of all the systematic uncertainties is provided and their estimates are shown here.

9. **Theoretical predictions** - A description of the theoretical predictions that are compared to the data is provided here.

10. **Cross Section Results** - The final comparisons of the theoretical predictions to data are shown and described here.

11. **Conclusion**

Chapter 2

The Standard Model

The Standard Model, first formulated in the 1960's and later finalized in the 1970's, provides today's current understanding of 3 out of the 4 fundamental forces of nature, i.e. the electromagnetic force, the weak force, and the strong force. The gravitational force is not described by the Standard Model. A representation of the particles in the Standard Model is shown in Figure 2.1. The Standard Model provides a road map of the elementary particles and how they interact, separating the particles into fermions which comprise matter, and bosons, the force carriers that govern interactions. The analysis in this thesis focuses on collisions producing a W boson and jets. Therefore, a particular background knowledge of the W boson, and strong interactions containing quarks and gluons, is necessary for understanding these events.

2.1 Elementary particles

First the particles in the Standard Model will be summarized. Fermions are the particles that comprise matter; they are divided up into leptons and quarks.

Leptons are comprised of three generations or flavors: electrons, muons, and taus which have -1 elementary charge ($1.602 \times 10^{-19}\,\mathrm{C}$), and each having a corresponding neutrino which is neutral. Having electric charge means that the electron-like leptons can interact electromagnetically through photons. All leptons also have weak isospin allowing them to

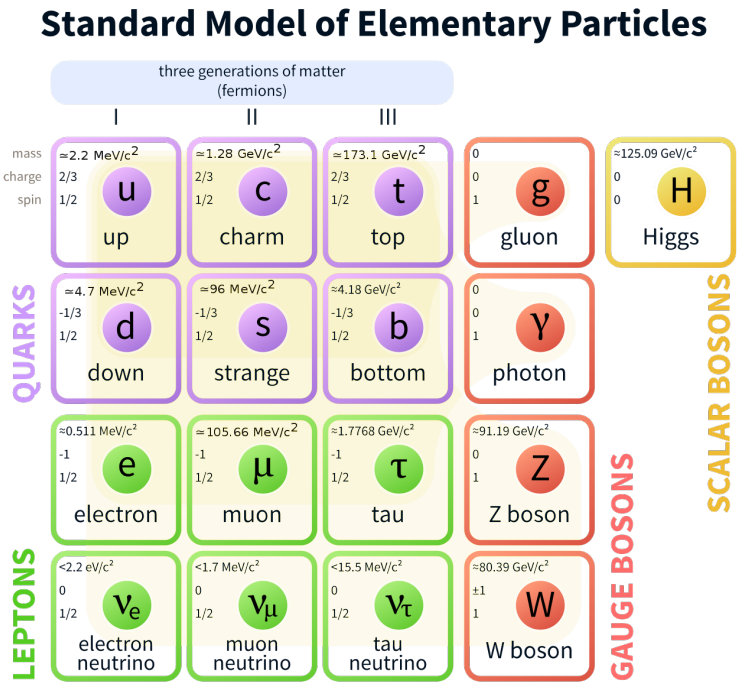


Figure 2.1: The elementary particles of the Standard Model are separated into quarks, leptons, gauge bosons, and the scalar Higgs boson. [3]

interact through the weak force (with W and Z bosons). Having no color charge, leptons have no coupling to gluons, and having relatively low mass, leptons interact only weakly with the Higgs boson. In high-energy environments like the LHC, the mass energy is negligible compared to kinetic energy and so usually these particles are treated as massless.

Similarly, there are three generations of **quarks**. The first generation contains the up-quark and the down-quark, the second, the charm-quark and the strange-quark, and the third, the top-quark and the bottom-quark. Often quarks are referred to by just their first letter, e.g. the bottom quark is often called a b-quark. Quarks and gluons have color charge, the conserved quantity relating to strong interactions. There are three color charges: red, blue, and green, that form a colorless bound state when combined (e.g. a proton is composed of a red, a blue, and a green quark). Although quarks do have electric charge and isospin, the electroweak coupling constant α is small in comparison to the strong coupling constant α_s.

This means that the quarks interact much more strongly with gluons.

The up-, down-, and strange-quarks, have relatively low mass such that it can be for the most part ignored while doing collider physics. However the charm-, bottom-, and top-quarks (sometimes referred to as heavy flavor quarks), are massive enough that the mass terms must be considered. Their mass can sometimes be used to the advantage of the experimentalist. b-quarks for example, because of their large mass, have decay products with relatively large momentum transverse to their original trajectory. The wider spread in its decay products, as well as its longer lifetime, can be used to help identify the b-quark itself. This is called b-tagging and is discussed later in Section 4.4.6

It should also be mentioned that for every particle in the Standard Model, there exists an anti-particle. Anti-particles have all the same attributes as their corresponding particles except that all of their conserved charges are reversed. For example, an up-quark has $+2/3$ of the elementary electric charge, and an up-antiquark has $-2/3$ of the elementary electric charge. If the same quark had red color charge, its corresponding anti-particle would have antired color charge.

As for bosons, there are photons, W and Z bosons, gluons, and Higgs bosons.

Photons have no charge themselves but interact with any particles containing electric charge including electrons, muons, taus, W bosons, and all quarks. Interactions involving the photon are explained using quantum electrodynamics (QED).

Gluons are the massless gauge bosons of the strong force. The mechanics of how quarks and gluons interact is referred to as quantum chromodynamics (QCD). Gluons always carry a color charge and an anti-color charge, and serve as the mediator of the force binding quarks within the proton. Because of the running of the α_s coupling, which is proportional to the strength of strong interactions, QCD has some unique properties known as confinement and

asymptotic freedom which are discussed more below.

***W* bosons** and ***Z* bosons** are the gauge bosons for weak interactions. While Z bosons are charge neutral, W bosons have electric charge ± 1. Both bosons are massive with the W around 80 GeV and the Z being 91 GeV. The W boson is of course the main focus of this analysis. 67.6 % of W bosons decay hadronically, i.e. into quarks and gluons. The rest decay leptonically i.e. into lepton pairs such as an electron and a neutrino. The electron neutrino branching ratio is 10.75 % [18] which is the main channel studied in this analysis.

The **Higgs** boson is responsible for giving the other particles mass. Any particle that has mass, therefore interacts with the Higgs boson. The Higgs is self-interacting and has a relatively large mass of 125 GeV. Unlike the other gauge bosons in the Standard Model which are all spin 1, the Higgs is scalar meaning it has 0 spin and therefore no possible polarizations.

The weak bosons, along with the Higgs boson are closely tied to the photon. Mathematically these forces derive from a single symmetry called electroweak symmetry. The Higgs boson breaks this symmetry through spontaneous electroweak symmetry breaking, from which the masses of all the Standard Model bosons, excluding gluons, are derived. Additionally, electroweak symmetry breaking gives rise to the property that the vacuum expectation value for the Higgs field is non-zero. More can be read about this subject in Reference [19].

2.2 Proton-proton collisions

In a deep-inelastic collision at high energies, it is not actually the protons which are scattering, but the constituent quarks and gluons. Usually the constituents of the proton are considered to be two up-quarks, a down-quark, and the gluons binding them together; how-

ever, in high energy collisions, the sea quarks must also be considered. Sea quarks are the virtual particles in the proton that arise from gluons splitting into quark-antiquark pairs. These quarks are short lived, usually annihilating back into gluons. There is a constant flux of splitting and annihilation such that there is a "sea" of quarks. The fraction of the proton's momentum that each flavor of quark holds is described by a parton distribution function (PDF).

A parton distribution function is the probability density of a quark or gluon $f(x, Q)$ to have a momentum fraction x, given an energy scale Q. Figure 2.2 shows an example of this, displaying the CT14 PDF set developed by the CTEQ collaboration. As a convention, the probability density is usually scaled by the momentum fraction when plotted ($xf(x, Q)$). This plot shows the PDFs for gluons, the up-, and down-quarks, and the up-, down-, and strange-antiquarks. As one might expect, the up-quark has the largest likelihood to carry most of the proton's momentum, since there are two valence up-quarks. Most of the remaining momentum is likely to be carried by the down-quark and gluons. Of the sea quarks, the up-antiquark has the next largest PDF since it is less massive, and requires less energy to create compared to the other sea quarks, followed by the down-, and strange-antiquarks for the same reasoning. Although the PDFs of the sea quarks are smaller, their contribution is non-negligible.

PDFs are fitted using data from deep inelastic scattering measurements from a variety of fixed-target experiments at research centers such as SLAC, CERN, FNAL, and DESY, as well as from collider measurements of Drell-Yan and jet production. There are a variety of errors associated with these measurements both statistical and systematic from sources such as energy calibrations. Regardless of how the fit is performed, the errors are reduced to a set of asymmetric eigenvectors which can be propagated through predictions and added using

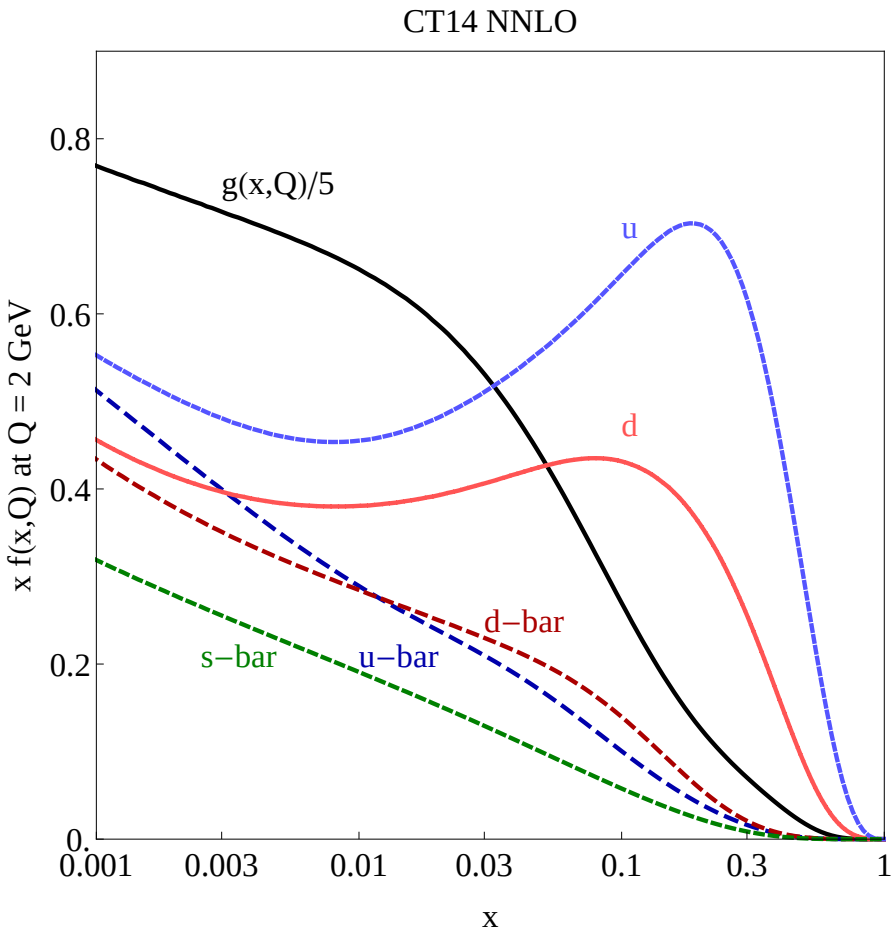


Figure 2.2: The central PDF $f(x, Q)$ for the gluons, and the up-, down-quarks, and up-, down-, and strange-antiquarks given by the CT14 PDF set as a function of the momentum fraction x. The PDF is by convention scaled by the momentum fraction [4]

.

the "master equation" to find the total PDF uncertainty in both directions. The master equation considers the up and down errors separately. It defines the up error $\Delta X^{+}_{\max}$ and the down error $\Delta X^{-}_{\max}$ to be:

$$\Delta X^{+}_{\max} = \sqrt{\sum_{i=1}^{N} \left[max\left(X_i^{+} - X_0, X_i^{-} - X_0, 0\right)\right]^2}, \tag{2.1}$$

$$\Delta X^{-}_{\max} = \sqrt{\sum_{i=1}^{N} \left[max\left(X_0 - X_i^{+}, X_0 - X_i^{-}, 0\right)\right]^2}, \tag{2.2}$$

Where i is an index over eigenvectors, X_0 is the central value, X^{+} is the positive error, and X^{-} is the negative error [20]. This PDF uncertainty is one of the two major errors in theoretical Monte Carlo predictions. These predictions are discussed later in Chapter 9

2.3 Color confinement and hadronization

The coupling constant for QED is relatively small at low energies, $\alpha = 1/137$, and increases slowly at high energies. In contrast, the strong coupling constant for QCD is small at high energies and increases at low energy. This behavior leads to two properties unique to QCD: asymptotic freedom at high energies, and confinement at low energies.

Asymptotic freedom means that the quarks and gluons behave as free particles in a high energy environment. In this region, calculations can be done perturbatively. Confinement prevents colored particles from ever being observed individually. For example, if a quark-antiquark pair is pulled apart, the energy from separation allows for the production of an additional quark-antiquark pair. As the colored objects are pulled farther apart, more and more of these colored pairs are produced. Eventually the colored objects combine to form

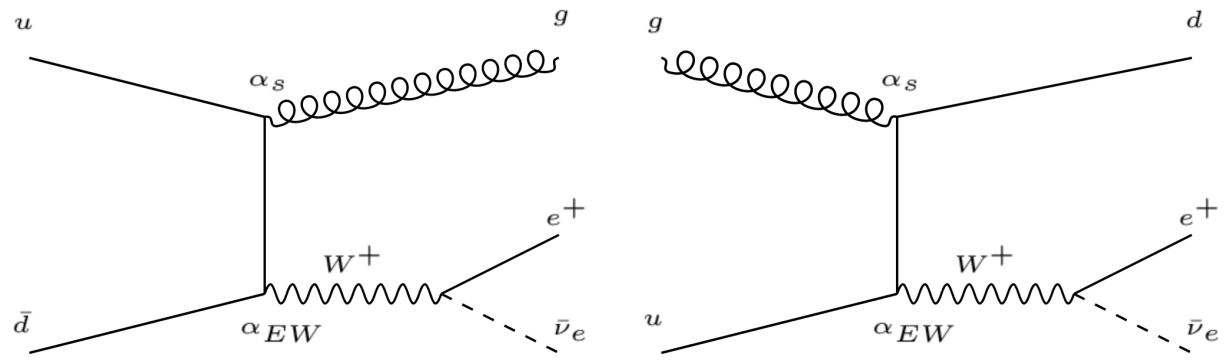
colorless particles. This process is called hadronization. In general, the colorless particles formed are mesons (quark-antiquark pairs which form a colorless bound state), and baryons (bound states made up of three quarks of different colors such that they all cancel). The umbrella term for a colorless bound state made up of quarks is a hadron.

In a collider, when a quark or gluon is produced through a hard proton-proton interaction, the quark or gluon is not directly observed. Instead, the quark or gluon hadronizes into a spray of collimated particles called a jet.

2.4 W + jets production

A W + jets event is any collision that produces a W boson and some number of jets. In this analysis, only events where the W boson decays into an electron and neutrino were selected. A few Feynman diagrams are shown in Figures 2.3 and 2.4. Figure 2.3 shows two possible diagrams of a W boson +1 jet event. In the first example, an up-quark and a down-antiquark collide and produce a W boson and a gluon. The W boson decays leptonically into an electron and a neutrino, and the gluon will hadronize into a jet. In the second example, an up quark interacts with a gluon in order to produce a down-quark and a W boson which again decays into an electron and a neutrino. It should be noted that the quark-quark-gluon interactions are strong interactions with α_s vertices, and the quark-quark-W boson interactions are weak interactions with α_{EW} vertices.

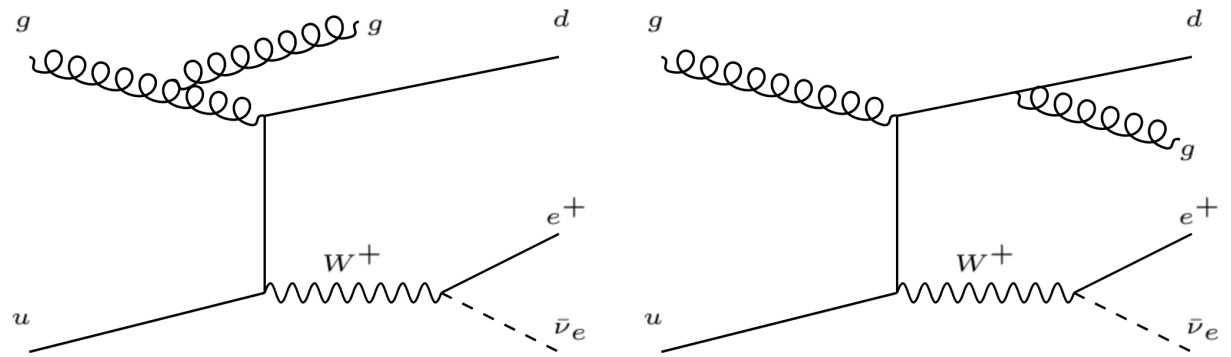
Figure 2.4 shows two examples of how $W + 2$ jet final states are formed. Both are actually the same process as Figure 2.3b, but with an additional gluon being radiated. Figure 2.4a shows the additional gluon being radiated from the initial state gluon, and Figure 2.4b shows it being radiated from the final state down-quark. For each $W + 1$ jet diagram, there



(a) (b)

Figure 2.3: Figure 2.3a on the left shows an up-quark and a down-antiquark colliding and producing a gluon and a W boson, which then decays into an electron and electron-neutrino. Figure 2.3b on the right shows an up-quark and a gluon producing a down-quark and a W boson which decays likewise.

are several more diagrams for $W + 2$ jets events. Therefore, while there are relatively few diagrams for $W + 1$ events, the number of possible diagrams grows quickly with the number of jets calculated. This makes calculations for higher jet multiplicities and higher order calculations more challenging.



(a) (b)

Figure 2.4: Figure 2.4a on the left shows the same diagrams as Figure 2.3b with the initial gluon radiating another gluon so that there are 2 jets in the final state. Figure 2.4b shows the same, except that it is the final state down quark that radiates the gluon.

Chapter 3

The LHC and ATLAS

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest and highest energy particle accelerator in the world. The circumference of the accelerator is 27 km, crossing the border between Switzerland and France. It was designed to collide protons at a center of mass energy of 14 TeV, and built for both precision measurements and discoveries. Two beams run opposite each other in separate adjacent beam pipes 100 m underground. The beams collide at four interaction points where there is a major detector constantly observing the collisions. Figure 3.1 shows the detectors ATLAS, ALICE, CMS, and LHCb around the ring.

There are over 1,200 superconducting dipole magnets, each producing a field strength of 8.4 Tesla, which are used to keep the beams going in a circle. Each of the two beam pipes is equipped with its own set of magnets. There are 392 quadrupole magnets, which are used to focus the beam horizontally and vertically. This is done by grouping the quadrupole magnets in pairs. Two quadruple magnets in close proximity have a focusing effect on charged particle beams. This is necessary since the beam pipes are quite narrow (6.3 cm in diameter). In order to accelerate the beams to their final energies, eight radio-frequency (RF) cavities are utilized in each beam pipe. The RF cavities serve not only to accelerate the protons, but also to focus the beam longitudinally. The beam is not a constant stream of protons but

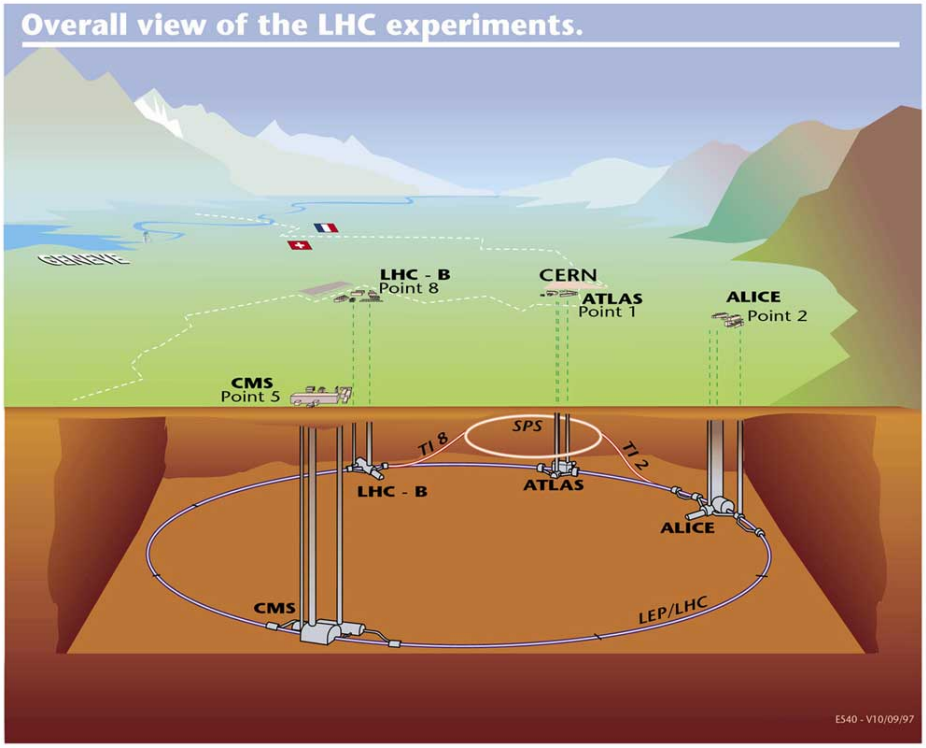


Figure 3.1: This graphic shows the LHC and the four major detectors around the accelerator: ATLAS, CMS, ALICE, and LHC-B. The actual detectors are underground, with access shafts leading up to their respective control centers above ground. The CERN campus, located above ground near the ATLAS control center, is also indicated [5].

instead it is composed of "bunches" of protons (many protons traveling together in a small time window). The bunch structure is required for RF cavities to accelerate the beams properly. Having the protons in bunches makes it more likely to have many collisions per crossing, and is useful in timing when the collisions occur [21]. In 2012, when the data for this analysis were collected, the bunches were 50 ns apart. A smaller bunch spacing allows for more bunches and thus more protons. The drawback is that less spacing between bunches can lead to more out-of-time pile-up. This is when the energy deposited in the detector from one bunch crossing overlaps with the energy from previous bunch crossings, potentially confusing the energy reconstruction. This is discussed further in Section 3.1.1.

The amount of data taken by the LHC is often quantified by the integrated luminosity L defined in equation 3.1, where DN/dt is the rate of events as a function of time t, and σ is the total proton-proton interaction cross-section.

$$L = \int \frac{1}{\sigma} \frac{dN}{dt} dt \tag{3.1}$$

Figure 3.2 shows the integrated luminosity for the data taken in 2012. The total integrated luminosity delivered by the LHC in 2012 was $22.8\,\text{fb}^{-1}$.

3.1.1 Pile-up

As mentioned earlier, a tighter bunch spacing is excellent for improving statistics, but comes at the cost of creating more out-of-time pile-up. Out-of-time pile-up occurs because it takes time for the energy in the detector to be measured. If two back-to-back bunch crossings deposit energy in the same part of the detector, the energy from one interaction may overlap with the energy from the previous one. This needs to be resolved in order to correctly

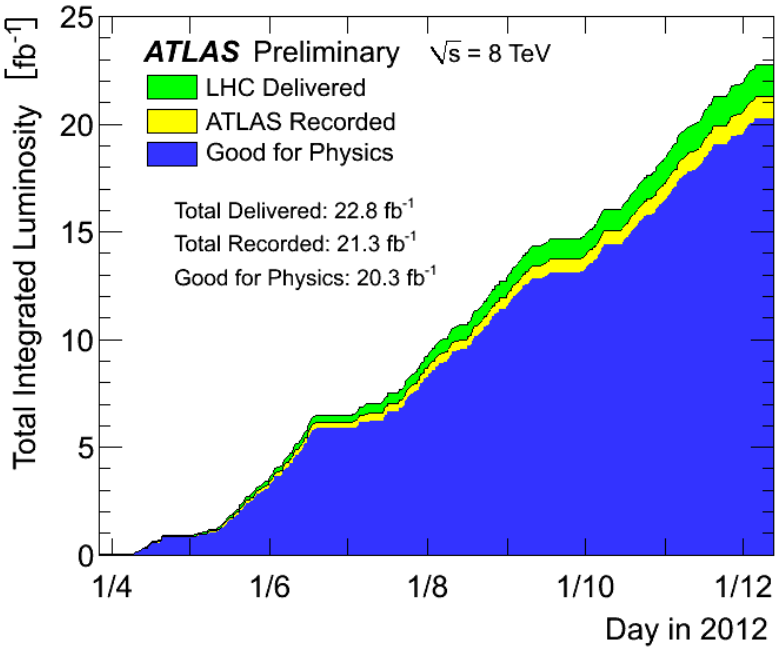


Figure 3.2: The total integrated luminosity delivered by the LHC, recorded by the ATLAS detector, and approved for physics in 2012 [6]

understand the kinematics of the particles from both interactions.

There is also another type of pile-up called in-time pile-up. In-time pile-up happens because the bunches themselves each contain on average 1.15×10^{11} protons and usually there are many collisions per bunch crossing. In 2012, when the data for this analysis were collected, there were on average 20.7 interactions per bunch crossing. These different interactions must be separated out in order to understand the underlying physics. This is achieved by taking very precise measurements of the trajectories of the particles and tracing them back to their interaction points. The interaction point for an individual proton-proton collision is called the primary vertex. Although, many collisions happen at the same time, there is a spatial separation between the primary vertices that can be used to differentiate between collisions and help mitigate this in-time pile-up.

3.2 The ATLAS detector

The ATLAS detector (**A** **T**oroidal **L**HC **A**pparatu**S**) is an all purpose detector located at interaction point 1 of the LHC at CERN. In order to utilize the full discovery potential of the LHC, ATLAS needs to be able to identify all the particles of the Standard Model. To do this, ATLAS is composed of four subdetectors: the inner detector, the liquid-argon electromagnetic calorimeter, the hadronic calorimeter, and the muon spectrometer. The subdetectors form concentric cylinders providing hermetic coverage of the collision, collecting information about the energy, momenta, and trajectory of the particles. The 7,000 tonne machine is the largest volume detector ever built for a particle collider. Many particles, such as muons, do not interact strongly and pass through most matter. Much of the volume is taken up by the muon detector. Because some of the interesting physics signatures produce very high momentum muons, a large magnetic volume is needed to produce a measurable curvature of the muon track. To accommodate this, ATLAS was built 25 m tall and 45 m long and rivals the weight of the Eiffel Tower. In order to get it down into the man-made cavern 100 m below ground, it was lowered down piece by piece through a service shaft seen in Figure 3.3 [5].

With the cylindrical shape of the detector, it is natural to use cylindrical coordinates to describe the geometry of the detector. Also, when describing collisions, it is often useful to use pseudorapidity (defined below in Equation 3.3), which is an approximation of rapidity. Rapidity is generally preferred because it is a more natural observable and behaves nicely under Lorentz transformations. Rapidity is a measure of relativistic velocity given by a

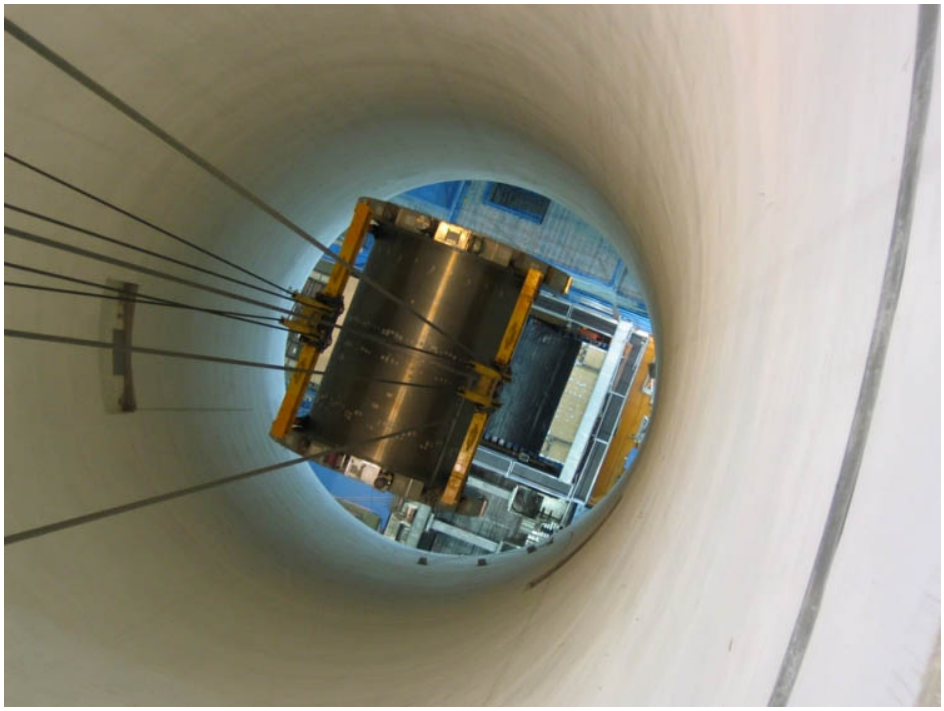


Figure 3.3: A picture of the liquid argon barrel being lowered down the service shaft for ATLAS [5]

particle's energy E and momentum along the beam access p_z:

$$y = \frac{1}{2} \ln \frac{E + p_z c}{E - p_z c} \tag{3.2}$$

Rapidity is dependent on mass which varies between particles, while pseudorapidity does not depend on mass and is easier to measure in the lab frame. In the limit of very high energy, rapidity and pseudorapidity become the same, and so at the large energies of the LHC, pseudorapidity is often an excellent approximation. Because of this, experimentalists often use the Lorentz invariant rapidity, y, and pseudorapidity, η, interchangeably. It should still be noted that for heavier particles like the W boson, this approximation breaks down. Throughout this thesis the following right-handed coordinate system described below, and shown in Figure 3.4, will be used unless otherwise stated.

- z, the vector along the beam line from the center of the detector
- ϕ is the azimuthal coordinate
- R is the radial distance from the center of the detector
- θ, the polar angle subtended from the z-axis
- η is the pseudorapidity defined as

$$\eta = -ln\left[tan\left(\frac{\theta}{2}\right)\right] \tag{3.3}$$

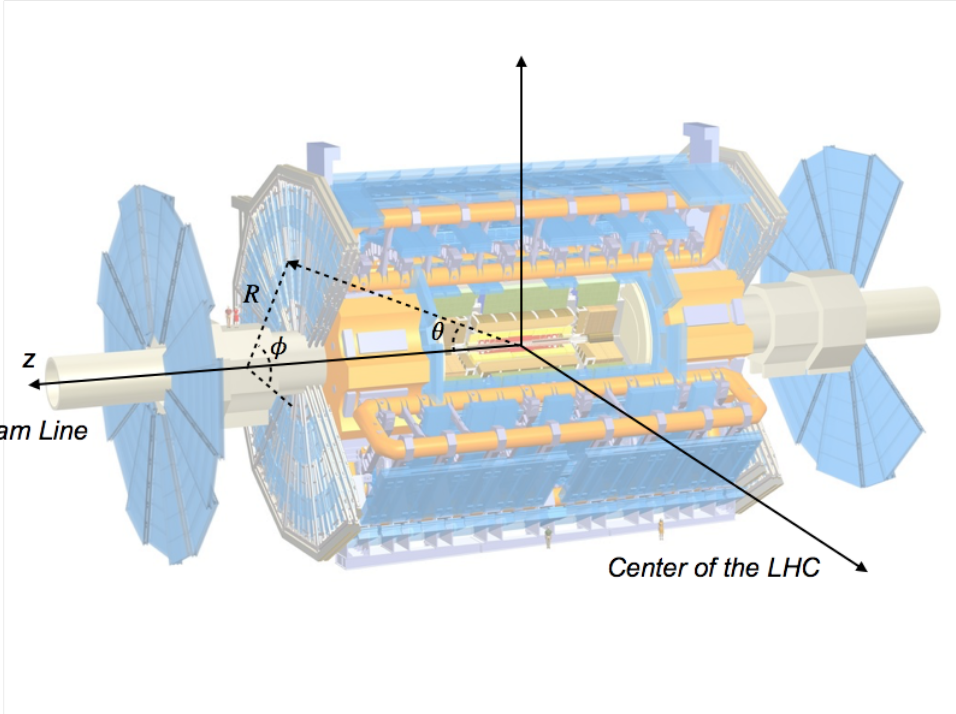
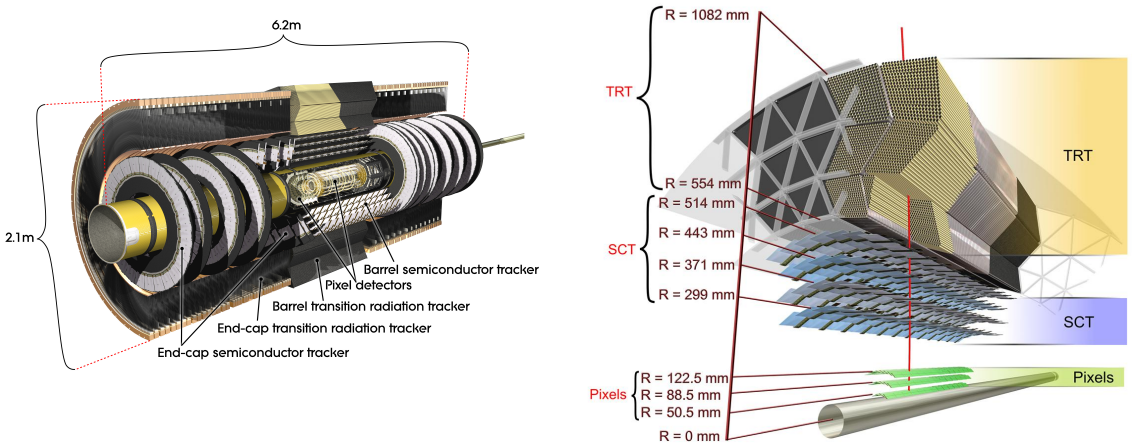


Figure 3.4: The figure above shows a cutaway view of the ATLAS detector with the right-handed coordinate system super-imposed over it. [5]

3.2.1 Inner Detector

The inner detector (ID), shown in Figure 3.5, lies closest to the interaction point. It is composed of three submodules: the pixel detector (pixel), the semiconductor tracker (SCT), and the transition radiation tracker (TRT). The main purpose of the ID is to provide detailed tracking information for the particles, not only to determine their trajectory and momenta, but also to reconstruct the primary and secondary vertices of collisions. Tagging primary vertices is useful for helping to mitigate pile-up (see Section 3.1.1), and identifying secondary vertices is useful for b-tagging (see Section 4.4.6). The entire detector is encased in a solenoid with the 2 Tesla field orientated along the beam line. Any charged particles originating from collisions will be subject to a Lorentz force, and the curvature can be used to determine the charge and momentum of the particle.



(a) (b)

Figure 3.5: Figure 3.5a shows a cutaway view of the inner detector showing the pixel, SCT, and TRT sub-detectors. Figure 3.5b shows a graphical representation of the head on view where the layers of each sub-detector are clearly visible [5].

3.2.1.1 Pixel Detector

The pixel detector is so-named because it is comprised of 1744 silicon pixel detectors (shown in figure 3.6) that provide 80.4 million readout channels. The pixel detector provides the finest granularity in the detector with each pixel being 50 µm in the ϕ direction and 400 µm in the z direction.

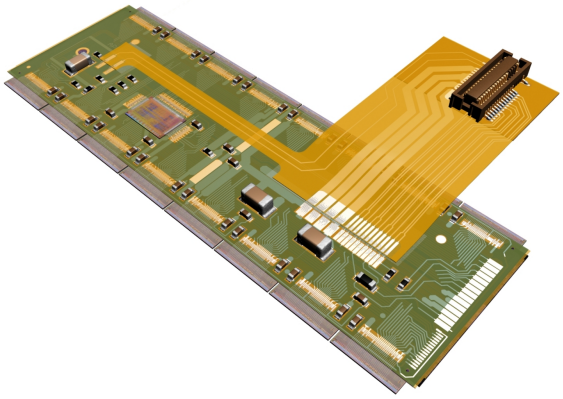


Figure 3.6: A rendered close up image of a pixel module. The silicon sensor is between the electronics containing over 46,000 pixels [7].

The pixels detect particles by measuring the ionizing energy deposited in them. As a particle passes through a pixel it leaves behind electron/hole pairs in the silicon. When an electric field is applied, a current is induced which can then be amplified and read out as a hit in the detector. There are three concentric layers of pixels at $R = 50.5$, 88.5, and 122.5 mm that cover a total z distance of 800 mm that make up the barrel region, and an additional three end caps on both sides of the detector at $z = 495$, 580, and 650 mm.

3.2.1.2 Semiconductor Tracker

The semiconductor tracker, like the pixel detector, uses silicon as the active material. Instead of pixels however, the SCT uses two layers of strips at the same radial distance in order to

measure all three spatial coordinates. One layer is parallel to the beam axis while the other is tilted by an angle of 40 mrad with respect to the beam axis. There are four coaxial sets of these double layers making up the barrel region, that covers the volume $255 < R < 549$ mm and $0 < z < 805$ mm, enabling up to 8 possible hits with a resolution of 17 µm in the R-ϕ plane and 580 µm along the z-axis. In the end-caps, there are again double layers, but with one set radial and the other angled by 40 mrad. There are 9 end-caps on either side that cover the volume from $275 < R < 560$ mm and $839 < |z| < 2735$ mm.

3.2.1.3 Transition Radiation Tracker

The TRT is the outermost module in the ID and is composed of drift tubes. The tubes, or straws, are filled with a gaseous mixture of Xe, CO_2, and O_2, and have an electrically grounded wire of tungsten placed axially in the center. The straws themselves are held at -1530 V so that there is a radial electric field within the tubes. When an ionizing particle passes through the tube, the resulting electrical charge, produced in the gas, travels down the wire and is read out as a signal current. The tubes in the barrel region, parallel to the beam line, do not provide any position measurement in the z-direction but have a resolution of 130 µm in R-ϕ plane. This is not as fine a resolution as the pixel or SCT, but the TRT has many more layers. There are 73 straw planes in the barrel region covering $563 < R < 1066$ mm and $0 < |z| < 712$ mm. Because the tubes are stacked hexagonally (see figure 3.7), a track that passes between straws in one plane will travel directly through the tubes in the next plane, so there will be at least 36 hits per track. Additionally, the forward end-cap region has 160 straw planes, where the straws are oriented radially, covering $644 < R < 1004$ mm and $848 < |z| < 2710$ mm. The many layers of the TRT combined with the precision of the pixel and SCT yield a robust pattern recognition for the track

reconstruction.

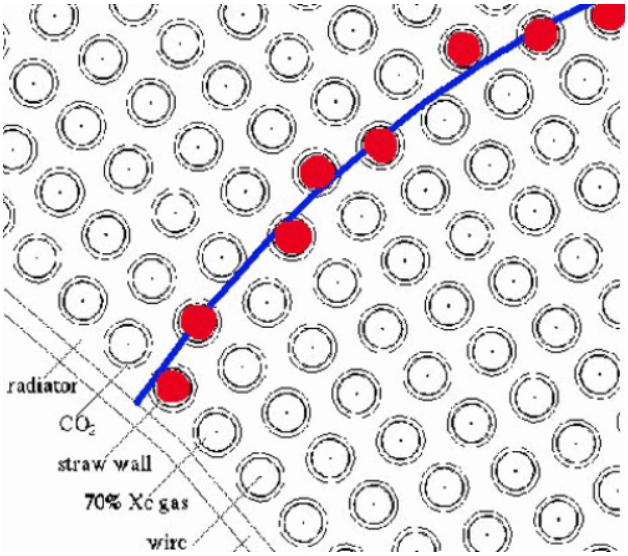


Figure 3.7: A sketch of the TRT with a particle traveling through [8].

The TRT has a second functionality in the identification of charged particles. When a charged high-energy particle crosses a boundary with differing dielectric constants, there can be transition radiation in the form of soft photons. The space between the tubes is filled with polypropylene fibers in the barrel region, and foils of the same material in the end-caps. Polypropylene has a different dielectric constant than the gaseous mixture in the tubes. The addition of transition radiation results in a much larger energy deposit in the tubes than just from the ionization. Lighter particles are much more likely to cause transition radiation. Typically, electrons are the only particles light enough at the LHC to create transition radiation and create many more hits with these large energy deposits than heavier charged particles.

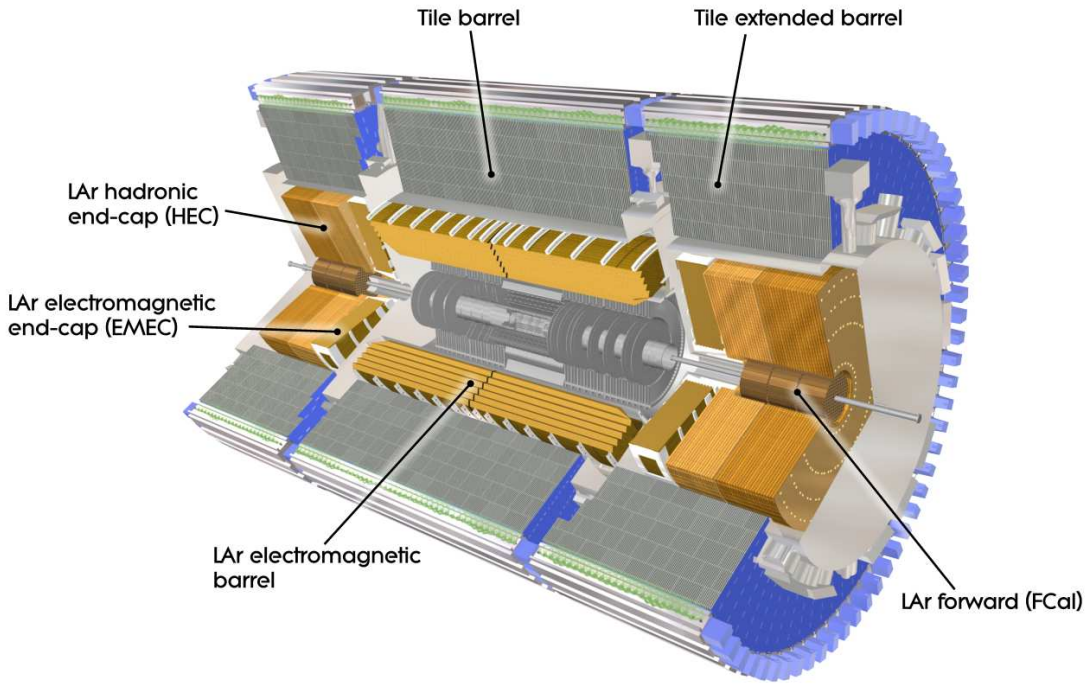


Figure 3.8: Depiction of the electromagnetic and hadronic calorimeters [9].

3.2.2 Electromagnetic Calorimeter

Outside the ID is the electromagnetic calorimeter (ECal). The ECal is a sampling calorimeter that uses liquid argon (LAr) as its active material to detect the energy deposited by electron and photon showers. The absorbing material used is lead, which is alternated with the LAr in an accordion style pattern shown in Figure 3.9. The unusual looking pattern was chosen to prevent any gaps in the ϕ coverage. When a photon or electron hits the lead it begins to shower. An electromagnetic shower is characterized by radiation length X_0, defined as the average distance a particle travels before it has radiated a factor of $1/e$ of its energy. Electrons and positrons will typically radiate photons, and photons will typically split into electron-positron pairs. When the energy of the photons falls below the threshold for pair production, the photons continue to deposit energy through processes such as Compton scattering. A high voltage is applied across the LAr, so the electron-ion pairs produce a

signal that is read out as a current. The current per daughter particle is the same, so the total current is proportional to the number of particles in a given region. Combining all the regions of a shower make it possible to infer the energy of the particle.

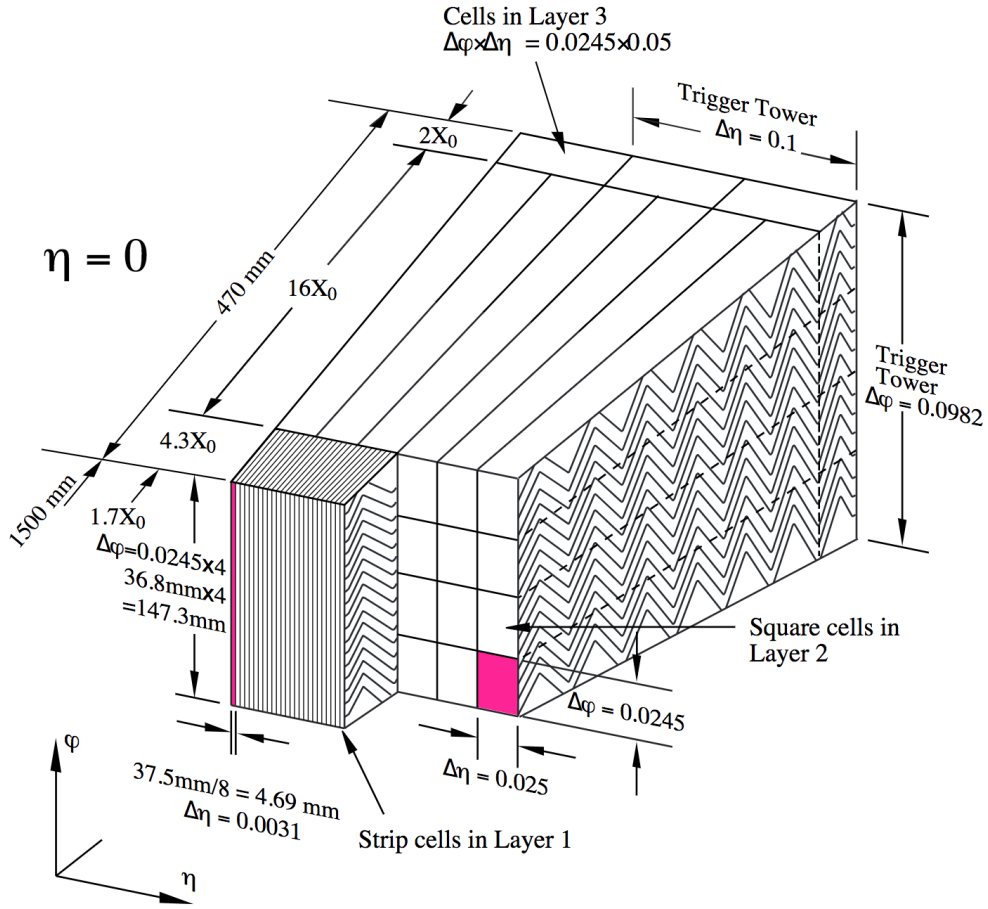


Figure 3.9: A sketch of the accordion structure of the lead-LAr layers in the barrel of the ECal [9].

The calorimeter has two main parts: the barrel covering $|\eta| < 1.475$ and the end-caps ($1.375 < |\eta| < 3.2$). The end-caps (EMEC) have a concentric outer wheel ($1.375 < |\eta| < 2.5$) and inner wheel ($2.5 < |\eta| < 3.2$). There is a presampler that covers the region $|\eta| < 1.8$, which is used to correct for the energy lost in the inner ATLAS detector. The barrel is made up of three radial layers with coarser granularity as R increases. Layer 1, having the finest coverage in η, where $\Delta\eta = 0.003$, also has the coarsest ϕ resolution $\Delta\phi = 0.1$. This

is compensated for in layer 2 and 3, however, where the ϕ resolution is 0.025. Layer 1 is $4.3X_0$ in the radial direction, while the second layer is $16X_0$ radially; most of the energy from an electromagnetic shower is deposited in layer 2. Layer 2 also has an η resolution of $\Delta\eta = 0.025$. Finally, the third and last layer has the largest granularity $\Delta\eta = 0.05$, and has the shortest radial dimension ($2X_0$). Generally, electromagnetic showers that make it to the third layer are wide enough where this is sufficient resolution. The outer wheel of the end-caps also has three layers to help extend the resolution of the barrel and the inner wheel also has two layers [10]. In the region of precision physics ($|\eta| < 2.5$), where the ECal has 3 layers, there is enough angular resolution to help distinguish between a photon (γ) and a neutral pion (π^0) π^0 almost always decays into a photon pair which can be distinguished from a single photon up until high p_{T} where the photons eventually become too collinear to distinguish.

3.2.3 Hadronic Calorimeter

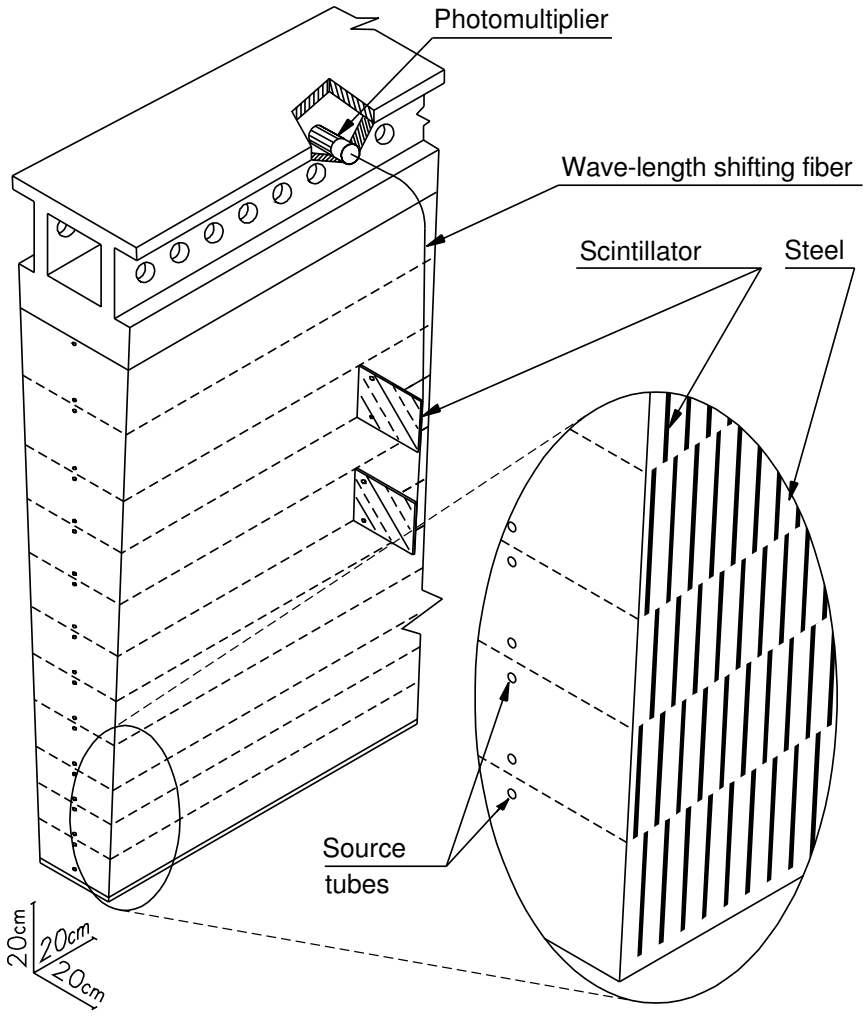
The hadronic calorimeter lies beyond the LAr detector as shown in Figure 3.8. The hadronic calorimeter is meant to measure the energy deposited from jets that was not contained in the ECal. It is a sampling calorimeter with the active material being scintillating tiles in a steel matrix, which act as the absorbing material. Thus, the calorimeter is often called the Tile Calorimeter or TileCal. Light travels out two sides of the scintillators through wavelength shifting fibers where it is read out by photomultiplier tubes (PMTS). A single section of the TileCal is made up of 64 modules in the ϕ-direction. Figure 3.10a shows one of these modules and the arrangement of the steel, tiles, fibers, and PMTS. The source tube holes shown are for calibration from the transit of a Cesium source. A set of 64 modules make up the barrel region. There are also 64 modules that make up either extended barrel totaling

192 modules in the TileCal.

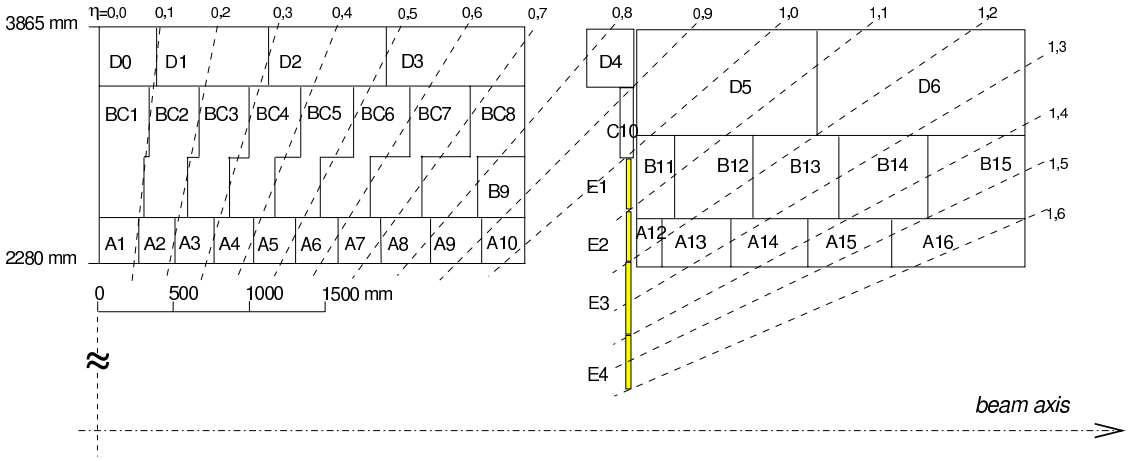
The modules are divided up into cells as shown in Figure 3.10b. There are three radial layers of cells, A, BC (just B in the extended barrel) and D, and there is one cell in the A and BC layer per 0.1 η. The D cells tend to vary in length, as do the cells in the gap region. The gap region is the space between the barrel and the extended barrel where there are electronics for the inner detector and other services. The barrel covers an η range of $|\eta| < 1.0$ and the extended barrel covers $0.8 < |\eta| < 1.7$. Similar to how electromagnetic showers are characterized by radiation length X_0, hadronic showers are characterized by interaction length λ_{int}. Interaction length is defined as the mean free path of a hadron before it interacts with atomic nuclei. It is typically much longer than radiation length. Layers A, B, and D are 1.5, 4.1, and 1.8 λ_{int} respectively in the central barrel, and 1.5, 2.6, and 3.3 λ_{int} in the extended barrel.

