Measurements of vector boson production in association with jets at the LHC using the ATLAS detector

Craig Anthony Sawyer
St. Cross College, Oxford

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Hilary Term, 2015

Abstract

This thesis presents measurements of the cross section for the production of a W boson in association with jets and the ratio of the cross sections for the production of a W boson and a Z boson in association with jets. Both measurements are performed in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS experiment at the Large Hadron Collider using a dataset corresponding to an integrated luminosity of 4.6 fb^{-1}. The dataset allows for the exploration of a large kinematic range, including jet production up to a transverse momentum of 1 TeV and multiplicities up to seven jets. Results are presented as a function of jet transverse momenta and rapidities and as a function of event variables such as the scalar sum of the transverse momenta of the jets. Results as a function of dijet angular variables are also presented. The measurements are compared to several state-of-the-art QCD predictions including next-to-leading-order perturbative calculations, resummation calculations and Monte Carlo generators. Finally, the effect of the results on parton distribution functions is explored.

To Samantha

Acknowledgements

My first thanks must go to the STFC for the funding they provided for the last 3.5 years, including a 12 month stay at CERN and travel to multiple conferences from Russia to Chicago. Additionally, thank you to St Cross College and the IOP which provided additional funding for me to travel to the LowX conference in Israel.

In the way of personal thanks, firstly I thank my supervisor, Amanda Cooper-Sarkar, for the advice and support she has offered me over the course of my DPhil and the ever watchful eye over my shoulder.

Also I thank the administrative team at Oxford for their invaluable help navigating the complexity of bureaucracy at CERN and Oxford. Much of my computing time (14 years) has been amassed on the Oxford batch system which has been incredibly speedy and reliable. For this, I have to thank the Oxford unix team who maintained the system and quickly responded to any issues, particularly when my poorly planned jobs took over large portions of the farm!

Next, a huge thanks must go to the people I have collaborated with on the R_{jets} and $W+$jets measurements who have helped take me from clueless first year DPhil student to a slightly less clueless fourth year DPhil student. In particular, Evelin Meoni, Hugo Beauchemin, Monica Dunford and Gerhard Brandt.

Thanks must also go to Caterina Doglioni who originally got me involved in the Jet/EtMiss group and John Backus Mayes with whom I worked very closely during the development and validation of the 2012 jet pile-up subtraction scheme.

My time at Oxford would not have been the same without the other Oxford ATLAS students - Jim, David, Shaun, Mireia, Rob and Kate in particular - a fantastic bunch of people who were always keen for a rant or to answer questions. A particular shout out to Jim and Gerhard without whom I would not have made it back in time for the RAL summer school.

Last, and by no means least, thank you to my family who supported me, not just over the last 3.5 years but also for all the years to get here and as I write this are currently spotting typo after typo. Particularly, thank you to Samantha for being willing to deal with either flying to Geneva straight after Friday work, or picking me up from the airport during my year at CERN.

Preface

Performing an analysis on an experiment such as ATLAS is, by its very nature, a collaborative effort and, as a result, much of the work presented in this thesis is dependent on work carried out by others in the collaboration. My initial involvement in the analysis began with a ratio analysis, R_{jets} [1], but following the amalgamation of two analysis groups the scope was extended to also include the $W + \text{jets}$ measurement [2]. My first task involved the development of the code used in the electron channel selection however this quickly expanded to include the integration of all detector level uncertainties into the analysis.

Once the base analysis was implemented, my focus shifted to the data-driven estimation of multi-jet backgrounds, the derivation of the various corrections required for comparison to theory and discussions with theorists to prepare the various theory predictions. In the meantime, a collaborator working on the muon channel left the field and I took over responsibilities for maintaining and running his part of the analysis.

Finally, as the analysis neared conclusion, I took over the development of the final combination step of the analysis and became responsible for running the whole analysis chain apart from $t\bar{t}$ background estimation and the unfolding, although all inputs to the unfolding were produced in the main analysis code. As a result, I have at some point been involved in most steps of the analysis.

In addition to the two analyses presented here, I qualified as an ATLAS author working on the pile-up part of the jet calibration procedure for 2012 data taking [3] and am still heavily involved in the Jet/EtMiss performance group. I also undertook Run Control shifts in the ATLAS Control Room during my stay at CERN.

Many of the plots used in this thesis have been published or made public and as such are labelled ***ATLAS*** or ***ATLAS*** Preliminary respectively. All plots have been produced by myself unless otherwise stated in the caption. Where possible I have referenced public documents however in a couple of cases plots were taken from internal ATLAS documentation in which case references may be hidden behind the ATLAS firewall.

Contents

Introduction

The Standard Model of Particle Physics describes the interactions of sub-atomic particles and has proven itself to provide incredibly accurate predictions of a wide range of phenomena. However, various other phenomena show us that there must be more than just the Standard Model. With the start-up of the Large Hadron Collider (LHC), the world's highest energy particle accelerator, a new kinematic region has been opened up and we can begin to try and answer the question of what theory looks like beyond the SM.

Broadly, the LHC physics program can be broken up into four categories. "Supersymmetry" (SUSY) searches look for experimental signatures of supersymmetric physics, one category of Beyond the Standard Model (BSM) models, whilst "exotic" searches aim to discover alternative BSM or non-SM signals such as quantum blackholes. "Higgs" analyses first aimed to discover the hypothesised Higgs boson however, following its discovery in 2012 [4, 5] such analyses now deal with BSM Higgs searches or the measurement of the properties of the new particle. Finally, "Standard Model" (SM) analyses perform precision measurements of SM processes, aiming to discover regions where the SM fails and to better understand the theory.

This thesis describes two such SM measurements, the precision measurement of W boson production in association with jets, $W+$jets, and R_{jets}, a ratio measurement of W and Z boson production in association with jets defined as

$$R_{\text{jets}} = \frac{\sigma(W + \text{jets})}{\sigma(Z/\gamma^* + \text{jets})}. \tag{1}$$

This ratio measurement takes advantage of the similarity in the W and Z production processes to cancel theoretical and experimental uncertainties producing a very precise SM measurement.

The use of these measurements is three-fold. Firstly, they provide a precision test of quantum chromodynamics (QCD) and allow us to test state-of-the-art QCD calculations. Secondly, BSM processes which mimic either W or Z production have the potential to produce tensions between predictions and measurements which may be observed in these SM analyses. Finally, $V+\text{jets}$ processes produce large, irreducible backgrounds to SUSY, exotic and Higgs analyses and so, improved understanding of such backgrounds can feed-back and improve such analyses. In particular, the ratio measurement provides a test of the assumptions made in background estimation methods which use the mapping between $Z/\gamma^*+\text{jets}$ and $W+\text{jets}$ events to provide data-driven estimates of $W + \text{jets}$ backgrounds.

Measurements of $W+\text{jets}$ production in proton-antiproton collisions have previously been carried out by the CDF and D0 collaborations at a centre-of-mass energy of $\sqrt{s} = 1.96$ TeV [6, 7]. More recently, measurements have been carried out at the LHC for $\sqrt{s} = 7$ TeV with an integrated luminosity of 35 pb^{-1} by the ATLAS collaboration [8] and with 5.0 fb^{-1} by the CMS collaboration [9]. The $W + \text{jets}$ measurement presented in this thesis updates and extends such measurements using an integrated luminosity of 4.6 fb^{-1} collected in 2011 and includes comparisons to new state-of-the-art QCD calculations. This measurement has been published in Ref. [2].

The R_{jets} measurement was first proposed in Ref. [10] and the first differential measurement of R_{jets} was carried out by the ATLAS collaboration using 33 pb^{-1}

of data at $\sqrt{s} = 7$ TeV [11], demonstrating the high precision obtainable by such measurements. This measurement was only performed as a function of the leading jet p_{T} in events with exactly one jet. R_{jets} has also been measured by the CMS collaboration using 36 pb^{-1} of data at $\sqrt{s} = 7$ TeV [12] as a function of the jet multiplicity. The R_{jets} measurement presented here extends the R_{jets} measurement to many more differential distributions using an integrated luminosity of 4.6 fb^{-1}. The measurement has been published in Ref. [1].

In this thesis, both measurements will be discussed simultaneously as much of the analysis is shared between the two. Firstly, an overview of the Standard Model will be presented in Chapter 1, including discussion of the state-of-the-art calculations which are used. Next, the experimental setup of the ATLAS detector will be outlined in Chapter 2 and the use of the detector to reconstruct and select physics objects in Chapter 3.

Moving on to the analyses themselves, the estimation of backgrounds will be discussed in Chapter 4 and the detector level uncertainties and results in Chapter 5. In Chapter 6, the procedures employed to correct the measurement and theory predictions will be described. Finally, the results of both the $W +$ jets and R_{jets} measurements will be shown and discussed in Chapter 7. The effects of these measurements on parton distribution function (PDF) fits will be briefly explored in Chapter 8 before final conclusions are drawn and the possibilities for future measurements discussed.

Chapter 1

Theoretical Background

The Standard Model of Particle Physics is a well established theory of high energy particle physics which has been successful in describing observed phenomenon up to an energy scale of $\sim$ 1 TeV. In this Chapter the parts of the Standard Model salient to this thesis will be discussed. In Section 1.1 the Standard Model will be outlined and in Section 1.2 the calculation of cross sections at hadron colliders will be discussed with particular attention to the quantum chromodynamic processes which prevail in hadron colliders. Finally, Section 1.3 outlines the state-of-the-art higher order calculations which are investigated in this thesis.

1.1 The Standard Model

The Standard Model (SM) is a gauge quantum field theory [13, 14] which describes the fundamental particles of nature and their interactions [15, 16]. The SM is described by the symmetries of its Lagrangian which is invariant under transformation of the $\mathrm{SU}(3) \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1)$ gauge group. Each part of this gauge group gives rise to a different force in the SM; the strong force, the weak force and the electromagnetic force respectively.

Each gauge group interaction describes the dynamics of the corresponding force

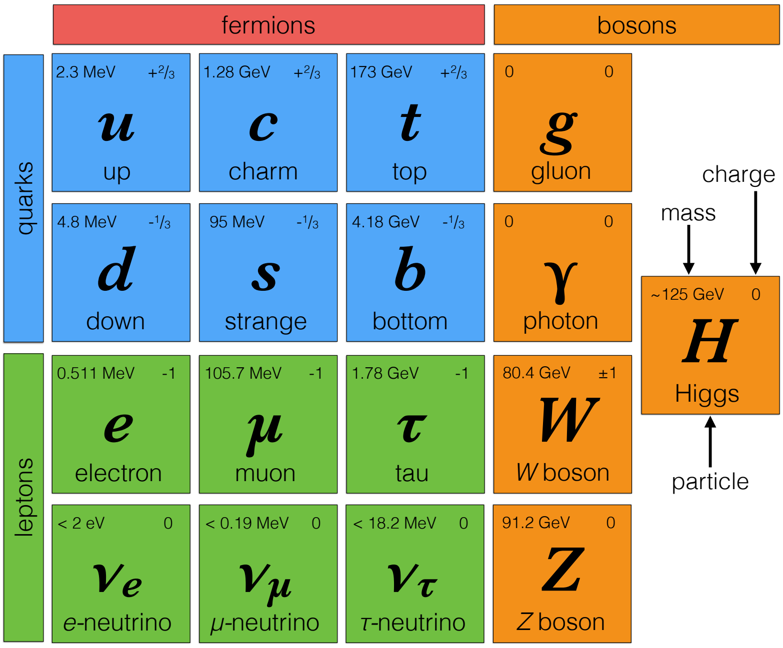


Figure 1.1: The particle content of the Standard Model separated into bosons and fermions. The fermions are also separated into quarks and leptons. Also shown is the mass and charge of each particle taken from Ref. [17].

of nature and describes the forces as mediated by gauge bosons. In the case of the strong force, or Quantum Chromodynamics (QCD), these gauge bosons are the eight gluons which couple to the colour charge of particles. The weak interaction is mediated by the $W^{\pm}$ and Z bosons and the electromagnetic interaction, or Quantum Electrodynamics (QED), is mediated by the photon which couples to electromagnetic charge.

Figure 1.1 shows the full particle content of the SM organised by properties of the particles. Particles are first separated according to their spin. Fermions (half-integer spin) form the matter-like particles in the theory whilst bosons (integer spin) are the charge carriers discussed above. In addition to the force-carrying bosons, an additional boson, the Higgs Boson, is included which comes from electroweak symmetry breaking [18]. The fermions are further separated into quarks and leptons, the former of which participate in strong interactions.

In fact, the electromagnetic and weak interactions are not as distinct as described above. The two are unified in the Glashow-Salam-Weinberg (GSW) model [19–21] which uses a $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ symmetry where the L refers to left-handed and the Y to "hypercharge". The $\mathrm{SU}(2)_L$ symmetry describes three gauge fields $\left(W_\mu^1, W_\mu^2, W_\mu^3\right)$ which couple to the "weak isospin", T, whilst the $\mathrm{U}(1)_Y$ gives rise to a single gauge field, B_μ, which couples to "weak hypercharge", Y. Only left-handed chiral particles couple to the W_μ^i fields [22], hence the L subscript.

In order to reproduce the observed weak and electromagnetic interactions, the observed W_μ^+, W_μ^-, Z_μ and A_μ (photon) fields are obtained by mixing the $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ fields by

$$W_\mu^\pm = \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt{2}}, \tag{1.1}$$

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu, \tag{1.2}$$

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu, \tag{1.3}$$

where θ_W is the Weinberg mixing angle.

In addition, it is observed that the photon is massless whilst the Z and W bosons have large masses (it is this large mass that gives the weak force its weak nature). In the gauge field theory of the SM this turns out to be a problem as the terms which must be added to the Lagrangian to give the field carriers mass break the gauge invariance. This is dealt with by invoking the Higgs mechanism [23–25]. This introduces a massive, spin-0 (scalar) boson with zero electric charge which spontaneously breaks the $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ symmetry, giving W and Z bosons a mass whilst preserving the photon as massless. The existence of such a boson has recently been confirmed at the Large Hadron Collider [4, 5].

In total therefore, the Standard Model is a gauge quantum field theory with a Lagrangian invariant under $\mathrm{SU}(3)_c \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ with an additional massive scalar field which spontaneously breaks the $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ symmetry.

1.2 Calculation of Hadron-Hadron Cross Sections

The calculation of the cross section of processes at hadron colliders can become a complicated procedure. Any such calculation can be separated into two parts. The first part deals with the observed characteristic of confinement [26]. This is the observation that strongly interacting particles (quarks and gluons) are never observed isolated but are always found in composite (colourless) hadrons such as protons. Thus, when performing a calculation of a process at a proton-proton collider one must first consider the probability of finding suitable gluons and quarks (partons) in the protons. The second part of such a calculation is the calculation of the partonic cross section, $\hat{\sigma}$, which describes the probability for the interaction of the partons to produce the process of interest. Combining these two parts, the cross section for a process initiated by two protons with momenta p_1 and p_2 can be expressed as [27]:

$$\sigma\left(p_1, p_2\right) = \sum_{i,j} \int \mathrm{d}x_1 \mathrm{d}x_2 f_i\left(x_1, \mu_F^2\right) f_j\left(x_2, \mu_F^2\right) \hat{\sigma}_{ij}\left(x_1 p_1, x_2 p_2, \alpha_S\left(\mu_R^2\right), \frac{Q^2}{\mu_R^2}, \frac{Q^2}{\mu_F^2}\right). \tag{1.4}$$

This equation will be discussed in the following sections. First, the strong coupling, α_S, will be discussed (Section 1.2.1). Next, the partonic cross section, $\hat{\sigma}\left(...\right)$, will be explained (Section 1.2.2) and the relation of this to the total proton-proton cross section discussed (Section 1.2.3). Finally, the technical implementation of such a calculation for event generation will be described in Section 1.2.4 and the simulation of detector effects in Section 1.2.5.

1.2.1 The Strong Coupling Constant

As will be explained in more detail in Section 1.2.2, the cross section calculation in a perturbative field theory (such as QCD) can be organised in orders of α_S and visualised in the form of Feynman diagrams. Such diagrams may contain loops of particles which often result in divergences in the calculation. Fortunately, such theories can be "renormalised" in such a way so that these divergences are absorbed into renormalised definitions of the coupling constants and particle masses, thereby taking into account loop effects on the physical constants themselves [28]. Such a procedure includes the introduction of an arbitrary scale, the renormalisation scale.

One effect of the choice of renormalisation scale, μ_R, is that the coupling constant, α_S, is no longer a constant but varies as a function of μ_R. Note that although couplings and masses vary, physical observables must not. This variation of α_S can be seen in Figure 1.2 where the running of the coupling constant is clearly visible as the variation in the measured value of α_S with the energy scale, Q. At leading order, one finds that the generic equation for such a running coupling constant can be formulated as [29]

$$\alpha\left(Q^2\right) = \frac{\alpha\left(\mu^2\right)}{1 + \alpha\left(\mu^2\right) b_0 \ln\left(\frac{Q^2}{\mu^2}\right)}. \tag{1.5}$$

In QED, $b_0 = -1/3\pi$ whilst, in QCD, $b_0 = \left(33 - 2n_f\right)/12\pi$ where n_f is the number of quark flavours. The direction of this running behaviour comes from the sign of b_0 which is negative in the QED case and positive in the QCD case (for $n_f < 17$). This sign change is what gives QCD its particular features (discussed below) and arises from the $33/12\pi$ part of b_0 coming from the self-interaction of gluons, itself due to the non-abelian nature of the SU(3) gauge group.

This dependence of α_S on the energy scale explains two of the important features of QCD. Firstly, at low energies (large distances) the value of α_S becomes larger. This means that as two quarks are separated, the potential energy of the system increases until there is enough energy to produce a quark-antiquark pair. This explains the observed phenomenon of confinement by which quarks are only observed in bound states of baryons (qqq) or mesons ($q\bar{q}$) [26]. Secondly, the value of α_S is observed to decrease at high energies (short distances). This is asymptotic freedom and means that short distance QCD interactions can be calculated

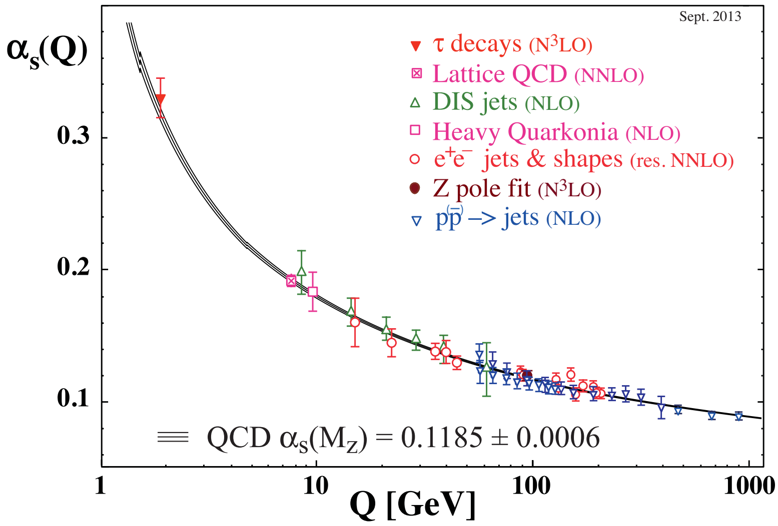


Figure 1.2: Summary of measurements of α_S as a function of the energy scale, Q. Taken from Ref. [17].

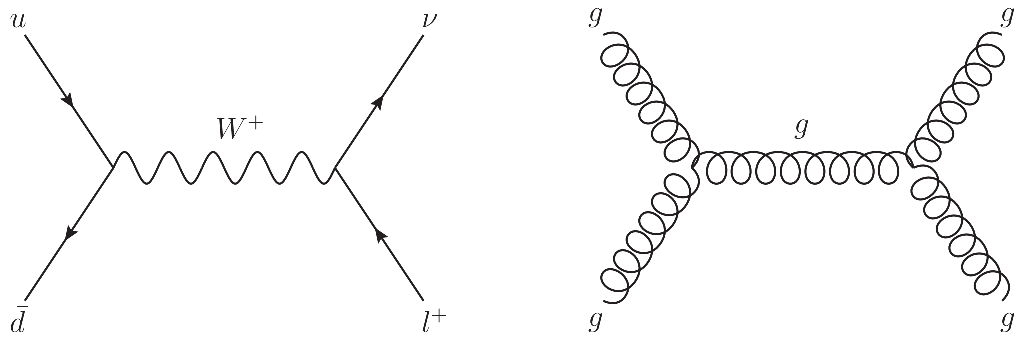
perturbatively [30, 31], as discussed in Section 1.2.2.

1.2.2 Perturbative QCD

When performing a calculation of the partonic cross section, $\hat{\sigma}\,(\ldots)$ in Equation 1.4, one can use Feynman diagrams to visualise the calculation. Such diagrams can then be used to separate the process into parts dependent on the number of factors of α_S which appear in the term related to each diagram as each strong vertex contributes one factor of α_S. The total partonic cross section is then reduced to [27]

$$\hat{\sigma} = \alpha_S^k \sum_{m=0}^{\infty} c^{(m)} \alpha_S^m, \tag{1.6}$$

where the $c^{(m)}$ defines the m-th order contribution to the cross section. In this equation, k is the number of strong vertices in the lowest order diagram (that with fewest strong vertices). For example $k = 0$ for W boson production whereas $k = 2$ for dijet production, the diagrams for which are shown in Figure 1.3. Equation 1.6 defines a perturbative expansion which converges if $\alpha_S \ll 1$. Such a calculation cannot be carried out to infinitely high m and so, the calculation is generally truncated at a



(a) $u\bar{d} \rightarrow l^+\nu$ (b) $gg \rightarrow gg$

Figure 1.3: Leading order Feynman diagrams for W^+ production (a) and dijet production (b) at a hadron collider. No strong interaction vertices are observed in W production ($\hat{\sigma} \propto \alpha_S^0$) whilst dijet production includes two strong interaction vertices ($\hat{\sigma} \propto \alpha_S^2$).

given value of m which defines the order of a calculation. The lowest order calculation is called leading order (LO), the next highest next-to-leading order (NLO), the next highest next-to-next-to-leading order (NNLO) and so on.

As discussed in Section 1.2.1, a method of dealing with divergences introduced by loops in Feynman diagrams introduces an arbitrary renormalisation scale, μ_R. Observable quantities such as cross sections cannot be dependent on the choice of this scale, however the truncation of the infinite perturbative series produces a residual dependence. Mathematically, this can be related to the size of missing higher order terms by

$$\frac{\mathrm{d}}{\mathrm{d}\mu_R}\hat{\sigma} = \alpha_S^k \frac{\mathrm{d}}{\mathrm{d}\mu_R} \sum_{m=0}^{\infty} c^{(m)} \alpha_S^m = 0, \tag{1.7}$$

from which it follows that if we truncate at $m = n$,

$$\frac{\mathrm{d}}{\mathrm{d}\mu_R} \sum_{m=0}^{m=n} c^{(m)} \alpha_S^m + \frac{\mathrm{d}}{\mathrm{d}\mu_R} \sum_{m=n+1}^{\infty} c^{(m)} \alpha_S^m = 0, \tag{1.8}$$

which can be rearranged to give

$$\frac{\mathrm{d}}{\mathrm{d}\mu_R} \sum_{m=0}^{m=n} c^{(m)} \alpha_S^m = \mathcal{O}\left(\alpha_S^{n+1}\right). \tag{1.9}$$

Thus, we see that the variation in the partonic cross section (calculated to order n)

due to the renormalisation scale choice gives an estimate of the size of the next term of the perturbative expansion.

1.2.3 Parton distribution functions

Having calculated the partonic cross section, this must be related to the total cross section for proton-proton interactions. This is done using the parton density functions, $f_i\left(x_1, \mu_F^2\right)$ and $f_j\left(x_2, \mu_F^2\right)$ in Equation 1.4, which define the probability to find a parton with a fraction x_n of the total proton momentum (this interpretation is valid at LO, at higher orders the interpretation is more complex). $xf_i\left(x_1, \mu_F^2\right)$ and $xf_j\left(x_2, \mu_F^2\right)$ are called the parton distribution functions (PDFs) and define the momentum distribution of partons in the proton. The PDFs are seen to be dependent on μ_F which is the factorisation scale. By analogy with the renomalisation scale, discussed above, this is an arbitrarily chosen scale of which the cross section is independent if the perturbative calculation is carried out to all orders.

The factorisation scale defines whether a parton emission is treated within the PDF or the partonic cross section part of Equation 1.4. A parton emitted at low transverse momentum ($< \mu_F$) is absorbed into the PDF definition and thereby treated as part of the hadronic structure of the proton whilst a parton emitted at high transverse momentum ($> \mu_F$) is treated in the partonic cross section. The simplifying assumption of a single scale is often used to set factorisation and renormalisation scales as equal.

PDFs cannot be calculated directly and, as such, a fit must be performed to data [29]. Generally, a physically motivated ansatz is chosen to parameterise the PDFs as a function of x at a given starting scale. The Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations which deal with parton splitting and emission can then be used to evolve the PDF to arbitrary scales [32–34]. These PDFs are then used to perform fits to data which constrains the parameters of the initial PDF. Such

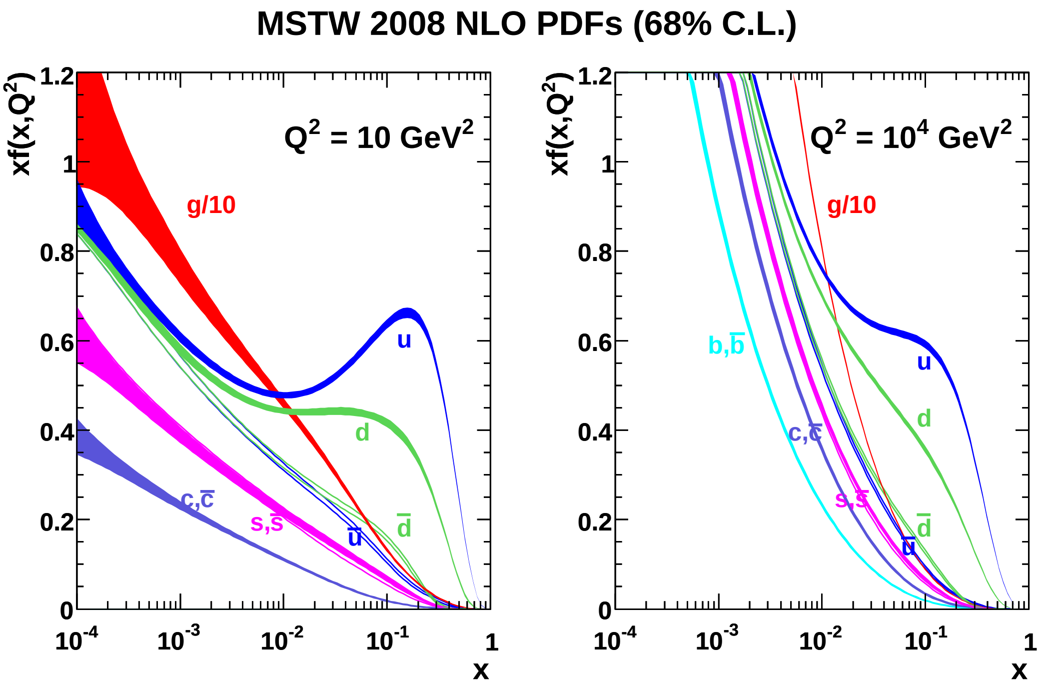


Figure 1.4: The MSTW 2008 NLO PDF set at two different scales. Taken from Ref. [35].

PDF fits are performed by global fitting groups such as MSTW [35], CTEQ [36] and NNPDF [37] and use a range of data and assumptions to constrain the PDFs. Much information can be extracted from data taken with the H1 and ZEUS experiments at the 6.3 km HERA accelerator, which collided 27.5 GeV electrons or positrons with protons up to 920 GeV. The data from these experiments allows a wide range of Q^2 and x values to be explored. A PDF set, HERAPDF, has been produced using these data alone [38, 39]. An example of a PDF set is shown in Figure 1.4 which shows the MSTW 2008 NLO PDF set at two scales.

To calculate the total observed cross section in Equation 1.4, one must integrate over x_1 and x_2 whilst applying any kinematic constraints such as energy conservation which relate x_1 and x_2. In addition, one must sum over all possible pairings of partons which can produce the required final state without increasing the order of the calculation.

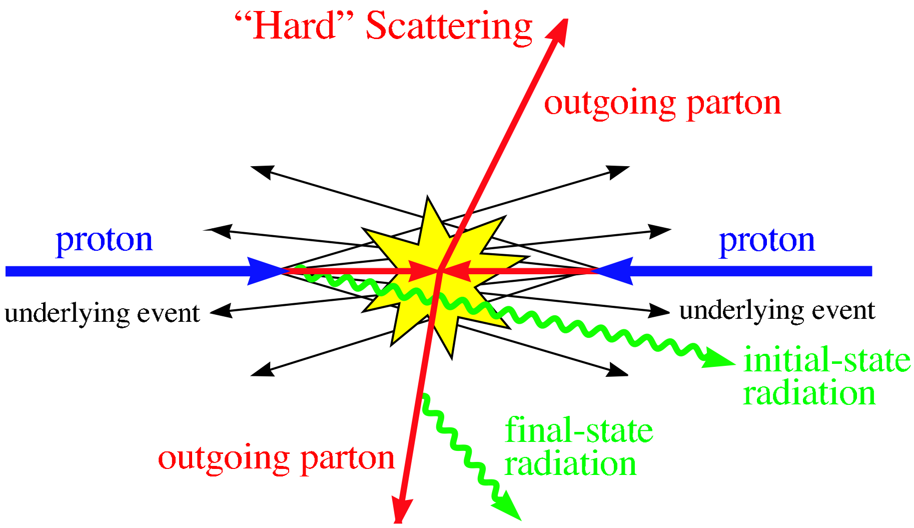


Figure 1.5: Schematic cartoon of a 2 → 2 hard scattering event generated using Monte Carlo. Taken from Ref. [41].

1.2.4 Monte Carlo Event Generation

The calculation described above (Equation 1.4) provides a prediction of a cross section however, for use in analyses, we require a more detailed simulation of physical observables [40]. For example, the perturbative calculation we have used so far may involve unstable particles which must be decayed further or isolated quarks which must be "hadronised" to form the particles found in the detector. In addition, proton remnants remain after the hard scatter and this must be related to observation. Such issues are generally dealt with by Monte Carlo generators. A schematic cartoon of an event generated using Monte Carlo is shown in Figure 1.5.

First, the hard scatter is calculated using perturbative QCD at a fixed order using incoming parton momenta taken from a chosen PDF. Next, a parton showering (PS) algorithm is used to simulate the effect of soft and collinear radiation on partons leaving the hard scatter. Sequential stochastic branching is used to model the splitting and emission of partons until the scale of the outgoing partons is evolved from the hard scatter scale to the hadronisation scale of 1 GeV. At this scale, a non-perturbative hadronisation model is employed to model the confinement (hadronisation) of coloured partons to form colourless hadrons. The combination of parton shower and hadronisation model allows Monte Carlo generators to model the collimated showers of hadrons ("jets") which are observed in hadron physics (although also produced at e^+e^- colliders). Such showers are characteristic of processes with strongly interacting partons (quarks or gluons) in the final state as such particles produce collimated showers of particles. Combining these showers into jets, as will be discussed in Section 3.4, allows the showers measured in the detector to be related back to the partons in the final state.

By analogy to the final state radiation (FSR) produced by the parton shower, initial state radiation (ISR) is similarly dealt with by parton showering. This models the effect of soft and collinear radiation coming from the initial state partons. The necessity for the partons entering the hard scatter perturbative calculation to have the correct momentum requires the parton shower evolution to be done in reverse.

Similarly to the QCD FSR which is dealt with by parton showering, there is a chance of charged particles emitting photons at any stage of the event generation. Such emissions are dealt with by the QED FSR routines in Monte Carlo generators.

Finally, the underlying event (UE) which includes the additional activity coming from the proton remnants and multiple parton interactions (MPI) must be included. This is done using non-perturbative QCD models which are tuned to experimental data.

Generally, the perturbative matrix element (ME) part of the Monte Carlo generation is calculated at LO, NLO or possibly NNLO and, although a parton shower describes soft and collinear emissions, it falls short when attempting to model hard, isolated partons. In order to improve modelling up to high jet multiplicities, such as those investigated in this thesis, Monte Carlos may be enhanced by the "multi-leg" procedure. In such generators, the usual ME is augmented by additional MEs with additional external legs. One problem produced by such an approach is that of double counting whereby a $W + j$ event can be produced either by a $W + j$ ME calculation (the calculation with one additional leg) or by a W ME calculation with an additional parton produced by the parton shower. Such double counting is overcome using a ME-PS matching procedure which includes a matching-scale which defines the transition between partons modelled in the ME and partons modelled by the PS.

In this thesis, two LO multi-leg Monte Carlo generators are used to model the $W+\text{jets}$ and $Z/\gamma^*+\text{jets}$ signal. The first of these is ALPGEN 2.13 [42] interfaced to HERWIG 6.520 [43] for (angular ordered) PS, JIMMY 4.31 [44] for MPI and using CTEQ6L1 [45] PDFs and the AUET2-CTEQ6L1 tune [46]. ALPGEN utilises the MLM [47] scheme to perform ME-PS matching and includes ME for up to 5 additional partons. The second is SHERPA 1.4.1 [48, 49] using CT10 PDFs [36] which uses the CKKW-L [50, 51] matching scheme, a dipole parton shower (ordered in transverse momentum) and up to 4 additional partons in the ME. In addition to the differences listed above, ALPGEN and SHERPA use different QED FSR models, PHOTOS [52] and a YFS-based method [53] respectively. Both ALPGEN and SHERPA utilise a cluster hadronisation model.

1.2.5 Detector and Pile-up Simulation

In order to simulate the interactions of particles as they propagate through the ATLAS detector, generated event records from Monte Carlo simulation are passed through a full simulation of the ATLAS detector using GEANT4 [54]. This provides a prediction which can be compared directly to data (detector level).

In addition, the effect of multiple proton-proton collisions in a single bunch crossing, "pile-up", is simulated by the addition of minimum bias (low p_{T} QCD) events with exactly one inelastic collision generated using PYTHIA 6 [55]. The number of overlaid minimum bias events is sampled according to a Poisson distribution with mean μ. The bunch-filling pattern surrounding the signal interaction is used to modulate this number so that only filled bunch-crossings are simulated. The number of simulated bunch crossings is different for each detector subsystem to take into account their different sensitivities to pile-up (some sub-detectors have a longer readout time than the bunch spacing and so are affected by both preceding and subsequent bunch crossings).

Monte Carlo is generated in 4 periods for comparison to the data, corresponding to changes in the detector and accelerator conditions. Monte Carlo events are reweighted to reproduce the distribution of the average number of interactions per bunch crossing, μ, observed in data (this distribution will be discussed in Section 2.1).

1.3 Higher order calculations

The measurements performed in this thesis explore a wide range of phase spaces and allow a stringent test of the state-of-the-art QCD calculations discussed here. These represent next-to-leading order (NLO) calculations from BlackHat+Sherpa (Section 1.3.1) approximate next-to-next-to-leading order[1] (nNLO) calculations (Sections 1.3.2 and 1.3.3), a resummation approach (Section 1.3.4) and rigorous combination of NLO and LO matrix element information in Monte Carlo (Section 1.3.5).

1.3.1 BlackHat+Sherpa

Next-to-leading order predictions are obtained from BlackHat+Sherpa [56–58] at parton level for inclusive $V+\geq n$ jets where n ranges from 0 to 4 for Z and from 0 to 5 for W. The virtual matrix element (the part of the NLO calculation which includes loops) is provided by the BlackHat program whilst tree-level diagrams (LO part and NLO real emission) are provided by Sherpa.

[1] lower case n is used to differentiate approximate NNLO from full NNLO which is not yet available

1.3.2 Exclusive sums

The BlackHat+Sherpa matrix element calculations are also used in an exclusive sums approach [59] whereby NLO information from different jet multiplicities is combined to produce an nNLO calculation. In this case, the $W+ \geq 1$ jet exclusive sum calculation is produced from the sum of $W + 1$ jet and $W+ \geq 2$ jets at NLO. This approach is useful to approximately include contributions from higher jet multiplicities that would not usually be present in a rigorous NLO calculation. The inclusion of such terms had previously been demonstrated to improve agreement between data and theory for several kinematic distributions [8].

1.3.3 LoopSim

The LoopSim technique [60] is another nNLO procedure which allows the merging of NLO samples of different jet multiplicities. LoopSim takes existing $W+1$ jet and $W + 2$ jet results and supplements them with a unitarity-based approximation of the missing higher order loop terms. Such a method is expected to provide predictions close to true NNLO when cross sections are dominated by large contributions associated with new scattering topologies that appear at NLO or beyond.

1.3.4 High Energy Jets

High Energy Jets (HEJ) [61, 62] is based on a LO matrix element but uses a perturbative resummation to give an approximation to the the hard-scattering matrix element for jet multiplicities of up to two or greater and to all orders in α_S. The approximation used becomes exact in the limit of large rapidity separation between partons.

1.3.5 MEPS@NLO

NLO particle level predictions are obtained from MEPS@NLO [48,49] which utilises virtual $W+1$ jet and $W+2$ jet virtual matrix elements from BlackHat merged with LO matrix elements for W events with up to four jets. The final states are matched to a PS and hadronised using Sherpa. Such a formalism is a rigorous method of combining NLO and LO matrix element information from different multiplicities in a Monte Carlo generator.

Chapter 2

The ATLAS Experiment

The ATLAS detector is a general purpose detector designed to study the collisions of high energy protons[1] provided by the Large Hadron Collider (LHC) at the European Organisation for Nuclear Research (CERN) in Geneva, Switzerland. In this Chapter, the LHC will be discussed (Section 2.1) and the ATLAS detector will be described (Section 2.2).

2.1 The Large Hadron Collider

The Large Hadron Collider [63] is a particle accelerator installed in a 27 km circular tunnel under the French-Swiss border, near Geneva. It is designed to accelerate protons or lead ions and collide them at four points around the accelerator. These four points house the four main detectors; ATLAS [64], CMS [65], LHC*b* [66] and ALICE [67]. The LHC is designed to provide collisions at a maximum centre of mass energy ($\sqrt{s}$) of 14 TeV with an instantaneous luminosity, L, of $\sim 10^{34}$ cm^{-2}s^{-1} where L is given by

$$L = \frac{1}{\sigma}\frac{\mathrm{d}N}{\mathrm{d}t}. \tag{2.1}$$

[1] or lead ions

σ is the interaction cross-section and N is the number of events detected in a certain time (t). The integrated luminosity, $\mathcal{L}$, is related to L and is given by

$$\mathcal{L} = \int L \, \mathrm{d}t \tag{2.2}$$

so that the total number of expected events, N_{tot} over a period of data-taking is

$$N_{\text{tot}} = \mathcal{L} \times \sigma. \tag{2.3}$$

In order to achieve energies of up to 14 TeV, protons are accelerated in stages through the LHC accelerator complex. Protons are initially produced from hydrogen gas by stripping the electrons off the atoms and are fed into the Linear Accelerator 2 (LINAC 2) which accelerates them to 50 MeV. From LINAC 2 the protons are injected into the Proton Synchrotron Booster which accelerates them to 1.4 GeV and then into the Proton Synchrotron (PS). Following the PS the protons now have an energy of 26 GeV and are injected into the 7 km Super Proton Synchrotron (SPS) and accelerated to 450 GeV. It is only after this chain that the protons are finally injected into the LHC and accelerated to their full energy. This accelerator chain is shown schematically in Figure 2.1.

To date, the LHC has not been run at full design energy or luminosity. In 2010, an integrated luminsoity of 0.047 fb^{-1} of proton-proton collisions was delivered at an energy of 7 TeV. In 2011, this dataset was extended by an additional 5.46 fb^{-1}. Finally, in 2012, the energy of the LHC was increased to 8 TeV and 22.8 fb^{-1} of proton-proton collisions was delivered [69]. The peak luminosity over these three periods of Run 1 data was 7.7×10^{33} $\text{cm}^{-2}\text{s}^{-1}$ in 2012.

One complication with the high luminosities produced by the LHC is that in each bunch crossing there is a high probability of multiple proton-proton collisions occurring. The number of proton-proton interactions per bunch crossing forms a

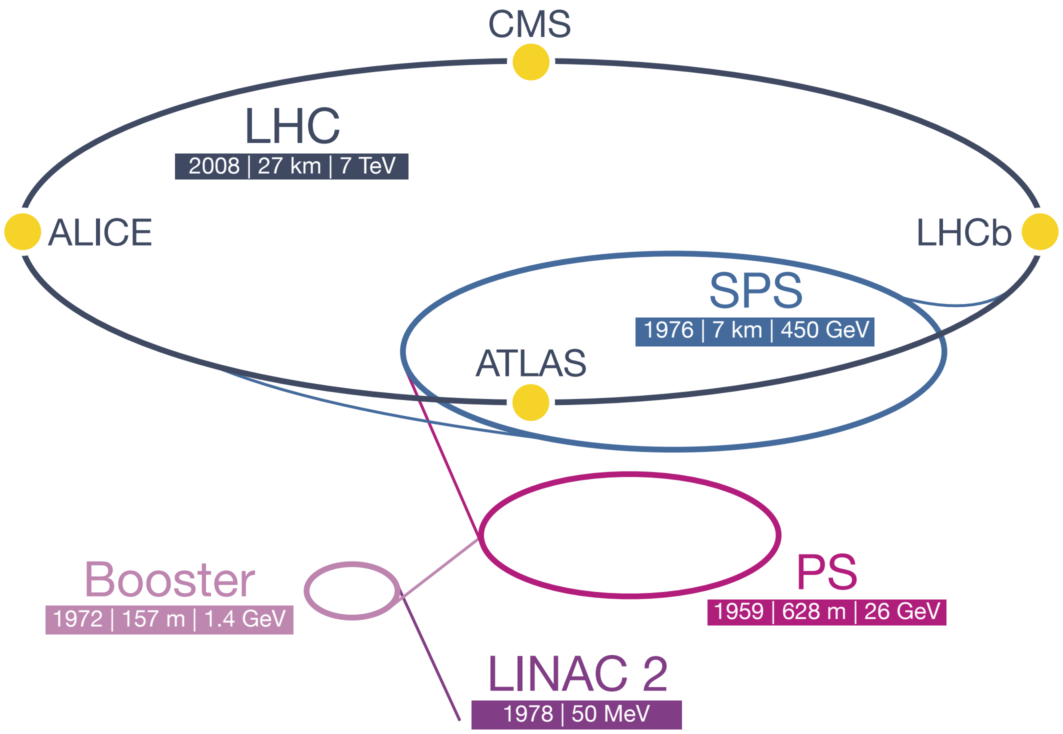


Figure 2.1: The LHC accelerator complex at CERN. Protons are accelerated in steps given by the energies indicated in the labelling. Also shown are the four main LHC detectors. Taken from Ref. [68].

Poisson distribution with mean μ. μ is related to the instantaneous luminosity, L, by

$$\mu = \frac{L \times \sigma_{\mathrm{inel}}}{n_{\mathrm{b}} \times f_{\mathrm{rev}}} \tag{2.4}$$

where σ_{inel} is the inelastic proton-proton cross section, n_{b} is the number of bunches per beam and f_{rev} is the revolution frequency of the LHC. Figure 2.2 shows the recorded luminosity as a function of μ for the 2011 and 2012 data-taking periods.

In this thesis, only the data taken at 7 TeV in 2011 were used, corresponding to 5.46 fb^{-1} of data [71, 72], of which 5.08 fb^{-1} were recorded by the ATLAS detector and 4.57 fb^{-1} have been deemed good for physics (all reconstructed physics objects of good quality). The total integrated luminosity as a function of time over the 2011 data-taking period is shown in Figure 2.3. This dataset has a peak luminosity of 3.6×10^{33} cm^{-2}s^{-1} and the mean number of interactions per bunch crossing for this period ranges from 2 to 19.

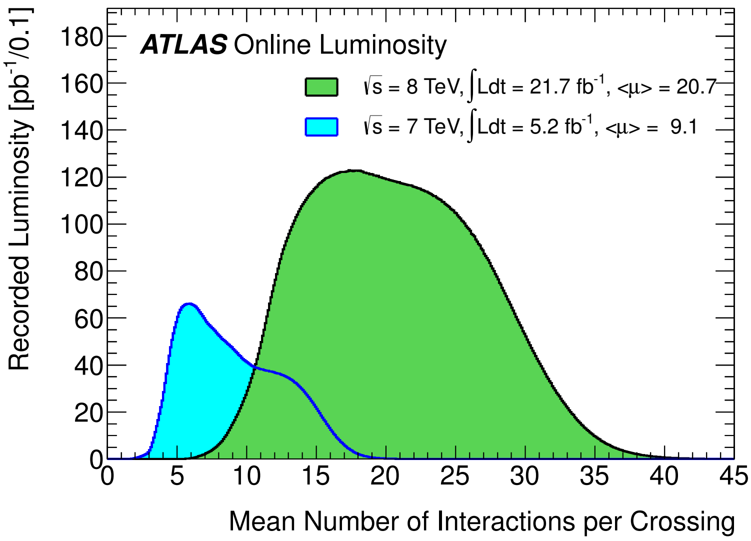


Figure 2.2: The mean number of interactions per bunch crossing, μ, for the 2011 and 2012 data-taking periods. Taken from Ref. [70].

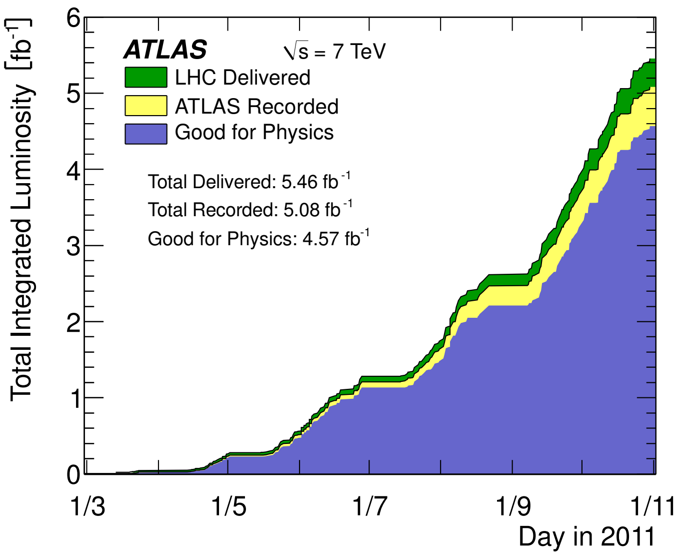


Figure 2.3: The total integrated luminosity as a function of time during 2011 data taking. The total delivered luminosity, the total recorded luminosity and the total luminosity recorded and suitable for analysis are shown. Taken from Ref. [70].

Subdetector	Resolution
Inner detector	$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = 0.03\% \times p_\mathrm{T} \oplus 1.5\%$
Electromagnetic calorimeter	$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 1\%$
Hadronic calorimeter	$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$
Muon spectrometer	$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = 4\%$ at $p_\mathrm{T} = 100$ GeV

Table 2.1: The performance of each subdetector of the ATLAS experiment in the central region. More details of each subdetector can be found in the relevant section of the text. In the calorimeter resolutions, the noise term has been omitted as this is highly dependent on pile-up.

2.2 The ATLAS Detector

The ATLAS detector [64] is a multi-purpose detector with a symmetric cylindrical geometry and nearly 4π coverage in solid angle. Figure 2.4 shows the overall layout of the detector. The separate components of the detector are discussed below in order of increasing radii. The inner tracking system is described in Section 2.2.1 and provides precision tracking of charged particles within the solenoid magnet. Outside the solenoid magnet is located the electromagnetic and hadronic calorimetry, discussed in Section 2.2.2. Finally, the muon spectrometer provides tracking of muons in a toroidal magnetic field and is described in Section 2.2.3.