Inside the extended barrel, beyond the ECal, lies the LAr hadronic end-cap calorimeter (HEC). The HEC is also a sampling calorimeter using LAr and shares the same cryostat as the EMEC, but instead of lead it uses copper as the absorption material. The flat plate design is shown in Figure 3.11. The calorimeter is made of two wheels: the front wheel HEC1 containing 24 copper plates of 25 mm thickness, and the rear wheel HEC2 containing 16 plates 50 mm thick. The thinner plates in HEC1 provide a finer sampling, yielding better resolution for the front end of the showers. The HEC covers an η range $1.5 < |\eta| < 3.2$, overlapping with the extended barrel of the TileCal.

Also sharing the same cryostat as the HEC and the EMEC, the forward calorimeter (FCal) also uses LAr as the active material. Composed of three wheels, the first wheel (FCal1) uses copper as the absorption material, and the second two (FCal2 and FCal3) use tungsten. In total, the three wheels cover the very forward range $3.1 < |\eta| < 4.9$. FCal1 is designed to



(a)



(b)

Figure 3.10: Figure 3.10a shows a scale representation of a TileCal module showing the tiles and readout. Figure 3.10b shows TileCal cells in one half of the barrel and extended barrel also to scale [9].

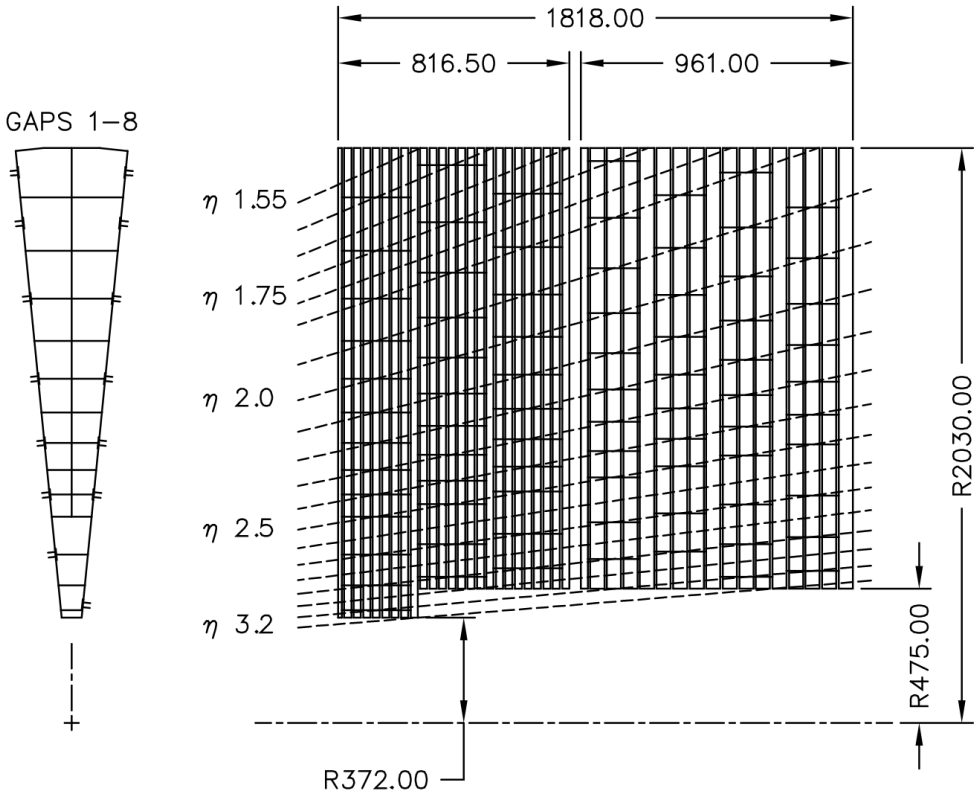
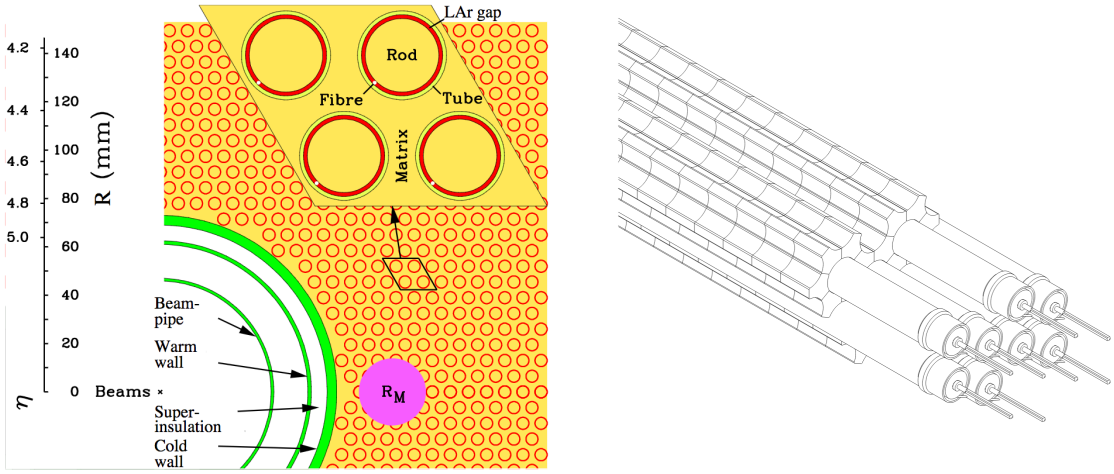


Figure 3.11: A sketch of the flat-plate design of the LAr hadronic end-cap [10]

pick up electromagnetic showers and is made of copper plates stacked perpendicular to the beam line with holes drilled through the plates for the electrodes (see Figure 3.12a). The electrodes consist of a coaxial copper rod inside a copper tube with a fiber wrapped around the rod to separate the two and LAr fills the gap. FCal2 and FCal3 are very similar except that the rods are tungsten and instead of the electrodes being embedded in copper plates they suspended in an array of tungsten slugs shown in Figure 3.12b.



(a) (b)

Figure 3.12: 3.12a shows a cross section of FCal1 with the electrodes a copper matrix. 3.12b shows the structure of FCal2 and FCal3 in a tungsten matrix [10].

3.2.4 Muon Spectrometer

Muons, because they decay weakly, and have longer decay times, tend to pass through the ID and calorimeters without depositing much energy. In order to detect and measure muons, the muon spectrometer comes into play. The muon spectrometer is the outermost and largest part of the detector, and is encased in a system of three toroidal magnetic fields; muons exiting the detector pass through the field and curve due to the Lorentz force. Measuring

the bend in the trajectory provides the momentum of the muon. In the barrel region of the muon system, there are three cylindrical layers concentric around the beam axis. The end-caps also have three chambers, but perpendicular to the beam axis.

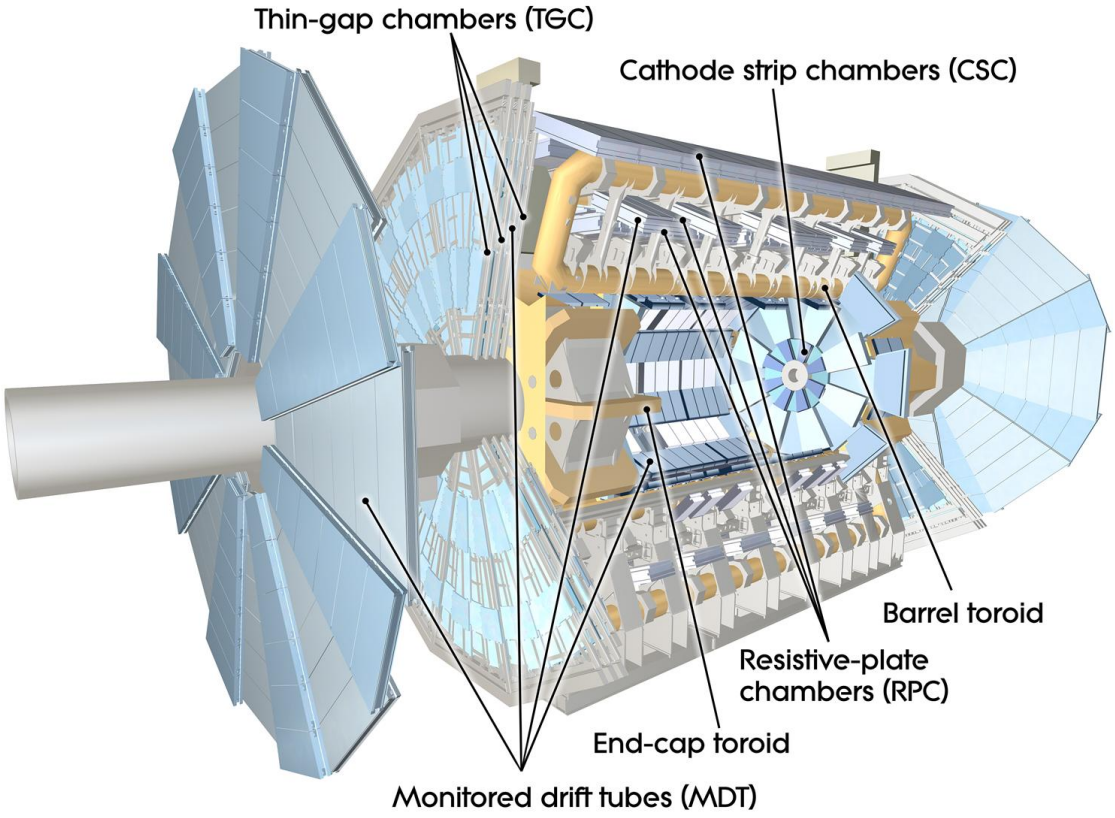


Figure 3.13: A cut-away view of the muon system [5]

There are four main components of the muon system (Figure 3.13). Two of the components are used for precision tracking information: the monitored drift tubes (MDT), which cover $|\eta| < 2.7$, provide a z-axis resolution of 35 µm, and the cathode strip chambers (CSC), covering $2.0 < |\eta| < 2.7$ provide an $R \times \phi$ resolution of 40 µm $\times$ 5 mm. The other two components are used for triggering. In the barrel ($|\eta| < 1.05$), resistive plate chambers (RPC) are used and in the end-cap ($1.05 < |\eta| < 2.7$), thin gap chambers (TGC) are used. The RPC's have a resolution of 10 mm and 10 mm in z and ϕ respectively, and the TGC

of 2-6 mm and 3-7 mm in R and ϕ. The trigger components provide not only additional tracking information in the direction orthogonal to that of the precision tracking, but also time resolution used for bunch-crossing identification.

Chapter 4

Object Reconstruction

4.1 Tracks

When a charged particle passes through the Inner Detector, it produces electron/hole pairs in the silicon of the pixel and SCT detectors, and ionizing radiation in the TRT. When it does this, the detector reads out a "hit" at the point the particle passes through. It is possible to have up to 3, 4, and 36[1] hits from the pixel, SCT, and TRT respectively. The hits alone do not provide much information, but if they are fit to a track, the charge and momentum of the particle can be determined, as the tracking system lies within a solenoidal magnetic field of 2 Tesla. In order to do this we make use of the Kalman filter algorithm.

4.1.1 Kalman filter algorithm

The Kalman filter algorithm [22] starts at the innermost part of the detector by creating track seeds from hits in the pixel and SCT. A window oriented in the direction of the track seed is then used to find other potential hits deeper in the ID. Hits within this window are collected and potential tracks are formed using a simplified Kalman filtering algorithm. The Kalman filtering algorithm uses the following variables to determine the quality of a track:

- $\frac{\mathbf{q}}{\mathbf{p}}$ - the charge of the particle divided by its momentum

[1] It is possible for low momentum tracks to curl in which case more hits are possible from hitting the same layer twice or more.

- ϕ_0 - the angle with the x-axis in the $x-y$ plane at the perigee point
- θ_0 - the angle with the beam line in the $r-z$ plane at the perigee point
- $\mathbf{z_0}$ - the longitudinal impact parameter, the z coordinate of the track at the perigee point
- $\mathbf{d_0}$ - the signed transverse impact parameter, the closest distance to the beam line in the transverse plane

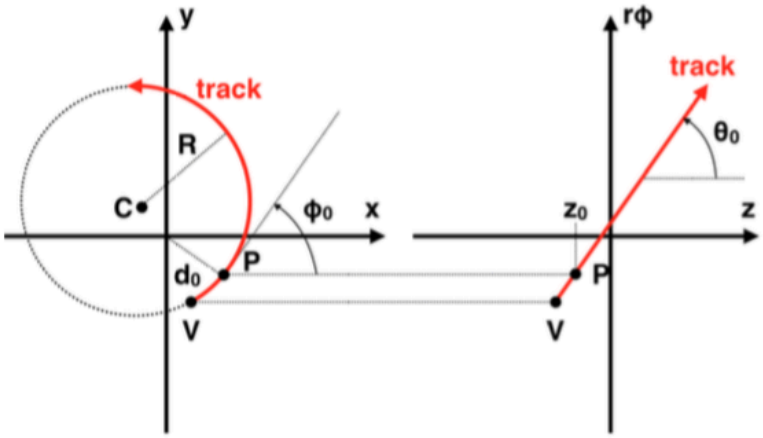


Figure 4.1: The track parameters for a charged particle track in the Inner Detector are shown. V indicates the vertex of the track and P is the point of closest approach (perigee). The transverse impact parameter d_0, and the azimuthal angle ϕ_0 are shown in the $x-y$ plane. The longitudinal impact parameter z_0, and the polar angle are shown in the $r-z$ plane [11].

The algorithm is iterative, at each layer, the track is re-evaluated and extrapolated to the next layer. The aim of the track fitting is to get precise measurements and track parameters for the inner and outermost points of the track.

Once tracks have been fitted, fake tracks that can result from insufficient hits or ambiguities must be vetted. To do this, tracks are scored based on whether they are fully reconstructed or remain as segments, and which detectors contribute to the track. The Pixel hits provide

higher scores since they are of higher precision than the SCT and TRT. The highest scoring tracks are kept and the lower scored tracks are discarded.

4.2 Electrons

Electron reconstruction in the central region of the detector starts by identifying energy clusters in the EM calorimeter, and matching them to tracks from the Inner Detector.

4.2.1 Electron cluster reconstruction

Electrons entering the EM calorimeter shower and deposit energy in groups of cells. The first step in identifying potential electrons is finding clusters of cells with energy deposited in the EM calorimeter. This is done using the sliding-window cluster algorithm.

The sliding-window algorithm starts by looking at the $\eta-\phi$ plane and dividing the calorimeter up into towers that can be summed longitudinally in the three layers of the ECal. The ECal cells form a grid in $\eta-\phi$ space that is 200×256 cells wide, with a resolution $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ (see Section 3.2.2). A 3×5 sliding-window is then moved around the grid and the transverse energy of the cells in the window are summed. As the window moves, it identifies local maxima above some threshold energy E_T^{thresh}, which are used to form preclusters. The position of the precluster is calculated as the energy weighted center of a smaller 3×3 window. The smaller window used for the position calculation is chosen to help reduce noise. Finally, to prevent the same cluster from being reconstructed twice, there is an isolation cut. If there are two preclusters with positions within a 2×2 window, the lower energy cluster is removed. The algorithm is complete once the entire 200×256 grid has been checked. The final clusters with a transverse energy above 2.5 GeV then move on

to track association.

4.2.2 Electron track reconstruction

Electron track reconstruction is performed in two parts: pattern recognition and track fitting. In 2012, an electron-specific pattern recognition and track fitting process was introduced, which takes into account the energy loss from bremsstrahlung and improves electron reconstruction. The process first assumes that the particle is a pion to estimate the energy loss in the material of the detector. If a track seed (three hits in the silicon detectors with transverse momentum greater than 1 GeV) cannot be extended to a full track (at least seven hits including the TRT), then pattern recognition is re-done assuming the particle to be an electron. The pattern recognition for the electron allows up to 30% energy loss at each material surface to account for bremsstrahlung.

The fitting is done with the ATLAS Global χ^2 Track Fitter [23] assuming the same particle from pattern recognition. If the fitting fails while assuming a pion, it is re-done assuming an electron. The track fit for the electron is the same as for the pion except that an extra term is used to account for bremsstrahlung.

Tracks from the ID are extrapolated to the middle layer of the ECal. These extrapolated tracks are compared to the positions of the clusters found using the sliding window algorithm. Tracks are considered loosely matched if one of the following requirements is met:

i The difference in η between the center of the EM cluster and the track must be within 0.05, and the difference in ϕ must be within either 0.2 on the side of the cluster the track is bent towards, or 0.05 in the opposite direction.

ii After rescaling the momentum of the track to the measured cluster energy, the track

must be within 0.05 in η, and the difference in ϕ must be within 0.1 on the side the track is bent towards, or 0.05 in the opposite direction.

Requirement *ii* allows for matching in cases where there is significant energy loss before the calorimeter. The bremsstrahlung is assumed to have occurred in the ID or the solenoid before the calorimeter, which will change its trajectory. Matched tracks at this point are re-evaluated using the Gaussian Sum Fitter (GSF) [24]. These tracks are then used to do the final track-cluster matching to build the electron candidates and provide information for particle identification.

4.2.3 Electron candidate reconstruction

The GSF refitted tracks are now required to match in ϕ within 0.1 (not 0.2). More than one track can be associated with a cluster. While all tracks are kept, the best matched track is chosen as the primary track to determine kinematics and charge, and to determine the electron identification decision. The primary track, in general, is the track closest to the barycenter of the cluster. In the case where there are two tracks whose distance to the barycenter is within 0.01 of each other in $\eta - \phi$, the track with more hits in the pixel detector is chosen. Hits in the first layer of the pixel detector are weighted twice as much, in order to prefer tracks with early hits. If the two tracks have the same number of hits in the pixel detector, the closer track is taken.

Finally, the energy of the cluster, and cells within the cluster, are calibrated using a multivariate analysis [25] trained on large samples of single-electron Monte Carlo events run through full detector simulation using the GEANT4 [26] simulation. Monte Carlo are software that use theoretical principles to produce events that model data.

4.2.4 Electron identification

Now that the electron candidates are reconstructed, the final step is to identify which electrons are suitable for analysis. For the analysis in this thesis, electron identification was done with a cut-based approach. It is possible for jets and photons to be mistakenly constructed as electrons instead. Variables describing the shape of the showers (outlined in Table 4.1) are used to discriminate against these background sources.

For the cut-based approach, each variable is sequentially required to be within some limit, and if it is not, the electron candidate is discarded. There are three levels of stringency: loose, medium, and tight, and electrons that pass these cuts are accordingly called loose, medium, and tight electrons. Tight electrons are a subset of medium electrons, and medium electrons are a subset of loose electrons. With increasing tightness, more variables are cut upon, and the cuts become stricter. The variables used in each type of electron cut are outlined in Table 4.2. The more rigid the cuts, the more likely it is the electrons selected are not background, but at the cost of discarding some signal electrons. Each set of cuts, loose, medium, and tight, are dependent on both η and E_T. The cuts must be energy dependent since the shape of the shower is energy dependent, e.g. more energetic showers are narrower, and also the cuts must vary in η due to non-uniformity in the amount of passive material the electron must pass through in the detector. There are 10 bins in η and 11 bins in E_T, on which the cuts are dependent.

Electron Discriminating Variables		
Type	Description	Name
Hadronic leakage	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\| < 0.8$ or $\|\eta\| > 1.37$)	R_{Had1}
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\| < 0.8$ or $\|\eta\| > 1.37$)	R_{Had}
Back layer of EM calorimeter	Ratio of the energy in the third later to the total layer	f_3
Middle layer of EM calorimeter	Lateral shower width, $\sqrt{\left(\Sigma E_i \eta_i^2\right)\left(\Sigma E_i\right) - \left(\left(\Sigma E_i \eta_i\right) / \left(\Sigma E_i\right)\right)^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated in a window of 3×5 cells	${W_\eta}^2$
	Ratio of the energy in 3×7 cells to the energy in 7×7 cells centered at the electron cluster position	R_η
Strip layer of EM calorimeter	Shower width, $\sqrt{\left(\Sigma E_i \left(i - i_{max}\right)^2\right) / \left(\Sigma E_i\right)}$, where i runs over all strips in a window of $\Delta\eta \times \Delta\phi \approx 0.0625 \times 0.2$, corresponding typically to 20 strips in η, and i_{max} is the index of the highest-energy strip	w_{stot}
	Ratio of the energy difference between the maximum energy deposit and the energy deposit in a secondary maximum in the cluster to the sum of these energies	E_{ratio}
Track quality	Number of hits in the b-layer (discriminates against photon conversions)	n_{Blayer}
	Number of hits in the pixel detector	n_{Pixel}
	Total number of hits in the pixel and SCT detectors	n_{Si}
	Transverse impact parameter	d_0
TRT	Total number of hits in the TRT	n_{TRT}
	Ratio of the number of high-threshold hits to the total number of hits in the TRT	F_{HT}
Track-cluster matching	$\Delta\eta$ between the cluster position in the strip layer and the extrapolated track	$\Delta\eta_1$
	$\Delta\phi$ between the cluster position in the middle layer and the extrapolated track	$\Delta\phi_2$
	Ratio of the cluster energy to the track momentum	E/p
Conversions	Veto electron candidates matched to reconstructed photon conversions	isConv

Table 4.1: Definition of electron discriminating variables [1].

Electron identification cuts			
Name	Loose	Medium	Tight
R_{Had1}	✓	✓	✓
f_3		✓	✓
$W_{\eta}2$	✓	✓	✓
R_{η}	✓	✓	✓
W_{stot}	✓	✓	✓
E_{Ratio}	✓	✓	✓
η_{Blayer}		✓	✓
η_{pixel}	✓	✓	✓
η_{Si}	✓	✓	✓
d_0		✓	✓
η_{TRT}		✓	✓
F_{HT}		✓	✓
$\Delta\eta_1$	✓	✓	✓
$\Delta\phi_2$			✓
E/p			✓
isConv			✓

Table 4.2: The variables used in the different selections of the electron identification menu [1].

4.3 Muons

Muons are able to interact with multiple subdetectors in ATLAS and so there are several different sets of criteria that can be used to reconstruct them. Information about muons can be used from the ID, the calorimeters, and of course the MS. Depending on which set of criteria is used, there are four different kinds of muons: stand-alone (SA) muons, segment-tagged (ST) muons, calorimeter-tagged (CaloTag) muons, and combined (CB) muons.

SA muons are reconstructed using only the MS. SA muons must pass through at least two layers of the MS chambers to provide a track measurement. Since the angular acceptance of the ID extends only to $|\eta| < 2.5$, SA muons can be used to extend the acceptance to $|\eta| < 2.7$.

ST muons combine tracks from ID with tracks from the MS. If a track from the ID is

extrapolated out to the MS and matches with a track segment from at least one layer of the MS, an ST muon can be reconstructed. The track segment from the MS must be from either the MDT or CSC chambers. ST tracks can also be used to increase acceptance where only one layer of the MS is hit.

CaloTag muons require a track from the ID to be matched with an energy deposition from the calorimeters. CaloTag muons have the lowest purity but can be used to gain statistics in the uninstrumented regions of the MS.

CB muons are the main type of muons reconstructed at ATLAS. CB muons are formed from the successful combination of a track from the ID and a track from the MS. These must be full tracks, not segmented tracks like those used to reconstruct ST muons. These muons have the highest purity and are used in most analyses at ATLAS. The advantage to using CB muons, in addition to high purity, is that the track parameters from the ID are available making the kinematics more reliable [27]. Additionally the track parameters can have further restrictions applied if the analysis requires it. CB muons are used in the cosmic muon analysis described in Appendix A.

4.4 Jets

Jets are the result of what happens when a parton goes through the process of parton showering and hadronization. When a colored object, such as a quark or gluon, is produced in a collision, it cannot exist as a final state particle. The parton radiates and produces a parton shower. After the shower evolves to a small scale (1 GeV), the quarks and gluons combine to form hadrons. Since the initial parton is highly boosted, the final state hadrons are collimated into a narrow cone which we call a jet. It is preferable to sum the momenta

of the jet's constituents and treat it as one hard object rather than many soft particles. For this task we use jet algorithms. Jet algorithms also attempt to sort the constituents such that, if there is another jet nearby, the appropriate particles are sorted into the proper jet.

4.4.1 Anti-k_T algorithm

Jets in this analysis were all constructed using the anti-k_T algorithm [28], so it will be described in detail here. The anti-k_T algorithm is a clustering algorithm. Clustering algorithms start with merging particles that are close in a variable called the "distance measure" to form a cone of a given radius. This process is iterated until all particles are clustered into jets. In general the distance measure for the k_T family of jet algorithms between particles i and j is

$$d_{ij} = \min\left(k_{Ti}^{2p}, k_{Tj}^{2p}\right)\frac{\Delta_{ij}^2}{R^2}, \tag{4.1}$$

where k_{Ti} is the transverse momentum of the i^{th} particle, Δ_{ij} is the spatial separation between the i^{th} and j^{th} particles in $\eta-\phi$ space, R is the given jet radius, and p is the parameter that determines the relationship between energy and spatial distance. The distance d_{iB} between particle i and the beam is

$$d_{iB} = k_{Ti}^{2p}. \tag{4.2}$$

All possible distance measures d_{ij} and d_{iB} are computed to find which is smallest. If the smallest is a d_{ij} measure, then the particles are combined and the distances are recalculated. If the smallest is a d_{iB} measure, then particle i is separated and called a new jet. The distance measures are continuously recalculated like this until all particles have been clustered.

The parameter R scales d_{ij} but not d_{iB}. This has the effect that any final pair of clustered jets will be spatially separated by at least R. This distance is often referred to as the cone size or jet radius. The parameter p is what determines what kind of algorithm is being used. In general it can be 1, 0, or -1. For the anti-k_T algorithm, the parameter p is set to -1. Setting p to 0 yields the Cambridge/Aachen algorithm which only takes into account the spatial separation of the particles. This has been shown to be especially useful in studying the substructure of jets. Setting p to 1, yields the k_T algorithm which takes into account the momentum of the particles. Finally setting p to -1 for the anti-k_T algorithm also takes into account the momentum of the particles but has the added effect of being less sensitive to pile-up and the underlying event.

4.4.2 Topological clustering algorithm

The analysis makes use of topological (topo) clusters as the inputs for the jet clustering algorithm. Topological clusters are 3-dimensional energy clusters meant to group cells from the hadronic or electromagnetic calorimeters into "energy blobs" based on their energy content and neighbor relations.

In order to form these clusters, the first step is to identify seed cells. Seed cells are determined by having an energy content with a signal to noise ratio above some threshold t_{seed}. The default ratio at ATLAS which is used in this analysis is 4. The noise is determined as the quadratic sum of both electronic noise and noise from pile-up (see Section 3.1.1).

$$\sigma_{total} = \sqrt{\sigma^2_{electronic} + \sigma^2_{pileup}}. \tag{4.3}$$

The electronic noise is measured when the beam is down and the pile-up noise is determined

from Monte-Carlo simulation. Once the seeds have been set, the cells directly neighboring the seed cells are checked to see if they have a ratio above a lower threshold $t_{neighbor}$ (the ATLAS default is 2). If they do, and the neighboring cell is not also a seed cell, they are then added to start to form clusters. If two clusters share a neighboring cell, the clusters are merged. After clusters have been formed, perimeter cells with a signal to noise ratio greater than t_{cell} are added. The ATLAS default is 0, so this includes all positive cells. A final step exists in the algorithm where any clusters with total energy below a threshold of E_{cut} are dropped. The default cut used is $|E_{\text{cut}}| = 0$. Because of the default thresholds, this noise suppression algorithm is often referred to as the 4-2-0 scheme.

Ideally these clusters would always be well isolated from each other; however this is not always the case. Cluster splitting is used to keep the clusters from growing too large and clustering different showered particles together. Local energy maxima are identified to split clusters. These are now called topo-clusters.

4.4.3 Jet Energy Calibration

Since the TileCal is a sampling calorimeter, not all of the energy is deposited into the active portion of the calorimeter. The actual kinematics of the jet must now be estimated.

4.4.3.1 Local cluster weighting

After the topo-clusters are constructed, a local cluster weighting (LCW) correction is used to form a second set of topo-clusters. The goal of the LCW method is correct for the non-compensating nature of the detector. The "non-compensating nature of the detector" refers to the fact that the calorimeters measure energy differently from electrons and photons, than from hadrons. The hadronic calorimeter is tuned to hadrons, but energy from electromag-

netic objects is still deposited in it. Without compensating for this fact, the energy will be mismeasured. LCW corrections are determined through Monte Carlo simulation of charged and neutral pions. The correction is applied probabilistically cluster by cluster. LCW is reviewed in depth in Ref. [29].

4.4.3.2 Origin correction

After the LCW corrections have been applied, the new topo-clusters are matched to the first primary vertex. By default, it is assumed that the jet originated from the center of the detector, but a better assumption is that it came from the first primary vertex of the interaction. The first primary vertex is defined as the vertex that has the largest Σp_T^2 of tracks with at least 400 MeV associated with it. The origin correction changes the four-vector of the jet such that the energy remains unchanged, but the trajectory now points toward the first primary vertex.

4.4.3.3 Pile-up correction

To reduce the effects from pile-up on jets, an area based subtraction method is used. Looking at the jets in the $\eta \times \phi$ plane, two variables are defined: A, the area of the jet in this plane, and ρ, the energy density. The area is defined using an active areas algorithm in which ghost particles (particles with infinitesimal momentum) are uniformly added to the event and the jets are reclustered. The number of ghost particles clustered into the jet is a measure of its area. The energy density of each jet is then simply the p_T/A and the energy of the event is the median energy density of the jets. The pile-up energy density is calculated using the k_{T} algorithm, for $R = 0.4$ jets reconstructed in the central region ($|\eta| < 2.0$). The correction is dependent on the variables associated with pile-up: the number of primary vertices (N_{PV})

and the average number of interactions per bunch crossing $\langle\mu\rangle$. The correction is given by

$$p_T^{corr} = p_T^{const} - \rho \times A - \alpha \times (N_{PV} - 1) - \beta \times \langle\mu\rangle, \tag{4.4}$$

where α and β are the jet size and algorithm based variables determined from Monte Carlo, and p_T^{const} is the jet p_{T} from the topo-cluster.

4.4.3.4 Jet energy scale correction

The jet energy scale (JES) correction is a scaling factor applied to the jet energy, derived from the comparison of reconstructed jet energy to the true jet energy using Monte Carlo. The JES factor is found using isolated jets from an inclusive jet Monte Carlo sample after the origin and pile-up corrections are applied. It was found that the JES correction has an angular dependence and so the JES factor is dependent on η [29].

4.4.3.5 Global Sequential Correction

It turns out that there is a response difference between jets initiated by quarks and jets initiated by gluons by up to 8% which is not taken into account by the JES correction. This difference in the past was responsible for large uncertainties in the JES correction. Part of this difference is due to "punch-through", which happens when a jet is energetic enough that not all of its energy is contained within the calorimeter. In order to improve this, further corrections are applied.

These corrections are done using Monte Carlo and are done in order based on the following variables:

1. The fraction of energy deposited in the first layer of the tile calorimeter.

2. The fraction of energy deposited in the third layer of the electromagnetic calorimeter.

3. The number of tracks with $p_{\mathrm{T}} > 1$ GeVassociated to a jet.

4. The p_{T}-weighted transverse width of the jet measured using tracks with $p_{\mathrm{T}} > 1$ GeV associated to the jet.

5. The amount of activity behind the jet as measured in the muon spectrometer.

The first two corrections are skipped in this analysis since they are already taken care of in the LCW correction. More can be read about the GSC correction in Ref. [29, 30].

4.4.4 Jet Vertex Fraction

With more pile-up, there is more energy from the hard scattering that is not associated with jets, and it becomes harder to reconstruct jets properly. The precision of the jet measurements suffer from this. Placing a cut on the jet vertex fraction (JVF) can help suppress the uncertainties associated with pile-up.

JVF is a measure of the amount of jet energy associated with a particular primary vertex. It is defined as the summed transverse momentum of all tracks matched with a particular jet $\mathrm{track}_k^{\mathrm{jet}}$ and primary vertex PV_0, divided by the summed momentum from all tracks associated with a jet $\mathrm{track}_l^{\mathrm{jet}}$ regardless of what PV_n to which they are matched [31].

$$\mathrm{JVF}(\mathrm{jet}, \mathrm{PV}_0) = \frac{\Sigma_k p_{\mathrm{T}}(\mathrm{track}_k^{\mathrm{jet}}, \mathrm{PV}_0)}{\Sigma_n \Sigma_l p_{\mathrm{T}}(\mathrm{track}_l^{\mathrm{jet}}, \mathrm{PV}_n)} \tag{4.5}$$

Therefore, JVF ranges from 0 to 1, where the larger it is, the more likely a jet is associated with a primary vertex. Jets are required to have a JVF > 0.5.

4.4.5 Jet Overlap

Electrons and jets are reconstructed separately from each other and so it is possible to mistakenly reconstruct the same object as both a jet and an electron. This happens not just from coincidence, but even more likely from an electron carrying a significant momentum fraction of the jet being reconstructed as a final state electron. In order to avoid this, an isolation cut is applied on jets to make sure they are not in proximity to electrons. Jets within a radius of $\Delta R < 0.2$ of a reconstructed electron are assumed to actually be electrons and are removed from the event. Then if there is a jet within $\Delta R < 0.4$ of an electron, the event is vetoed.

In the 2011 $W+$jets analysis, jets were only considered as signal jets if they did not overlap with the selected electron within $\Delta R < 0.5$. In 2014 the ATLAS recommendation for jet-electron isolation changed [32]. The new recommendation suggests to veto events with jet-electron overlap, in contrast to simply excluding jets that do not make the isolation cut. A study was conducted comparing these two methods. To be precise for this study: when applying the 2014 recommendation, jets with an electron within $\Delta R < 0.2$ are first removed and then events with an overlap $\Delta R < 0.4$ are vetoed. As mentioned before, the $\Delta R < 0.2$ removal is to protect against double counting since the electrons are treated as jet structures early in the analysis code.

The comparison of the 2014 treatment with respect to the previous 2011 $W+$jets treatment is shown in Figure 4.2 for the exclusive jet multiplicity, lead jet p_{T}, H_{T}, and ΔR separation between the two lead jets. The effect is more pronounced at larger jet multiplicity where the effect of having more jets is that there is more likely one near an electron, but the effect is negligible at low jet multiplicity. The same is true for larger energies which generally

correspond to larger jet multiplicities. As a result of this study, the new 2014 treatment for jet-electron overlap removal was adapted for this analysis.

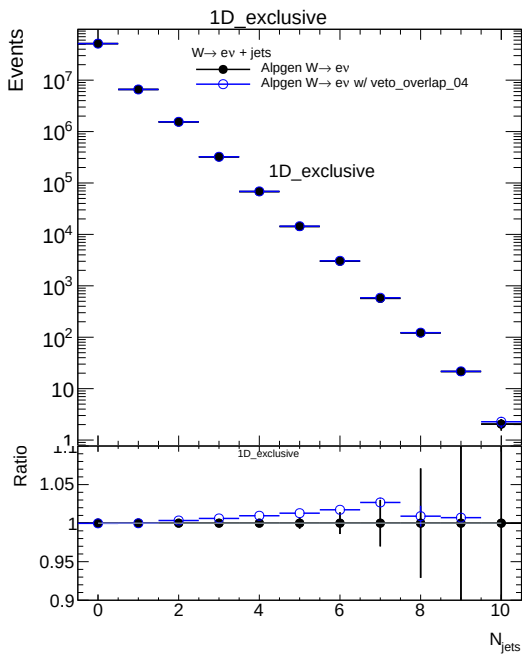
4.4.6 b-tagging

It is often useful to differentiate the b-quark jets (b-jets) from other types of jets. In this analysis, events with b-jets are rejected in order to help reduce the background from events with a top and antitop ($t\bar{t}$) (detailed further in Section 5.3). This is not generally easy since jets appear similar in the detector. However, there are some attributes of b-jets that can help distinguish them. For example, hadrons containing a b-quark have a lifetime on the order of 1.5 ps before they decay and hadronize. This means that tracks traced back to the vertex of the b-jet will be a measurable distance away from the primary vertex of the collision. As such, observing the impact parameter, or identifying a jet that comes from a secondary vertex will help identify jets that are likely from b-quarks. Additionally since the b-quark is more massive relative to its decay products than lighter jets, the constituents of the jet tend to be more spread out in the detector [33].

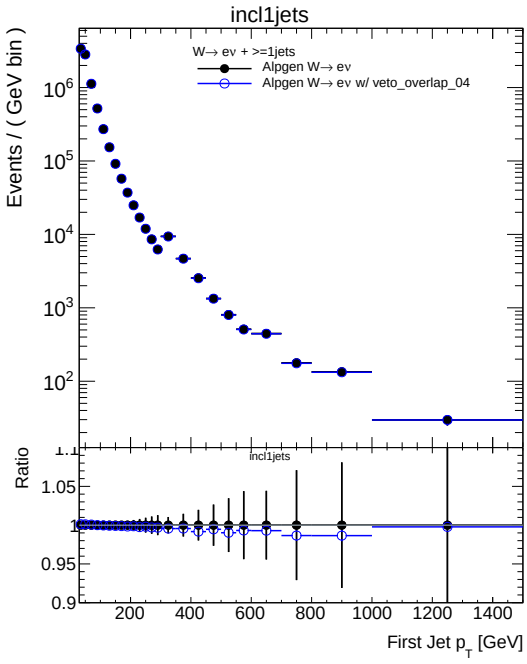
This analysis uses a multivariate tagging algorithm called MV1. The most discriminating variables that describe the impact parameter, secondary vertices, and jet topology are run through neural networks, and weighted probabilities are output. The multivariate tagging algorithm uses these probabilities to tag which jets are b-jets [34].

4.5 Missing transverse energy

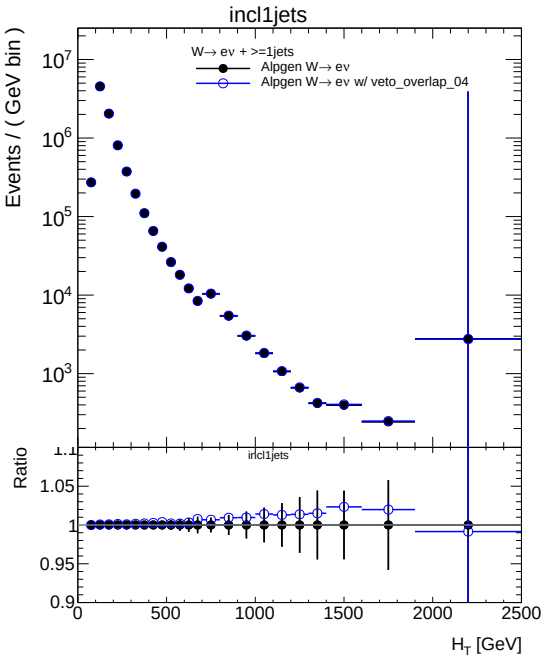
The final state particles should have the same total momentum as the initial particles from the collision (conservation of momentum). Since the initial state particles are the quarks and



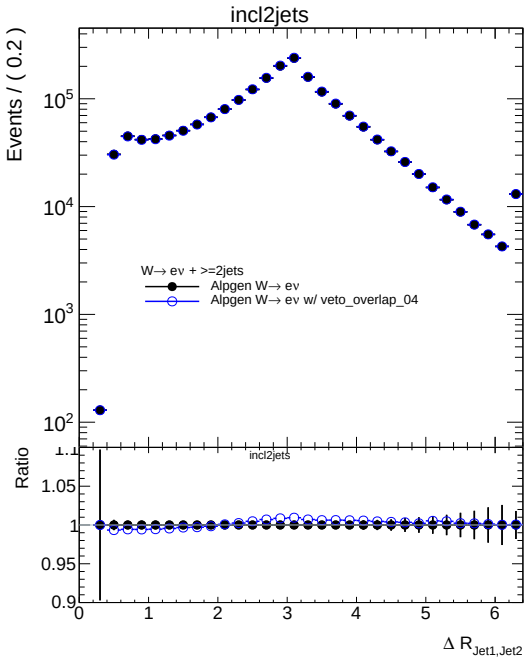
(a) Exclusive Jet Multiplicity



(b) Lead Jet p_T for inclusive 1 jet production



(c) H_T for inclusive 1 jet production



(d) ΔR separation between two leading jets

Figure 4.2: The comparison of removing jets within $\Delta R < 0.5$ of electrons to vetoing events containing an overlap within $\Delta R < 0.4$

gluons within the protons, the amount of z-momentum that each one carried is unknown, but there is little initial momentum in the transverse direction. Therefore, the transverse components of all the final state particles should add to zero, within the measurement resolution of the detector. Any non-zero transverse momentum is indicative that some energy was not reconstructed in the detector; this energy is labeled missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ (MET). If this energy is significant, it is likely from neutrinos which do not deposit energy in the detector, since they are charge neutral and interact only weakly. When computing $E_{\mathrm{T}}^{\mathrm{miss}}$, all energy deposited in the detector is included in the vector sum, including energy not associated with reconstructed objects (soft terms). $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined then as the negative vector sum of all the transverse momentum in the detector.

Chapter 5

Data and backgrounds

The data used in this analysis were recorded by ATLAS in 2012 during the proton-proton collision run with a center of mass beam energy of 8 TeV. During this run the bunch crossings were separated by 50 ns. The average number of inelastic proton-proton collisions per bunch crossing $\langle\mu\rangle$ was 20.7 and the peak instantaneous luminosity[1] was $7.7 \times 10^{33}\,\mathrm{cm}^2\mathrm{s}^{-1}$. Excluding parts of the run that did not meet data-quality requirements, the integrated luminosity was $20.2\,\mathrm{fb}^{-1}$.

Most of the collisions do not produce events relevant to this analysis, namely those with a W boson decaying to an electron and neutrino as well as some number of jets. In order to select the events of interest, a series of cuts are performed as described in detail in Section 5.1. Many of the events that pass the selection are actually background. For example, an event with only jets and no W boson produced can potentially have a jet mismeasured as an electron with enough missing transverse energy to look like a W boson decay. This event might pass the event selection despite not being a signal event. In order to account for this, background events are estimated using various data-driven techniques and Monte Carlo described in more detail in Sections 5.2-5.4.

[1] Instantaneous luminosity is the number of protons passing through a given area in a given amount of time

5.1 Data selection

Events are first selected using a single electron trigger requiring either that the event contains at least one isolated electron with a transverse momentum greater than 24 GeV or any electron with at least 60 GeV. For this trigger algorithm, the electron is considered to be isolated if a variable called isolation momentum is less than 10% of the electron's transverse momentum. Isolation momentum is defined as the sum of the p_{T} of reconstructed charged-particle tracks with p_{T} greater than 1 GeV within a cone of size $\Delta R < 0.2$ around the electron (not including the track of the electron). The isolation cut reduces the trigger rate, allowing the threshold p_{T} to be set lower at 24 GeV. The higher 60 GeV trigger is set high to compensate for the lack of an isolation requirement.

All events are required to have at least one reconstructed vertex with at least three associated tracks each having p_{T} greater than 400 MeV. The vertex with the highest Σp_{T}^2 associated with it is considered to be the primary vertex (PV).

5.1.1 Electron selection

Reconstructed electrons must meet the *tight* electron ID requirements introduced in Section 4.2.4 (and further discussed in Ref [1]). Electrons must then have a transverse momentum of at least 25 GeV and have a pseudorapidity within the range $|\eta| < 2.47$ excluding $1.37 < |\eta| < 1.52$. This η-range is excluded due to it being the region where the barrel calorimeter transitions to the endcap, and there is a gap containing dead material. The dead material increases the uncertainty associated with reconstructing the electron energy. The electron needs to originate from the PV. To check that this is the case, the impact parameter d_0 must be less than five times its uncertainty ($|d_0|/\sigma_{d_0} < 5$) and $|z_0 \cdot \sin\theta|$ must be smaller than

0.5 mm, where z_0 is the longitudinal impact parameter and θ is the polar angle (see Section 4.1.1).

Furthermore a track-based and a calorimeter-based isolation restriction is placed on the the electron as well. Similar to the isolation from the trigger, the sum of p_{T} of the tracks within $\Delta R < 0.3$ of the electron (excluding the electron track) are defined as p_{T}^{cone30}. The tracks contributing to the sum are required to have $p_{\mathrm{T}} > 400\,\mathrm{MeV}$ and impact parameters $|d_0| < 1.5\,\mathrm{mm}$ and $|z_0| < 1.0\,\mathrm{mm}$. This variable p_{T}^{cone30} is required to be less than 7% of the electron's transverse energy. For the calorimeter based isolation, the sum of the positive EM topological clusters in a radius $\Delta R < 0.3$ is called isolation energy E_T^{cone30}. The energy from the electron, corresponding to the 5×7 cells in the barycenter, is subtracted off the sum, and the energy is corrected to compensate for energy leakage as well as in-, and out-of-time, pile-up. The isolation energy is then required to be less than 14% of the electron's transverse energy.

5.1.2 Jet selection

Jets are constructed using the anti-k_{T} algorithm described in Section 4.4.1, and are required to have $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and a rapidity $|y| < 4.4$. To suppress jets from out-of-time pile-up, jets with $p_{\mathrm{T}} < 50\,\mathrm{GeV}$ are required to have at least 50% of their transverse momentum arise from tracks associated with the primary vertex (JVF > 0.5 see Section 4.4.4). This requirement is applied to jets within the acceptance of the inner detector ($|\eta| < 2.4$), where tracking information is available.

5.1.3 Event selection

Events are required to have exactly one electron passing all cuts. If a second electron exists passing the *medium* identification, with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.47$ (excluding $1.37 < |\eta| < 1.52$), the event is rejected. To avoid an already identified electron being counted as a jet, any jet within a cone $\Delta R < 0.2$ of an electron is removed from the event. This is to suppress background from $Z \to ee$. After this, if a jet is within $\Delta R < 0.4$ of an electron, the event is rejected. This ensures that particles are well isolated and prevents energy contamination.

The event is required to have $E_{\mathrm{T}}^{\mathrm{miss}} > 25\,\mathrm{GeV}$, consistent with the presence of a neutrino. Furthermore, the transverse mass of the W boson, defined as $m_T = \sqrt{2p_{\mathrm{T}}^{e}p_{\mathrm{T}}^{\nu}\left(1-\cos\left(\phi^{e}-\phi^{\nu}\right)\right)}$, must be greater than $40\,\mathrm{GeV}$, where p_{T}^{e} and p_{T}^{ν} are the transverse momentum of the electron and neutrino (same as the $E_{\mathrm{T}}^{\mathrm{miss}}$) respectively, and ϕ^{e} and ϕ^{ν} are the ϕ-coordinates of the electron and neutrino respectively.

Finally, to suppress background from $t\bar{t}$ decays, jets containing b-quarks are identified using the b-tagging process described in Section 4.4.6. If the event contains a b-jet, the event is vetoed. The event selection is detailed in Table 5.1.

5.2 Multi-jet Background

The multi-jet (MJ) background comes from events containing only jets, where one of the jets is misidentified as an electron. There are three main ways this can happen: A light flavor jet (one originating from an up, down, or strange quark) is misidentified as an electron, a jet containing a charm or bottom quark has the hadron decay into an electron, or an electron

Pre-selection	
Vertex	Primary vertex with $N_{\text{tracks}} \geq 3$
Electron selection	
Pseudorapidity	$\lvert\eta\rvert < 2.47$ (excluding $1.37 < \lvert\eta\rvert < 1.52$)
Electron ID	Tight
Electron p_{T}	$p_{\text{T}} > 25\,\text{GeV}$
Electron Isolation	$p_{\text{T}}^{cone30}/p_{\text{T}} < 0.07$ and $E_{\text{T}}^{cone30}/E_{\text{T}} < 0.14$
Electron impact parameters	$\lvert d_0/\sigma_{d_0}\rvert < 5$ and $\lvert z_0 \cdot \sin(\theta)\rvert < 0.5$
$W \to e\nu$ selection	
Z veto	no second electron fulfilling *Medium* identification, as well as pseudorapidity requirement, and $p_{\text{T}} > 20\,\text{GeV}$
Electron-jet distance	ΔR(selected electron,signal jet) ≥ 0.4
Missing energy	$E_{\text{T}}^{\text{miss}} > 25\,\text{GeV}$
Transverse mass	$m_{\text{T}} > 40\,\text{GeV}$
B-jet veto	No b-jet ($p_{\text{T}} > 20\,\text{GeV}$)

Table 5.1: Requirements for electron selection and event selection defining the signal region. Events not passing these cuts are discarded from the analysis.

from a photon conversion passes the signal selection. Additionally the event must contain $E_{\text{T}}^{\text{miss}}$ from either a mismeasurement of energy deposited by jets, or from neutrinos from heavy flavored decays (charm or bottom decays). Even though the rate of jets faking electrons is low, the cross-section of MJ events is large enough that it is the largest background for signal events containing 3 jets or fewer.

To estimate the MJ background, a data-driven method is used for two reasons. The first is that Monte Carlo simulation of particles faking other particles is not sufficiently accurate. The second is that, due to the low fake rate, the Monte Carlo samples are not large enough to produce sufficient statistics for a reliable background estimate.

The estimate is carried out by observing the $E_{\text{T}}^{\text{miss}}$ distribution. The total measured cross-section is the combination of the MJ background, plus the signal and all other backgrounds combined (denoted as signal+EW backgrounds). A well reconstructed multi-jet event is not

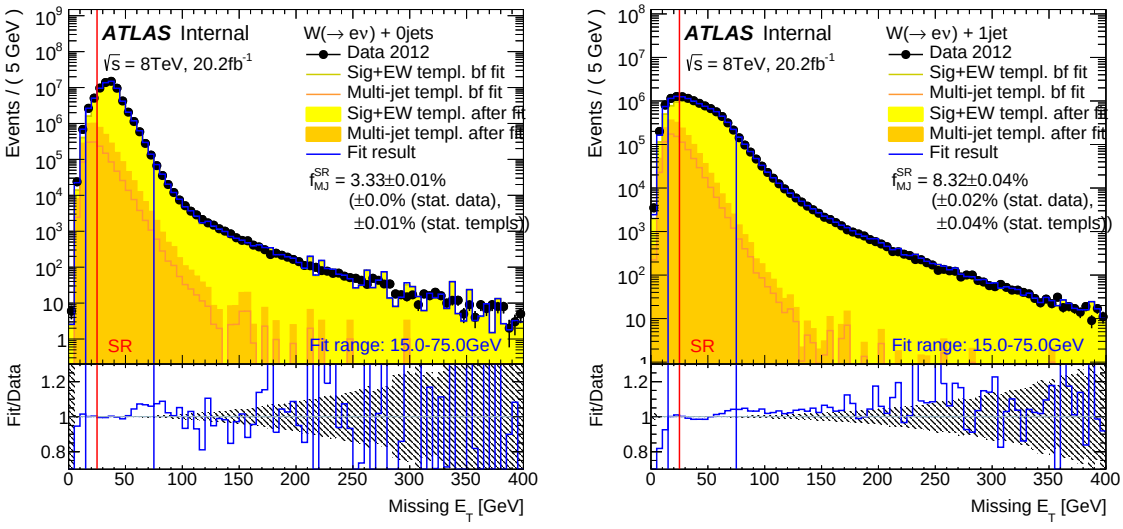
expected to have a significant amount of $E_{\mathrm{T}}^{\mathrm{miss}}$, since there are no high p_{T} neutrinos to carry away energy outside of the detector. As discussed previously though, mismeasurement of an MJ event can result in $E_{\mathrm{T}}^{\mathrm{miss}}$. However, the distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ for the combination of signal+EW backgrounds does contain real neutrinos and is expected to peak at 40 GeV (since it should have the energy of half the mass of the W boson). The MJ background is therefore estimated by fitting two templates: the first describing the MJ background, and the second describing the signal+EW backgrounds. For the fit, the requirement $E_{\mathrm{T}}^{\mathrm{miss}} > 25\,\mathrm{GeV}$ is removed. All other requirements, including $m_T > 40\,\mathrm{GeV}$, are kept.

The shape of the MJ template is determined by selecting data with fake electrons. The region of the data selected is referred to as the MJ control region (MJ-CR). The event selection is kept the same except for the electron ID, the isolation, and the impact parameter cuts which are reversed. The MJ control region is summarized in Table 5.2. There exists contamination from the signal+EW backgrounds which is estimated using Monte Carlo, and is generally less than 1.5% (4-5% where no jets are selected). The shape of the signal+EW backgrounds is also taken from Monte Carlo.