The performance of each subdetector system of the ATLAS detector is shown in Table 2.1. The tracking and muon spectrometer performance is given in terms of the fractional resolution on the momentum measurement whilst the calorimeter performance is given in terms of energy resolution. In the resolution for the calorimeters the noise term which dominates at low energy has been omitted as this is highly dependent on the level of pile-up.

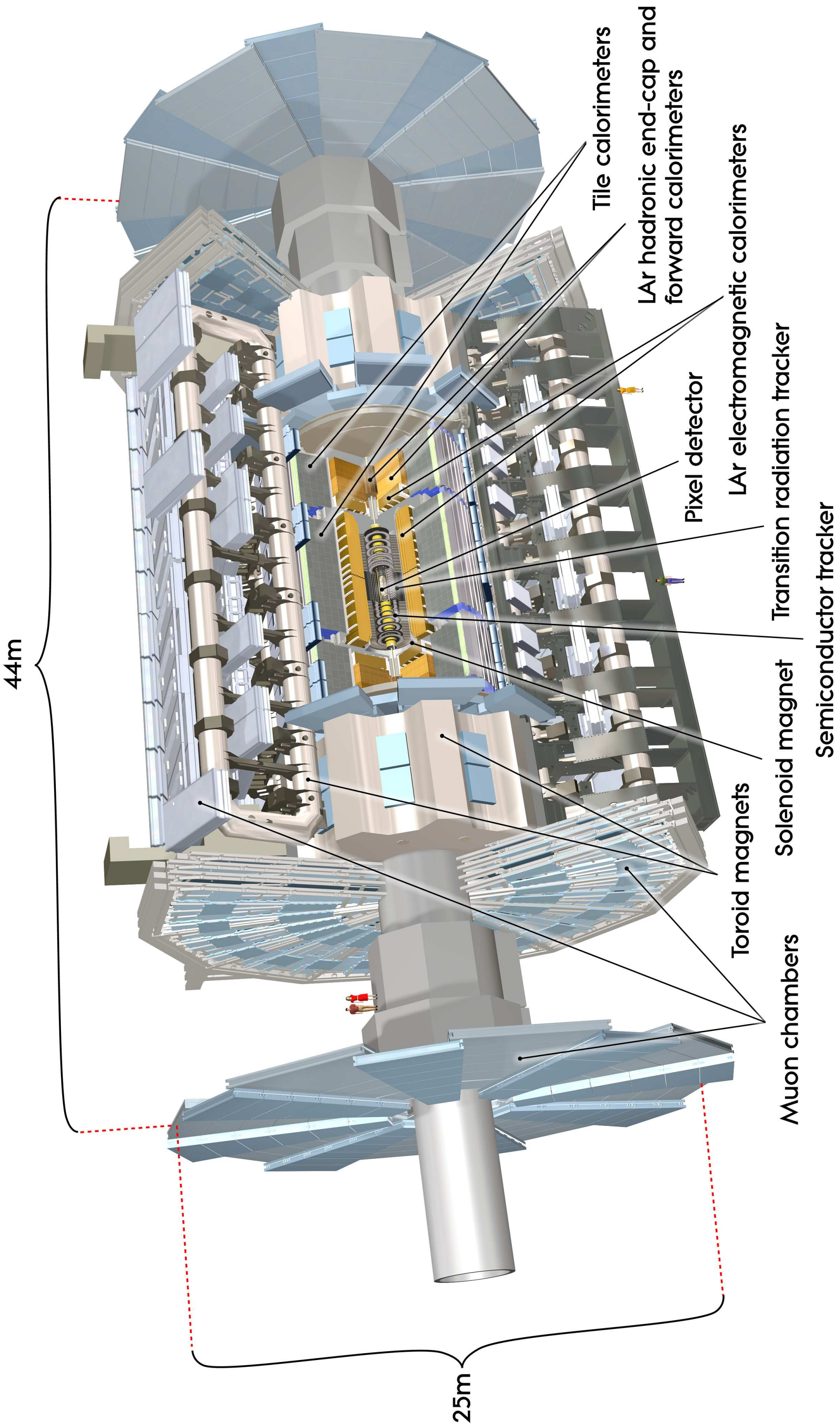


Figure 2.4: Cut-away of the ATLAS detector. The overall weight of the detector is approximately 7000 tonnes. Taken from Ref. [64].

2.2.1 Charged particle tracking and solenoid

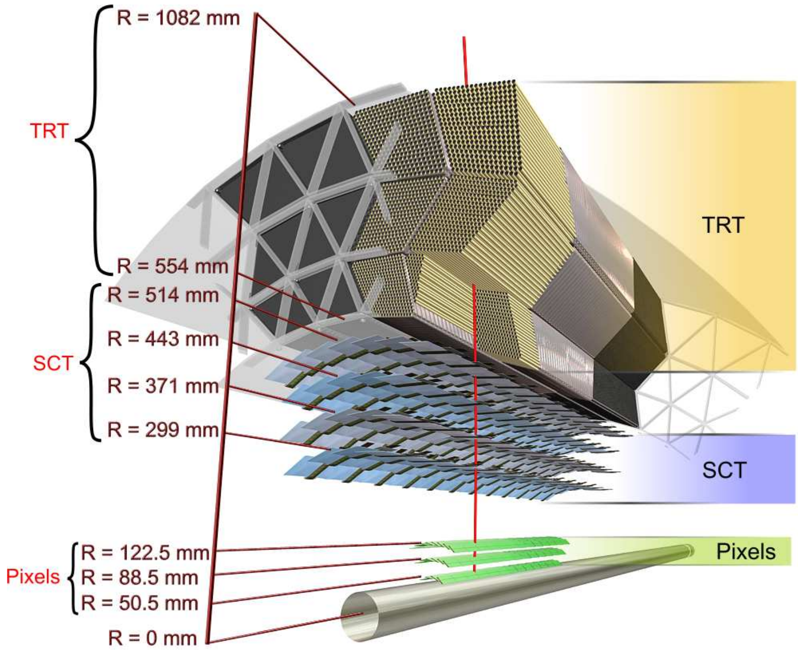
The layout of the Inner Detector (ID) [73,74] is shown in Figure 2.5. The ID records three-dimensional coordinates along the path of charged particles in the pseudorapidity range $|\eta| < 2.5$. Such points are used by pattern recognition algorithms to reconstruct particle tracks. These can be used to derive the charge and momentum of the particle due to curvature in a solenoidal magnetic field. In addition, such tracks can be used to reconstruct primary and secondary vertices which can be used for heavy flavour jet tagging and removing the effects of pile-up.

The ID consists of three independent sub-detectors, which are discussed below, surrounded by a superconducting solenoid magnet producing a 2 T magnetic field. The design of the ID is a balance between a low mass detector (to avoid unwanted interactions of particles with the detector material) and high spatial resolution. The detector elements closest to the interaction point have the highest granularity to overcome the higher particle flux whilst those at larger radius can afford a coarser granularity.

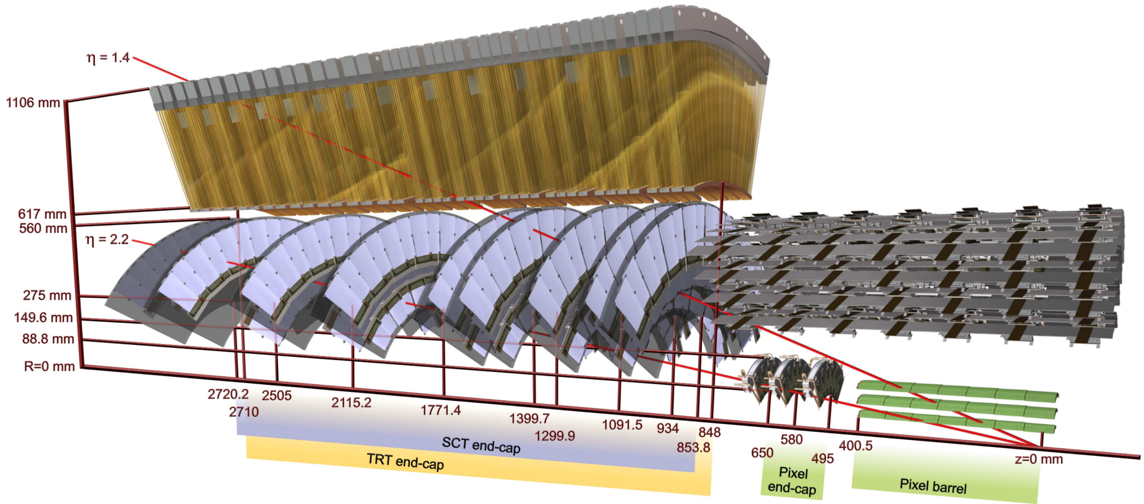
At inner radii close to the interaction point, high-resolution positional measurements are made possible by silicon pixel detectors [75]. The system provides at least three precision measurements per track over the full ID acceptance. In the central region the pixel system consists of three barrels at radii of 50.5 mm, 88.5 mm and 122.5 mm. Three disks on each side form the pixel end-caps and complete the angular coverage.

The nominal pixel size is dictated by the readout pitch of the front-end (FE) electronics and is 50×400 μm^2 for 90% of the pixels or 50×600 μm^2 for pixels in the regions containing the FE module chips. The pixel detector is built from 1744 pixel sensors with 47232 pixels on each sensor.

At intermediate radii a silicon microstrip detector, the semiconductor tracker



(a) Barrel



(b) End-cap

Figure 2.5: Drawing showing the sensors and structural elements traversed by a charged particle in the inner detector in (a) the barrel region ($\eta = 0.3$) and (b) the end-cap region ($\eta = 1.4$ and $\eta = 2.2$). Taken from Ref. [64].

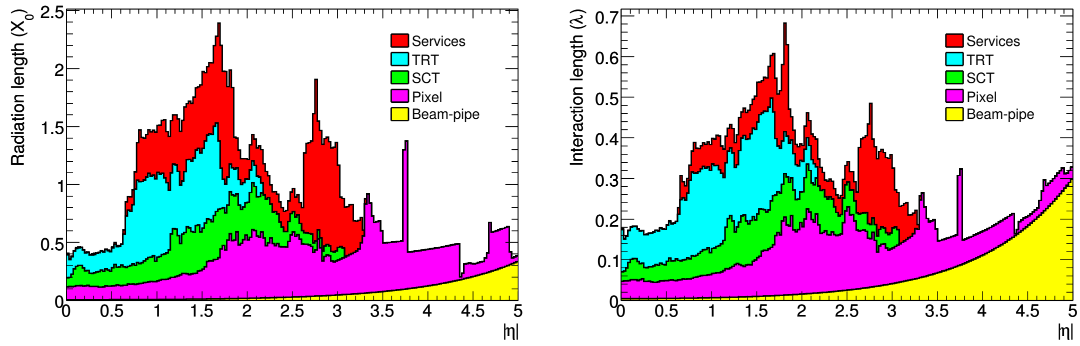
(SCT), provides eight precision measurements per track [76, 77]. The barrel region consists of four layers, each with silicon microstrip detectors on both sides (forming eight single layers of silicon) at 299 mm, 371 mm, 443 mm and 514 mm. The end-cap region consists of nine disks which are also double-sided.

The strip pitch in the SCT is determined by required digitising precision, granularity, occupancy and noise performance. In the barrel the strip pitch is 80 μm with two 6 cm strips connected to form strips of 12 cm. Each SCT sensor has 768 active strips and two such detectors are glued back-to-back at a small angle to provide z coordinate measurement. $R\phi$ coordinates are provided by the segmentation of the strips themselves. The SCT comprises 15912 sensors in total.

The outer layer of the ID is formed of the Transition Radiation Tracker (TRT), built of 4 mm diameter drift (straw) tubes [78]. The barrel region is built of modules containing between 329 and 793 straws covering the radial range from 554 mm to 1082 mm with straws arranged parallel to the beam. In the end-cap region, straws are arranged radially. As a result the TRT provides tracking in the $R - \phi$ plane in the barrel and the $z - \phi$ plane in the end-cap. The straws are filled with a Xe-CO_2-O_2 mixture and polypropylene fibres (foils) are placed between the straws in the barrel (end-cap).

The addition of polypropylene between TRT straws produces transition radiation (TR) [79]. TR is produced in proportion to $\frac{1}{\text{mass}}$ for a fixed energy and so, significantly more TR is produced by electrons compared to heavier particles such as pions. The photons from TR are absorbed by the Xe gas in the straws and produce larger signals than other particles. The front end electronics can discriminate between low and high threshold hits, allowing the TRT to be used for both tracking and electron identification purposes.

Figure 2.6 shows the material distribution in the ID as a function of pseudorapid-



(a) Radiation length (b) Interaction length

Figure 2.6: The material distribution in the inner detector as a function of pseudorapidity in terms of (electromagnetic) radiation length and (nuclear) interaction length. The breakdown indicates the contributions of external services and of individual sub-detectors. Taken from Ref. [64].

ity in terms of (electromagnetic) radiation length[2] and (nuclear) interaction length[3]. The most striking features are the onset of non-active and structural material at the interface between barrel and end-cap and the pixel services at $|\eta| > 2.7$ which leave along the beam-pipe. This also shows that a large fraction of the services and structural material is external to the active ID envelope and therefore does not affect tracking but does affect the calorimetry.

2.2.2 Calorimetry

The ATLAS calorimeters are shown in Figure 2.7. The calorimeter system consists of an electromagnetic (EM) calorimeter covering the region $|\eta| < 3.2$, a hadronic barrel calorimeter covering the region $|\eta| < 1.7$, hadronic end-cap calorimeters covering the regions $1.5 < |\eta| < 3.2$ and forward calorimeters covering the regions $3.1 < |\eta| < 4.9$ [80, 81].

In order to contain hadronic showers the calorimeter system is designed to have ~ 10 interaction lengths of material across the full range of pseudorapidity. The amount of material as a function of pseudorapidity is shown in Figure 2.8.

[2] Radiation length is the distance an electron must travel to loose $\frac{1}{e}$ of its energy.
[3] Interaction length is the distance a hadron must travel to loose $\frac{1}{e}$ of its energy.

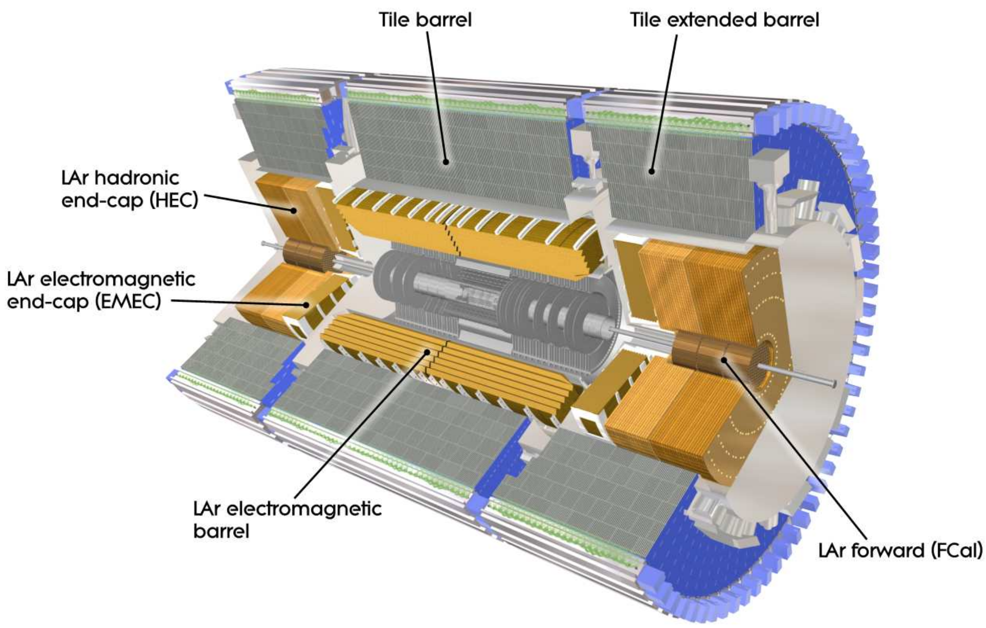


Figure 2.7: Cut-away view of the ATLAS calorimeter system. Taken from Ref. [64].

Electromagnetic calorimeter

The electromagnetic calorimeter (ECal) is divided into a barrel part covering the region $|\eta| < 1.475$ [82] and two end-caps covering $1.375 < |\eta| < 3.2$ [83]. The barrel is split into two halves with a small gap at $z = 0$ whilst the end-caps are separated into wheels covering $1.375 < |\eta| < 2.5$ (outer wheel) and $2.5 < |\eta| < 3.2$ (inner wheel).

The ECal uses a lead absorber and liquid argon (LAr) active material. The electrodes have an accordion geometry which ensures uniform coverage in ϕ. In the barrel region, the accordion waves are axial and run in ϕ with the folding angles of the waves increasing with radius to keep the LAr gap constant. This is shown schematically in Figure 2.9. In the end-caps, the waves are parallel to the radial direction and run axially.

The total thickness of the ECal ranges from 22 to 38 radiation lengths dependent on η and therefore is expected to fully contain electromagnetic showers from electrons and photons. It is read out in three layers of $\eta - \phi$ cells with varying degrees of

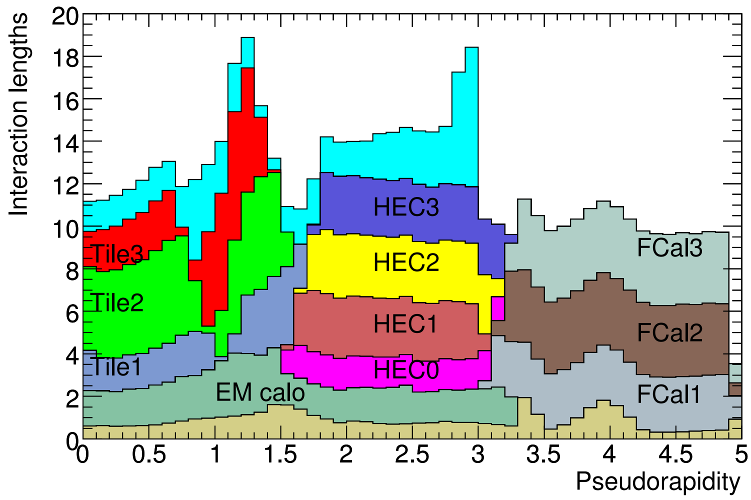


Figure 2.8: Cumulative amount of material in units of interaction length as a function of pseudorapidity in the ATLAS calorimeter system. Included is the material budget in front of the electromagnetic calorimeter and the total amount of material in front of the first active layer of the muon spectrometer. Taken from Ref. [64].

segmentation. In particular, the first layer has very fine η segmentation to allow identification of $\pi^0 \rightarrow \gamma\gamma$.

Within the inner radius of the ECal is a separate 11mm LAr layer. This acts as a presampler and enables an estimation of the energy lost by electrons and photons before encountering the ECal. This is possible as it has no dedicated absorber and so measures the showers produced by material closer to the interaction point. This presampler covers the region $|\eta| < 1.8$ [84].

Hadronic calorimeters

Outside the ECal lies the hadronic calorimeter (HCal). In the barrel region ($|\eta| < 1.7$) the HCal is formed of the tile calorimeter which is a sampling calorimeter using steel as the absorber and scintillator as the active medium [85]. The tile calorimeter is subdivided into a central barrel and two extended barrel regions [86]. Azimuthal coverage is maximised by orienting the scintillator tiles radially and normal to the beamline.

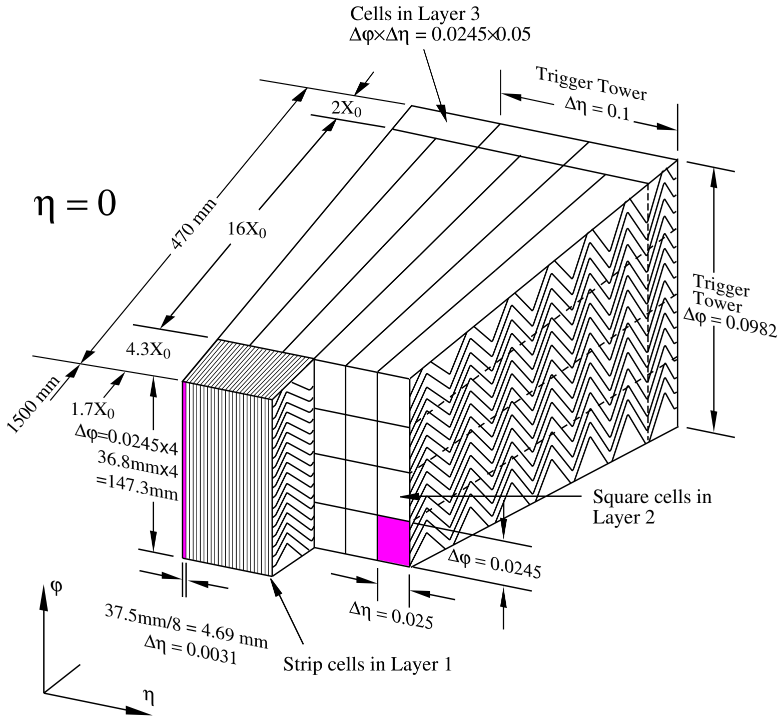


Figure 2.9: Schematic of a barrel module of the electromagnetic liquid argon calorimeter. The granularity in η and ϕ of the cells in each of the three layers is shown. Taken from Ref. [64].

Scintillator tiles are read out using wavelength-shifting fibres and are grouped together at photo-multiplier tubes (PMTs) at the edge of each module. The fibre grouping defines a three dimensional cell structure with dimensions $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the inner two layers and 0.2×0.1 in the third layer. The $R - \eta$ segmentation of the tile calorimeter is shown schematically in Figure 2.10.

In the region covering $1.5 < |\eta| < 3.2$, hadronic calorimetry is provided by the hadronic end-cap (HEC) calorimeter [87]. This is a LAr based sampling calorimeter with copper used as the absorber. This LAr calorimeter has a flat-plate design with the absorber plates arranged perpendicular to the beam axis. The HEC is formed of two end-cap wheels, HEC1 and HEC2, which provide a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for $|\eta| < 2.5$ and 0.2×0.2 outside this.

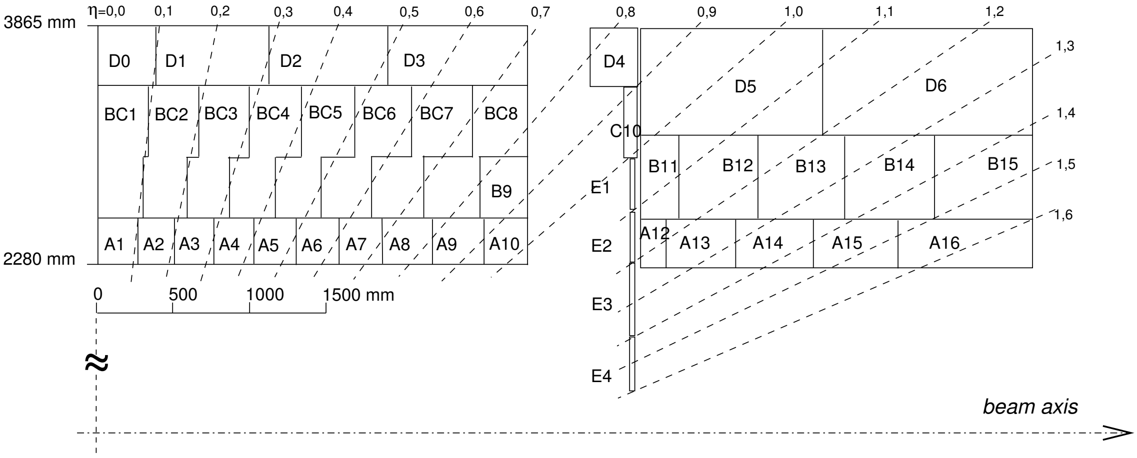


Figure 2.10: Segmentation of the tile calorimeter modules in the central (left) and extended (right) barrels. Taken from Ref. [64].

Forward calorimeter

Forward calorimetry is provided by the forward calorimeter (FCal) which covers the regions $3.1 < |\eta| < 4.9$ [88]. Due to the very forward position of the FCal it is exposed to high particle fluxes. An electrode structure is used with small-diameter rods centred in tubes parallel to the beamline as shown in Figure 2.11. The active medium is again LAr but with much smaller LAr gaps to avoid ion build-up.

The FCal is split into three modules, one electromagnetic (FCal1) and two hadronic (FCal2 and FCal3) as shown in Figure 2.12. FCal1 uses a copper absorber due to its heat removal capabilities which are advantageous given the high particle flux. FCal2 and FCal3 primarily use tungsten absorbers due to their high density which limits the lateral spread of hadronic showers. The FCal granularity is $\sim 0.2 \times 0.2$ in $\eta - \phi$.

2.2.3 Muon spectrometer and toroid

Outside the calorimetry systems is located the muon spectrometer (MS) which detects charged particles exiting the calorimeter and provides a measure of their momentum in the region $|\eta| < 2.7$ [89]. The muon spectrometer is located on and

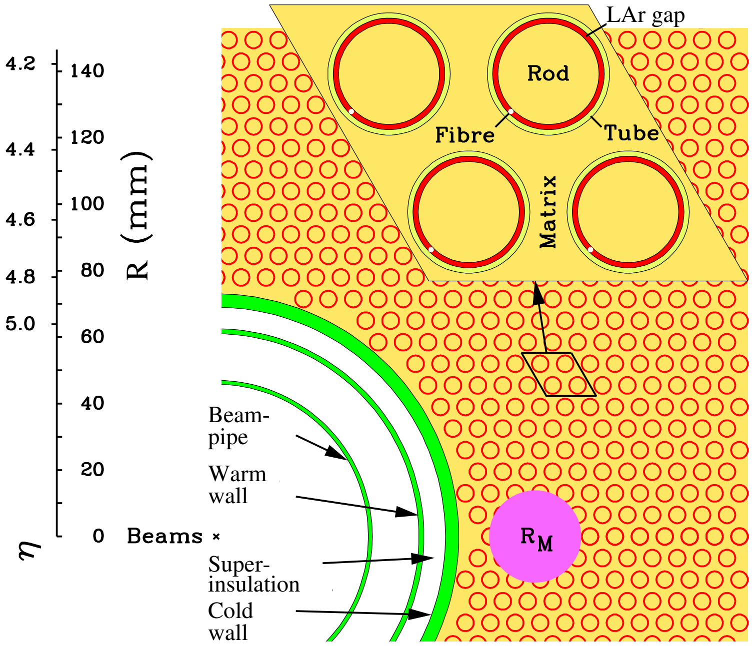


Figure 2.11: Schematic diagram showing the electrode structure of FCal1 with the matrix of copper plates, tubes and rods. Taken from Ref. [64].

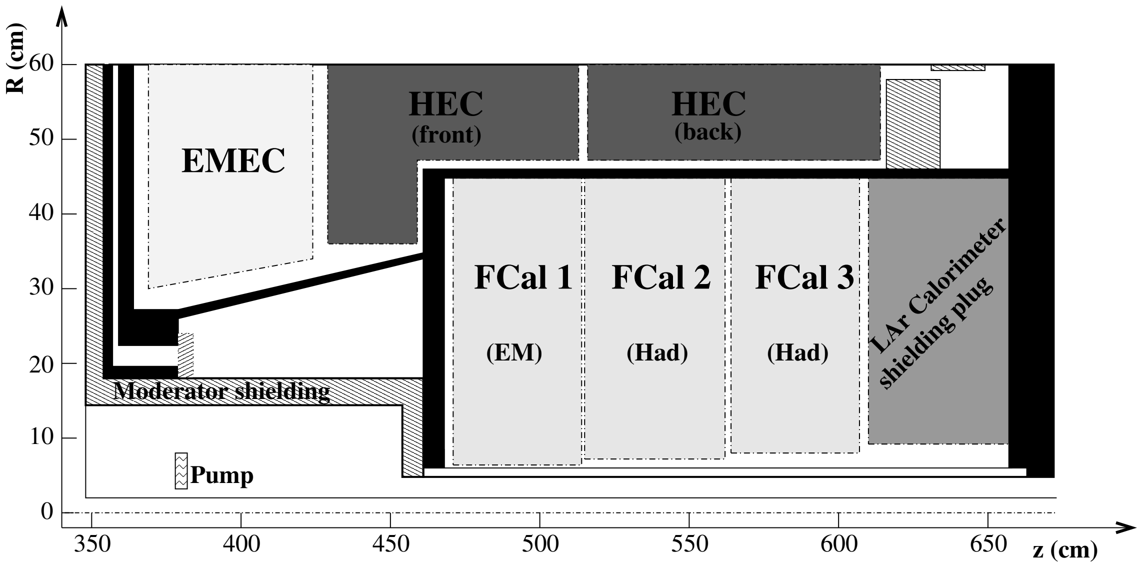


Figure 2.12: Schematic diagram showing the positioning of the three FCal modules. Taken from Ref. [64].

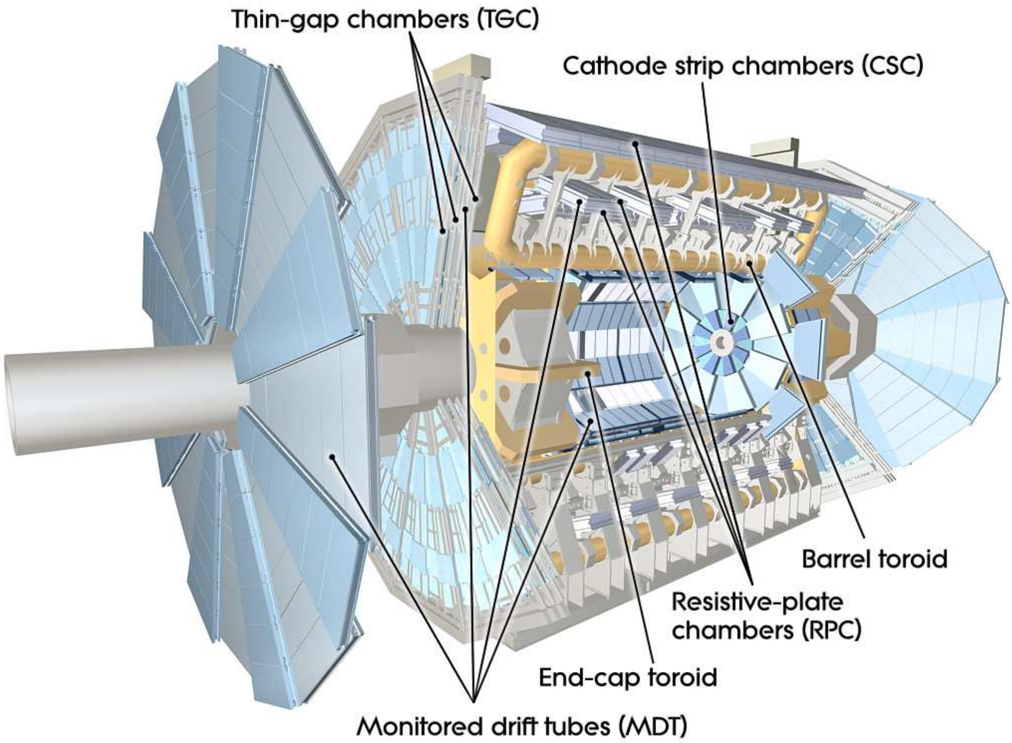


Figure 2.13: Cut-away view of the ATLAS muon system including toroid magnets. Taken from Ref. [64].

between the eight coils of a superconducting toroid magnet in the barrel region and in front of and behind the end-cap toroid magnets. It is this toroidal magnetic field (0.5 T in the barrel, 1 T in the end-caps) which enables the muon spectrometer to measure the momentum of charged particles. Figure 2.13 shows the layout of the muon spectrometer sub-detectors and the toroid magnets.

In the barrel region muon chambers are arranged in three cylindrical layers at 5 m, 7.5 m and 10 m from the beam axis. In the end-caps the chambers take the form of wheels perpendicular to the beam-line located at distances of 7.4 m, 10.8 m, 14 m, and 21.5 m from the interaction point.

Precision measurements are performed by the Monitored Drift Tube (MDT) chambers which cover the range $|\eta| < 2.7$. These chambers consist of three to eight layers of drift tubes. The exception to this is the innermost end-cap layer covering $2.0 < |\eta| < 2.7$ which uses Cathode Strip Chambers (CSCs) due to their higher rate capabilities. Each CSC is a multiwire proportional chamber with orthogonally

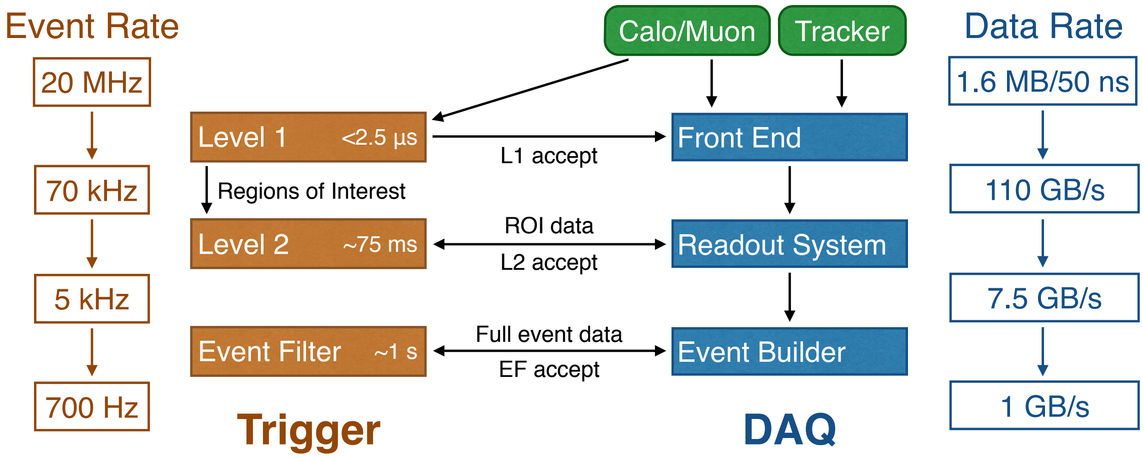


Figure 2.14: Schematic diagram of the ATLAS trigger and data acquisition system. Typical event and data rates during the 2012 run are included. Black arrows indicate the direction of the flow of information. Taken from Ref. [68].

segmented cathode planes.

The muon system is also capable of fast triggering (see Section 2.2.4) on muon tracks [90] requiring chambers capable of delivering track information with a latency of a few tens of nanoseconds. In the barrel region ($|\eta| < 1.05$) Resistive Plate Chambers (RPCs) perform this requirement whilst in the end-cap ($1.05 < |\eta| < 2.4$) Thin Gap Chambers (TGCs) are used. An RPC is a gaseous parallel electrode-plate detector operated in avalanche mode whilst a TGC is a multi-wire proportional chamber with a wire-to-cathode distance (1.4 mm) smaller than the wire-to-wire distance (1.8 mm) and operated in a quasi-saturated mode.

2.2.4 Trigger

The ATLAS data taking system is not technically capable of writing out data at the LHC collision rate of 20 MHz. To reduce the rate of events the ATLAS trigger system is employed [91]. This system is designed to quickly discard uninteresting events and save those events which may contain physics of interest.

Figure 2.14 shows a schematic diagram of the trigger and data acquisition (DAQ) system. The trigger system is split into three levels which reduce the event rate in steps.

The first level, Level 1 (L1) [92] reduces the event rate from 20 MHz to 70 kHz using coarse information from the calorimeter and the RPC and TGC parts of the muon system. The L1 system is hardware-based and must make a decision within 2.5 μs based on the presence of high p_{T} objects, high missing transverse energy or high total activity. Whilst the L1 decision is being made, information is stored in pipelines on the front end (FE) electronics and is only transmitted off detector if an L1 accept signal is received.

The Level 2 (L2) trigger is the first stage of the High Level Trigger (HLT) [93] and receives Regions of Interest (RoI) coordinates from the L1 system based on the objects which were identified at L1. At L2 these RoIs are used to request a subset of the full detector information ($\sim 2\%$) so that these regions can be fully reconstructed to inform the L2 decision. The L2 system reduces the event rate by a factor of ~ 14 to 5 kHz.

Finally the second stage of the HLT, the Event Filter (EF), reconstructs the full event and makes a final decision on whether an event is to be kept or not. This final decision is made in about 1s and reduces the event rate to a more manageable 700 Hz which can be successfully written to tape.

Occasionally it is advantageous to use low p_{T} triggers which would usually result in an event acceptance rate which is too high to sustain. For such triggers, a prescale can be applied which results in only one event being saved for every N events which pass such a trigger. These prescales are carefully set and varied during running to ensure that enough data are taken to fulfil the requirements of a broad range of analyses whilst not saturating the DAQ systems.

Data period	Luminosity delivered (pb^{-1})	Maximum luminosity (10^{30} cm^{-2} s^{-1})
D	186	659
E	52.7	832
F	160	1100
G	572	1263
H	287	1264
I	416	1887
J	240	1995
K	685	2328
L	1625	3252
M	1184	3848

Table 2.2: The stable luminosity delivered by the LHC as a function of the 2011 data taking period. Also shown is the maximum instantaneous luminosity for each data period.

2.2.5 Data Periods

Over the course of data-taking, the condition and running of the detector changes. For example, during part of 2011 a region of the liquid argon detector was non-functional due to a failure in the readout electronics. In addition, as the instantaneous luminosity of the LHC varied, the trigger menu was varied to prevent unmanageable increases in the dataflow rate.

In order to track such changes in an analysis, the 2011 dataset has been split into data periods designated alphabetically. In this thesis, the data taken in the periods from D to M are used. Table 2.2 shows the stable luminosity delivered by the LHC in each data period and the maximum instantaneous luminosity.

2.2.6 Coordinate System and Variables

ATLAS uses a right-handed coordinate system with it's origin at the interaction point. In this system, x points to the centre of the LHC ring, y points upwards whilst z points along the beam line. Due to the spherical nature of interactions, spherical (r,θ,ϕ) coordinates are often used.

Often the polar angle, θ, is replaced by the pseudorapidity, η, given by

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{2.5}$$

The pseudorapidity is attractive to high energy physicists because pseudorapidity differences are Lorentz invariant whilst polar angle differences are not. In addition, particle production at a collider is approximately isotropic in $\eta - \phi$ space.

A related angular variable is that of rapidity, y, given by

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \tag{2.6}$$

where E is the energy of a particle and p_z the momentum of the particle along the beam axis. In the relativistic limit ($E \gg m$) the pseudorapidity and rapidity of a particle converge.

Other important variables used in this thesis are the transverse momentum of a particle, p_{T}, the impact parameter in the $x-y$ plane between a particle and a point in the detector, d_0, and the angular separation in $\eta - \phi$ space between two objects, ΔR, given by

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{2.7}$$

Chapter 3

Object Reconstruction and Selection

The first stage in an analysis is to reconstruct the raw data from the detector. This process is carried out by reconstruction algorithms which take the raw data and convert them into physics objects. W and Z bosons are reconstructed from their decay products via the decay channels $W \to e\nu$, $W \to \mu\nu$, $Z \to ee$ and $Z \to \mu\mu$. As such, algorithms must be used to reconstruct electrons, muons, jets and missing transverse energy (due to the neutrinos). In addition, tracks in the inner detector must be reconstructed as input to the physics object reconstruction alogorithms.

In this Chapter, the reconstruction of tracks (Section 3.1), electrons (Section 3.2), muons (Section 3.3), jets (Section 3.4) and missing transverse energy (Section 3.5) will be discussed. In addition, the identification requirements applied to the objects used in this analysis will be discussed and the event selection outlined in Section 3.6. Note that reconstructed tau and photon physics objects are included in the calculation of missing transverse energy however are not used independently in the analysis. As such their reconstruction will not be discussed here but details can be found in Refs. [94] and [95] respectively.

3.1 Tracks

Charged particle track reconstruction is described in Ref. [96]. The first step in the reconstruction is the creation of three dimensional silicon hits called spacepoints. This is trivial in the pixel detector given the radial position of the layers and the two dimensional segmentation of the pixels. In the SCT, the creation of spacepoints is more detailed as a single SCT hit can only provide positioning orthogonal to the strip direction (in ϕ) and at a radius defined by the SCT layer. Thus, hits on either side of a module must be combined to define a z position (due to the small stereo angle between them) and thereby a full three-dimensional spacepoint.

Spacepoints in the pixel detector and inner layer of the SCT are used to form seed tracks using a Kalman filter [97]. These seeds are then extended through the SCT. At this stage the tracks are analysed for various quality requirements such as goodness-of-fit and the number of silicon layers traversed with no associated hit. If a track passes these quality requirements, the track is further extended into the TRT and the fit and quality requirements repeated with the full set of spacepoints from all three parts of the ID.

This inside-out algorithm is followed by a second outside-in algorithm which begins with unused track segments in the TRT and extends them into the SCT and pixel to find any associated silicon hits for track fitting. This second outside-in step of the track finding algorithm improves the track finding efficiency for tracks arising from secondary vertices or long-lived particles.

Tracks with $p_{\mathrm{T}} > 400$ MeV are extended to the centre of the detector where they can be used for primary vertex finding [98, 99] (based on χ^2 minimisation) which is followed by secondary vertex finding. In this way, the tracks can be associated to particular vertices and thereafter used to remove objects arising from additional pile-up vertices. The primary vertex in an event is defined to be that with the largest p_{T}

Data Period(s)	L1 Trigger	L2 Trigger	HLT Trigger
D-J	EM14	L2_e20_medium	EF_e20_medium
K	EM16	L2_e22_medium	EF_e22_medium
L-M	EM16VH	L2_e22vh_medium1	EF_e22vh_medium1

Table 3.1: Electron triggers used in this analysis. The number in the names refers to the E_{T} threshold in GeV, medium refers to a trigger identification requirement close to the offline medium requirement (see Section 3.2.3), medium1 refers to a trigger identification requirement using reoptimised cuts (similar to medium++) and "vh" indicates the use of a variable threshold requirement and a hadronic leakage requirement (as discussed in the text). The alphabetical data periods are explained in Section 2.2.5.

associated to it when summing over all associated tracks. In this analysis, an event is required to have at least one good vertex with at least three tracks associated to it.

3.2 Electrons

3.2.1 Trigger

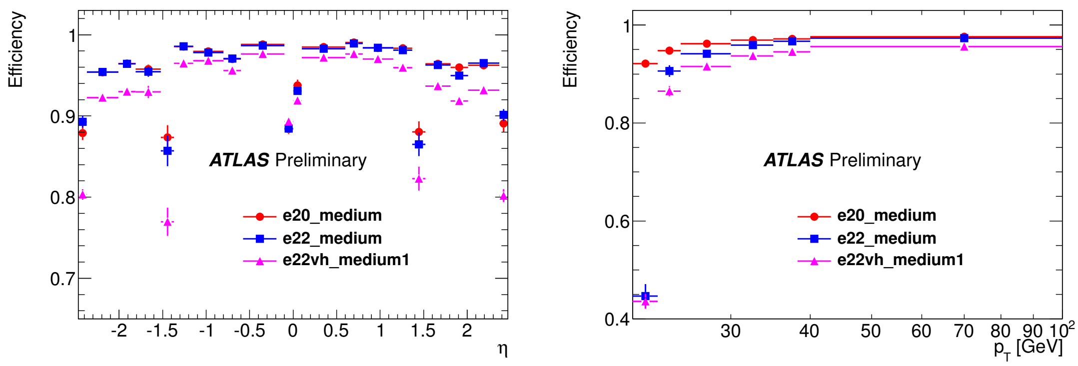
As discussed in Section 2.2.4 events are only read out from the detector and saved for analysis if they pass one of the many triggers. For the $W \to e\nu$ and $Z \to ee$ events used in this thesis, events are triggered by the presence of high-p_{T} electrons. The lowest p_{T} non-prescaled triggers are used. Over the 2011 run the instantaneous luminosity increased and as such, different triggers were used over the data-taking periods. Table 3.1 shows the triggers used over the data-taking period and the L1 and L2 trigger requirements used.

The L1 electron triggers utilise coarse granularity calorimeter information based on regions (towers) of $\Delta\eta \times \Delta\phi \sim 0.1 \times 0.1$. EM clusters are formed by identifying local maxima using a 4×4 group of trigger towers as a sliding window. The window's core 2×2 region must contain one pair of neighbouring towers with a combined energy above threshold. The transverse energy of towers at L1 is calculated with a precision of 1 GeV.

The L2 electron triggers only use the second layer of the ECal due to timing constraints. The cell in the RoI from L1 with the largest deposited energy is found and used to seed the rest of the algorithm. The final cluster position is found as the energy weighted average cell position on a 3×7 grid centred on this seed. At this point a cluster reconstruction similar to that used offline (see Section 3.2.2) and tracking information based on fast tracking algorithms run in the RoI are used to produce the final L2 decision

Finally, the Event Filter (EF) electron triggers are able to perform electron reconstruction in the most similar way to the offline algorithms as possible. Calorimeter quantities are reconstructed using offline algorithms run on cell information from a region expanded slightly from the original RoI. An adapted version of the offline track reconstruction can also be employed at EF level.

Additional requirements were added to the triggers over the course of 2011 data-taking. In particular in the L-M periods, when the instantaneous luminosity was at its highest, "vh" requirements were added to the trigger. The v in this refers to a variable threshold trigger whereby the actual threshold is varied dependent on η so that larger thresholds are used in regions where cluster energy response agrees best with offline cluster energy, thereby reducing event rates whilst maintaining high selection efficiency. Without this variable threshold the reduction in event rates, necessitated by the increased instantaneous luminosity, would have had to be acheived by increasing the trigger thresholds globally and thereby reducing trigger efficiencies. The h refers to a hadronic leakage requirement whereby a veto was included on hadronic energy of more than 1 GeV behind the EM-cluster. Both these requirements were included to maintain the L1 rate at acceptable rates without having to significantly increase the threshold and reduce acceptance.



(a) ϵ_{trig} vs. η (b) ϵ_{trig} vs. p_{T}

Figure 3.1: The electron trigger efficiency, ϵ_{trig}, as a function of (a) η and (b) p_{T} for offline medium++ electrons. The different colours show the different electron triggers used in this analysis. Taken from Ref. [100].

The efficiency of the triggers used in this analysis has been studied centrally using tag-and-probe methods [100] whereby known resonances such at $Z \to ee$ are used to select unbiased electrons (probes) by using strict selection criteria on the second electron produced by the decay of resonances (tags). Figure 3.1 shows the trigger selection efficiency for the triggers used in this analysis as a function of offline electron p_{T} and η. Mis-modelling of this efficiency in Monte Carlo is corrected by the application of centrally produced p_{T} and η dependent scale factors (SF_{trig}).

3.2.2 Offline Reconstruction

Offline electrons are reconstructed using track information from the ID and cluster information from the ECal [101]. The reconstruction of electrons with $|\eta| < 2.47$ is seeded using clusters in the ECal. A sliding window algorithm is used with a window of 3×5 in units of 0.0025×0.025 in $\eta \times \phi$ space. A seed cluster is located by the sliding window if the total transverse energy in the window is greater than 2.5 GeV and is a local maximum.

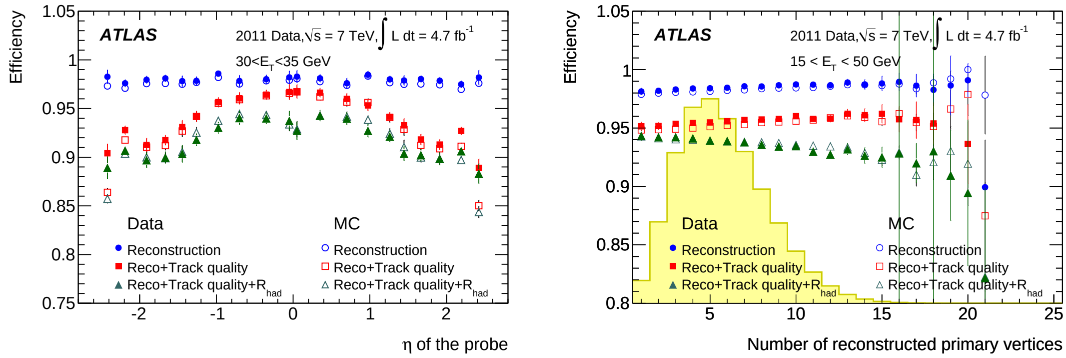
This seed cluster becomes a reconstructed electron if it can be loosely matched to a track in the ID. The track impact point (the point the extrapolated track intersects the middle layer of the ECal) is required to be within $\Delta\eta < 0.05$ of the cluster and within $\Delta\phi < 0.05$. A looser requirement of $\Delta\phi < 0.1$ is applied on the side where the track bends in the magnetic field to account for bremsstrahlung. In the case of multiple tracks being associated to a single cluster the track with the smallest $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ is chosen.

Once a cluster-track match is performed the cluster is rebuilt using 3×7 (5×5) towers of cells in the barrel (endcap) regions. The cluster energy is then determined using a combination of the estimated energy deposit before the ECal (based on the presampler), the measured energy deposited in the cluster and the energy lost laterally and longitudinally (based on shower shape). Such energy deposits are estimated by comparison to Monte Carlo simulation.

The final reconstructed electron four momentum is computed using both cluster and track information, taking the energy from the cluster energy and the direction (η and ϕ) defined by the parameters of the associated track. The parameters used in the reconstruction procedure have been optimised using Monte Carlo simulation. In order to validate these using data, tag-and-probe methods [102] similar to those discussed above in Section 3.2.1 are used to derive the electron reconstruction efficiency shown in Figure 3.2. The differences observed between data and Monte Carlo are corrected by the application of p_{T} and η dependent scale factors provided centrally (SF_{reco}).

3.2.3 Offline Identification

Following the electron reconstruction discussed above many electrons are reconstructed from background electrons (from photon conversions or in-flight decays of unstable particles) and fake electrons (from jets faking electrons). In order to remove such electrons and identify only real, prompt electrons a cut-based selection is used based on calorimeter and tracking information. Three levels of electron



(a) ϵ_{reco} vs. η (b) ϵ_{reco} vs. N_{PV}

Figure 3.2: The electron reconstruction efficiency, ϵ_{reco}, in data and Monte Carlo as a function of (a) η and (b) N_{PV}. The different colours relate to different event selections of all reconstructed electrons, electron candidates with additional track quality requirements (matching $J/\psi \to ee$ tag-and-probe requirements) and candidates with additional track quality and hadronic leakage requirements (matching $Z \to ee$ and $W \to e\nu$ tag-and-probe requirements). Taken from Ref. [101].

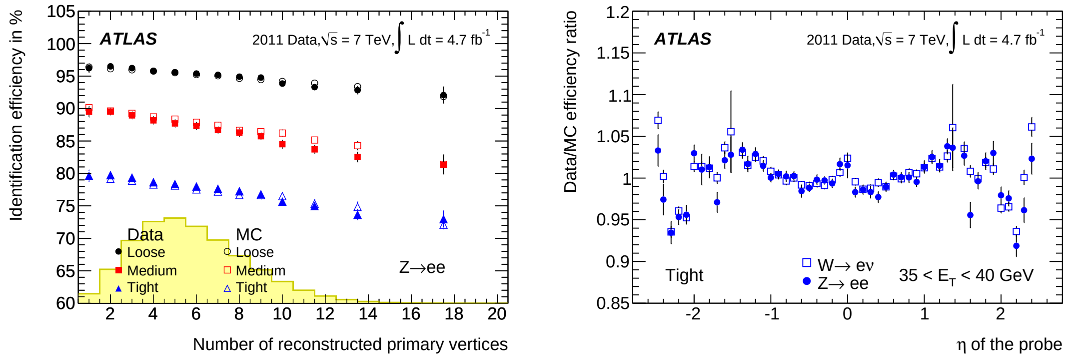
identification are defined with various degrees of background rejection and real electron identification efficiency. These are loose++, medium++ and tight++[1] which are updated versions of the loose, medium and tight identification criteria used in 2010 [102]. The updated ++ requirements are outlined in Ref. [101].

The loose criterion is based on hadronic leakage (the amount of energy deposited in the first layer of the HCal behind the electron cluster) and the lateral shower width[2] of the cluster in the middle layer of the calorimeter [103]. The medium criterion builds on loose by applying cuts based on shower width[3] in the ECal strip layer and the difference between the highest and second highest energy deposits in the cluster. Medium also includes additional track quality requirements and a tighter track-cluster matching than that used in the electron reconstruction. Finally, tight inherits all the medium requirements and in addition requires particle ID in the

[1] The ++ terminology will be used in this thesis to differentiate the updated 2011 requirements from those used in 2010.

[2] lateral shower width is given by $\sqrt{(\sum E_i \eta_i^2) / (\sum E_i) - ((\sum E_i \eta_i) / (\sum E_i))^2}$ where E_i and η_i are the energy and pseudorapidity of cell i and the sum over i covers a window of 3×5 cells.

[3] shower width is given by $\sqrt{\left(\sum E_i \left(i - i_{\max}\right)^2\right) \left(\sum E_i\right)}$ where E_i us the energy of strip i, i runs over strips in a window of $\Delta\eta \times \Delta\phi \sim 0.0625 \times 0.2$ and $i_{\max}$ is the index of the highest-energy strip.



(a) ϵ_{ID} vs. N_{PV} (b) Scale factor vs. η

Figure 3.3: The loose++, medium++ and tight++ identification efficiency, ϵ_{ID}, measured in data and Monte Carlo using tag-and-probe in $Z \to ee$ events and the efficiency scale factors derived for tight++ using $Z \to ee$ and $W \to e\nu$ tag-and-probe. Taken from Ref. [101].

TRT and requires a b-layer (inner layer of the pixel detector) hit. The track-cluster matching requirements are further tightened and electron candidates matched to reconstructed photon conversions are vetoed.

The updated loose++ identification updates loose with the addition of shower-shape variables in the first ECal layer and some track quality and track-cluster matching requirements which improve the hadronic background rejection by ~ 5. Such a tightening of requirements is also inherited by medium++ and tight++ identification criteria.

Similarly to the reconstruction efficiencies explained above, identification efficiencies can be studied using tag-and-probe methods. Figure 3.3 shows the identification efficiency as a function of N_{PV} for loose++, medium++ and tight++ identification requirements. Also shown in Figure 3.3 are the p_{T} and η dependent scale factors which are provided centrally and applied to correct for the measured differences between data and Monte Carlo.

3.2.4 Offline Calibration

The baseline calibration has been derived using testbeam measurements carried out on a single module of the LAr calorimeter [104] whereby known energy electrons are fired into the module and the measured response used to derive calibration factors. This initial calibration is then corrected using Monte Carlo based simulation studies. At the cluster level, corrections are applied to correct for energy loss due to absorption by dead material in the calorimeter and leakage whereby some of the electron energy is deposited outside the reconstructed cluster. Finally, additional fine corrections are applied depending on the η and ϕ position of the electron to compensate for any additional energy modulation.

Following these corrections a final residual correction is derived and validated in-situ using $Z \rightarrow ee$ and $J/\psi \rightarrow ee$ [102]. In these studies the line-shape of the resonances is compared between Monte Carlo and data and a fit is performed allowing the parameters of the residual calibration to float.

3.2.5 Isolation

To further reduce background electron selection, requirements can be included on the electron candidates to ensure that they are isolated from other deposits in the detector. This can either be done using calorimeter information, track information or a combination of the two.

Isolation requirements in the calorimeter are based on the $E_{\mathrm{T}}^{\mathrm{coneXX}}$ variable [105]. This is calculated by summing all the energy deposits in the calorimeter in a cone of radius 0.XX around the electron in $\eta \times \phi$ space as shown in Figure 3.4. From the total, the calorimeter deposits in the electron cluster of 5×7 cells are subtracted and Monte Carlo is used to subtract the leakage of electron energy into the surrounding cells. Isolation requirements are then applied based on the ratio of $E_{\mathrm{T}}^{\mathrm{coneXX}}$ to the

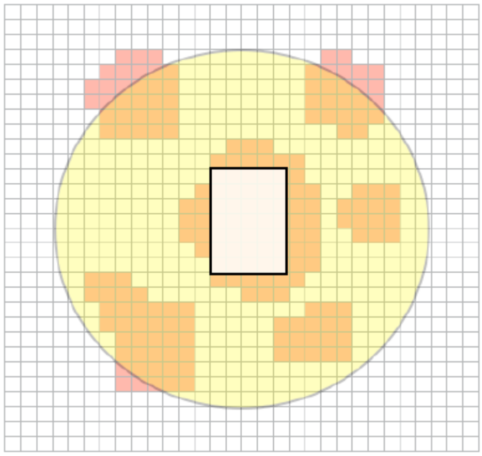


Figure 3.4: Sketch illustrating the calorimeter isolation requirements. The central white rectangle shows the 5×7 window which defines the electron cluster. The yellow circle shows a cone of $\Delta R = 0.4$. All cells outside the white rectangle and within the yellow circle are included in the calculation of $E_{\mathrm{T}}^{\mathrm{cone40}}$. Taken from Ref. [105].

electron E_{T}.

In the ID a similar definition of the isolation is used but using tracks instead of calorimeter clusters. The equivalent of $E_{\mathrm{T}}^{\mathrm{coneXX}}$ is denoted as $p_{\mathrm{T}}^{\mathrm{coneXX}}$ in this thesis. This is the sum p_{T} of all tracks in a cone of radius 0.XX except for the track associated to the electron.

In this analysis, a complex calorimeter isolation requirement is applied (denoted `iso98Etcone20`) in which an η dependent cut is applied to $E_{\mathrm{T}}^{\mathrm{cone20}}/E_{\mathrm{T}}$. This cut is chosen such that the isolation selection efficiency is 98% for signal electrons. In addition, a similar track based requirement is used (`iso97Ptcone40`) which uses $p_{\mathrm{T}}^{\mathrm{cone40}}/p_{\mathrm{T}}$ to achieve a selection efficiency of 97%.

3.2.6 Selection

Following the trigger requirement shown in Table 3.1, electrons are required to pass the offline tight++ identification requirement (Section 3.2.3) and pass the `iso97Ptcone40` and `iso98Etcone20` requirements (Section 3.2.5). The electrons are calibrated using centrally produced calibration factors (Section 3.2.4) and in addition, the p_{T} of electrons in Monte Carlo is smeared to improve the agreement of energy resolution between Monte Carlo and data.

Electrons are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.47$. Electrons falling in the crack region ($1.37 < |\eta| < 1.52$) are not included. Finally, quality cuts are applied which ensure that any electrons affected by noise bursts or failures in the LAr calorimeter are removed and any cells included in the electron clusters are behaving well.

Following this selection, efficiency scale factors are applied to correct mis-modelling of selection efficiencies in the electron selection. The total applied efficiency scale factor is given by

$$SF_{TOT}(p_{\mathrm{T}}, \eta) = SF_{reco}(p_{\mathrm{T}}, \eta) \times SF_{ID}(p_{\mathrm{T}}, \eta) \times SF_{trig}(p_{\mathrm{T}}, \eta). \tag{3.1}$$

In addition, the application of a scale factor to correct for the mis-modelling of the isolation efficiency has been considered. Such a discrepancy has previously been measured in ATLAS [106] and the resulting scale factors are found to be small (corrections are $< 1\%$ with a $\sim 1\%$ uncertainty). As such, no isolation scale factors have been applied.

3.3 Muons

3.3.1 Trigger

In the muon channel, events are triggered by the presence of high-p_{T} muons. Similar to in the electron channel (Section 3.2.1), over the course of 2011 data-taking the increasing instantaneous luminosity necessitated the use of tighter trigger requirements in the latter stages of the year. Table 3.2 summarises the triggers used in the muon channel and the L1, L2 and EF parts of them.

Data Period(s)	L1 Trigger	L2 Trigger	HLT Trigger
D-I	MU10	L2_mu18_MG	EF_mu18_MG
	MU10	L2_mu18	EF_mu18
L-M	MU11	L2_mu18_MG_medium	EF_mu18_MG_medium
	MU11	L2_mu18_medium	EF_mu18_medium

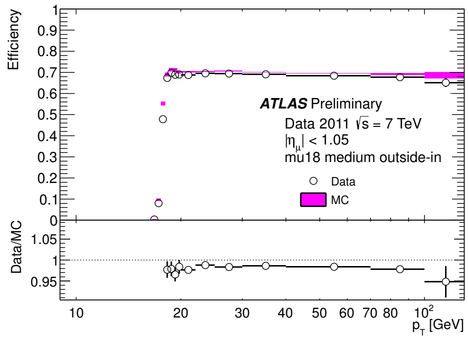
Table 3.2: Muon triggers used in this analysis. The number in the names refers to the p_{T} threshold in GeV, MG refers to the use of an inside-out algorithm rather than an outside-in used for muon reconstruction and medium refers to the triggers being seeded by MU11 rather than MU10 at L1. The two triggers in each period are combined by requiring either trigger to fire.

The L1 trigger decision uses information from the RPC and TGC parts of the detector. The signals from these subdetectors are searched for hits using pre-calculated paths (roads) at defined p_{T} thresholds [92]. The MU10 trigger requires two (three) hits in the barrel (endcap) regions. As instantaneous luminosity increased, this requirement was increased to three hits in both regions for the MU11 trigger.

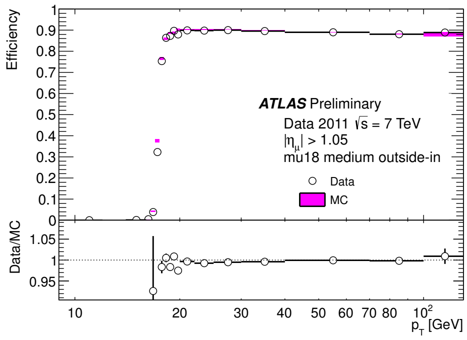
At L2, information in the RoI defined by the L1 candidate is read out from the MDTs. This information is used to refine the track fit using fast fitting algorithms and look up tables. The L2 combined algorithm also uses tracks in the ID to refine the momentum resolution and reject muons produced by decays in-flight. The EF level trigger refines this further using information from the whole detector (ID, RPCs, TGCs and MDTs).

Two muon reconstruction algorithms are used in the trigger system. One, labelled "MG", uses an inside-out approach whereby in the HLT ID tracks are used as seeds to the muon fitting and are extrapolated out to the muon system. The alternative uses an outside-in approach whereby muon segments are used as seeds and extended into the ID to reconstruct the full muon track. An OR of both these algorithms is used in this analysis.

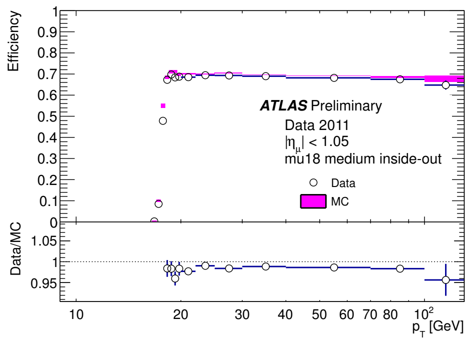
Studies have been performed in an updated version of Ref. [106] studying the efficiency of these triggers in both data and Monte Carlo. These provide a more precise version of the results shown in Figure 3.5. Scale factors, SF_{trig}, are applied



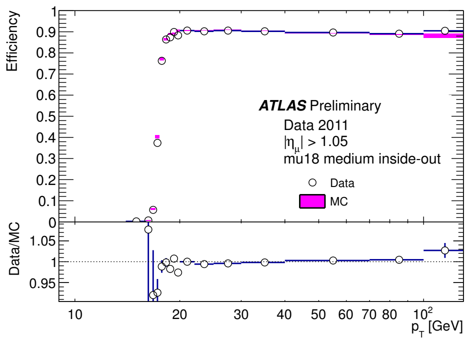
(a) mu18_medium (barrel)



(b) mu18_medium (endcap)



(c) mu18_MG_medium (barrel)



(d) mu18_MG_medium (endcap)

Figure 3.5: Muon trigger efficiency in 2011 data-taking for the mu18_medium and mu18_MG_medium triggers as a function of muon p_{T} derived using tag-and-probe in $Z \to \mu\mu$ events. The similar performance between the two algorithms in terms of efficiency and agreement between data and Monte Carlo is clearly visible. Taken from Ref. [107].

to correct the Monte Carlo modelling of the trigger efficiency.

3.3.2 Offline Reconstruction

Once an event has been triggered by one of the muon triggers discussed in Section 3.3.1, the event is reconstructed offline. To reconstruct muons in this analysis the Combined (CB) algorithm [108] is employed using the "Chain 1" combination procedure. Such muons are referred to as STACO muons.

In the STACO algorithm, two separate tracks are reconstructed - that of the muon in the ID and that of the muon in the muon spectrometer (MS). Such tracks

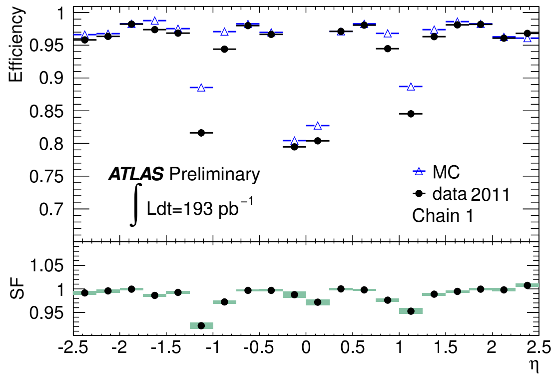


Figure 3.6: Muon reconstruction efficiency as a function of η for STACO combined muons. The plot shows preliminary results from early 2011 data, although in the final analysis efficiencies derived from the entire dataset are used. Taken from Ref. [110].

are compared and pairs of tracks which are deemed similar are combined into a single muon candidate. The "Chain 1" combination procedure refers to how the ID and MS track information is combined. This is done using a statistical combination of the track parameters in the ID and MS. At low p_{T} ($\leq$ 80 GeV) the ID track dominates this statistical combination and at high p_{T} ($\geq$ 100 GeV) the MS measurement dominates. In the overlap region the two measurements have approximately equal weight [109].

Using tag-and-probe in $Z \to \mu\mu$ the reconstruction efficiency for STACO muons has been studied in both data and Monte Carlo, an example of which is shown in Figure 3.6. Centrally produced scale factors, SF_{reco} are applied to correct the mis-modelling of the muon reconstruction efficiency [108].

3.3.3 Offline Identification

In order to reduce the number of background muons from in-flight decays of hadrons, further identification requirements are included in the analysis. At least one hit is required in the b-layer, at least two hits in the pixel detector and at least 6 hits in the SCT. At most, two active pixel or SCT sensors are allowed to have been traversed with no hits. In addition, at least 9 TRT hits are required within the region of full

TRT acceptance.

These identification requirements are included on the muons used for the reconstruction efficiency studies discussed in Section 3.3.2 and so, no identification specific efficiencies or scale factors are derived.

3.3.4 Offline Calibration

In a similar way as described in Section 3.2.4, the muon calibration in Monte Carlo is corrected to agree with data. The calibration factors are derived using studies of $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ line shapes [108]. Separate scale and resolution corrections are applied to the Monte Carlo simulation.

3.3.5 Selection

Following the trigger requirement shown in Table 3.2, muons are selected based on the identification criteria discussed in Section 3.3.3 and calibrated as discussed in Section 3.3.4. In addition, muons are required to be isolated from additional tracks by requiring $p_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.1$ and a transverse impact parameter[4] with respect to the primary vertex of $|d_0/\sigma(d_0)| < 3$ where $\sigma(d_0)$ is the uncertainty in d_0. The mis-modelling of such quantities has been found to result in very small isolation scale factors [106] and as such, no such scale factors are applied.

Muons are required to have $p_{\mathrm{T}} > 25$ GeV and fall in the region $|\eta| < 2.4$. Following these selections, p_{T} and η dependent scale factors are applied to the Monte Carlo to correct for mis-modelling of selection efficiencies. The total applied efficiency scale

[4] d_0 is defined as the minimum distance between the extrapolated path of the muon and the primary vertex in the plane perpendicular to the beam axis, $\sqrt{x_0^2 - y_0^2}$, where x_0 and y_0 are the coordinates of the point of closest approach.

factor is given by

$$SF_{TOT}(p_{\mathrm{T}}, \eta) = SF_{reco}(p_{\mathrm{T}}, \eta) \times SF_{trig}(p_{\mathrm{T}}, \eta). \tag{3.2}$$

3.4 Jets

Emission of strongly interacting particles from collisions results in showers of hadronic particles. In order to group such showers into single objects, jet algorithms have been developed which group the showers into a single object, a jet. In order to allow comparisons between jets reconstructed from partons in theoretical calculations, truth particles following Monte Carlo showering and objects measured in the detector, such jet algorithms must be infrared and collinear emission safe [111]. Infrared and collinear (IRC) safety is the requirement that if an event is modified by additional soft radiation or a collinear splitting, the final jets are unchanged. IRC safety is attractive in particle physics as it:

- reduces the dependence of jets on the non-perturbative effects which cause infrared or collinear emission.
- removes the divergent tree-level matrix elements associated to infrared and collinear emission in fixed-order perturbative calculations.

In this analysis, the anti-k_t algorithm [112] is employed with a radius parameter of $R = 0.4$. The jets are built using the FastJet software package [113]. In this algorithm, d_{ij} and d_{iB} are introduced as the distance between entities (particles and pseudojets) and the distance between entities and the "beam" respectively. These are defined as:

$$d_{ij} = \min\left(p_{\mathrm{T},i}^{-2}, p_{\mathrm{T},j}^{-2}\right) \frac{\Delta R_{ij}^2}{R^2}, \tag{3.3}$$

$$d_{iB} = p_{\mathrm{T},i}^{-2}, \tag{3.4}$$

where ΔR_{ij} is the distance between entities i and j in $\eta - \phi$ space and R is the radius parameter (in this case 0.4). The clustering proceeds by identifiying the smallest of all d_{ij} and d_{iB} variables and, in the case of a d_{ij} being the smallest, combines entities i and j into one entity. If instead a d_{iB} is the smallest, entity i is called a jet and removed from the list.

For a given ΔR, such an algorithm is dominated by the $\min\left(p_{\mathrm{T},i}^{-2}, p_{\mathrm{T},j}^{-2}\right)$ term. As a result, hard entities dominate the recombination and collect up the soft entities. Generally this algorithm leads to conical shaped jets except in the case where two hard entities are close together ($R < \Delta R < 2R$). In this case, if $p_{\mathrm{T},i} \sim p_{\mathrm{T},j}$ then the jet boundary is flat whilst if there is a large difference between the p_{T} of the entities the higher p_{T} entity will dominate and remain conical whilst distorting the other entity.

3.4.1 Topo-clustering

In order to suppress the effects of calorimeter noise, before jet clustering the energy deposits are grouped into topological clusters [114] (topo-clusters) of calorimeter cells taking advantage of the high granularity ATLAS calorimeter systems to follow the shower development. The seed of topological clustering is any cell with a signal over noise ratio (S/N) above 4. The noise includes the contribution from electronic noise (estimated from events triggered in random bunch crossings in 2010 data) and pile-up noise (estimated from Monte Carlo simulation). Next, any cells neighbouring the seed cell with $S/N > 2$ are added iteratively. Finally, all calorimeter cells neighbouring the topo-cluster are added ($S/N > 0$). This algorithm and the choice of the 4/2/0 scheme has been optimised to most effectively suppress calorimeter noise [115].

Following a final part of the algorithm which splits topo-clusters dependent on local maxima to separate close-by calorimeter deposits, topo-clusters are defined to

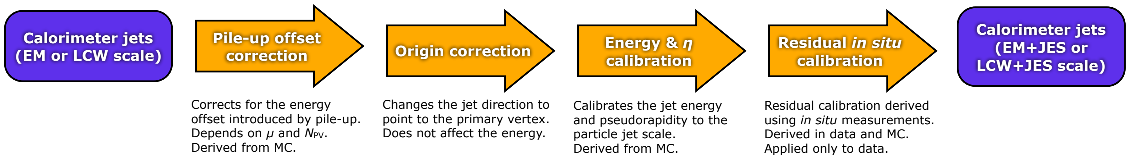


Figure 3.7: Overview of the ATLAS jet calibration scheme employed in 2011 data-taking. The electromagnetic (EM) scale calibration is used in this analysis. Taken from Ref. [116].

have an energy equal to the sum of the energies of all included clusters. They are defined to be massless and have a direction given by an energy weighted average of the η and ϕ positions of the included clusters. At this point the topo-clusters are passed to the jet algorithm to perform jet finding.

3.4.2 Calibration

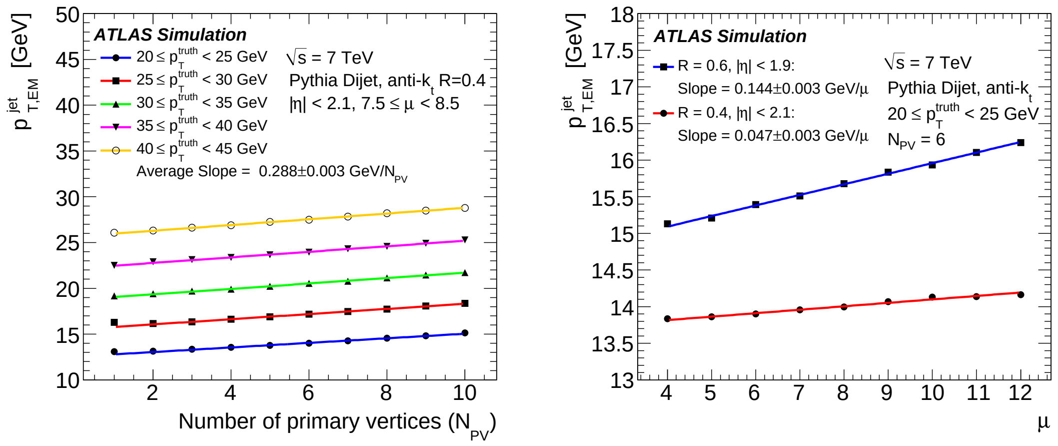
Following the jet reconstruction from topo-clusters, the calibration scheme shown schematically in Figure 3.7 is used to restore the jet energy scale to that of jets reconstructed from stable simulated particles in Monte Carlo [116].

The calibration scheme has 4 distinct stages:

1. **Pile-up offset correction**: the jets formed from topo-clusters are corrected for the effects of additional proton-proton collisions (pile-up). The correction is dependent on the number of primary vertices, N_{PV}, and the average number of interactions per crossing, μ (see Equation 2.4) and is applied as an additive correction:

$$p_{\mathrm{T}}^{\mathrm{corr.}} = p_{\mathrm{T}} - \alpha\left(\eta\right) \times \left(N_{PV} - N_{PV}^{\mathrm{ref.}}\right) - \beta\left(\eta\right) \times \left(\mu - \mu^{\mathrm{ref.}}\right). \tag{3.5}$$

   α and β are derived from the average slopes of linear fits to the N_{PV} and μ dependence of p_{T} in Monte Carlo (and validated in data), as demonstrated in Figure 3.8.



(a) p_{T} vs. N_{PV} (b) p_{T} vs. μ

Figure 3.8: The average p_{T} of anti-k_t, $R = 0.4$ jets as a function of N_{PV} and μ, the latter also including anti-k_t, $R = 0.6$ jets. The average slope of the N_{PV} (μ) dependence defines the α (β) terms in Equation 3.5. Taken from Ref. [116].

2. **Origin correction**: the jets are corrected to point back to the primary vertex in the event rather than the centre of the ATLAS detector.

3. **Monte Carlo based jet calibration**: the energy and pseudorapidity of the jets are next calibrated using a simple correction derived from the relation of the reconstructed and truth variables in Monte Carlo simulation. Figure 3.9 shows the average jet energy response before this correction, defined as the ratio of reconstructed and truth level energy. The energy calibration function is the inverse of this response.

4. **Residual in-situ corrections**: following the Monte Carlo based calibration, dijet events are used to perform an in-situ inter-calibration between jets in the forward region ($|\eta| > 0.8$) and jets in the central region ($|\eta| < 0.8$). After this, further in-situ methods using Z bosons recoiling from jets, photons recoiling from jets and multi-jet systems recoiling from high p_{T} jets are used to compare the jet response in Monte Carlo simulation and data in the central region, $|\eta| < 0.8$. These in-situ response methods are compared and combined to define a residual in-situ correction to the jet calibration which is applied only to data.

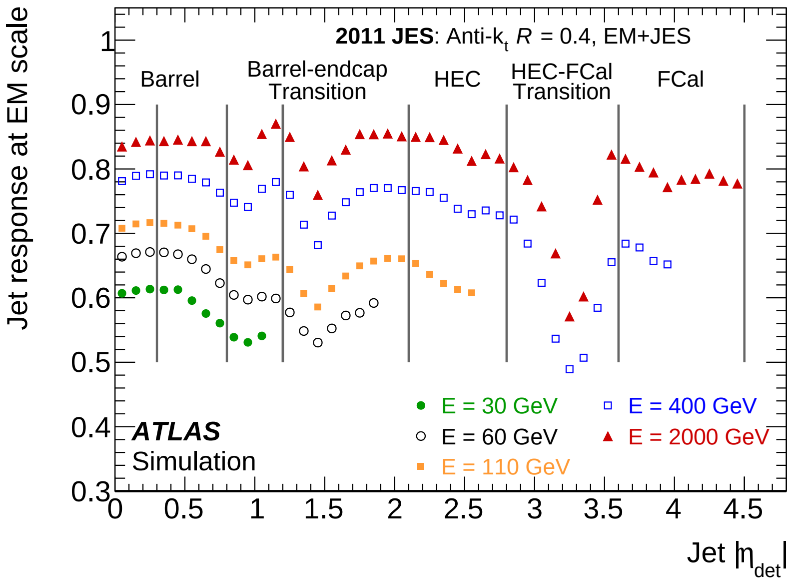


Figure 3.9: The average energy response of jets in Monte Carlo as a function of $|\eta|$ for various energy jets. The inverse of this response defines the jet calibration factors applied to calorimeter jets. Taken from Ref. [116].

3.4.3 Selection

Following the calibration of jets described in Section 3.4.2, jets are required to pass the "BadLooser" quality condition. This is a cleaning criteria which has been developed to remove jets arising from non-collision backgrounds or detector noise [116]. In addition, jets within the ID ($|\eta| < 2.4$) are checked to be consistent with arising from the primary vertex using the Jet Vertex Fraction (JVF). This is a p_{T} weighted measure of the fraction of tracks associated with a jet which are associated with the primary vertex,

$$\mathrm{JVF}(\mathrm{jet}_i, \mathrm{PV}_j) = \frac{\sum_k p_{\mathrm{T}}(\mathrm{track}_k^{\mathrm{jet}_i}, \mathrm{PV}_j)}{\sum_n \sum_l p_{\mathrm{T}}(\mathrm{track}_l^{\mathrm{jet}_i}, \mathrm{PV}_n)}. \tag{3.6}$$

JVF is demonstrated schematically in Figure 3.10. To reject pile-up jets a requirement of $|\mathrm{JVF}| > 0.75$ is applied in this analysis.

Following quality and JVF requirements being applied to the jets, any jets falling into the LAr Hole (a region of the ECal where readout electronics failed during 2011)

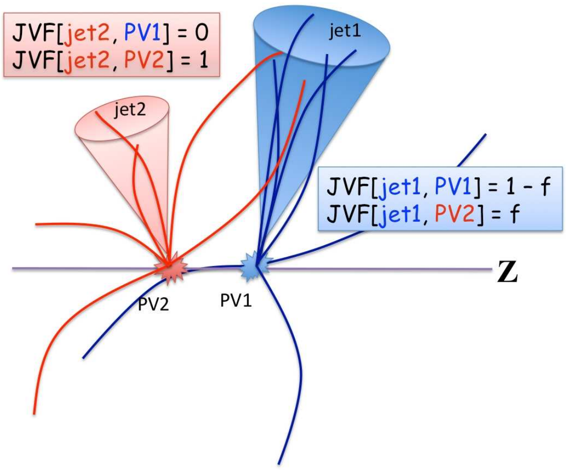


Figure 3.10: Schematic representation of the jet vertex fraction (JVF) method for separating pile-up jets from hard scatter jets. The blue jet is associated to the hard scatter by virtue of the large number of tracks associated to it coming from the primary vertex. The red jet is labelled as a pile-up jet due to the low fraction of associated tracks coming from the primary vertex. Taken from Ref. [3].

are found. If any jet above 20 GeV is found in this region ($-0.1 < \eta < 1.5$ and $-0.9 < \phi < -0.5$) the event is rejected as this effects the missing transverse energy measured in the event (see Section 3.5). In addition any event with a "BadLooser" identified jet above 20 GeV is also removed for the same reason.

Finally, jets passing all the quality criteria are required to have $|y| < 4.4$ and $p_{\mathrm{T}} > 30$ GeV.

3.5 Missing transverse energy

$W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ both have a neutrino in the final state. Due to the very weakly interacting nature of neutrinos, these particles will pass through the detector without depositing any energy and so cannot be directly reconstructed. Instead, their presence is inferred by invoking the principle of conservation of momentum in the plane perpendicular to the beam (as the initial momentum parallel to the beam is unknown at a hadron collider) by reconstructing the whole event and assigning any imbalance in the total transverse momentum as the transverse momentum of any neutrinos. At ATLAS, the missing transverse momentum is historically called missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, due to the primary input being calorimeter information.

In this thesis, the "RefFinal" reconstruction algorithm [117] is used to reconstruct the missing transverse energy. This algorithm uses the best calibration scheme available for each calorimeter deposition by associating calorimeter cells with reconstructed physics objects and using the appropriate calibration scheme. The total missing transverse energy is then calculated as the sum of terms,

$$E_{x(y)}^{\mathrm{miss}} = E_{x(y)}^{\mathrm{miss},e} + E_{x(y)}^{\mathrm{miss},\gamma} + E_{x(y)}^{\mathrm{miss},\tau} + E_{x(y)}^{\mathrm{miss,jets}} + E_{x(y)}^{\mathrm{miss,softjets}} + E_{x(y)}^{\mathrm{miss,CellOut}} + E_{x(y)}^{\mathrm{miss},\mu}, \quad (3.7)$$

where each term is given by:

- $E_{x(y)}^{\mathrm{miss},e}$: this is calculated from electrons which pass the medium electron identification requirement (Section 3.2.3) with $p_{\mathrm{T}} > 10$ GeV and calibrated with the standard electron calibration (Section 3.2.4).

- $E_{x(y)}^{\mathrm{miss},\gamma}$: this is calculated from photons which pass the "tight" photon identification requirement [118] with $p_{\mathrm{T}} > 10$ GeV and calibrated with the standard photon calibration.

- $E_{x(y)}^{\mathrm{miss},\tau}$: this is calculated from τ jets which pass the "tight" τ identification requirement [119] calibrated with the local calibration weighting (LCW) scheme[5] [120] and required to have $p_{\mathrm{T}} > 10$ GeV.

- $E_{x(y)}^{\mathrm{miss,jets}}$: this is calculated from anti-k_t, $R = 0.4$ jets with $p_{\mathrm{T}} > 20$ GeV calibrated with the LCW scheme and the same procedure outlined in Section 3.4.2.

- $E_{x(y)}^{\mathrm{miss,softjets}}$: this is calculated from LCW jets with $7 < p_{\mathrm{T}} < 20$ GeV (without jet calibration applied). Note that this is separated from the $E_{x(y)}^{\mathrm{miss,jets}}$ as the

[5]LCW uses cluster properties to calibrate clusters to the hadronic level by classifying clusters as electromagnetic or hadronic and correcting for dead material and unclustered calorimeter energy.

jet calibration procedure is not expected to perform well at low p_{T}. This separation has been optimised centrally.

- $E_{x(y)}^{\mathrm{miss,CellOut}}$: this is calculated from LCW topo-clusters outside reconstructed objects and from low p_{T} reconstructed tracks to recover the contribution from particles which do not reach the calorimeter or are too low energy to seed topo-clusters.

- $E_{x(y)}^{\mathrm{miss},\mu}$: this is calculated from the momenta of muon tracks reconstructed in the region $|\eta| < 2.7$. Combined muons (requiring an associated ID track, see Section 3.3.2) are used in the central ($|\eta| < 2.5$) region.

3.6 Event selection

In this analysis, events are first required to pass a Good Runs List (GRL) requirement to ensure that all parts of the detector were functioning correctly and taking data. This ensures good reconstruction of all physics objects. In addition, any event affected by LAr noise bursts or data corruption (as defined centrally) are removed.

Following the identification of electrons and muons outlined in Sections 3.2 and 3.3, W and Z events are selected. W events first require the presence of exactly one selected lepton (electron or muon). The selected lepton is required to be consistent with the object which fired the lepton trigger (trigger matched). In addition, the missing transverse energy (see Section 3.5) is required to be $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV. The transverse mass, m_{T}, is defined as

$$m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{l} E_{\mathrm{T}}^{\mathrm{miss}} \left(1 - \cos\left(\Delta\phi\left(l, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)\right)}, \tag{3.8}$$

where p_{T}^{l} is the lepton transverse momentum and $\Delta\phi\left(l, E_{\mathrm{T}}^{\mathrm{miss}}\right)$ is the angle in the plane transverse to the beam between the missing transverse energy and the lepton. The requirement $m_{\mathrm{T}} > 40$ GeV is included. For the Z selection, exactly two selected leptons are required with isolation requirements removed on the sub-leading lepton. The leptons are required to be of opposite sign and the dilepton mass is required to fall in the range $66 < m_{ll} < 116$ GeV. Finally, the leptons are required to be separated by $\Delta R(ll) > 0.2$. The full selection in the electron and muon channels is summarised in Tables 3.3 and 3.4 respectively.

Once W and Z events have been selected, the jets are selected as outlined in Section 3.4. In addition, the jets are required to be isolated from the selected leptons by the requirement $\Delta R(l, \text{jet}) > 0.5$. If a jet does not satisfy this requirement the jet is removed from the selection. The full jet selection is summarised in Table 3.5.

Pre-selection	
Vertex	≥ 1 good vertex: $N_{\text{tracks}} \geq 3$
Data quality	remove events affected by LAr noise and data corruption
LAr Hole	Veto events with jet in LAr hole
$E_{\mathrm{T}}^{\mathrm{miss}}$ cleaning	$E_{\mathrm{T}}^{\mathrm{miss}}$ cleaning cuts (event rejected if any BadLooser jet found with $p_T > 20$ GeV and $\lvert\eta\rvert < 4.5$)
Electron selection	
Author[6]	1 or 3
Transverse momentum	$p_{\mathrm{T}} > 25$ GeV
Rapidity	$\lvert\eta\rvert < 2.47$ (excluding $1.37 < \lvert\eta\rvert < 1.52$)
Electron ID	Tight++
Electron Isolation*	Passes `iso97Ptcone40` && `iso98Etcone20`
Calorimeter problems[7]	OQ&1446 == 0
	* Applied to W electron and leading electron in Z selection
$Z \to ee$ event selection	
Triggers	EF_e20_medium (D-J), EF_e22_medium (K), EF_e22vh_medium1 (L-M)
Electrons	Exactly 2 selected electrons
Charge	Opposite sign
Invariant Mass	$66 < m_{ee} < 116$ GeV
Electron separation	$\Delta R(ll) > 0.2$
$W \to e\nu$ event selection	
Triggers	EF_e20_medium (D-J), EF_e22_medium (K), EF_e22vh_medium1 (L-M)
Electrons	Exactly 1 selected electron
Missing energy	$E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV
Transverse mass	$m_{\mathrm{T}} > 40$ GeV

Table 3.3: Criteria used for the selection of $W \to e\nu$ and $Z \to ee$ events.

[6] Electron author defines the algorithm which has been used to reconstruct an electron. 1 means the electron has been found using the standard cluster-based reconstruction described in the text, 2 means the electron has been found using an alternative soft, track-based algorithm and 3 means the electron has been reconstructed by both algorithms. Here, requiring 1 or 3 is equivalent to requiring the electron to be reconstructed by the cluster-based algorithm.

[7] The electron OQ defines an object quality bitmask. The OQ&1446 == 0 requirement removes bad electrons, defined as electrons with clusters affected by dead front end boards (FEBs), dead high voltage supplies (HV) or cells which are masked due to being dead or noisy.

Pre-selection	
Vertex	≥ 1 good vertex: $N_{\text{tracks}} \geq 3$
Data quality	remove events affected by LAr noise and data corruption
LAr Hole	Veto events with jet in LAr hole
$E_{\text{T}}^{\text{miss}}$ cleaning	$E_{\text{T}}^{\text{miss}}$ cleaning cuts (event rejected if any BadLooser jet found with $p_T > 20$ GeV and $\lvert\eta\rvert < 4.5$)
Muon selection	
Reconstruction	STACO combined
Transverse momentum	$p_{\text{T}} > 25$ GeV
Rapidity	$\lvert\eta\rvert < 2.4$
Quality	> 0 b-layer hits > 1 pixel hits > 5 SCT hits < 3 pixel and SCT holes > 8 TRT hits
Isolation*	$p_{\text{T}}^{\text{cone20}}/p_{\text{T}} < 0.1$
Impact parameter*	$\lvert d_0/\sigma(d_0)\rvert < 3$
	* Applied to W electron and leading electron in Z selection
$Z \to \mu\mu$ event selection	
Triggers	EF_mu18_MG or EF_mu18 (D-I) EF_mu18_MG_medium or EF_mu18_medium (J-M)
Muons	Exactly 2 selected muons
Charge	Opposite sign
Invariant Mass	$66 < m_{\mu\mu} < 116$ GeV
Muon separation	$\Delta R(ll) > 0.2$
$W \to \mu\nu$ event selection	
Triggers	EF_mu18_MG or EF_mu18 (D-I) EF_mu18_MG_medium or EF_mu18_medium (J-M)
Muons	Exactly 1 selected muon
Missing energy	$E_{\text{T}}^{\text{miss}} > 25$ GeV
Transverse mass	$m_{\text{T}} > 40$ GeV

Table 3.4: Criteria used for the selection of $W \to \mu\nu$ and $Z \to \mu\mu$ events.

Jet selection	
Jet isolation	$\Delta R(\ell, \text{jet}) > 0.5$ (jet veto)
Jet rapidity	$\lvert y\rvert < 4.4$
Pileup-jet rejection	$\lvert\text{JVF}\rvert \leq 0.75$ if $\lvert\eta\rvert < 2.4$
Transverse momentum	$p_{\text{T}} > 30$ GeV

Table 3.5: Criteria used for the selection of jets.

Chapter 4

Background Estimation

In this Chapter, the background estimation methods and uncertainties are discussed. Some backgrounds were estimated using Monte Carlo simulation, as discussed in Section 4.1. Multi-jet and $t\bar{t}$ backgrounds are challenging to model correctly in Monte Carlo and, particularly for the case of top quark pair background, Monte Carlo is known to not describe the data well at high multiplicity. As such, these two background sources were estimated using data-driven methods and are described in 4.2 and 4.3.

4.1 Monte Carlo Background Estimation

As summarised in Table 4.1, some backgrounds were estimated using Monte Carlo simulation. In Section 1.2.4 the Monte Carlo simulation of signal samples using Alpgen and Sherpa was discussed. These same samples are used to estimate "cross-talk" backgrounds ($W+$ jets events reconstructed as $Z/\gamma^* +$ jets and vice-versa). In the Z channel and at low jet multiplicity (less than three) in the W channel, $t\bar{t}$ Monte Carlo was used, using Alpgen 2.13 [42] interfaced to Herwig 6.520 [43] and Jimmy 4.31 [44] with CTEQ6L1 [45] PDFs and the AUET2-CTEQ6L1 tune [46]. This sample includes up to 5 additional partons in the final

Process	Generator and PDFs	σ (nb)	Note
$W \to e\nu$	ALPGEN (+HERWIG),CTEQ6L1	10.46	
$W \to \mu\nu$	ALPGEN (+HERWIG),CTEQ6L1	10.46	
$W \to e\nu$	SHERPA, v1.4, CT10	10.46	
$W \to \mu\nu$	SHERPA, v1.4, CT10	10.46	
$W \to \tau\nu$	ALPGEN (+HERWIG),CTEQ6L1	10.46	
$Z/\gamma^*(\to ee)$	ALPGEN (+HERWIG),CTEQ6L1	1.07	$m_{ll} > 40$ GeV
$Z/\gamma^*(\to \mu\mu)$	ALPGEN (+HERWIG),CTEQ6L1	1.07	$m_{ll} > 40$ GeV
$Z/\gamma^*(\to \tau\tau)$	ALPGEN (+HERWIG),CTEQ6L1	1.07	$m_{ll} > 40$ GeV
$Z/\gamma^*(\to ee)$	SHERPA, v1.4, CT10	1.07	$m_{ll} > 40$ GeV
$Z/\gamma^*(\to \mu\mu)$	SHERPA, v1.4, CT10	1.07	$m_{ll} > 40$ GeV
WZ	HERWIG, MRST	$17.5 \cdot 10^{-3}$	filter eff.: 0.30986
ZZ	HERWIG, MRST	$6.49 \cdot 10^{-3}$	filter eff.: 0.21152
WW	HERWIG, MRST	$4.50 \cdot 10^{-2}$	filter eff.: 0.38863
$t\bar{t}$	ALPGEN (+HERWIG), CTEQ6L1	0.1773	filter eff.: 0.54295
single top, t-channel	ACER + PYTHIA, MRST	$6.97 \cdot 10^{-3}$	
single top, s-channel	ACER + PYTHIA, MRST	$5.0 \cdot 10^{-4}$	
single top, Wt	ACER + PYTHIA, MRST	$1.57 \cdot 10^{-2}$	

Table 4.1: Samples of simulated events used in the analysis. The cross sections quoted are the ones used to normalise the estimates of expected number of events.

state. Single t production was simulated using ACERMC 3.8 [121] interfaced to PYTHIA [55] with MRST LO* PDFs [122]. Diboson WW, WZ and ZZ production was simulated using HERWIG 6.510 and JIMMY 4.3 with MRST LO* PDFs and the AUET2-LO* [46] tune.

These samples were normalised to the highest order perturbative QCD prediction available. V + jets samples were normalised to the NNLO prediction calculated with FEWZ [123] and the MSTW2008 NNLO PDFs [35]. The $t\bar{t}$ samples were normalised to the cross section calculated at NNLO + NNLL [124–129] and the single t samples were normalised to the cross sections calculated at NNLO + NNLL [130–132]. The diboson samples were normalised to cross sections calculated at NLO using MCFM [133] with the MSTW2008 PDF set.

4.1.1 Monte Carlo Background Uncertainties

Background predictions taken from Monte Carlo were normalised to their respective (N)NLO cross sections as discussed above. The uncertainty on these cross sections was used as a systematic for each background. The cross sections of the $V+\text{jets}$ samples were assigned a 5% uncertainty stemming from scale and PDF uncertainties in the NNLO prediction [106]. The uncertainty on the cross section of diboson processes was taken as 5% for WW and ZZ and 7% for WZ [133]. The uncertainty on the single t was taken as 3.4% for the t-channel, 4% for the s-channel and 7% for the Wt-channel [130–132]. Where the $t\bar{t}$ estimate was taken from Monte Carlo (in $W+\text{jets}$ for 0, 1 and 2 jet multiplicities and in $Z/\gamma^*+\text{jets}$ for all multiplicities), a 6% normalisation uncertainty was used [124–129]. An additional shape uncertainty on the $t\bar{t}$ Monte Carlo estimation was included by comparing the nominal Alpgen prediction with that generated using Powheg+Pythia [134].

4.2 Multi-jet backgrounds

Multi-jet backgrounds to $V+\text{jets}$ production arise primarily from jets faking leptons (dominant in the electron channel) or real leptons coming from heavy flavour decays within jets (dominant in the muon channel). Due to the large uncertainty on when an isolated lepton will be reconstructed from a jet and the very large Monte Carlo samples which would be required to simulate such backgrounds, data-driven techniques were employed.

In order to carry out a data-driven estimate of the multi-jet backgrounds a template method was employed. First, a set of events was taken from the data using a selection designed to give a sample enriched in background events. This selection was designed to be as close as possible to the nominal selection used in the analysis but with carefully selected cuts having been changed to reduce the efficiency of signal selection whilst reducing the background rejection. The assumption was made that the shape of the background in the nominal selection was the same as that in the background enriched selection. This assumption is one of the reasons that cut variations had to be carefully chosen to avoid biasing the background template shape.