MJ-CR	
Reverted electron ID	Medium, but not Tight
Anti-isolation	$p_{\mathrm{T}}^{cone30}/p_{\mathrm{T}} > 0.07$
Impact parameters	None

Table 5.2: MJ-CR requirements which differ from the signal selection as listed in 5.1.

An extended maximum likelihood fit is used in the range of 15-75 GeV, to avoid fitting to the behavior in the tail, and is done separately in exclusive jet bins up to 5 jets. Events with 6 or more jets are done inclusively. The fit for the 0 and 1 jet exclusive bins are shown in Figure 5.1.



(a) No jets

(b) 1 jet

Figure 5.1: Fitted $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in the SR for the jet multiplicities 0 and 1. The blue line indicates the total fit result, while the yellow and orange shapes represent the signal+EW backgrounds and multi-jet templates, respectively. The colored lines show the templates before the fit, while the colored areas represent the templates normalized to the fit result. The SR with the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement $> 25\,\mathrm{GeV}$ is marked by the vertical red line. The vertical blue lines mark the fit range. The lower plots show the ratio of the combined fit result w.r.t. data.

The scale factor is the factor applied to the template to get the fitted result. This is done to get the correct normalization. In order to estimate the MJ background in the signal region (SR), the scale factor is applied and the requirement that $E_{\mathrm{T}}^{\mathrm{miss}} > 25\,\mathrm{GeV}$ is applied again. This is done for all kinematic distributions. The results of the fit are summarized in Table 5.3

n_{jet}	$f_{\mathrm{MJ,SR}}$	$\mathrm{SF}_{\mathrm{MJ}}$
0	3.33	3.30
1	8.32	2.09
2	14.71	2.64
3	15.95	2.64
4	16.31	2.70
5	16.57	3.08
$\geq$6	14.72	3.08

Table 5.3: The fraction of multi-jet events in the signal region in % ($f_{\mathrm{MJ,SR}}$), and the MJ-scale factors (SF) for the different fitted jet multiplicities (n_{jet}).

The statistical uncertainty for the multi-jet fit is found during the fitting procedure. First, the likelihood is assumed to be parabolic around its maximum and minimum. The fitted parameter is then adjusted by hand until the likelihood is changed by $\Delta \log L$, and the amount the parameter is adjusted is then taken as the uncertainty.

5.3 $t\bar{t}$ Background

The top-antitop ($t\bar{t}$) background is modeled using Monte Carlo; however, a normalizing factor of 1.086 is applied to account for an observed difference in the overall normalization with respect to data. The background comes mainly from $t\bar{t} \to b\bar{b}q\bar{q}e\nu$, which is why the

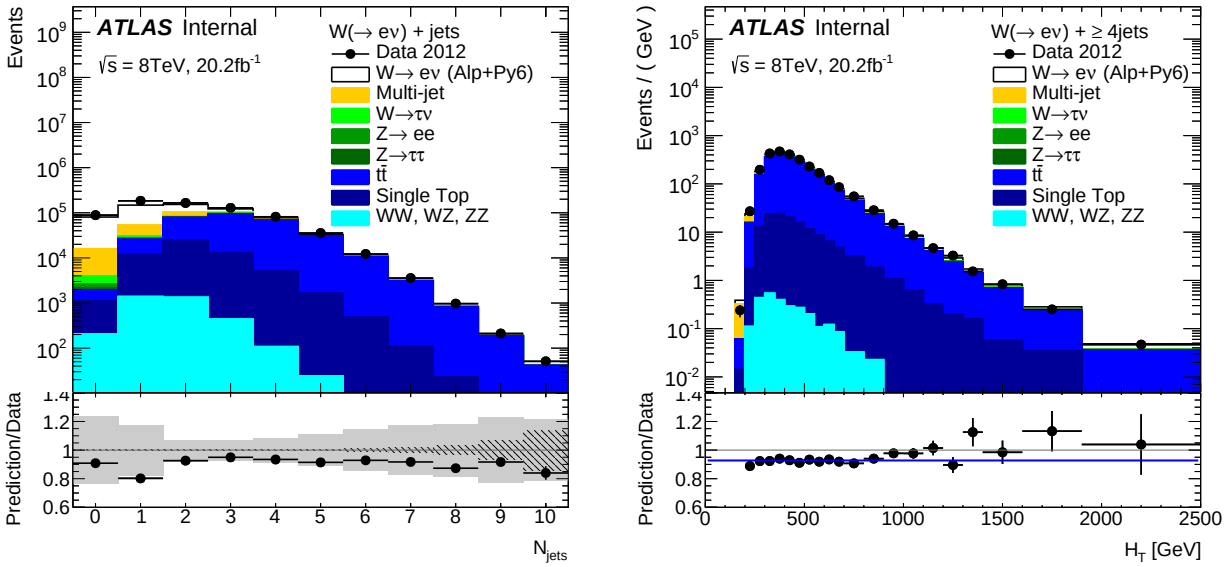
background becomes significant starting at 4 jets. The Monte Carlo is cross-checked using an enriched sample from the data, similar to the multi-jet control region used in the MJ background. This region is called the validation region and uses all the same cuts as for the signal region (see Table 5.1) but the b-jet veto is reversed to require at least one b-jet. The multi-jet contribution for the validation region is found using the same method described in Section 5.2.

Figure 5.2 shows the comparison of data to Monte Carlo in the validation region for the exclusive jet multiplicity and the H_{T} distribution[2] in the presence of at least 4 jets. There is a global offset similar to what is observed in the $t\bar{t}$ cross-section measurement by ATLAS [35] The correction factor is found by fitting the ratio of prediction to data for the H_{T} distribution in the presence of at least 4 jets in the validation region. The fit is a constant single parameter fit. The fit was also tried using a non-zero slope, but the difference was minor. The parameter found from the fit translates to the normalization 1.086 mentioned earlier.

5.4 Other Backgrounds

All other backgrounds are estimated using Monte Carlo. These include backgrounds from diboson events (WW, WZ, ZZ) where for example one boson decays leptonically and the other hadronically. Additionally there are backgrounds from vector boson plus jet events ($W \to \tau\nu$, $Z \to ee$, $Z \to \tau\tau$) mainly where τ-leptons decay leptonically via $\tau \to \nu_\tau \bar{\nu}_e e$, or else extra leptons are missed. Lastly, single top events can look similar to signal events when the top decays to $t \to be^+\nu$ and the bottom quark is not tagged as a b-jet. Table 5.4

[2] H_{T} is the scalar sum of the p_{T} of all the hard objects in the event including jets, the electron and the $E_{\mathrm{T}}^{\mathrm{miss}}$.



(a) N_{jets} (b) H_{T}

Figure 5.2: Data-to-prediction comparison in the $t\bar{t}$ validation region as a function of the exclusive jet multiplicity (a) and H_{T} in the presence of at least 4 jets (b). The lower panel displays the ratio of the prediction over data. The gray and hashed bands in (a) indicate the fractional size of detector level systematic and data statistics, respectively. The uncertainties in the ratio in (b) are determined via error propagation from the numerator and denominator uncertainties, treating them as uncorrelated. The ratio in (b) is fitted with a constant. For more details see Section 5.3 [2].

shows the estimated fraction of the number of background events compared to signal events generated from Monte Carlo, as well as the estimates using data-driven methods for the MJ background and $t\bar{t}$ background.

	No jets	1 jet	2 jets	3 jets	4 jets	5 jets	6 jets	7 jets	8 jets	9 jets	10 jets
WW, WZ, ZZ	<0.1	0.4	1.0	1.1	0.8	0.5	0.3	0.2	<0.1	<0.1	<0.1
Single t	<0.1	0.2	0.8	1.3	1.7	1.8	1.9	1.4	3.1	2.8	0.9
$t\bar{t}$	<0.1	0.2	1.3	6.4	16.3	26.6	35.9	42.8	46.2	55.1	57.8
$Z \to \tau\tau$	0.1	0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.5	<0.1
$Z \to ee$	0.3	2.6	4.5	5.7	5.9	5.7	5.5	5.4	4.6	3.6	4.2
$W \to \tau\nu$	1.8	2.0	2.0	1.9	1.8	1.4	1.5	1.0	2.1	0.6	<0.1
Multi-jet	3.5	8.5	14.9	16.2	16.4	16.2	14.2	13.9	13.7	11.6	20.0
$\boldsymbol{W \to e\nu}$	94.4	86.1	75.3	67.1	56.8	47.4	40.5	35.0	30.0	25.8	17.2

Table 5.4: Fraction of the number of background, and W+jets signal events, expected from MC simulation and from the data-driven method in the SR. The numbers are in % w.r.t. the total SM prediction for the different exclusive jet multiplicities [2].

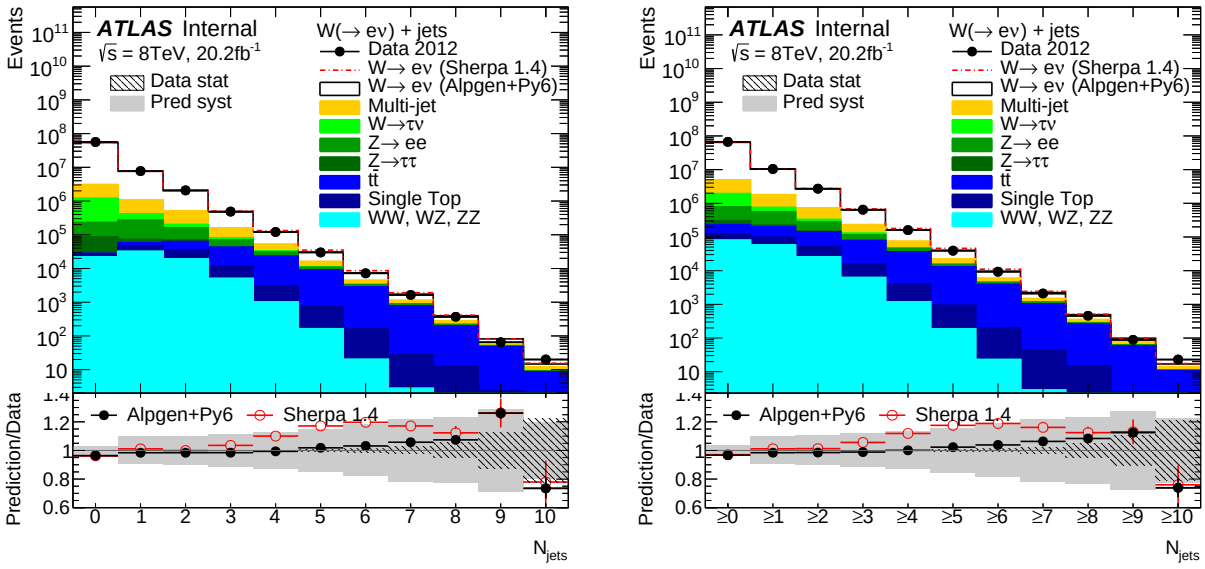
Chapter 6

Detector Level Results

The yield of W+jets events is determined by subtracting the background estimates described in Sections 5.2-5.4 from the event count in data. At this point, these detector level results can be compared to Monte Carlo. The number of events are listed in Table 6.1 for each exclusive jet multiplicity. Jet multiplicity is the number of jets in an event. Exclusive means there are exactly some number n jets in an event, as opposed to inclusive jet multiplicity, which means there are at least n jets in an event. A graphical representation of both the inclusive and exclusive jet multiplicity distributions are shown in Figure 6.1 comparing the data to the Monte Carlos ALPGEN [36] and SHERPA [37]. The label `Alpgen+Py6` refers to the fact that PYTHIA6 [38] is used to shower the particles after ALPGEN calculates the matrix elements. More details about the Monte Carlo are discussed in Chapter 9. The gray uncertainty bands are the detector level systematic uncertainties which are described in detail in Section 8. Bins are divided by bin width to smooth out distributions and make it easier to compare bins of different width. For this reason some data bins can have a fractional number of events [2].

	No jets	1 jet	2 jets	3 jets	4 jets	5 jets
WW, WZ, ZZ	22 102 ± 76	31 894 ± 92	19 352 ± 71	5025 ± 36	1006 ± 16	163 ± 6.5
Single t	3549 ± 28	13 168 ± 56	16 243 ± 65	6443 ± 53	2030 ± 33	561 ± 18
$t\bar{t}$	2028 ± 12	11 678 ± 29	26 586 ± 44	30 415 ± 47	19 543 ± 38	8093 ± 25
$Z \to \tau\tau$	56 760 ± 270	15 030 ± 110	5642 ± 64	1510 ± 32	365 ± 15	96.7 ± 6.5
$Z \to ee$	144 400 ± 690	194 890 ± 750	92 050 ± 470	27 250 ± 240	7106 ± 86	1732 ± 31
$W \to \tau\nu$	954 900 ± 2800	148 960 ± 560	39 850 ± 250	9320 ± 110	2154 ± 45	437 ± 20
Multi-jet	1 876 900 ± 2700	643 500 ± 1200	304 490 ± 910	77 530 ± 460	19 710 ± 230	4940 ± 120
$\mathbf{W \to e\nu}$	51 249 000 ± 22 000	6 552 700 ± 3700	1 533 800 ± 1400	321 150 ± 490	68 280 ± 190	14 422 ± 75
Total predicted	54 310 000 ± 22 000	7 611 700 ± 4000	2 038 000 ± 1700	478 640 ± 720	120 190 ± 320	30 450 ± 150
Data observed	56 342 232	7 735 501	2 070 776	486 158	120 943	29 901
Multi-jet/(W + jets)	0.037	0.098	0.199	0.241	0.289	0.343
	6 jets	7 jets	8 jets	9 jets	10 jets	
WW, WZ, ZZ	20 ± 2.3	2.8 ± 0.8	0.1 ± 0.1	0 ± 0	0 ± 0	
Single t	138.6 ± 9.2	24.6 ± 3.4	12.3 ± 3.3	2.3 ± 1.3	0.1 ± 0.1	
$t\bar{t}$	2670 ± 14	743.2 ± 7.5	182.5 ± 3.7	45.2 ± 1.9	8.5 ± 0.8	
$Z \to \tau\tau$	20.8 ± 2.7	5.1 ± 2	1.4 ± 0.6	0.4 ± 0.3	0 ± 0	
$Z \to ee$	408 ± 11	93 ± 5.6	18.1 ± 1.7	3.0 ± 0.8	0.6 ± 0.4	
$W \to \tau\nu$	110 ± 10	18.2 ± 3.3	8.1 ± 3.7	0.5 ± 0.5	0 ± 0	
Multi-jet	1057 ± 51	241 ± 12	54 ± 2.6	9.5 ± 0.5	2.9 ± 0.1	
$\mathbf{W \to e\nu}$	3007 ± 29	608 ± 13	118.7 ± 5.5	21.1 ± 2.2	2.5 ± 0.5	
Total predicted	7430 ± 63	1735 ± 20	395.2 ± 8.9	82 ± 3.4	14.7 ± 1.1	
Data observed	7204	1641	368	65	20	
Multi-jet/(W + jets)	0.351	0.396	0.455	0.451	1.163	

Table 6.1: Number of W + jets events expected from MC simulation and from data-driven methods, as well as the number of events observed in data, for the different exclusive jet multiplicities. The uncertainties listed here are the statistical uncertainties in the predictions only. The event counts have been rounded to 2 significant digits on the associated error, apart from very low event counts where only 1 significant digit is given.



(a) Exclusive jet multiplicity

(b) Inclusive jet multiplicity

Figure 6.1: Jet multiplicity distribution in the W+jets signal region, (a) exclusive and (b) inclusive in the number of jets, for the SM prediction compared to data at the detector level. The lower plot shows the ratio of the prediction over data, where the hashed band gives the fractional size of the data statistical uncertainties, the gray band shows the fractional size of the total uncertainty in the prediction (total systematic, luminosity and statistical) and the error bars on the ratio indicate the fractional size of the statistical uncertainties in the prediction only. For a list of the systematic uncertainties contained in the gray band see Chapter 8.

A total of 67 million events in data pass cuts to make it into the signal region. About 16% of these contain at last one jet. Of these events, the fraction of signal events ranges from 94% in the events with no jets, down to 17% in the events with 10 jets. The multi-jet background (Section 5.2) is the largest background at low jet multiplicities until the $t\bar{t}$ background takes over starting at the exclusive 5 jet bin (or the inclusive 4 jet bin).

Compared to a similar analysis done on 2011 data, the multi-jet background is lower for the 0 and 1 jet bin, but still slightly higher for 2 or more jets. This shows the success of the suppression of the multi-jet background despite the increased pile-up in the 2012 run [16].

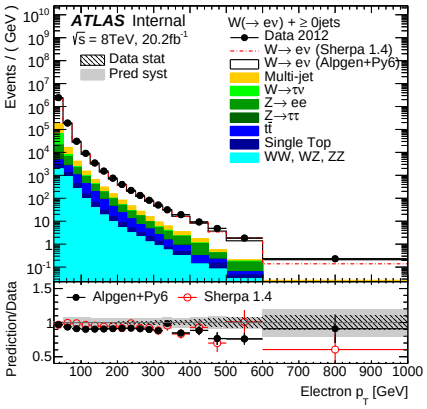
6.1 Kinematic Distributions

In this section the differential distributions are shown for various kinematics. This is not a complete list of the final distributions shown at particle level in Chapter 10, but they serve as a benchmark in the analysis before performing the unfolding and calculating uncertainties. Figure 6.2 shows the kinematics relating to the W-boson independent of the jet multiplicity:

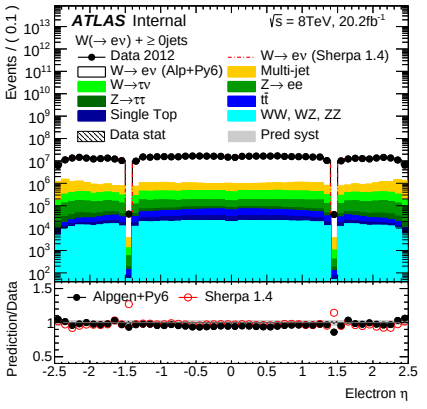
- The transverse momentum of the electron from the W (Electron p_{T}),
- The pseudorapidity of the electron (Electron η),
- The $E_{\mathrm{T}}^{\mathrm{miss}}$, which should correspond to the p_{T} of the neutrino from the W,
- The transverse mass of the W-boson (m_{T}),
- The transverse momentum of the W (Wp_{T}),
- And the H_{T}, the sum of the p_{T} of all hard objects in the event.

The same distributions for the W-related kinematics as well as the jet-related distributions for the events containing at least one jet are shown in Figures 6.3-6.4. The same for the events containing at least 2 jets in Figures 6.5-6.7. The additional jet kinematics are:

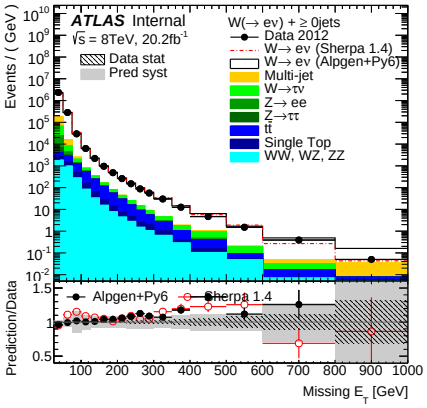
- The transverse momentum of the jet with the highest p_{T} (leading jet p_{T}),
- The rapidity of the jet with the highest p_{T} (leading jet y),
- The transverse momentum of the jet with the second highest p_{T} (second leading jet p_{T}),
- The rapidity of the jet with the second highest p_{T} (second leading jet y),
- The dijet invariant mass of the two leading jets (jet invariant mass $m_{1,2}$),
- The ΔR separation between the two lead jets, where $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ (Angular separation $\Delta R_{1,2}$).



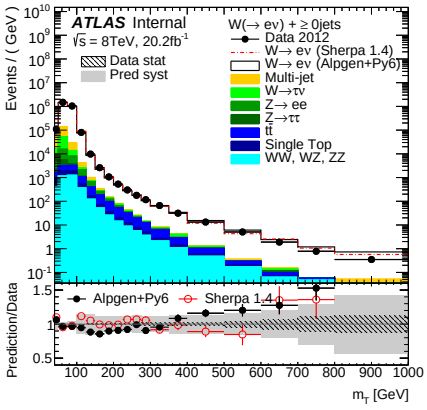
(a) Electron p_{T}



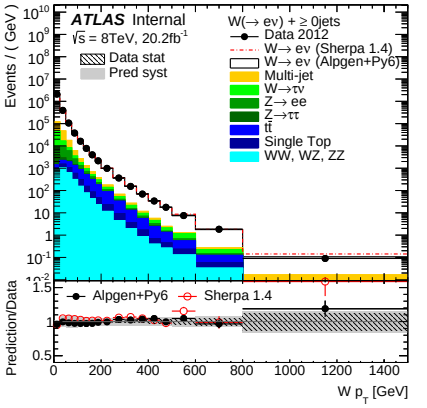
(b) Electron η



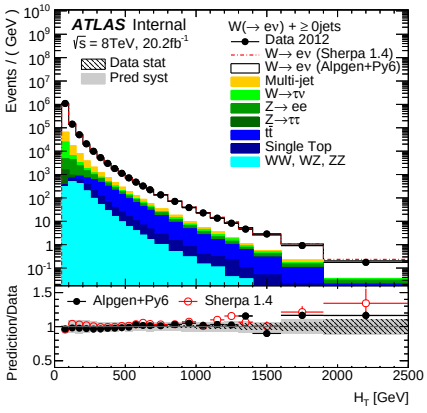
(c) $E_{\mathrm{T}}^{\mathrm{miss}}$



(d) m_{T}

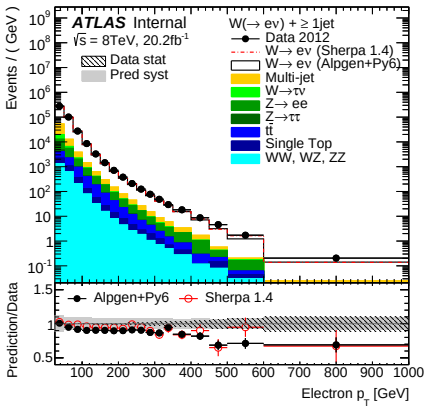


(e) W p_{T}

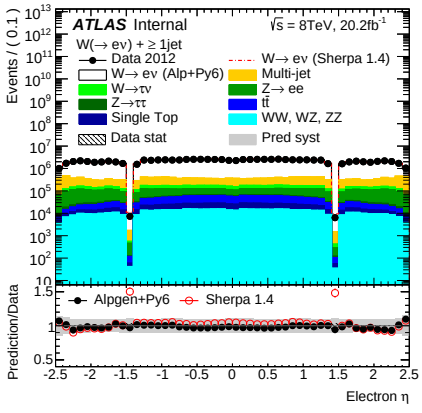


(f) H_{T}

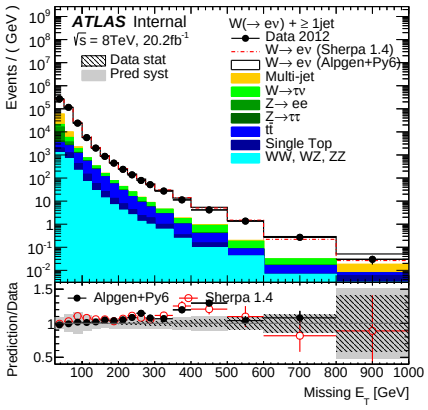
Figure 6.2: Kinematic distributions at the detector level: Electron p_{T}(a) and η(b), $E_{\mathrm{T}}^{\mathrm{miss}}$(c), m_{T}(d), the W p_{T}(e) and H_{T}(f), for any number of jets produced. The lower panels are defined as in Figure 6.1.



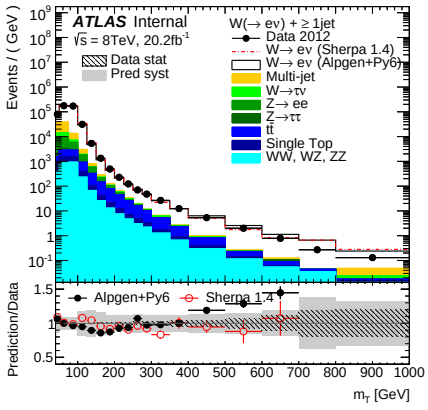
(a) Electron p_{T}



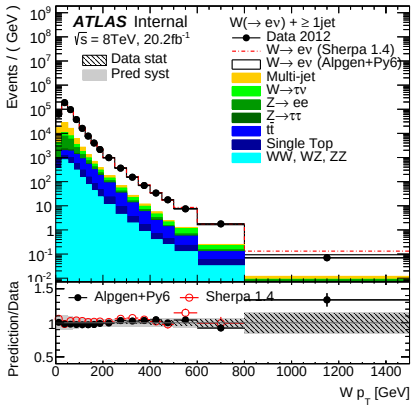
(b) Electron η



(c) $E_{\mathrm{T}}^{\mathrm{miss}}$

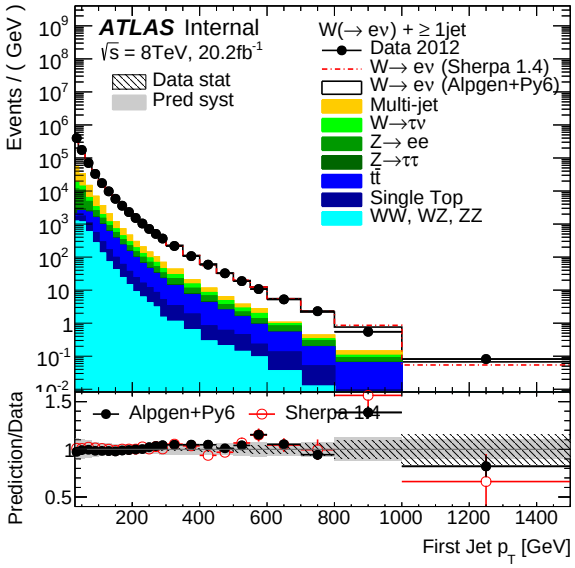


(d) m_{T}

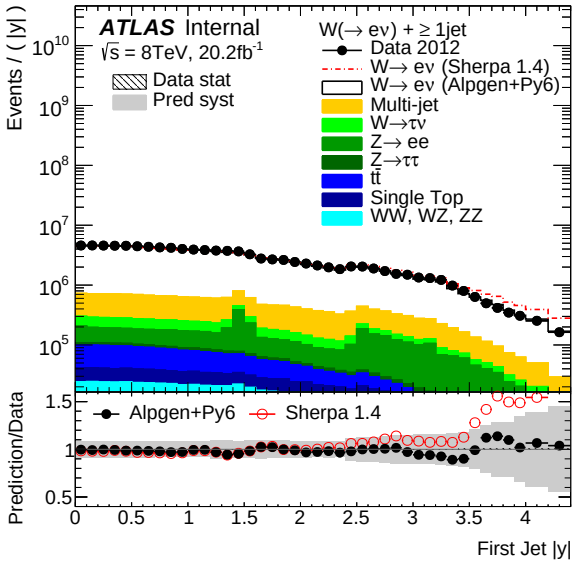


(e) W p_{T}

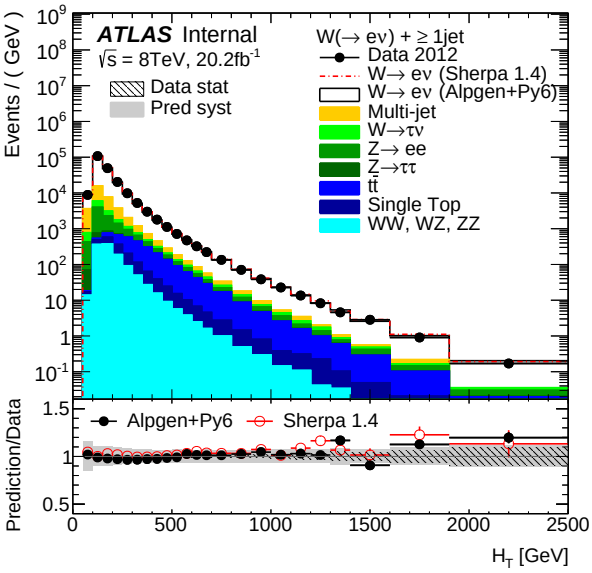
Figure 6.3: W-related kinematic distributions at the detector level: Electron p_{T}(a) and η(b), $E_{\mathrm{T}}^{\mathrm{miss}}$(c), m_{T}(d) and the W p_{T}(e), with at least 1 jet produced in association with the W boson. The lower panels are defined as in Figure 6.1.



(a) Leading jet p_{T}

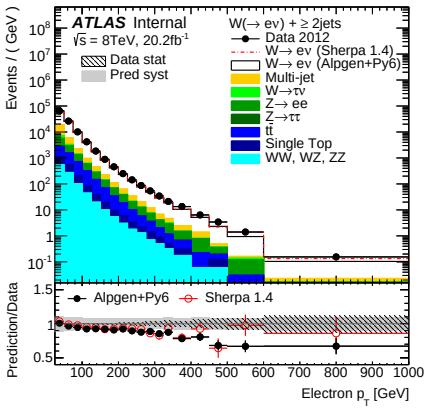


(b) Leading jet $|y|$

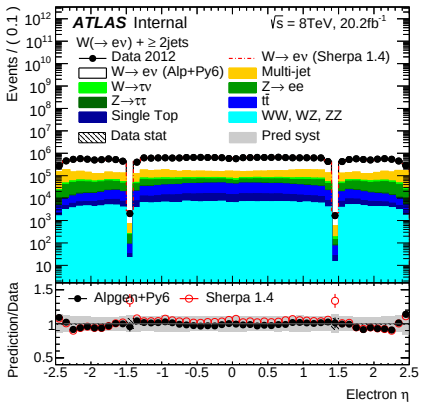


(c) H_{T}

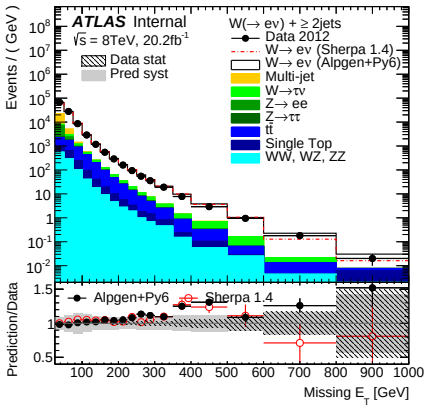
Figure 6.4: Jet-related kinematic distributions at the detector level: Leading jet p_{T}(a) and rapidity y (b), as well as H_{T}(c), and the symmetrized version of the leading jet rapidity, $|y|$, with at least 1 jet produced in association with the W boson. The lower panels are defined as in Figure 6.1.



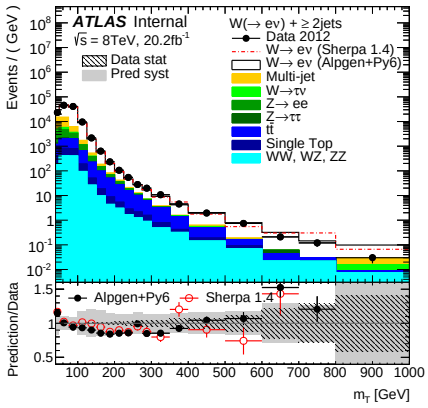
(a) Electron p_{T}



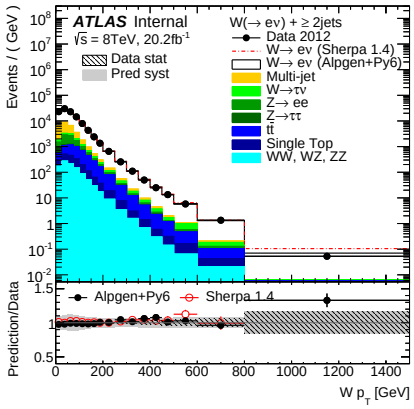
(b) Electron η



(c) $E_{\mathrm{T}}^{\mathrm{miss}}$

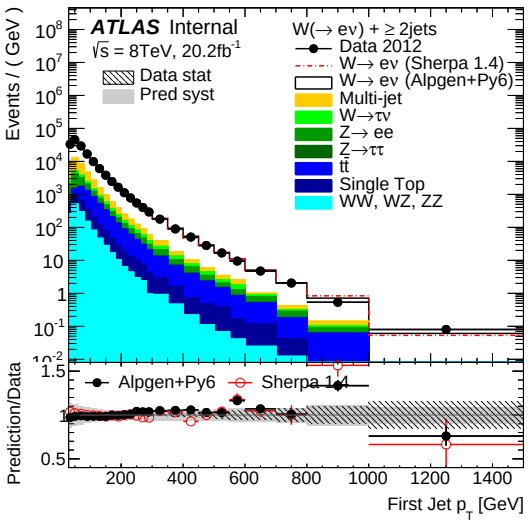


(d) m_{T}

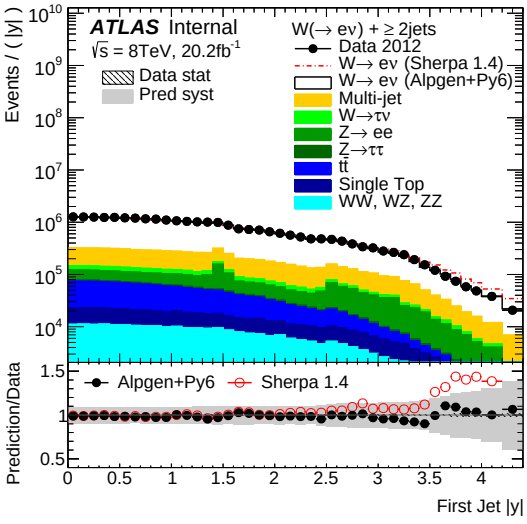


(e) W p_{T}

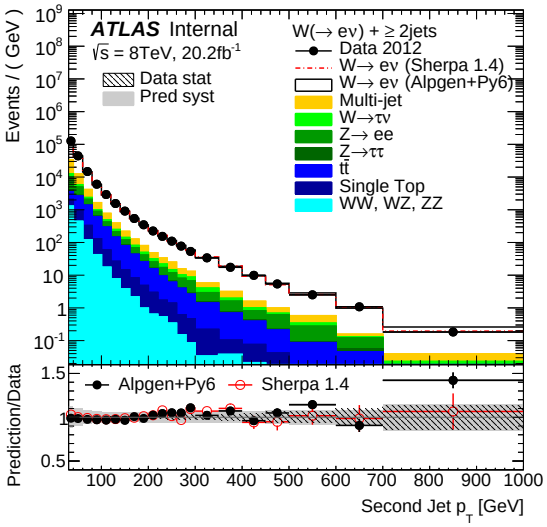
Figure 6.5: W-related kinematic distributions at the detector level: Electron p_{T}(a) and η(b), $E_{\mathrm{T}}^{\mathrm{miss}}$(c), m_{T}(d) and the W p_{T}(e), with at least 2 jets produced in association with the W boson. The lower panels are defined as in Figure 6.1.



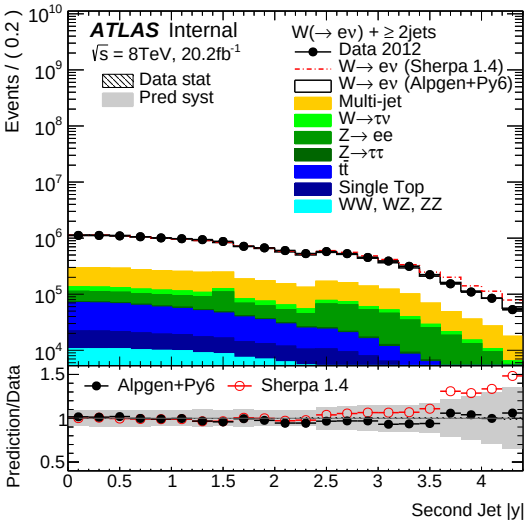
(a) Leading jet p_{T}



(b) Leading jet $|y|$

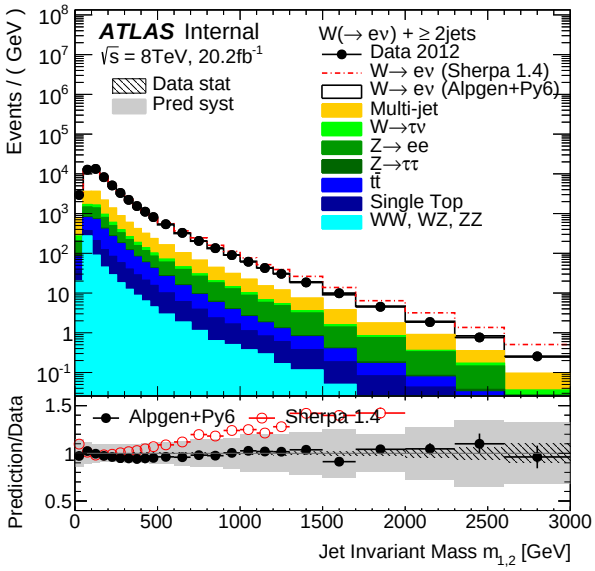


(c) Second leading jet p_{T}

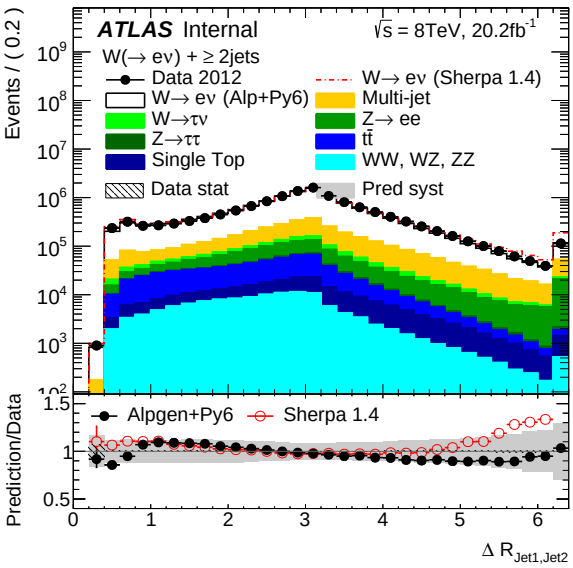


(d) Second leading jet $|y|$

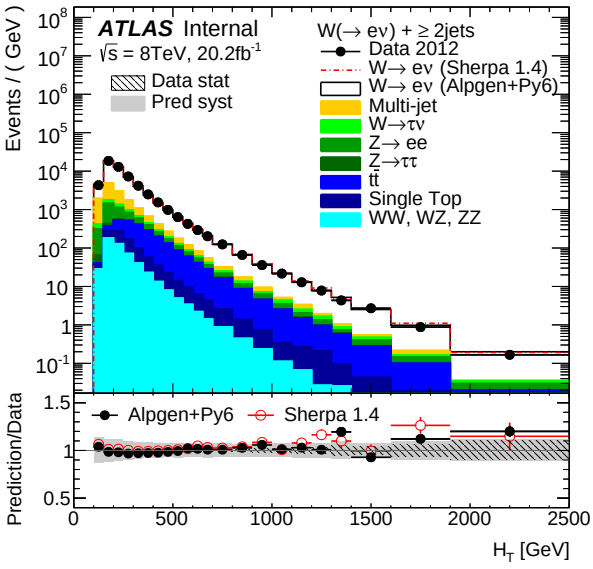
Figure 6.6: Jet-related kinematic distributions at the detector level: Leading jet p_{T}(a) and absolute rapidity $|y|$(b), second leading jet p_{T}(c) and absolute rapidity $|y|$(d), third leading jet p_{T}(e), with at least 2 jets produced in association with the W boson. The lower panels are defined as in Figure 6.1.



(a) Invariant mass $m_{1,2}$



(b) Angular separation $\Delta R_{1,2}$



(c) H_{T}

Figure 6.7: Further jet-related kinematic distributions at the detector level: Invariant mass $m_{1,2}$ (a) and angular separation $\Delta R_{1,2}$ (b) of the two leading jets, as well as H_{T}(c), with at least 2 jets produced in association with the W boson. The lower panels are defined as in Figure 6.1.

The ALPGEN+PYTHIA6 generator, using the CTEQ6L1 PDF set [39], describes the data within uncertainties in most distributions. The SHERPA Monte Carlo however, does not match well with the data at high jet rapidities, at high energies in the dijet invariant mass, or at high angular separation between jets (see Figures 6.4b, 6.7b, and 6.7a respectively). Because of this, ALPGEN is chosen as the main W+jets event generator and is used for unfolding (discussed in Chapter 7), but SHERPA is still used for cross-checks and comparisons. In general, the cross section should not be dependent on detector effects. For this reason, detector level results are corrected to particle level (discussed in Chapter 7), which is independent of the detector effects. These distributions are described in greater depth in Chapter 10. Detector level distributions for the W^+ and W^- selections are shown in Appendix B.

Chapter 7

Unfolding

At this point the detector level results are unfolded to particle level. That is, the effects from detector resolution and efficiencies are corrected to obtain the true kinematic distribution. These distributions are more easily compared to theoretical predictions since detector effects do not modify the observed distributions anymore.

In order to do this, an iterative Bayesian unfolding technique is performed to determine this correction (discussed in Section 7.2). Before this is done however, the phase space for theoretical predictions must first be defined.

7.1 The fiducial region

When applying selection cuts to Monte Carlo events more information is available compared to the data. For example, the analysis can specify that the selected electron comes from the W boson, and not worry about electrons from jets. As such, the following requirements are used to define the phase space for theory predictions, also called the fiducial region. Note that this is similar to, but not the same as, the event selection for data, detailed in Section 5.1.

- The **electron** must have come from the W boson and is required to have $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$. Photons within $\Delta R < 0.1$ of the electron are added to it. This is called

a dressed electron.

- The **neutrino** must also come from the W boson and is identified as $E_{\mathrm{T}}^{\mathrm{miss}}$.

- **Jets** are clustered from all stable final state particles using the anti-k_{T} algorithm with radius parameter $R = 0.4$. Jets can not overlap with an electron within a radius of $\Delta R < 0.2$, and must have $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|y| < 4.4$.

After objects are selected, events are vetoed if any of the following requirements are not satisfied: the event must contain exactly one electron and must contain $E_{\mathrm{T}}^{\mathrm{miss}}$ greater than 25 GeV. The transverse mass of the W must be greater than 40 GeV. There must not be a jet within $\Delta R < 0.4$ of an electron. The fiducial region is summarized in Table 7.1. This can be compared to the event selection used for data in Table 5.1 (Note again that these are not exactly the same). The goal of the unfolding is to correct from the signal region that is at detector level, to this fiducial region. In contrast to the signal region, the fiducial region has no b-jet veto, as well as no electron isolation. Additionally the accepted pseudorapidity range of the electron is extended to 2.5 and has no gap. The pseudorapidity range in the data selection corresponds to the LAr barrel and end-cap geometry, while the truth level selection corresponds to the inner detector. This selection was kept general enough so that it can be easily handled by many different Monte Carlo generators, which may handle particles differently. If the muon channel were to be studied as well, the same selection could apply by simply changing the electron cuts to muon cuts. Confining muons to the LAr geometry would not make sense, since they generally do not interact with the EM calorimeter.

$W \to e\nu$ fiducial region	
Electron	Exactly 1 electron with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $\lvert\eta\rvert < 2.5$
Electron-jet distance	ΔR(selected electron,truth jet) ≥ 0.4
Missing energy	$E_{\mathrm{T}}^{\mathrm{miss}} > 25\,\mathrm{GeV}$
Transverse mass	$m_{\mathrm{T}} > 40\,\mathrm{GeV}$

Table 7.1: Fiducial region for the cross section measurement of $W \to e\nu$+jets

7.2 Unfolding method

The method used to correct from detector level (also referred to as reconstructed level) to particle level (also called truth level) is an iterative Bayesian unfolding technique [40, 41], named such because the algorithm is based on Bayes' theorem. Consider a kinematic variable like H_{T}. It is possible that at the detector level, an event fills a particular bin, but at the particle level there is slightly more (or less) H_{T} and the event fills a different bin. Bayes' theorem states that the probability $P(\mathbf{x_T}|\mathbf{x_D})$ for a true measured value $\mathbf{x_T}$, given a reconstructed or data value $\mathbf{x_D}$ is:

$$P(\mathbf{x_T}|\mathbf{x_D}) \propto P(\mathbf{x_D}|\mathbf{x_T}) \cdot P(\mathbf{x_T}). \tag{7.1}$$

$P(\mathbf{x_D}|\mathbf{x_T})$ is the probability that the value $\mathbf{x_D}$ is reconstructed, when the true value $\mathbf{x_T}$ is generated. The measured values in this case are the discrete bins in the distributions. $P(\mathbf{x_D}|\mathbf{x_T})$ is found from the response matrix discussed below. $P(\mathbf{x_T})$ is the prior, i.e. the assumption for the probability of the true distribution prior to the unfolding. The true distribution, obtained from Monte Carlo, is used as the initial prior.

The estimate of the unfolded distribution $\hat{N}(\mathbf{x_T})$ is simply:

$$\hat{N}(\mathbf{x_T}) \propto \sum_{\mathbf{x_D}} N(\mathbf{x_D}) \cdot P(\mathbf{x_T}|\mathbf{x_D}) \tag{7.2}$$

where $N(\mathbf{x_D})$ is the measured distribution. The measured distribution is of course determined by the background subtracted from the data, $N(\mathbf{x_D}) = N_{\text{data}}(\mathbf{x_D}) - N_{\text{bkgd}}(\mathbf{x_D})$. The sum in this case is a sum over the bins in the distribution.

Before the unfolding is performed, however, a correction must be applied to account for the fact that not all events in the signal region come from the fiducial region. This fiducial correction $c_{\text{fid}}(\mathbf{x_D})$ is applied multiplicatively. Similarly, after unfolding, another correction is applied to account for the reconstruction efficiency $\varepsilon_{\text{reco}}(\mathbf{x_T})$. This takes into account that not all events in the fiducial region are reconstructed in the signal region. Details on these corrections are given below.

The full estimate for the unfolded distribution is then given as:

$$\hat{N}(\mathbf{x_T}) \propto \frac{1}{\varepsilon_{\text{reco}}(\mathbf{x_T})} \sum_{\mathbf{x_D}} \left[\left(N_{\text{data}}(\mathbf{x_D}) - N_{\text{bkgd}}(\mathbf{x_D}) \right) \cdot c_{\text{fid}}(\mathbf{x_D}) \right] \cdot P(\mathbf{x_T}|\mathbf{x_D}). \tag{7.3}$$

The choice of a prior can have a varying influence on the unfolding. In order to reduce the influence of the prior, the unfolding can be iterated; the unfolded distribution of one iteration can be used as the prior in the next iteration. However, each iteration increases the statistical uncertainty, so these effects must be balanced. In this analysis, the use of 2 iterations was found to optimize these two effects.

7.2.1 Unfolding matrices

The response matrices (also called migration matrices) represent the relation between events at the reconstructed level and truth level. Often, an event that fills a particular bin at the reconstructed level, can fill a different bin at truth level. These matrices can be found by generating Monte Carlo at truth and reconstructed level and then plotting a 2D histogram

with one axis being the reconstructed bins and the other being the truth level bins. The off-diagonal terms represent the migrations between bins. Events that populate these matrices must pass both the signal region cuts from Section 5.1, as well as those defining the fiducial region in Section 7.1. A perfectly diagonal histogram would correspond to there being no correction from unfolding, truth level and reconstructed level distributions would be identical. These bin migrations can occur across bins of a particular kinematic variable, like H_{T} or rapidity, but they can also occur across jet multiplicity. This is to say that an event may have a different number of jets at reconstructed level and particle level. This is important since distributions are shown for particular jet multiplicities. For this reason the migration matrices must therefore be 4-dimensional, with two axes containing bins of the kinematic variable and the other two axes containing bins of jet multiplicity. For example, the migration matrix for H_{T} contains one axis for each: the number of jets reconstructed $N_{\mathrm{jets}}^{\mathrm{reco}}$, the number of jets at truth level $N_{\mathrm{jets}}^{\mathrm{truth}}$, the H_{T} reconstructed $H_{\mathrm{T}}^{\mathrm{reco}}$, and the H_{T} at truth level $H_{\mathrm{T}}^{\mathrm{truth}}$. The exception to this is of course jet multiplicity itself, which is 2-dimensional.

The unfolding matrices used in this analysis are generated from Alpgen since it agrees better with data than Sherpa. The unfolding matrix for the exclusive jet multiplicity distribution is shown as an example in Figure 7.1. Here the x-axis shows the reconstructed level jet multiplicity and the y-axis shows the same at truth level. The z-axis, represented by color, shows the number of Monte Carlo generated events in each bin, which will be normalized before being applied to Equation 7.3.

Figure 7.2 shows similar three-dimensional lego plots for several observables at their most relevant jet multiplicities. These plots show the diagonal terms $N_{\mathrm{jets}}^{\mathrm{truth}} = N_{\mathrm{jets}}^{\mathrm{reco}} = 1$ in Figure 7.2a and 7.2d, and the same for the 2 jet multiplicity terms in Figure 7.2b. The x-axes now show the reconstructed observable and the y-axes show the observable at truth level.

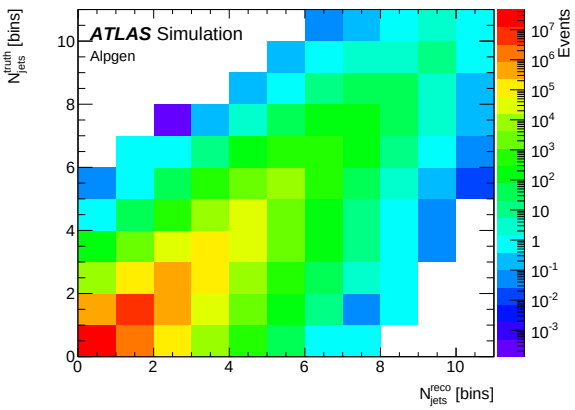


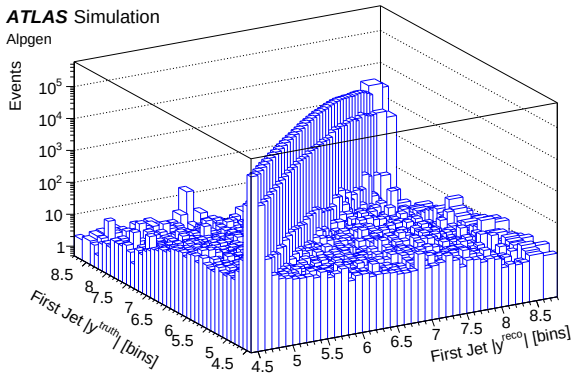
Figure 7.1: The migration matrix for the exclusive jet multiplicity N_{jets} determined from ALPGEN used in the unfolding of the 2012 data.