Having selected a suitable background template from the data, a fit was performed using the background enriched template taken from data and a template taken from Monte Carlo predicting the shape of signal and other background events in the nominal selection. These two templates were fitted to the distribution in data arising from the nominal selection.

Such a template fit allowed the normalisation of the background template (taken from data) to be ascertained and so, the shape and normalisation of the background in the nominal selection. Such a fit was carried out on a carefully chosen kinematic distribution for which different shapes were expected between the data driven and Monte Carlo templates. This shape difference is a requirement of such template methods as the fit is only successful given a shape difference between the two templates.

4.2.1 $W \to e\nu$ channel

To normalise the multi-jet template in the $W \to e\nu$ channel, a fit was carried out using the multi-jet template and a template formed from the sum of $W \to e\nu$ signal (from ALPGEN) and background ($t\bar{t}$, diboson, $W \to \tau\nu$, $Z \to \tau\tau$ and single-t) Monte Carlo contributions in the signal region. The normalisation of the two templates was varied to obtain the best fit to the data in the signal region.

The fit was carried out using the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution before the application of the $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV cut to take advantage of the large shape differences at low values of $E_{\mathrm{T}}^{\mathrm{miss}}$. The fit was carried out in exclusive jet multiplicity bins up to $N_{\mathrm{jet}} = 5$ above which the multi-jet template had limited statistics and so, the scale factor for $N_{\mathrm{jet}} > 5$ was taken from the $N_{\mathrm{jet}} = 5$ fit.

The number of expected multi-jet events with $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV returned from the

$E_{\mathrm{T}}^{\mathrm{miss}}$ fit was then used to scale the multi-jet template in other distributions.

Template Selection

The control sample in the $W \rightarrow e\nu$ channel was selected using data accepted by electron triggers using looser identification requirements than in the signal selection. These triggers are listed in Table 4.2. The jet based triggers, *_EFFS, which trigger on a loose electron in association with additional jets were included in later periods to increase the template statistics.

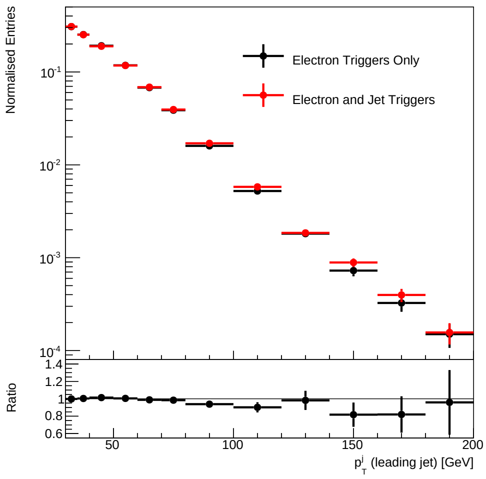
The jet based triggers can introduce a shape bias[1] compared to just using loose electron triggers. To understand the effect of this bias, the fit is rerun with and without the jet triggers included. The fitted multi-jet fraction was found to be consistent with and without triggers for at least two jets and up and so, jet triggers were only employed for events with at least two jets when building the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution.

The bias was also investigated in differential distributions such as leading jet p_{T} and rapidity, as shown in Figures 4.1 and 4.2. Biases are observed in the jet p_{T} distributions for leading and second leading jet and in the jet rapidity distributions for leading, second leading and third leading jet. As a result, the jet triggers were only employed in events with at least 3 jets for p_{T}-based distributions and at least 4 jets for jet rapidity based distributions.

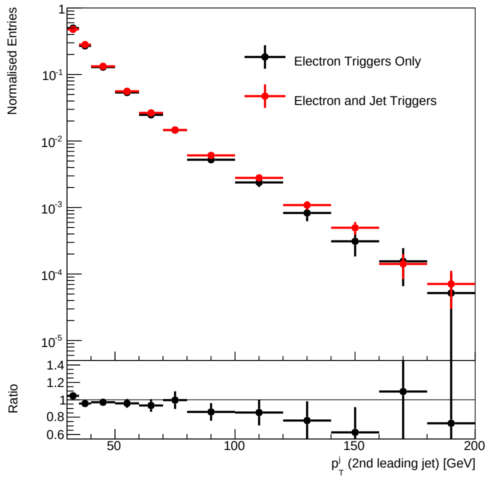
Note that, although these control triggers have looser electron criteria compared to the signal triggers, the number of events in the multi-jet templates is reduced by the necessity of applying large prescales. The control triggers therefore correspond to an integrated luminosity of between 113 pb^{-1} (excluding jet triggers) and 196 pb^{-1} (including jet triggers).

Due to variations in the prescales applied to these triggers during 2011, the

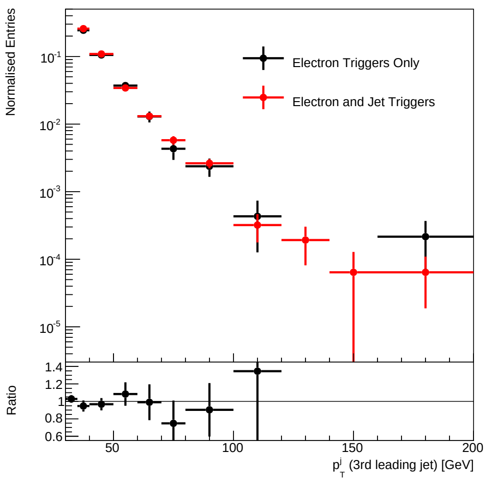
[1]Note that an assumption is made here that the use of loose electron triggers provides a good description of the multi-jet shapes in the signal region. This assumption is expected to be fair given the good agreement at detector level demonstrated in Refs. [106] and [8].



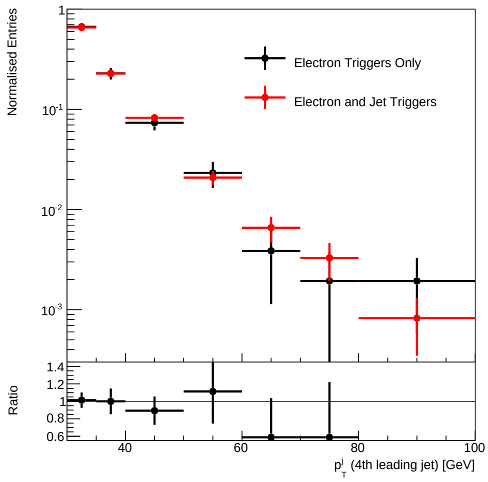
(a) Leading jet p_T



(b) 2nd leading jet p_T

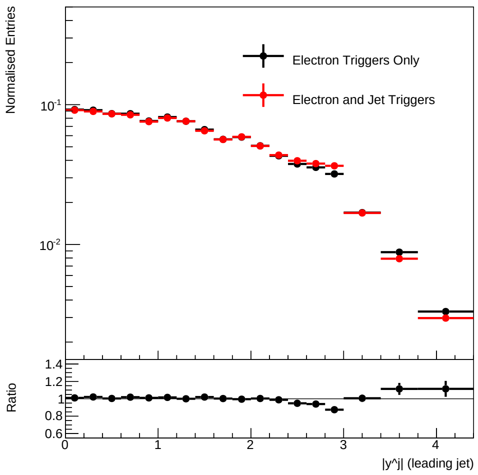


(c) 3rd leading jet p_T

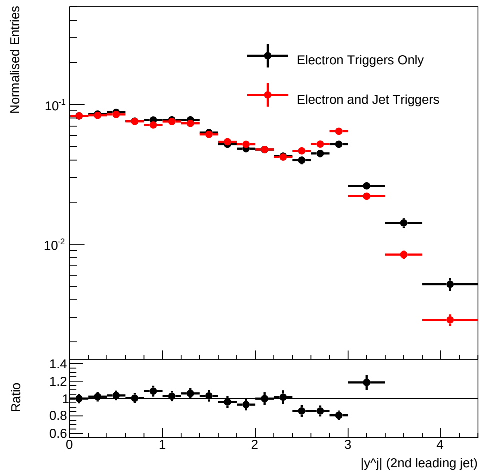


(d) 4th leading jet p_T

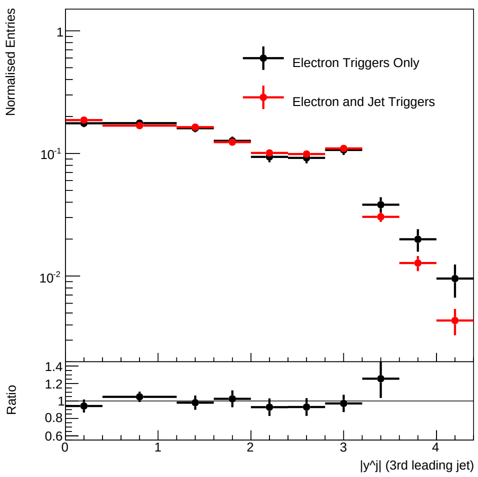
Figure 4.1: Comparison of jet p_T template shapes in periods L and M for multi-jet templates using just electron triggers and templates using both electron and jet triggers.



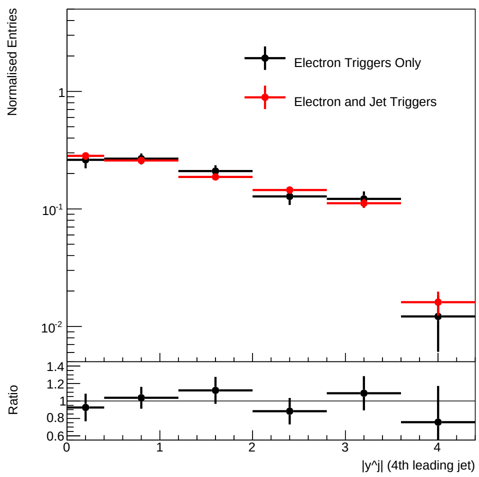
(a) Leading jet y



(b) 2nd leading jet y



(c) 3rd leading jet y



(d) 4th leading jet y

Figure 4.2: Comparison of jet y template shapes in periods L and M for multi-jet templates using just electron triggers and templates using both electron and jet triggers.

Data Period(s)	Triggers
L5, M	EF_e22vh_loose1, EF_e22vh_looseTrk, EF_e22vh_loose_4j15_a4tc_EFFS, EF_e22vh_loose_3j20_a4tc_EFFS, EF_e22vh_loose_2j20_a4tc_EFFS
L1,L2,L3,L4	EF_e22vh_loose1, EF_e22vh_looseTrk
K	EF_e22_looseTrk, EF_e22_loose1, EF_e22_loose
J2	EF_e22_looseTrk, EF_e22_loose1, EF_e22_loose, EF_e20_looseTrk, EF_e20_loose1, EF_e20_loose
G, H, I, J1	EF_e20_looseTrk, EF_e20_loose1, EF_e20_loose
D, E, F	EF_e20_loose

Table 4.2: Triggers used to select a multi-jet enriched template in the $W \to e\nu$ channel. The jet-based triggers, *_EFFS, were only used for events in higher jet multiplicity bins as outlined in the text.

relative contribution to the multi-jet templates from the various data periods was not proportional to the luminosity recorded in the signal selection. In addition, pile-up conditions varied dramatically over the year which influenced the multi-jet background shape and rate. Figure 4.3 shows that the data fall into two distinct regimes - early 2011 data (periods D-K) and late 2011 data (periods L-M)[2]. As a result of this observation, separate templates were formed for early and late data and fitted separately. These fit results were then summed after fitting to give the total multi-jet background for the full data period.

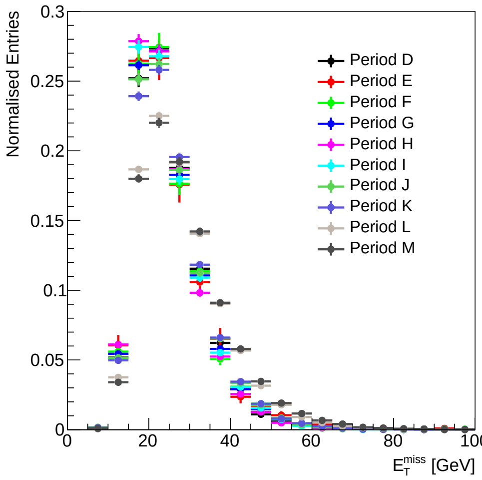
Offline electron identification requirements were also relaxed in the multi-jet template by requiring the electron to pass loose identification cuts from Ref. [102] and a subset of the medium identification cuts (see Section 3.2.3):

- TrackPixel_Electron (number of hits in the pixel detector)
- TrackSi_Electron (number of hits in the pixel and SCT detectors)
- TrackMatchEta_Electron (η match between track and cluster)

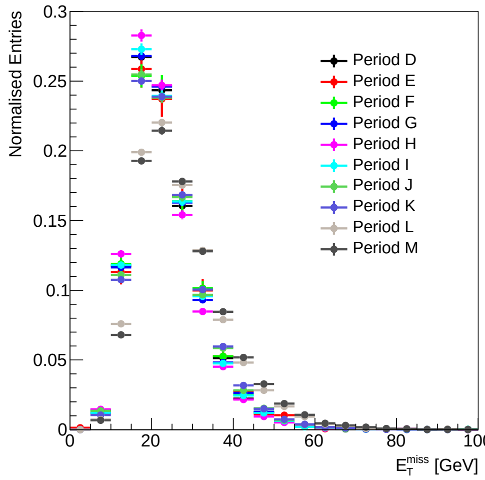
The electron was also required to fail the tight++ identification used in the signal selection. At least one of the tight requirements was required to be failed omitting:

- ConversionMatch_Electron (electron matched to photon conversion)

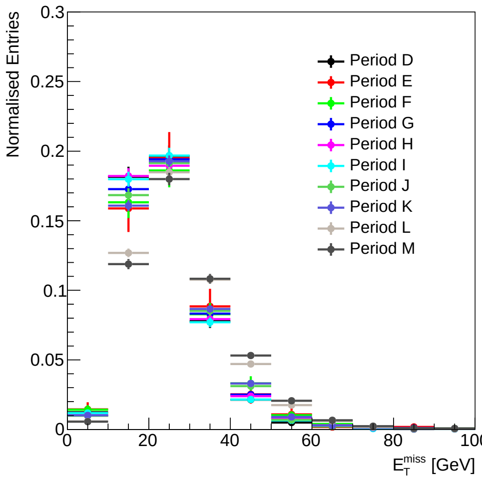
[2]See Section 2.2.5 for a description of the alphabetical data periods.



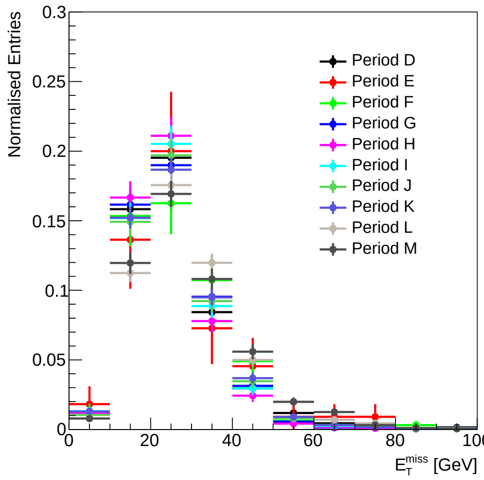
(a) $N_{\text{jet}} = 0$



(b) $N_{\text{jet}} = 1$



(c) $N_{\text{jet}} = 2$



(d) $N_{\text{jet}} = 3$

Figure 4.3: Comparison of $E_{\text{T}}^{\text{miss}}$ template shapes between different periods of 2011 data taking.

- TrackMatchPhi_Electron (ϕ match between track and cluster)
- TrackMatchEtaTight_Electron (tighter η match between track and cluster)

Electrons were required to be anti-isolated in the calorimeter by requiring the energy deposited in a cone of $\Delta R = 0.3$ ($E_{\mathrm{T}}^{\mathrm{cone30}}$) to be larger than 20% of the total transverse energy of the electron.

This isolation requirement was chosen using two quantities of merit:

1. χ^2/ndf for the fit defined as:

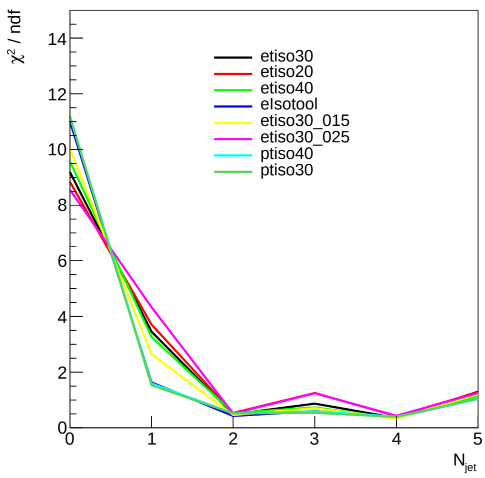
$$\chi^2 = \sum_{\mathrm{bins}} \frac{\left(N_{\mathrm{data}} - N_{\mathrm{MC}} - N_{\mathrm{multi-jet}}\right)^2}{\left(\sigma_{\mathrm{MC}}^{\mathrm{stat}}\right)^2 + \left(\sigma_{\mathrm{data}}^{\mathrm{stat}}\right)^2 + \left(\sigma_{\mathrm{multi-jet}}^{\mathrm{stat}}\right)^2}. \tag{4.1}$$

2. Contamination: the percentage of the multi-jet template coming from other backgrounds as predicted by Monte Carlo.

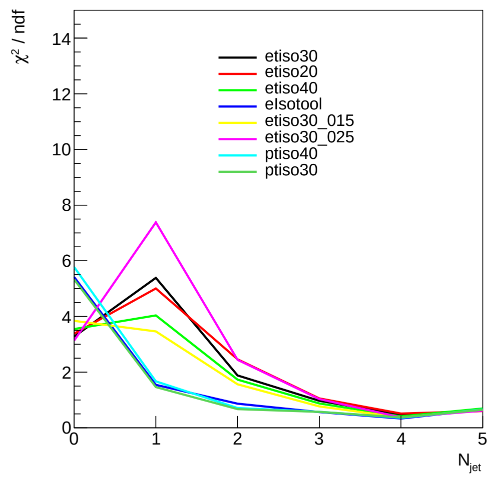
Figure 4.4 shows these two quantities as a function of jet multiplicity for a range of different templates[3]:

1. $E_{\mathrm{T}}^{\mathrm{cone20}}/E_{\mathrm{T}} < 0.20$ (etiso20)
2. $E_{\mathrm{T}}^{\mathrm{cone30}}/E_{\mathrm{T}} < 0.20$ (etiso30)
3. $E_{\mathrm{T}}^{\mathrm{cone40}}/E_{\mathrm{T}} < 0.20$ (etiso40)
4. $E_{\mathrm{T}}^{\mathrm{cone30}}/E_{\mathrm{T}} < 0.15$ (etiso30_015)
5. $E_{\mathrm{T}}^{\mathrm{cone30}}/E_{\mathrm{T}} < 0.25$ (etiso30_025)
6. $p_{\mathrm{T}}^{\mathrm{cone30}}/p_{\mathrm{T}} < 0.20$ (ptiso30)
7. $p_{\mathrm{T}}^{\mathrm{cone40}}/p_{\mathrm{T}} < 0.20$ (ptiso40)

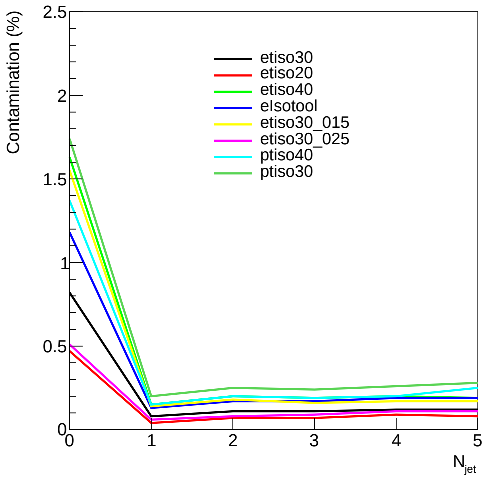
[3] $E_{\mathrm{T}}^{\mathrm{coneXX}}$ denotes the amount of energy in the calorimeter in a cone of $\Delta R = XX/100$ around the electron whilst $p_{\mathrm{T}}^{\mathrm{coneXX}}$ denotes the sum of the track momenta in a cone of $\Delta R = XX/100$ around the electron as discussed in Section 3.2.5.



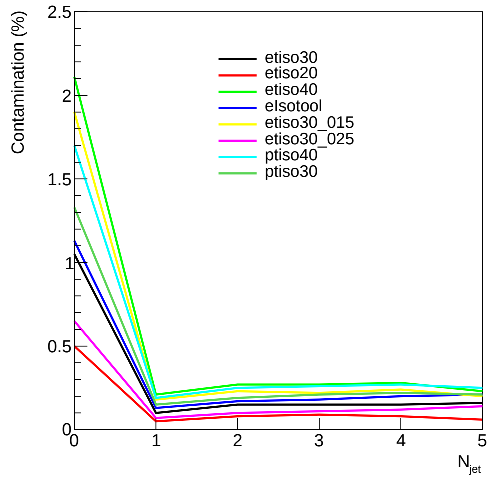
(a) χ^2 - D to K



(b) χ^2 - L to M



(c) Contamination - D to K



(d) Contamination - L to M

Figure 4.4: Comparison of χ^2 and contamination of various template definitions in the electron channel multi-jet estimate.

8. reversed signal isolation (eIsotool)

It is observed that the χ^2 values are large for low jet multiplicity (0, 1 and 2 jets). Note that the large χ^2 values in the 0 jet bin are not significant to the measurements as these events are not included in the unfolded differential results. In the 1 and 2 jet bins, it is observed that the calorimeter isolation based templates have larger χ^2 values. This is attributed to mis-modelling at low missing energy ($E_{\mathrm{T}}^{\mathrm{miss}} < 20$ GeV) as raising the lower edge of the fit range improves the χ^2 with little change in the predicted number of multi-jet events (see below). Little value is given to this observation as the final measurement applies a cut of $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV.

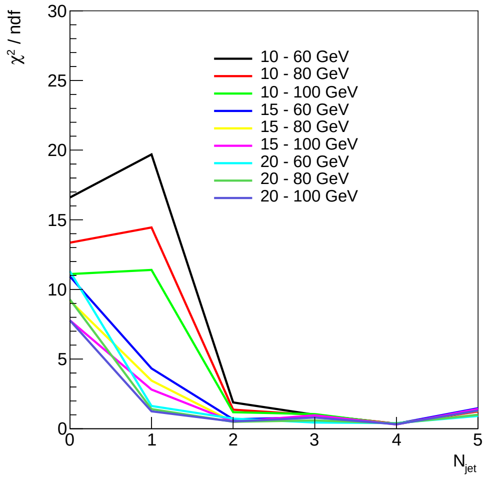
The contamination of the multi-jet templates is seen to be small for events with 1 or more jets in all template choices. There is little to choose between templates in terms of contamination, although some discrimination is seen.

From these studies, the etiso30 template was chosen as the nominal template as it has reasonably low contamination and competitive χ^2. Furthermore, this choice is aligned with other ongoing analyses in ATLAS. A less contaminated template was not chosen as the least contaminated templates have fewer events in them. Therefore, this choice represented a balance between statistical accuracy and signal contamination.

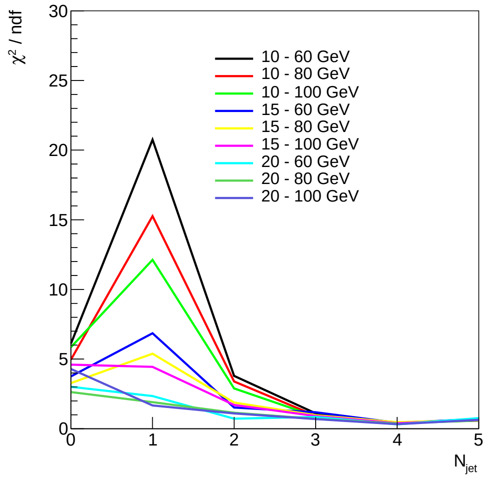
Fit Range Choice

To find the best range in $E_{\mathrm{T}}^{\mathrm{miss}}$ to perform the fit, the fit was carried out over a set of possible ranges using the nominal template. Figure 4.5 shows the χ^2 and multi-jet fraction for different fit ranges.

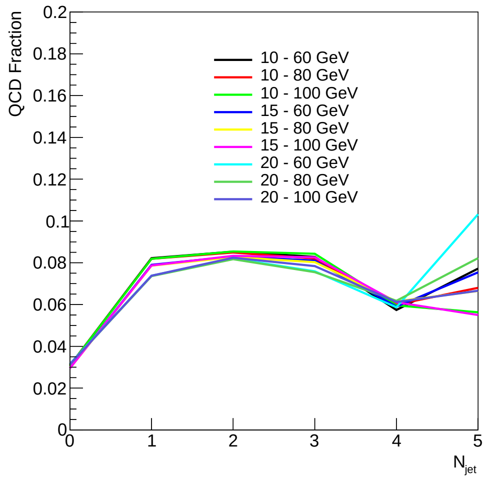
It is observed that the large χ^2 seen in calorimeter isolation based templates in the 1 and 2 jet bins (see above) is reduced when the lower limit of the fit range is increased, thus demonstrating that the larger χ^2 values arise due to mis-modelling



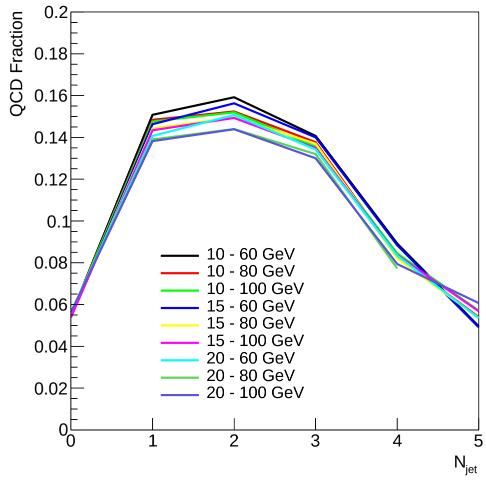
(a) χ^2 - D to K



(b) χ^2 - L to M



(c) Multi-jet Fraction - D to K



(d) Multi-jet Fraction - L to M

Figure 4.5: Comparison of χ^2 and multi-jet fraction (QCD fraction) of various choices of fit range in the electron channel multi-jet estimate.

at low $E_{\mathrm{T}}^{\mathrm{miss}}$. Away from the 1 and 2 jet bins, there is little to choose between fit range choices.

The nominal fit range was chosen to be $15 < E_{\mathrm{T}}^{\mathrm{miss}} < 80$ GeV as this fit range gives reasonable χ^2 values and represents a balance between including many bins to increase the statistical precision of the fit whilst avoiding regions of poor modelling which may bias the fit. Nevertheless, all the fit ranges shown here were used to assign a fit range systematic as discussed below.

Figures 4.6 and 4.7 show the central fit results from the multi-jet estimate.

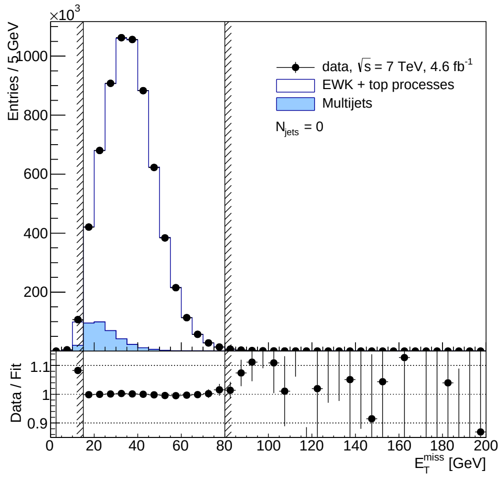
Systematic Uncertainties

The systematic uncertainty on the multi-jet background contribution was estimated by varying the procedure in a number of ways.

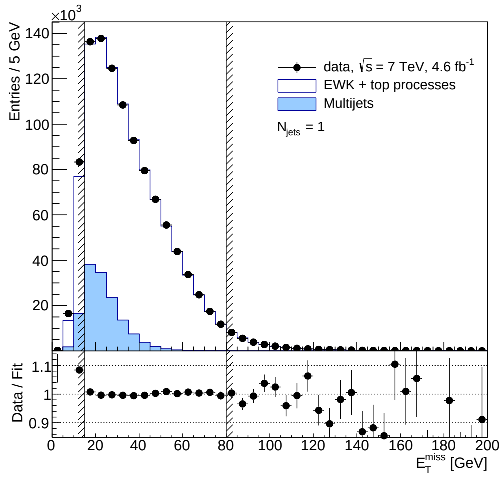
Electron Identification To assess the uncertainty associated to the choice of which tight requirements are allowed to be passed, alternative templates were created removing the requirement to fail selected parts of the tight identification.

Isolation Requirement Alternative isolation templates, as listed above, were used to estimate the uncertainties associated to the choice of isolation requirement. The effect of the choice of cone size was studied using the etiso20 and etiso40 templates. The effect of the choice of isolation cut was investigated using the etiso30_015 and etiso30_025 templates. Finally, the difference between the calorimeter- and track-based isolation templates was included by comparing the nominal template to the eIsotool template with reversed signal isolation.

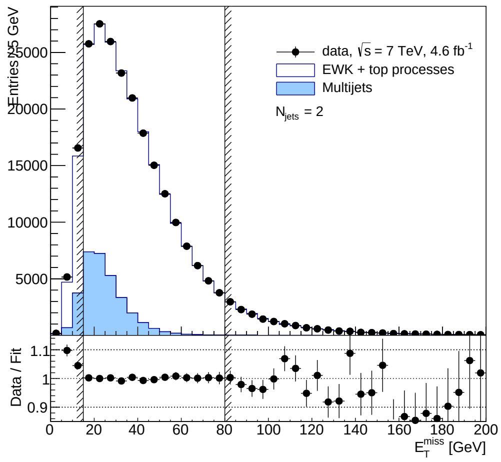
Fit Range The nominal fit range of 15-80 GeV was varied using alternative minima of 10 or 20 GeV and alternative maxima of 60 or 100 GeV to investigate the



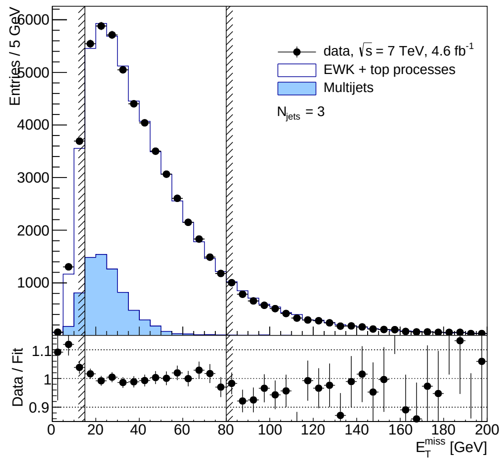
(a) $N_{\text{jet}} = 0$



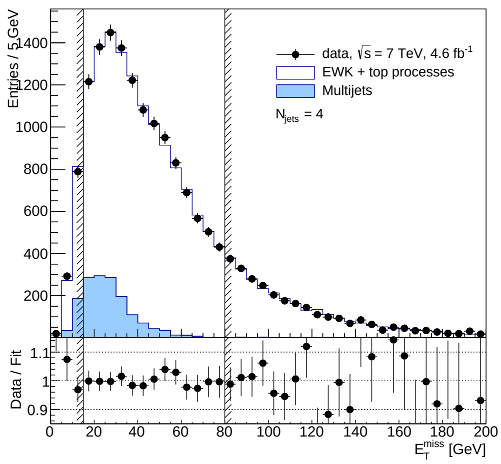
(b) $N_{\text{jet}} = 1$



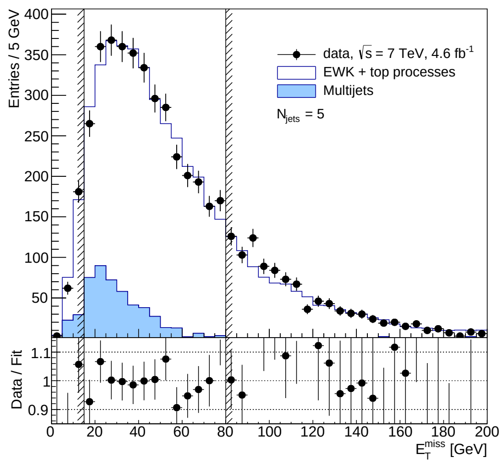
(c) $N_{\text{jet}} = 2$



(d) $N_{\text{jet}} = 3$

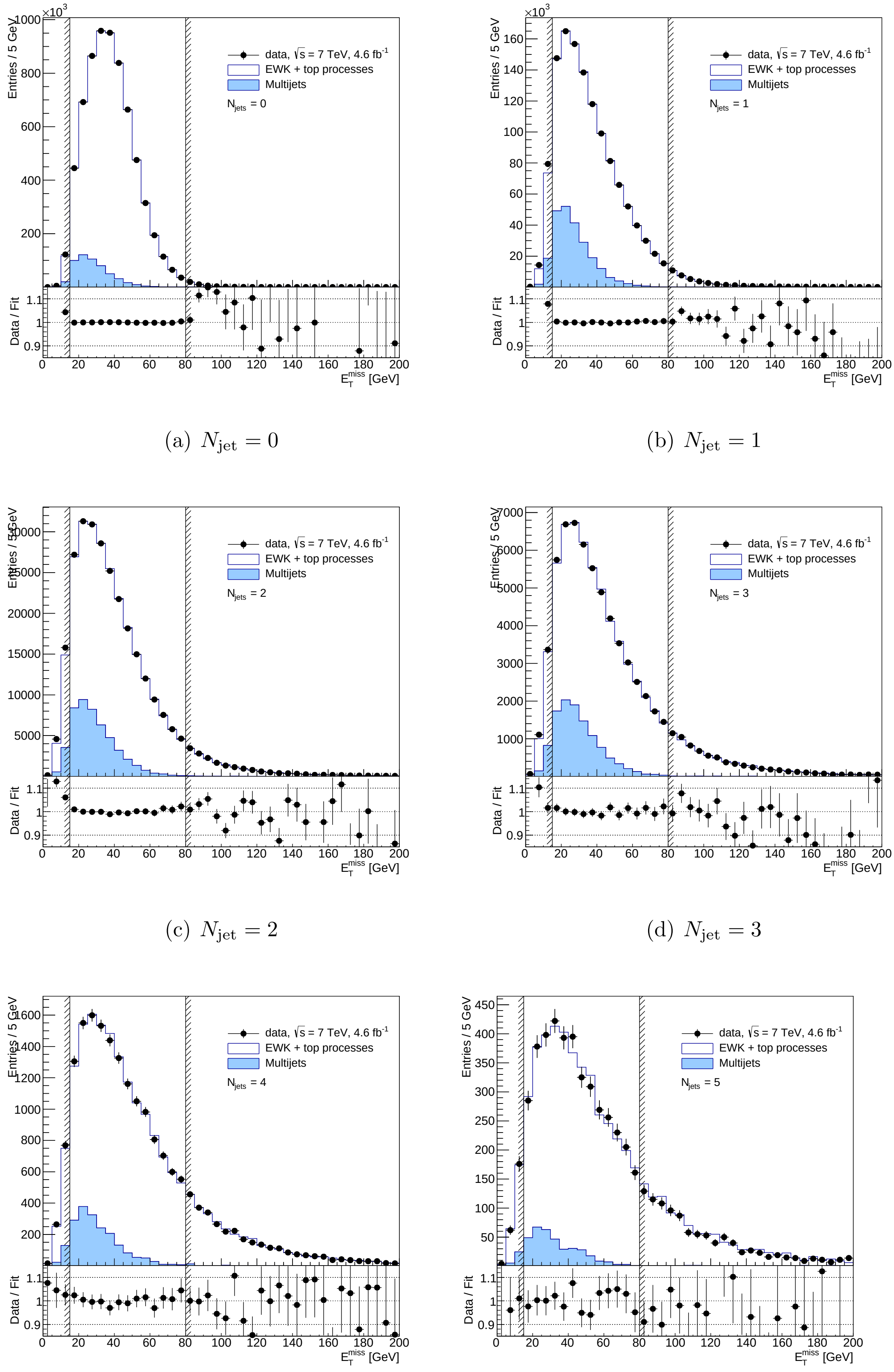


(e) $N_{\text{jet}} = 4$



(f) $N_{\text{jet}} = 5$

Figure 4.6: Nominal fits used in the multi-jet estimate in the $W \to e\nu$ channel for periods D to K. The vertical lines show the range over which the fit is performed.



(a) $N_{\mathrm{jet}} = 0$

(b) $N_{\mathrm{jet}} = 1$

(c) $N_{\mathrm{jet}} = 2$

(d) $N_{\mathrm{jet}} = 3$

(e) $N_{\mathrm{jet}} = 4$

(f) $N_{\mathrm{jet}} = 5$

Figure 4.7: Nominal fits used in the multi-jet estimate in the $W \to e\nu$ channel for periods L to M. The vertical lines show the range over which the fit is performed.

effect of changing the fit range.

Monte Carlo Template Modelling The fit was also carried out using SHERPA W and Z Monte Carlo in place of the default ALPGEN samples. The result of this fit was compared to the nominal fit to ascertain the size of the effect of Monte Carlo modelling on the fit results.

Statistical Uncertainty The statistical uncertainty on the fit result was also included as an uncertainty on the multi-jet background. This is only significant for high jet multiplicity bins.

The total uncertainty on the multi-jet background was formed by taking the envelope of all variations in each of the above categories and adding the resultant uncertainties in quadrature. In the cases where the variations cause a shape change in the multi-jet template (isolation and identification variations), the shape change was also propagated and included in the formation of the envelope. When adding the multi-jet backgrounds from the two periods the systematics for the two periods were treated as fully correlated.

In addition to the above multi-jet specific uncertainties, the Monte Carlo based template was varied according to the detector level systematic variations (to be discussed in Chapter 5) so that detector level systematics are propagated fully through the data-driven method.

Table 4.3 shows the estimated multi-jet fraction and the fractional uncertainty on the multi-jet background estimation as a function of jet multiplicity.

Periods D-K

	$N_{\mathrm{jet}} = 0$	$N_{\mathrm{jet}} = 1$	$N_{\mathrm{jet}} = 2$	$N_{\mathrm{jet}} = 3$	$N_{\mathrm{jet}} = 4$	$N_{\mathrm{jet}} = 5$
Multi-jet fraction	0.0295	0.0786	0.0827	0.0803	0.0615	0.0665
Statistics	0.0156	0.0142	0.0268	0.0514	0.1148	0.2174
Range up	0.0666	0.0418	0.0262	0.0292	0.0036	0.2362
Range down	0.0002	0.0631	0.0123	0.0598	0.0418	0.1735
Modelling up	-	-	-	-	-	0.0844
Modelling down	0.0599	0.0019	0.0051	0.0548	0.0183	-
Isolation up	0.1073	0.0203	0.0130	0.0211	0.0352	0.0913
Isolation down	0.1202	0.0672	0.0756	0.0591	0.0667	0.1013
ID up	0.0474	0.0508	0.0834	0.0703	0.0628	0.1966
ID down	0.0002	0.0631	0.0123	0.0598	0.0418	0.1735
Total up	0.1358	0.0704	0.0924	0.0942	0.1356	0.3964
Total down	0.1384	0.0964	0.0897	0.1194	0.1423	0.2997

Periods L-M

	$N_{\mathrm{jet}} = 0$	$N_{\mathrm{jet}} = 1$	$N_{\mathrm{jet}} = 2$	$N_{\mathrm{jet}} = 3$	$N_{\mathrm{jet}} = 4$	$N_{\mathrm{jet}} = 5$
Multi-jet fraction	0.0540	0.1443	0.1496	0.1370	0.0821	0.0536
Statistics	0.0182	0.0149	0.0253	0.0452	0.1062	0.2964
Range up	0.0448	0.0289	0.0443	0.0215	0.0763	0.0586
Range down	0.0026	0.0365	0.0376	0.0364	0.0586	0.0845
Modelling up	0.0298	0.0155	0.0003	0.0085	0.0573	0.4069
Modelling down	-	-	-	-	-	-
Isolation up	0.0387	0.0155	0.0204	0.0152	-	0.0390
Isolation down	0.1280	0.0904	0.0936	0.1081	0.0924	0.0737
ID up	0.0382	0.0383	0.0384	0.0332	0.1445	0.1159
ID down	0.0029	0.0242	0.0366	0.0294	0.0572	-
Total up	0.0786	0.0548	0.0670	0.0625	0.2031	0.5213
Total down	0.1294	0.1015	0.1102	0.1261	0.1629	0.3169

Combination

	$N_{\mathrm{jet}} = 0$	$N_{\mathrm{jet}} = 1$	$N_{\mathrm{jet}} = 2$	$N_{\mathrm{jet}} = 3$	$N_{\mathrm{jet}} = 4$	$N_{\mathrm{jet}} = 5$
Multi-jet fraction	0.0420	0.1147	0.1187	0.1112	0.0724	0.0599
Total up	0.0983	0.0596	0.0752	0.0729	0.1761	0.4538
Total down	0.1325	0.1000	0.1036	0.1239	0.1546	0.3076

Table 4.3: The estimated multi-jet fraction as a function of jet multiplicity in the $W \to e\nu$ channel and the fractional uncertainty on the predicted number of events arising from the various sources discussed in the text.

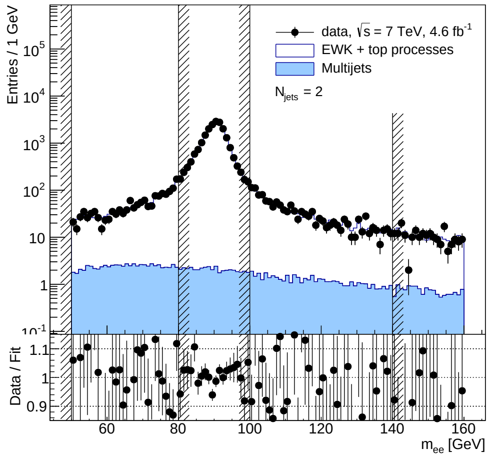


Figure 4.8: Nominal fit in the $N_{\mathrm{jet}} = 2$ bin used in the multi-jet estimate in the $Z \to ee$ channel. The vertical lines show the range over which the fit is performed (note the peak region is excluded).

4.2.2 $Z \to ee$ channel

The multi-jet background in the $Z \to ee$ background was also estimated using a template based method with the template selection similar to that employed in Ref. [135] and the fit performed using the m_{ee} distribution.

The multi-jet enriched template was built using events which passed a loose dielectron trigger (EF_2g20_loose). Events were required to pass the loose offline electron selection from Ref. [101] but fail the nominal selection.

The fit was carried out in an extended mass window of [50,140] GeV and the peak region was removed from the fit. The fit was carried out in exclusive jet multiplicity bins up to 2 jets. Above this, limited statistics result in a poor fit and so, the scaling factor for 2 jets was applied to higher jet bins.

Figure 4.8 shows the nominal fit in the 2 jet bin of the $Z \to ee$ channel.

Systematic Uncertainties

The systematic uncertainties were estimated by varying fit and selection parameters and comparing the shifted results to the nominal result:

Fit Range The nominal fit range of [50,140] GeV was varied to minima of 45 and 55 GeV and maxima of 135 and 145 GeV. The variations from the nominal fit result were used to form the envelope of the variations which was taken as the systematic uncertainty.

Monte Carlo Template Modelling The fit was also carried out using Sherpa W and Z Monte Carlo in place of the default Alpgen samples. The result of this fit was compared to the nominal fit to ascertain the size of the effect of Monte Carlo modelling on the fit results.

Statistical Uncertainty The statistical uncertainty on the multi-jet estimate due to statistics was also included as an uncertainty on the multi-jet background.

The total uncertainty on the multi-jet background was formed by summing the above systematic variations in quadrature. Table 4.4 summarises the estimated multi-jet background and associated uncertainties in the $Z \to ee$ channel.

In addition to the above multi-jet specific uncertainties, the Monte Carlo based template was varied according to the detector level systematic variations (to be discussed in Chapter 5) so that detector level systematics were propagated fully through the data-driven method.

	$N_{\mathrm{jet}} = 0$	$N_{\mathrm{jet}} = 1$	$N_{\mathrm{jet}} = 2$
Multi-jet fraction	0.002	0.002	0.004
Statistics	0.29	0.80	0.90
Range up	0.10	0.06	0.14
Range down	0.18	0.10	0.39
Modelling up	-	-	-
Modelling down	0.87	0.49	0.74
Total up	0.31	0.80	0.91
Total down	0.94	0.94	1.23

Table 4.4: The estimated multi-jet fraction as a function of jet multiplicity in the $Z \to ee$ channel and the fractional error from the various uncertainty sources discussed in the text.

4.2.3 $W \to \mu\nu$ channel

The multi-jet background in the $W \to \mu\nu$ channel was estimated in a similar way as in the $W \to e\nu$ channel as described in Section 4.2.1 - by performing a two component fit to the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution.

Template Selection

To find a suitable multi-jet template, 7 template selections were investigated:

1. AntiD0WithIso - muon fails $|d_0/\sigma(d_0)| < 3$ but includes signal isolation cut.
2. AntiD0WithIso_D0Window - muon fails $|d_0/\sigma(d_0)| < 3$ but includes signal isolation cut and additional $0.1 < |d_0| < 0.4$ cut.
3. AntiD0noIso - muon fails $|d_0/\sigma(d_0)| < 3$ and signal isolation requirement is removed.
4. AntiIsoNoD0 - muon fails signal isolation requirement and $|d_0/\sigma(d_0)| < 3$ requirement is removed.
5. AntiIsoWithD0 - muon fails signal isolation requirement but includes $|d_0/\sigma(d_0)| < 3$ requirement.

6. RestrAntiIso - muon isolation falls in window $0.1 < p_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.5$ and $|d_0/\sigma(d_0)| < 3$ requirement is removed.

7. RestrAntiIsoWithD0 - muon isolation falls in window $0.1 < p_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.5$ but includes $|d_0/\sigma(d_0)| < 3$ requirement.

Figure 4.9 shows the χ^2 of the fit procedure and the contamination of the multi-jet template by signal, $t\bar{t}$, single t and diboson processes as determined using Monte Carlo.

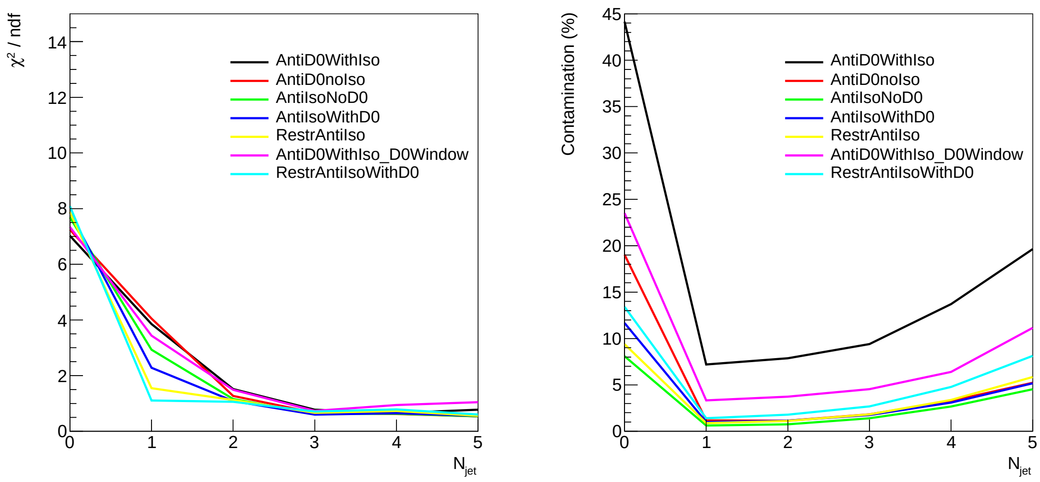
The contamination studies show that all the templates contain a significant number of events arising from non-multi-jet processes. As a result, the contamination was subtracted from the multi-jet template using Monte Carlo simulation before the template was used in the multi-jet fit. To minimise the effect of Monte Carlo mis-modelling on multi-jet background shapes, the most contaminated templates were removed from consideration.

The χ^2 values from the fit show the same large χ^2 values for 0 jet events as was observed in the $W \to e\nu$ channel. Again this was not considered a significant effect due to this bin not being used in any differential results. The χ^2 values for more than 1 jet show very little differentiation between templates however for 1 jet the χ^2 value favours the use of the RestrAntiIsoWithD0 template. Due to this and the low estimated contamination of this template, this template was chosen as the nominal template.

Fit Range Choice

To find the best range in $E_{\mathrm{T}}^{\mathrm{miss}}$ to perform the fit, the fit was carried out over a set of possible ranges using the nominal template. Figure 4.10 shows the χ^2 and multi-jet fraction for different fit ranges.

The results of the fitting procedure are seen to be very similar across all range



(a) χ^2 (b) Contamination

Figure 4.9: Comparison of χ^2 and signal contamination of various template definitions in the muon channel multi-jet estimate.

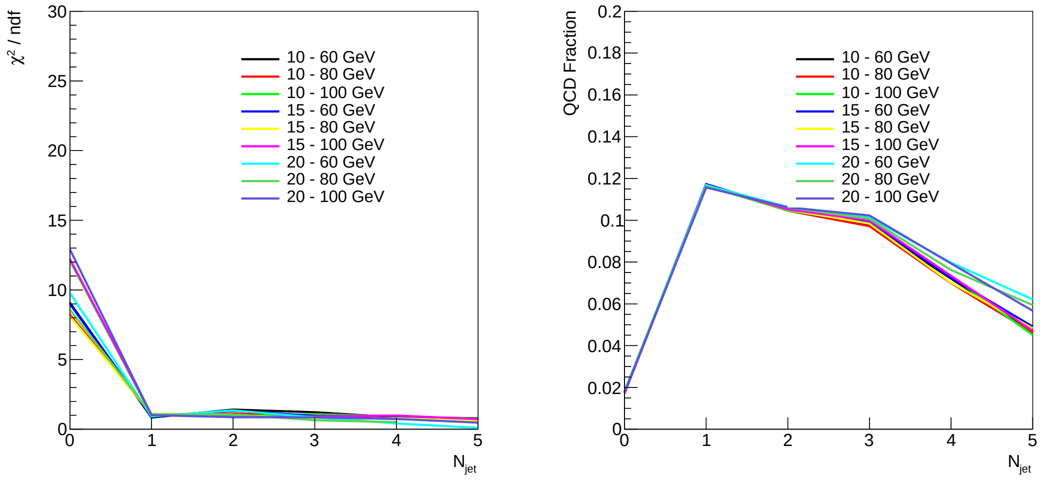
variations except for some small variation in the highest jet bins where the statistical uncertainty is larger. For the nominal fit range the range $15 < E_{\mathrm{T}}^{\mathrm{miss}} < 80$ GeV was chosen to parallel the range used in the electron channel.

Figure 4.11 shows the nominal fit results from the multi-jet estimate. The nominal multi-jet fractions are shown in Table 4.5.

Systematic Uncertainties

The systematic uncertainty on the multi-jet background contribution was estimated using the same variations as used above in the $W \to e\nu$ channel to derive fit range, Monte Carlo template modelling and statistical uncertainties. In addition, a $W \to \mu\nu$ specific multi-jet uncertainty (Template Choice) was derived resulting in four multi-jet uncertainties:

Fit Range The nominal fit range of 15-80 GeV was varied using alternative minima of 10 or 20 GeV and alternative maxima of 60 or 100 GeV to investigate the effect of changing the fit range.



(a) χ^2

(b) Multi-jet Fraction

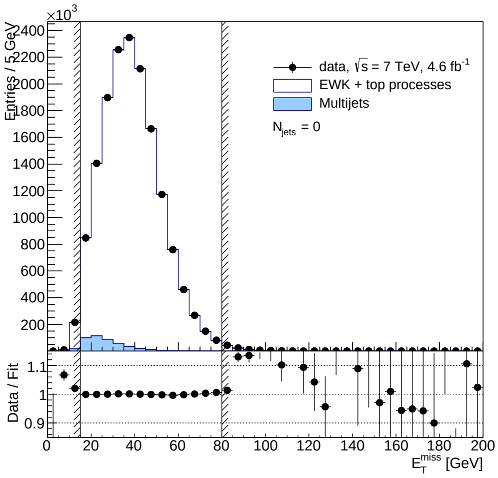
Figure 4.10: Comparison of χ^2 and multi-jet fraction (QCD fraction) of various choices of fit range choice in the muon channel multi-jet estimate.

Monte Carlo Template Modelling The fit was also carried out using SHERPA W and Z Monte Carlo in place of the default ALPGEN samples. The result of this fit was compared to the nominal fit to ascertain the size of the effect of Monte Carlo modelling on the fit results.

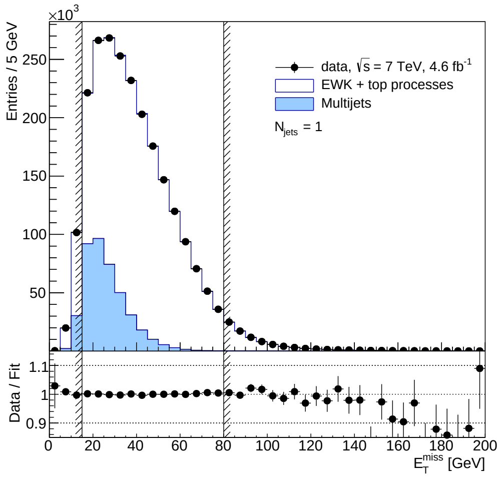
Statistical Uncertainty The statistical uncertainty on the fit result was also included as an uncertainty on the multi-jet background. This is only significant for high jet multiplicity bins.

Template Choice The uncertainty due to the choice of template was estimated by comparison of the nominal template with the AntiD0WithIso_D0Window and RestrAntiIso templates defined above. The total template choice uncertainty was defined as the envelope of these variations.

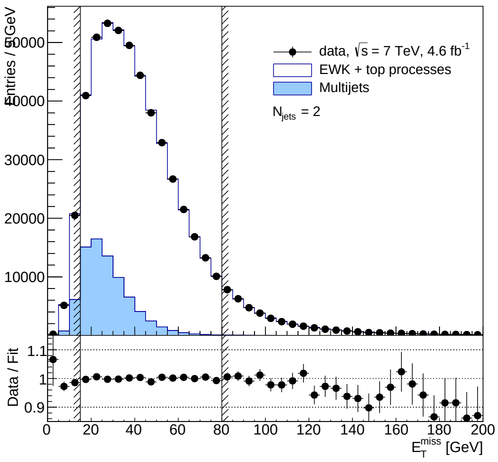
These multi-jet specific uncertainties were combined in quadrature to form the final multi-jet uncertainty. The template choice was also propagated as a shape systematic using the variation of the shape of differential distributions between templates. In addition to these specific uncertainties, the detector level systematics



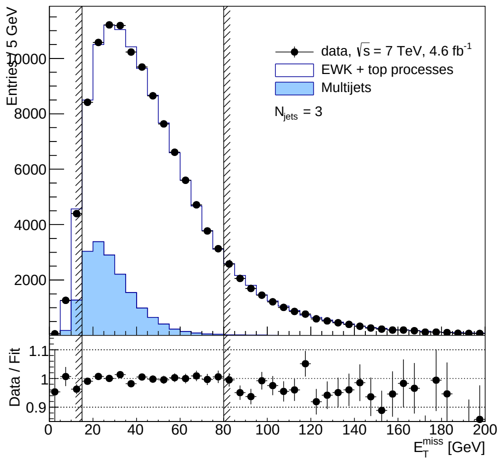
(a) $N_{\text{jet}} = 0$



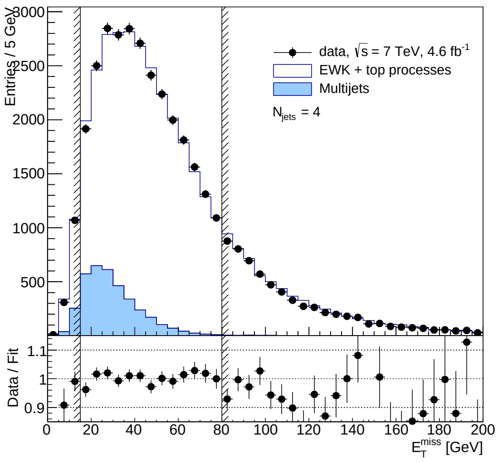
(b) $N_{\text{jet}} = 1$



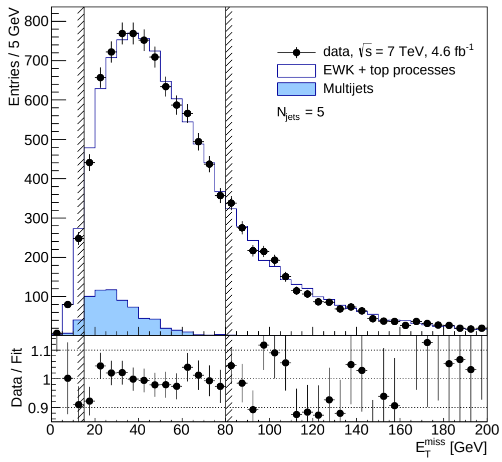
(c) $N_{\text{jet}} = 2$



(d) $N_{\text{jet}} = 3$



(e) $N_{\text{jet}} = 4$



(f) $N_{\text{jet}} = 5$

Figure 4.11: Nominal fits used in the multi-jet estimate in the muon channel. The vertical lines show the range over which the fit is performed.

	$N_{\text{jet}} = 0$	$N_{\text{jet}} = 1$	$N_{\text{jet}} = 2$	$N_{\text{jet}} = 3$	$N_{\text{jet}} = 4$	$N_{\text{jet}} = 5$
Multi-jet fraction	0.0174	0.1168	0.1051	0.0984	0.0701	0.0483
Statistics	0.0186	0.0061	0.0136	0.0281	0.0689	0.1773
Range up	0.0393	0.0069	0.0092	0.0231	0.0863	0.2306
Range down	0.0289	0.0058	0.0049	0.0122	0.0033	0.0428
Modelling up	0.0685	-	-	-	-	-
Modelling down	-	0.0307	0.0116	0.0811	0.0326	0.0313
Template up	0.0314	0.0327	0.0418	0.0405	0.0717	-
Template down	-	-	-	-	-	0.0665
Total up	0.0870	0.0340	0.0449	0.0544	0.1317	0.2909
Total down	0.0344	0.0318	0.0185	0.0867	0.0763	0.1966

Table 4.5: The estimated multi-jet fraction in the $W \to \mu\nu$ channel as a function of the jet multiplicity and the fractional uncertainty on the predicted number of events arising from the various sources discussed in the text.

(to be discussed in Chapter 5) were also propagated fully through the data-driven estimation.

Table 4.5 shows the estimated multi-jet fraction and its associated uncertainty as a function of jet multiplicity.

4.2.4 $Z \to \mu\mu$ channel

The multi-jet background in the $Z \to \mu\mu$ channel was estimated using a template fit using the $m_{\mu\mu}$ distribution. Due to the lower statistics in the Z channel, the multi-jet enriched template was built reversing the muon isolation criteria without requiring a reduced window of d_0 significance requirement similar to what was used in Ref. [135].

The fit was carried out in an extended mass window of [50,140] GeV in exclusive jet multiplicity bins up to 2 jets. Above this, limited statistics result in a poor fit and so, the scaling factor for 2 jets was applied to higher jet bins.

Figure 4.12 shows the nominal fit in the 2 jet bin for the multi-jet estimate in the $Z \to \mu\mu$ channel whilst Table 4.6 shows the multi-jet fractions estimated from

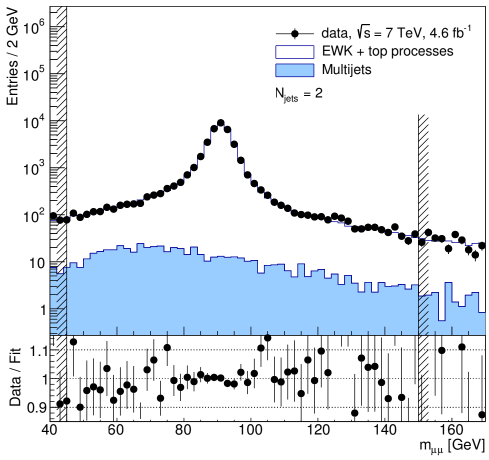


Figure 4.12: Nominal fit used in the $N_{\text{jet}} = 2$ bin of the multi-jet estimate in the $Z \to \mu\mu$ channel. The vertical lines show the range over which the fit is performed.

these fits.

Systematic Uncertainties

As in the electron channel, the systematic uncertainties were estimated by varying fit and selection parameters and comparing shifted results to the nominal:

Fit Range The nominal fit range of [50,140] GeV was varied to minima of 45 and 55 GeV and maxima of 135 and 145 GeV. The variations from the nominal fit result were used to form the envelope of the variations which was taken as the systematic uncertainty.

Template Selection The uncertainty arising from the choice of template was evaluated by comparison to alternative templates with an inverted d_0 significance cut and no isolation requirement and an inverted isolation cut including the d_0 significance requirement. The total uncertainty was taken as the envelope of these variations.

	$N_{\mathrm{jet}} = 0$	$N_{\mathrm{jet}} = 1$	$N_{\mathrm{jet}} = 2$
Multi-jet fraction	0.004	0.005	0.009
Statistics	0.08	0.16	0.20
Range up	-	0.02	0.04
Range down	0.01	0.01	0.01
Template up	0.22	0.04	0.04
Template down	-	0.01	0.08
Modelling up	-	-	-
Modelling down	0.40	0.35	0.60
Total up	0.24	0.17	0.21
Total down	0.41	0.38	0.63

Table 4.6: The estimated multi-jet fraction as a function of jet multiplicity in the $Z \rightarrow \mu\mu$ channel and the fractional error from the various uncertainty sources discussed in the text.

Monte Carlo Template Modelling The fit was also carried out using SHERPA W and Z Monte Carlo in place of the default ALPGEN samples. The result of this fit was compared to the nominal fit to ascertain the size of the effect of Monte Carlo modelling on the fit results.

Statistical Uncertainty The uncertainty on the multi-jet estimate due to statistics was also included as an uncertainty on the multi-jet background.

The total uncertainty on the multi-jet background was formed by summing the above systematic variations in quadrature. Table 4.6 summarises the sizes and combination of these sources to form a total uncertainty on the multi-jet estimate used in the analysis.

In addition to the above multi-jet specific uncertainties, the Monte Carlo based template was varied according to the detector level systematic variations (to be discussed in Chapter 5) so that detector level systematics were propagated fully through the data-driven method.

	$N_{\text{jet}} = 3$	$N_{\text{jet}} = 4$	$N_{\text{jet}} = 5$
Purity ($t\bar{t}$ fraction)	0.61	0.79	0.86
Contamination ($W \to e\nu$ fraction)	0.13	0.07	0.05

Table 4.7: The purity and contamination of the nominal $t\bar{t}$ data-driven template as a function of the exclusive jet multiplicity in the electron channel.

4.3 $t\bar{t}$ backgrounds

For higher jet multiplicities in the $W+$jets channel, the $t\bar{t}$ background becomes important. The background from this source is larger than the signal in the 4 jet bin and above. In the previous $W+$jets measurement [8], Monte Carlo was used to estimate the expected background. However, due to large uncertainties on the top production cross section, the shape of kinematic distributions and the jet energy scale, an uncertainty of almost 40% on the total number of events with 4 jets was found. The data-driven method employed here greatly reduces this background uncertainty.

For the $t\bar{t}$ data-driven estimate, a similar template method was employed as was discussed in Section 4.2 using a template enriched in $t\bar{t}$ events instead of a template enriched in multi-jet background events.

To select a template from data which is enriched in $t\bar{t}$ events, the decay of the top quark was exploited whereby top quarks almost always decay via $t \to Wb$. As such, a top template was extracted from data by requiring at least one b-jet in an event using the MV1 b-tagging algorithm [136] at a 70% efficiency working point. This working point was chosen on the basis of purity (the percentage of the template arising from $t\bar{t}$ events) and contamination (the percentage of the template arising from $W+$jets events). The purity and contamination of the resulting top template is shown in Table 4.7.

As the b-tagged template is not purely composed of $t\bar{t}$ events, the contributions from non-$t\bar{t}$ events (excluding multi-jets) were subtracted from the template using Monte Carlo simulation. The contribution from multi-jet events was subtracted using data-driven estimates, obtaining the normalisation using the same methods as outlined above but with the b-tagging requirement applied to the multi-jet template.

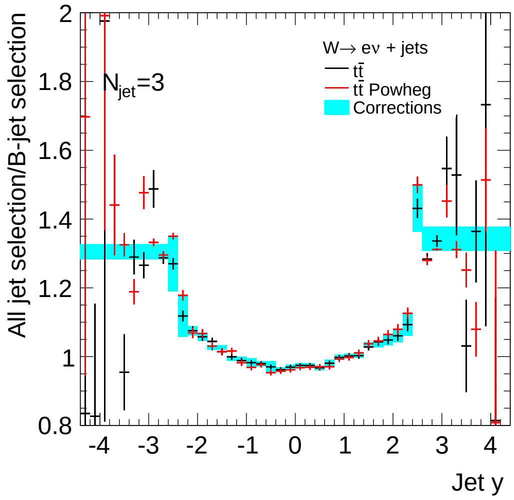
Bias between the b-tagged sample and the inclusive $t\bar{t}$ sample was evaluated in Monte Carlo simulation using POWHEG and ALPGEN+HERWIG samples. The ratio of the number of events with no b-tagging requirement to that with the b-tagging requirement as a function of the rapidity of the leading jet is shown in Figure 4.13. For large rapidities the bias is as large as 30% due to the lack of b-tagging beyond the tracker acceptance. ALPGEN+HERWIG $t\bar{t}$ samples were used to derive a correction factor for the b-tagging induced bias which is applied to the $t\bar{t}$ template extracted from data. The uncertainty on this correction factor was taken as the statistical uncertainty or the difference to the correction factor taken from POWHEG if this difference was statistically significant.

Having derived the $t\bar{t}$ template from data, the normalisation of the control sample must be estimated. This is challenging because, since the top decays via a W, kinematic distributions are often very similar between $t\bar{t}$ and W + jets and fits struggle to converge. A variable which was found to separate the templates is that of transformed aplanarity which is defined as e^{-8A} where A is the aplanarity, defined as 1.5 times the size of the smallest eigenvalue of the momentum tensor [138] given by:

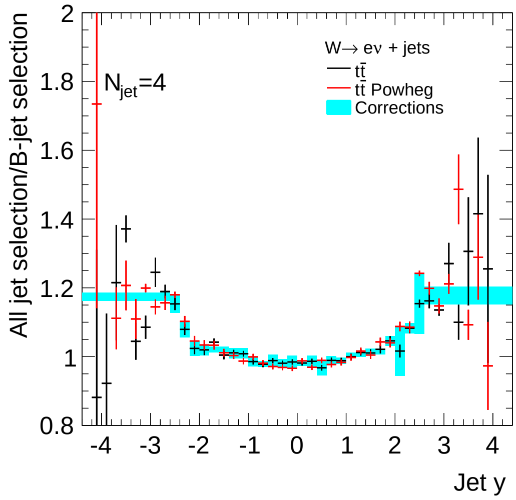
$$M_{xyz} = \sum_i \begin{pmatrix} p_{xi}^2 & p_{xi}p_{yi} & p_{xi}p_{zi} \\ p_{yi}p_{xi} & p_{yi}^2 & p_{yi}p_{zi} \\ p_{zi}p_{xi} & p_{zi}p_{yi} & p_{zi}^2 \end{pmatrix}, \tag{4.2}$$

where the sum runs over all jets and leptons used in the measurement. This variable was previously used in top cross section measurements to separate the W+jets events from $t\bar{t}$ [139].

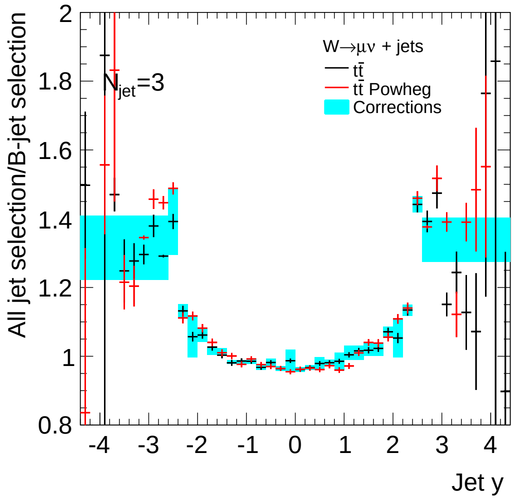
The transformed aplanarity was found to have good separation between W and



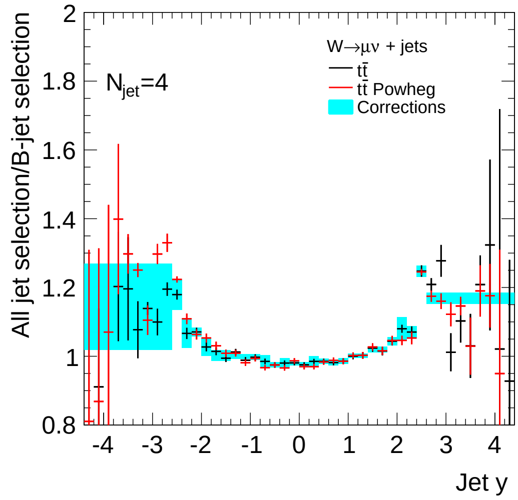
(a) $W \to e\nu$, $N_{\mathrm{jet}} = 3$



(b) $W \to e\nu$, $N_{\mathrm{jet}} = 4$



(c) $W \to \mu\nu$, $N_{\mathrm{jet}} = 3$



(d) $W \to \mu\nu$, $N_{\mathrm{jet}} = 4$

Figure 4.13: b-tagging correction factors applied to the $t\bar{t}$ data-driven templates. Black points show the correction derived from Alpgen+Herwig and red points from Powheg. The blue band shows the final correction factor and uncertainty which is applied. Taken from Ref. [137].

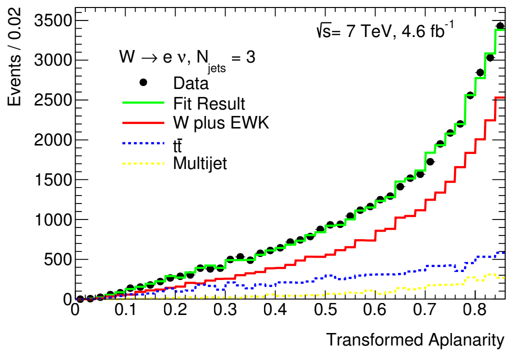
$t\bar{t}$ template shapes and to be stable to variations of the W template from jet energy scale. For values of transformed aplanarity above 0.9, the effect of the jet energy scale uncertainty was found to be large and so, in order to exclude this region, the fit was performed between 0 and 0.85.

The fit was performed using a maximum likelihood fit with three templates in the signal region. The first template was composed of $W+$jets signal and electroweak backgrounds (including single t) and derived from Monte Carlo. The second ($t\bar{t}$) template was taken from data, as described above, following subtraction of multi-jet and b-tagged signal and application of the b-tagging bias corrections. Finally, a multi-jet template was included from the non-b-tagged multi-jet templates with the normalisation fixed to that derived in Section 4.2. Figure 4.14 shows example fitted transposed aplanarity distributions in which it is seen that the data (black points) are well described by the template fit (green line).

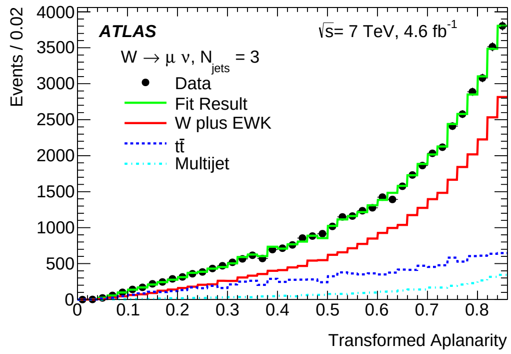
4.3.1 $t\bar{t}$ Uncertainties

In this Section, the uncertainties considered for the data-driven $t\bar{t}$ background are discussed.

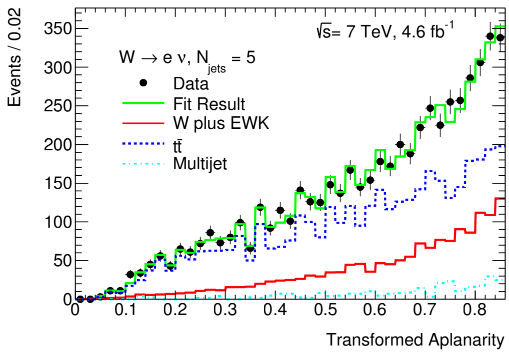
Statistical Uncertainty The statistical uncertainty returned by the maximum likelihood fit is used. There is however a caveat here as the $t\bar{t}$ template and the data are correlated as the former is a subset of the latter. This correlation is somewhat broken by the application of b-tagging and the subtraction of template contamination. This results in an underestimate of the statistical uncertainty by the maximum likelihood fit. To evaluate the size of this, a toy Monte Carlo method was employed by which toy templates and pseudo-data were produced (taking correlations into account) and the fit re-run in each case. The width of the pull distribution in this case gives a true estimate of the statistical uncertainty. Correction factors were applied



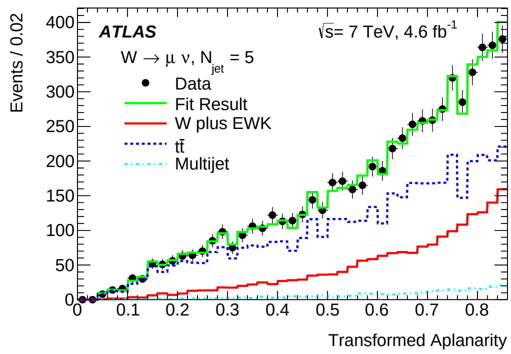
(a) $W \to e\nu$, $N_{\mathrm{jet}} = 3$



(b) $W \to \mu\nu$, $N_{\mathrm{jet}} = 3$



(c) $W \to e\nu$, $N_{\mathrm{jet}} = 5$



(d) $W \to \mu\nu$, $N_{\mathrm{jet}} = 5$

Figure 4.14: Templates and results of the data-driven $t\bar{t}$ fit in the $W \to e\nu$ and $W \to \mu\nu$ channels in the $N_{\mathrm{jet}} = 3$ and $N_{\mathrm{jet}} = 5$ bins. The black points are the data, the red line is the W+jets and electroweak template, the blue line is the $t\bar{t}$ template and the yellow line is the multi-jet template. The uncertainties shown are statistical only. Taken from Ref. [137].

to the errors returned from the fit to inflate the uncertainty accordingly. These vary from a 15% inflation at three jets to a 40% inflation at seven jets.

Model Uncertainty In the lower jet multiplicities there is significant contamination from Wc (W production with one charm quark), Wcc (W production with two charm quarks) and Wbb (W production with two b quarks). ALPGEN simulation was used to subtract these events from the templates, with the heavy flavour distributions normalised according to the theoretical cross section. Comparisons between Wc, Wcc and Wbb shows that the Wcc and Wbb have similar shaped differential distributions whilst Wc has a different shape. To cross-check the normalisations applied, data and Monte Carlo were compared in the 1-jet and 2-jet bins which are dominated by Wc and Wcc/Wbb respectively. These two bins were fit separately by varying the Wc and Wcc/Wbb normalisations. The Wc normalisation was found to be 1.24 (1.27) in the electron (muon) channel and the Wcc/Wbb normalisation was found to be 0.99 (0.93). As a conservative estimate of this normalisation the fits were repeated with these derived normalisation factors applied. The resulting uncertainty is less than 1%.

Fit Range The upper value of the fit range was varied down to 0.83 (from 0.85) and up to 0.87. The variation in the $t\bar{t}$ estimate with this variation was found to be less than 1%.

Uncertainty on the b-tagging Corrections As discussed above, the uncertainty on the bias corrections was taken as the statistical uncertainty on the nominal ALPGEN corrections or the difference between the POWHEG and ALPGEN corrections if statistically significant. The resulting uncertainty is 1-3% on the $t\bar{t}$ estimate.

b-tagging Uncertainties Centrally provided b-tagging uncertainties were applied to the Monte Carlo samples used in the contamination subtraction procedure. These include the variations associated to b-tagging efficiency, c-tagging efficiency and the light-tagging inefficiency. These variations were applied and propagated separately and combined in quadrature. The resulting uncertainty on the $t\bar{t}$ is less than 2%.

Detector Uncertainties Uncertainties relating to jet energy scale, jet energy resolution, lepton energy scale and lepton energy resolution (to be discussed in Chapter 5) were applied to the Monte Carlo templates used in the fit procedure. These uncertainties were propagated separately through the analysis so that they could be correctly correlated in the final results.

Signal Template Shape The difference between using a $W+$jets template derived from Alpgen and Sherpa was investigated. It was found that the change in signal modelling resulted in large variations in the number of predicted $t\bar{t}$ events (up to 30%). The reason for this was investigated and found to arise from the poor statistical sample available for Sherpa and the Sherpa Monte Carlo mis-modelling the transformed aplanarity in general. It was found that the shape differences between Alpgen and Sherpa were within the systematic uncertainty bands already defined and so, a comparison to Sherpa was not included as an additional systematic.

4.4 Multi-jet and $t\bar{t}$ Correlations

As described above, the multi-jet estimate was carried out using $t\bar{t}$ Monte Carlo followed by the $t\bar{t}$ estimate being carried out using this multi-jet template. In order to understand if this procedure resulted in correlations between the two backgrounds, an iterative cross-check was performed. Following the $t\bar{t}$ estimation described above, the data-driven $t\bar{t}$ estimate was used in a second multi-jet fit. This second multi-jet fit was then used to rederive the $t\bar{t}$ estimate.

In the case of the $t\bar{t}$ estimate, comparing the first iteration with the second iteration showed that there was no significant effect on the estimate in any of the jet bins. Similarly, comparison of the multi-jet estimates derived from the fit using $t\bar{t}$ Monte Carlo and that using data-driven $t\bar{t}$ showed no significant difference. As such, the procedure outlined above (using Monte Carlo $t\bar{t}$ in the multi-jet estimate) was used.

Chapter 5

Detector Level Uncertainties and Results

In this Chapter, the final steps of the analysis to produce results at detector level are discussed. In Section 5.1 the uncertainties associated with jets, electrons, muons and $E_{\mathrm{T}}^{\mathrm{miss}}$ are discussed and then, in Section 5.2, distributions are shown at detector level comparing Monte Carlo and data-driven based predictions with data.

5.1 Detector Level Uncertainties

In this Section, an account of all the systematics uncertainties derived at detector level is given. First, common (electron and muon) jet systematics are detailed, followed by channel specific systematics and finally common $E_{\mathrm{T}}^{\mathrm{miss}}$ specific uncertainties. All the systematic shifts were applied to Monte Carlo predictions of signal and background processes. In the case of the data-driven backgrounds, as discussed in Chapter 4, fits were repeated using the systematically shifted Monte Carlo templates and the variations properly correlated to the associated systematic variation of Monte Carlo estimates.

5.1.1 Jet Energy Scale

Five groups of centrally provided uncertainties were considered for the Jet Energy Scale (JES) uncertainty [116]:

In-situ calibration uncertainties are those propagated from the in-situ γ+jet, Z+jet, dijet and multi-jet studies which were used to derive the residual calibration and to validate the jet energy calibration (see Section 3.4.2). This class of uncertainty is provided in the form of 57 nuisance parameters, each of which corresponds to a possible systematic shift of the jet energy scale. One such nuisance parameter deals with the uncertainty on the dijet inter-calibration which calibrates forward jets ($|\eta| > 0.8$) to central ($|\eta| < 0.8$) jets. This uncertainty is propagated by coherently shifting the p_{T} and E of all the jets up and down by an amount defined as a function of p_{T} and $|\eta|$,

$$E \rightarrow (1 + \alpha\,(p_{\mathrm{T}}, |\eta|))\,E, \tag{5.1}$$

$$p_{\mathrm{T}} \rightarrow (1 + \alpha\,(p_{\mathrm{T}}, |\eta|))\,p_{\mathrm{T}}, \tag{5.2}$$

where α defines the size and direction of the nuisance parameter.

55 nuisance parameters arise from the γ+jet, Z+jet and multi-jet related uncertainties. The last in-situ nuisance parameter arises from the uncertainty on single hadron test-beam studies [140] which are the only way to understand high p_{T} uncertainties (above the reach of other in-situ methods). These 56 uncertainties were propagated by coherently shifting the p_{T} and E of all the jets up and down by an amount defined as a function of p_{T},

$$E \rightarrow (1 + \beta_i\,(p_{\mathrm{T}}))\,E, \tag{5.3}$$

$$p_{\mathrm{T}} \rightarrow (1 + \beta_i\,(p_{\mathrm{T}}))\, p_{\mathrm{T}}, \tag{5.4}$$

where β_i defines the size and direction of the i-th nuisance parameter.

Non-closure uncertainties cover the differences arising because the jet calibration was derived in a different Monte Carlo generation campaign than those samples used in the analysis. Small differences were made in the simulation between such campaigns and so, some small differences in the simulated jet response were observed. These uncertainties were formulated identically to the in-situ uncertainties of Equations 5.1 and 5.2.

Pile-up based uncertainties are those defined by the uncertainty on the correction applied to jets to remove the effects of pile-up. These uncertainties were propagated by coherently shifting the p_{T} and E of all the jets up and down by an amount defined as a function of p_{T}, $|\eta|$, μ and N_{PV}:

$$E \rightarrow (1 + \gamma_i\,(p_{\mathrm{T}}, |\eta|, \mu, N_{PV}))\, E, \tag{5.5}$$

$$p_{\mathrm{T}} \rightarrow (1 + \gamma_i\,(p_{\mathrm{T}}, |\eta|, \mu, N_{PV}))\, p_{\mathrm{T}}, \tag{5.6}$$

where γ_i defines the size and direction of the i-th uncertainty term.

Close-by uncertainties account for the effect of nearby jets on the response and therefore calibration of jets. They were propagated by coherently shifting the p_{T} and E of all the jets up and down by an amount defined as a function of p_{T} and the ΔR to the closest jet,

$$E \rightarrow (1 + \delta_i\,(p_{\mathrm{T}}, \min(\Delta R_{jj})))\, E, \tag{5.7}$$

$$p_{\mathrm{T}} \rightarrow \left(1 + \delta_i \left(p_{\mathrm{T}}, \min(\Delta R_{jj})\right)\right) p_{\mathrm{T}}, \tag{5.8}$$

where δ_i defines the size and direction of the i-th nuisance parameter.

Flavour uncertainties account for the different responses to gluon-initiated and quark-initiated jets[1] as observed in Monte Carlo. Ideally, one would apply different calibration factors dependent on flavour but there is no definitive way to separate gluon and quark jets in data so identical calibration factors must be employed. One uncertainty accounts for the difference in the flavour response of jets and a second uncertainty accounts for the flavour composition of the signal. These uncertainties are formulated identically to the in-situ uncertainties of Equations 5.1 and 5.2 and the nominal flavour composition was derived from Alpgen Monte Carlo. The uncertainty on the composition was derived by comparison to the flavour composition derived from Sherpa Monte Carlo. An additional b-tagged jet energy scale uncertainty was also included which was applied to the b-tagged jets used in the $t\bar{t}$ data-driven estimation.

In order to reduce the number of JES uncertainty nuisance parameters which must be propagated through the analysis, the covariance matrix of 55 of the 57 in-situ calibration uncertainties was diagonalised and reduced to 6 nuisance parameters (β_i'), resulting in a total of 8 nuisance parameters describing the in-situ based uncertainties (the 6 from the reduction and the η-intercalibration and single-hadron uncertainties not included in the reduction). Therefore, a total of 15 components was used to propagate the JES uncertainties:

- 6 from the reduction of the nuisance parameters (β_i'-like)
- 1 from the uncertainty of the η-intercalibration of jets (α-like)
- 1 from the uncertainty on single-hadron response measurements used for high

[1]defined in Monte Carlo by geometric matching of the jet to partons in the truth record.

p_{T} jets (β_i-like)

- 1 from the different response of jets in separate generation campaigns (α-like)
- 2 from the uncertainty on terms used in the pile-up correction of jets (γ_i-like)
- 1 from the effect of close-by jets (δ_i-like)
- 3 from the flavour response and composition uncertainties (α-like)

These 15 systematic shifts were propagated through the analysis and then combined in quadrature to ascertain the total JES uncertainty.

5.1.2 Jet Energy Resolution

The Jet Energy Resolution (JER) has been studied in data and Monte Carlo using two distinct in-situ methods [141]. It was shown that the JER is well modelled in Monte Carlo and so, unlike in the electron or muon cases, no correction was applied to Monte Carlo for the JER. In order to propagate the level of knowledge of the JER, as found in Ref. [141], jets in Monte Carlo were smeared by an additional amount given by the uncertainty of the JER measurements. Jets were smeared 5 times using a different random seed and these 5 alternatives were propagated through the analysis. The total JER uncertainty was then taken as the mean of the effect of the 5 variations.

5.1.3 Electron Uncertainties

Energy Calibration

Uncertainties on the electron energy scale derived using in-situ Z tag-and-probe studies [95] were provided with the calibration described in Section 3.2.4. They were separated into p_{T} and η dependent terms relating to:

- Statistical uncertainty of the in-situ calibration
- Calibration method uncertainty
- Uncertainty from the generator used in the calibration
- Uncertainty on the material budget assumed in the calibration
- Uncertainty on the use of the presampler to correct for energy deposited in front of the calorimeter

These terms were propagated separately and combined in quadrature.

The electron energy resolution uncertainty (also from tag-and-probe) was propagated by varying the smearing applied to Monte Carlo upwards and downwards according to the uncertainty on the electron energy resolution measurement. The total uncertainty was assigned as the averaged combination of these two variations.

Efficiency Scale Factors

Uncertainties on reconstruction, identification and trigger scale factors were derived centrally [101]. Reconstruction and identification scale factor uncertainties were separated into one uncorrelated term and 6 correlated terms by a similar covariance matrix diagonalisation method as was employed to reduce the terms in the JES uncertainty in Section 5.1.1. The dominant systematics were found to arise from the uncorrelated terms and so, ideally a toy Monte Carlo method would be used to propagate this uncertainty instead of the naive up and down variations used here. However, since such uncertainties are subdominant in jet based measurements, the up and down variations were used as a reasonable estimate and to reduce computation time. Trigger scale factor uncertainties were propagated using simple up and down variations.

5.1.4 Muon Uncertainties

Momentum Calibration

Systematics on the muon momentum scale and resolution derived using tag and probe [108] were provided for the calibration procedure described in 3.3.4. These were propagated using three p_{T} and η dependent components relating to:

- The uncertainty on the total p_{T} scale from the in-situ calibration
- The uncertainty on the muon p_{T} arising from the resolution of the inner detector (ID)
- The uncertainty on the muon p_{T} arising from the resolution of the muon spectrometer (MS)

These components were propagated separately and combined in quadrature to give the total uncertainty from the muon p_{T} calibration.

Efficiency Scale Factors

Uncertainties were also provided centrally for the reconstruction and efficiency scale factors applied. For the reconstruction efficiency, a single uncertainty was propagated through the analysis using simple up and down variations of one standard deviation. For the trigger efficiency, a statistical uncertainty was propagated using up and down variation of the scale factor and a systematic uncertainty was propagated via the inclusion of an additional ϕ dependence of the trigger scale factor which was not included in the nominal scale factor.

5.1.5 Missing Transverse Energy

The missing transverse energy was reconstructed by combining all hard reconstructed objects (electrons, muons, jets etc.) with any soft calorimeter deposits not already included in the hard objects, as outlined in Section 3.5. Any systematic shifts in the hard objects, as discussed above, were therefore propagated through to the measurement of the missing transverse energy by rebuilding the missing transverse energy using the systematically varied objects. Any remaining uncertainty arises from the calibration of the soft terms.

First, the uncertainty on the scale of the soft terms was derived by varying the scale separately up and down according to centrally defined uncertainties [117]. Secondly, the effect of the soft term resolution uncertainty was propagated by smearing the soft terms. All these variations were propagated through the analysis and then combined in quadrature.

5.2 Results and distributions at detector level

The detector level uncertainties discussed above were propagated through the analysis in parallel to the background uncertainties arising from the data driven methods for multi-jet and $t\bar{t}$ backgrounds (discussed in Sections 4.2 and 4.3 respectively) and the normalisation uncertainties on the Monte Carlo backgrounds (discussed in Section 4.1.1).

A previous analysis [135] demonstrated that the ALPGEN samples shown here predict a boson rapidity distribution which is too wide. This results in a mismodelling of the lepton phase space selection. The agreement is improved by reweighting the ALPGEN sample to CT10 PDFs. As this acceptance correction is dependent on the boson p_{T}, an acceptance correction was applied dependent on

Number of partons	$W \to e\nu$	$Z \to ee$	$W \to \mu\nu$	$Z \to \mu\mu$
0	1.0423	1.0676	1.0446	1.0693
1	1.0386	1.0600	1.0413	1.0618
2	1.0285	1.0441	1.0303	1.0452
3	1.0210	1.0323	1.0225	1.0334
4	1.0171	1.0262	1.0185	1.0276
5	1.0145	1.0243	1.0154	1.0234

Table 5.1: Acceptance corrections applied to Alpgen samples to correct for the mis-modelled boson rapidity.

the number of partons in the matrix element, as shown in Table 5.1. This acceptance correction was only applied to the detector level results as this is the only part of the analysis where such factors are relevant.

Tables 5.2 and 5.3 show the sizes of the signal and background contributions at detector level for the electron and muon channels respectively. It is observed that the $Z \to ee$ and $Z \to \mu\mu$ selections are significantly more pure than the $W \to e\nu$ and $W \to \mu\nu$ selections. This can be explained by the dilepton requirement and invariant mass cut, which helps to reduce the background significantly. The importance of an accurate $t\bar{t}$ background prediction in the $W \to e\nu$ and $W \to \mu\nu$ channels at high multiplicity is also evident given that the $t\bar{t}$ background becomes larger than the signal above a jet multiplicity of 4.

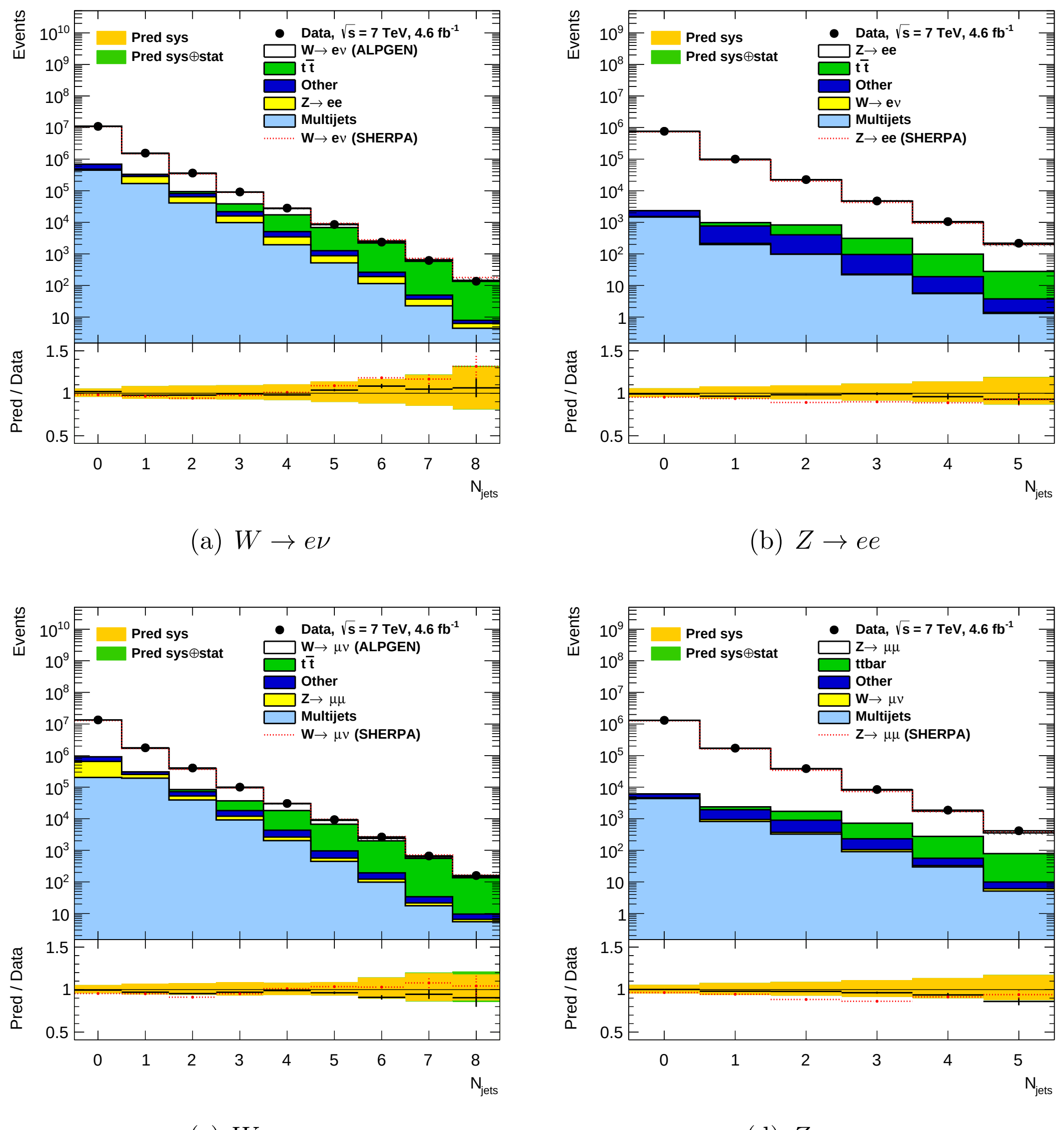
Figures 5.1, 5.2, 5.3 and 5.4 show the detector level distributions for exclusive jet multiplicity, leading jet p_{T}, leading jet rapidity and H_{T}. In all distributions the predicted and measured distributions agree well. $V+$jets signal samples generated with Alpgen rather than Sherpa generally describe the data better.

N_{jet}	0	1	2	3	4	5	6	7
				$W \to e\nu$				
$W \to e\nu$	94%	78%	73%	58%	37%	23%	14%	11%
$Z \to ee$	$< 1\%$	8%	7%	7%	5%	4%	3%	3%
Multi-jet	4%	11%	12%	11%	7%	6%	5%	4%
$t\bar{t}$	$< 1\%$	$< 1\%$	3%	18%	46%	62%	76%	80%
Single top	$< 1\%$	$< 1\%$	2%	3%	4%	3%	2%	2%
τ, diboson	2%	3%	3%	3%	2%	1%	1%	1%
Total Predicted	11 100 000	1 510 000	354 000	89 500	28 200	8550	2530	572
	± 640 000	± 99 000	± 23 000	± 5600	± 1400	± 440	± 200	± 61
Data Observed	10 878 398	1 548 000	361 957	91 212	28 076	8514	2358	618
				$Z \to ee$				
$Z \to ee$	100%	99%	96%	93%	90%	-	-	-
$W \to e\nu$	$< 0.1\%$	$< 0.1\%$	$< 0.1\%$	$< 0.1\%$	$< 0.1\%$	-	-	-
Multi-jet	0.2%	0.2%	0.4%	0.5%	0.5%	-	-	-
$t\bar{t}$	$< 0.1\%$	0.2%	2%	5%	8%	-	-	-
Single top	$< 0.1\%$	$< 0.1\%$	0.1%	0.2%	0.1%	-	-	-
τ, diboson	0.1%	0.5%	1%	1%	1%	-	-	-
Total Predicted	754 000	96 500	22 100	4700	1010	-	-	-
	± 47 000	± 6900	± 1700	± 390	± 93	-	-	-
Data Observed	761 280	99 991	22 471	4729	1050	-	-	-

Table 5.2: The approximate size of the signal and backgrounds, expressed as a percentage of the total number of predicted events. They are derived from either data-driven estimates or simulations for exclusive jet multiplicities for the $W \to e\nu$ selection (upper table) and for the $Z \to ee$ selection (lower table).