7.2.2 Reconstruction efficiency

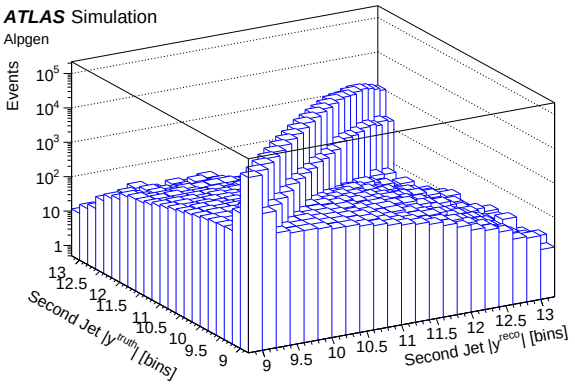
The reconstruction efficiency is defined as the fraction of events in the fiducial phase space (FR) which also pass the signal region (SR) requirements at reconstruction level.

$$\varepsilon_{\text{reco}} = \frac{N \text{ events passing both SR and FR cuts}}{N \text{ events passing FR cuts}} \tag{7.4}$$

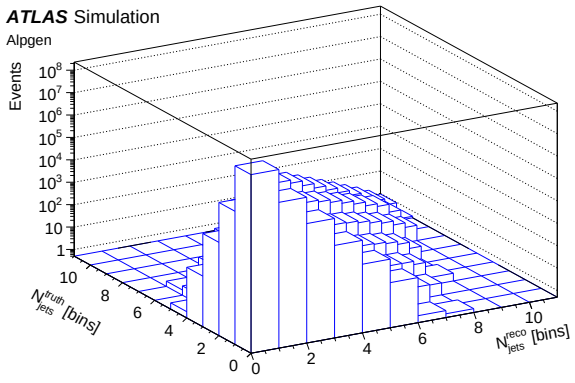
The efficiency is a function of the truth level observable and the jet multiplicity. The corrections were done with ALPGEN using SHERPA as a cross-check. The difference is considered as a systematic uncertainty in the unfolding. Figure 7.3b shows the reconstruction efficiency for inclusive and exclusive jet multiplicity, and Figures 7.4-7.5 show the efficiencies for several observables at their most relevant jet multiplicities. In general the efficiency is around 50%. The main loss in efficiency comes from the electron identification *tight* requirement in the signal region. About 25% of events from the fiducial region are rejected in the reconstruction due to this cut. This is expected based on efficiency studies done by ATLAS [1]. About 14% of the efficiency is lost due to the electron isolation and impact parameter cuts. 5% loss comes from the tightened $|\eta|$ requirement on the electron, and 7% from the reconstructed



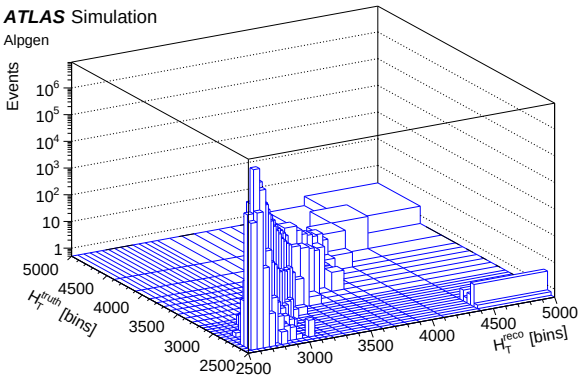
(a) Leading jet rapidity



(b) Second leading jet rapidity



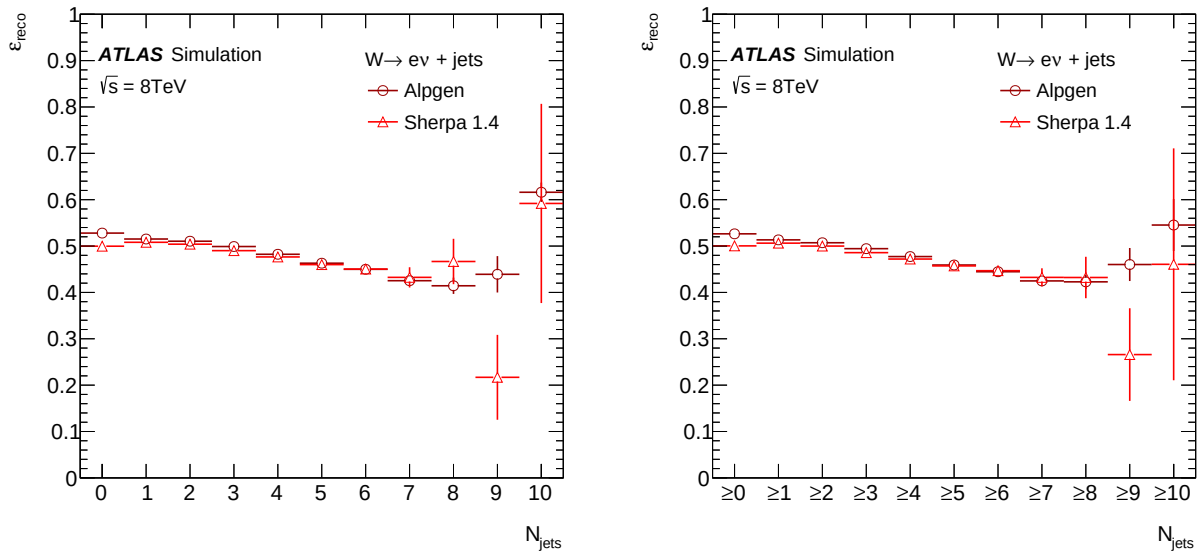
(c) N_{jets}



(d) H_{T}, 1jet

Figure 7.2: Extracts of the migration matrices for the leading and second leading absolute jet rapidities (a) and (b) with $N_{\text{jets}} = 1$ or $N_{\text{jets}} = 2$, respectively, and for the N_{jets} distribution (c) and H_{T} with $N_{\text{jets}} = 1$ (d). The matrices are drawn as three dimensional lego plots to give a better impression of the event number difference on and off the diagonal.

$E_{\mathrm{T}}^{\mathrm{miss}}$. These are the most significant losses. The remaining few percent come from other signal requirements.



(a) Exclusive N_{jets}

(b) Inclusive N_{jets}

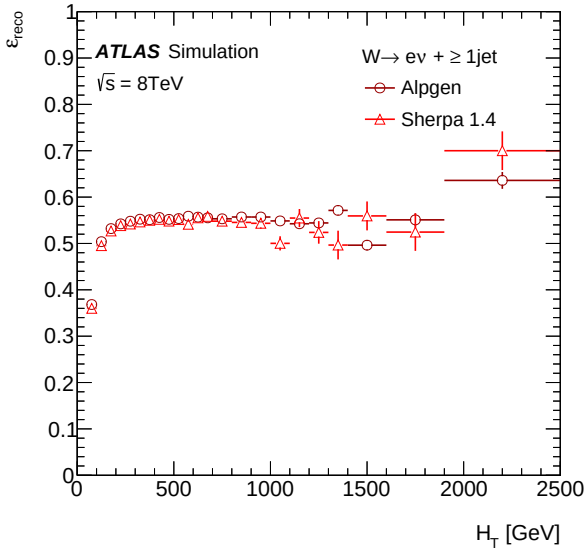
Figure 7.3: Reconstruction efficiency $\varepsilon_{\mathrm{reco}}$ for the exclusive and inclusive jet multiplicity (a) and (b). Predictions both from ALPGEN and SHERPA are shown. The values on the x-axis correspond to the values of the respective observable at the truth level.

7.2.3 The fiducial correction

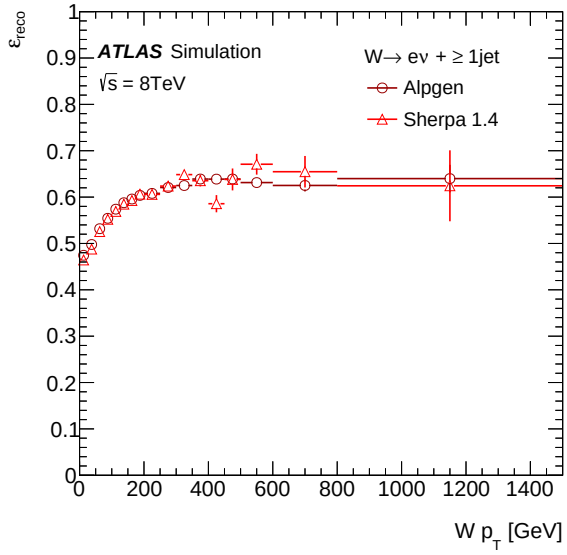
The fiducial correction is defined as the fraction of reconstructed events that come from the fiducial phase space.

$$c_{\mathrm{fid}} = \frac{N \text{ events passing both SR and FR cuts}}{N \text{ events passing SR cuts}} \tag{7.5}$$

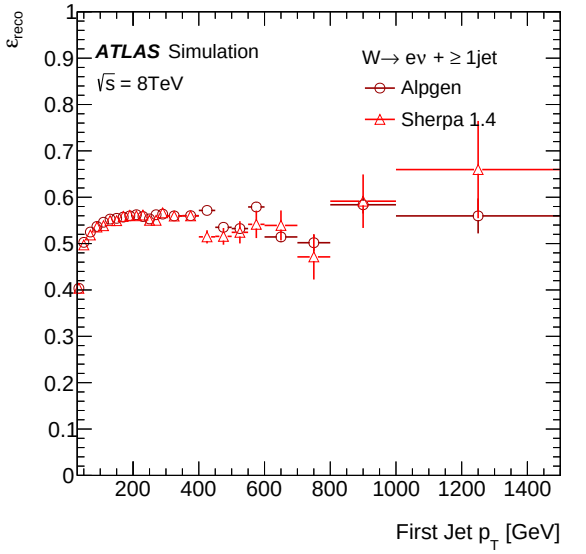
Since the fiducial region has generally looser cuts, the expectation would be that this fraction is 1. A fiducial correction below 1 implies that events from outside the fiducial region are able to pass the signal region cuts. These are often referred to as fakes. The fraction of fakes



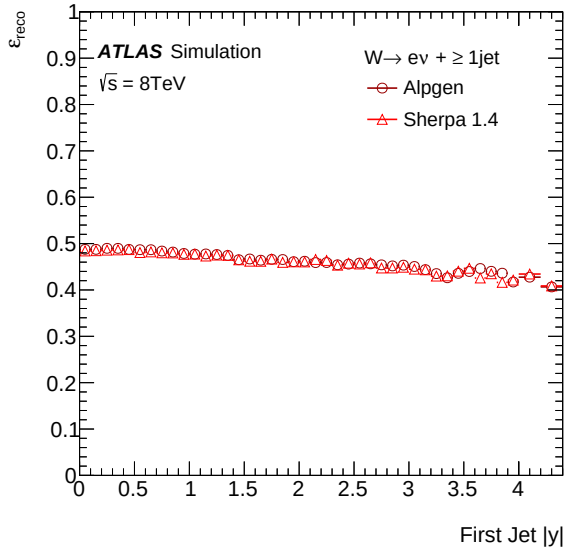
(a) H_{T}



(b) Wp_{T}

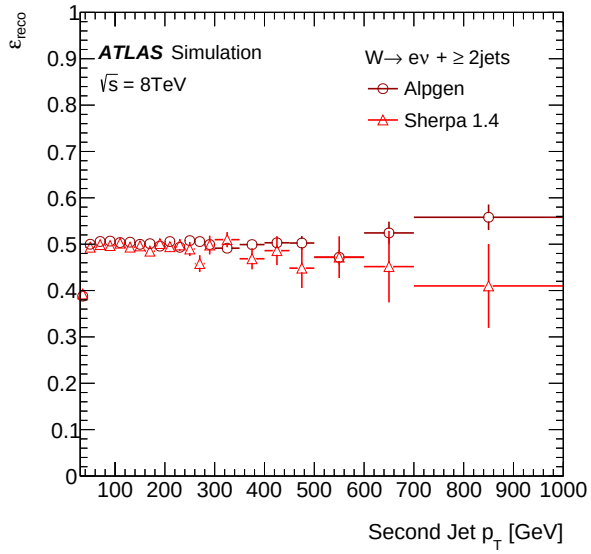


(c) Leading jet p_{T}

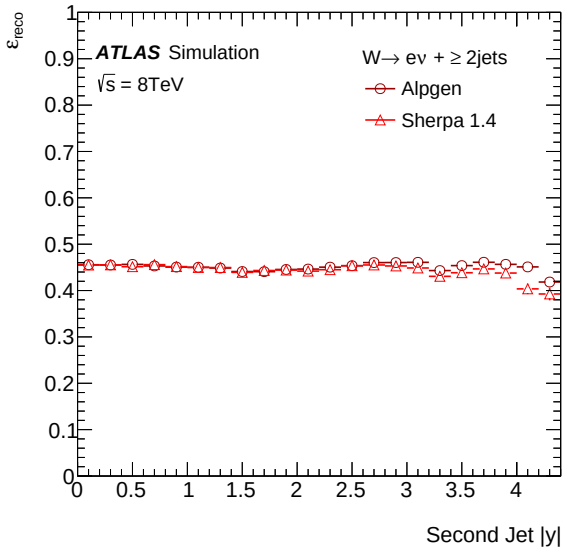


(d) Leading jet rapidity

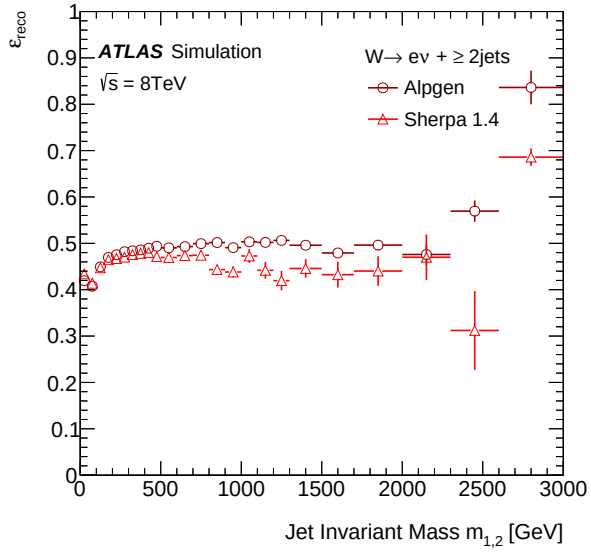
Figure 7.4: Reconstruction efficiency $\varepsilon_{\mathrm{reco}}$ for H_{T} (a), W boson p_{T} (b), and leading jet p_{T} and absolute rapidity $|y|$ (c) and (d). Predictions both from ALPGEN and SHERPA are shown. $\varepsilon_{\mathrm{reco}}$ is given for the multiplicity $N_{\mathrm{jets}} \geq 1$. The values on the x-axis correspond to the values of the respective observable at the truth level.



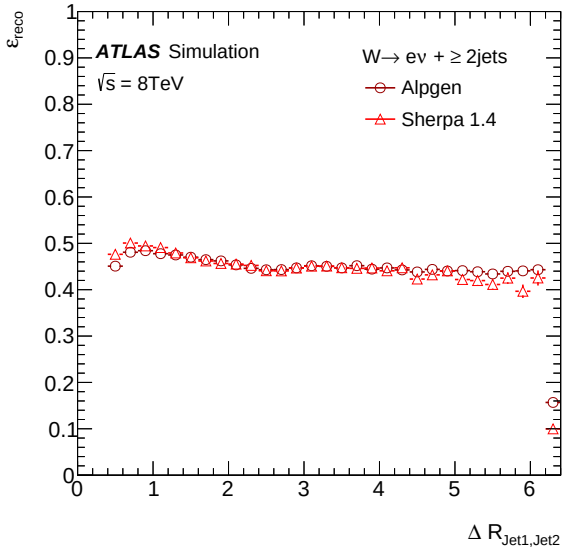
(a) Second leading jet p_{T}



(b) Second leading jet $|y|$



(c) Invariant mass $m_{1,2}$

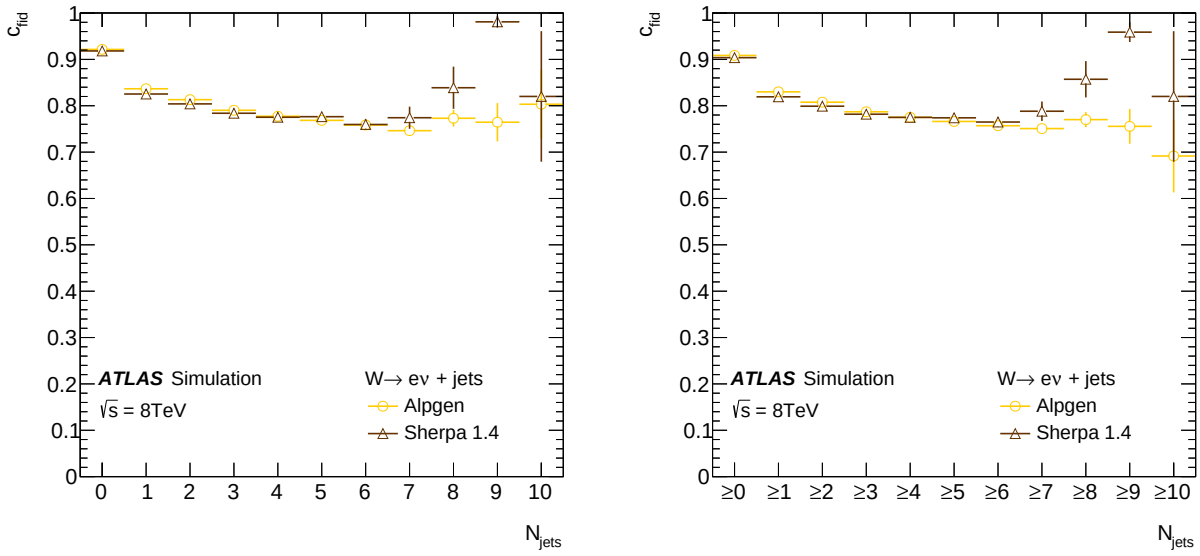


(d) Angular separation $\Delta R_{1,2}$

Figure 7.5: Reconstruction efficiency $\varepsilon_{\mathrm{reco}}$ for the second leading jet p_{T} and absolute rapidity $|y|$ (a) and (b), and the invariant mass $m_{1,2}$ and angular separation $\Delta R_{1,2}$ of the leading two jets (c) and (d). Predictions both from ALPGEN and SHERPA are shown. $\varepsilon_{\mathrm{reco}}$ is given for the multiplicity $N_{\mathrm{jets}} \geq 2$. The values on the x-axis correspond to the values of the respective observable at the truth level.

f_{fake} is related to the fiducial correction by $c_{\text{fid}} = 1 - f_{\text{fake}}$.

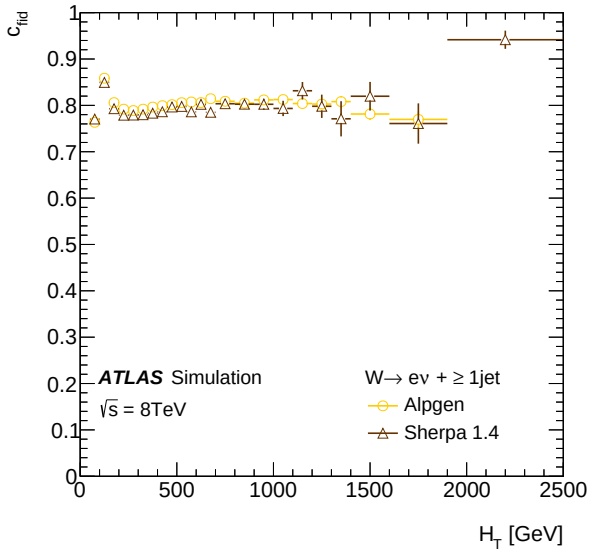
Analogous to the reconstruction efficiency, the fiducial correction is dependent on the bin of the reconstructed observable and jet multiplicity. Again the correction is done with ALPGEN using SHERPA as a cross-check. Figure 7.6 shows the fiducial correction for the jet multiplicity distributions. Figures 7.7-7.8 show the corrections to some of the observables at their most relevant jet multiplicities.



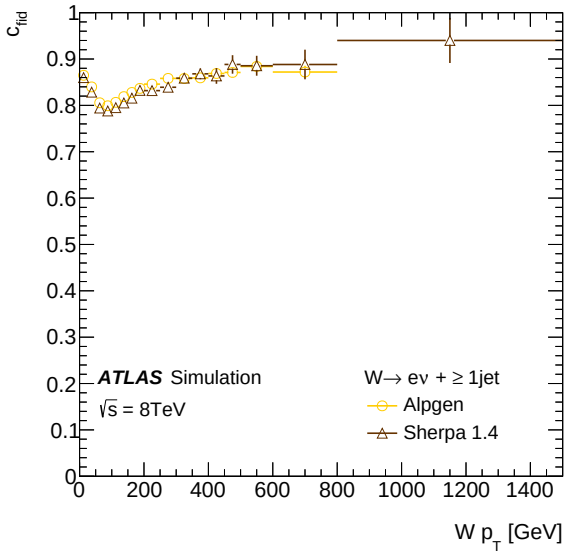
(a) Exclusive N_{jets} (b) Inclusive N_{jets}

Figure 7.6: Fiducial correction c_{fid} for the exclusive and inclusive jet multiplicity (a) and (b). Predictions both from ALPGEN and SHERPA are shown. The values on the x-axis correspond to the values of the respective observable at the reconstructed level.

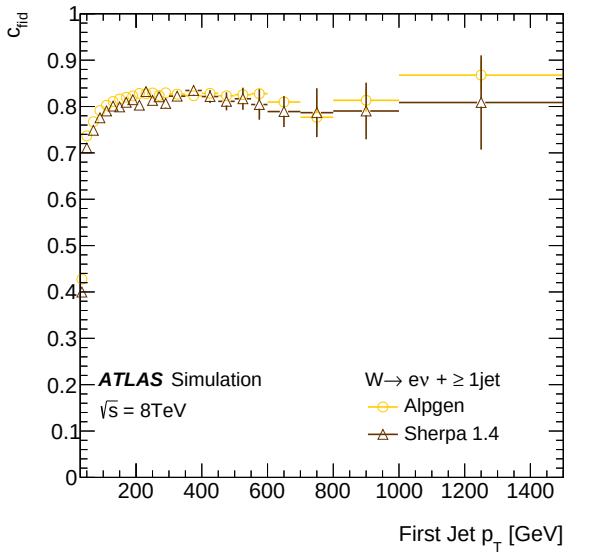
The corrections are generally high, greater than 0.75, except in the case where two jets are involved. The fraction of reconstructed events drops low, below 0.5, especially in the region of high jet rapidity where there are a lot of pile-up jets. This indicates there is a large contribution from fake jets, likely from pile-up jets which are not present in true W + jets events. The jump at $|\eta| = 2.5$ in Figures 7.7d and 7.8b, corresponds to the forward region where tracking information is unavailable. This is likely a factor in the sharp increase of fake



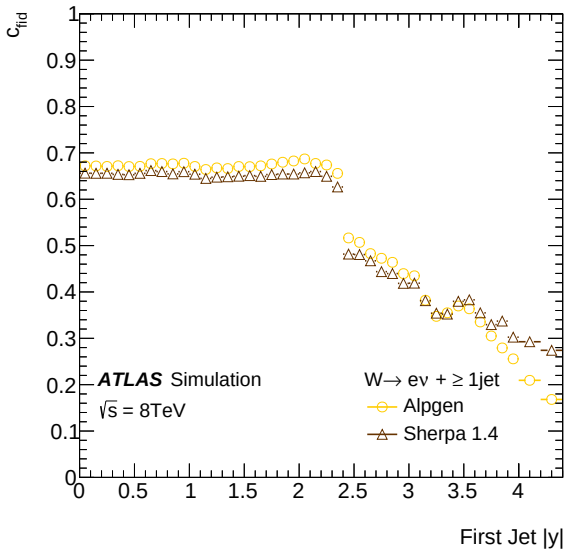
(a) H_{T}



(b) Wp_{T}

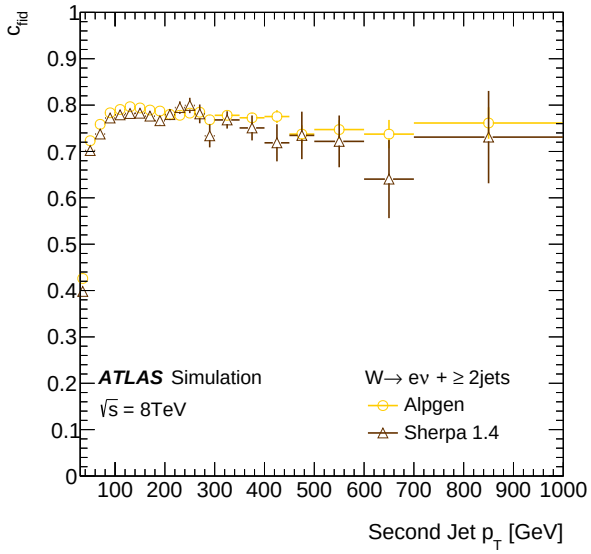


(c) Leading jet p_{T}

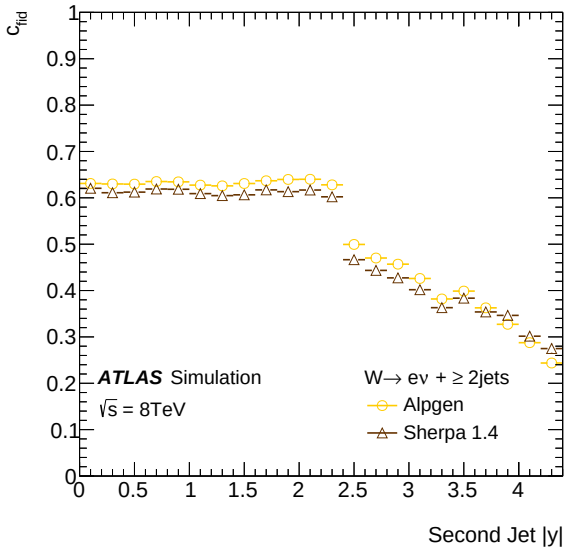


(d) Leading jet rapidity

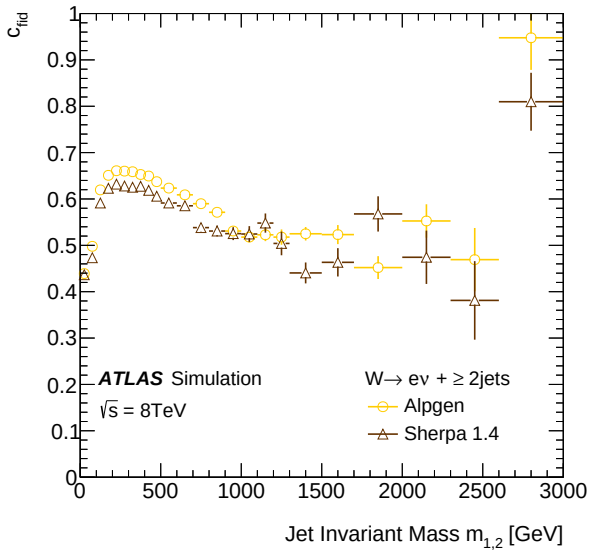
Figure 7.7: Fiducial correction c_{fid} for H_{T} (a), W boson p_{T} (b), and leading jet p_{T} and absolute rapidity $|y|$ (c) and (d). Predictions both from ALPGEN and SHERPA are shown. c_{fid} is given for the multiplicity $N_{\mathrm{jets}} \geq 1$. The values on the x-axis correspond to the values of the respective observable at the reconstructed level.



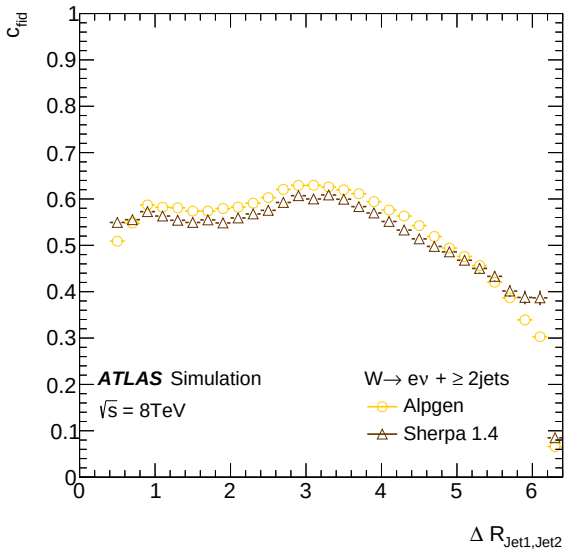
(a) Second leading jet p_{T}



(b) Second leading jet $|y|$



(c) Invariant mass $m_{1,2}$



(d) Angular separation $\Delta R_{1,2}$

Figure 7.8: Fiducial correction c_{fid} for the second leading jet p_{T} and absolute rapidity $|y|$ (a) and (b), and the invariant mass $m_{1,2}$ and angular separation $\Delta R_{1,2}$ of the leading two jets (c) and (d). Predictions both from ALPGEN and SHERPA are shown. c_{fid} is given for the multiplicity $N_{\mathrm{jets}} \geq 2$. The values on the x-axis correspond to the values of the respective observable at the reconstructed level.

events.

7.3 Validation

Two major tests are performed to validate that the unfolding is done properly and they are outlined here.

7.3.1 Closure test

Unfolding is performed on the same Monte Carlo that the inputs for unfolding are taken from; in this case, ALPGEN. Since the migration matrices, reconstruction efficiencies, and fiducial corrections were all generated from these events, the result should be exactly the particle level prediction from the same generator.

This closure test was performed on all unfolded distributions and the difference was found to be exactly zero in all distributions except for the angular separation between the lead two jets. The difference of a few percent occurs in the highest and lowest ΔR bins. This is likely due to poor numerical conversion during phase-wrapping of separations greater than 2π (ΔR can be larger than 2π since it is dependent on $\Delta\eta$ which is not an angle) This difference is mitigated by the jet radius being 0.4 as well as large uncertainties at high angular separation. When the distribution is projected down to just the inclusive 2 jet multiplicity, the difference is negligible (shown in Figure 7.9). This figure also shows that the statistical uncertainty increases due to the process of unfolding. This behavior is expected since it is propagated.

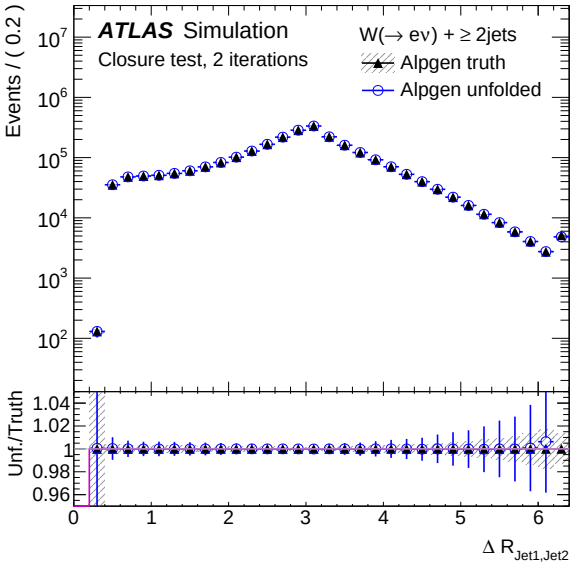


Figure 7.9: Closure test for the angular separation $\Delta R_{1,2}$ of the leading two jets $N_{\text{jets}} \geq 2$. The true ALPGEN distribution for the corresponding jet multiplicity is compared to the unfolded ALPGEN distribution. The MC inputs in the unfolding are taken from ALPGEN as well. The lower panels show the ratios of the unfolded to the true distribution. The hashed band indicates the fractional uncertainty in the true prediction, the error bars the fractional uncertainty in the unfolded result.

7.3.2 Bias in unfolding

This analysis uses ALPGEN to construct all the relevant corrections in the unfolding procedure, that is: the reconstruction efficiency, the fiducial correction, and the migration matrices. This introduces a potential bias in these corrections based on how well ALPGEN models reconstruction and truth level events. If SHERPA is used, these corrections may be slightly different.

The differences in reconstruction efficiency and fiducial correction between ALPGEN and SHERPA have already been studied in Section 7.2.2 and 7.2.3, and provide one source of uncertainty in the unfolding. It has also been discussed that the influence of the prior falls off quickly with the number of iterations. The impact on the migration matrix, must still be discussed, however.

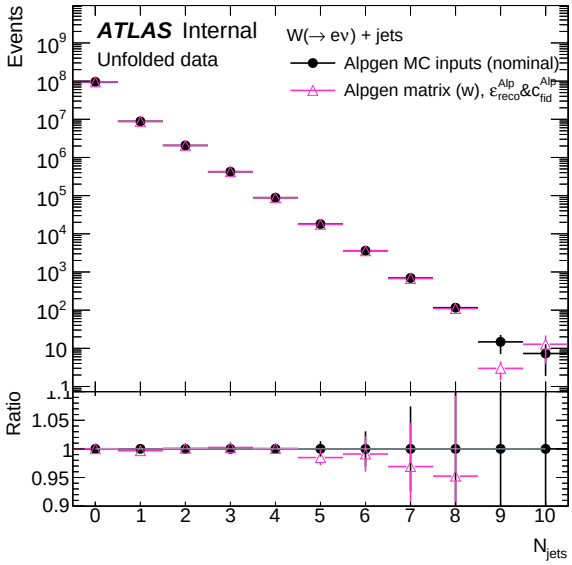
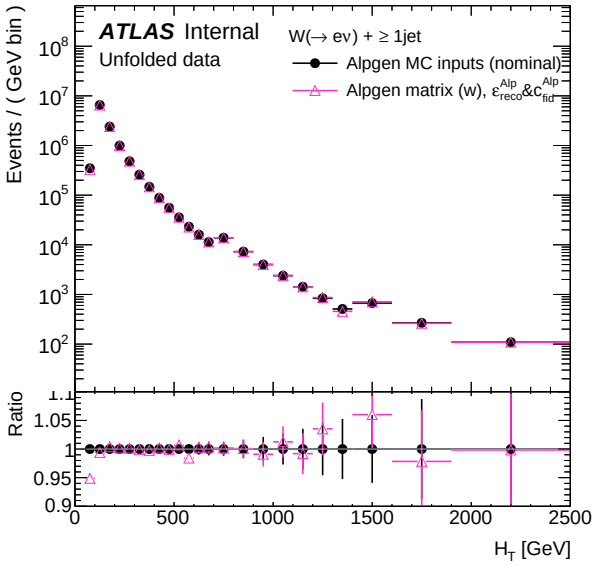
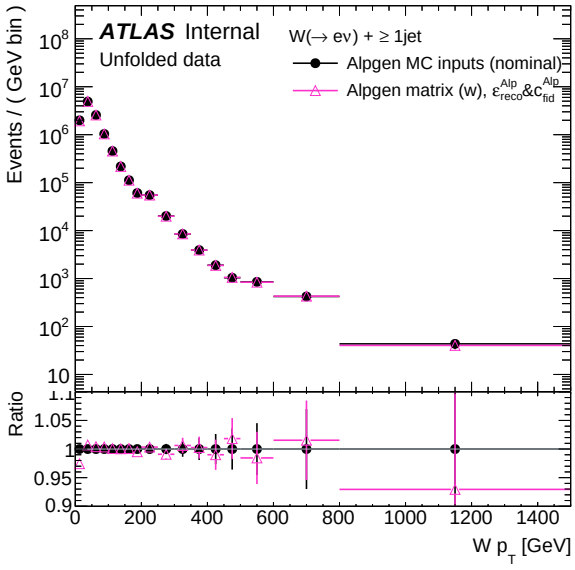


Figure 7.10: Result of the unfolding of data as a function of N_{jets}, using either the nominal ALPGEN matrix (black) or the ALPGEN matrix reweighted for better agreement with data at detector level (pink). The lower panel gives the ratio w.r.t. the nominal data unfolding, with the errors indicating the statistical uncertainty in the unfolded distribution.

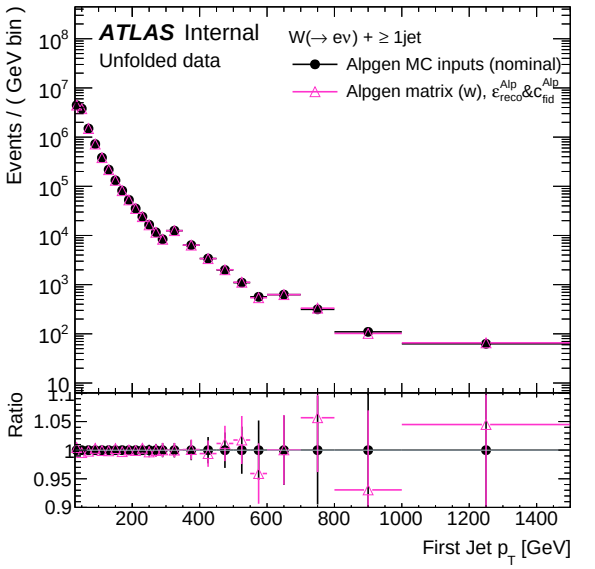
Ultimately, the bias is evaluated by reweighting the migration matrix such that the projection of the reconstructed distribution agrees better with data. These weights are applied to the truth axis of the migration matrix. The weights used to achieve this, are the ratio of data to Alpgen at the detector level. Figures 7.10 - 7.12 show the difference between the unfolding done with the reweighted migration matrices and the nominal unfolding. The differences are generally small. Note that most discontinuities in the slope are due to the change in bin width. Statistically significant differences are considered as systematic uncertainties, but their impact is close to negligible.



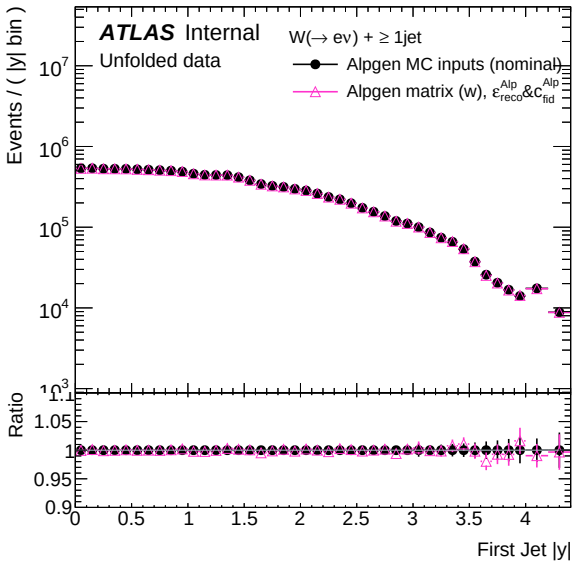
(a) H_{T}



(b) W p_{T}

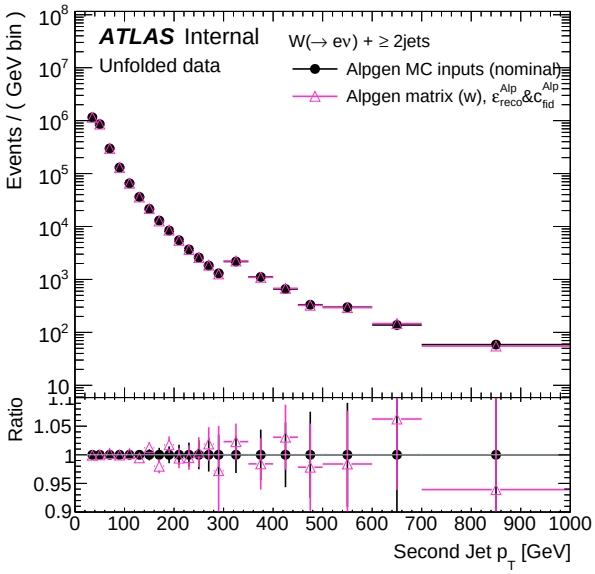


(c) Leading jet p_{T}

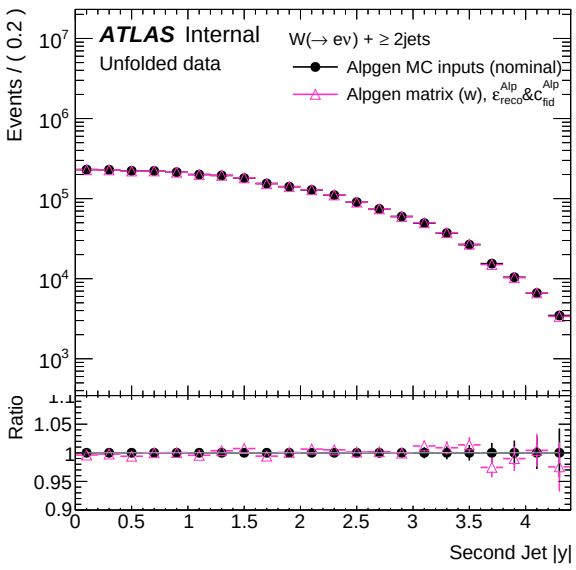


(d) Leading jet $|y|$

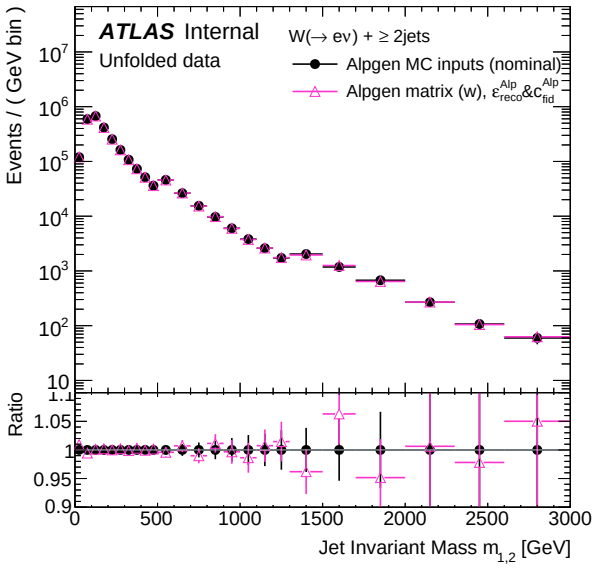
Figure 7.11: Result of the unfolding of data as a function of H_{T} (a), W p_{T} (b), leading jet p_{T} (c) and leading jet absolute rapidity (d) in the presence of at least 1 jet, using either the nominal ALPGEN matrix (black) or the ALPGEN matrix reweighted for better agreement with data at detector level (pink). The lower panel gives the ratio w.r.t. the nominal data unfolding, with the errors indicating the statistical uncertainty in the unfolded distribution.



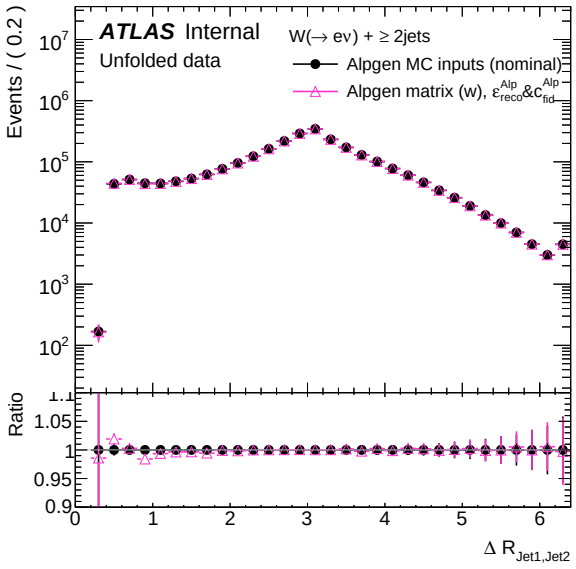
(a) Second leading jet p_{T}



(b) Second leading jet $|y|$



(c) Invariant mass $m_{1,2}$



(d) Angular separation $\Delta R_{1,2}$

Figure 7.12: Result of the unfolding of data as a function of the second leading jet p_{T} (a) and absolute rapidity (b), and the invariant mass $m_{1,2}$ (c) and angular separation $\Delta R_{1,2}$ of the leading two jets (d) in the presence of at least two jets, using either the nominal ALPGEN matrix (black) or the ALPGEN matrix reweighted for better agreement with data at detector level (pink). The lower panel gives the ratio w.r.t. the nominal data unfolding, with the errors indicating the statistical uncertainty in the unfolded distribution.

Chapter 8

Systematic Uncertainties

There are systematic uncertainties that affect distributions starting at the reconstruction level and those that arise afterwards during unfolding. This chapter outlines the systematics considered in this analysis, how they are estimated, and how they are propagated through to the final cross sections.

8.1 Detector level systematic uncertainties

There are systematic uncertainties associated in the reconstruction of all physics objects in the detector including jets, electrons, and $E_{\mathrm{T}}^{\mathrm{miss}}$. Additionally, there are uncertainties resulting from the reweighting of Monte Carlo in order to match the average number of interactions per bunch crossing $\langle\mu\rangle$, and from the estimates of backgrounds. To determine these uncertainties, calibrations are varied in Monte Carlo according to the prescription of the specific uncertainty. These variations are propagated through the data-driven multi-jet estimate (Section 5.2) and unfolding procedures (Chapter 7) and the difference in the result from nominal is the systematic uncertainty due to the specified source. The upwards and downwards uncertainties from each source are propagated through individually, but the final results are averaged and taken as a symmetric uncertainty.

$$\Delta_{\mathrm{sys}} = \frac{1}{2}\left(\Delta_{\mathrm{sys,up}} + \Delta_{\mathrm{sys,down}}\right) \tag{8.1}$$

In doing so this helps reduce statistical fluctuations in the uncertainties. The total systematic uncertainty is simply the individual contributions summed in quadrature.

8.1.1 Jet uncertainties

Jet systematic uncertainties are split up into those associated with the jet energy scale (JES), the jet energy resolution (JER), and other jet uncertainties.

8.1.1.1 Jet energy scale uncertainty

The jet energy scale uncertainty is the uncertainty associated with how well-understood the calorimeter's response is to a particle of a given energy level. This uncertainty is dependent on transverse momentum and pseudorapidity. There are originally 68 different individual sources that comprise the JES uncertainty deriving from sources such as: in-situ measurements of the JES, pseudorapidity calibrations, pile-up corrections, flavor response, punch-through, and more. In order to reduce the number of uncertainties considered, the covariance matrix of many of these uncertainties can be diagonalized. The most impactful eigenvectors are kept and the rest are discarded. This reduces the 56 uncertainties related to the in-situ measurements of the JES, to 6 uncertainties. For more information see Reference [42]. There are 12 remaining uncertainties that cannot be diagonalized. In this way the number of relevant sources is reduced to 18, while losing very little information about the uncertainties. For the total JES uncertainty, the differences determined from these variations are added in quadrature. This has been done according to the recommendation of an internal ATLAS study [43].

8.1.1.2 Jet energy resolution uncertainty

The jet energy resolution uncertainty is the uncertainty in how precisely the energy of a jet is measured. This is estimated through in-situ measurements, and includes estimates for electronic noise, and pile-up. The procedure for propagating the systematic uncertainty comes from the ATLAS recommendation for 2012 data [44]. Jets from Monte Carlo are smeared by $+1\sigma$ of the estimated JER uncertainty. That is, the transverse momentum of each jet is varied by a positive factor drawn randomly from a Gaussian with σ equal to the JER uncertainty.

To limit the statistical uncertainty, each jet is smeared several times per event. That is, for each smearing, the event was counted but with the randomly smeared jet p_{T}. To correct for the double counting of events, the distribution is divided by the number of smearings per jet. The uncertainty is then the difference in the cross section between the smeared events, and the nominal. The JER uncertainty was investigated with 1, 10, and 50 smearings per jet. The difference was found to be minimal. To balance the statistical impact and run-time, the systematic uncertainty for the final cross sections was calculated with 10 smearings.

8.1.1.3 Other jet uncertainties

There are two other types of uncertainties associated with jets. The first is the uncertainty that stems from the jet vertex fraction cut (see Section 4.4.4). The systematic uncertainty for this cut is found by varying the cut as a function of jet p_{T} and η according to the recommendation from Reference [45].

The second uncertainty comes from the scale factors applied as a correction to the b-jet veto (see Section 4.4.6). b-tagging, c-tagging, and light-jet mistag scale factors are applied to correct the Monte Carlo predictions to data in the signal region where the b-jet veto

is applied, as well as in the $t\bar{t}$ validation region where a b-jet is required. The effect of these scale factors is estimated by varying these three scale factors independently by their uncertainties. These scale factors are only derived in a limited energy range up to a certain p_{T}, but are extrapolated to higher energies. This extrapolation introduces an additional uncertainty which is varied independently. These scale factors, their uncertainties, and the extrapolation and its uncertainty are all taken from Reference [46]. The total uncertainty from these scale factors and their extrapolation is their sum in quadrature.

8.1.2 Electron uncertainties

Similarly as for jets, there are systematic uncertainties for the electron energy scale and electron energy resolution. Additionally, electrons have scale factor corrections to correct for the efficiencies of reconstruction, identification, triggering, isolation cuts, and charge misidentification.

8.1.2.1 Electron energy scale and resolution uncertainties

The electron energy scale uncertainty comes from three components. The first is the uncertainty in the energy scale measurement which is based on $Z \to ee$ events. Second is uncertainty in the calibration which is dependent on the calorimeter presampler energy. Last is the uncertainty from calibration which is dependent on the ratio of the calorimeter energy in layer-1 to layer-2. This last calibration is intended to help correct for the detector material in front of the calorimeter. These three components are varied separately and then the outcomes are added in quadrature.

To estimate the uncertainty of the electron energy resolution, the width of the Gaussian used for electron energy smearing in Monte Carlo is varied. These uncertainties are taken from

and are detailed in Reference [47].

8.1.2.2 Other electron uncertainties

There are scale factors used to correct for the efficiencies of electron reconstruction, identification, triggering and isolation in Monte Carlo to values measured in data. To find the systematic uncertainty in these scale factors, the scale factors themselves are varied by their uncertainties to measure their effects on the W + jets measurement. The scale factors for reconstruction and identification and their uncertainties are taken from the ATLAS electron efficiency study detailed in Reference [1]. The scale factors for the electron triggers and electron isolation are taken from Reference [48].

The uncertainty due to misidentifying the electron charge is only relevant to the W^+ and W^- cross sections but is found the same way by varying the corresponding scale factors by their uncertainties and re-evaluating the W^+ and W^- cross sections.

8.1.3 Missing transverse momentum uncertainties

Recall from Section 4.5 that the $E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed based on the energies of the hard jets and electrons in the event, as well as the energy not associated with reconstructed objects (soft terms). The $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ can be divided into a hard component (the part derived from the vector sum of the hard jets and electrons), and soft component (the part derived from the soft terms). The uncertainty from the hard component is contained within the jet and electron uncertainties described above. The uncertainties from the soft terms must still be considered however.

The soft component is split into subcomponents parallel to the hard component and transverse. The advantage of doing this is that the parallel term is sensitive to both the energy

scale and resolution of the soft term, but the transverse term is only affected by the energy resolution. To find the systematic uncertainty of the soft term energy resolution, the resolution is smeared once for the parallel and transverse soft term components separately, and once simultaneously. For almost all distributions, the sum in quadrature of the individual smearings was larger than the simultaneous smearing. So the sum in quadrature was taken as the systematic uncertainty for the soft term energy resolution. These uncertainties are derived from a $Z \to \mu\mu$ study, detailed in Reference [49]. The soft term energy scale is varied in Monte Carlo by its uncertainty in order to obtain the systematic uncertainty on the final cross section.

8.1.4 Pile-up rescaling uncertainties

In order to best account for the underlying event, the average number of interactions $\langle\mu\rangle$ is scaled by a factor of 1.09 in Monte Carlo to better match data. The systematic uncertainty is found by varying this factor by ± 0.04, and repeating the full analysis. The scale factor and its uncertainty are taken from the ATLAS recommendation for 2012 data and Monte Carlo [50].

8.1.5 Background uncertainties

The background systematic uncertainties are broken up into the same backgrounds discussed in Chapter 5.

8.1.5.1 Multi-jet uncertainties

See Section 5.2 for the description of how the multi-jet estimate was performed. The following systematics uncertainties are considered for the multi-jet background:

- The statistical uncertainty of the fit itself.

- The choice of the fit range. The fit was performed in the range of 15-75 GeV. The upper and lower fit boundaries were varied individually by ± 5 GeV. This counts as two uncertainties, one for the upper boundary and one for the lower boundary.

- The choice of the control region. Two alternate control regions were used to check the effect of the reverted electron ID, and the effect of the anti-isolation used in the multi-jet control region. The first alternate control region uses a different reverted electron ID, and the second uses an alternative anti-isolation. These were varied independently and count as two separate uncertainties.

- The choice of Monte Carlo. To check against the possible bias in using ALPGEN, the multi-jet fit was performed again using SHERPA instead.

- The choice of binning. The binning in the fit was changed from bins of 5 GeV to 10 GeV.

For each variation, the difference between the variation and the nominal is taken as the systematic uncertainty. Results from the multi-jet systematics are outlined in Table 8.1. More information can be found about the fit and its uncertainties in Reference [2].

8.1.5.2 $t\bar{t}$ uncertainties

Recall in Section 5.3, that the top background is modeled with Monte Carlo and a normalization factor is applied to correct to the data in the validation region. This normalization is removed, and the difference is taken as a systematic uncertainty. The potential bias from choice of Monte Carlo (in this case POWHEG [51]) is explored by comparing to an alter-

	No jets	1 jet	2 jets	3 jets	4 jets	5 jets	≥ 6 jets
Statistical	0.40	0.48	0.79	1.60	3.16	6.61	13.46
Lower fit range	0.66	7.84	3.38	3.25	4.50	4.62	5.24
Upper fit range	0.05	0.98	0.48	0.74	1.02	4.22	4.90
Anti-isolation in MJ-CR	1.54	1.07	0.47	0.66	1.03	1.14	5.18
Reverted electron ID in MJ-CR	13.33	3.82	7.99	7.40	2.82	5.05	1.18
W + jets MC generator	5.26	3.53	2.89	6.55	2.71	2.46	5.10
Modified binning	0.12	0.14	0.17	0.16	0.07	0.48	11.21
Multi-jet fit	14.43	9.53	9.21	10.58	6.90	10.77	20.31

Table 8.1: Systematic uncertainties in the multi-jet prediction in % from the MJ fit for the different jet multiplicities.

nate prediction using MC@NLO+Herwig [52]. The effect was found to be minimal and no additional uncertainty was added due to this bias.

8.1.5.3 Monte Carlo cross section uncertainties

When using Monte Carlo, there is a theoretical uncertainty on the cross section generated which must be accounted for. To evaluate the systematic uncertainty, the cross section generated from Monte Carlo is varied by its uncertainty and the final background estimate is re-evaluated, including the re-fitting of the multi-jet background. The systematic uncertainty for $t\bar{t}$ and single top quark production are found by simultaneously varying their cross sections by their uncertainties, $\pm 6.8\,\%$ [53, 54, 55, 56]. For single Z production, the uncertainty is $\pm 5\,\%$ [57], and for diboson production (WW, WZ, ZZ) the cross sections are varied together by $\pm 7\,\%$ [58]. The uncertainties for Z boson, diboson, and single top quark production are small compared to other uncertainties. The $t\bar{t}$ uncertainty becomes significant in jet multiplicities greater than 5, where $t\bar{t}$ production is large in comparison to other cross sections.

8.1.6 Heavy flavor uncertainties

In addition to experimental uncertainties from b-jet tagging efficiencies, there is an additional theoretical uncertainty associated with the modeling of the b and c jets. There is a small mis-modeling observed in comparisons of data to Monte Carlo in phase spaces rich with heavy jets, such as the $t\bar{t}$ validation region. A mis-modeling may have a small effect when making extrapolations of the signal region (containing a b-jet veto) to the fiducial region (having no veto). This can be estimated by applying a scale factor of 1.8 to the $W + c$ jet contribution, and a scale factor of 0.5 to the $W + c\bar{c}$ and $W + b\bar{b}$ jet contributions [2]. This correction is not applied in the main analysis, but its impact on the final cross section compared to the nominal is taken as a systematic uncertainty.