N_{jet}	0	1	2	3	4	5	6	7
				$W \to \mu\nu$				
$W \to \mu\nu$	93%	82%	78%	62%	40%	25%	17%	11%
$Z \to \mu\mu$	3%	4%	3%	3%	2%	1%	1%	1%
Multi-jet	2%	11%	10%	9%	7%	5%	4%	3%
$t\bar{t}$	$< 1\%$	$< 1\%$	3%	19%	46%	64%	75%	83%
Single top	$< 1\%$	$< 1\%$	2%	3%	4%	3%	2%	2%
τ, diboson	2%	3%	3%	3%	2%	1%	1%	$< 1\%$
Total Predicted	13 300 000	1 710 000	384 000	96 700	30 100	8990	2400	627
	± 770 000	± 100 000	± 24 000	± 6100	± 1600	± 480	± 180	± 66
Data Observed	13 414 400	1 758 239	403 146	99 749	30 400	9325	2637	663
				$Z \to \mu\mu$				
$Z \to \mu\mu$	100%	99%	96%	91%	84%	-	-	-
$W \to \mu\nu$	$< 0.1\%$	0.1%	0.1%	0.2%	0.2%	-	-	-
Multi-jet	0.3%	0.5%	0.9%	1%	2%	-	-	-
$t\bar{t}$	$< 0.1\%$	0.3%	2%	6%	13%	-	-	-
Single top	$< 0.1\%$	$< 0.1\%$	0.1%	0.2%	0.2%	-	-	-
τ, diboson	0.1%	0.5%	1%	1%	1%	-	-	-
Total Predicted	1 300 000	168 000	37 800	8100	1750	-	-	-
	± 79 000	± 12 000	± 2800	± 660	± 160	-	-	-
Data Observed	1 302 010	171 200	38 618	8397	1864	-	-	-

Table 5.3: The approximate size of the signal and backgrounds, expressed as a percentage of the total number of predicted events. They are derived from either data-driven estimates or simulations for exclusive jet multiplicities for the $W \to \mu\nu$ selection (upper table) and for the $Z \to \mu\mu$ selection (lower table).



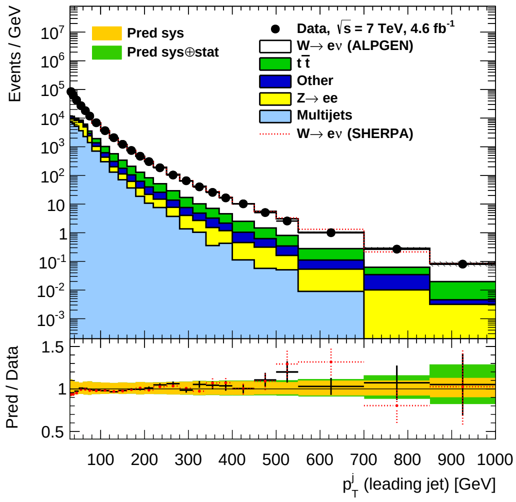
(a) $W \to e\nu$

(b) $Z \to ee$

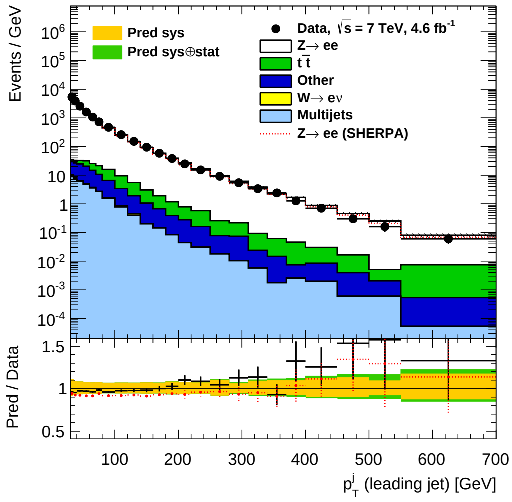
(c) $W \to \mu\nu$

(d) $Z \to \mu\mu$

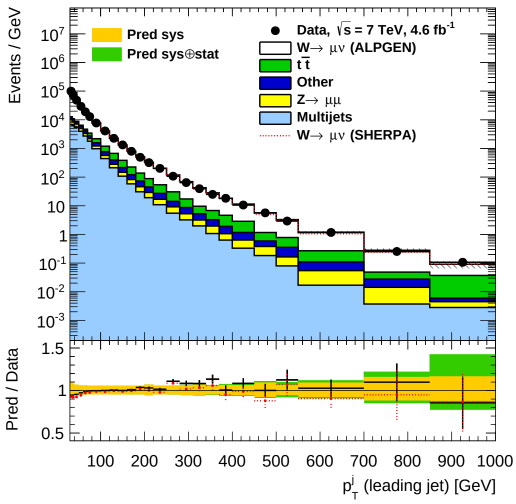
Figure 5.1: Distribution of events passing the $W \to e\nu$, $Z \to ee$, $W \to \mu\nu$ and $Z \to \mu\mu$ selection as a function of the exclusive jet multiplicity. Predictions from data-driven methods and Monte Carlo simulation are shown. Simulation of signal is shown using Alpgen (black) or Sherpa (red). The experimental systematic uncertainties are shown by the yellow (inner) band and the combined statistical and systematic uncertainties are shown by the green (outer) band.



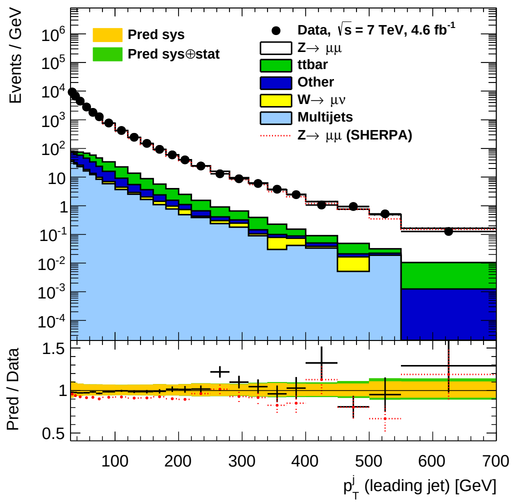
(a) $W \rightarrow e\nu$



(b) $Z \rightarrow ee$

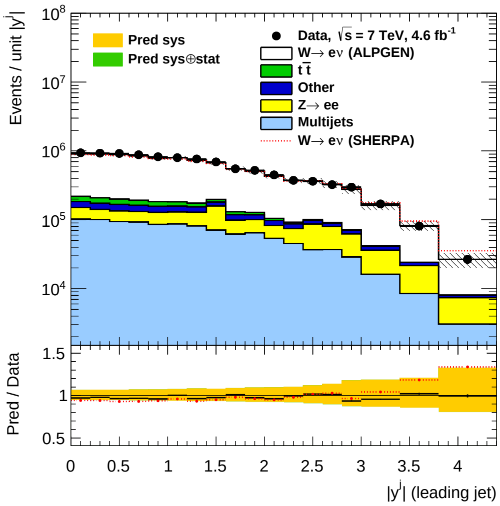


(c) $W \rightarrow \mu\nu$

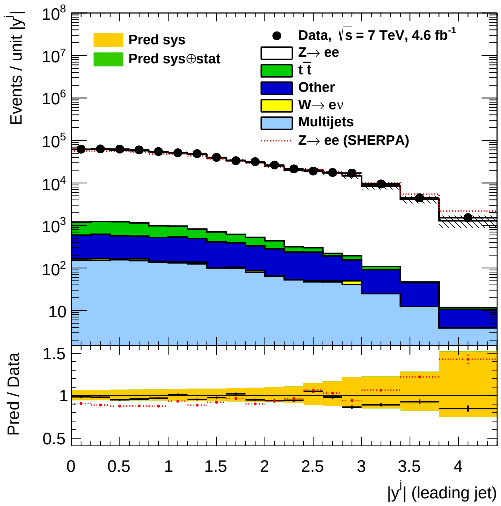


(d) $Z \rightarrow \mu\mu$

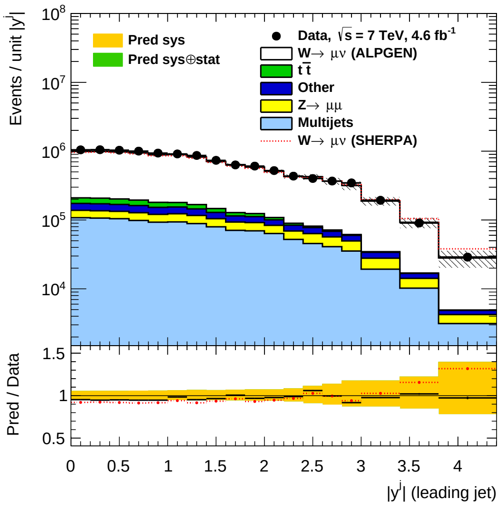
Figure 5.2: Distribution of events passing the $W \rightarrow e\nu$, $Z \rightarrow ee$, $W \rightarrow \mu\nu$ and $Z \rightarrow \mu\mu$ selection as a function of the leading jet p_{T}. Predictions from data-driven methods and Monte Carlo simulation are shown. Simulation of signal is shown using ALPGEN (black) or SHERPA (red). The experimental systematic uncertainties are shown by the yellow (inner) band and the combined statistical and systematic uncertainties are shown by the green (outer) band.



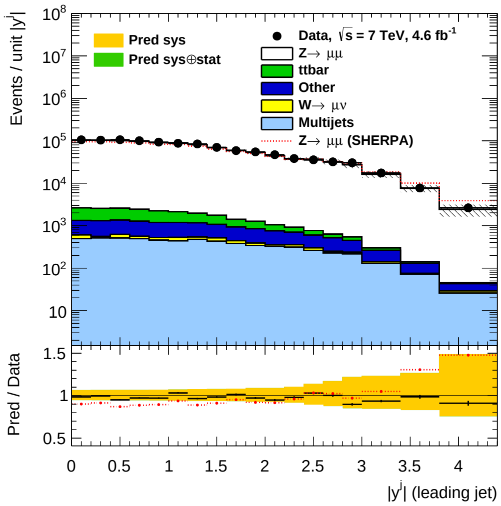
(a) $W \to e\nu$



(b) $Z \to ee$

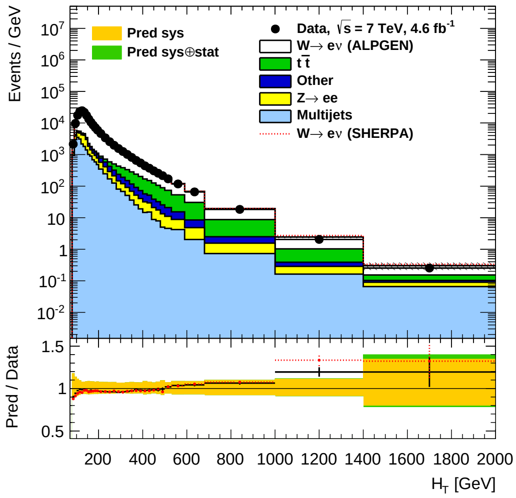


(c) $W \to \mu\nu$

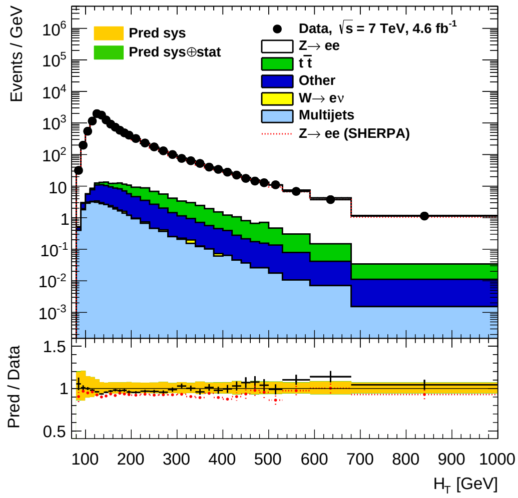


(d) $Z \to \mu\mu$

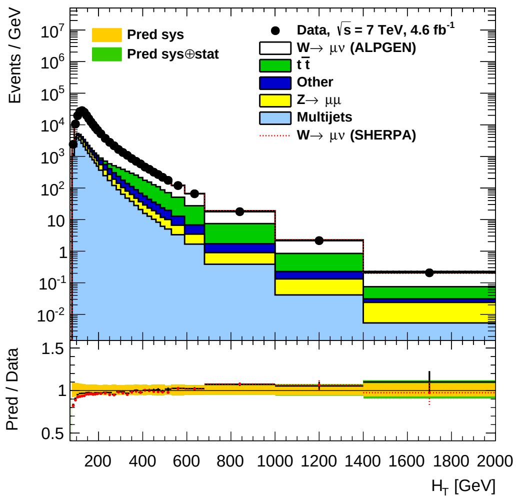
Figure 5.3: Distribution of events passing the $W \to e\nu$, $Z \to ee$, $W \to \mu\nu$ and $Z \to \mu\mu$ selection as a function of the leading jet rapidity. Predictions from data-driven methods and Monte Carlo simulation are shown. Simulation of signal is shown using ALPGEN (black) or SHERPA (red). The experimental systematic uncertainties are shown by the yellow (inner) band and the combined statistical and systematic uncertainties are shown by the green (outer) band.



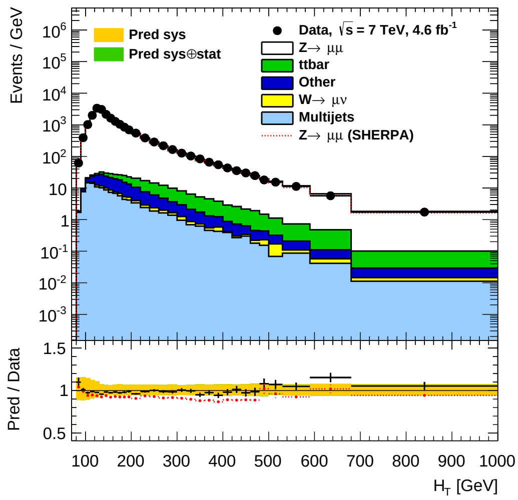
(a) $W \to e\nu$



(b) $Z \to ee$



(c) $W \to \mu\nu$



(d) $Z \to \mu\mu$

Figure 5.4: Distribution of events passing the $W \to e\nu$, $Z \to ee$, $W \to \mu\nu$ and $Z \to \mu\mu$ selection as a function of the scalar sum p_{T} of all identified objects, H_{T}. Predictions from data-driven methods and Monte Carlo simulation are shown. Simulation of signal is shown using ALPGEN (black) or SHERPA (red). The experimental systematic uncertainties are shown by the yellow (inner) band and the combined statistical and systematic uncertainties are shown by the green (outer) band.

Chapter 6

Measurement and Theory Corrections

A number of corrections were applied to the measurement and theory predictions to bring them to consistent phase space and object definitions. Throughout this procedure, the particle level ("truth") phase space definition of the W boson, Z boson and jets is the same as that at reconstructed level, as outlined in Table 6.1.

Lepton kinematic variables were calculated at truth level using final state leptons from the W or Z decay. Photons radiated by the boson decay products, within a cone of $\Delta R = 0.1$ of the lepton, were added to the lepton four momentum. This procedure is known as "dressing" and the resulting leptons, "dressed". Particle level jets were reconstructed using the same anti-k_t algorithm with a radius parameter of 0.4 as described in Section 3.4. All final state particles with a lifetime greater than 30 ps were included in the jet clustering, except for neutrinos, electrons or muons from boson decays and the photons included in the dressing. Finally, the missing transverse energy was calculated using the four momentum of the neutrino from the decay of the W boson.

Firstly, the measurements were unfolded to the fiducial space of the electron and muon channels separately, as defined in Table 6.2. This procedure will be discussed in Section 6.1.

Next, these measurements were corrected to the common phase space, also de-

W boson
$E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV $m_{\mathrm{T}} > 40$ GeV

Z boson
$66 < m_{ll} < 116$ GeV $\Delta R(l, l) > 0.2$

Jet Selection
$p_{\mathrm{T}} > 30$ GeV $\lvert y \rvert < 4.4$ $\Delta R(\mathrm{jet}, l) > 0.5$

Table 6.1: Truth boson and jet definitions.

Electron Channel
$p_{\mathrm{T}}^{e} > 25$ GeV $\lvert \eta^{e} \rvert < 2.47$ (excl. $1.37 < \lvert \eta^{e} \rvert < 1.52$)

Muon Channel
$p_{\mathrm{T}}^{\mu} > 25$ GeV $\lvert \eta^{\mu} \rvert < 2.4$

Common Channel
$p_{\mathrm{T}}^{l} > 25$ GeV $\lvert \eta^{l} \rvert < 2.5$

Table 6.2: Truth phase space definition in the electron and muon phase spaces.

fined in Table 6.2, using corrections derived in Section 6.3. At this stage, the electron and muon measurements were compared and combined as outlined in Section 6.4.

Theory predictions were provided in the common phase space for comparison to the combined measurement, however certain effects were not included in the theory calculations and so, were applied in the form of bin-by-bin corrections. These include corrections for QED radiation from leptons (FSR), discussed in Section 6.5, and the non-perturbative (NP) corrections from Section 6.6.

6.1 Corrections for detector effects

Signal event yields at detector level were first determined by subtracting the estimated background contributions of Chapter 4 from the data. The resulting distributions still include detector effects such as selection efficiency, resolution effects and residual mis-calibrations.

The aim is to perform a measurement which is independent of detector effects and so, can be compared to theory predictions without having to pass theory predictions though a complex and computationally intensive detector simulation. The process by which the detector level distributions, shown in Chapter 5, were corrected is called unfolding. The individual channels were unfolded to particle level in the phase spaces defined in Table 6.2.

The ratio measurement at particle level was obtained by first unfolding the W and Z measurements separately and then forming the ratio of the particle level distributions, treating the uncertainties as correlated or uncorrelated (Monte Carlo and data statistical uncertainties and data-driven background uncertainties) where appropriate.

6.1.1 Unfolding method

As a first approximation, the unfolding could have been carried out using a simple bin-by-bin correction whereby the correction was formed from Monte Carlo using:

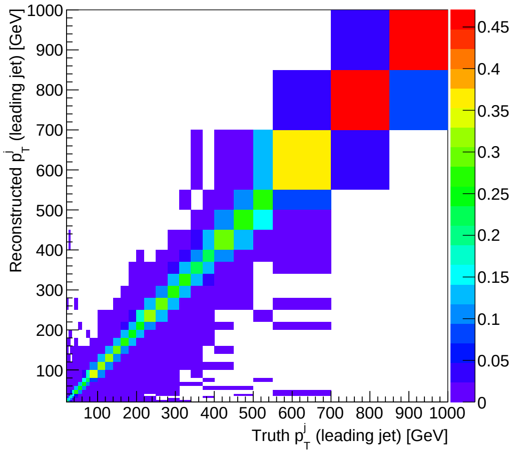
$$C = \frac{N_{\text{true}}}{N_{\text{reco}}} \tag{6.1}$$

where N_{true} is the number of events in a given bin at particle (truth) level and N_{reco} is the number of events at detector level. Such an approach does not account for migrations between bins in the distributions and so, is only valid if migrations are small.

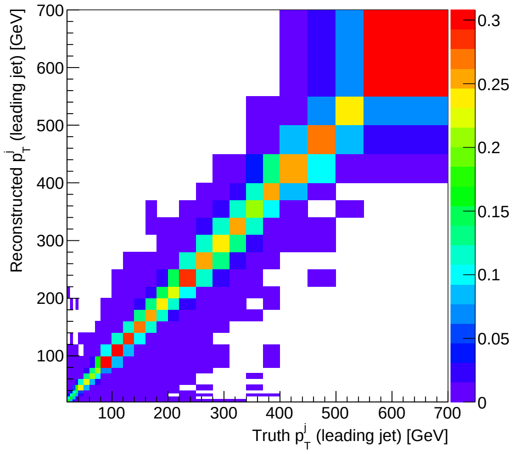
An extension of a bin-by-bin unfolding procedure would be to use matrices whereby the matrix is formed to describe the relationship between particle level and detector level quantities. This matrix can then be inverted and used to carry out the unfolding correction on the measured detector level distributions. Such a method has limitations of being strongly dependent on the prior distribution in Monte Carlo and can be unstable in the presence of statistical fluctuations.

To include the effects of migrations, an iterative Bayesian unfolding technique [142] was employed, as implemented in the RooUnfold package [143]. This method treats the response matrices as describing the probability of observed data given the true distribution. The unfolding is iterated such that the first iteration uses the Monte Carlo truth distribution as a prior whilst subsequent iterations use the result from the previous iteration as a prior. In this way, by increasing the number of iterations the dependence on the prior is reduced but the statistical uncertainty increases.

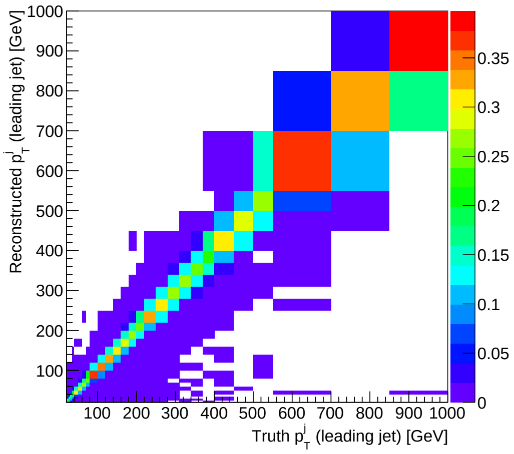
Response matrices, as shown in Figure 6.1, were filled using events which pass both the detector and particle level selections using Alpgen+Herwig Monte Carlo samples. The detector level values are shown on the y-axis whilst particle level values



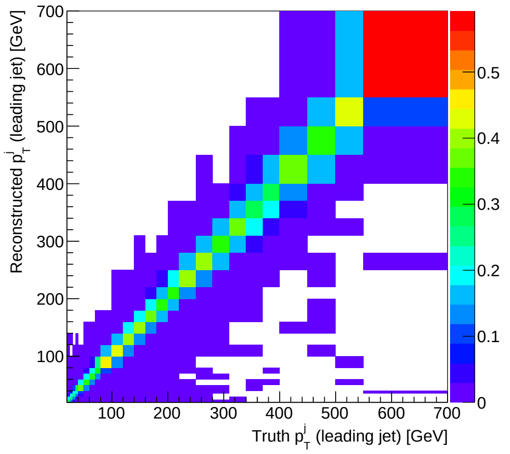
(a) $W \to e\nu$



(b) $Z \to ee$



(c) $W \to \mu\nu$



(d) $Z \to \mu\mu$

Figure 6.1: Example response matrices showing reconstructed level variables versus truth level variables. Each column is normalised to the number of events in the truth level bin so that the sum of each column gives the reconstruction efficiency.

are shown on the x-axis. Each column is normalised to the total number of truth events falling in that bin so that the sum across each column corresponds to the reconstruction efficiency.

In order to include migrations across the jet p_{T} threshold of 30 GeV, for example due to a 30 GeV jet being reconstructed as a 29 GeV jet at detector level, the jet p_{T} cut was reduced to 20 GeV (at both reconstructed and truth level) when unfolding jet p_{T} distributions. The inclusive jet multiplicity distribution was obtained by summing up the unfolded exclusive jet multiplicity distribution.

Two further corrections also had to be applied to correctly unfold these measurements. Firstly, events may pass the detector level selection but not pass the particle level selection. This "fake" correction was defined as a bin-by-bin correction (binned in reconstructed variables) given by

$$C_{\text{fake}} = \frac{P(\text{reco \& truth})}{P(\text{reco})}, \tag{6.2}$$

where $P(\text{reco \& truth})$ is the probability for an event to pass both detector level and particle level selection and $P(\text{reco})$ is the probability for an event to pass detector level selection. Figure 6.2 shows the fake corrections as a function of leading jet p_{T} which are applied before unfolding by RooUnfold.

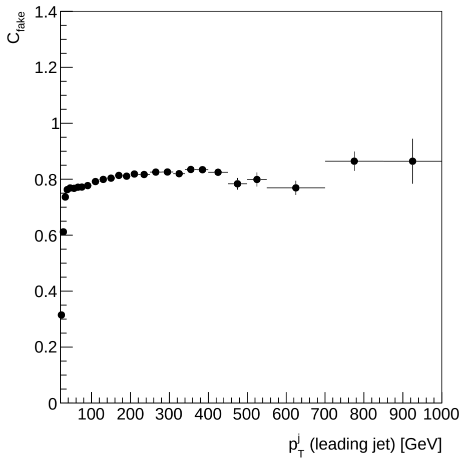
Secondly, events may pass the particle level selection but not the detector level selection. This "efficiency" correction was defined as a bin-by-bin correction (binned in truth variables) given by

$$C_{\text{eff}} = \frac{P(\text{reco \& truth})}{P(\text{truth})}, \tag{6.3}$$

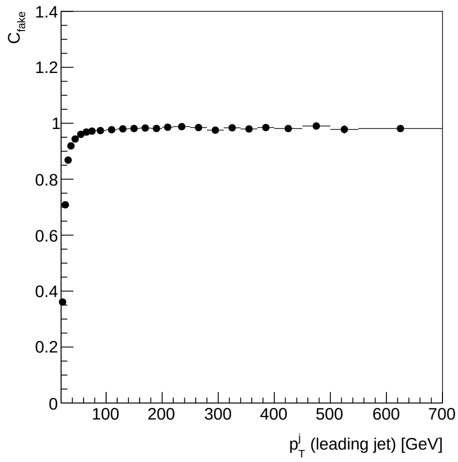
where $P(\text{reco \& truth})$ is the probability for an event to pass both detector level and particle level selection and $P(\text{truth})$ is the probability for an event to pass particle level selection. The efficiency corrections were applied internally in RooUnfold during the unfolding.

6.1.2 Closure tests

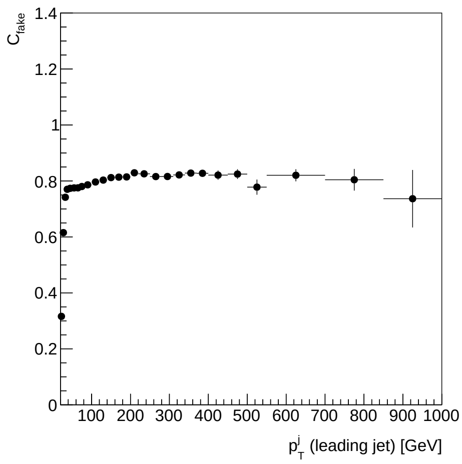
To validate the unfolding method, two closure tests were performed. Firstly, the detector level distributions from Alpgen+Herwig samples were used as pseudo-data and the unfolding was performed using corrections and matrices derived from the same sample. Secondly, the same corrections and matrices were used to unfold



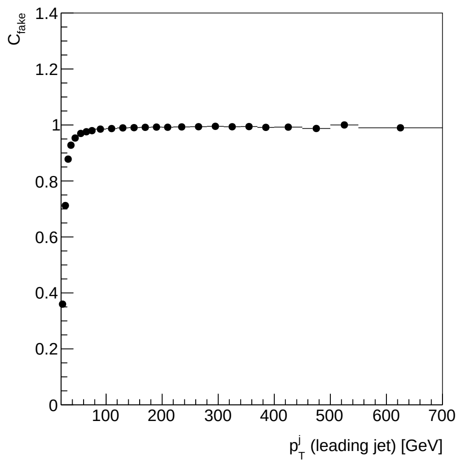
(a) $W \to e\nu$



(b) $Z \to ee$



(c) $W \to \mu\nu$



(d) $Z \to \mu\mu$

Figure 6.2: Example fake corrections as a function of leading jet p_{T}.

the SHERPA detector level distributions.

The results of the first of these tests are shown in Figure 6.3. It is observed that the agreement between unfolded ALPGEN and truth ALPGEN is perfect, as would be expected given that the sample emulating data and the sample used to unfold are exactly the same.

The results of the second test are shown in Figure 6.4 where good closure is observed when unfolding SHERPA with ALPGEN. Some small differences are observed and the effect of subsequent iterations inflating statistical fluctuations can also be

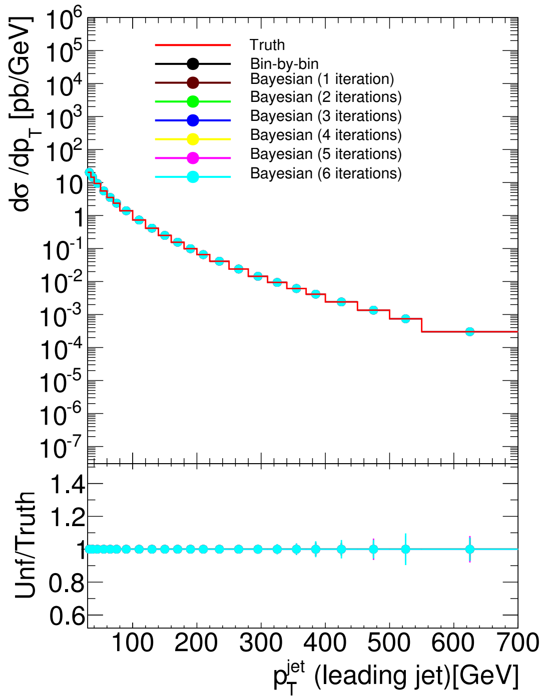


Figure 6.3: Unfolding closure test unfolding ALPGEN with ALPGEN in the $W \to \mu\nu$ channel. Different unfolding options of bin-by-bin and Bayesian iterative with 1-6 iterations are shown. Note that all the points lie on top of one another and demonstrate that the unfolding is behaving perfectly. Taken from Ref. [137].

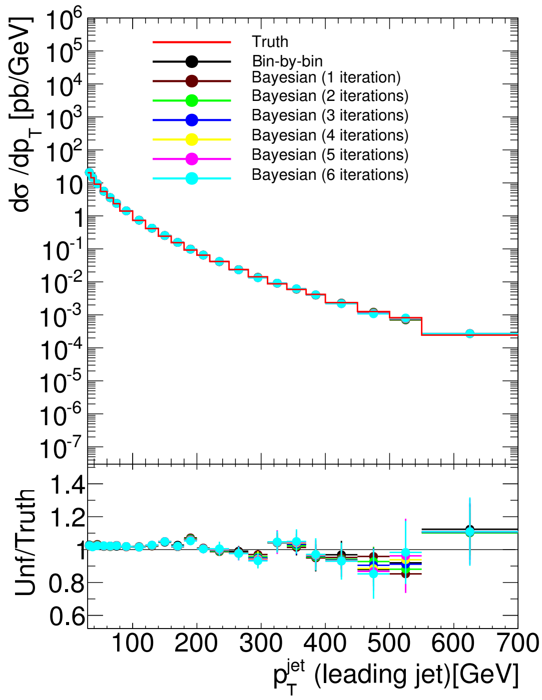


Figure 6.4: Unfolding closure test unfolding SHERPA with ALPGEN in the $W \to \mu\nu$ channel. Different unfolding options of bin-by-bin and Bayesian iterative with 1-6 iterations are shown. Taken from Ref. [137].

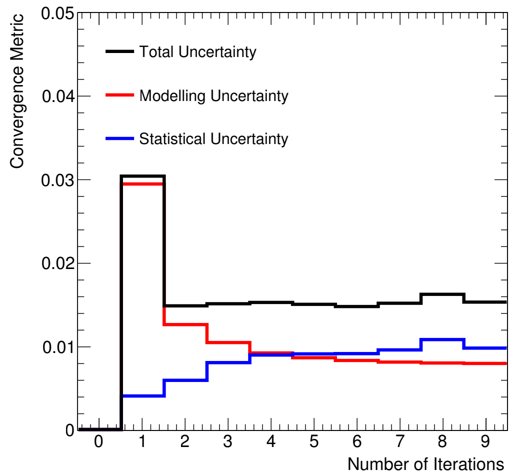


Figure 6.5: Convergence metric (total unfolding uncertainty averaged over all bins) against number of iterations in the $W \to \mu\nu$ channel for exclusive jet multiplicity. The mean unfolding modelling uncertainty (red) and the mean unfolding statistical uncertainty (blue) are shown along with the mean total unfolding uncertainty (black). Taken from Ref. [137].

seen.

Such closure tests could have been used to define the optimum number of iterations to be used in the unfolding procedure, however they were inconclusive. Instead, a convergence metric was used to optimise the number of iterations.

The total uncertainty on the unfolding procedure is discussed below in Section 6.1.3 and is a combination of a statistical and a modelling term. The modelling term refers to the prior used and so, generally reduces with an increase in the number of iterations whilst the statistical term increases. In order to define the optimum number of iterations, the unfolding was repeated for different numbers of iterations and the number of iterations for which the total unfolding uncertainty (averaged across the bins of a distribution) is smallest was found.

Figure 6.5 shows the change in the mean unfolding uncertainty (convergence metric) with the number of iterations. For the distribution shown, jet multiplicity, the minimum uncertainty is observed to be for 2 iterations. Generally the optimum number of iterations was found to be between 1 and 3 and as such, the final choice of the number of iterations was chosen to be 2 across all distributions and channels.

6.1.3 Uncertainties

Uncertainties on the unfolding procedure arise primarily from two sources; the modelling of observables in Monte Carlo and the statistical fluctuations of the Monte Carlo samples. These were evaluated as described below.

Modelling uncertainties were estimated using ALPGEN samples with varied generation parameters. In particular, variations of the minimum parton p_{T} (changed from 15 to 25 GeV), the parton-jet matching cone size used in the MLM matching scheme [47] (changed from 0.7 to 0.4), the functional form of the QCD scales (changed from[1] $\sqrt{m_W^2 + \sum m_{\mathrm{T}}^2}$ to $\sqrt{m_W^2 + \sum p_{\mathrm{T}}^2}$ where m_W is the mass of the W and m_{T} is the transverse mass of the hard partons) and the size of the renormalisation/factorisation scales (doubled and halved from the nominal). These samples were produced without the full detector simulation and were used to reweight the nominal ALPGEN samples dependent on jet multiplicity and boson p_{T} at truth level. These reweighted samples were then used to perform the unfolding and the difference between these results and the nominal were used to assign the modelling systematic.

Statistical uncertainties arising from the limited Monte Carlo samples used to form the unfolding matrices were propagated using toy Monte Carlo. The unfolding inputs were fluctuated using Gaussian distributions with width given by the statistical error in a given bin. 100 toys were created in this way and the unfolding was repeated for each. The RMS of the distribution of results was then taken as the statistical uncertainty.

[1] the sum runs over all hard partons in the matrix element.

6.2 Total Experimental Uncertainty

After the unfolding, the total uncertainty on the measurement in the electron and muon channels is given by the quadratic sum of:

- The Monte Carlo normalisation uncertainties discussed in Section 4.1.1
- The multi-jet background uncertainties discussed in Section 4.2
- The $t\bar{t}$ background uncertainties discussed in Section 4.3.1
- The detector level uncertainties discussed in Chapter 5
- The unfolding systematics discussed in Section 6.1.3

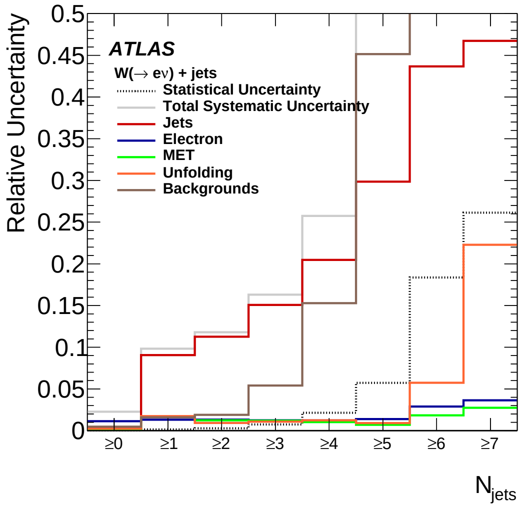
Figure 6.6 shows the various uncertainty sources as a function of the exclusive jet multiplicity for the $W +$ jets and R_{jets} measurements in the electron and muon channels[2]. Figures 6.7 and 6.8 show the uncertainty as a function of leading jet p_{T} and rapidity respectively.

At low jet multiplicity in the $W +$ jets measurement, it is observed that the jet uncertainties from jet energy scale and jet energy resolution dominate the measurement uncertainty. At higher jet multiplicity the background uncertainty begins to dominate as a result of the increasing importance of the $t\bar{t}$ background. It is also observed that the jet uncertainty is slightly larger in the electron channel than the muon channel. This effect arises due to the $Z \to ee$ background being more dominant in the $W \to e\nu$ channel than the $Z \to \mu\mu$ in the $W \to \mu\nu$ background. This difference is driven by the increased likelihood of electrons to fake jets and one of the electrons to be mis-identified as a jet. The cancellation of the jet uncertainties in the R_{jets} measurement is very good as expected and is more significant in the muon channel (due to the larger jet systematics in $W \to e\nu$ compared to $W \to \mu\nu$).

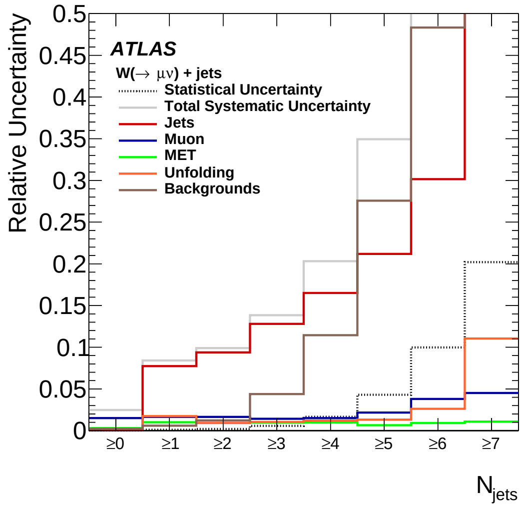
[2]Recall that R_{jets} is the measurement of the ratio of W+jets to Z/γ^*+jets defined in Equation 1.

As a function of jet p_{T}, the $W+$jets uncertainty is again driven predominantly by the jet uncertainties except at higher jet p_{T} where the unfolding and background uncertainties become more important. The bump observed in the jet uncertainty at 80 GeV in the electron channel arises from the uncertainty on the $Z \to ee$ background, as discussed above. Again the cancellation in the R_{jets} measurement is very good however above 250 GeV the statistical uncertainty becomes the dominant uncertainty and the measurement is statistically limited. This statistical uncertainty is primarily driven by the Z/γ^*+jets part of the ratio which has an order of magnitude fewer events.

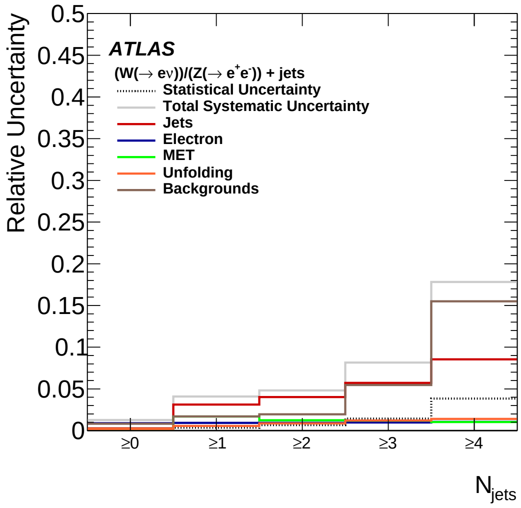
Finally, as a function of leading jet rapidity the $W+$jets measurement is dominated by the jet energy scale uncertainty across the range. The cancellation of systematics observed in the R_{jets} measurement is particularly impressive in this distribution with uncertainties of 38% in the most forward region cancelling to 4.5% in the ratio.



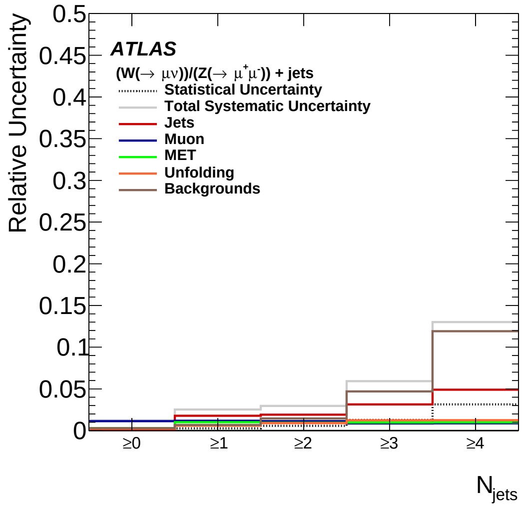
(a) $W \to e\nu$



(b) $W \to \mu\nu$

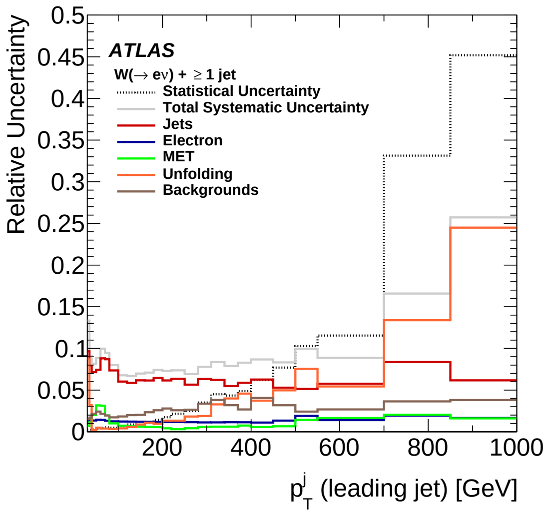


(c) R_{jets} (electron)

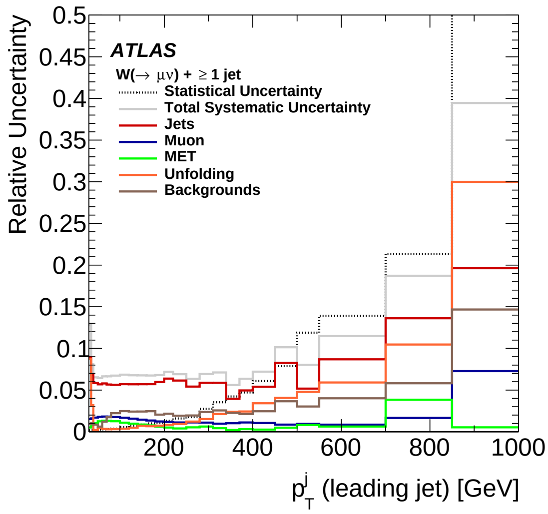


(d) R_{jets} (muon)

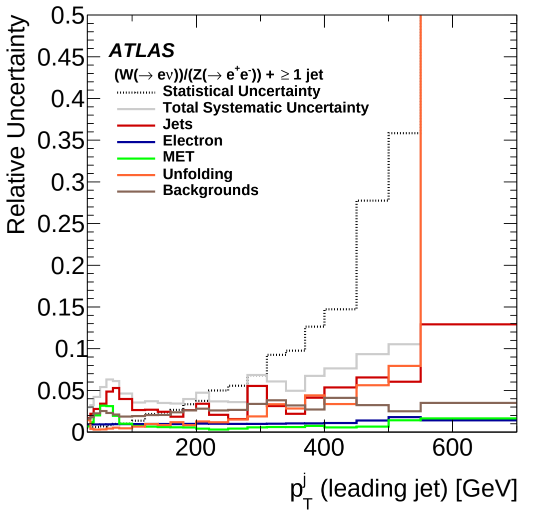
Figure 6.6: The fractional (relative) uncertainty on the $W+$jets and R_{jets} measurements in the electron and muon channels as a function of the inclusive jet multiplicity. The "Jets" category shows the combined effect of JES and JER uncertainties whilst the "Backgrounds" category includes background Monte Carlo uncertainties and the uncertainties arising from data-driven backgrounds.



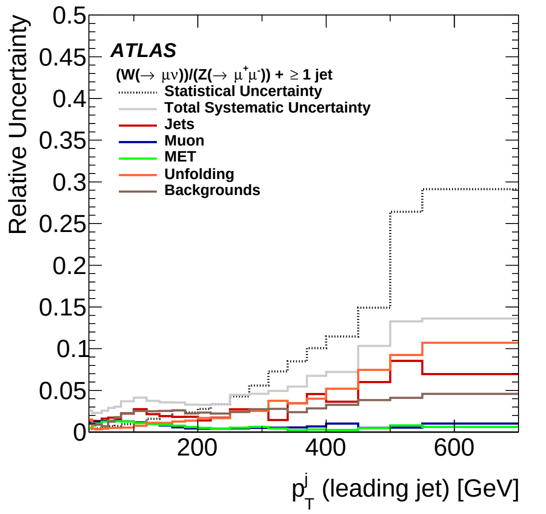
(a) $W \to e\nu$



(b) $W \to \mu\nu$

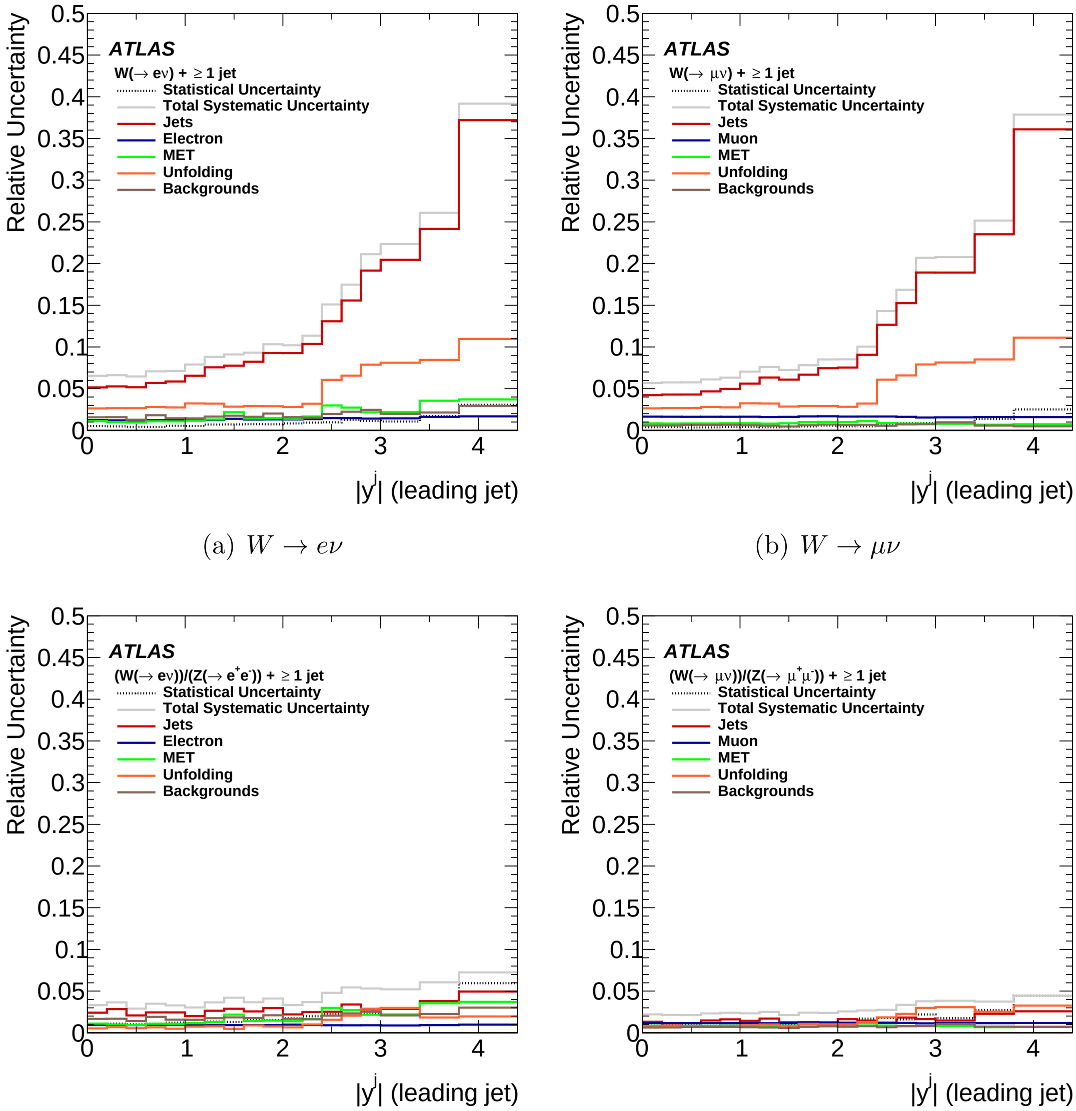


(c) R_{jets} (electron)



(d) R_{jets} (muon)

Figure 6.7: The fractional (relative) uncertainty on the $W+$jets and R_{jets} measurements in the electron and muon channels as a function of the leading jet p_{T} in events with at least one jet. The "Jets" category shows the combined effect of JES and JER uncertainties whilst the "Backgrounds" category includes background Monte Carlo uncertainties and the uncertainties arising from data-driven backgrounds.