8.1.7 Luminosity uncertainty

The integrated luminosity of the data taken in 2012 is $L_{\text{int}} = 20.2(4)\,\text{fb}^{-1}$ [59]. This corresponds to a relative uncertainty of 1.9 %. This uncertainty is applied to all signal predictions and background estimates from Monte Carlo.

8.1.8 Smoothing of systematic uncertainties

To reduce statistical fluctuations in systematic uncertainties, a smoothing procedure based on rebinning is used. The statistical error in the systematic difference (the difference between the systematically varied result and the nominal), is calculated from the statistical error in the nominal and systematically varied distributions, taking the correlation between the two into account. The fractional difference between the nominal and the systematically varied result, is calculated for each bin for each observable, i.e., H_{T}, p_{T}, rapidity etc.. Based on its

error, the fractional difference between the nominal and varied distribution is then rebinned starting from the highest value of the observable.

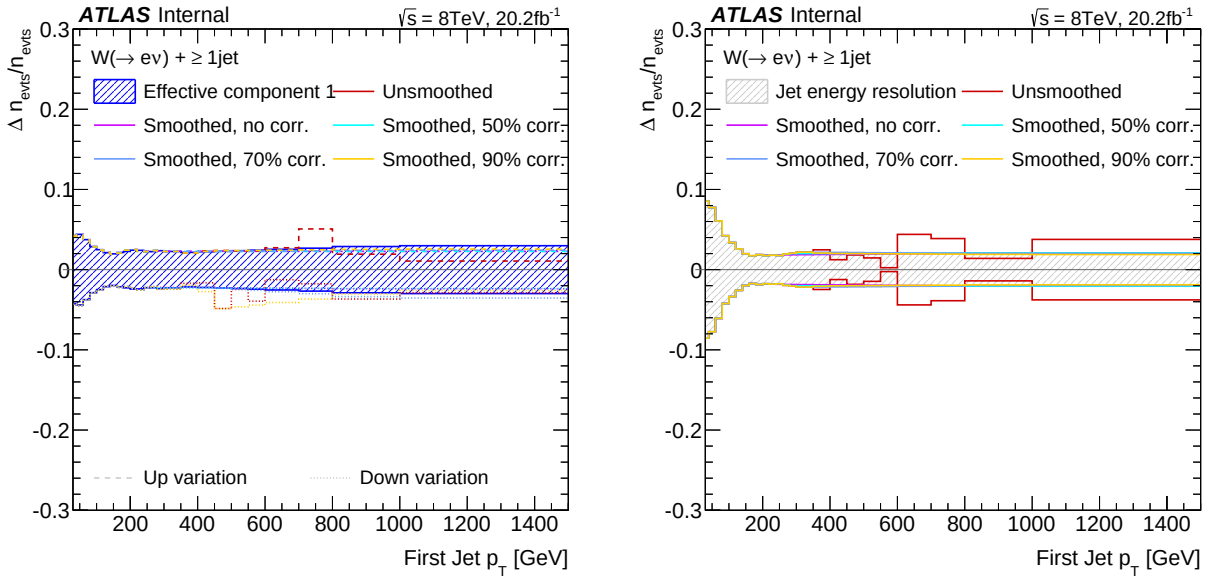
For each rebinning, the weighted mean of the factional difference f_{syst} is calculated:

$$\bar{f}_{\text{syst}} = \left(\sum_{\text{bins } i} f_{\text{syst},i} \cdot w_i \right) / \left(\sum_{\text{bins } i} w_i \right), \tag{8.2}$$

where $w_i = 1/(\Delta f^{\text{stat}}_{\text{syst},i})^2$ is the inverse of the squared statistical error in the fractional systematic difference. If this weighted mean of the fractional difference is larger than twice the weighted standard deviation $\sigma_{f_{\text{syst}}} = 1/\sqrt{\sum_{\text{bins } i} w_i}$, then the fractional systematic difference is considered as significant. Else, the calculation of the values is repeated including the next bin at lower observable values, until it reaches a significant value. All insignificant bins used are then combined. Typically, this results in a combination of several bins at large energy scales, while at low energy scales the systematic difference is already significant in each bin and remains unchanged.

The correlation between the nominal and varied distributions is used as a parameter in the smoothing process. For most distributions it is set to 70 %. For the uncertainty in choice of Monte Carlo (discussed later in Section 8.2.2.2), there is no correlation since the difference is the result of two different generators. Similarly, the correlation is set to zero for the uncertainty from reweighting of data (discussed later in Section 8.2.2.3), in order to get strong rebinning which helps suppress the large statistical fluctuations. In all other cases, the impact of the correlation was found to be small, justifying one global default. The smoothing is done separately for each jet multiplicity. Two examples are shown in Figure 8.1 displaying the effect of the smoothing and the impact of varying the correlation for the lead jet p_{T} distribution. The rebinning nicely smoothes the statistical fluctuations observed

at higher values of p_{T}.



(a) JES, effective component 1

(b) JER, 10 iterations

Figure 8.1: Smoothing of the systematic uncertainties in the leading jet p_{T} in the presence of at least 1 jet for the JES, effective component 1 (a) and the JER uncertainty (b). The hashed area shows the uncertainty that is further used in the analysis, the lines display the unsmoothed systematic uncertainty and four smoothed versions with different correlation assumptions. The symmetrized and with 70 %-correlation smoothed systematic difference determines the hashed area. For the JES, both up- and down variations are displayed, for the JER, the version where each jet is smeared 10 times is displayed. The plots here are done with uncertainties in the cross section, but the same applies for uncertainties at the detector level.

Several uncertainties are not smoothed. The uncertainty due to limited Monte Carlo statistics is left unchanged since it is itself a statistically driven uncertainty. Also, the uncertainties found from scaling the nominal distribution by a scaling factor are exempt since they are global across bins. These include the background cross section uncertainties, the heavy flavor jet cross section uncertainties, $t\bar{t}$ normalization uncertainties, most multi-jet uncertainties, and the uncertainty from the integrated luminosity.

8.1.9 Combined systematic uncertainties at detector level

The systematic uncertainties are estimated as described above for the total Standard Model prediction and then summed in quadrature. The jet, electron, $E_{\mathrm{T}}^{\mathrm{miss}}$, pile-up, background, and heavy flavor uncertainties, are given in Tables 8.2-8.3 as their fraction of the total Standard Model prediction ($\Delta n_{\mathrm{evts}}/n_{\mathrm{evts}}$).

For events with at least one jet, the jet systematic uncertainty dominates. This is expected and has historically been the case [60]. The main sources for the jet uncertainty are from the jet energy scale (JES) and the jet energy resolution (JER) at lower jet multiplicities, and just the JES at higher jet multiplicities.

	No jets	1 jet	2 jets	3 jets	4 jets	5 jets	6 jets	7 jets	8 jets	9 jets	10 jets
Jet	1.01	7.83	8.59	10.58	11.43	12.80	15.09	18.50	18.28	20.86	23.41
Electron	1.02	1.16	1.02	0.96	0.90	0.92	0.92	0.92	0.92	0.92	0.92
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.91	1.64	2.65	3.14	3.44	3.24	3.97	3.46	3.46	3.46	3.46
Pileup	0.36	0.32	0.40	0.35	0.20	0.20	0.20	0.20	0.20	0.20	0.20
Multi-jet bkg	0.50	0.81	1.38	1.71	1.14	1.61	2.83	2.77	2.72	2.34	3.94
$t\bar{t}$	0.00	0.01	0.13	0.63	1.58	2.50	3.29	3.82	4.08	4.73	5.20
Other background	0.00	0.06	0.21	0.67	1.50	2.29	2.96	3.39	3.73	4.26	4.55
W+heavy flavor jets	1.72	5.11	4.24	2.55	1.17	0.41	0.24	0.71	1.23	0.81	0.80
Total systematic	2.49	9.60	10.10	11.54	12.28	13.76	16.50	19.73	19.66	22.24	25.00

Table 8.2: Systematic uncertainties in the total SM prediction at the detector level in % for exclusive jet multiplicities. The uncertainty contributions from jet, electron, $E_{\mathrm{T}}^{\mathrm{miss}}$, pileup, background and W+heavy flavor jet related systematic uncertainties are listed and the total is the sum in quadrature.

	Inclusive	≥ 1 jet	≥ 2 jets	≥ 3 jets	≥ 4 jets	≥ 5 jets	≥ 6 jets	≥ 7 jets	≥ 8 jets	≥ 9 jets	≥ 10 jets
Jet	0.56	8.14	9.10	10.84	11.88	13.53	15.89	18.49	18.83	21.55	27.76
Electron	1.03	1.12	1.00	0.94	0.90	0.91	0.91	0.91	0.91	0.91	0.91
$E_{\mathrm{T}}^{\mathrm{miss}}$	1.06	1.94	2.78	3.21	3.43	3.38	3.84	3.56	3.56	3.56	3.56
Pileup	0.25	0.34	0.38	0.31	0.20	0.20	0.20	0.20	0.20	0.20	0.20
Multi-jet bkg	0.49	0.91	1.39	1.52	1.15	1.76	2.81	2.75	2.69	2.52	3.42
$t\bar{t}$	0.02	0.09	0.32	0.94	1.87	2.72	3.43	3.91	4.23	4.83	5.33
Other background	0.02	0.13	0.37	0.94	1.74	2.48	3.08	3.49	3.85	4.32	4.65
W+heavy flavor jets	2.21	4.76	3.74	2.15	0.93	0.22	0.35	0.74	0.83	0.79	0.69
Total systematic	2.77	9.74	10.39	11.72	12.75	14.56	17.24	19.78	20.22	22.96	29.09

Table 8.3: Systematic uncertainties in the total SM prediction at the detector level in % for inclusive jet multiplicities. The uncertainty contributions from jet, electron, $E_{\mathrm{T}}^{\mathrm{miss}}$, pileup, background and W+heavy flavor jet related systematic uncertainties are listed and the total is the sum in quadrature.

8.2 Systematic uncertainties in unfolding

The systematic uncertainties in the final unfolded distributions have two contributions. The first is from the detector level uncertainties described above, which are propagated through the unfolding. The second is from the uncertainties from the unfolding itself.

8.2.1 Propagating systematic uncertainties through unfolding

The detector level uncertainties are propagated through unfolding by varying the input, and repeating the unfolding. The difference between the unfolded distribution with the varied input and nominal unfolding is taken as the systematic uncertainty. The detector level uncertainties are separated into three categories:

1. Uncertainties in the particle calibrations, like the electron, $E_{\mathrm{T}}^{\mathrm{miss}}$, and jet uncertainties, as well as pileup uncertainties, which influence both the background estimate and the unfolding matrix.

2. Uncertainties in the background estimate like the Monte Carlo cross section and multi-

jet fit uncertainties, which only impact the background estimate.

3. Uncertainties in the cross section normalization of the $W+$ heavy flavor jets which only impact the unfolding matrix.

For uncertainties which affect the unfolding matrix (cases 1 and 3), a new matrix is determined by applying the systematic shift in the particle, Monte Carlo description, or $W+$ heavy flavor jet cross section and re-evaluating the matrix. The reconstruction efficiency $\varepsilon_{\text{reco}}$ and the fiducial correction c_{fid} are also re-evaluated using the shifted result. For uncertainties affecting the background estimate (cases 1 and 2), the difference $N_{\text{data}}(\mathbf{x_D}) - N_{\text{bkgd}}(\mathbf{x_D})$ is re-evaluated using the shifted background and unfolded using the nominal unfolding matrix (case 2) or the new matrix (case 1).

The systematically varied unfolded event estimate is therefore given by:

$$\hat{N}_{\text{syst}}(\mathbf{x_T}) \propto \sum_{\mathbf{x_D}} \left(N_{\text{data}}(\mathbf{x_D}) - N^{\text{syst}}_{\text{bkgd}}(\mathbf{x_D}) \right) \cdot P^{\text{syst}}(\mathbf{x_T}|\mathbf{x_D}) \tag{8.3}$$

which can be compared to Equation 7.2 [2]. The systematic uncertainty in the unfolded result is then:

$$\Delta\hat{N}^{\text{syst}}(\mathbf{x_T}) = -\left(\hat{N}^{\text{syst}}(\mathbf{x_T}) - \hat{N}(\mathbf{x_T}) \right) \tag{8.4}$$

The negative sign here is put in by hand as a convention to provide consistency in sign with other ATLAS approaches for propagating uncertainty.

8.2.2 Systematic uncertainties in the unfolding procedure

There are three sources of uncertainty in the unfolding procedure that are considered:

1. The uncertainty due to limited statistics of the Monte Carlo in the inputs generated from Monte Carlo.

2. The uncertainty due to the choice of Monte Carlo in the inputs generated from Monte Carlo.

3. The uncertainty due to a potential bias of the reweighting from the data distributions.

8.2.2.1 Uncertainty due to limited Monte Carlo statistics

The number of events generated in Monte Carlo is what determines the statistical uncertainty in the generated distributions that are used for constructing the reconstruction efficiency, the fiducial correction and the migration matrix. Bins of higher energy and higher jet multiplicity are especially sensitive, since the differential cross section is low and these bins are less populated.

The impact of the finite Monte Carlo statistics is estimated by varying unfolding inputs and repeating the unfolding procedure. For this, the number of events in each bin of the migration matrix (which pass both the SR and FR selections), the missed events (which pass FR, but not SR selections), and the fake events (which pass SR, but not FR selection) are varied. The events are varied using a Gaussian with the bin entry as the mean, and the bin uncertainty as Gaussian σ. To prevent matrices with negative values, the content of the bin matrix is required to be three times the bin uncertainty, else a Poisson variation is used. The variations in the missed and fake events are used to redetermine the reconstruction efficiency and fiducial correction.

For each variation, a new migration matrix, reconstruction efficiency, and fiducial correction is produced and the unfolding procedure is applied to the detector level results. This is

repeated 300 times to create 300 final cross section distributions. The RMS of the central 90 % of the result is taken as the systematic uncertainty due to limited Monte Carlo statistics for each bin. This uncertainty is considered uncorrelated between bins, so to calculate uncertainty in inclusive jet multiplicity bins (e.g. bins containing ≥ 1 jet), the uncertainties from the summed multiplicities can simply be added in quadrature.

8.2.2.2 Uncertainty due to choice of Monte Carlo

The reconstruction efficiency, fiducial correction, prior, and migration matrix are all calculated from Alpgen. Any modeling attributes unique to Alpgen introduce a potential bias due to the choice of using this generator. To estimate the uncertainty introduced by this bias, the unfolding is repeated with all inputs calculated from Sherpa instead. The difference between the result from Sherpa and the nominal result is taken as the systematic uncertainty in choice of Monte Carlo.

8.2.2.3 Uncertainty due to bias from data distributions

Although Alpgen provides a very reasonable description of the data, any differences between Alpgen and the data distribution introduces a potential bias. To estimate the uncertainty introduced by this bias, a reweighting scheme is used.

This reweighting scheme is used to reweigh the Alpgen migration matrix, and is described in detail in Section 7.3.2. The difference in the resulting cross section between the nominal result and the result from using the reweighted matrix is taken as the systematic uncertainty. This uncertainty is negligible except in only a few bins.

8.3 Combined unfolded systematic uncertainties

The systematic uncertainties are estimated as described above and combined in quadrature for the total systematic uncertainty. The fractional size of the detector level uncertainties propagated through unfolding, as well as the unfolding systematic uncertainties are listed in Tables 8.4 and 8.5.

	Inclusive	≥ 1 jet	≥ 2 jets	≥ 3 jets	≥ 4 jets	≥ 5 jets	≥ 6 jets	≥ 7 jets
Jet	0.51	11.63	14.23	18.53	24.60	34.89	49.97	70.21
Electron	1.11	1.41	1.42	1.52	1.80	2.14	2.14	2.14
$E_{\mathrm{T}}^{\mathrm{miss}}$	1.11	2.61	4.24	5.54	7.07	8.80	12.22	13.20
Pileup	0.27	0.55	0.63	0.53	0.41	0.41	0.41	0.41
Multi-jet background	0.51	1.26	2.13	2.56	2.46	4.72	8.84	11.89
$t\bar{t}$ background	0.02	0.14	0.55	1.81	4.22	7.61	11.66	16.79
Other background	0.02	0.20	0.62	1.77	3.90	6.83	10.30	14.71
W+heavy flavor jets	0.07	0.24	0.75	0.96	1.37	1.92	1.44	1.53
Unfolding	4.71	4.14	4.86	4.36	4.03	5.10	5.19	5.58
Total systematic	5.03	12.77	15.88	20.23	26.75	38.16	54.77	76.04

Table 8.4: Systematic uncertainties in the measured cross sections in % for inclusive jet multiplicities. The total is the sum in quadrature of the separately listed sources.

The jet uncertainties dominate in jet multiplicities greater than or equal to one. In the zero jet bin it is the unfolding systematic uncertainties that dominate. For events with exactly 1 jet, the total systematic uncertainty is about 12 %, most of which is the combined jet uncertainty. As the jet multiplicity rises, the contribution from the soft contributions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty, and the uncertainty from the various background estimates rise, but still remain well below the jet uncertainty.

One uncertainty that is significantly reduced after unfolding is the uncertainty from the $W+$ heavy flavor jets cross section. Since the uncertainty only propagates through the migration

	No jets	1 jet	2 jets	3 jets	4 jets	5 jets	6 jets	7 jets
Jet	1.00	10.89	13.19	17.11	22.19	31.18	45.08	65.73
Electron	1.07	1.41	1.40	1.46	1.76	1.76	1.76	1.76
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.94	2.13	3.91	5.14	6.64	8.06	11.76	11.76
Pileup	0.37	0.53	0.66	0.57	0.39	0.39	0.39	0.39
Multi-jet background	0.50	1.10	2.05	2.73	2.24	3.98	8.13	11.30
$t\bar{t}$ background	0.01	0.03	0.24	1.18	3.38	6.66	10.56	15.45
Other background	0.00	0.09	0.34	1.22	3.16	5.98	9.28	13.44
W+heavy flavor jets	0.05	0.10	0.69	0.85	1.23	2.04	1.41	1.47
Unfolding	5.77	3.94	4.99	4.46	3.85	4.73	4.85	5.25
Total systematic	6.06	11.93	14.88	18.78	24.13	34.10	49.63	70.98

Table 8.5: Systematic uncertainties in the measured cross sections in % for exclusive jet multiplicities. The total is the sum in quadrature of the separately listed sources.

matrix, only shape differences between $W+$ light flavor and heavy flavor jets have an impact, but not normalization. The shape differences are small, so the uncertainty is reduced.

8.3.1 Systematic uncertainties in the W^+ and W^- cross sections

Systematic uncertainties for the W^+ and W^- cross sections are determined in the same way as for the charge insensitive cross section. The only difference is that there is one additional uncertainty from the possibility of misidentifying the charge of the electron. The misidentification is corrected for by using a scale factor, the uncertainty of which is small and included in the electron uncertainty.

The fractional size of the combined uncertainties for the W^+ and W^- cross sections are listed in Tables 8.6 and 8.7 for the inclusive and exclusive jet multiplicities separately. The uncertainties for the charge sensitive distributions are only carried out to jet multiplicities of 6 jets. At 7 jets, the uncertainty exceeds 100 % and is therefore not given.

	Inclusive	≥ 1 jet	≥ 2 jets	≥ 3 jets	≥ 4 jets	≥ 5 jets	≥ 6 jets
				W^+			
Jet	0.46	11.33	13.37	16.98	22.26	30.65	45.32
Electron	1.14	1.39	1.40	1.50	1.75	1.75	1.75
$E_{\mathrm{T}}^{\mathrm{miss}}$	1.15	2.28	3.53	4.50	5.64	6.79	10.62
Pileup	0.28	0.51	0.56	0.47	0.37	0.37	0.37
Multi-jet background	0.40	0.96	1.45	1.64	2.85	7.43	18.64
$t\bar{t}$ background	0.02	0.12	0.46	1.47	3.34	5.90	9.69
Other background	0.02	0.17	0.53	1.45	3.10	5.30	8.43
W+heavy flavor jets	0.08	0.21	0.62	0.83	1.32	2.15	1.80
Unfolding	4.94	4.26	4.97	4.04	3.93	5.11	5.24
Total systematic	5.24	12.45	14.87	18.30	24.01	33.73	52.09
				W^-			
Jet	0.58	12.05	15.51	20.89	28.48	43.62	67.46
Electron	1.06	1.43	1.43	1.54	2.08	2.08	2.08
$E_{\mathrm{T}}^{\mathrm{miss}}$	1.06	3.09	5.36	7.23	9.33	12.01	15.97
Pileup	0.26	0.61	0.74	0.63	0.42	0.42	0.42
Multi-jet bkg	0.68	1.90	3.65	4.32	4.87	10.68	15.39
$t\bar{t}$ background	0.02	0.18	0.70	2.35	5.70	10.88	17.59
Other background	0.03	0.23	0.76	2.26	5.20	9.63	15.42
W+heavy flavor jets	0.04	0.34	1.02	1.24	1.40	1.45	1.02
Unfolding	4.42	3.89	4.57	4.66	4.38	4.46	4.72
Total systematic	4.76	13.27	17.55	23.32	31.73	48.98	74.95

Table 8.6: Systematic uncertainties in the measured cross sections in % for inclusive jet multiplicities for the W^+ and W^- selections separately. The total is the sum in quadrature of the separately listed sources.

Uncertainties behave the same way for the W^+ and W^- cross sections as for the charge independent cross section. The jet energy scale and jet energy resolution still dominate for events with at least one jet, and scale similarly with jet multiplicity. Additionally the unfolding is dominant in the presence of no jets. Uncertainties are larger for the W^- production due to the process having less cross section overall at a proton-proton collider.

8.3.2 Systematic uncertainties in the W^+ / W^- ratio

Systematic uncertainties in the W^+ / W^- ratio are calculated from the W^+ and W^- systematic uncertainties. The uncertainties for the ratio are split into correlated and uncorrelated uncertainties. Correlated uncertainties are calculated by re-evaluating the ratio r with the systematic variations applied to the numerator and denominator together r^{syst}. The uncertainty is then simply $\Delta r = |r^{\text{syst}} - r|$.

The uncorrelated uncertainties are propagated in the traditional way by adding the relative uncertainties from the numerator and denominator in quadrature:

$$\Delta r^{\text{uncorr}} = r \cdot \sqrt{\left(\frac{\Delta\sigma^{\text{uncorr}}_{W^+}}{\sigma_{W^+}}\right)^2 + \left(\frac{\Delta\sigma^{\text{uncorr}}_{W^-}}{\sigma_{W^-}}\right)^2}. \tag{8.5}$$

The uncorrelated uncertainties are those that result from statistical sources i.e. statistical uncertainty from the multi-jet fit, statistical uncertainty in the background estimates, and the unfolding uncertainty due to limited Monte Carlo statistics. All other systematic uncertainties are treated as correlated.

The systematic uncertainties for the W^+ / W^- ratio are listed in Tables 8.8 and 8.9 for the inclusive and exclusive jet multiplicities. There is a large amount of cancellation in many of the uncertainties in the ratio. In the presence of at least one jet the total systematic

	No jets	1 jet	2 jets	3 jets	4 jets	5 jets	6 jets
				W^+			
Jet	0.99	10.75	12.46	15.66	20.17	27.41	41.06
Electron	1.11	1.39	1.38	1.46	1.65	1.65	1.65
$E_{\mathrm{T}}^{\mathrm{miss}}$	1.03	1.91	3.27	4.19	5.33	6.40	8.68
Pileup	0.37	0.49	0.58	0.50	0.38	0.38	0.38
Multi-jet bkg	0.38	0.88	1.52	1.82	2.16	5.36	17.08
$t\bar{t}$ background	0.01	0.04	0.21	0.97	2.69	5.13	8.84
Other background	0.00	0.08	0.30	1.00	2.51	4.60	7.58
W+heavy flavor jets	0.06	0.09	0.57	0.70	1.09	2.23	1.59
Unfolding	6.02	4.06	5.22	4.08	3.71	4.87	5.04
Total systematic	6.31	11.78	14.08	16.96	21.71	30.00	47.11
				W^-			
Jet	1.01	11.11	14.27	19.31	25.26	38.33	60.69
Electron	1.02	1.43	1.42	1.44	1.94	1.94	1.94
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.83	2.45	4.90	6.72	8.72	11.16	15.70
Pileup	0.37	0.58	0.77	0.68	0.40	0.40	0.40
Multi-jet bkg	0.72	1.53	3.50	4.54	3.74	9.69	14.66
$t\bar{t}$ background	0.01	0.05	0.31	1.56	4.55	9.48	15.94
Other background	0.00	0.11	0.40	1.56	4.19	8.34	13.94
W+heavy flavor jets	0.01	0.15	0.96	1.20	1.39	1.54	1.22
Unfolding	5.43	3.70	4.55	4.86	4.08	4.19	4.52
Total systematic	5.74	12.16	16.26	21.70	28.08	43.26	67.96

Table 8.7: Systematic uncertainties in the measured cross sections in % for exclusive jet multiplicities for the W^+ and W^- selections separately. The total is the sum in quadrature of the separately listed sources.

uncertainty reduces from 11-13 % in the W^+ and W^- cross sections down to only 1.5 % in the ratio. Most of this is due to the reduction in the combined jet uncertainties from 11-12 % in the individual cross sections down to less than 1 % in the ratio.

	Inclusive	≥ 1 jet	≥ 2 jets	≥ 3 jets	≥ 4 jets	≥ 5 jets	≥ 6 jets
Jet	0.13	0.74	2.08	3.83	6.72	14.85	27.30
Electron	0.09	0.14	0.13	0.08	0.47	0.47	0.47
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.09	0.82	1.87	2.80	3.84	5.52	6.16
Pileup	0.02	0.10	0.18	0.16	0.06	0.06	0.06
Multi-jet bkg	0.32	1.24	2.85	3.21	5.90	14.66	27.09
$t\bar{t}$ background	0.02	0.08	0.28	0.96	2.61	5.88	10.23
Other background	0.01	0.06	0.23	0.81	2.12	4.38	7.20
W+heavy flavor jets	0.03	0.13	0.39	0.40	0.09	0.70	0.78
Unfolding	0.55	0.50	0.61	0.74	1.27	1.81	2.65
Total systematic	0.66	1.75	4.09	5.93	10.38	22.88	41.01

Table 8.8: Systematic uncertainties in the measured W^+/W^- cross section ratio in % for inclusive jet multiplicities. The total is the sum in quadrature of the separately listed sources.

Most detector related uncertainties see high cancellation except the multi-jet fit uncertainty (Section 8.1.5.1). At low jet multiplicities this is dominated by the potential Monte Carlo bias uncertainty, where there is a difference in how Sherpa and Alpgen model the individual W^+ and W^- cross sections which causes imperfect cancellation in the ratio. At higher multiplicities the uncertainty from the choice of reverted electron ID in the control region dominates because it suffers from low statistics. Additionally the statistical uncertainty on the multi-jet fit itself is treated as uncorrelated and thus does not benefit from the cancellation in the ratio, and so this also becomes significant at high jet multiplicities. In general, the systematic uncertainties are larger than the statistical uncertainties in the data,

	Exclusive	1 jet	2 jets	3 jets	4 jets	5 jets	6 jets
Jet	0.07	0.45	1.75	3.50	5.39	12.42	23.53
Electron	0.10	0.14	0.16	0.11	0.41	0.41	0.41
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.20	0.54	1.66	2.59	3.51	5.01	8.46
Pileup	0.01	0.08	0.19	0.19	0.06	0.06	0.06
Multi-jet bkg	0.40	0.83	2.84	3.20	3.79	12.19	25.36
$t\bar{t}$ background	0.02	0.05	0.18	0.69	2.10	5.24	9.42
Other background	0.00	0.03	0.11	0.56	1.68	3.78	6.52
W+heavy flavor jets	0.05	0.07	0.39	0.49	0.30	0.68	0.37
Unfolding	0.63	0.67	0.91	1.00	1.18	1.82	2.80
Total systematic	0.79	1.30	3.87	5.60	8.04	19.33	37.52

Table 8.9: Systematic uncertainties in the measured W^+/W^- cross section ratio in % for exclusive jet multiplicities. The total is the sum in quadrature of the separately listed sources.

but of the same order.

Chapter 9

Theoretical predictions

With the unfolding completed, data can now be easily compared to theoretical predictions, both from fixed order theory and from parton shower Monte Carlo. Predictions can be generally categorized into the order in α_s that the matrix element is calculated to using perturbative QCD. In general, higher orders are more difficult to calculate and require drastically more computation time. Each additional jet introduces an additional vertex and another order of α_s. As such, this analysis compares predictions from a number of fixed order, including those calculated to leading order (LO), next to leading order (NLO), and next to next to leading order (NNLO). These fixed order predictions can either be used to produce their own predictions at the parton level, or be used as inputs to parton shower Monte Carlo programs. These predictions are briefly outlined in this chapter and shown in comparison to data in Chapter 10.

Monte Carlo predictions shower the final state partons to particle level and then hadronize them. Showered partons can sometimes form new jets, and so, events may include more jets than are calculated by the matrix element as described above. Often one program is used to calculate the matrix element, and another program is used to shower the resultant final state partons. Table 9.1 summarizes the predictions that are compared to data in Chapter 10.

There are a few remaining effects not taken into account by the unfolding that must be

investigated. The corrections from the underlying event and hadronization cannot be modeled using perturbative QCD. These are categorized as non-perturbative corrections. Parton shower Monte Carlos model them, but fixed order predictions do not and an independent correction must be applied. Additionally, while the prediction is calculated to a specified order in α_s, some corrections from QED are not taken into account in general. This effect is also discussed below along with the effect of "dressing" the electrons.

Program	Order in α_s	$N_{\text{partons}}^{\text{max}}$ at highest order	PDF set	PS	NPC	Comments
N_{jetti}	NNLO	1	CT14		✓	Not shown for N_{jets}, $\Delta R_{\text{jet1,jet2}}$, and $m_{\text{jet1,jet2}}$
Blackhat+Sherpa	NLO	1, 2, or 3	CT10		✓	
MCFM 6.8	NLO	1	CT10+ 3 more		✓	Figure 10.6 only
Powheg+Pythia8	NLO	1	CT10	✓		Figure 10.6 only
Sherpa 2.2.1	NLO	2	NNPDF 3.0 NNLO	✓		Including NLO EW corrections in Figure 10.6
Sherpa 2.2.1	LO	3	NNPDF 3.0 NNLO	✓		
Alpgen+Pythia6	LO	5	CTEQ6L1 (LO)	✓		
Alpgen+Herwig	LO	5	CTEQ6L1 (LO)	✓		
Sherpa 1.4.1	LO	4	CT10	✓		

Table 9.1: Summary of $W+$jets predictions which are compared to the measured differential $W+$jets cross sections. Included is the order in α_s, the maximum number of partons included in the matrix element calculation, which PDF set is used, whether the parton shower (PS) is included, and whether the non-perturbative corrections (NPC) (see Section 9.5 below) are applied.

9.1 Theoretical uncertainty

There are two main uncertainties associated with theoretical predictions. The first is the uncertainty from the PDF used, which is outlined in Section 2.2. The second is the scale uncertainty. When calculating the cross section, an appropriate scale must be chosen for factorization and renormalization. An all-orders calculation would be insensitive to the values of these scales, but predictions to finite order retain a sensitivity, which decreases as the order of the calculation increases. More can be read about this subject in reference [19]. A good choice of scale, related to the physical process, will lead to a more reliable prediction. The ATLAS prescription for this analysis is to use a dynamic scale of $H_{\mathrm{T}}/2$. In order to find the scale uncertainty, the scales are varied individually up and down by a factor of two, and the envelope of the variation is taken as the uncertainty.

9.2 LO predictions

The leading order predictions used in this analysis are Alpgen v2.14 and Sherpa v1.4.1. Alpgen events were produced with up to 5 partons in the final state and showered with Pythia v6.426 (abbreviated as Alpgen+PY6 in the figures). Pythia also models the hadronization and underlying event, so no non-perturbative corrections need to be applied. Electromagnetic final state radiation is modeled using Photos [61] and Tauola [62], and the PDF set CTEQ6L1 is used to model the proton structure. Alpgen is also shown using an alternate parton showering model from Herwig [63] with Jimmy [64].

Sherpa events were generated with up to 4 partons in the final state. Sherpa does its own modeling of the parton shower along with the hadronization, underlying event, and electromagnetic final state radiation.

9.3 NLO predictions

NLO predictions are generated with BlackHat+Sherpa [65] (abbreviated BH+S in the figures). Often a prediction is referred to as the generator that does the matrix element calculation "+" the generator that does the showering, e.g. Alpgen+Pythia6. In this case however, BlackHat provides the NLO virtual matrix element corrections and Sherpa calculates the tree-level diagrams and provides the phase-space integration. Tree-level diagrams are the LO parts of the cross-section calculation, and the virtual matrix element corrections are the higher order terms. Although BlackHat+Sherpa provides predictions with up to 5 jets at NLO, only $W + 1$, 2, and 3 jets are used in this analysis as it focuses on 1 and 2 jet differential cross-sections. The proton structure is modeled using the PDF set CT10 NLO. In addition to the vanilla BlackHat+Sherpa predictions, the exclusive sum approach is also shown for the $W + 1$ jet predictions. The exclusive sum adds the NLO correction for the $W + 2$ jet prediction to the full $W + 1$ jet inclusive prediction. This has the advantage over the standard fixed order prediction of containing information about additional partons, but lacks complete theoretical rigor. This is useful in observables that are sensitive to additional jets such as inclusive lead jet p_{T} or H_{T}.

The MCFM generator [66] is also used in this analysis to model one jet at NLO and 2 jets at LO. Four choices of PDF are shown for MCFM: CT10 [67], HERAPDF 1.5 [68], MSTW 2008 [69], and NNPDF 2.3 [70].

As mentioned above, Sherpa 1.4.1 is used for LO predictions. The later release of Sherpa 2.2.1 gives us NLO predictions as well. Sherpa 2.2.1 calculates the matrix element with up to 2 partons in the final state at NLO and also 3 partons at LO. Sherpa provides its own modeling of the parton shower, hadronization, the underlying event. The PDF set CT10 was

used to model the proton structure. The LO prediction is shown as well but the theoretical uncertainty is left out of the figures.

Finally the NLO $W + 1$ jet cross sections are also calculated using POWHEG r2129 showered with PYTHIA 8 (abbreviate PWHG+PY8 in the figures). The CT14 PDF set is used for the POWHEG calculation, and CTEQ6L1 is used for the parton shower. The POWHEG cross sections are scaled by a factor of 1.1 for events with at least 1 jet to match the integrated number of events in the data. Only statistical uncertainties are shown.

9.4 NNLO predictions

The N_{jetti} [71] [72] program is capable of calculating the W+jets cross section at NNLO. It is named after the N-jettiness subtraction technique it uses to control the infrared singularities of final state partons. This sample was generated using the CT14 NNLO PDF set. The kinematic selection defined for theory predictions in Table 7.1 is used with the exception of the jet rapidity cut. This sample was generated requiring $|y| < 2.5$ instead of 4.4. This stricter rapidity cut was necessary due to the inherent larger uncertainties of the technique when dealing with higher rapidity jets. In order to make comparisons to data and the other samples, a correction is applied. The correction was estimated using ALPGEN+PYTHIA6. A straight forward bin by bin ratio is taken of events generated using the $|y| < 4.4$ cut to the events generated using the $|y| < 2.5$ cut. This is done separately for all observables. The effect of this correction is about 10 to 15 % at low transverse momentum of the W boson, jets, and H_{T}, and decreases to 0 at higher energies. The systematic uncertainty for the correction is taken as the statistical uncertainty from the ALPGEN+PYTHIA sample as well as the potential bias from Monte Carlo sample. This is estimated by repeating the correction

with SHERPA 1.4.1 and taking the difference as the uncertainty.

9.5 Non-Perturbative Corrections

There are two corrections that must be applied that cannot be calculated from perturbative QCD. The first is from hadronization, the process where colored partons turn into colorless jets. Some of the energy from hadronization is deposited outside of the jet cone and thus is lost. The second correction is from the underlying event. When the partons collide the remnants of the protons are unaccounted for but will inevitably deposit some additional energy into the detector. This is called the underlying event. It is by happy coincidence that these two effects directly oppose each other for the $R = 0.4$ jet size used, and are of approximately the same order. As a result the combined non-perturbative correction (NPC) is small.

The non-perturbative corrections are determined with SHERPA 2.2.1 at leading order. The matrix element calculation is carried out with emissions of up to 2 jets using the default PDF NNPDF 3.0 NNLO. The generated events are selected using the same criterion as the nominal analysis from Table 7.1.

Four samples are generated for the NPC:

1. Parton level with no hadronization (Had) or underlying event (UE) (Fragmentation off for hadronization, multiple parton interactions off and intrinsic transverse beam momentum off for underlying event).

2. Particle level with non-perturbative effects on.

3. With hadronization on and the underlying event off.

4. With hadronization off and the underlying event on.

The corrections are done separately on a bin-by-bin basis and are defined as:

$$\delta^{had} = \frac{\text{Had On, UE Off}}{\text{Had Off, UE Off}} \qquad \text{and} \qquad \delta^{UE} = \frac{\text{Had Off, UE On}}{\text{Had Off, UE Off}} \tag{9.1}$$

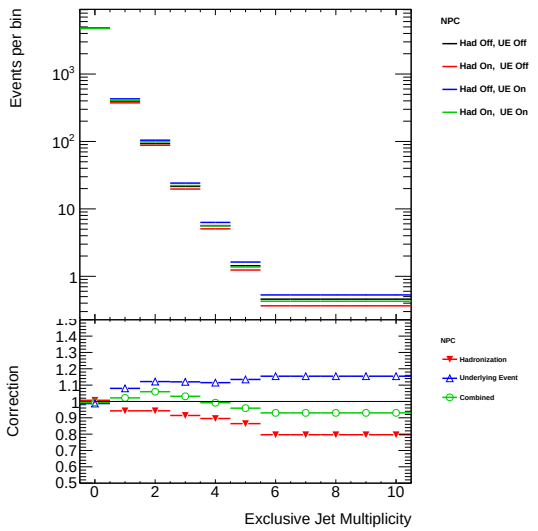
Or with the two corrections combined:

$$\delta^{comb} = \frac{\text{Had On, UE On}}{\text{Had Off, UE Off}} \tag{9.2}$$

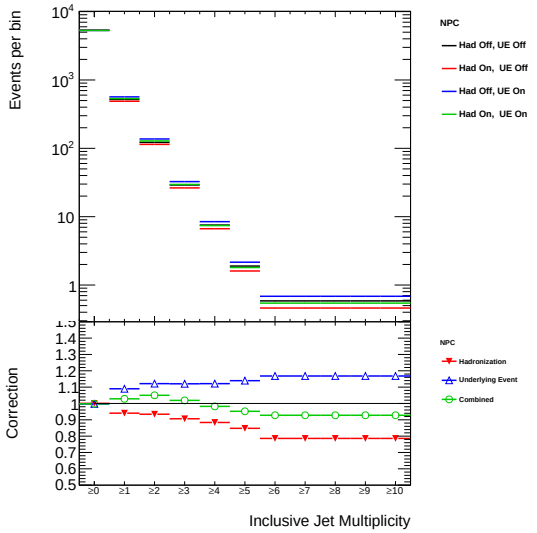
100 million events are generated for each sample which provides adequate statistics in the low energy regions where the individual corrections are large. In the higher energy bins, low statistics can create large ratios between samples despite this being the region where the corrections should be closest to one. To compensate for this, high energy bins are merged together and the ratio of each sub-bin is fixed to the ratio of the merged bin. The merging is done iteratively until the fractional statistical uncertainty of the merged bin is less than 2%.

The systematic uncertainties of the NPC are determined by taking the envelope of the following varied parameters: the starting scale of the parton shower, the recoil scheme, the mode of shower evolution, and the number of emitted partons from the matrix element. Each of these samples are generated separately. More can be read about these shower parameters in References [73] and [74]. The statistical and systematic uncertainties are added to the theoretical uncertainty in the predictions they are applied to.

The non-perturbative corrections are shown for the charge-insensitive distributions in Figures 9.1 - 9.5. The top half of the figures show the four samples generated and the bottom half

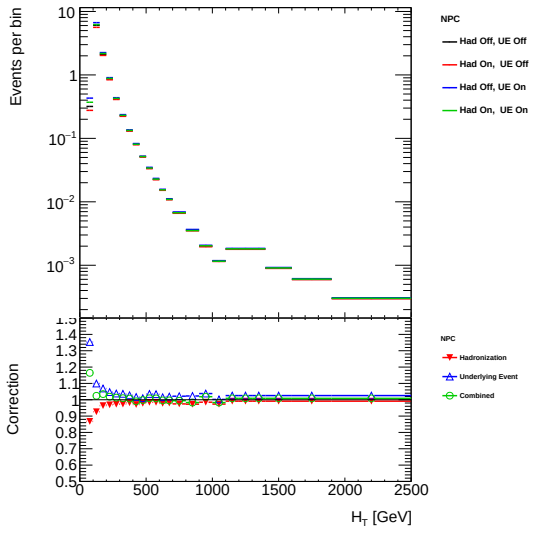


(a) Exclusive Jet Multiplicity

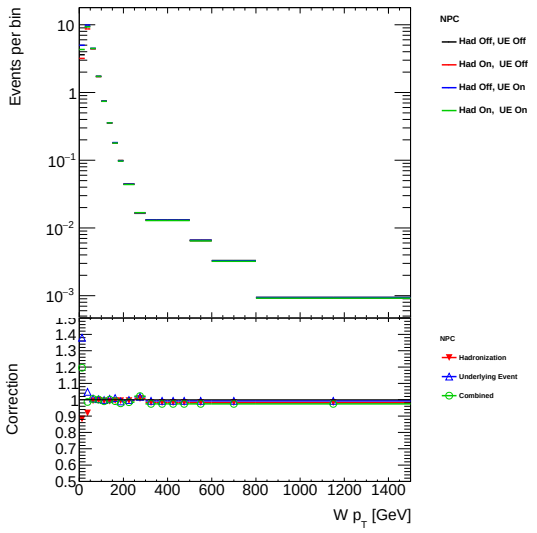


(b) Inclusive Jet Multiplicity

Figure 9.1: The hadronization and underlying event corrections for the exclusive and inclusive jet multiplicity.

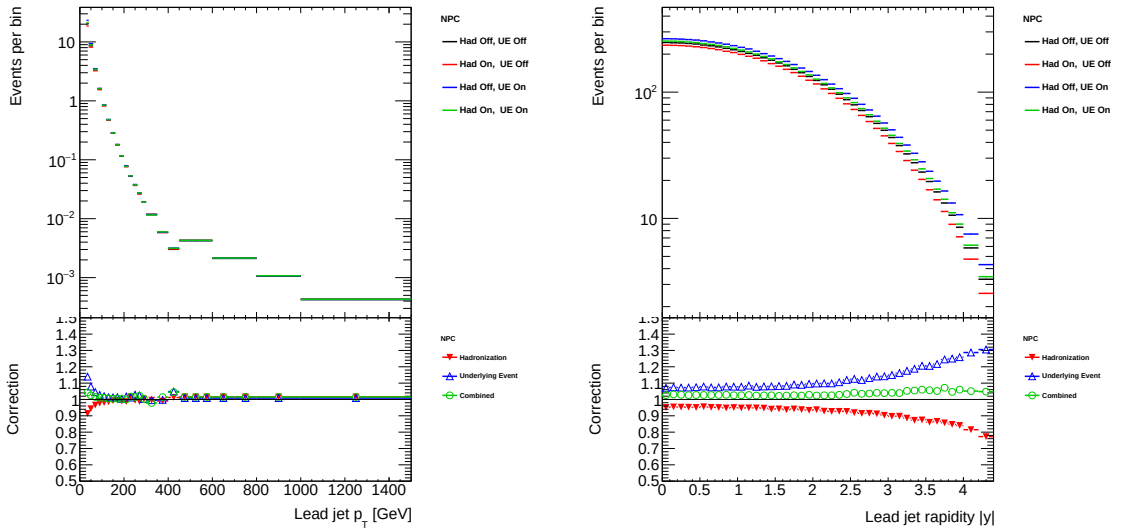


(a) $H_{\mathrm{T}} \geq 1$ jet



(b) W $p_{\mathrm{T}} \geq 1$ jet

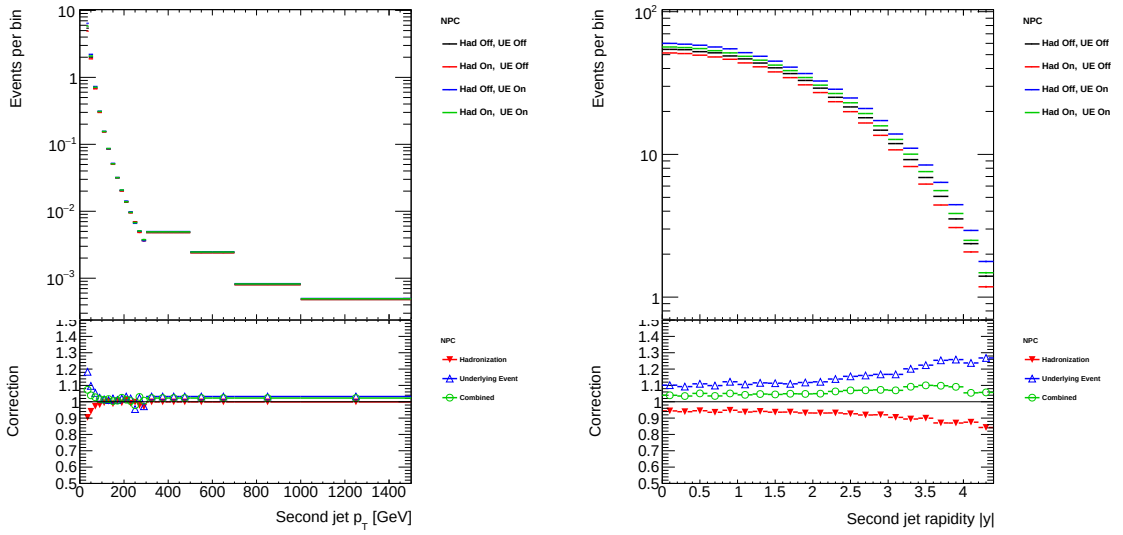
Figure 9.2: The hadronization and underlying event corrections for the H_{T} and W boson transverse momentum.



(a) Lead jet $p_{\mathrm{T}} \geq 1$ jet

(b) Lead jet rapidity ≥ 1 jet

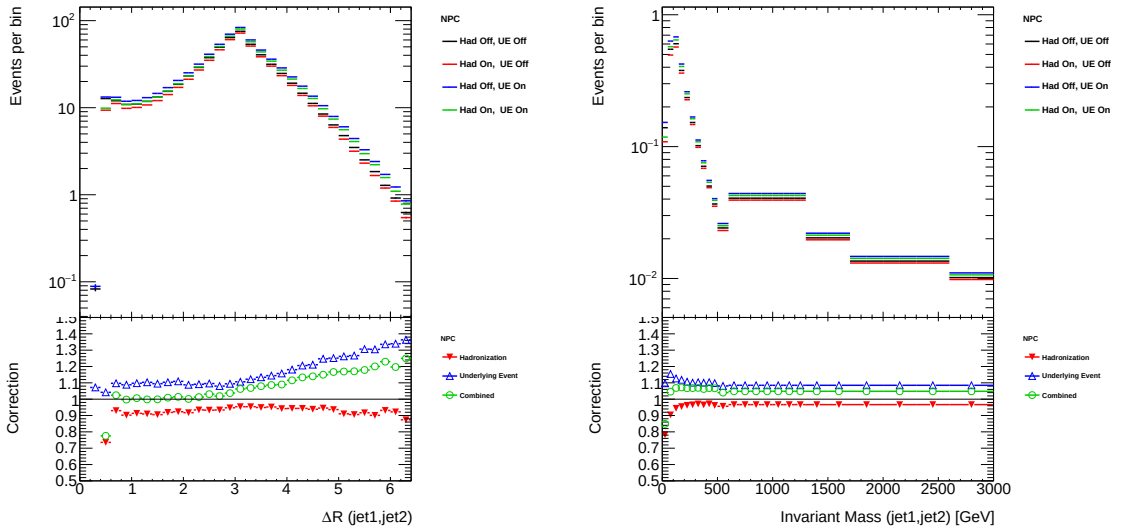
Figure 9.3: The hadronization and underlying event corrections for the lead jet p_{T} and rapidity.



(a) Second leading jet $p_{\mathrm{T}} \geq 2$ jet

(b) Second leading jet rapidity ≥ 2 jet

Figure 9.4: The hadronization and underlying event corrections for the second leading jet p_{T} and rapidity.



(a) Angular separation $\Delta R_{\text{jet1,jet2}}$

(b) Invariant mass of two lead jets ≥ 2 jet

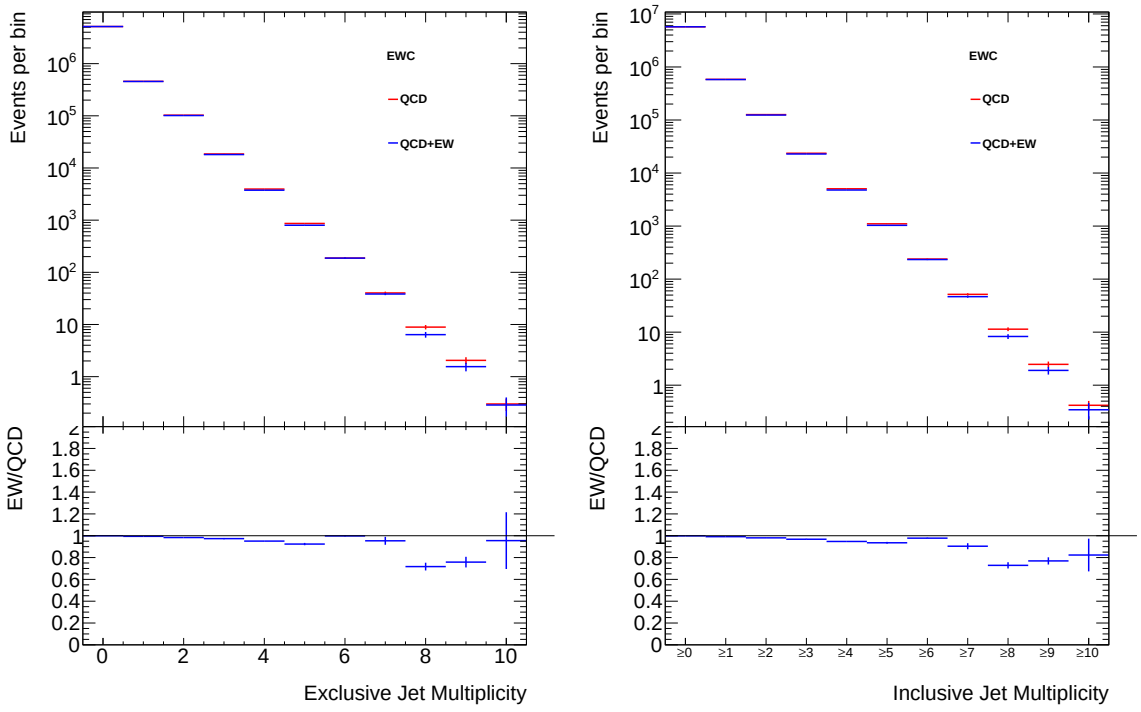
Figure 9.5: The hadronization and underlying event corrections for the angular separation of the two lead jets, and the invariant mass of the two lead jets.

show the corrections defined in Equations 9.1 and 9.2. The corrections are in general on the order of 2-3%. Unsurprisingly, the corrections grow with the number of jets as seen in Figure 9.1. More jets means that there are more opportunities for energy to be lost from hadronization and gained from the underlying event.

9.6 Electroweak Corrections

In general, calculations are referred to by their order in α_s; however this doesn't include higher order calculations of the electroweak coupling α_{EW}. SHERPA 2.2.1 has the capability of including these NLO electroweak terms in the cross section calculation. These extra terms can have a significant impact especially on certain observables such as the transverse momentum of the W boson [75].

Having these corrections available makes it possible to study the impact of these additional



(a) Exclusive Jet Multiplicity

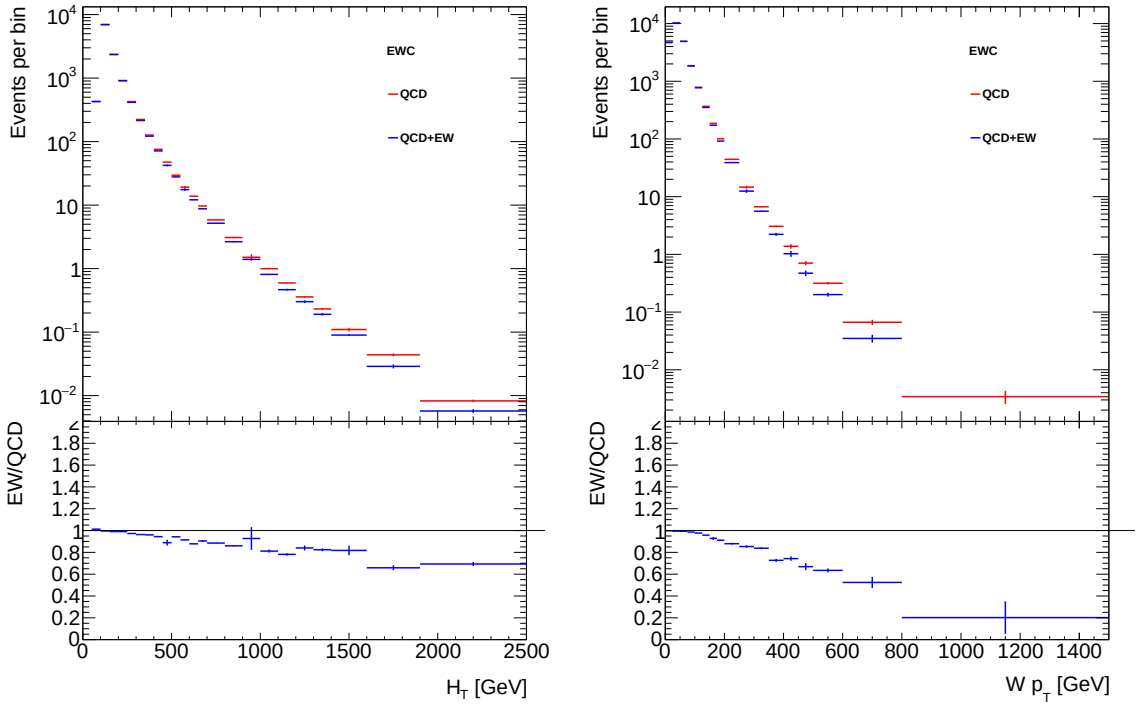
(b) Inclusive Jet Multiplicity

Figure 9.6: The electroweak corrections for the inclusive and exclusive jet multiplicity generated with SHERPA 2.2.1 using the cuts from Table 7.1

terms on W + jets kinematics. This is done in a straight forward manner, simply running SHERPA with and without these terms included with all other parameters the same as for the nominal analysis. Figures 9.6 - 9.10 show the impact of these corrections.

The impact in general grows with energy. Energy related distributions like H_{T} and lead jet p_{T} are impacted minimally at low energies and grow to 20-30% at higher values. Angular distributions remain mostly unaffected. The largest impact is on the transverse momentum of the W boson (Figure 9.7b) which quickly grows to tens of percent.

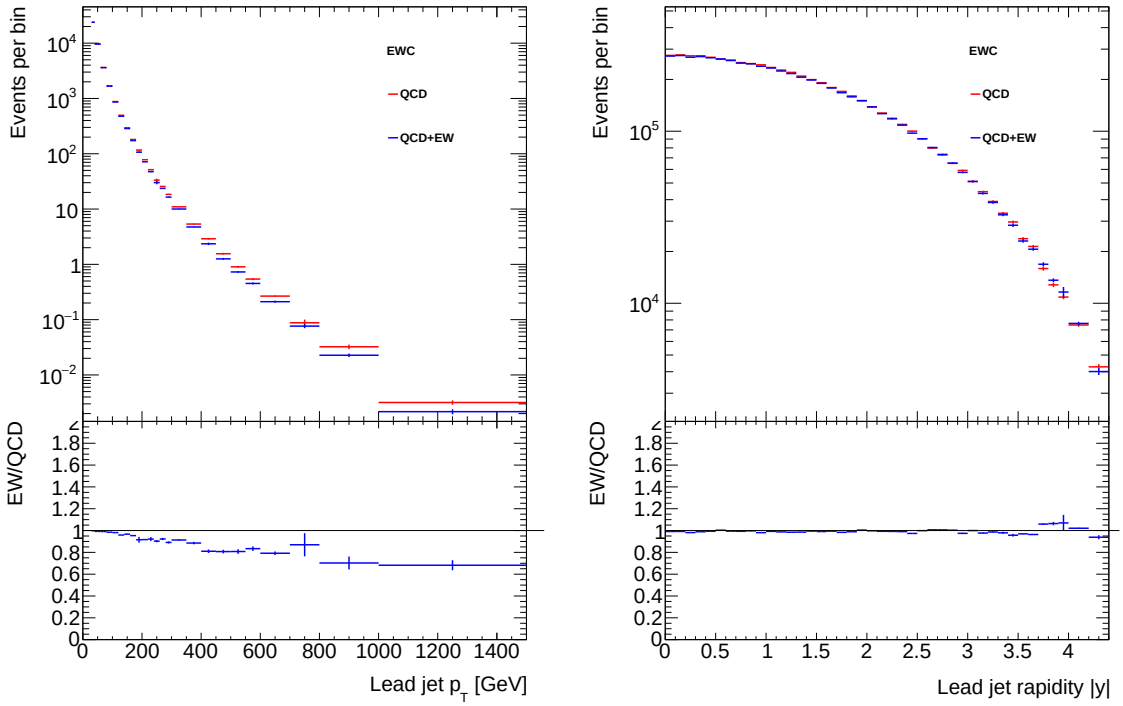
The electroweak corrections are included in the distributions shown for SHERPA 2.2.1 at NLO, but are not applied to the other Monte Carlo predictions.



(a) $H_{\mathrm{T}} \geq 1$ jet

(b) W $p_{\mathrm{T}} \geq 1$ jet

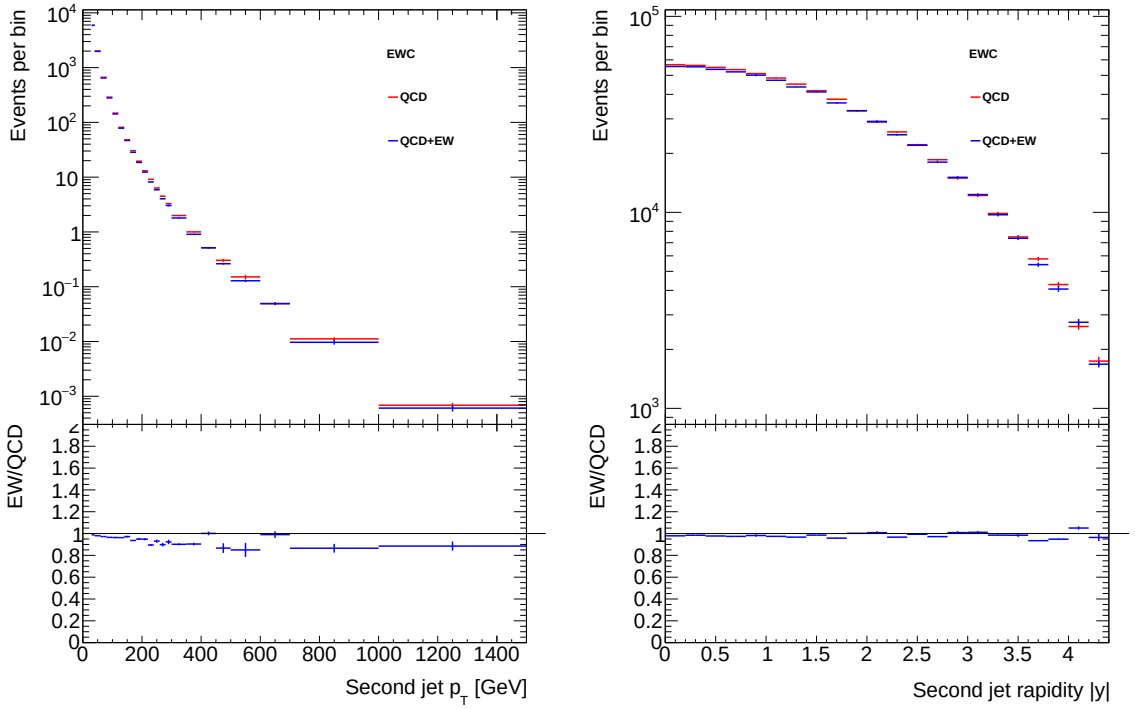
Figure 9.7: The electroweak corrections for H_{T} and W boson transverse momentum generated with Sherpa 2.2.1 using the cuts from Table 7.1



(a) Lead jet $p_{\mathrm{T}} \geq 1$ jet

(b) Lead jet rapidity ≥ 1 jet

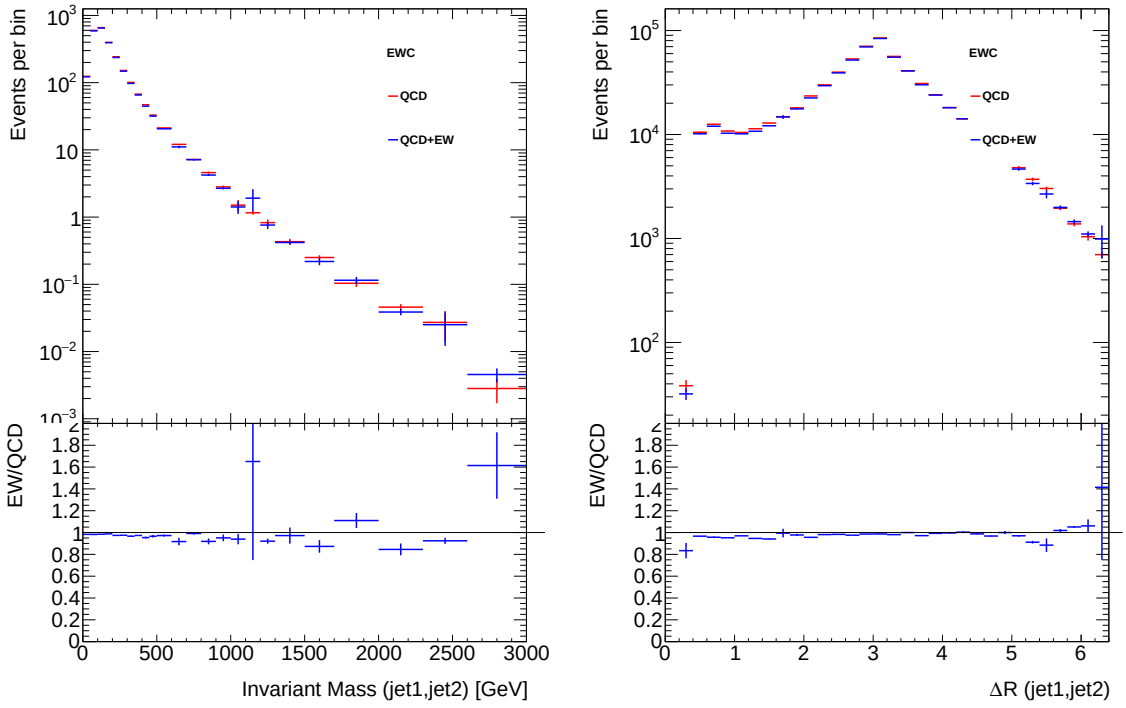
Figure 9.8: The electroweak corrections for lead jet transverse momentum and rapidity generated with SHERPA 2.2.1 using the cuts from Table 7.1



(a) Second leading jet $p_{\mathrm{T}} \geq 2$ jet

(b) Second leading jet rapidity ≥ 2 jet

Figure 9.9: The electroweak corrections for second lead jet transverse momentum and rapidity generated with SHERPA 2.2.1 using the cuts from Table 7.1



(a) Invariant mass of two lead jets ≥ 2 jet

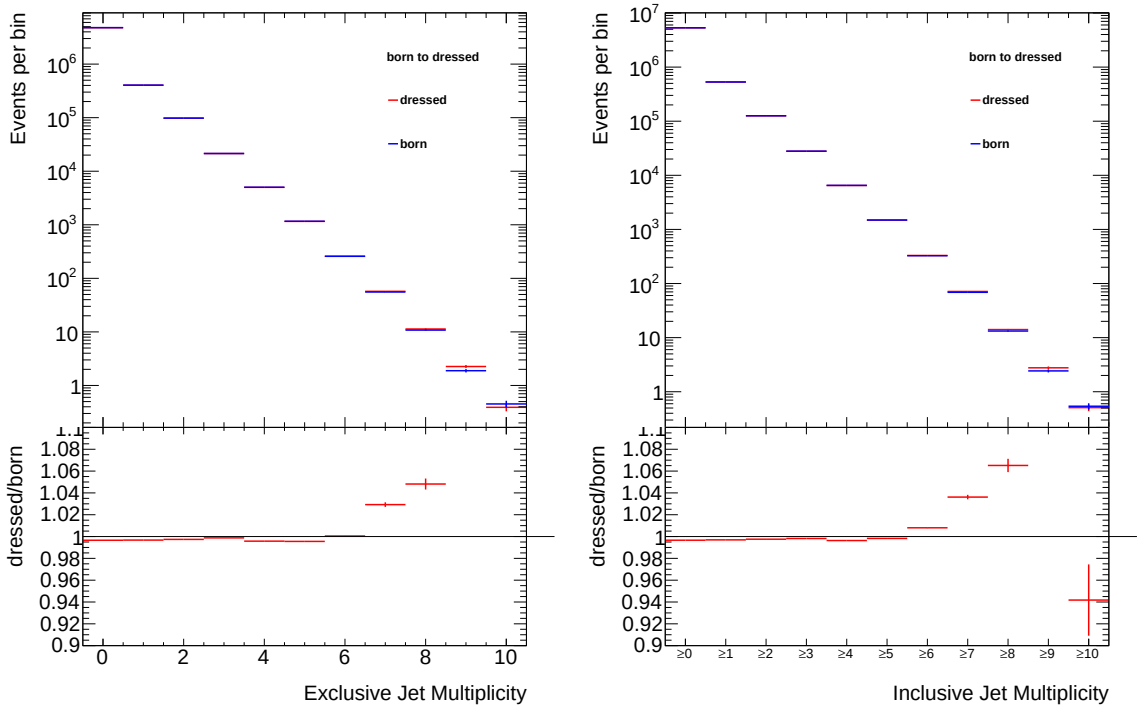
(b) Angular separation $\Delta R_{\text{jet1,jet2}}$

Figure 9.10: The electroweak corrections for dijet invariant mass and the angular separation between the two lead jets generated with SHERPA 2.2.1 using the cuts from Table 7.1

9.7 Dressed electron corrections

The final effect to consider is dressing the electron. High energy electrons tend to radiate energy through photon emissions. This creates a showering effect similar to the way that quarks radiate gluons. This radiation is called electromagnetic final state radiation. Modeling these emissions and summing their energy back into the electron is called dressing the electron. For keeping track of the nomenclature, the electron before radiation is called the Born level electron, and the electron after is called the dressed electron.

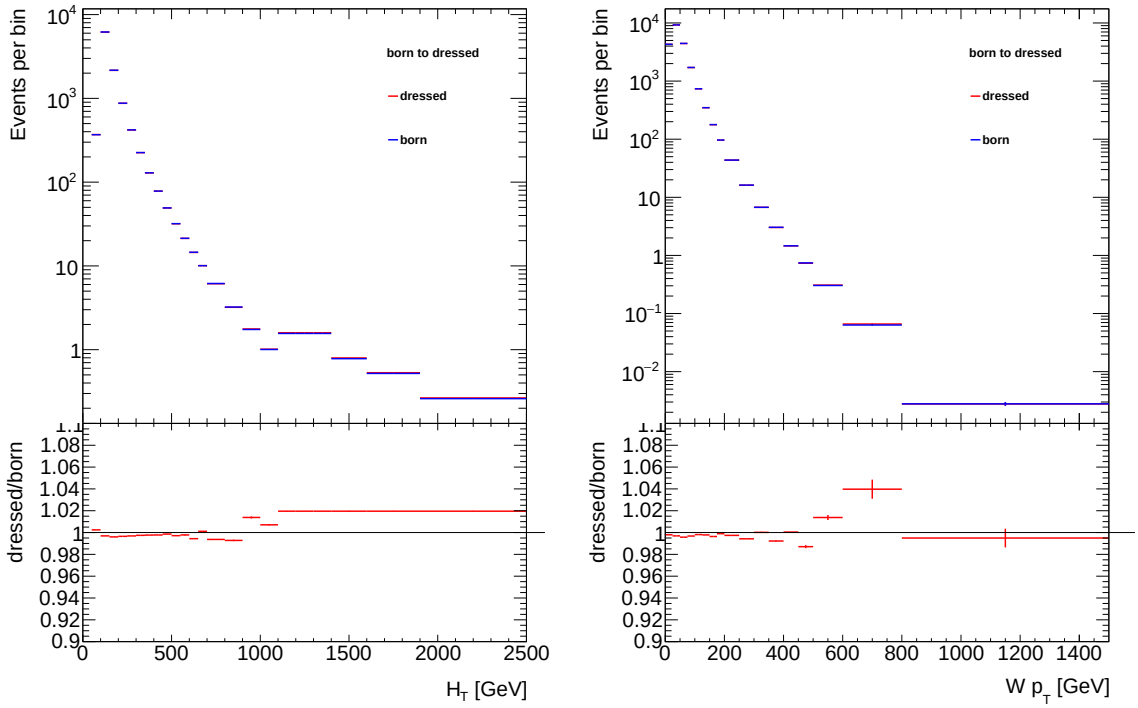
Sherpa 2.2.1 is used for the making the correction from Born level to the dressed. Similar to the electroweak correction, events are generated both with the radiation and without, and the ratio of the Born level to dressed level is used as the bin by bin correction. The effect is shown for the charge insensitive observables in Figures 9.11 - 9.15. The impact is very small (generally less than 1%) and was not applied to the predictions in favor of not adding an additional systematic uncertainty. The same smoothing using for the non-perturbative corrections is used here to smooth the low statistics bins.



(a) Exclusive Jet Multiplicity

(b) Inclusive Jet Multiplicity

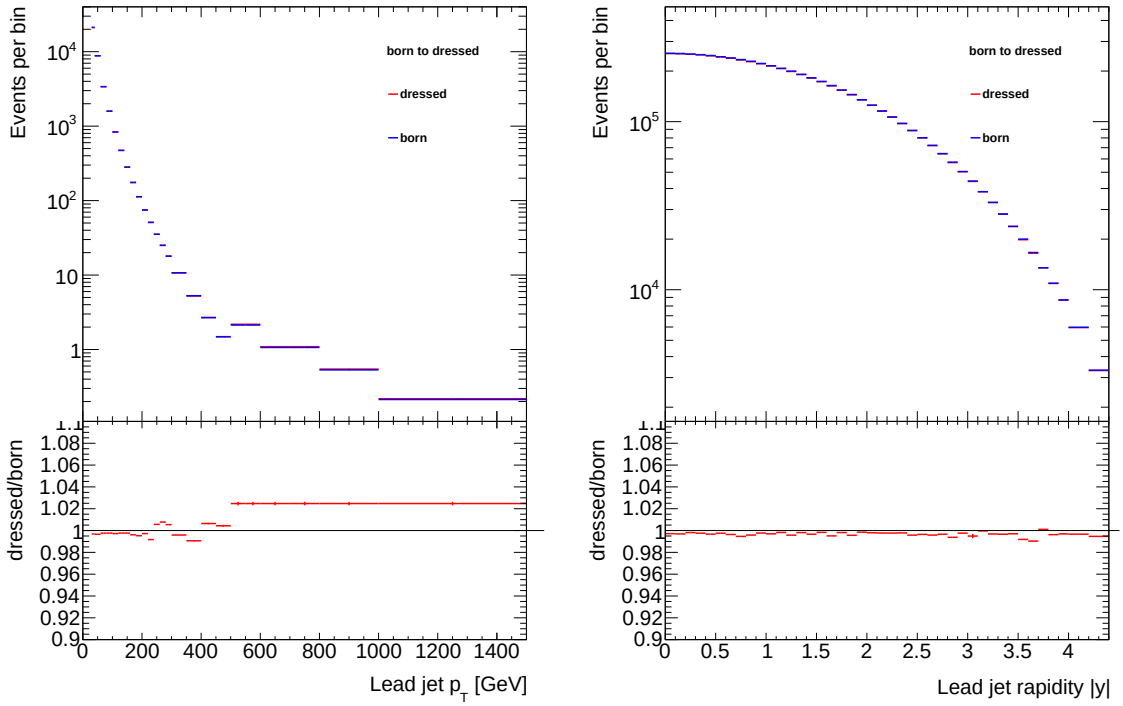
Figure 9.11: The Born to dressed level corrections for the inclusive and exclusive jet multiplicity generated with SHERPA 2.2.1 using the cuts from Table 7.1



(a) $H_{\mathrm{T}} \geq 1$ jet

(b) W $p_{\mathrm{T}} \geq 1$ jet

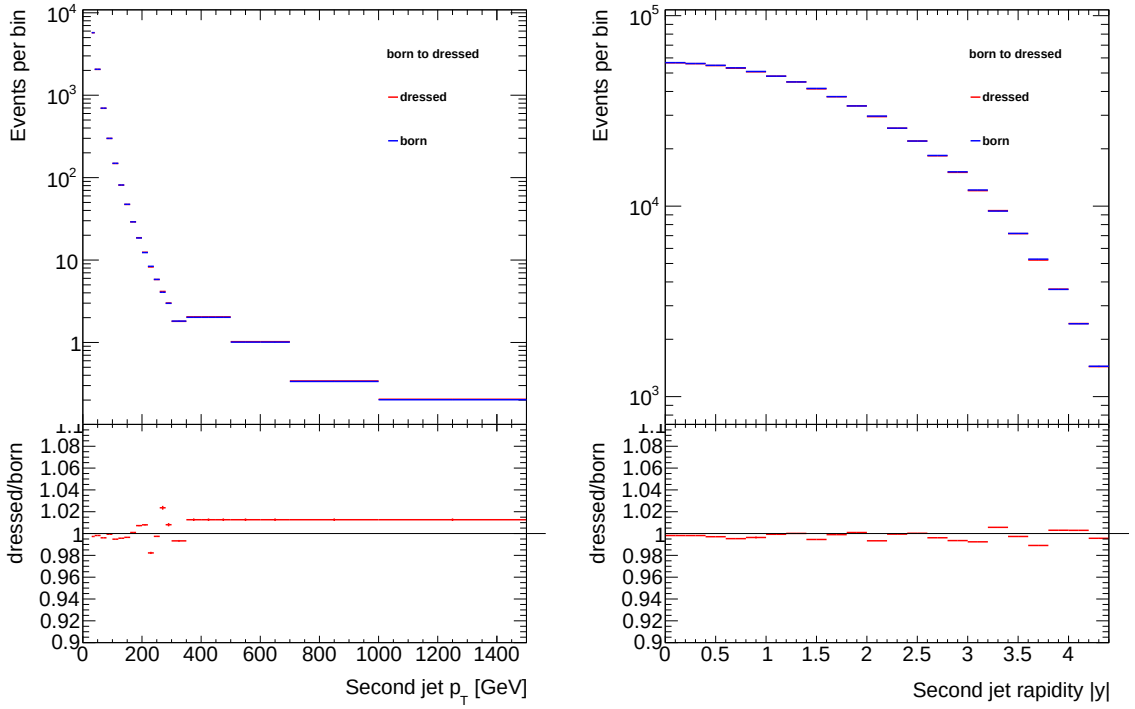
Figure 9.12: The Born to dressed level corrections for H_{T} and W boson transverse momentum generated with SHERPA 2.2.1 using the cuts from Table 7.1



(a) Lead jet $p_{\mathrm{T}} \geq 1$ jet

(b) Lead jet rapidity ≥ 1 jet

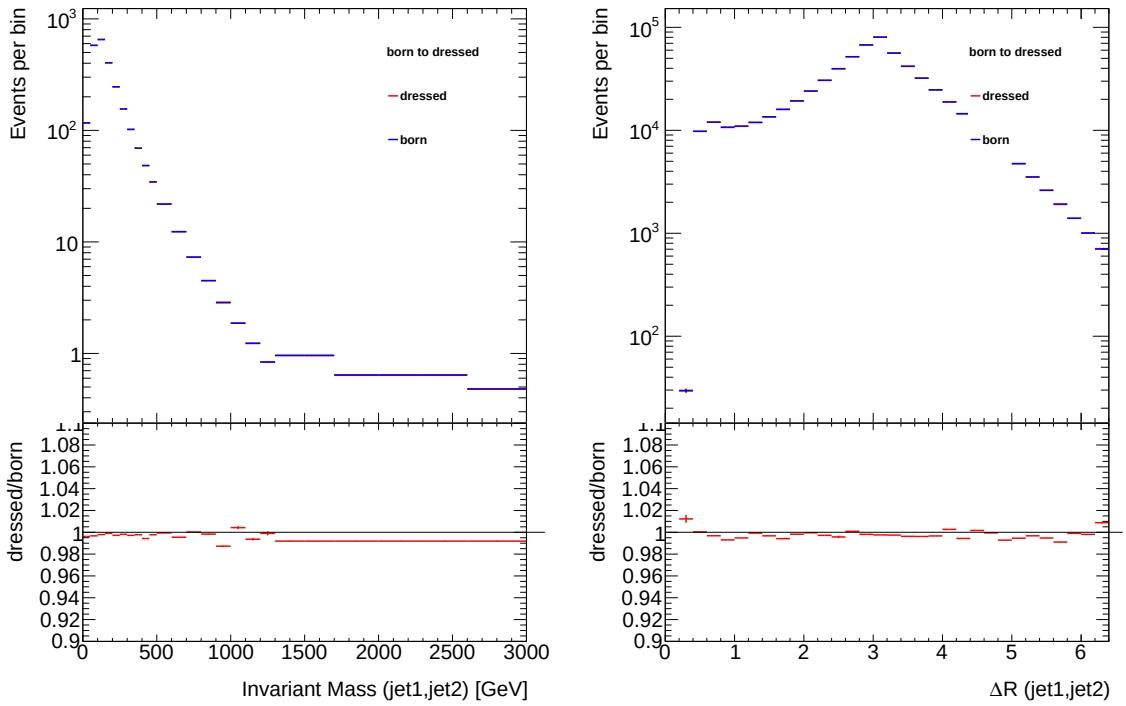
Figure 9.13: The Born to dressed level corrections for lead jet transverse momentum and rapidity generated with SHERPA 2.2.1 using the cuts from Table 7.1



(a) Second leading jet $p_{\mathrm{T}} \geq 2$ jet

(b) Second leading jet rapidity ≥ 2 jet

Figure 9.14: The Born to dressed level corrections for second lead jet transverse momentum and rapidity generated with SHERPA 2.2.1 using the cuts from Table 7.1



(a) Invariant mass of two lead jets $\geq$ 2 jet

(b) Angular separation $\Delta R_{\text{jet1,jet2}}$

Figure 9.15: The Born to dressed level corrections for dijet invariant mass and the angular separation between the two lead jets generated with SHERPA 2.2.1 using the cuts from Table 7.1

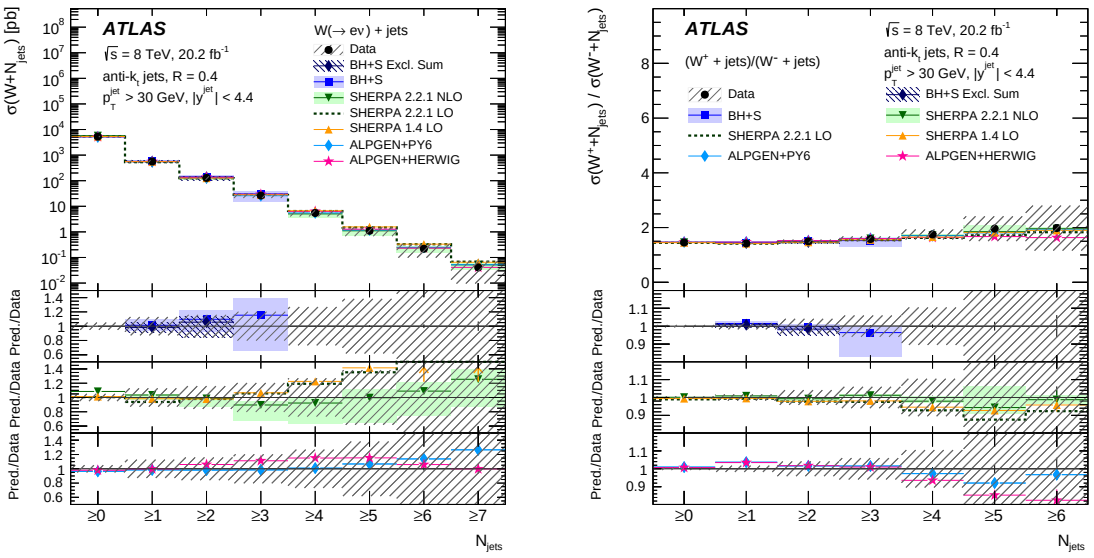
Chapter 10

Cross Section Results

This chapter presents the measured cross sections for $W \to e\nu$ production and the ratio of the W^+ to W^- cross sections, calculated from the W^+ and W^- cross sections measured separately. For the W^+ and W^- distributions themselves, see Appendix C. The data are compared to the theoretical predictions outlined in Chapter 9. The observables discussed are separated into the jet multiplicity distributions, the inclusive $W+ \geq 1$ jet distributions, and the inclusive $W+ \geq 2$ jets distributions.

10.1 Jet multiplicity distribution

The inclusive jet multiplicity cross sections are shown in Figure 10.1 for inclusive W production and the W^+/W^- ratio. The fixed order calculations are compared to each other; BlackHat+Sherpa using the CT10 PDF set is NLO for 1, 2 and 3 jets, and N_{jetti} using the CT14 PDF set is NNLO for 1 jet. Table 9.1 can be reviewed to reference the order in α_s, the PDF used, and the number of partons calculated in the matrix element for all the predictions shown. The fixed order calculations demonstrate the current understanding of perturbative QCD, and have the lowest theoretical uncertainties. Overall there is agreement between the predictions and the data. Both LO Sherpa predictions start to diverge at high jet multiplicity but the NLO Sherpa predictions provide a much better description of the data. There are two parton shower models shown for Alpgen, both of which provide



(a) (b)

Figure 10.1: Cross section for the production of W bosons (left) and the W+/W- ratio (right) for different inclusive jet multiplicities. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].

consistent predictions within experimental uncertainties. The theoretical uncertainties for the LO predictions (which are large due to the LO nature of the calculation) are not shown. In general, the trends for W^+ and W^- production are the same. Taking the ratio improves agreement between data and prediction indicating there is cancellation of some systematic effects, both theoretical and experimental. For example, scale uncertainties will mostly cancel in the ratio.

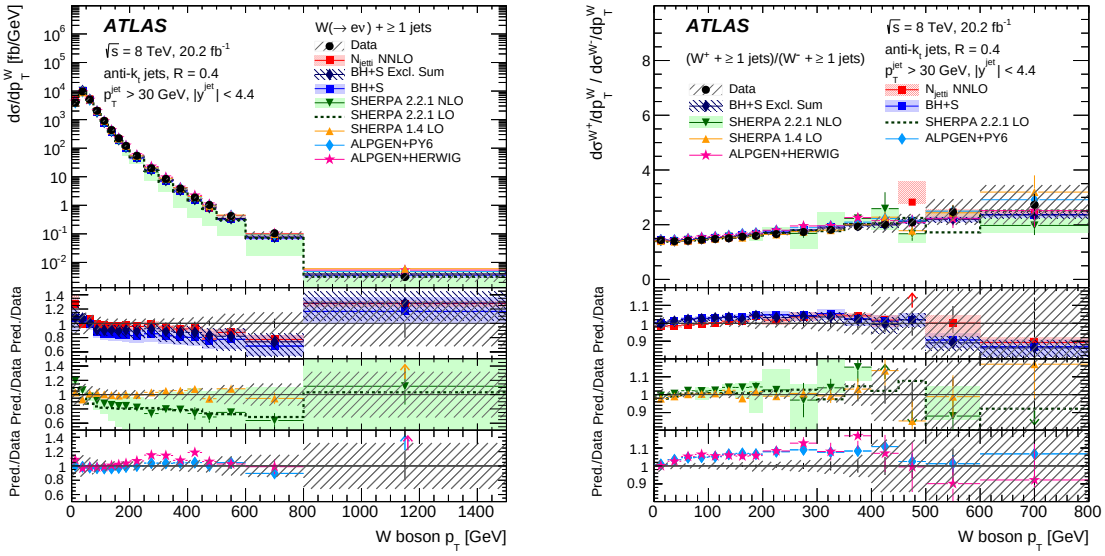
10.2 Distributions for $N_{\text{jets}} \geq 1$

Figure 10.2 shows the differential cross section of W production and the W^+/W^- ratio for $N_{\text{jets}} \geq 1$ as a function of the p_{T} of the W boson. This is one of the more difficult distributions to model, as at high p_{T} the electroweak corrections become significant (see Section 9.6). Sherpa 2.2.1 does not do as well as 1.4.1 at either LO or NLO, but it should be noted that Sherpa 1.4.1 includes up to 4 partons in the matrix element calculation, and Sherpa 2.2.1 at LO includes up to 3 partons, while Sherpa 2.2.1 at NLO includes only up to 2 partons. Sherpa 2.2.1 does benefit by having smaller theoretical uncertainty, and as noted before, the theory uncertainties at LO are typically much larger. Thus, the better agreement at LO can not be considered that significant. For the fixed order predictions, N_{jetti} does much better than BlackHat+Sherpa. This is unsurprising since at high W p_{T}, dijet production with a W emission becomes dominant, which is described better by N_{jetti}, being NNLO in inclusive $W+1$ jet production. The W^+/W^- ratio has much greater experimental precision. Most predictions with the exception of N_{jetti} and Sherpa 1.4.1 overestimate the cross section by one to nearly four times the uncertainty. Alpgen has the greatest overestimation, which is consistent with the mismodeling seen in Figure 10.1b.

The distributions of the p_{T} of the lead jet for events with $N_{\mathrm{jets}} \geq 1$ is shown in Figure 10.3. The N_{jetti}, ALPGEN, and LO SHERPA 1.4.1 describe well both the W production and the W^+/W^- ratio. SHERPA 2.2.1 LO and NLO, as well as the BLACKHAT+SHERPA predictions tend to underestimate the cross section. This is in contrast to what was seen in W+dijet production [76] where the predictions overestimated the cross section. In the W+dijet analysis, the event selection required a larger leading jet p_{T} as well as a dijet invariant mass of at least 500 GeV. This shows how different phase spaces can effect the predictions.

The distribution of lead jet rapidity is shown in Figure 10.4. At forward rapidity, there is a sharp drop off in the data at $|y| \approx 3.6$, compared to the smoothly falling predictions. In this region there is also a large experimental uncertainty from the uncertainty due to Monte Carlo bias in the unfolding, which stems from a difference in how ALPGEN and SHERPA model this region. This combined with the large uncertainty in JES and JER as well, cover the discrepancy between data and the theoretical predictions (see Chapter 8 for a description of systematic uncertainties). Overall, most predictions overestimate the cross section in the forward region but remain within 1-2 times the experimental uncertainty. The showering model seems to have a significant effect which can seen in the comparison of ALPGEN+HERWIG to ALPGEN+PYTHIA 6. This mismodeling in the forward region largely cancels out in the W^+/W^- ratio.

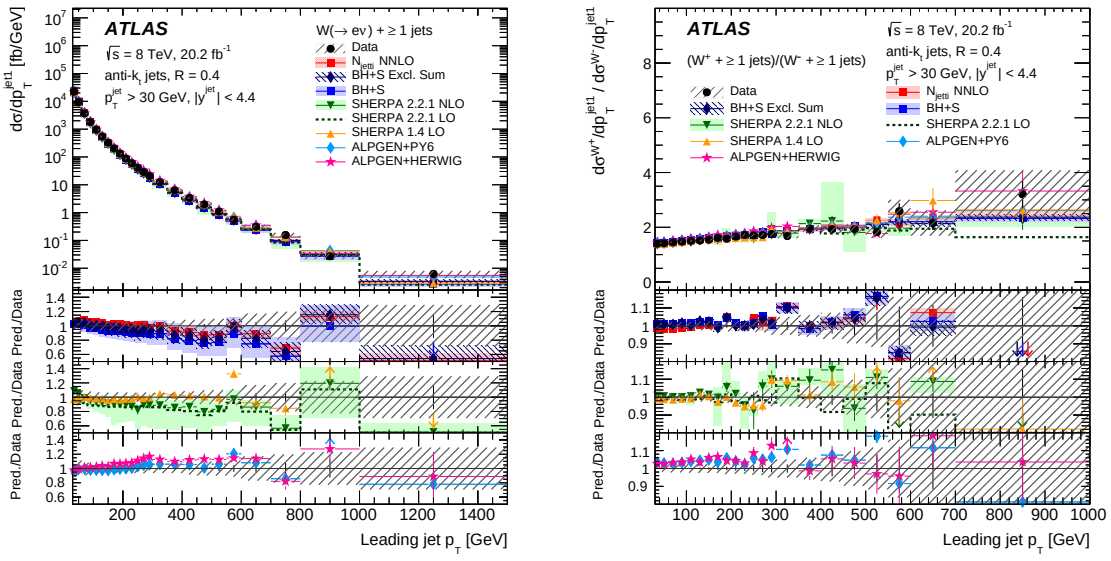
The differential cross section for W production and the W^+/W^- ratio for $N_{\mathrm{jets}} \geq 1$ is shown as a function of H_{T} in Figure 10.5. The leading order predictions, SHERPA and ALPGEN describe the data best. These predictions benefit from having multiple partons in the matrix element calculation (see Table 9.1), which is important especially in the H_{T} distribution. In general higher energies are more sensitive to the higher multiplicity matrix



(a) (b)

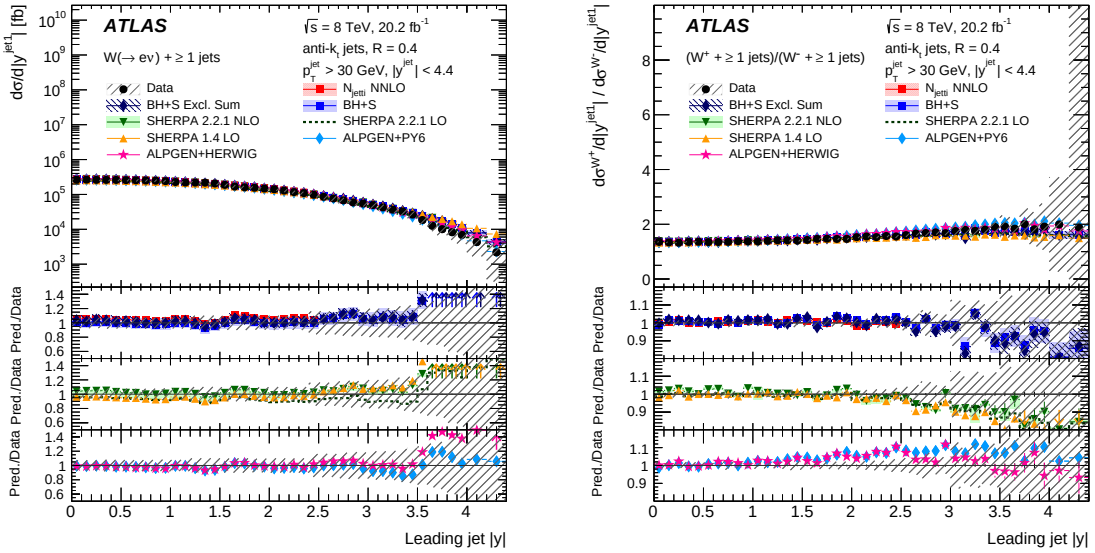
Figure 10.2: Differential cross sections for the production of W bosons (left) and the W^+/W^- ratio (right) as a function of the transverse momentum of the W boson for events with $N_{\text{jets}} \geq 1$. The last bin in the left figure includes values beyond the shown range. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].

element contributions. The BlackHat+Sherpa predictions underestimate the data at high H_{T}, which is expected since the additional jets are important in this region. The BlackHat+Sherpa using the exclusive sum approach and the N_{jetti} prediction both do better benefiting from higher order jet corrections in the matrix element calculation, either approximate for the case of the exclusive sum, or exact for the case of N_{jetti}. Statistics decrease at larger values of H_{T} where the cross section is small and so the uncertainty increases. In the ratio, systematic uncertainties also increase in the same region where there is imperfect cancellation.



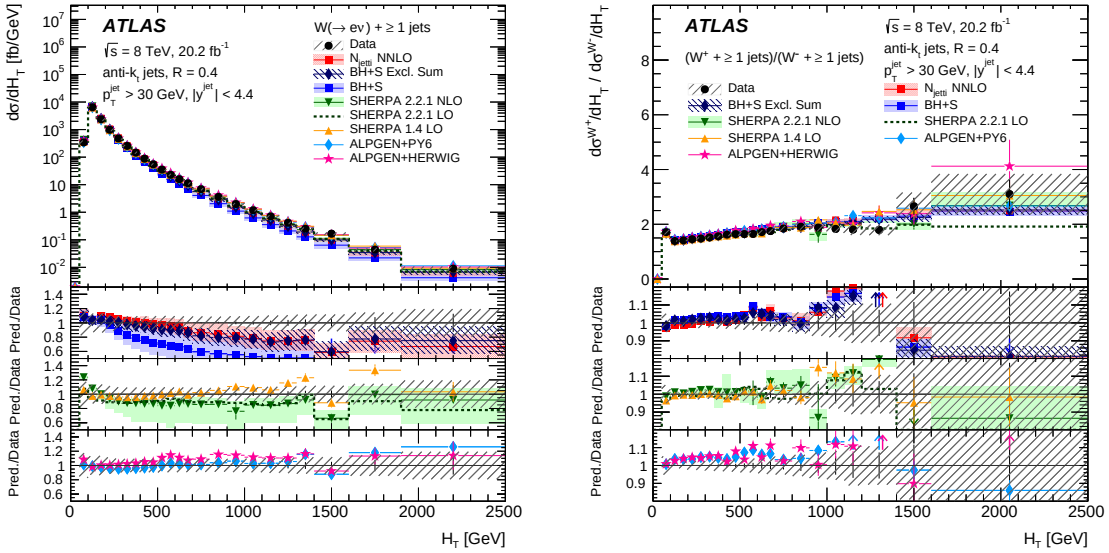
(a) (b)

Figure 10.3: Differential cross sections for the production of W bosons (left) and the W^+/W^- ratio (right) as a function of the lead jet p_{T} for events with $N_{\mathrm{jets}} \geq 1$. The last bin in the left figure includes values beyond the shown range. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



(a) (b)

Figure 10.4: Differential cross sections for the production of W bosons (left) and the W^+/W^- ratio (right) as a function of the lead jet rapidity for events with $N_{\text{jets}} \geq 1$. The last bin in the left figure includes values beyond the shown range. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



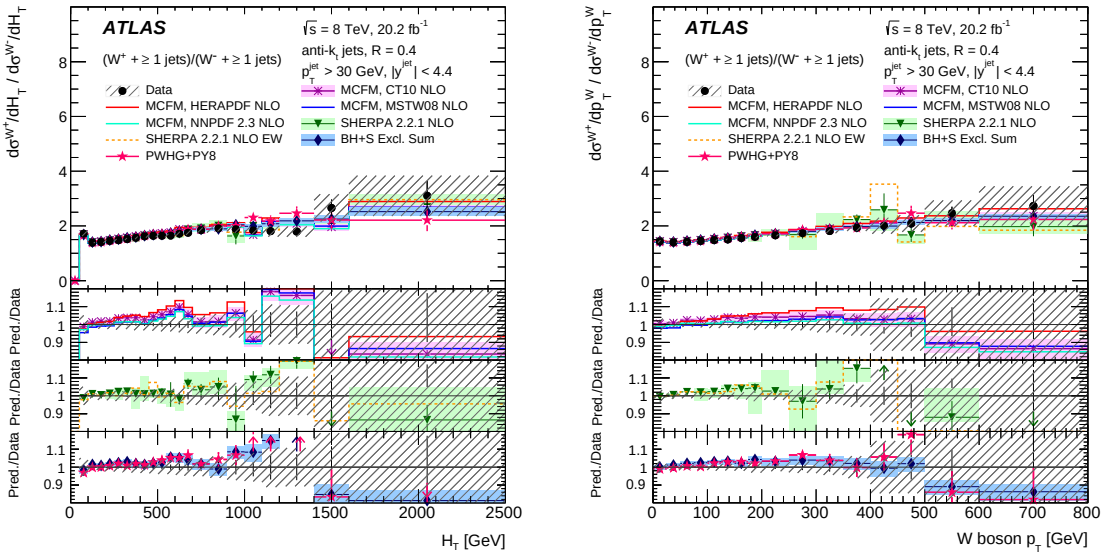
(a) (b)

Figure 10.5: Differential cross sections for the production of W bosons (left) and the W^+/W^- ratio (right) as a function of H_{T} for events with $N_{\mathrm{jets}} \geq 1$. The last bin in the left figure includes values beyond the shown range. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].

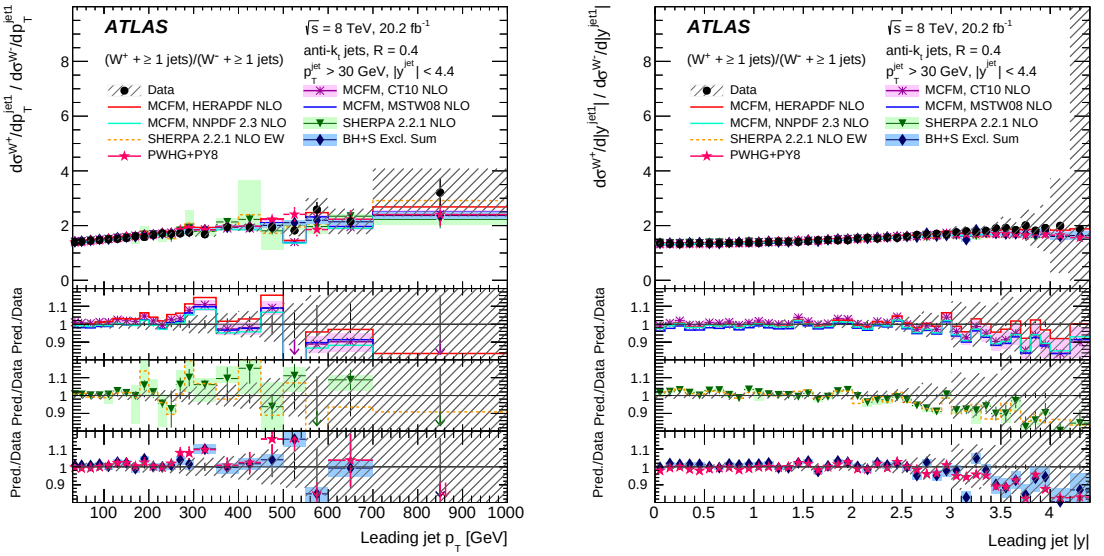
Figure 10.6 shows the differential cross sections of the W^+/W^- ratio for the above four observables again but compared to the MCFM predictions with four different PDF sets. The PDF sets compared are CT10, HERAPDF 1.4, MSTW 2008, and NNPDF 2.3, which are all NLO PDF sets (MCFM is also NLO). The PDF uncertainties shown are for the CT10 PDF set. As can be seen from the comparisons, the PDF set can have a non-trivial effect on the prediction. The most significant effects are for the p_{T} of the W boson, and in the forward region of the jet rapidity.

Figure 10.6 also shows the predictions from Powheg+Pythia 8, the BlackHat+Sherpa exclusive sum (both using the CT10 PDF set), and Sherpa 2.2.1 with and without the electroweak corrections using the PDF set NNPDF 3.0 NLO. The predictions from Powheg+Pythia8 are comparable to the BlackHat+Sherpa exclusive sum method. The effect of the second jet emission being calculated at NLO in BlackHat+Sherpa is similar in magnitude (but smaller in uncertainty) to the correction provided by the parton showering in Pythia 8.

Also included in Figure 10.6 are the Sherpa 2.2.1 predictions with the NLO electroweak corrections applied (these corrections are not applied in the other plots). Here in the ratio, the EW corrections should cancel, but the effect on the W^+ and W^- distributions is also shown individually in Figures C.7-C.14 of Appendix C. The electroweak corrections scale larger with W boson and jet p_{T}, and H_{T} observables. The corrections can be quite large (see Section 9.6) and have the effect of reducing the cross section such that there is larger disagreement with the data. The electroweak correction is as large as 30-80% depending on the observable, and largely cancels out with the NLO correction from α_s.

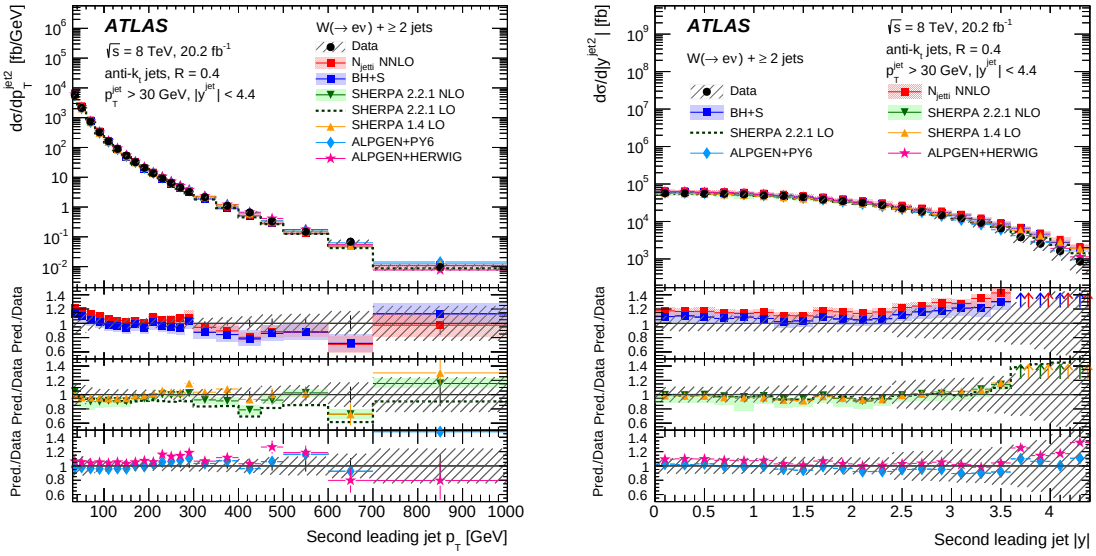


(a) (b)



(c) (d)

Figure 10.6: W^+ to W^- cross-section ratio as a function of H_{T} (top left), W p_{T} (top right), leading jet p_{T} (bottom left) and leading jet rapidity (bottom right) for events with $N_{\mathrm{jets}} \geq 1$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



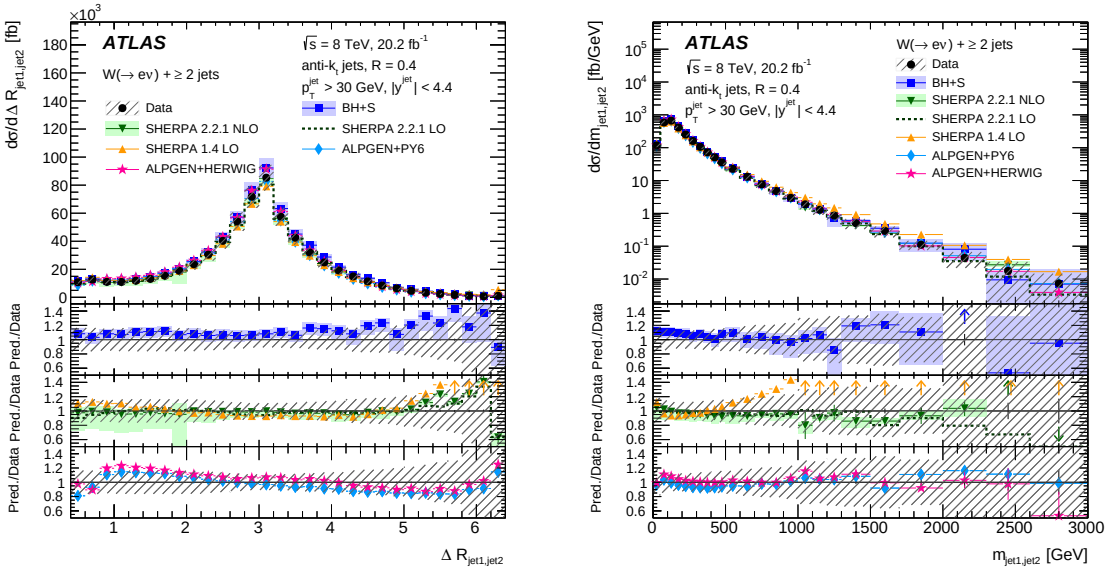
(a) (b)

Figure 10.7: Differential cross sections for the production of W bosons as a function of the second lead jet p_{T} (left) and rapidity (right) for events with $N_{\mathrm{jets}} \geq 2$. The last bin in the left figure includes values beyond the shown range. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].

10.3 Distributions for $N_{\mathrm{jets}} \geq 2$

The differential cross section for events with at least two jets is shown as a function of second leading jet p_{T} and second leading jet rapidity in Figure 10.7. The N_{jetti} prediction is NNLO for the first jet, and contains NLO matrix elements for the second jet. As such, the BlackHat+Sherpa predictions, which are also NLO for the second jet, and the N_{jetti} predictions contain a similar level of agreement with respect to data. The second lead jet is well modeled by all predictions up to a rapidity of $|y| \approx 2.5$, after which the predictions, with the exception of Alpgen tend to over predict the data.

Figure 10.8 shows the differential cross section for $W + 2$ jet production as a function of the angular separation between the two lead jets and the invariant mass of the two lead jets. These observables are sensitive to hard parton radiation at large angles and the matrix element/parton shower matching schemes of the predictions. BlackHat+Sherpa describes both distributions well even at large dijet mass, only slightly overpredicting the data at low invariant mass. This leads to a small offset in the ΔR angular separation, since it is dominated by low dijet invariant mass. The Sherpa 1.4.1 generator predicts too large a cross section at large angular separation and high dijet invariant mass. Sherpa 2.2.1 provides a better description at both LO and NLO, especially for the dijet invariant mass. The Alpgen prediction describes the large dijet invariant mass well, but less so for small and large angular separations between the two lead jets. Additional distributions are provided in Appendix C



(a) (b)

Figure 10.8: Differential cross sections for the production of W bosons as a function of the angular separation between the two lead jets and dijet mass of the two lead jets for events with $N_{\text{jets}} \geq 2$. The last bin in the left figure includes values beyond the shown range. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].