(a) $W \to e\nu$

(b) $W \to \mu\nu$

(c) R_{jets} (electron)

(d) R_{jets} (muon)

Figure 6.8: The fractional (relative) uncertainty on the $W+$jets and R_{jets} measurements in the electron and muon channels as a function of the leading jet rapidity in events with at least one jet. The "Jets" category shows the combined effect of JES and JER uncertainties whilst the "Backgrounds" category includes background Monte Carlo uncertainties and the uncertainties arising from data-driven backgrounds.

6.3 Acceptance corrections

The unfolded results for electron and muon channels were measured in the channel dependent phase spaces, defined in Table 6.2. In order to compare and combine the two lepton channels, they were first extrapolated to the common phase space defined in Table 6.2. Furthermore, theory predictions were only provided in the common phase space and so, for separate comparison to electron and muon unfolded results[3], these must be extrapolated to specific lepton phase spaces.

Both these extrapolations were carried out by application of a bin-by-bin correction computed using ALPGEN+HERWIG Monte Carlo samples. These corrections were defined at particle level and given by:

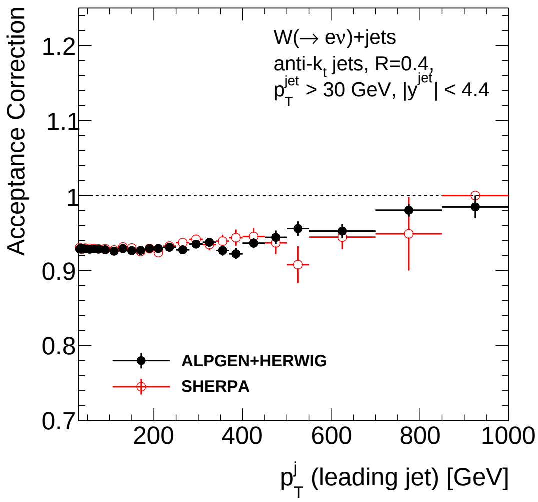
$$\delta^{\text{acc.}} = \frac{\text{lepton phase space}}{\text{common phase space}}. \tag{6.4}$$

This correction is the correction applied for correcting theory to lepton phase spaces, the inverse of this was used to correct unfolded results to the common phase space.

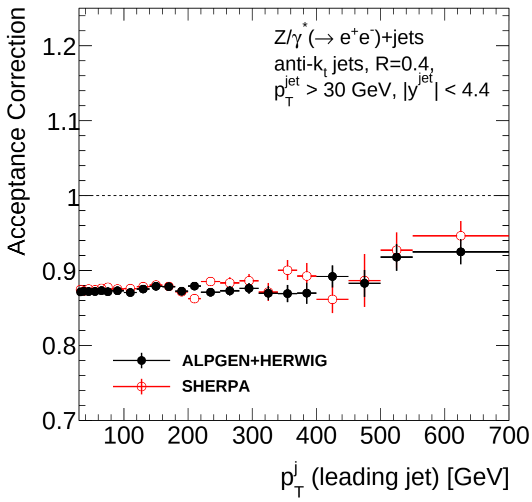
The systematic uncertainty on these corrections was taken as the difference between the nominal correction and the correction derived using SHERPA 1.4 and was taken to be symmetric around the nominal value.

Figures 6.9 and 6.10 show the acceptance corrections for electron and muon channels respectively.

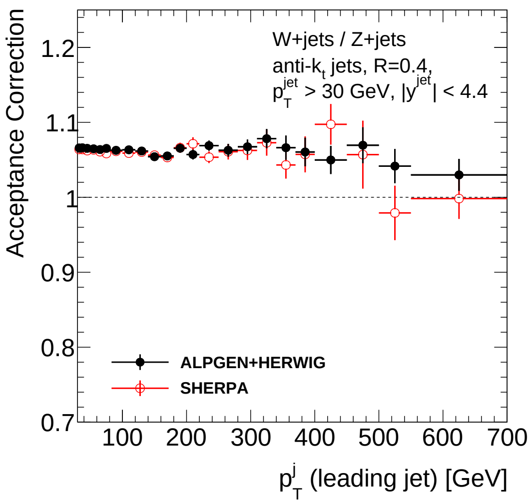
[3]not shown in this thesis.



(a) W+jets

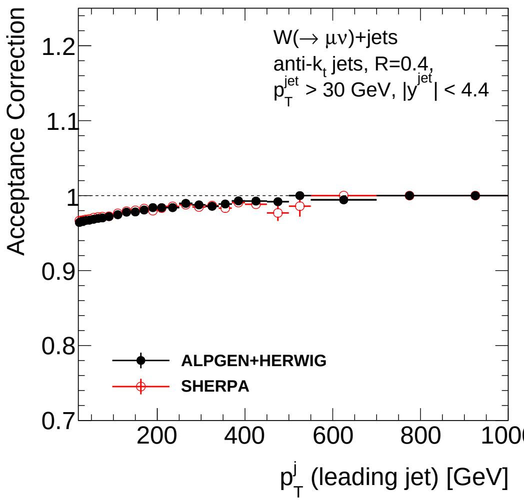


(b) Z+jets

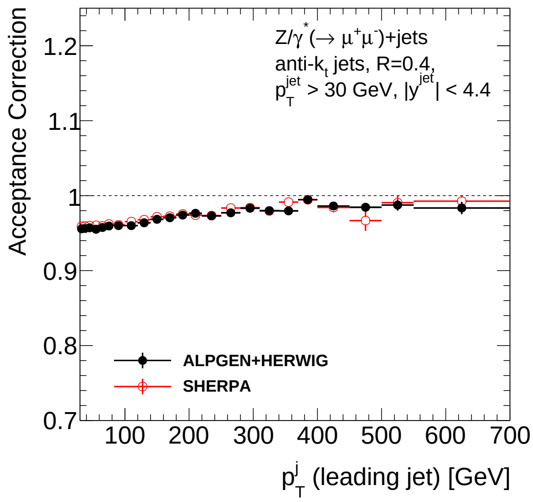


(c) R_{jets}

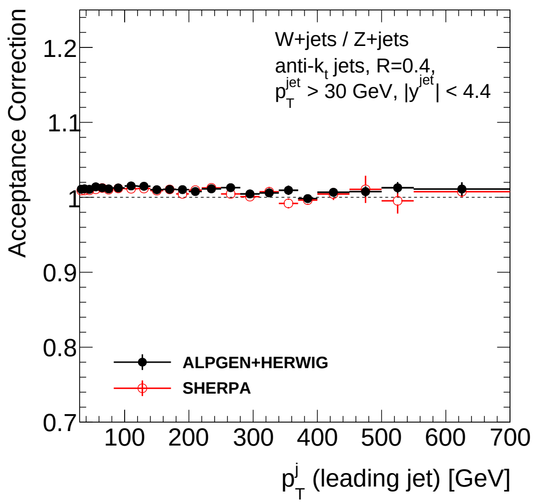
Figure 6.9: The acceptance correction as a function of leading jet p_{T} from Alpgen+Herwig and Sherpa for the electron channel.



(a) W+jets



(b) Z+jets



(c) R_{jets}

Figure 6.10: The acceptance correction as a function of leading jet p_{T} from Alpgen+Herwig and Sherpa for the muon channel.

6.4 Combination of electron and muon channels

The cross sections measured in the electron and muon channels can be combined in order to benefit from a decrease in statistical fluctuations and a cross-calibration of systematic errors. In order to carry out such a combination, both channels were first extrapolated to the common phase space using the acceptance corrections, discussed in Section 6.3.

The combination was carried out using the averaging procedure described in Ref. [144, 145]. Each distribution was combined separately using a minimisation of the χ^2 function:

$$\chi^2 = \sum_{k,i} \frac{[m^i - \sum_j \gamma^i_{j,k} m^i b_j - \mu^i_k]^2}{(\delta^i_{\mathrm{stat},k})^2 \mu^i_k (m^i - \sum_j \gamma^i_{j,k} m^i b_j) + (\delta^i_{\mathrm{uncor},k} m^i)^2} + \sum_j b_j^2, \tag{6.5}$$

where the sums run over all measurement sets k (electrons and muons) and points i (bins) considered. The χ^2 function depends on the combined measurements m^i and the shifts b_j of the correlated (bin-to-bin) systematic error sources, where b_j is expressed in units of standard deviations. The deviations of m^i from the original measurements μ^i_k are minimised. The correlated error sources can also shift ($b_j \neq 0$), and such shifts incur a χ^2 penalty of b_j^2. The relative statistical uncertainties of a specific measurement are labelled $\delta^i_{\mathrm{stat},k}$ and are assumed to be proportional to the inverse of the square root of the number of events. The relative uncorrelated systematic uncertainties on the measurement i in the experimental data set k are given by $\delta^i_{\mathrm{uncor},k}$. The relative correlated systematic uncertainties are given by $\gamma^i_{j,k}$, which quantifies the influence of the correlated systematic error source j on the measurement i in the experimental data set k.

Due to statistical fluctuations inherent in the large set of systematics propagated through this analysis, systematic terms were grouped into single uncertainties for use in the combination procedure. The treatment of these systematic groupings is summarised in Table 6.3 and detailed below.

Statistical uncertainties on the data and Monte Carlo sets used were treated as statistical uncertainties on the measurements.

Lepton uncertainties were treated as correlated bin-to-bin but include no correlation between electron and muon channels as there is no cross-talk of the uncertainties.

Jet and missing energy uncertainties were treated as correlated between channels and bins but were not included in the combination procedure as they were observed to destabilise the combination result. The uncertainty on the combined result due to these sources was propagated by taking a weighted quadratic average of the uncertainty with weights given by:

$$w = \frac{1}{\sigma^2_{\text{stat}} + \sigma^2_{\text{uncor}}} \tag{6.6}$$

Multi-jet background uncertainties were included in the combination as uncorrelated uncertainties. By their nature, these uncertainties have uncorrelated and correlated parts which cannot be separated simply. It was found that including the multi-jet uncertainties as correlated causes instabilities in the combination result. In order to stabilise such fluctuations, the (small) multi-jet uncertainties were treated as uncorrelated in the combination.

$t\bar{t}$ **background statistical** uncertainties were treated as uncorrelated between bins and channels and were included in the combination.

$t\bar{t}$ **background systematic** uncertainties were treated as correlated between bins and channels and were included in the combination.

Other background uncertainties including background estimation of single top, diboson and $W \to \tau\nu$ and $Z \to \tau\tau$ production were included in the combination and treated as correlated between bins and between channels.

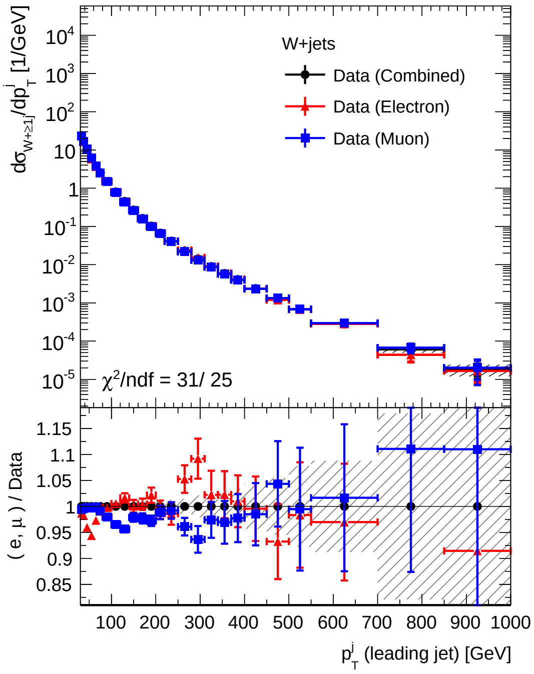
Cross-talk uncertainties such as the background of $W \to e\nu$ events in the $Z \to ee$ event selection were treated in the same way as jet and missing energy uncertainties whereby the uncertainties were not included in the combination but were included on the final result using the weighted average with weights defined as in Equation 6.6.

Unfolding uncertainties were included in the combination. The statistical component of the unfolding systematic was included as uncorrelated whilst the unfolding modelling uncertainty was treated as fully correlated between bins and channels.

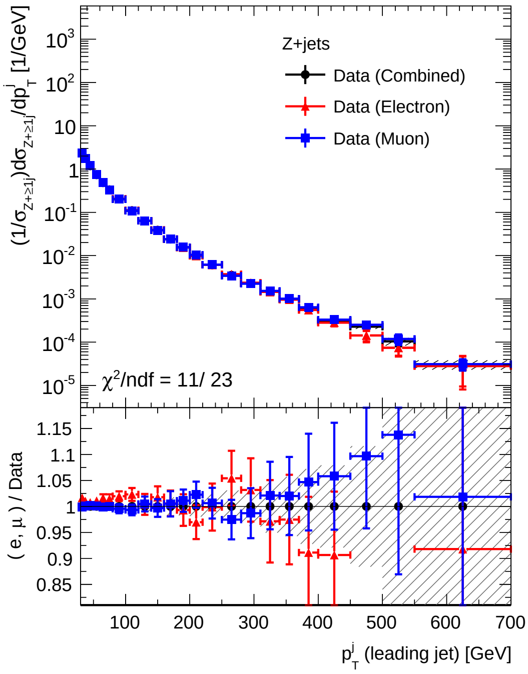
Figure 6.11 shows the combinations for the leading jet p_{T} distributions. The separate (acceptance corrected) distributions for electron and muon channels are shown along with the combined distribution. Generally the combined result falls between the electron and muon results and the resultant χ^2 values demonstrate that the two separate measurements are compatible. To cross check the consistency of the combined result in $W+\mathrm{jets}$ and $Z/\gamma^*+\mathrm{jets}$, the integral of the combined distribution was compared to the cross section in the corresponding bin of the unfolded jet multiplicity distributions. The integral was found to be consistent with the multiplicity distributions in all cases.

Uncertainty	Channel Correlation	Bin-to-bin Correlation	Included in Combination
Statistical	-	-	✓
Electron	-	✓	✓
Muon	-	✓	✓
JES (baseline)	✓	✓	-
JES (additional)	✓	✓	-
JER	✓	✓	-
MET	✓	✓	-
Multi-jet syst.	-	-	✓
Multi-jet stat.	-	-	✓
$t\bar{t}$ syst.	✓	✓	✓
$t\bar{t}$ stat.	-	-	✓
Single t	✓	✓	✓
Diboson	✓	✓	✓
$W \to \tau\nu$ and $Z \to \tau\tau$	✓	✓	✓
V+jets cross-talk	✓	✓	-
Unfolding model	✓	✓	✓
Unfolding stat.	-	-	✓
Luminosity	✓	✓	✓

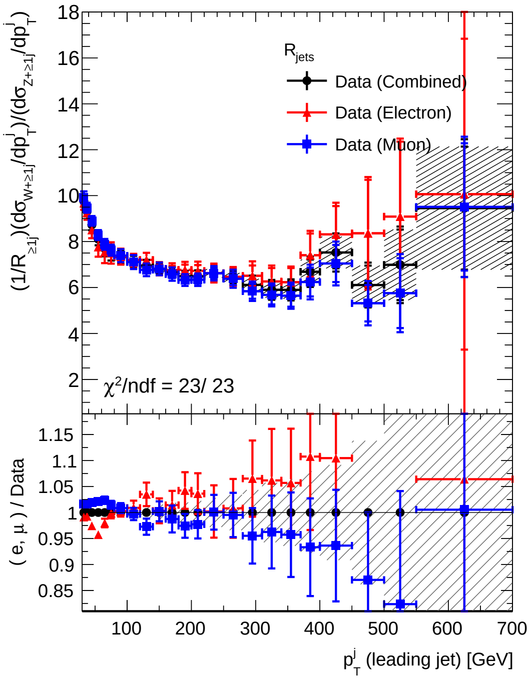
Table 6.3: Schematic of the treatment of systematics in the combination procedure. Each line represents a single uncertainty in the combination. Channel correlation refers to whether the uncertainty is treated as correlated between electron and muon channels and bin-to-bin correlation refers to whether the uncertainty is treated as correlated between bins within the distribution. The final column refers to whether the uncertainty is included in the combination. An uncertainty that is not included in the combination procedure is included on the combined result using a weighted average as discussed in Section 6.4.



(a) $W + \text{jets}$



(b) $Z/\gamma^* + \text{jets}$



(c) R_{jets}

Figure 6.11: Comparison of the $W + \text{jets}$, $Z/\gamma^* + \text{jets}$ and R_{jets} measurements as a function of the leading jet p_{T} in electron, muon and combined channels. The errors shown only include the statistical uncertainty.

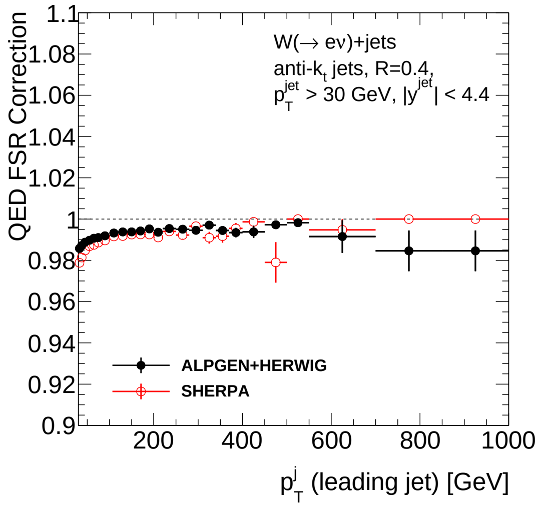
6.5 QED final state radiation corrections

Corrections for QED final state radiation (FSR) account for the differences obtained using the lepton four-momentum at the W/Z vertex before radiation (born) compared to using the lepton four-momentum after radiation and dressing (dressed) as discussed in Section 6.1. This is necessary as the theory predictions were provided using born leptons whilst the results were unfolded to dressed level. FSR corrections were computed using samples generated with Alpgen 2.13 [42] interfaced to Herwig 6.520 [43] and Jimmy 4.31 [44] using CTEQ6L1 [45] PDFs and the AUET2-CTEQ6L1 tune [46]. These samples use Photos [52] to calculate QED FSR. The FSR corrections were calculated as the bin-by-bin ratio of the distributions with dressed leptons and that with born leptons:

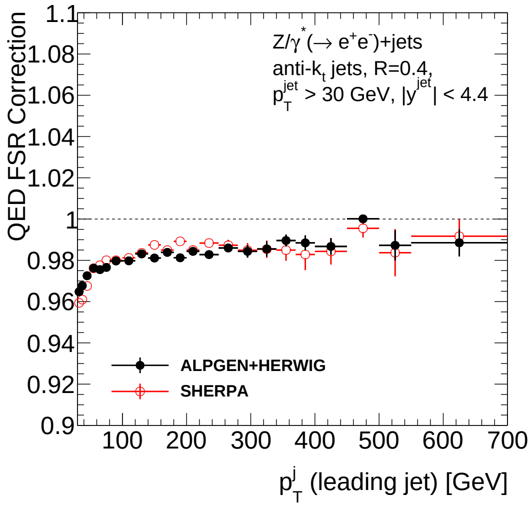
$$\delta^{\mathrm{FSR}} = \frac{\text{dressed leptons}}{\text{born leptons}}. \tag{6.7}$$

The systematic uncertainty on this correction was assigned by comparison to the correction derived using Sherpa 1.4 samples which uses a photon radiation model based on the YFS method [53]. The uncertainty was defined as the difference between the correction from Alpgen+Herwig and Sherpa and was taken to be symmetric around the value of the nominal correction.

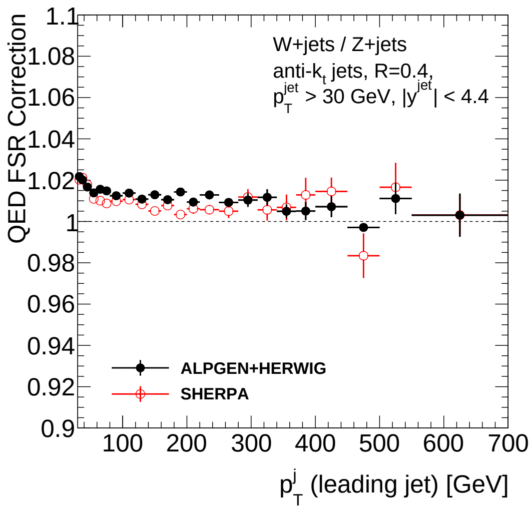
Figures 6.12 and 6.13 show the FSR correction for electron and muon channels respectively. The FSR correction reduces the average jet energy as dressed leptons tend to have a lower p_{T} than born leptons and so, low p_{T} jets are more likely to be recoiling against a truth lepton which is below the p_{T} threshold. It is observed that the correction is of order 1-2% in the W+jets channel and approximately 2-4% in the Z+jets channel as one might expect given the presence of two leptons in the Z boson final state. This results in a 1-2% correction in the R_{jets} measurement but in the opposite direction to the individual channels.



(a) W+jets

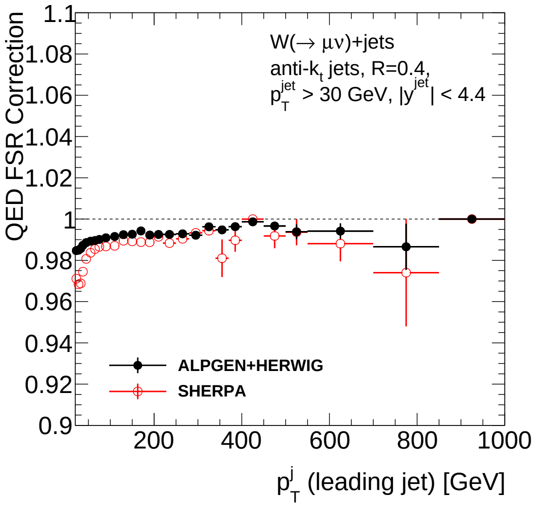


(b) Z+jets

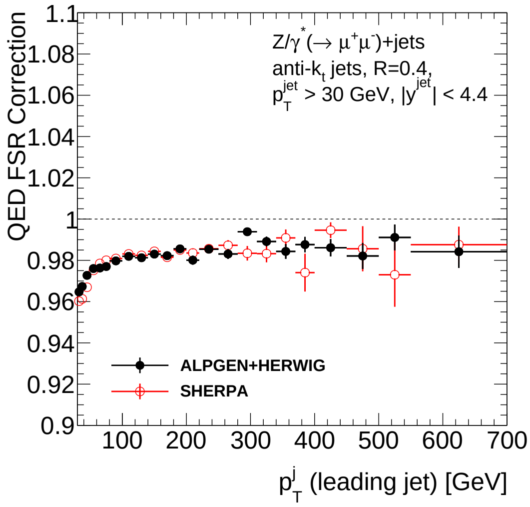


(c) R_{jets}

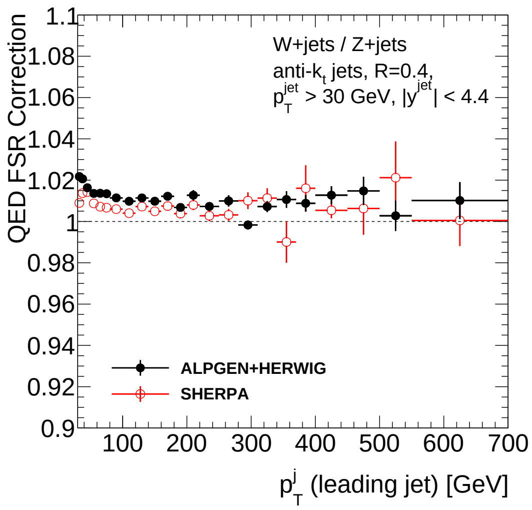
Figure 6.12: The QED final state radiation correction as a function of leading jet p_{T} from Alpgen+Herwig and Sherpa for the electron channel.



(a) W+jets



(b) Z+jets



(c) R_{jets}

Figure 6.13: The QED final state radiation correction as a function of leading jet p_{T} from ALPGEN+HERWIG and SHERPA for the muon channel.

6.6 Non-perturbative QCD corrections

Non-perturbative (NP) corrections account for contributions from the underlying event (UE) and hadronisation. The nominal corrections were estimated using $W+$ jets and $Z/\gamma^* +$ jets samples generated with Alpgen 2.13 [42] interfaced to Herwig 6.520 [43] and Jimmy 4.31 [44] using CTEQ6L1 [45] PDFs and the AUET2-CTEQ6L1 tune [46] with the underlying event turned on and off.

UE corrections were calculated as the bin-by-bin ratio of the distributions with UE turned on and off, having reconstructed jets from partons (before hadronisation) and using born leptons:

$$\delta^{\mathrm{UE}} = \frac{\text{parton jets, UE on, born leptons}}{\text{parton jets, UE off, born leptons}}. \tag{6.8}$$

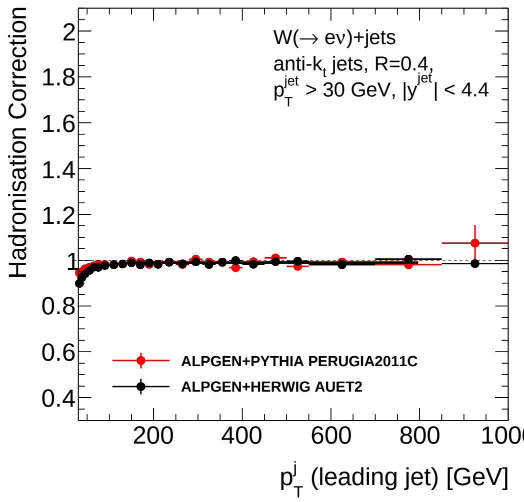
The hadronisation correction was computed as the bin-by-bin ratio of the distributions reconstructing jets using partons and reconstructing jets using hadrons, in both cases with the UE turned on and using born leptons:

$$\delta^{\mathrm{had}} = \frac{\text{hadron jets, UE on, born leptons}}{\text{parton jets, UE on, born leptons}}. \tag{6.9}$$

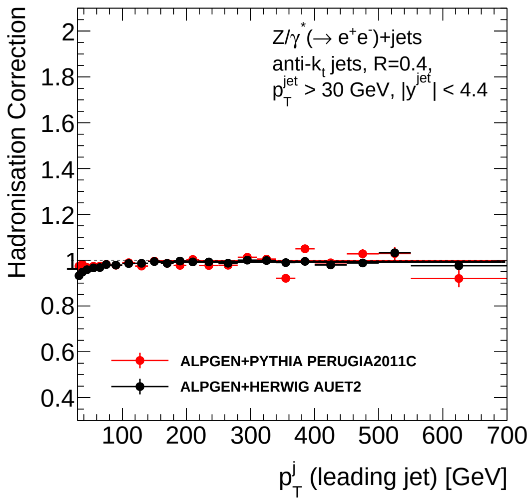
The systematic uncertainty on these corrections was assigned by comparison to corrections derived from samples generated with Alpgen 2.14 [42] interfaced to Pythia 6.425 [55] using the Perugia2011C tune [146] which employs both a different hadronisation model (Lund string) and UE tune. The systematic uncertainty was taken as the difference between the total correction ($\delta^{\mathrm{UE}} \times \delta^{\mathrm{had}}$) from Alpgen+Herwig and Alpgen+Pythia and was taken to be symmetric around the value of the nominal corrections.

Figures 6.14 and 6.15 show the hadronisation and underlying event corrections respectively for leading jet p_{T}. The hadronisation correction reduces the average jet energy included by the jet clustering due to increased out-of-cone radiation whilst the UE correction increases the average jet energy due to including remnants of the protons in the jet clustering. In the high jet p_{T} region large statistical fluctuations are observed which are mitigated by performing a linear fit in the high kinematic region. This fitting procedure is carried out in all p_{T}-like distributions.

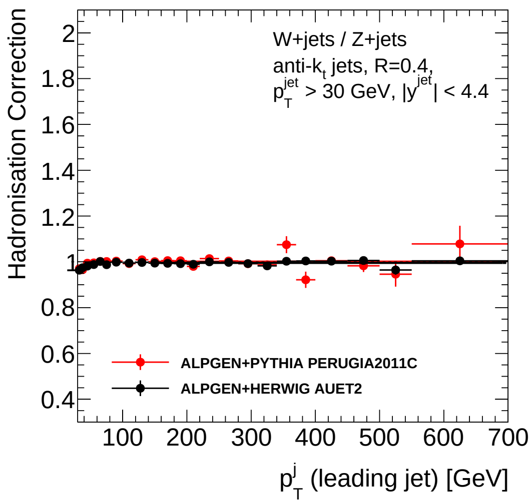
Figure 6.16 shows the combined NP correction for leading jet p_{T}. The combined correction on W+jets and Z+jets is about 15% at low p_{T} and quickly falls with increasing p_{T}. Due to the similarity in the corrections in $W+\text{jets}$ and $Z/\gamma^{*}+\text{jets}$ channels, the correction for the ratio measurement is very close to unity across the whole p_{T} range.



(a) W+jets

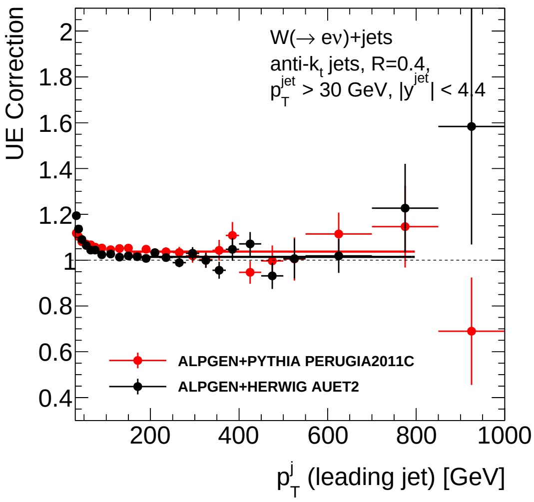


(b) Z+jets

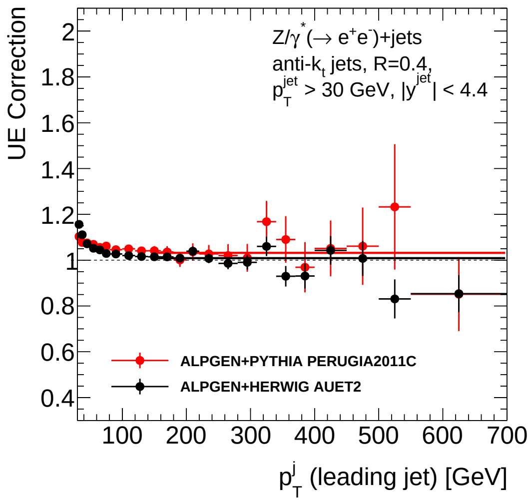


(c) R_{jets}

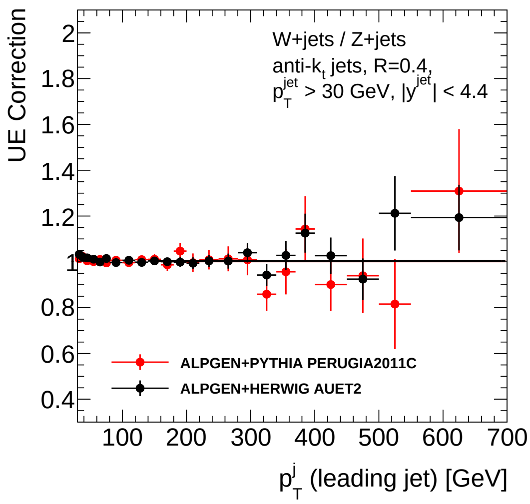
Figure 6.14: The hadronisation correction as a function of leading jet p_{T} from ALPGEN+HERWIG and ALPGEN+PYTHIA.



(a) W+jets

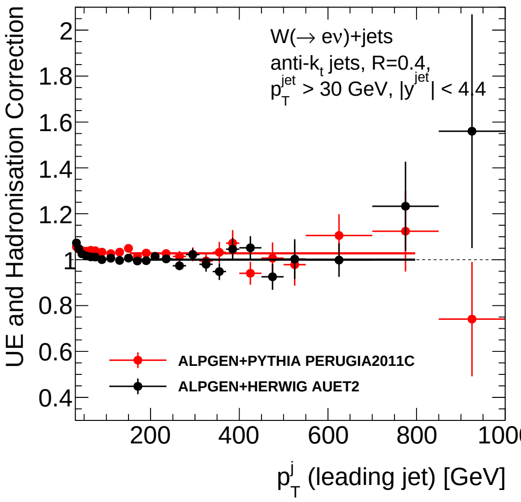


(b) Z+jets

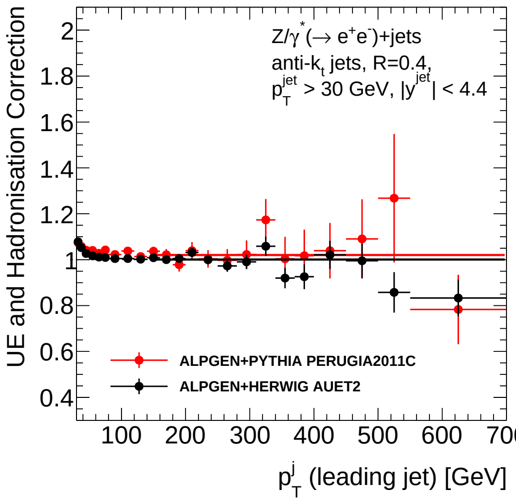


(c) R_{jets}

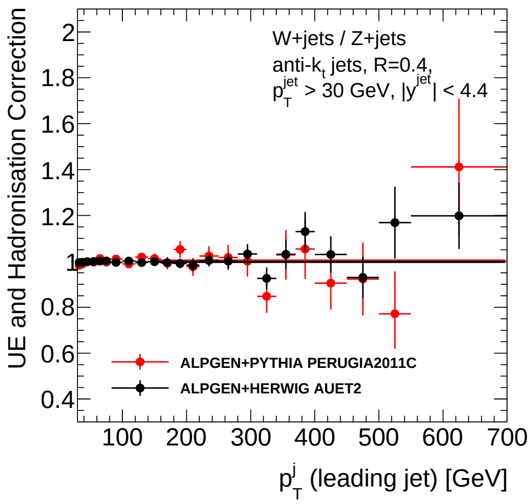
Figure 6.15: The underlying event correction as a function of leading jet p_{T} from Alpgen+Herwig and Alpgen+Pythia.



(a) W+jets



(b) Z+jets



(c) R_{jets}

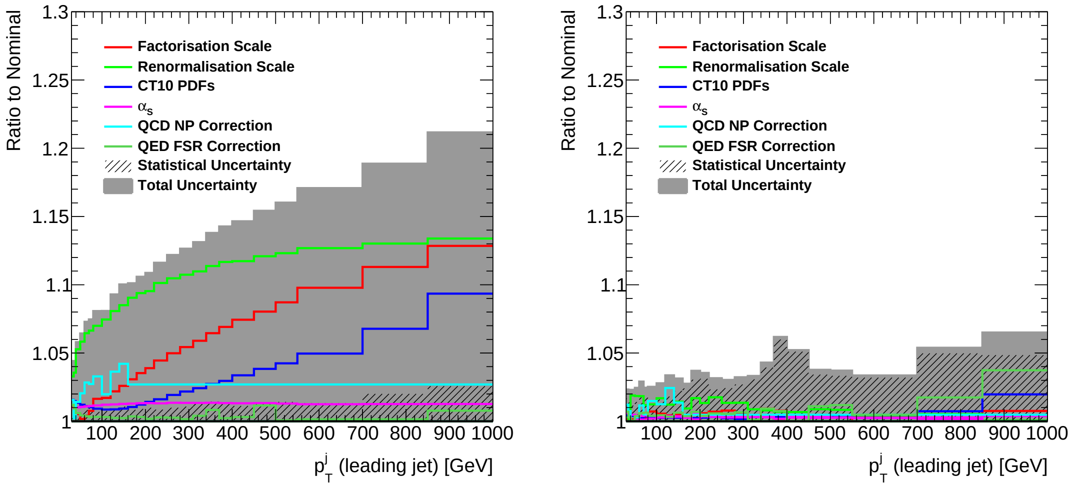
Figure 6.16: The total non-perturbative correction as a function of leading jet p_{T} from ALPGEN+HERWIG and ALPGEN+PYTHIA.

6.7 Total Theoretical Uncertainty

For comparison to the combined results, the BlackHat+Sherpa theory predictions (provided in the combined phase space) have a total uncertainty given by the quadratic sum of

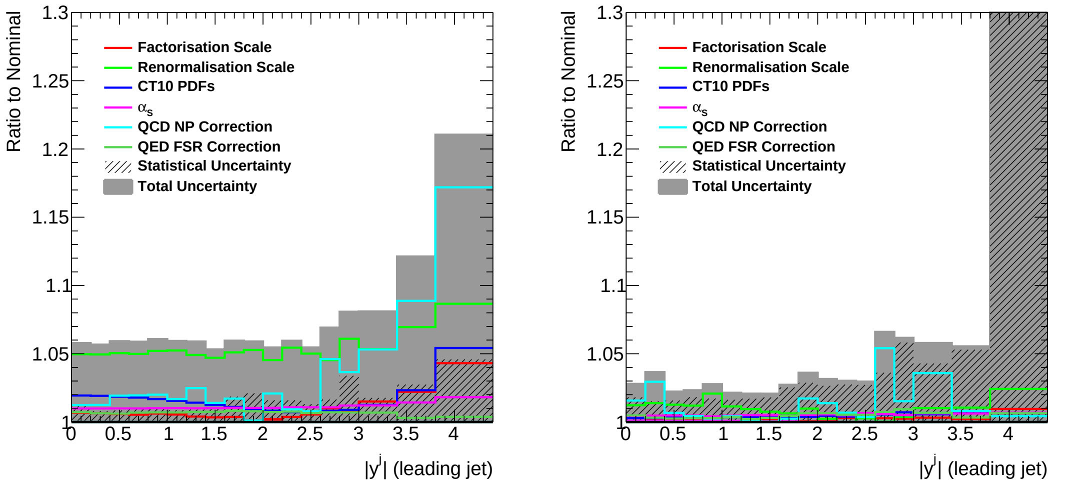
- The renormalisation scale uncertainty (derived by doubling and halving the nominal scale choice)
- The factorisation scale uncertainty (derived by doubling and halving the nominal scale choice)
- The PDF uncertainty (derived using the Hessian method [36] on the 52 CT10 eigenvectors)
- The α_S uncertainty (derived by varying α_S used in the calculation)
- The QED FSR correction uncertainty, discussed in Section 6.5
- The non-perturbative correction uncertainty, discussed in Section 6.6
- The statistical uncertainty

Figures 6.17 and 6.18 show the BlackHat+Sherpa uncertainties as a function of leading jet p_{T} and rapidity respectively. In results plots, the same uncertainty prescription is used to build the uncertainty band for BlackHat+Sherpa, LoopSim and BlackHat+Sherpa exclusive sums. In the case of HEJ (the resummation calculation described in Section 1.3.4), the non-perturbative corrections described above are not a-priori applicable as the resummation procedure may already encompass some of the hadronisation effects. As such, for HEJ, the uncertainties are built using the same prescription excluding the non-perturbative uncertainty. Note that HEJ uses leading order matrix elements and so, the scale uncertainties are significantly larger than the other predictions. MEPS@NLO uncertainties only include the statistical uncertainty on the prediction whilst SHERPA and ALPGEN include a 5% normalisation uncertainty [106] and the statistical uncertainty.



(a) $W \to e\nu$ (b) R_{jets}

Figure 6.17: The uncertainty on the $W +$ jets and R_{jets} predictions from BLACKHAT+SHERPA as a function of the leading jet p_{T}.



(a) $W \to e\nu$ (b) R_{jets}

Figure 6.18: The uncertainty on the $W +$ jets and R_{jets} predictions from BLACKHAT+SHERPA as a function of the leading jet rapidity.

Chapter 7

Results

The unfolded cross sections and ratio in the combined phase space, defined in Chapter 6, are compared to a number of theoretical predictions corrected for QED FSR and non-perturbative QCD using the corrections derived in Sections 6.5 and 6.6 respectively. Note that these corrections ensure that all the predictions are directly comparable to the measurement, except for those from HEJ as non-perturbative QCD corrections were not available for this prediction.

The ratio measurement is compared to BlackHat+Sherpa NLO predictions and LO multi-leg Monte Carlo predictions from Alpgen+Herwig and Sherpa. The W+jets results are also compared to these predictions and, in addition, are compared to predictions from HEJ (resummation), MEPS@NLO (NLO Monte Carlo), LoopSim (approximate NNLO) and BlackHat+Sherpa exclusive sums. More details of these calculations can be found in Section 1.3 and the predictions are summarised in Table 7.1.

Section 7.1 presents the unfolded cross sections for W+jets distributions whilst Section 7.2 presents the unfolded ratio measurements.

Program	Max. no. of partons at approx. NNLO ($\alpha_s^{N_{\text{jet}}+2}$)	NLO ($\alpha_s^{N_{\text{jet}}+1}$)	LO ($\alpha_s^{N_{\text{jet}}}$)	Parton/Particle level	Distributions shown
LoopSim	1	2	3	parton level all corrections	Leading jet p_{T} and H_{T} for $W + \geq 1$ jet
BlackHat+Sherpa	–	5	6	parton level all corrections	All
BlackHat+Sherpa exclusive sums	1	2	3	parton level all corrections	Leading jet p_{T} and H_{T} for $W + \geq 1$ jet
HEJ	all orders, resummation			parton level QED FSR corr.	All for $W + \geq 2$, 3, 4 jets
MEPS@NLO	–	2	4	particle level	All
Alpgen	–	–	5	particle level	All
Sherpa	–	–	4	particle level	All

Table 7.1: Summary of theoretical predictions, including the maximum number of partons at each order in α_s, whether or not the predictions are calculated at parton or particle level and the distributions for which they are shown. Also shown are which of QCD hadronisation, QCD underlying event and QED FSR corrections are applied.

7.1 W+jets

7.1.1 Jet multiplicity

Figures 7.1 and 7.2 show the cross section for $W \to l\nu$ as a function of inclusive and exclusive jet multiplicity. Both inclusive and exclusive variables are investigated in this thesis as they probe different parts of the theoretical predictions. For example, exclusive jet multiplicity probes the theory of jet vetoes such as those used in Higgs physics to improve signal to background ratios [147]. The data are in good agreement with the predictions from BlackHat+Sherpa and HEJ. Both Alpgen and Sherpa agree with the data within the experimental systematic uncertainties however they exhibit different behaviour at high jet multiplicity. Alpgen has a tendency to underestimate the data at high multiplicity whilst Sherpa tends

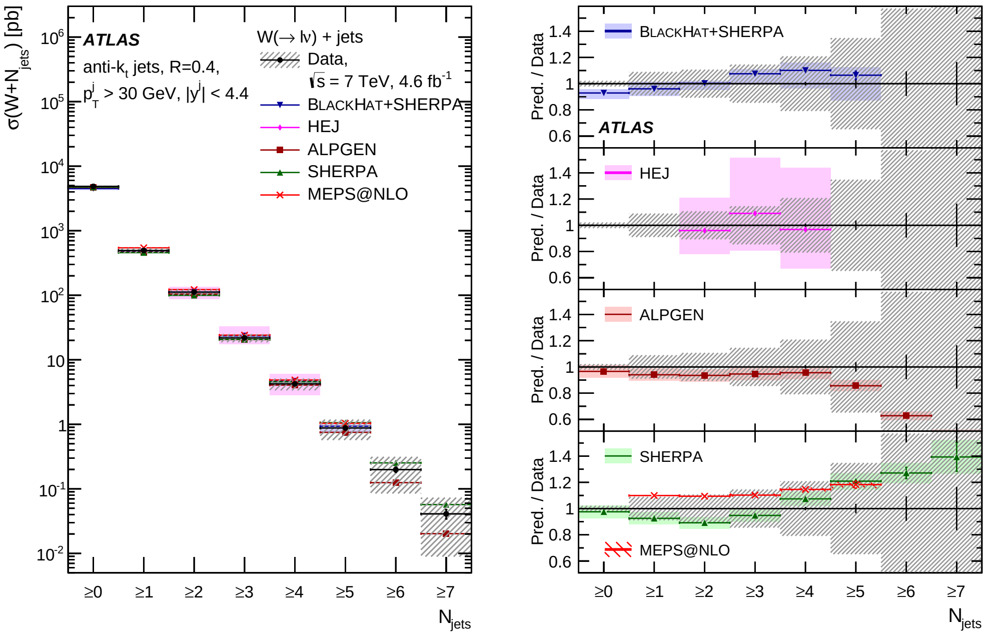


Figure 7.1: Cross section for the production of $W \to l\nu$ as a function of inclusive jet multiplicity. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

to overestimate the cross section. MEPS@NLO shows a shape improvement over SHERPA however a 10% overestimation of the cross section is observed.

These results demonstrate that no prediction perfectly agrees with the measurement of the cross section as a function of jet multiplicity. In order to avoid this imperfect normalisation of the predictions affecting the understanding of the shape differences in differential distributions, in the following plots all the theoretical predictions are scaled to the measured W + jets cross section in the corresponding jet multiplicity bin. In this way, normalisation differences are removed and the shapes of the distributions can be compared. Table 7.2 shows the scale factors applied to the predictions.

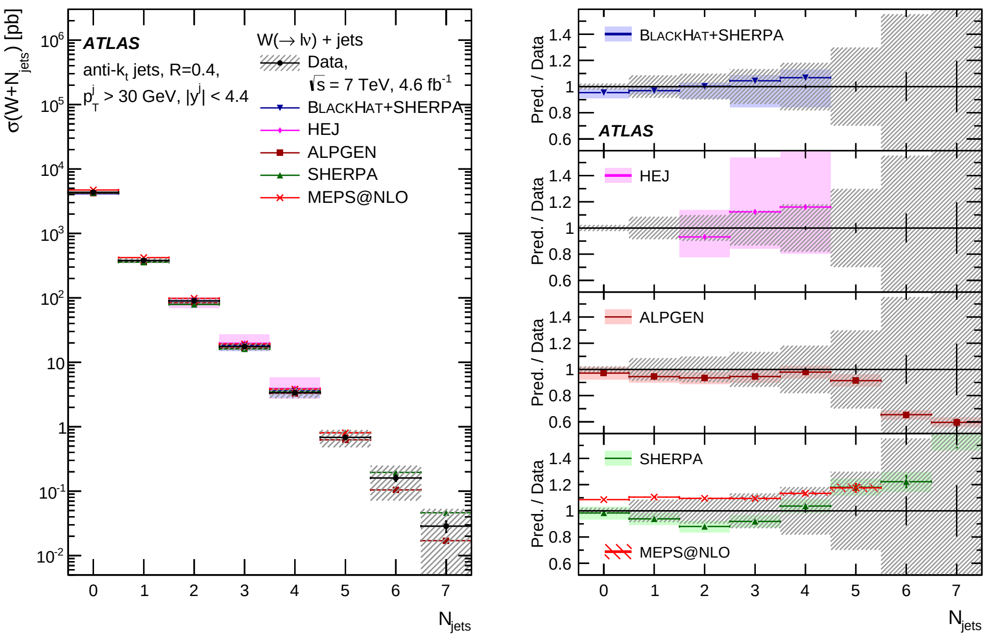


Figure 7.2: Cross section for the production of $W \to l\nu$ as a function of exclusive jet multiplicity. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

$N_{\rm jet}$	≥ 1	$= 1$	≥ 2	$= 2$	≥ 3	$= 3$	≥ 4	≥ 5
LoopSim	1.029	-	-	-	-	-	-	-
BlackHat+Sherpa	0.960	0.969	1.003	1.002	1.075	1.044	1.101	1.064
HEJ	-	-	0.960	0.932	1.091	1.123	0.968	-
MEPS@NLO	1.099	1.105	1.094	1.095	1.103	1.094	1.146	1.183
Alpgen	0.940	0.945	0.936	0.935	0.946	0.946	0.960	0.856
Sherpa	0.925	0.939	0.892	0.880	0.948	0.919	1.074	1.209

Table 7.2: The scale factors which are applied to the theoretical predictions in differential cross section distributions.