Chapter 11

Conclusion

This thesis has presented the cross section measurement of W boson production accompanied by jets as well as the W^+ to W^- ratio of those cross sections for the data collected by ATLAS in 2012. During this time, the center of mass energy for the proton-proton collisions was $\sqrt{s} = 8\,\text{TeV}$ and the total integrated luminosity of this run was $20.2\,\text{fb}^{-1}$. The differential distributions shown are for observables that are sensitive to perturbative QCD, the modeling of the parton shower, and parton structure of the proton. Taking the ratio of W^+ to W^-production has the advantage that many systematic and theoretical uncertainties cancel out, allowing for high precision measurements.

What was observed was that higher order predictions are able to describe data well, and have improved accuracy and uncertainty in certain regions of phase space. However these fixed order calculations often over or underestimate the data in regions of high transverse momentum, high rapidity, or large angular separation. This is due to the lack of jets in these regions which are needed to properly describe the data. Often the LO generators, which include many partons in the matrix element, model the data better in these regions thanks to the inclusion of additional jets. This comes at the cost of having larger theoretical uncertainties, which makes definite conclusions difficult. None of the generators describe all observables well, but many describe certain observables better than others. The hardest observables to model are the H_{T}, jet rapidity, and dijet invariant mass, suggesting that a

better description of high energy and high rapidity jets is needed. In the W^+/W^- ratio, increased sensitivity to the PDF is observed. In some cases the choice of PDF can modify the prediction by as much as the experimental uncertainty. This indicates that the data can be useful in providing additional information for parton distribution function fits, and in fact, this data is being used thusly within the ATLAS collaboration. These distributions help provide a better understanding of perturbative QCD, and the impact on parton distribution functions. They can also serve as a reference for other analyses and future studies of W production in association with jets.

APPENDICES

APPENDIX A

Cosmic Muon Analysis

Introduction

In order to earn authorship in the ATLAS collaboration, one must complete a service project that contributes to maintenance or upgrade of the detector. This service project involved using cosmic ray muons to help calibrate the hadronic calorimeter, TileCal. Cosmic ray data has been used in the past to cross check the calibration as well to inter-calibrate the parts of the TileCal. The goal of this study is to reproduce the distributions made in the past for the 2008, 2009, and 2010, using the 2011 and 2012 data for side-by-side comparison.

Event selection

Muon trajectories were constructed from the Inner Detector using tracking information from the silicon detectors (SCT and Pixel). The quality of the extrapolation is ensured by requiring at least eight hits between the two detectors. A cut on the impact parameters is also used to select paths with semi-projective geometry: the transverse component $d_0 \leq 380$ mm and longitudinal $z_0 \leq 800$ mm.

In order to reduce the effect of muon radiative energy losses in the detector, an upper limit of 30 GeV is placed on the momentum of the muon. Additionally, to reduce errors from multiple scattering, a lower limit on the momentum is placed at 10 GeV, as well as a cut on

the azimuthal angle of the path to be within the volume of the module (all of which have a half width of $\Delta\phi = 0.049$):

$$|\phi_{track} - \phi_{cell}| < 0.045 \tag{A.1}$$

The track path length dx is defined as the distance of the extrapolated track between where it enters and where it exits a cell. In this analysis, only cells with a path length $dx > 20$ cm are considered. To reduce noise contribution to the signal, cells are required to have an energy of at least 60MeV.

Vertical tracks in the TileCal are poorly measured due to the vertical orientation of the scintillating tiles (see Section 3.2.3). As such, it is required that the polar angle of the track be greater than 0.13 radians measured from the vertical.

Cell energy response

Figure A.1 shows the cell response as a function of azimuthal angle phi for several of the individual modules in the bottom of the detector in the BC layer. The response corresponding to each module is shown in a different color and marker style and the total energy response is superimposed as well in black with full circles. That is to say, a muon passing through the detector at some angle, ϕ, may deposit energy in more than one module. For a reasonably straight muon, most of the energy should be deposited in the module of the corresponding ϕ value, but some energy will bleed into adjacent cells. More importantly, this distribution should be uniform across modules, which is the case. All the requirements for event selection mentioned in section A Event selection are applied here except for equation A.1 and the additional requirement $|d_0| < 100$ mm and $|z_0| < 300$ mm were applied in order to obtain a more projective track geometry.

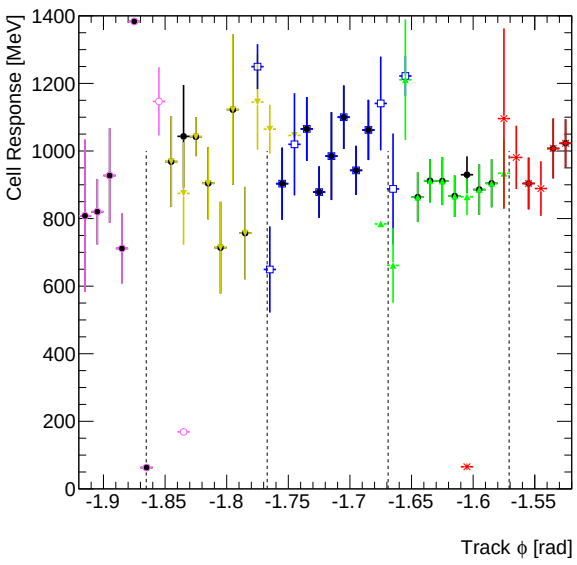


Figure A.1: The cell response in the BC layer as a function of ϕ in the central bottom region of the calorimeter. The average cell response for each module is shown in a unique color and marker style, and the total response is shown in full black circles. The nominal edges of the modules are indicated by the dashed black lines

The cell response as a function of path length in BC layer can be found in Figure A.2. Here, the requirement that the path length through the cell must be greater than 20 cm is ignored. The gray is a scatter plot with each point corresponding to a cell in a single event. The red circles indicate the average cell response for a particular path length bin and the solid black line indicates the linear best fit of the average. The large number of events at the edge of 840 mm is a relic of the radial length of the BC layer; since most cosmic muons are vertical, many tend to travel trough the entire radial length of the cells. The fact that cell response scales linearly with path length tells us that the slope dE/dx is a good quantity for studying inter-calibration of the detector.

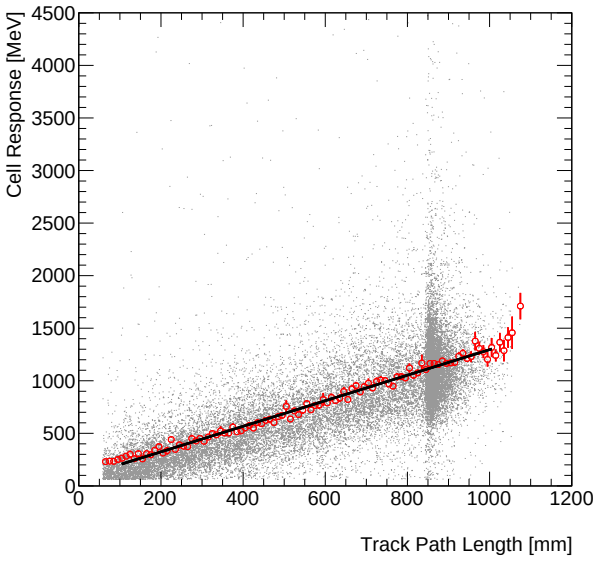


Figure A.2: Mean energy response as a function of path length in the BC layer. Red solid line indicates linear best fit. Large statistics at 840 mm is a result of most muons crossing the detector vertically and traveling the full radial length of the layer.

Results

Uniformity in cells

The uniformity in cell response is shown in Figure A.3 for each radial layer. For each cell that meets the requirements, the truncated mean is evaluated and recorded across the given layer. Because there is a 1% truncated mean as the estimator for this study, it is required that each cell have at least 100 events in order to contribute to the figure. This severely limits the number of measured cells that contribute, but still there are enough statistics to at least say there is consistency between each of the layers.

Uniformity in pseudorapidity

To observe the uniformity of response in pseudorapidity, the normalized truncated mean is calculated for each η bin. Results shown in Figure A.4 are separated by layer. This is useful

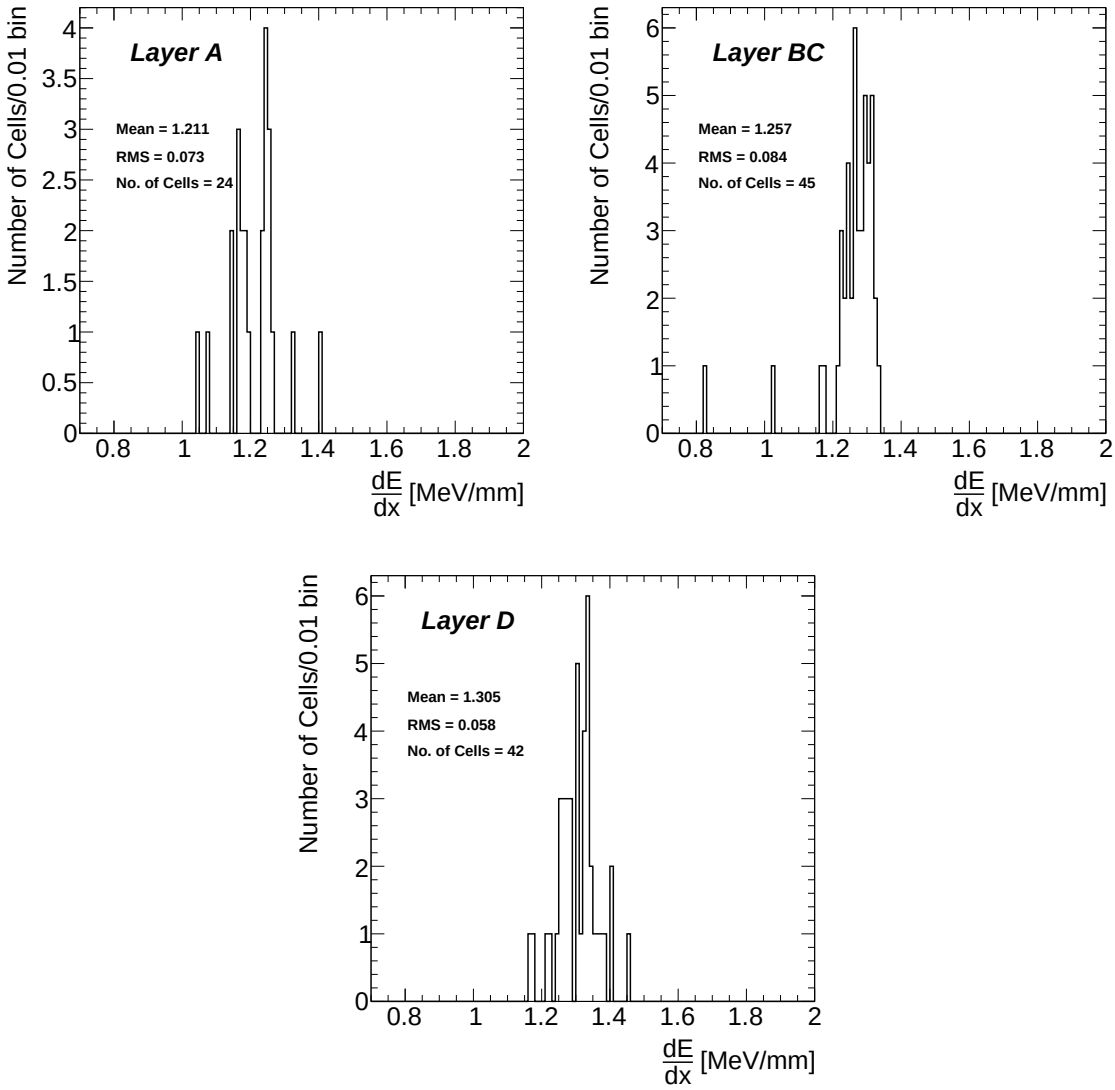


Figure A.3: Distribution of the truncated mean of dE/dx of the cells in each of the separate layers. At least 100 events per cell are required in order for the the cell to contribute to the distribution.

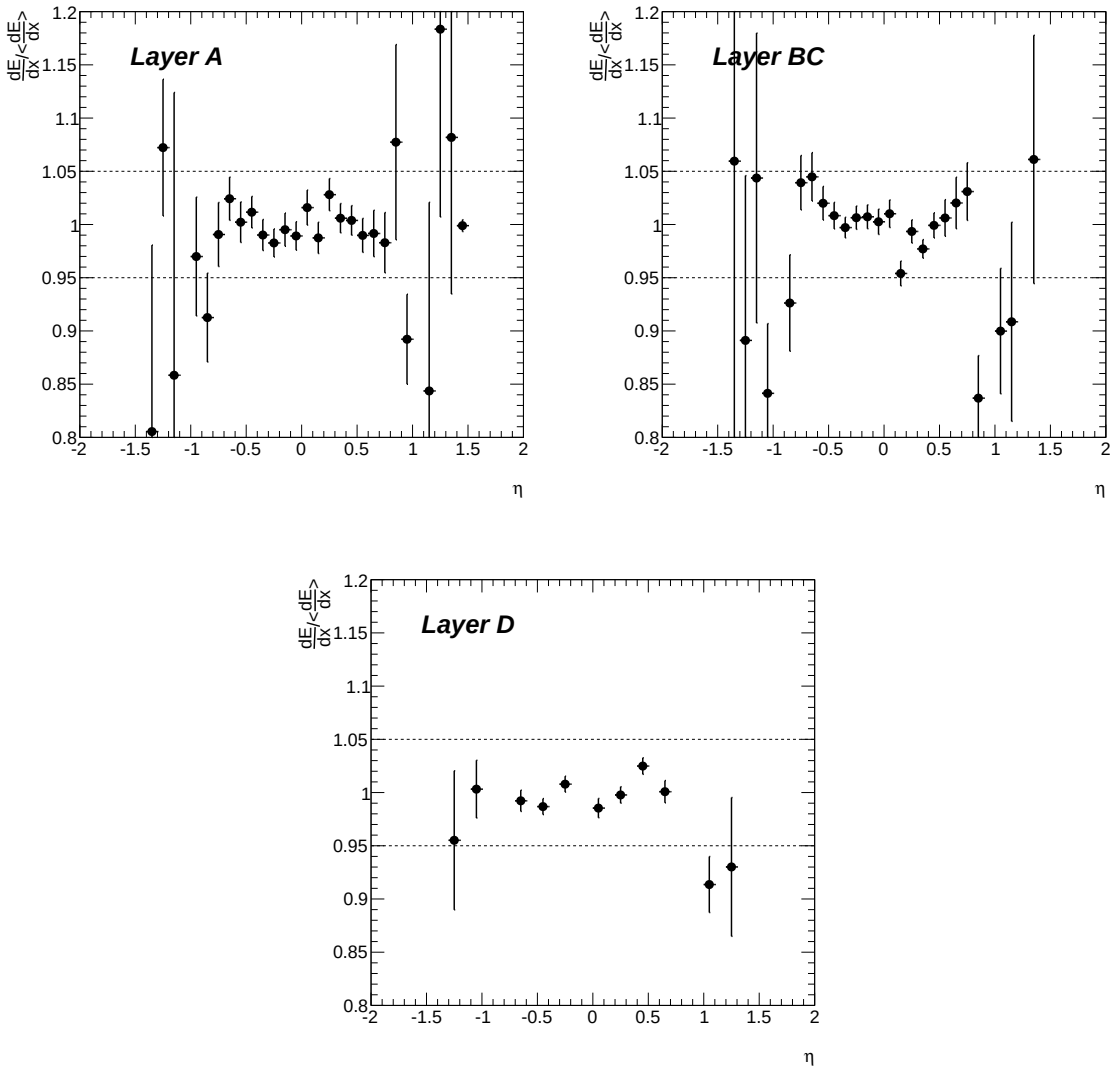


Figure A.4: Uniformity of the normalized truncated mean of dE/dx as a function of of pseudorapidity η in each of the radial layers. Statistics are lower in large $|\eta|$ due to vertical orientation of most cosmic ray muons.

since muons cross each layer at a slightly different η-coordinate as a result of cosmic rays not being projective muons. Most cosmic muons tend to cross the detector vertically especially from angles coming from the access shafts (see Figure 3.1), so statistics are best in the long barrel, where variations are generally within 5% within each layer.

Uniformity in modules

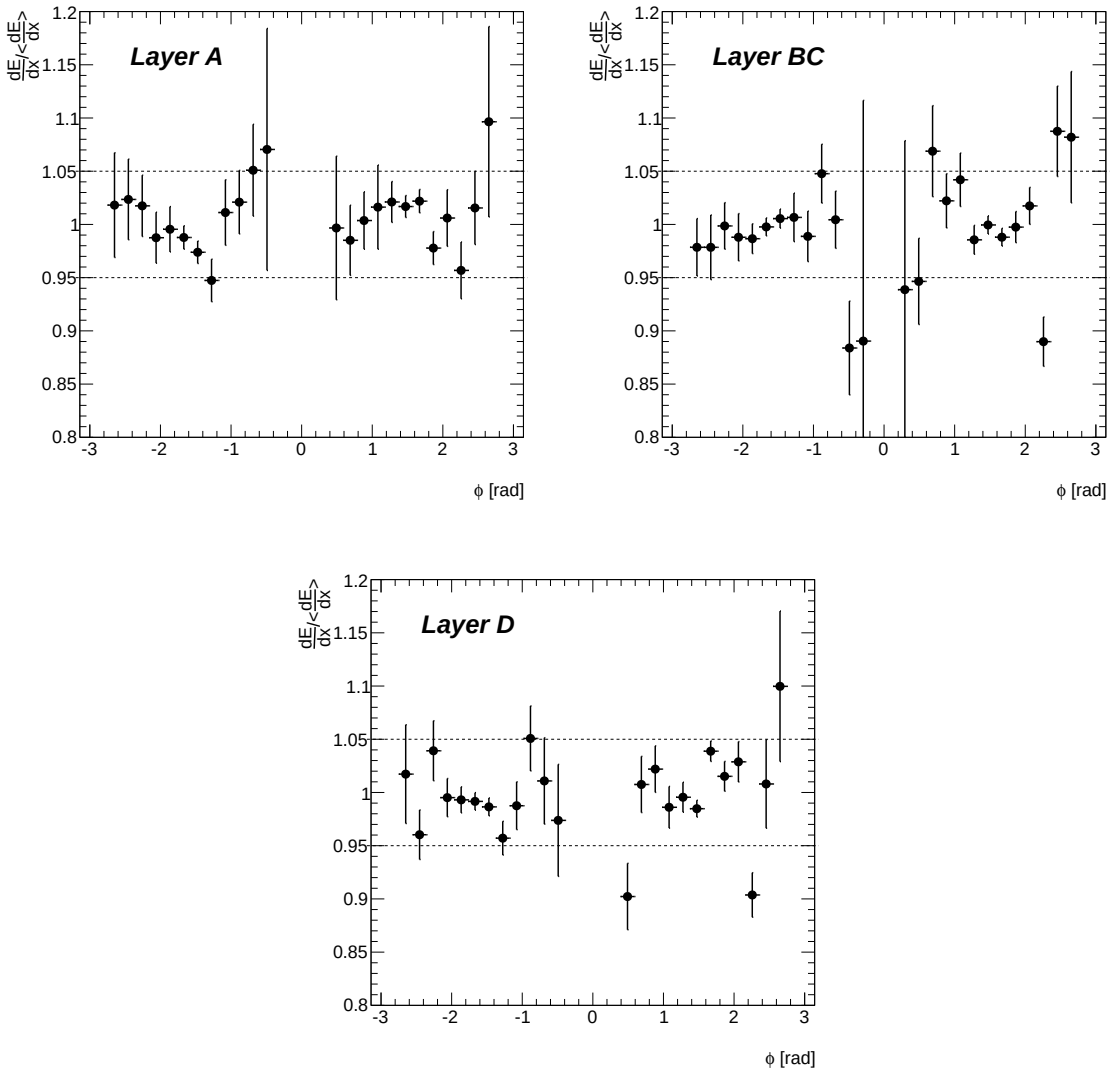


Figure A.5: Uniformity of the normalized truncated mean of dE/dx as a function of azimuthal angle ϕ in each of the radial layers. The gap around $\phi = 0$ is from the lack of horizontal cosmic rays.

As was done for pseudorapidity, the normalized truncated mean for each ϕ bin was also calculated. The results are shown in Figure A.5. All the same cuts were applied as mentioned in Section A. Large variations near the horizontal modules around $\phi = 0, \phi = \pm\pi$ is likely due to both poor statistics in the extended barrel as well as a difference in muon momentum in events passing the selection criteria. That said, variance within the order of 5% for the more vertically oriented modules, was seen.

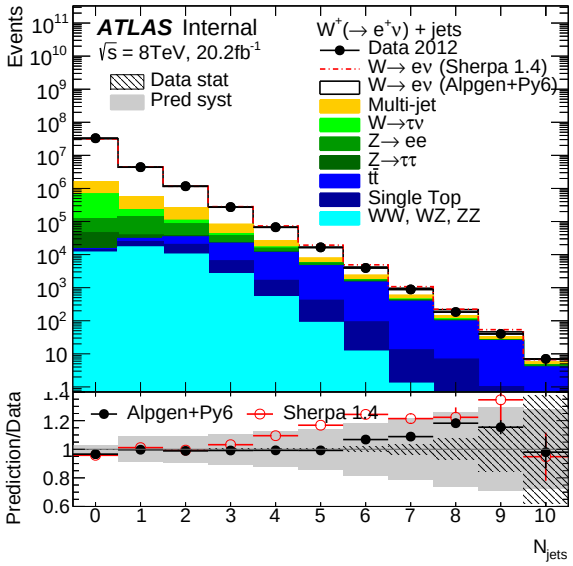
Conclusion

The statistics for the 2011 and 2012 data are much worse than in previous studies. There were 100k events that passed event selection in 2008 cosmic rays, where as in the 2011 and 2012 combined study there were only 10k. That said, the statement can still be made, that the variation in the central regions of the detector where statistics are generally stronger are generally within 5%. Additionally, for the plots shown, there is general consistency with the those made in the past.

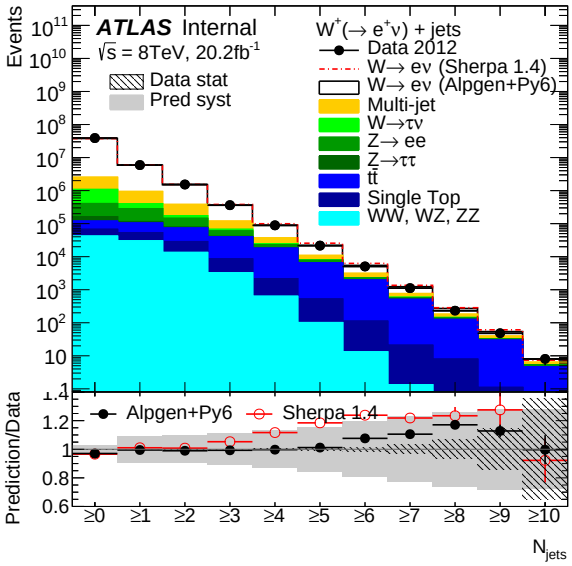
APPENDIX B

Detector level distributions for W^+ and W^-

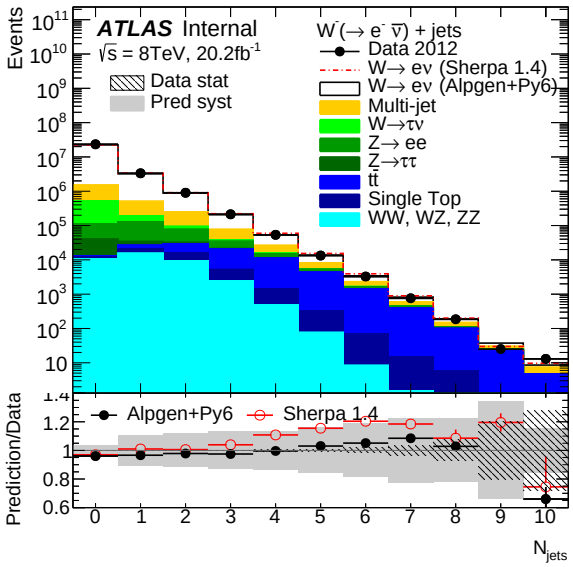
Below are the detector level distributions for W^+ and W^- selections. The corresponding plots for the charge indiscriminant selection is in Chapter 6. Figure B.1 shows the exclusive and inclusive jet multiplicities for W^+ and W^-. The W-related kinematic distributions for the W^+ are in Figure B.2 and for the W^- in Figure B.3. The W^+-related kinematics for 1 jet inclusive events are shown in Figure B.4 and the jet-related kinematics are in Figure B.5. Similarly, the W^--related kinematics for 1 jet inclusive events are shown in Figure B.6 and the jet-related kinematics are in Figure B.7. Lastly, the jet-related kinematics for the 2 jet inclusive events for the W^+ are shown in Figure B.8 and for the W^- in Figure B.9.



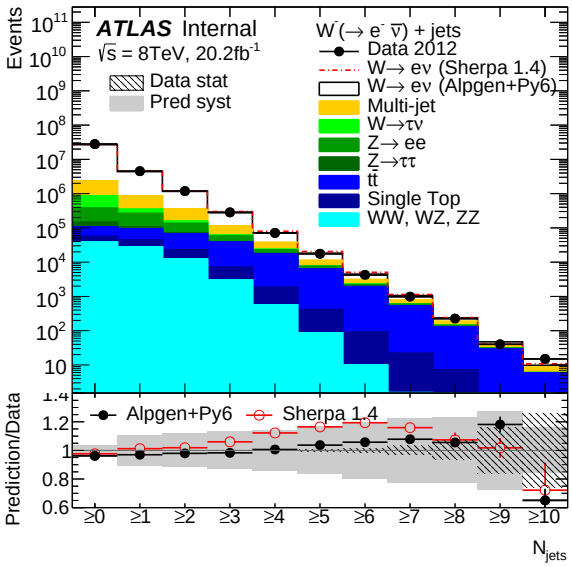
(a) W^+, exclusive jet multiplicity



(b) W^+, inclusive jet multiplicity

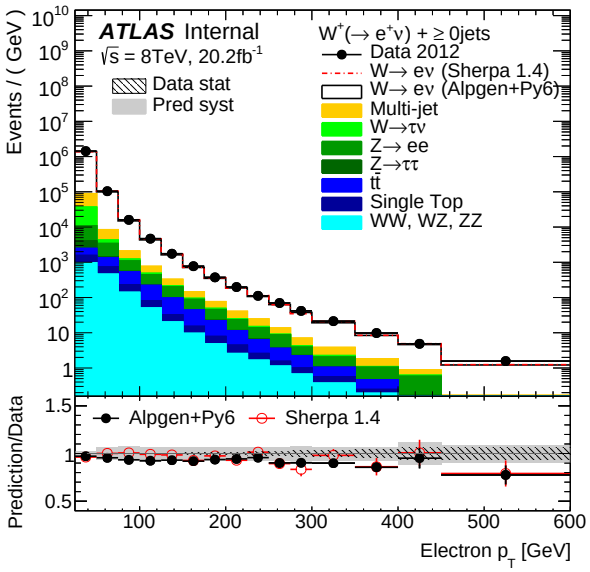


(c) W^-, exclusive jet multiplicity

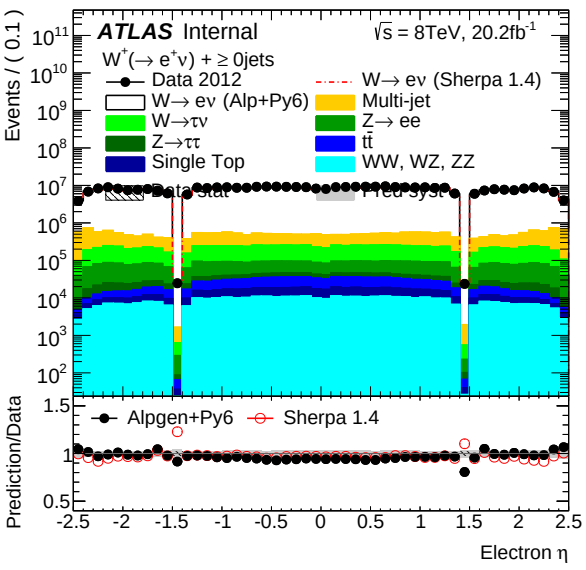


(d) W^-, inclusive jet multiplicity

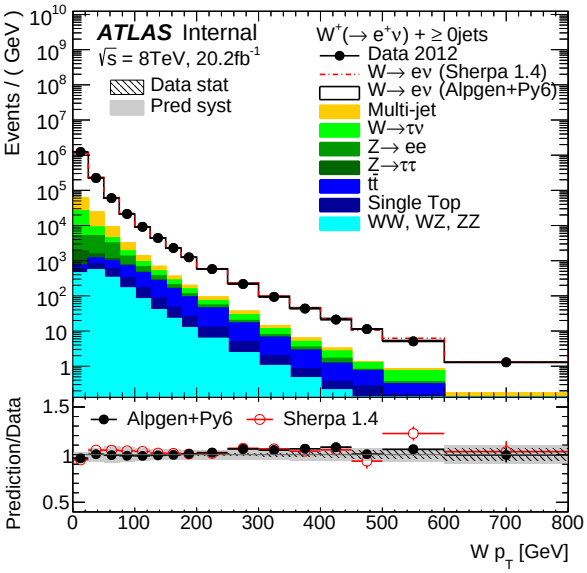
Figure B.1: Jet multiplicity distribution in the W+jets signal region for the W^+ (top) and W^- (bottom) selections, exclusive (left) and inclusive (right) in the number of jets, for the SM prediction compared to data at the detector level. The lower panel and the uncertainties are as in 6.1.



(a) Positron p_{T}

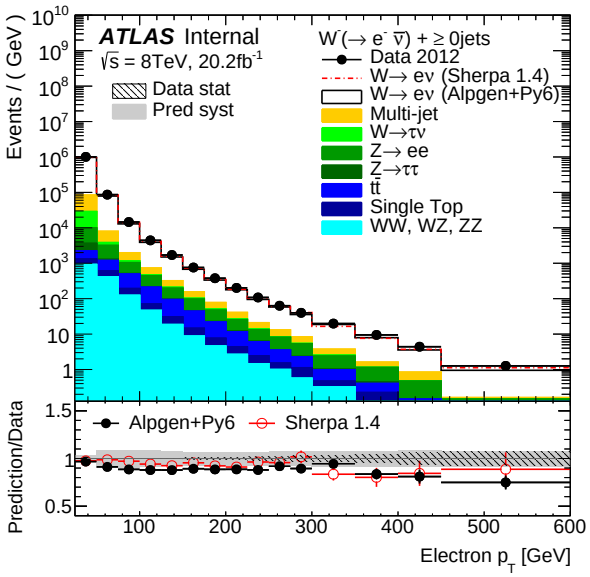


(b) Positron η

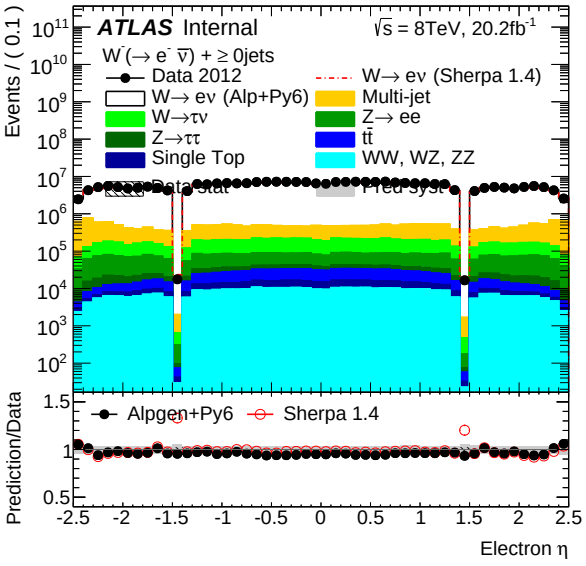


(c) W^+ p_{T}

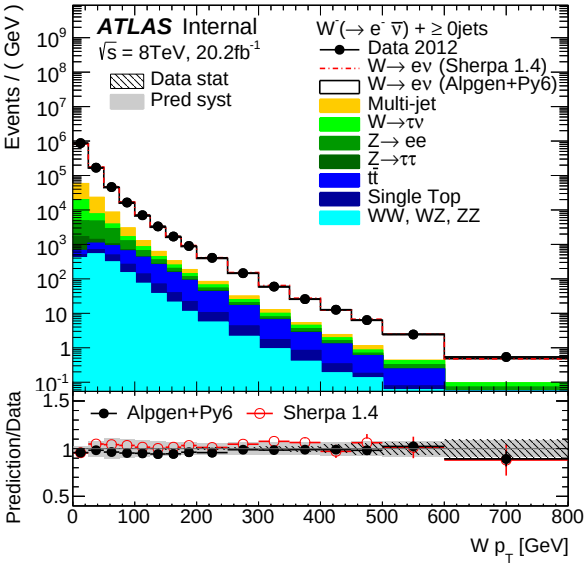
Figure B.2: W-related kinematic distributions at the detector level for the W^+ selection: Positron (e^+) p_{T}(a) and η(b), as well as W^+ p_{T}(c), for any number of jets produced. The lower panels are defined as in Figure 6.1.



(a) Electron p_{T}

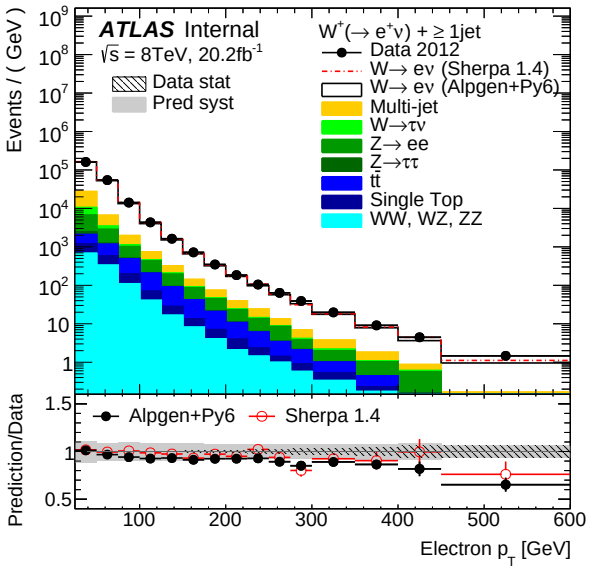


(b) electron η

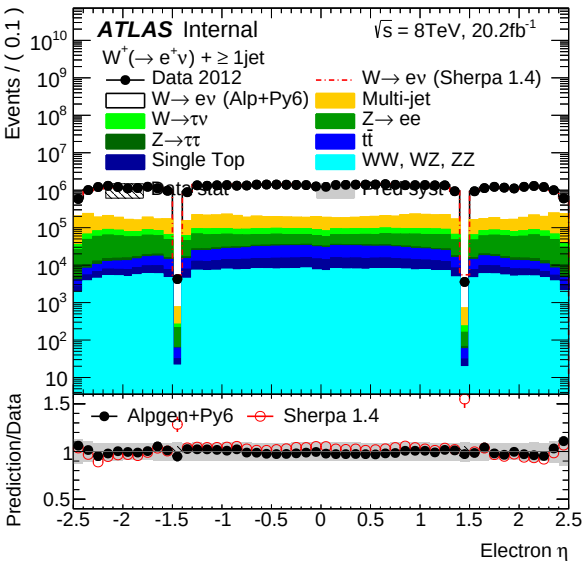


(c) W^- p_{T}

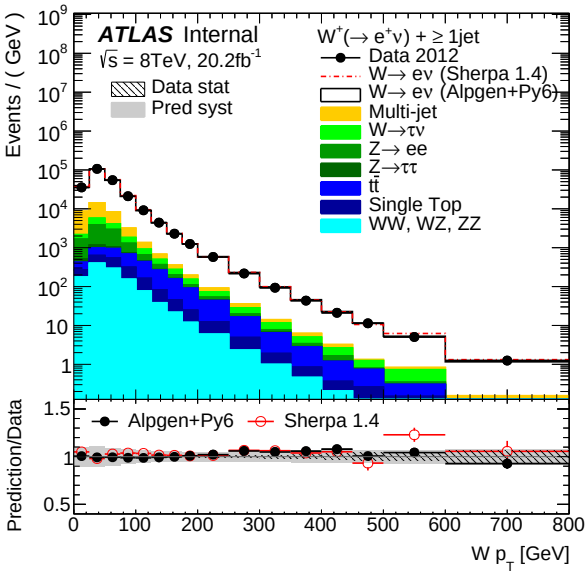
Figure B.3: W-related kinematic distributions at the detector level for the W^- selection: Electron (e^-) p_{T}(a) and η(b), as well as W^- p_{T}(c), for any number of jets produced. The lower panels are defined as in Figure 6.1.



(a) Positron p_{T}

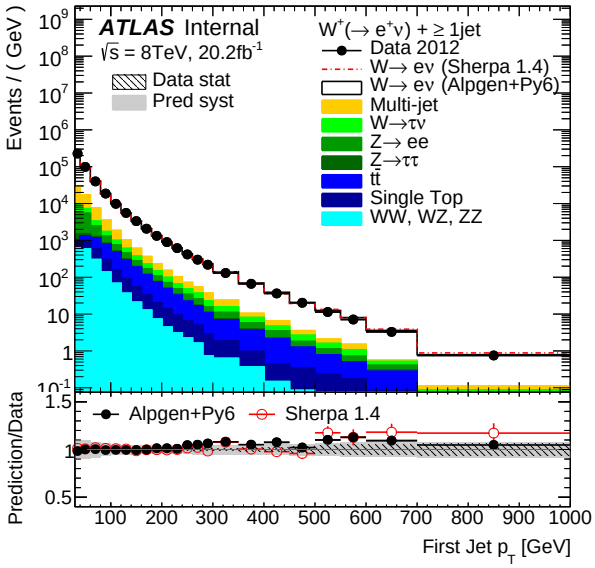


(b) Positron η

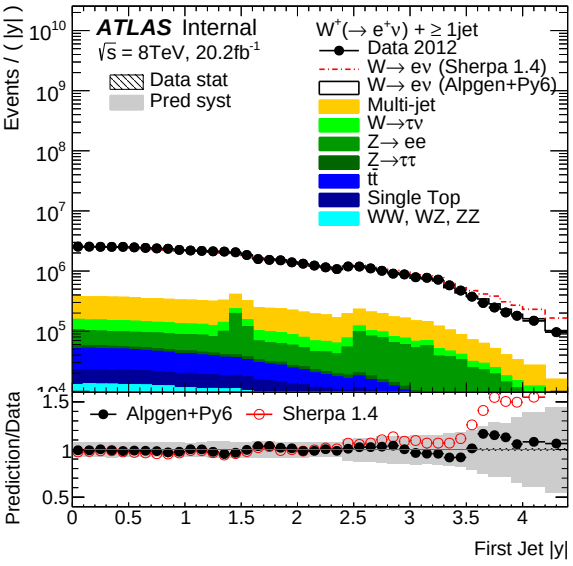


(c) W^+ p_{T}

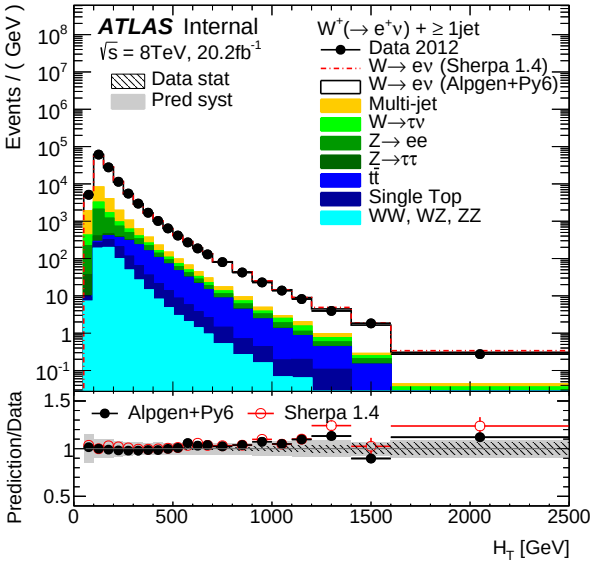
Figure B.4: W-related kinematic distributions at the detector level for the W^+ selection: Positron (e^+) p_{T}(a) and η(b), as well as W^+ p_{T}(c), with at least 1 jet produced in association. The lower panels are defined as in Figure 6.1.



(a) Leading jet p_{T}

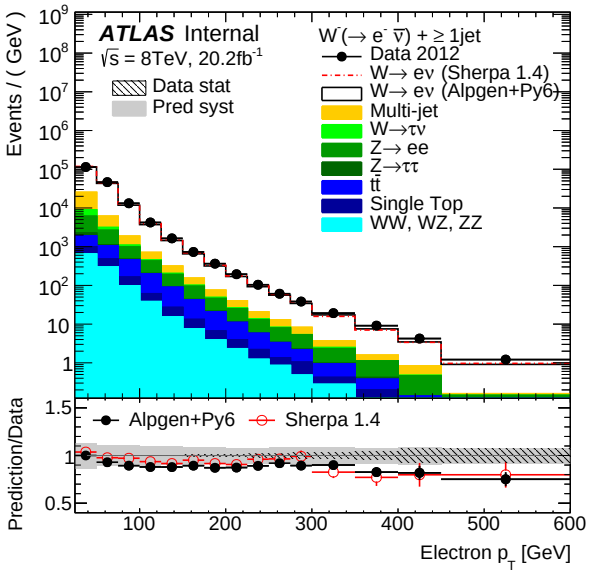


(b) Leading jet $|y|$

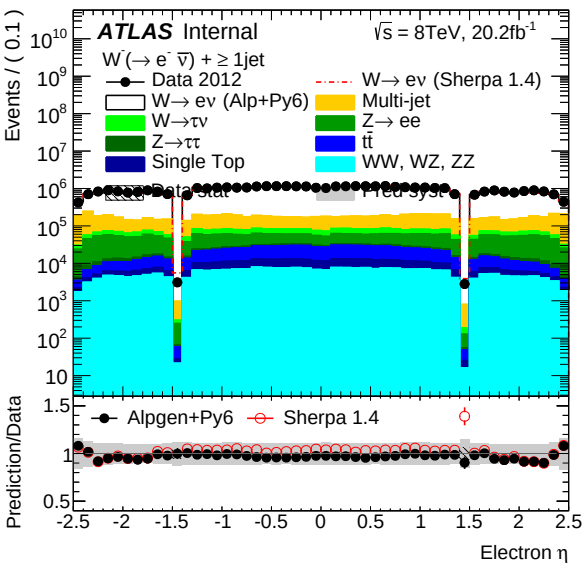


(c) H_{T}

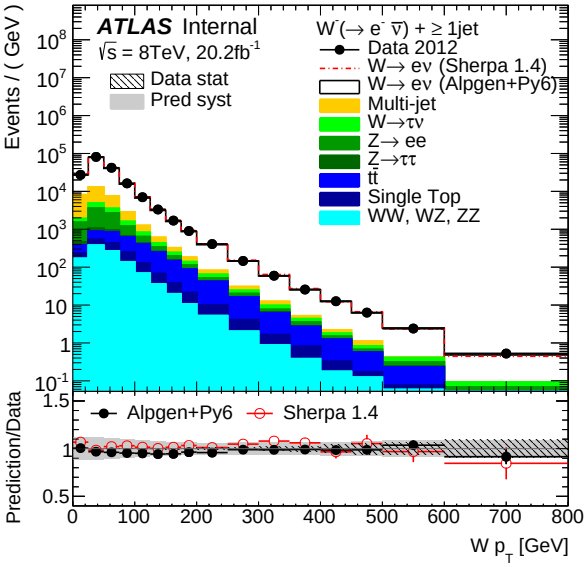
Figure B.5: Jet-related kinematic distributions at the detector level for the W^+ selection: Leading jet p_{T}(a) and absolute rapidity $|y|$(b), as well as H_{T}(c), with at least 1 jet produced in association. The lower panels are defined as in Figure 6.1.



(a) Electron p_{T}

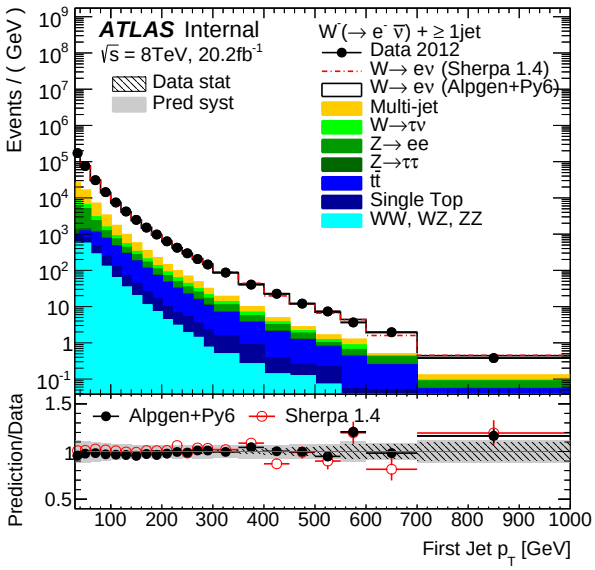


(b) Electron η

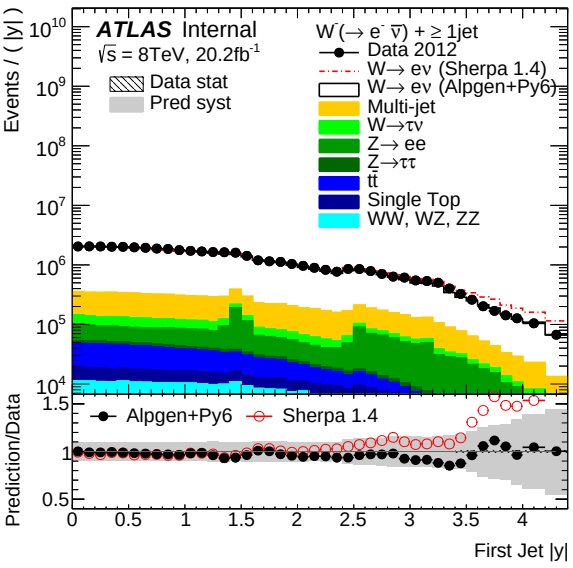


(c) W^- p_{T}

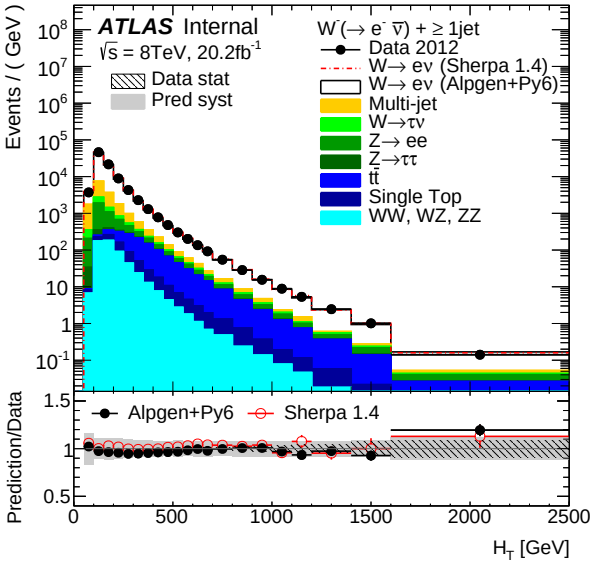
Figure B.6: W-related kinematic distributions at the detector level for the W^- selection: Electron (e^-) p_{T}(a) and η(b), as well as W^- p_{T}(c), with at least 1 jet produced in association. The lower panels are defined as in Figure 6.1.



(a) Leading jet p_{T}

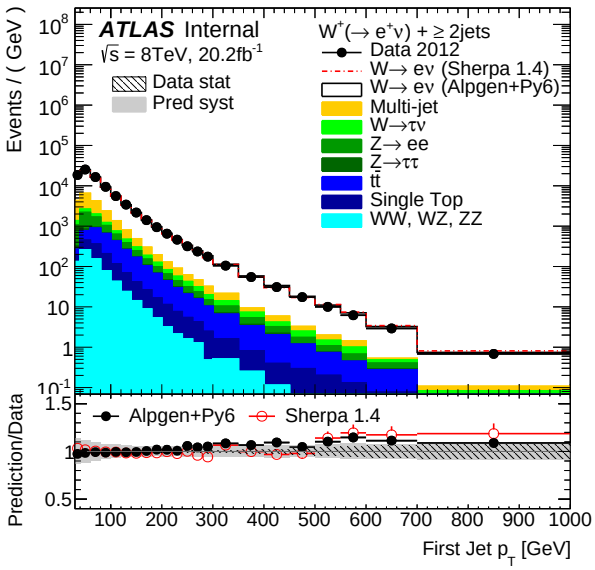


(b) Leading jet y

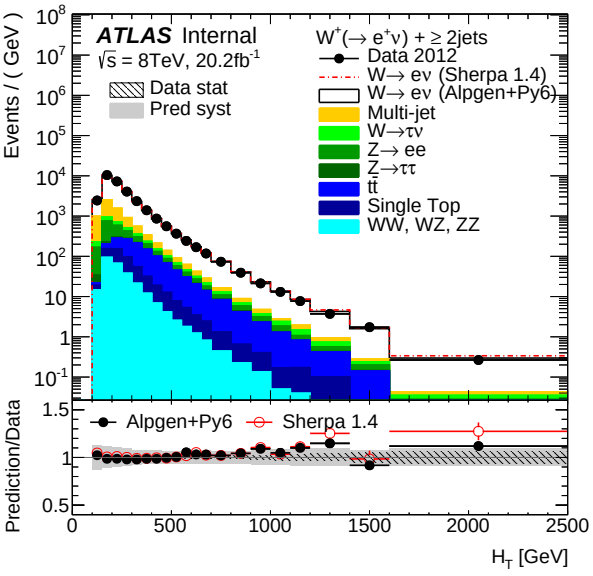


(c) H_{T}

Figure B.7: Jet-related kinematic distributions at the detector level for the W^- selection: Leading jet p_{T}(a) and absolute rapidity $|y|$(b), as well as H_{T}(c), with at least 1 jet produced in association. The lower panels are defined as in Figure 6.1.

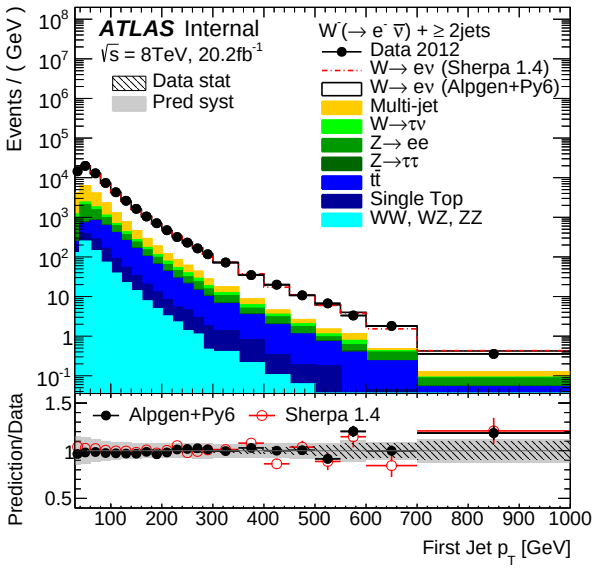


(a) Leading jet p_{T}

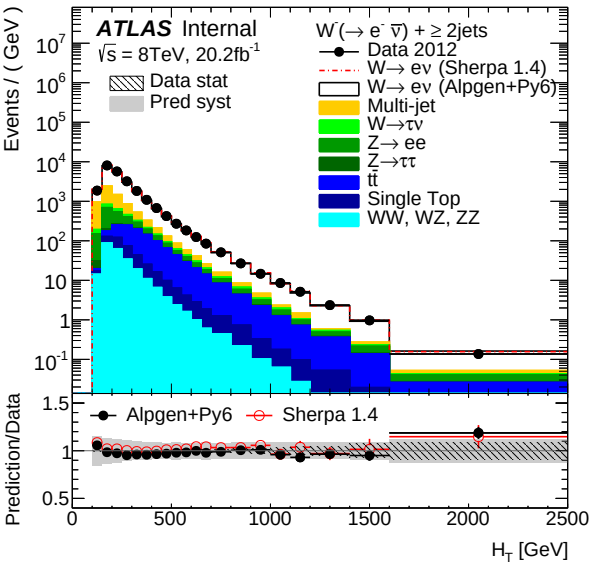


(b) H_{T}

Figure B.8: Jet-related kinematic distributions at the detector level for the W^+ selection in the presence of at least two jets: Leading jet p_{T}(a) and H_{T}(b). The lower panels are defined as in Figure 6.1.



(a) Leading jet p_{T}



(b) H_{T}

Figure B.9: Jet-related kinematic distributions at the detector level for the W^- selection in the presence of at least two jets: Leading jet p_{T}(a) and H_{T}(b). The lower panels are defined as in Figure 6.1.

APPENDIX C

Additional cross section distributions

Distributions for events with $N_{\text{jets}} \geq 2$

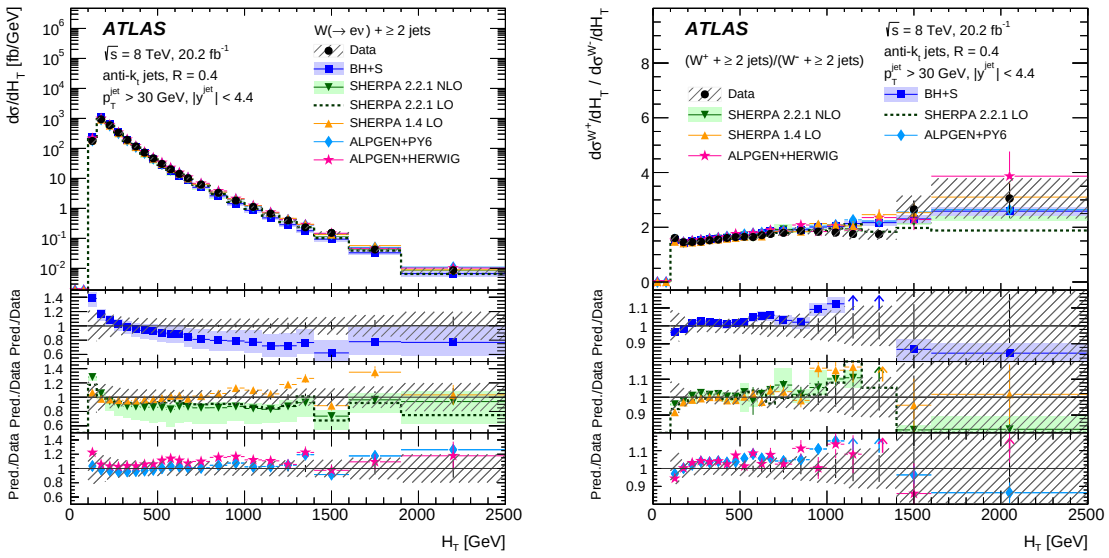
The following section contains additional distributions not shown in Chapter 10 for W production and the W^+/W^- ratio with at least 2 jets. The cross section is shown as a function of H_{T} (Figure C.1), W boson p_{T} (Figure C.2), lead jet p_{T} (Figure C.3).

Pseudorapidity of the electron

Figures C.4 and C.5 in this section show the W, W^+, W^-, and W^+/W^- ratio cross sections as a function of the pseudorapidity of the electron for events with $N_{\text{jets}} \geq 0$ and $N_{\text{jets}} \geq 1$.

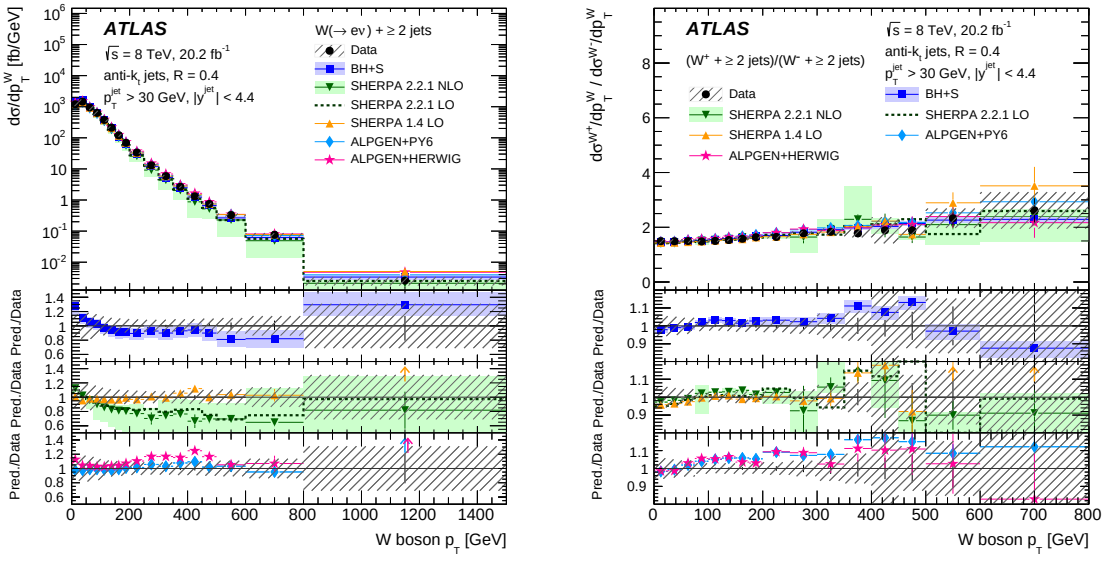
W^+ and W^- cross sections

This section contains all the distributions shown above and in Chapter 10 for W^+ and W^- production separately. These are the distributions used to calculate the W^+/W^- ratio. Figures C.6-C.10 show the default sets of distributions for $N_{\text{jets}} \geq 1$, while Figures C.11-C.14 show the same distributions generated using MCFM and different PDF sets. Finally, Figures C.15-C.17 show the default set of predictions for events with $N_{\text{jets}} \geq 2$.



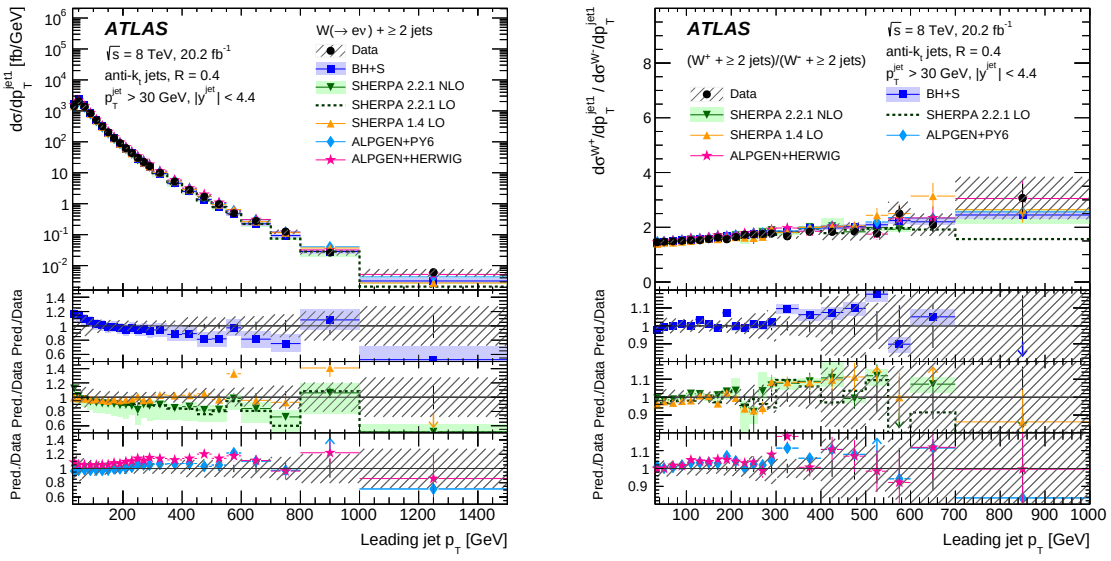
(a) (b)

Figure C.1: Differential cross sections for the production of W bosons (left) and the W^+/W^- ratio (right) as a function of H_{T} for events with $N_{\mathrm{jets}} \geq 2$. The last bin in the left figure includes values beyond the shown range. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



(a) (b)

Figure C.2: Differential cross sections for the production of W bosons (left) and the W^+/W^- ratio (right) as a function of W p_{T} for events with $N_{\mathrm{jets}} \geq 2$. The last bin in the left figure includes values beyond the shown range. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



(a) (b)

Figure C.3: Differential cross sections for the production of W bosons (left) and the W^+/W^- ratio (right) as a function of lead jet p_{T} for events with $N_{\mathrm{jets}} \geq 2$. The last bin in the left figure includes values beyond the shown range. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].

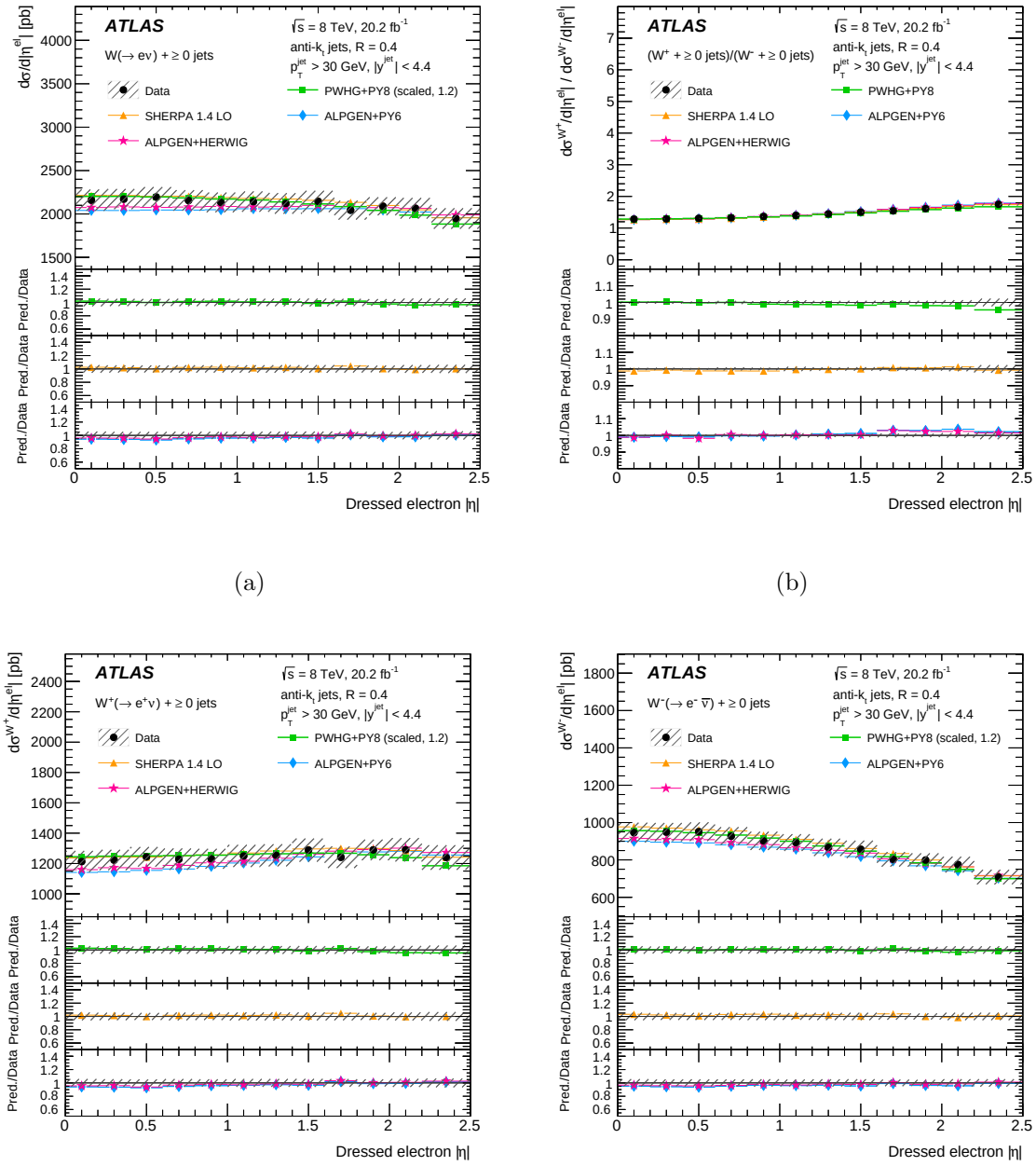


Figure C.4: Differential cross sections for the production of W bosons (top left), W^+ (bottom left), W^- (bottom right) and the W^+/W^- ratio (top right) as a function of the electron η for events with $N_{\text{jets}} \geq 0$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].

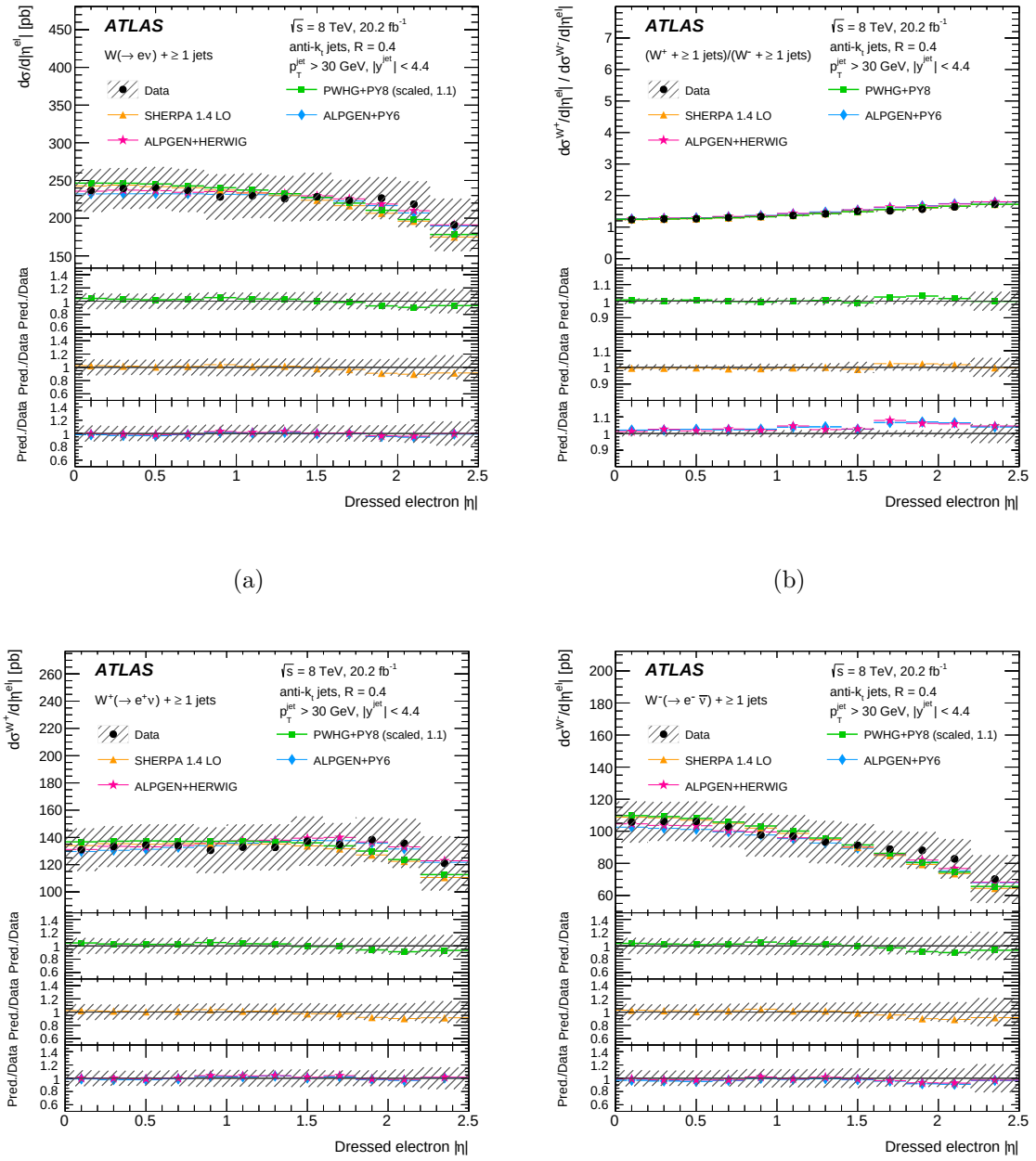
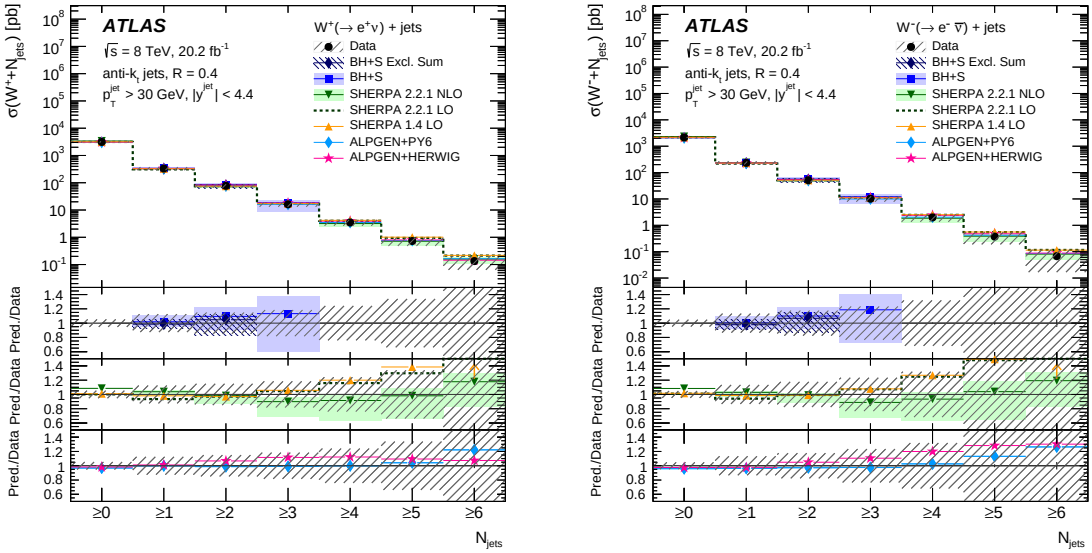
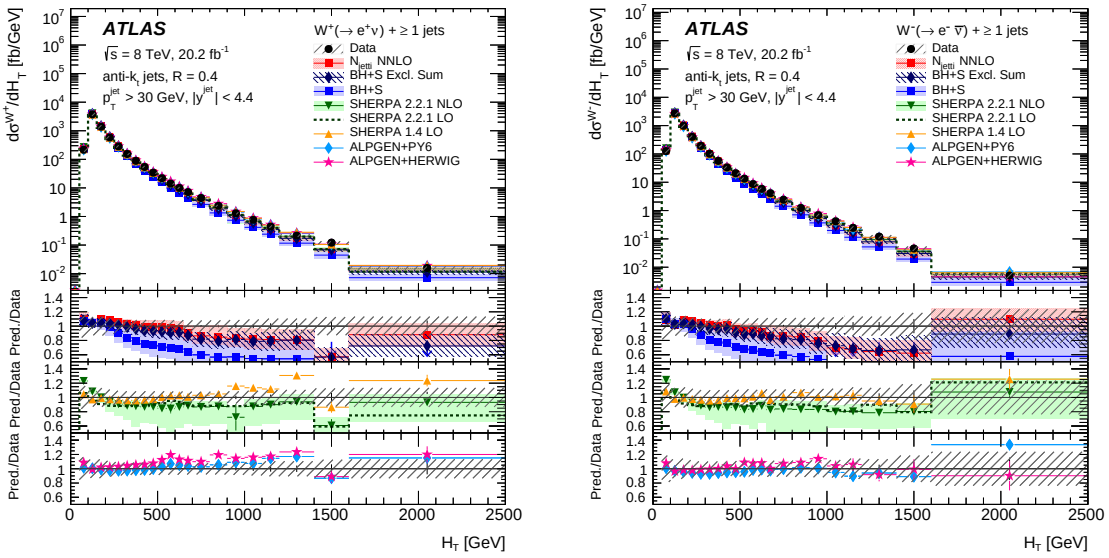


Figure C.5: Differential cross sections for the production of W bosons (top left), W^+ (bottom left), W^- (bottom right) and the W^+/W^- ratio (top right) as a function of the electron η for events with $N_{\text{jets}} \geq 1$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



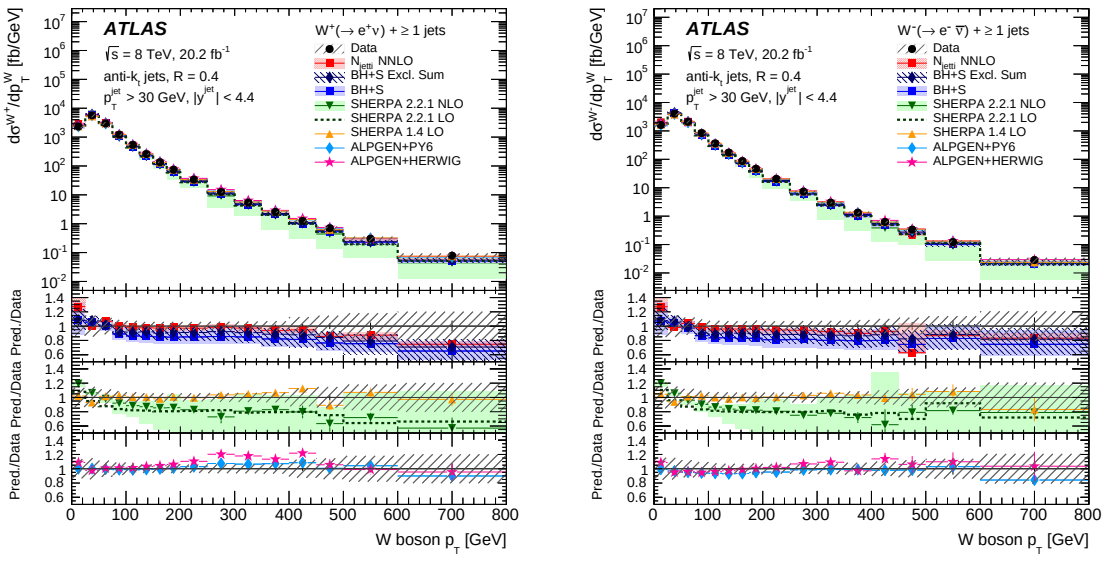
(a) (b)

Figure C.6: Differential cross sections for the production of W^+ (left) and W^- (right) as a function of the inclusive jet multiplicity. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



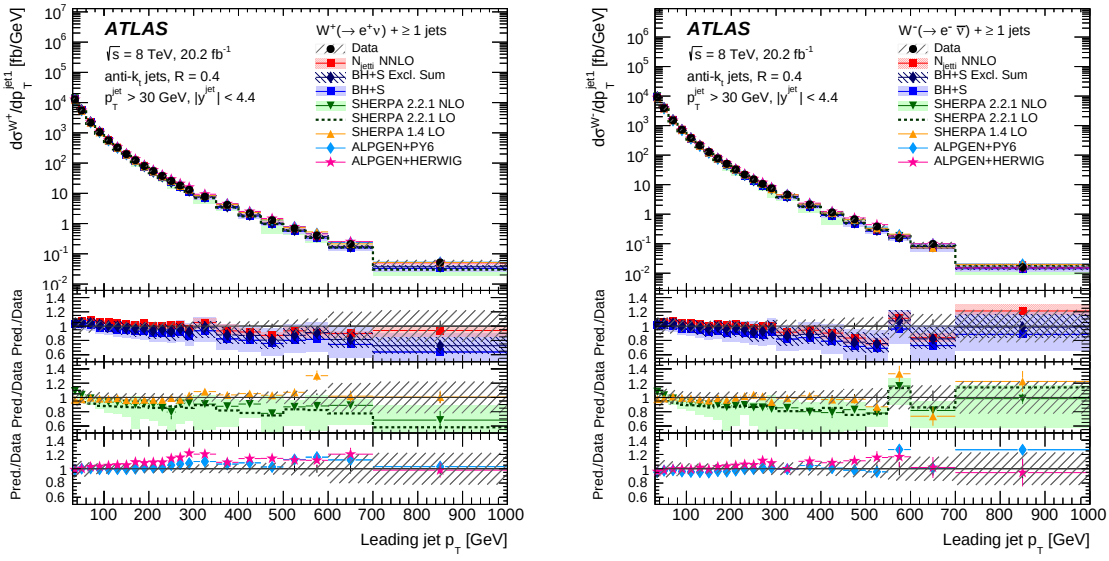
(a) (b)

Figure C.7: Differential cross sections for the production of W^+ (left) and W^- (right) as a function of the H_{T} for events with $N_{\mathrm{jets}} \geq 1$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



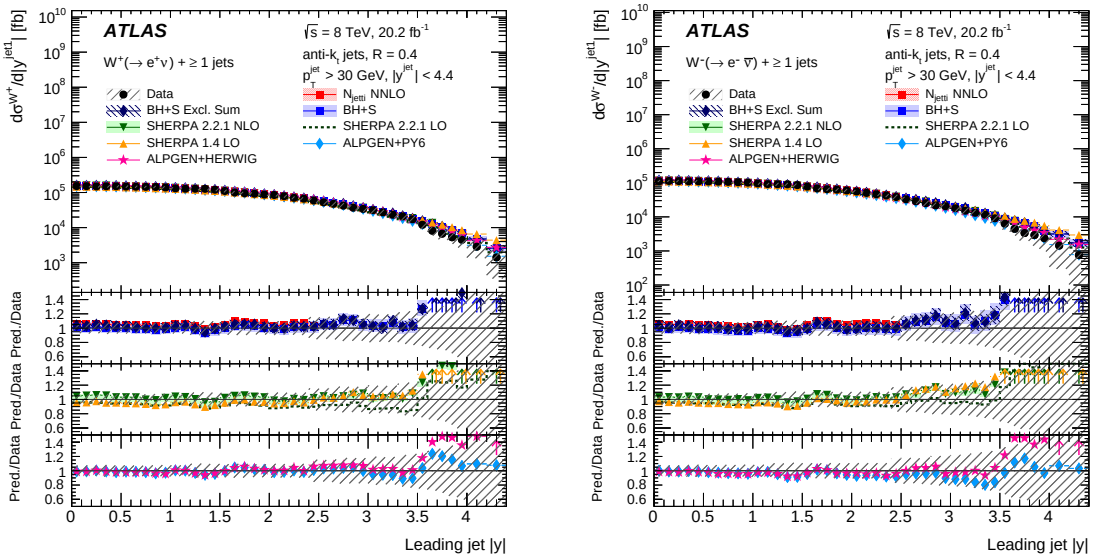
(a) (b)

Figure C.8: Differential cross sections for the production of W^+ (left) and W^- (right) as a function of the p_{T} of the W boson for events with $N_{\mathrm{jets}} \geq 1$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



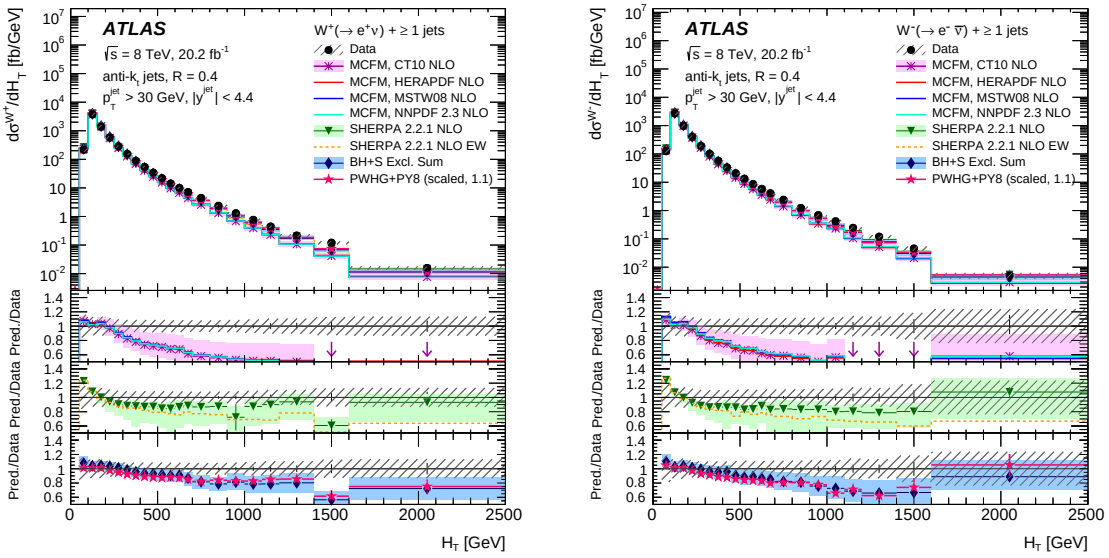
(a) (b)

Figure C.9: Differential cross sections for the production of W^+ (left) and W^- (right) as a function of the p_{T} of the lead jet for events with $N_{\mathrm{jets}} \geq 1$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



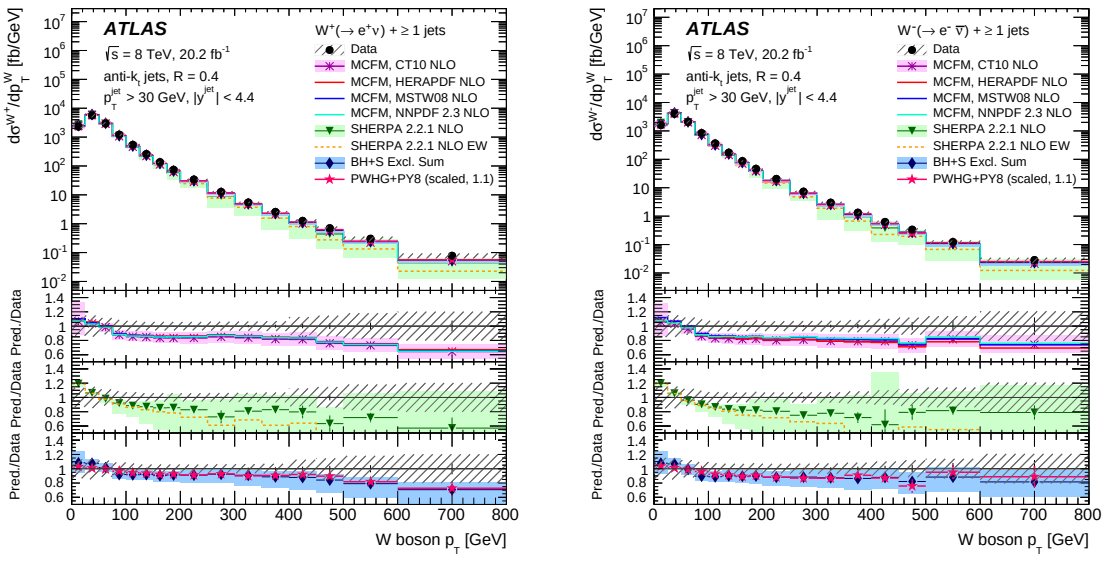
(a) (b)

Figure C.10: Differential cross sections for the production of W^+ (left) and W^- (right) as a function of the rapidity of the lead jet for events with $N_{\text{jets}} \geq 1$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



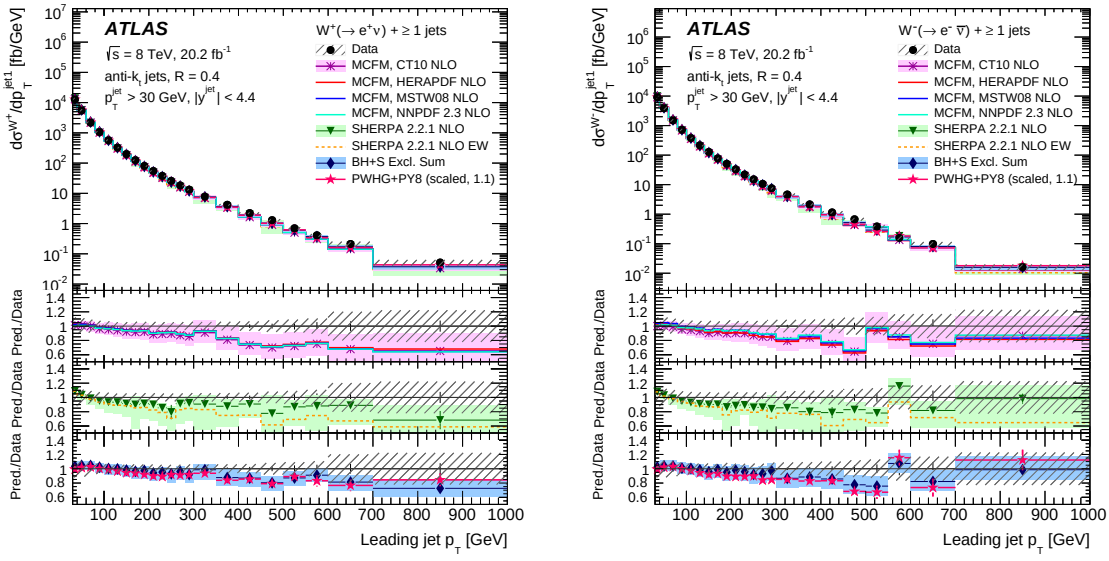
(a) (b)

Figure C.11: Differential cross sections for the production of W^+ (left) and W^- (right) as a function of the H_{T} for events with $N_{\mathrm{jets}} \geq 1$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



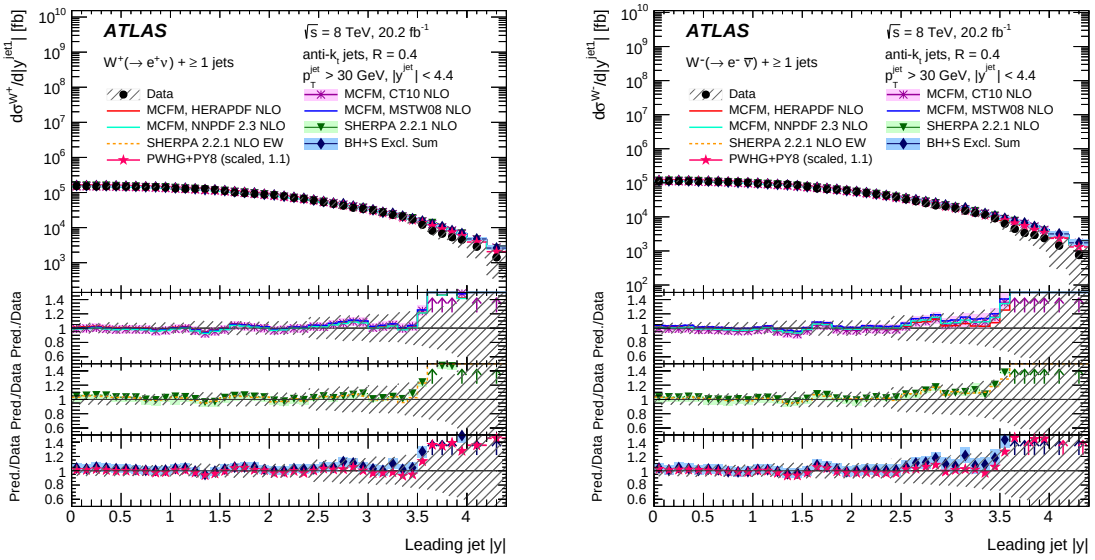
(a) (b)

Figure C.12: Differential cross sections for the production of W^+ (left) and W^- (right) as a function of the p_{T} of the W boson for events with $N_{\mathrm{jets}} \geq 1$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



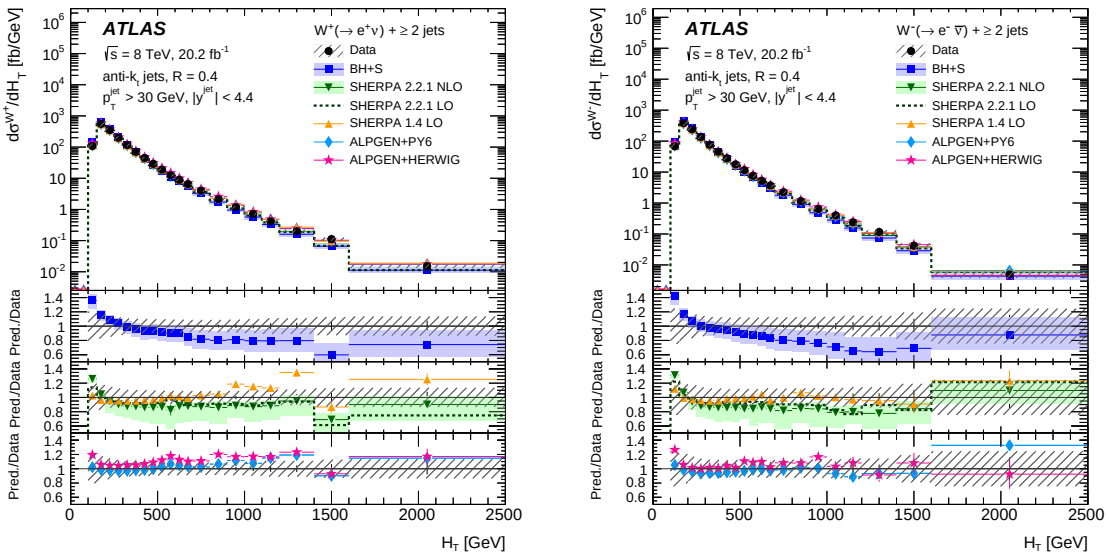
(a) (b)

Figure C.13: Differential cross sections for the production of W^+ (left) and W^- (right) as a function of the p_{T} of the lead jet for events with $N_{\mathrm{jets}} \geq 1$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



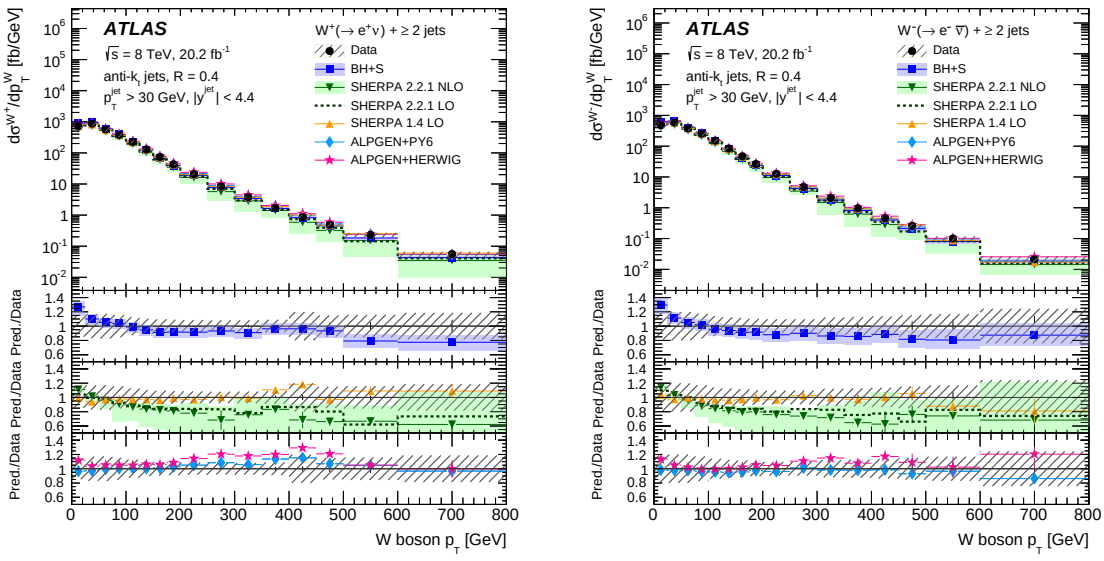
(a) (b)

Figure C.14: Differential cross sections for the production of W^+ (left) and W^- (right) as a function of the rapidity of the lead jet for events with $N_{\text{jets}} \geq 1$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



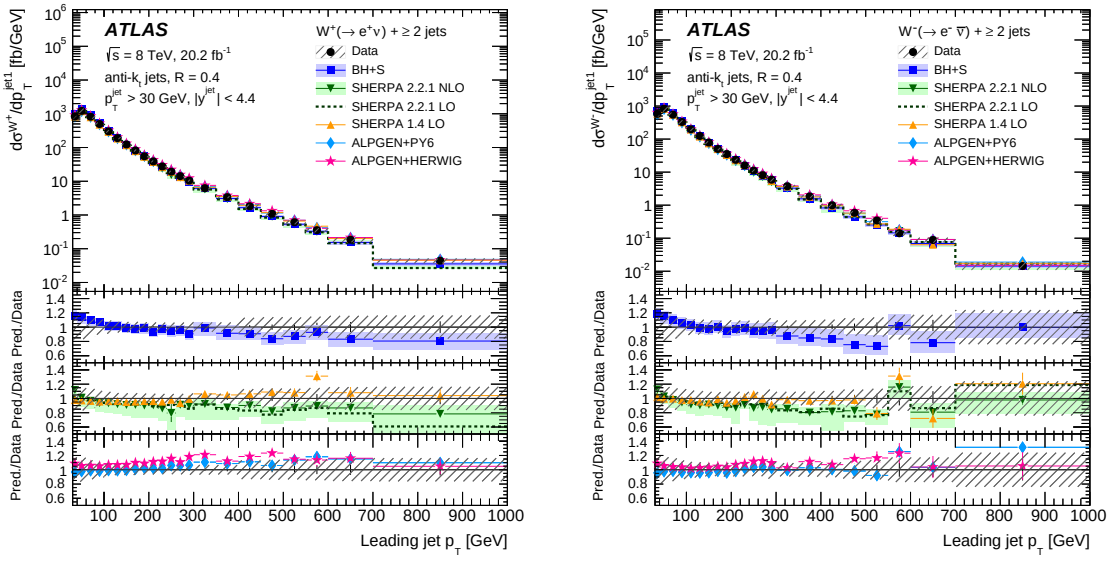
(a) (b)

Figure C.15: Differential cross sections for the production of W^+ (left) and W^- (right) as a function of the H_{T} for events with $N_{\mathrm{jets}} \geq 2$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



(a) (b)

Figure C.16: Differential cross sections for the production of W^+ (left) and W^- (right) as a function of the p_{T} of the W boson for events with $N_{\mathrm{jets}} \geq 2$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].



(a) (b)

Figure C.17: Differential cross sections for the production of W^+ (left) and W^- (right) as a function of the p_{T} of the lead jet for events with $N_{\mathrm{jets}} \geq 2$. For the data, the statistical uncertainties are indicated as vertical bars, and the combined statistical and systematic uncertainties are shown by the hatched bands. The uppermost panel in each plot shows the differential cross sections, while the lower panels show the ratios of the predictions to the data. The theoretical uncertainties on the predictions are described in the text. The arrows on the lower panels indicate points that are outside the displayed range [12].

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using 2012 LHC proton-proton collision data*, ArXiv e-prints (2016) , arXiv:1612.01456 [hep-ex].

[2] G. Arabidze, B. Bernarding, M. Dunford, J. Huston, V. S. Lang, A. Tricoli, M. C. Mondragon, and T. P. Steinsberger, *Measurement of the W boson production cross section and the W^+/W^- ratio in association with jets in pp collisions at $\sqrt{s} = 8\,TeV$: Supporting note*, Tech. Rep. ATL-COM-PHYS-2016-124, CERN, Geneva, Feb, 2016. https://cds.cern.ch/record/2131742.

[3] *The Durham HepData Project*, http://hepdata.cedar.ac.uk/pdf/pdf3.html.

[4] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. **D93** (2016) no. 3, 033006, arXiv:1506.07443 [hep-ph].

[5] ATLAS Collaboration, *ATLAS Experiment at CERN*, http://atlas.cern/.

[6] ATLAS Collaboration, *Atlas Experiment-Public Results: Luminosity Public Results*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[7] *CERN Document Server: rendered ATLAS Pixel Module closeups*, https://cds.cern.ch/record/854105.

[8] *CERN Document Server: The Colliderscope: a real-time show*, https://cds.cern.ch/journal/CERNBulletin/2010/05/News

[9] ATLAS Collaboration, *Atlas Public Twiki*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/.

[10] ATLAS Collaboration Collaboration, *ATLAS detector and physics performance: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1999. https://cds.cern.ch/record/391176.

[11] A. Chegwidden, *A multivariate analysis search for a single-top quark produced in association with missing energy in proton-proton collisions at the Large Hadron Collider at a center-of-mass energy of $\sqrt{s} = 8$ TeV with the ATLAS Detector.* PhD thesis, Michigan State University., 2017. https://doi.org/10.25335/M56T53.

[12] ATLAS Collaboration, M. Aaboud et al., *Measurement of differential cross sections and W^+/W^- cross-section ratios for W boson production in association with jets at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **05** (2018) 077, arXiv:1711.03296 [hep-ex].

[13] C. F. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, H. Ita, D. A. Kosower, and D. Maitre, *An Automated Implementation of On-Shell Methods for One-Loop Amplitudes*, Phys. Rev. **D78** (2008) 036003, arXiv:0803.4180 [hep-ph].

[14] H. Ita, Z. Bern, L. J. Dixon, F. Febres Cordero, D. A. Kosower, and D. Maitre, *Precise Predictions for Z + 4 Jets at Hadron Colliders*, Phys. Rev. **D85** (2012) 031501, arXiv:1108.2229 [hep-ph].

[15] R. Boughezal, C. Focke, X. Liu, and F. Petriello, *W-Boson Production in Association with a Jet at Next-to-Next-to-Leading Order in Perturbative QCD*, Phys. Rev. Lett. **115** (2015) 062002. http://link.aps.org/doi/10.1103/PhysRevLett.115.062002.

[16] ATLAS Collaboration, *Study of jets produced in association with a W boson in pp collisions at s=7TeV with the ATLAS detector*, prd **85** (2012) no. 9, 092002, arXiv:1201.1276 [hep-ex].

[17] ATLAS Collaboration, *Measurement of the transverse momentum distribution of W bosons in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D **85** (2012) 012005.

[18] Particle Data Group Collaboration, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.

[19] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995. http://www.slac.stanford.edu/ mpeskin/QFT.html.

[20] D. Bourilkov, R. C. Group, and M. R. Whalley, *LHAPDF: PDF use from the Tevatron to the LHC*, in *TeV4LHC Workshop - 4th meeting Batavia, Illinois, October 20-22, 2005*. 2006. arXiv:hep-ph/0605240 [hep-ph].

[21] *CERN Accelerating science*, https://home.cern/.

[22] *TrkKalmanFitter*. https://twiki.cern.ch/twiki/bin/view/AtlasComputing/TrkKalmanFitter.

[23] T. G. Cornelissen, M. Elsing, I. Gavrilenko, J. F. Laporte, W. Liebig, M. Limper, K. Nikolopoulos, A. Poppleton, and A. Salzburger, *The global chi**2 track fitter in ATLAS*, J. Phys. Conf. Ser. **119** (2008) 032013.

[24] ATLAS Collaboration Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, Tech. Rep. ATLAS-CONF-2012-047, CERN, Geneva, May, 2012. http://cds.cern.ch/record/1449796.

[25] ATLAS Collaboration Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, European Physical Journal C **74** (2014) 3071, arXiv:1407.5063 [hep-ex].

[26] GEANT4 Collaboration, *Geant4 a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) no. 3, 250 – 303. http://www.sciencedirect.com/science/article/pii/S0168900203013688.

[27] ATLAS Collaboration, S. Leontsinis, *Performance of the muon identification and reconstruction with the ATLAS detector*, EPJ Web Conf. **95** (2015) 05009.

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

[29] ATLAS Collaboration, *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*, Tech. Rep. ATLAS-CONF-2015-037, CERN, Geneva, Aug, 2015. https://cds.cern.ch/record/2044941.

[30] ATLAS Collaboration, *Jet global sequential corrections with the ATLAS detector in proton-proton collisions at sqrt(s) = 8 TeV*, Tech. Rep. ATLAS-CONF-2015-002, CERN, Geneva, Mar, 2015. https://cds.cern.ch/record/2001682.

[31] D. W. Miller, A. Schwartzman, and D. Su, *Jet-Vertex Association Algorithm*, Tech. Rep. ATL-COM-PHYS-2008-008, CERN, Geneva, Jan, 2008. https://cds.cern.ch/record/1082880.

[32] D. e. a. Adams, *Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014*, Tech. Rep. ATL-PHYS-INT-2014-018, CERN, Geneva, Jul, 2014. https://cds.cern.ch/record/1743654.

[33] ATLAS Collaboration, G. Aad et al., *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11** (2016) no. 04, P04008, arXiv:1512.01094 [hep-ex].

[34] *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2014-004, CERN, Geneva, Feb, 2014. http://cds.cern.ch/record/1664335.

[35] ATLAS Collaboration, *Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, ArXiv e-prints (2015) , arXiv:1511.04716 [hep-ex].

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

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

[38] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026, arXiv:hep-ph/0603175.

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

[40] G. D'Agostini, *A Multidimensional unfolding method based on Bayes' theorem*, Nucl. Instrum. Meth. **A362** (1995) 487–498.

[41] D'Agostini, G., *Improved iterative Bayesian unfolding*, arXiv:1010.0632.

[42] ATLAS Collaboration, *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*, ATLAS-CONF-2015-037, 2015. https://cds.cern.ch/record/2044941.

[43] *Recommendations for Jet Energy Scale Uncertainties in 2012.* https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetUncertainties2012Final.

[44] *Recommendations for Jet Energy Resolution.* https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ApplyJetResolutionSmearing.

[45] *JVF Twiki.* https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JVFUncertaintyTool.

[46] *B-tagging scale factors.* https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/BTaggingCalibrationDataInterface.

[47] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C **74** (2014) 3071, arXiv:1407.5063 [hep-ex].

[48] *Electron Efficiency Measurements in 2012 Twiki.* https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EfficiencyMeasurements2012.

[49] ATLAS Collaboration, *Reconstruction and Performance of Missing Transverse Momentum in the ATLAS Detector using Proton-Proton Collisions at* $\sqrt{s}$=8 TeV, ATLAS-PERF-2014-04-001, 2015. https://cds.cern.ch/record/2056659.

[50] *ATLAS Tracking Twiki for 2012.* https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TrackingCPMC12.

[51] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **0411** (2004) 040, arXiv:hep-ph/0409146 [hep-ph].

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

[53] J. Ferrando and D. Wendland, *Reference* $t\bar{t}$ *production cross sections for use in ATLAS analyses*, Tech. Rep. ATL-COM-PHYS-2014-112, CERN, Geneva, Feb, 2014. https://cds.cern.ch/record/1662536.

[54] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. **D81** (2010) 054028, arXiv:1001.5034 [hep-ph].

[55] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. **D83** (2011) 091503, arXiv:1103.2792 [hep-ph].

[56] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a* W^- *or* H^-, Phys. Rev. **D82** (2010) 054018, arXiv:1005.4451 [hep-ph].

[57] ATLAS Collaboration Collaboration, *ATLAS simulation of boson plus jets processes in Run 2*, Tech. Rep. ATL-PHYS-PUB-2017-006, CERN, Geneva, 2017. https://cds.cern.ch/record/2261937.

[58] ATLAS Collaboration Collaboration, *Multi-Boson Simulation for 13 TeV ATLAS Analyses*, Tech. Rep. ATL-PHYS-PUB-2017-005, CERN, Geneva, 2017. https://cds.cern.ch/record/2261933.

[59] ATLAS Collaboration, ATLAS Collaboration, *Luminosity determination in pp collisions at* $\sqrt{s}$ *= 8 TeV using the ATLAS detector at the LHC*, arXiv:1608.03953 [hep-ex].

[60] ATLAS Collaboration, G. Aad et al., *Measurements of the W production cross sections in association with jets with the ATLAS detector*, Eur. Phys. J. **C75** (2015) no. 2, 82, arXiv:1409.8639 [hep-ex].

[61] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED correctionsin Z and W decays*, European Physical Journal C **45** (2006) 97–107, hep-ph/0506026.

[62] S. Jadach, Z. Was, R. Decker, and J. Khn, *The decay library TAUOLA, version 2.4*, Computer Physics Communications **76** (1993) no. 3, 361 – 380. http://www.sciencedirect.com/science/article/pii/001046559390061G.

[63] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, arXiv:hep-ph/0011363 [hep-ph].

[64] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. **C72** (1996) 637–646, arXiv:hep-ph/9601371 [hep-ph].

[65] C. F. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, T. Gleisberg, H. Ita, D. A. Kosower, and D. Maitre, *Next-to-Leading Order QCD Predictions for W+3-Jet Distributions at Hadron Colliders*, Phys. Rev. **D80** (2009) 074036, arXiv:0907.1984 [hep-ph].

[66] J. M. Campbell and R. K. Ellis, *An Update on vector boson pair production at hadron colliders*, Phys. Rev. **D60** (1999) 113006, arXiv:hep-ph/9905386 [hep-ph].

[67] M. Guzzi, P. Nadolsky, E. Berger, H.-L. Lai, F. Olness, and C. P. Yuan, *CT10 parton distributions and other developments in the global QCD analysis*, arXiv:1101.0561 [hep-ph].

[68] ZEUS, H1 Collaboration, H. Abramowicz et al., *Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data*, Eur. Phys. J. **C75** (2015) no. 12, 580, arXiv:1506.06042 [hep-ex].

[69] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[70] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. **B867** (2013) 244–289, arXiv:1207.1303 [hep-ph].

[71] R. Boughezal, C. Focke, X. Liu, and F. Petriello, *W-boson production in association with a jet at next-to-next-to-leading order in perturbative QCD*, Phys. Rev. Lett. **115** (2015) no. 6, 062002, arXiv:1504.02131 [hep-ph].

[72] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, A. Huss, and T. A. Morgan, *Precise QCD predictions for the production of a Z boson in association with a hadronic jet*, Phys. Rev. Lett. **117** (2016) no. 2, 022001, arXiv:1507.02850 [hep-ph].

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

[74] *Sherpa Manual Version 2.2.1.* https://sherpa.hepforge.org/doc/SHERPA-MC-2.2.1.html.

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

[76] ATLAS Collaboration, M. Aaboud et al., *Measurements of electroweak Wjj production and constraints on anomalous gauge couplings with the ATLAS detector*, Eur. Phys. J. **C77** (2017) no. 7, 474, arXiv:1703.04362 [hep-ex].