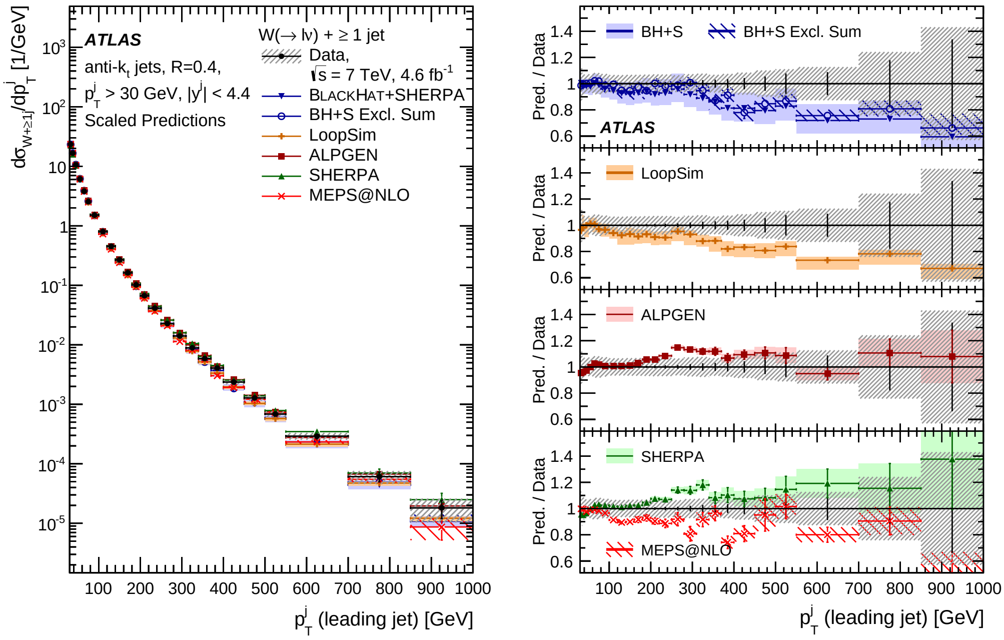


Figure 7.3: Cross section for the production of $W \to l\nu$ as a function of leading jet p_{T} in events with $N_{\mathrm{jet}} \geq 1$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

7.1.2 Jet transverse momentum

Figure 7.3 shows the cross section as a function of leading jet p_{T} in events with one or more jets. It is observed that the fixed order prediction at NLO from BlackHat+Sherpa underestimates the data at high transverse momentum. It is expected in this region that there are contributions from higher-order $W+ \geq 2$ jets however approximate NNLO predictions from LoopSim and BlackHat+Sherpa exclusive sums do not show any improvement in the description of the data. These discrepancies with the data are expected to be further exacerbated by the higher order electroweak corrections which are expected to be large (20%) at high p_{T} [148]. Predictions from Alpgen, Sherpa and MEPS@NLO are generally within one standard deviation of the data although there are regions where systematic deviations are seen over a range of bins such as at low p_{T} for MEPS@NLO.

Figure 7.4 shows the cross section as a function of leading jet p_{T} in events with

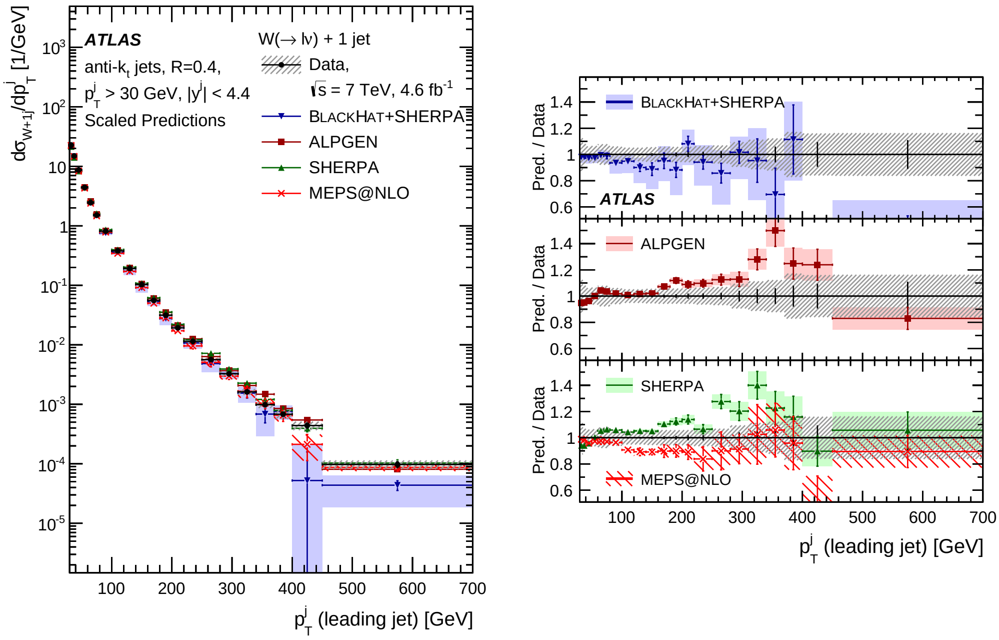


Figure 7.4: Cross section for the production of $W \to l\nu$ as a function of leading jet p_{T} in events with $N_{\mathrm{jet}} = 1$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

exactly one jet. In this case, the agreement between NLO predictions and the data is better (within the large statistical uncertainties) as already observed in Z+jet events [135]. This agreement is unexpected given that at high leading jet p_{T} there is a large scale separation between the leading jet p_{T} and the second jet p_{T} ($p_{\mathrm{T}} < 30$ GeV) and resummation effects are expected to be important. An explanation for this effect has been put forward in Ref. [149]. The surprisingly good agreement appears to arise from the details of the jet isolation requirement whereby jets within $\Delta R < 0.5$ of leptons are vetoed but the event is not. As a result, the measurement includes a piece arising from dijet production with the emission of a boson close to one of the jets. Such a process has an accompanying "giant K-factor" [60] and may explain the surprisingly good agreement observed here.

Figure 7.5 shows the cross section as a function of leading jet p_{T} in events with two or more jets. The SHERPA prediction systematically overestimates the data at high jet p_{T} whist the ALPGEN prediction shows similar agreement as for the

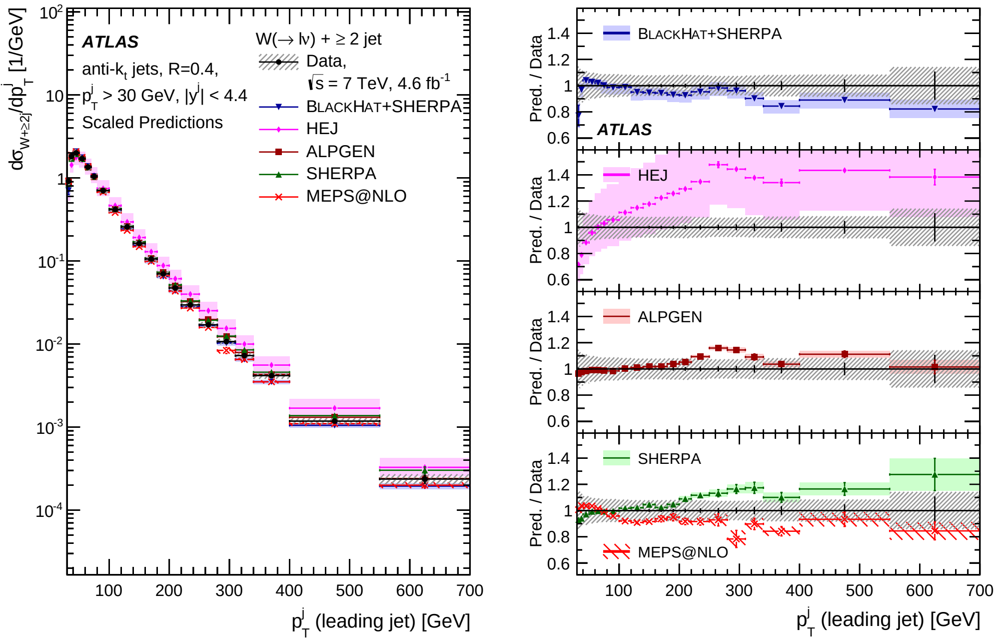


Figure 7.5: Cross section for the production of $W \to l\nu$ as a function of leading jet p_{T} in events with $N_{\mathrm{jet}} \geq 2$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

inclusive leading jet p_{T} above. BLACKHAT+SHERPA and MEPS@NLO generally agree well with the data. From this jet multiplicity upwards, the HEJ predictions are also available. For $N_{\mathrm{jet}} \geq 2$ HEJ predicts a harder jet spectrum than measured in data albeit with large leading-order scale uncertainties[1]. Figure 7.6 shows the cross section as a function of leading jet p_{T} in events with three or more jets and it is observed that all predictions describe the data well within the uncertainties.

Figure 7.7 shows the cross section as a function of the second leading jet p_{T} in events with two or more jets. It is observed that ALPGEN and SHERPA agree with the data best whilst BLACKHAT+SHERPA generally underestimates the data. MEPS@NLO predicts the shape at high p_{T} ($p_{\mathrm{T}} > 100$ GeV) well but the shape differs significantly at low p_{T}. HEJ predicts a harder jet spectrum, similar to what was observed for the leading jet p_{T} above.

Figure 7.8 shows the cross section as a function of the third leading jet p_{T} in

[1]recall HEJ has no non-perturbative QCD corrections applied.

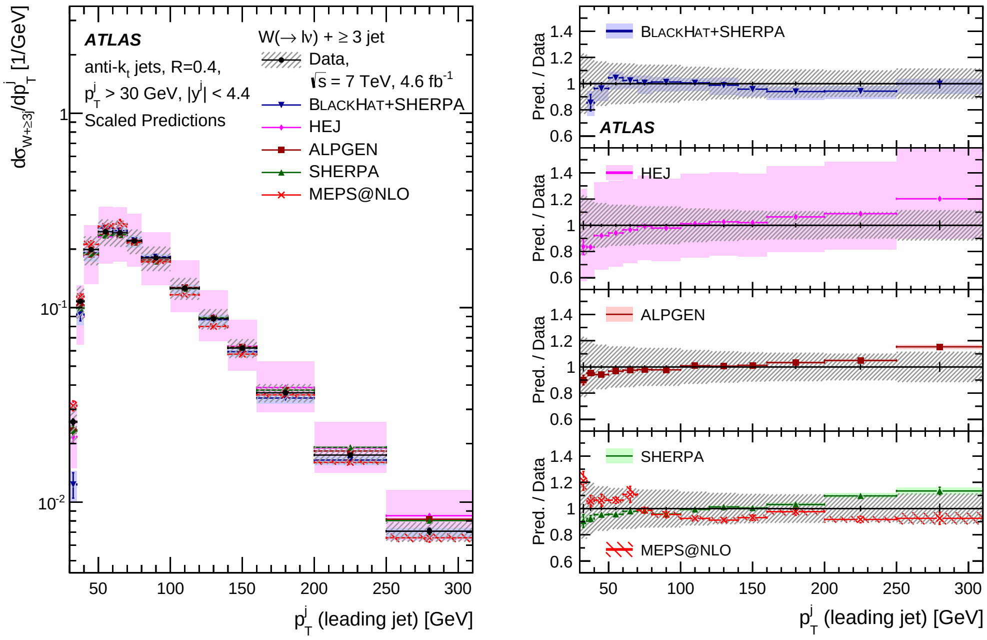


Figure 7.6: Cross section for the production of $W \to l\nu$ as a function of leading jet p_{T} in events with $N_{\mathrm{jet}} \geq 3$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

events with three or more jets. Sherpa tends to overestimate the data at high p_{T} whilst all other predictions agree with the data within one standard deviation.

Figure 7.9 shows the cross section as a function of the fourth leading jet p_{T} in events with four or more jets. Predictions from HEJ generally provide a better description than for lower jet multiplicities as, with increasing jet multiplicity, jets are more likely to have similar transverse momenta and a larger rapidity separation between the most forward and backward jets. In such topologies the resummation approximations used in HEJ are expected to work better. Alpgen, Sherpa and BlackHat+Sherpa describe the data within the experimental uncertainties but exhibit shape differences at high p_{T}.

Figure 7.10 shows the cross section as a function of the fifth leading jet p_{T}. All the available predictions agree with the data within the experimental uncertainties.

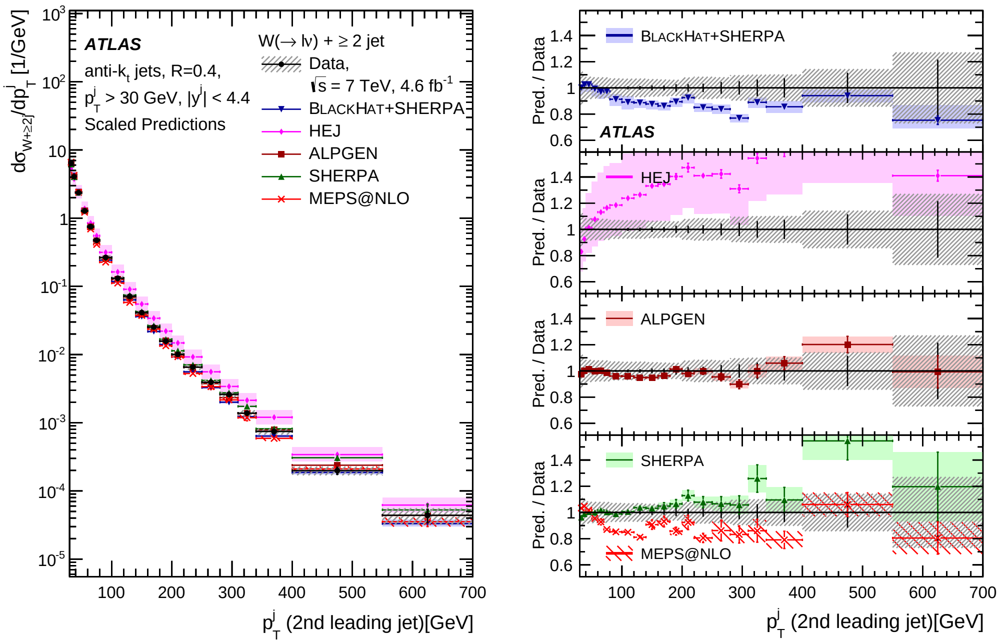


Figure 7.7: Cross section for the production of $W \to l\nu$ as a function of second leading jet p_{T} in events with $N_{\mathrm{jet}} \geq 2$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

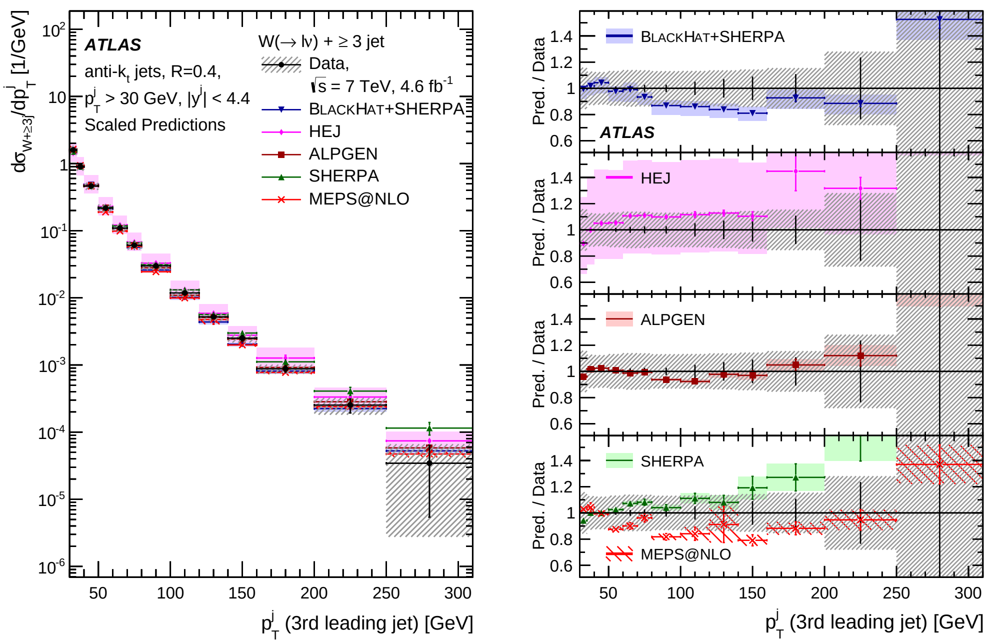


Figure 7.8: Cross section for the production of $W \to l\nu$ as a function of third leading jet p_{T} in events with $N_{\mathrm{jet}} \geq 3$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

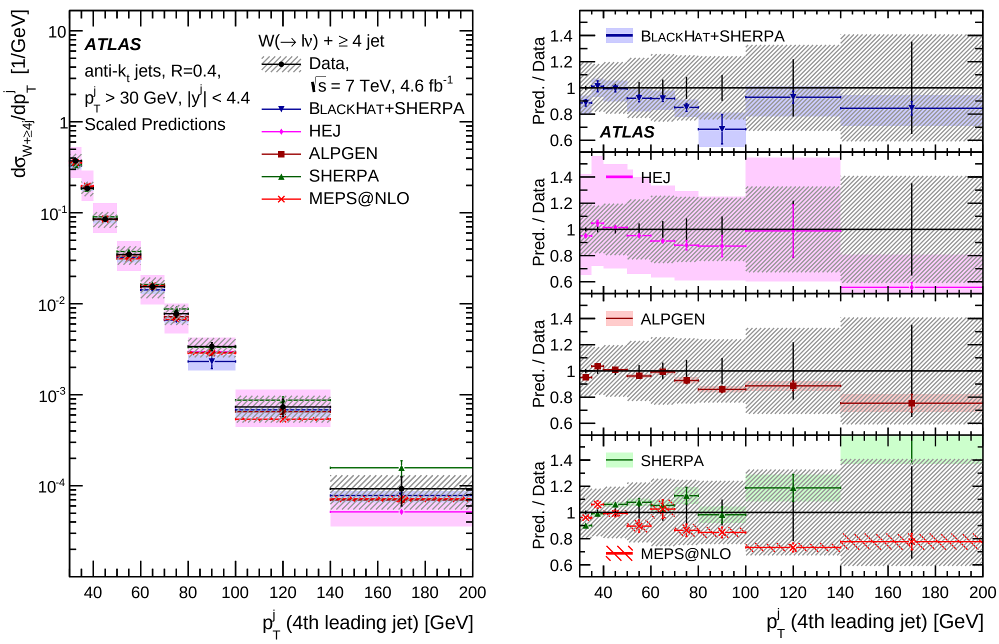


Figure 7.9: Cross section for the production of $W \to l\nu$ as a function of fourth leading jet p_{T} in events with $N_{\mathrm{jet}} \geq 4$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

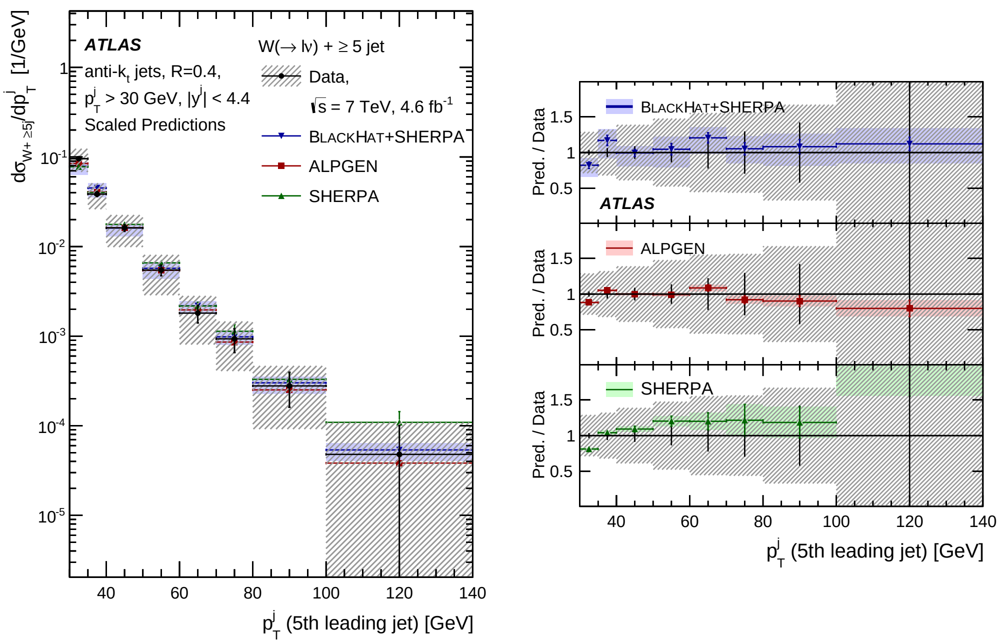


Figure 7.10: Cross section for the production of $W \to l\nu$ as a function of fifth leading jet p_{T} in events with $N_{\mathrm{jet}} \geq 5$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

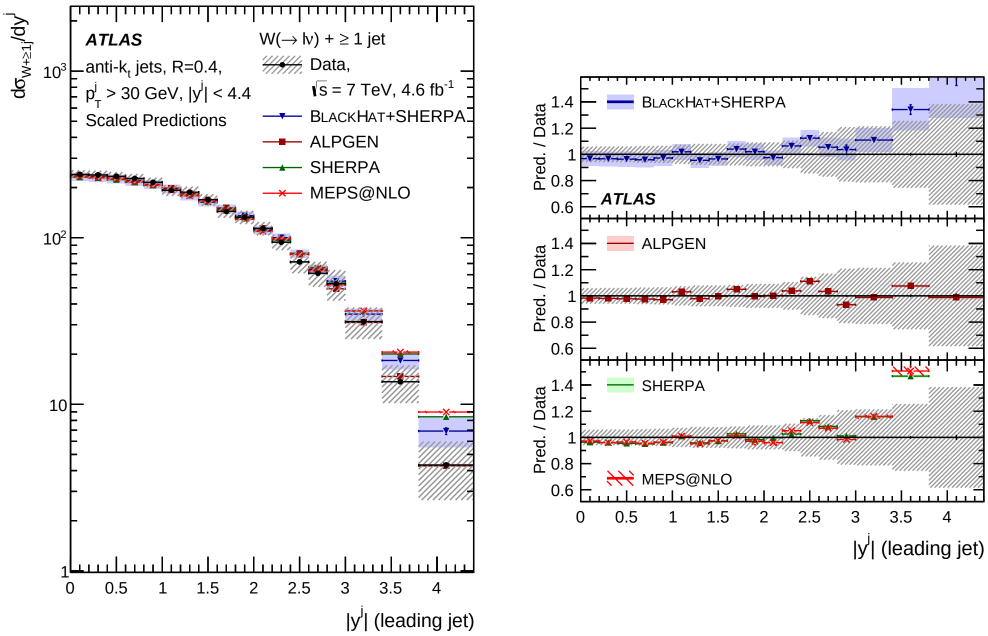


Figure 7.11: Cross section for the production of $W \to l\nu$ as a function of leading jet rapidity in events with $N_{\text{jet}} \geq 1$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

7.1.3 Jet rapidity

Figure 7.11 shows the cross section as a function of the leading jet rapidity in events with one or more jets. Overall, good agreement is observed in the central region, although in the forward region ($|y| > 3.5$) only ALPGEN shows good agreement with the data. BLACKHAT+SHERPA overestimates the leading jet rapidity by one standard deviation whilst SHERPA overestimates the data by two or more standard deviations. No improvement is observed over SHERPA in the predictions from MEPS@NLO. Reweighting ALPGEN to the same (CT10) PDFs used in the other predictions has little effect on the prediction.

Figure 7.12 shows the cross section as a function of the second leading jet rapidity in events with one or more jets. Similar effects are seen as for the leading jet rapidity discussed above. In the forward region SHERPA shows an improved description of the second leading jet as compared to the leading jet rapidity whilst MEPS@NLO

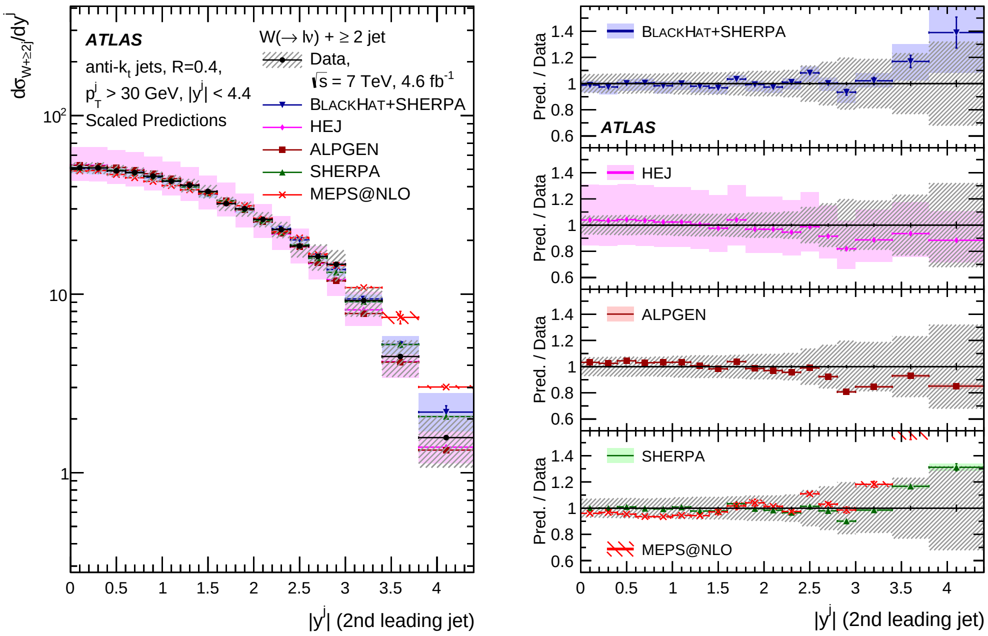


Figure 7.12: Cross section for the production of $W \to l\nu$ as a function of second leading jet rapidity in events with $N_{\text{jet}} \geq 2$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

continues to disagree with the data to a similar extent as in the leading jet rapidity distribution. At this jet multiplicity the HEJ predictions are available and provide a good description of the data over the full rapidity range.

Further plots of the cross section as a function of higher jet multiplicity rapidities (up to the fifth leading jet) are shown in Appendix A.

7.1.4 Inclusive variables

Inclusive variables such as H_{T} (the scalar sum of neutrino p_{T}, lepton p_{T} and jet p_{T}) and S_{T} (the scalar sum of jet p_{T}) are of interest as they are often used in new physics searches [150–152]. New physics processes, such as supersymmetry, are expected to produce events with large numbers of high p_{T} objects resulting in excesses of events at high H_{T} or S_{T} and so, good understanding of such variables up to high values is important.

Figure 7.13 shows the cross section as a function of H_{T} (the scalar sum of neutrino p_{T}, lepton p_{T} and jet p_{T}) in events with at least one jet. It is observed that the NLO prediction from BlackHat+Sherpa underestimates the data at high values of H_{T} due to missing higher order terms. H_{T} is one of the kinematic variables most influenced by subprocesses such as $qq \to qqW$ (dijet production followed by emission of a W boson from one of the quarks) [60]. The partial inclusion of such higher order terms in approximate NNLO predictions from LoopSim and BlackHat+Sherpa exclusive sums significantly improves the agreement of the predictions with data.

Figure 7.14 shows the cross section as a function of H_{T} in events with exactly one jet. Much better agreement between BlackHat+Sherpa and data is observed as the influence of $qq \to qqW$ subprocesses is reduced by requiring exactly one jet.

Figures 7.15, 7.16 and 7.17 shows the cross section as a function of H_{T} in events with two or more jets, three or more and four or more jets. The predictions are generally in good agreement with the data, especially the predictions from BlackHat+Sherpa and MEPS@NLO. Generally Sherpa and HEJ overestimate the cross sections at high H_{T} but the significance of the deviation reduces with increasing jet multiplicity.

Figure 7.18 shows the cross section as a function of S_{T} (the scalar sum of jet p_{T}) in events with at least one jet. The conclusions drawn are the same for S_{T} as for

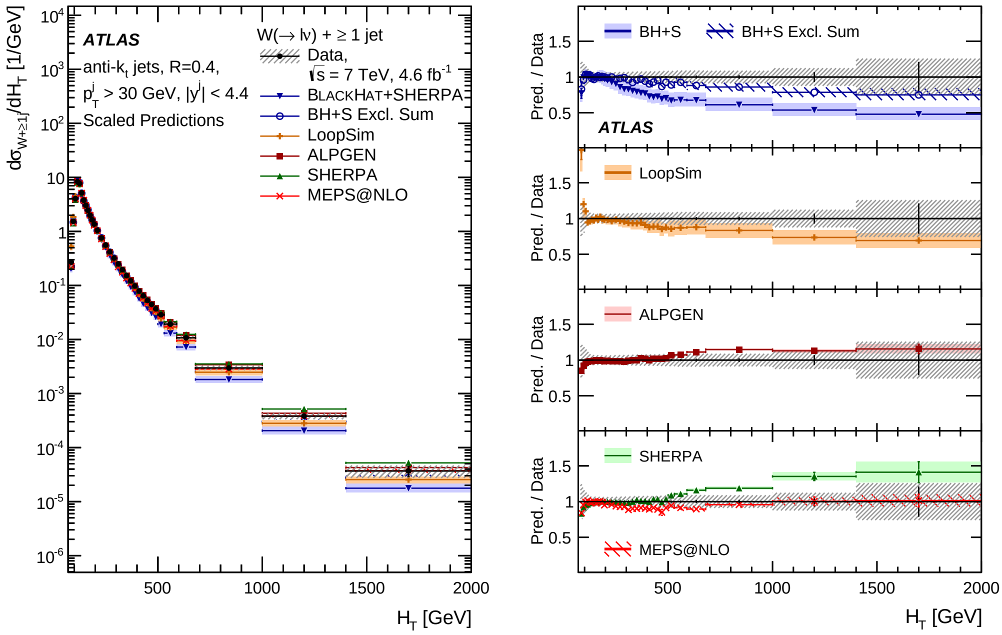


Figure 7.13: Cross section for the production of $W \to l\nu$ as a function of H_{T} in events with $N_{\mathrm{jet}} \geq 1$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

H_{T}, although the agreement between data and prediction is generally better at low S_{T} than low H_{T}.

Further plots of the cross section as a function of H_{T} and S_{T} for a range of jet multiplicities are shown in Appendix A.

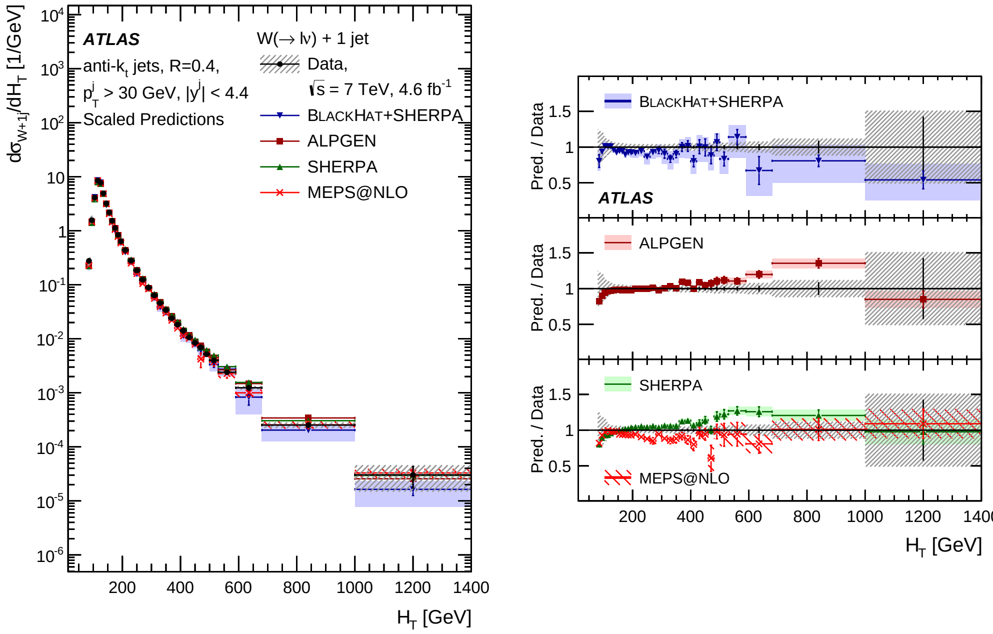


Figure 7.14: Cross section for the production of $W \to l\nu$ as a function of H_{T} in events with $N_{\mathrm{jet}} = 1$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

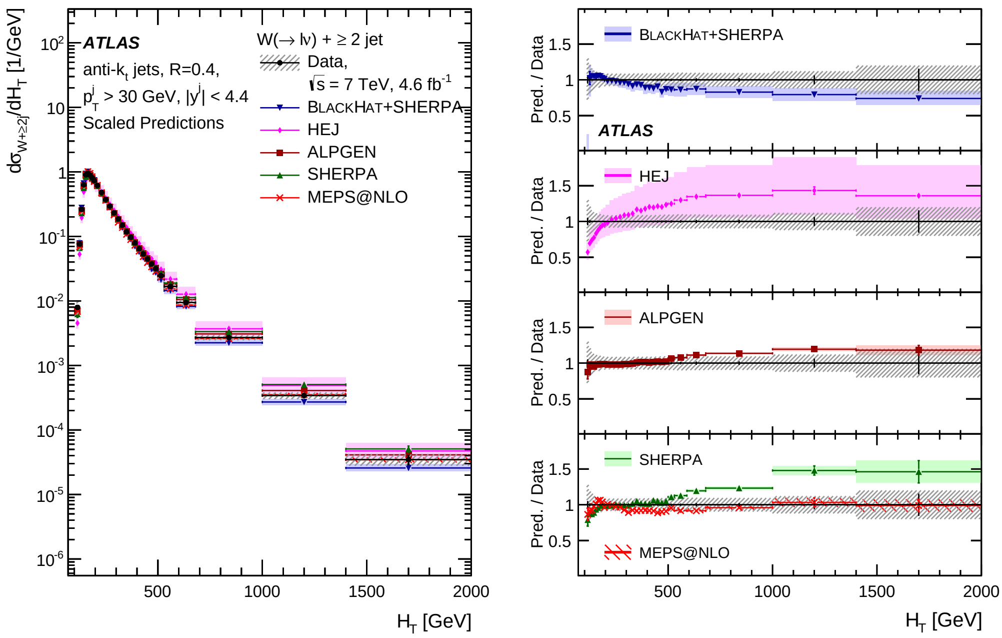


Figure 7.15: Cross section for the production of $W \to l\nu$ as a function of H_{T} in events with $N_{\mathrm{jet}} \geq 2$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

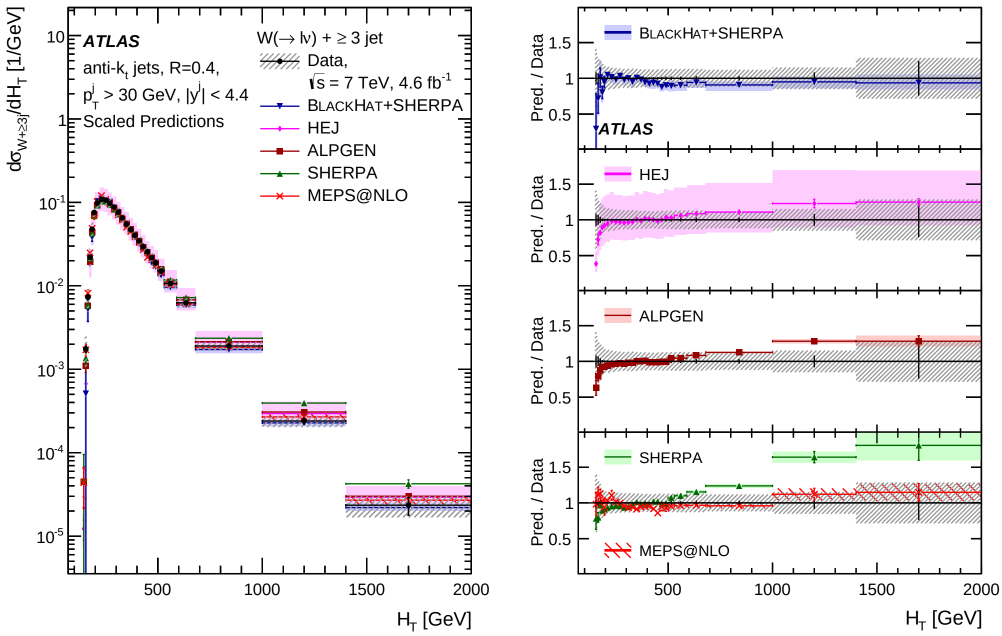


Figure 7.16: Cross section for the production of $W \to l\nu$ as a function of H_T in events with $N_{\text{jet}} \geq 3$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

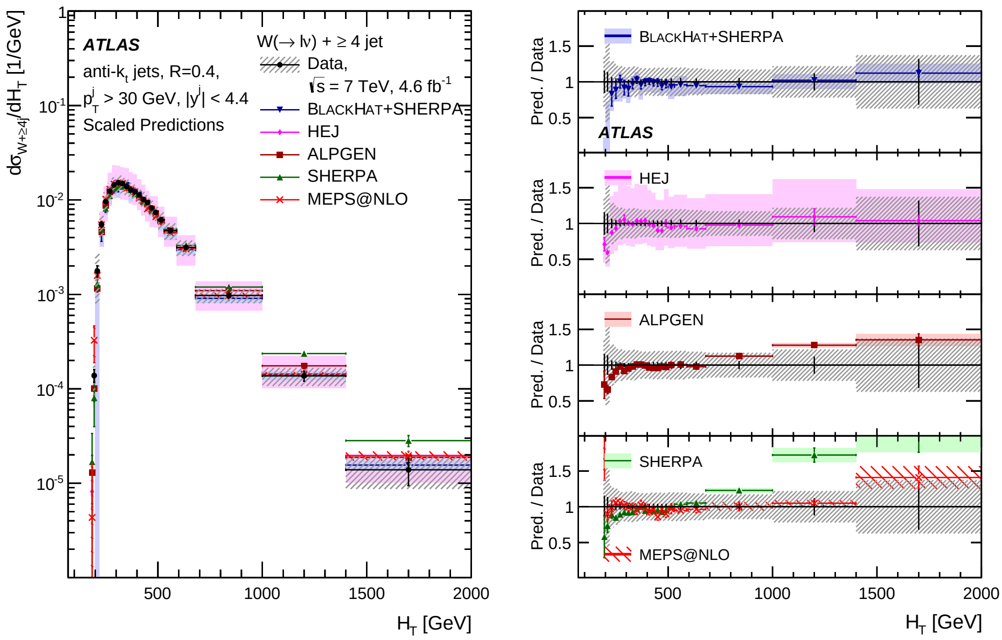


Figure 7.17: Cross section for the production of $W \to l\nu$ as a function of H_T in events with $N_{\text{jet}} \geq 4$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

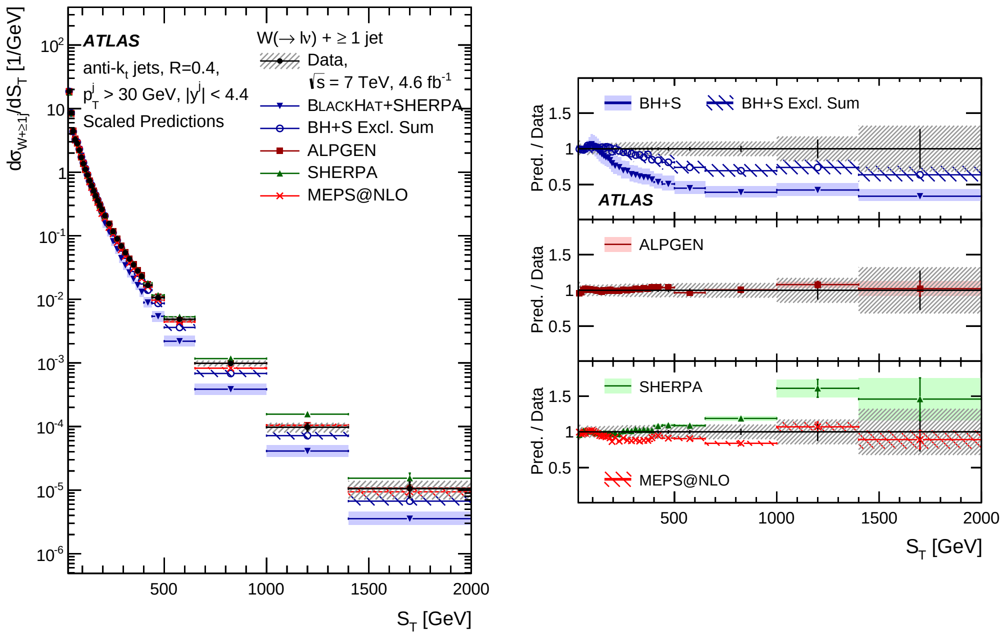


Figure 7.18: Cross section for the production of $W \to l\nu$ as a function of S_T in events with $N_{jet} \geq 1$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

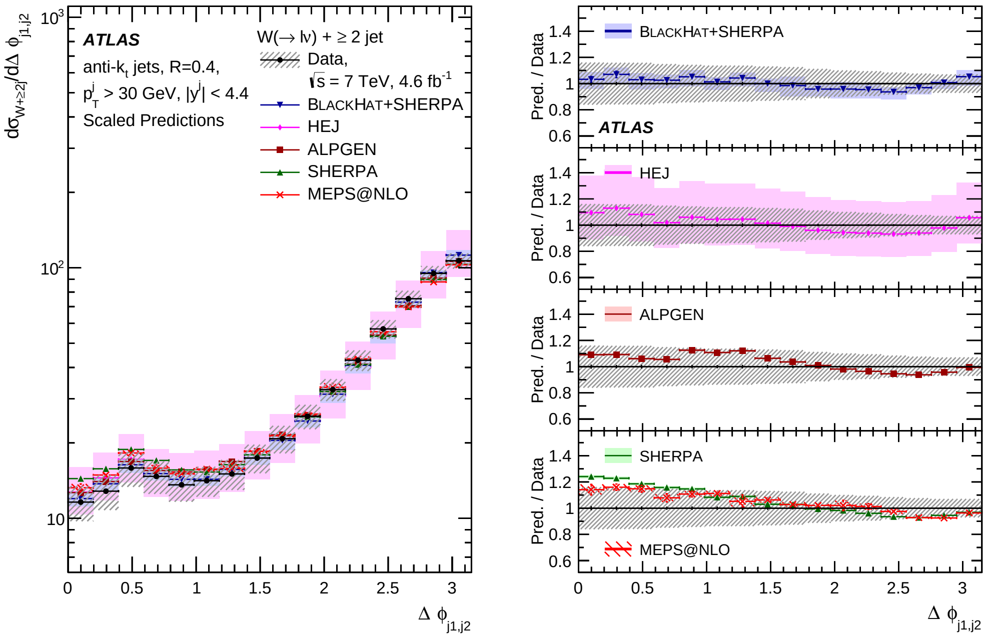


Figure 7.19: Cross section for the production of $W \to l\nu$ as a function of $\Delta\phi_{j1,j2}$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

7.1.5 Dijet variables

Dijet variables such as the dijet invariant mass and angular separation between jets are of interest to analyses using vector boson fusion (VBF) processes either in Higgs physics [153] or new physics searches [154]. The VBF process results in two jets which are generally of high invariant mass and angular separation and so, understanding of Standard Model processes in these phase spaces are particularly important.

Figure 7.19 shows the cross section as a function of the azimuthal angle between the two leading jets ($\Delta\phi_{j1,j2}$). All the predictions agree reasonably with the data, although SHERPA overestimates the cross section at small azimuthal separation. This small tension is improved upon by the MEPS@NLO methodology.

Figure 7.20 shows the cross section as a function of the rapidity separation be-

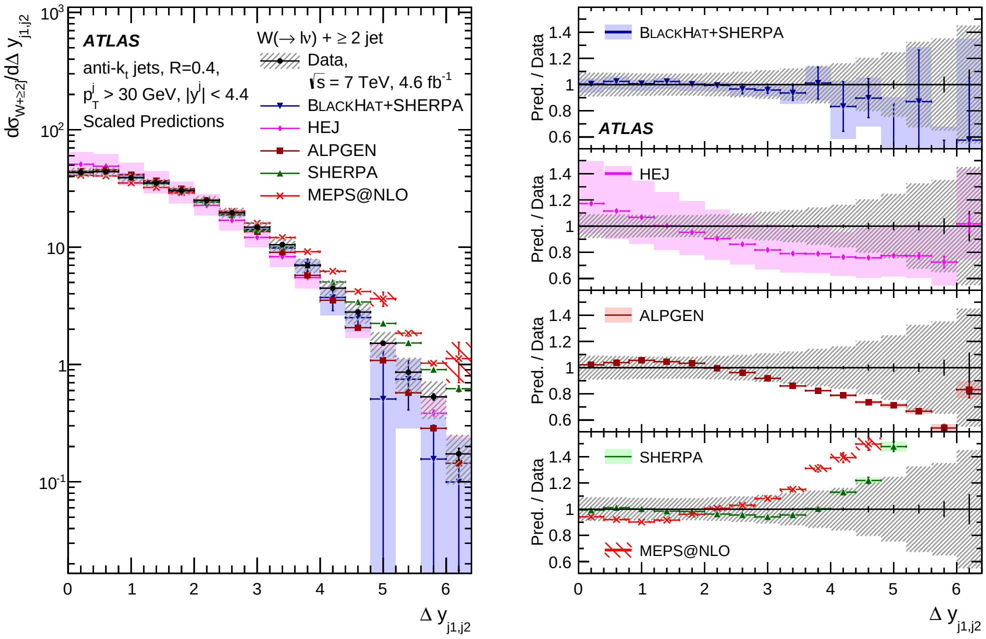


Figure 7.20: Cross section for the production of $W \to l\nu$ as a function of $\Delta y_{j1,j2}$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

tween the two leading jets ($\Delta y_{j1,j2}$). BlackHat+Sherpa predicts the data well, although with large statistical uncertainties in the large separation region, and Alpgen tends to underestimate the data at large rapidity separation. Sherpa overestimates the cross section in the large rapidity region and this tension is increased in the updated MEPS@NLO methodology. Finally, HEJ appears to underestimate the data at large separation however the shape in this region ($\Delta y_{j1,j2} > 3.5$) agrees well with data. This is the region in which the HEJ resummation approximation is most valid.

Figure 7.21 shows the cross section as a function of the $\eta - \phi$ separation between the two leading jets ($\Delta R_{j1,j2}$). The comparison of predictions and data reflects the observations discussed above in terms of the azimuthal and rapidity separation with BlackHat+Sherpa showing the best agreement and all predictions demonstrating very different shapes.

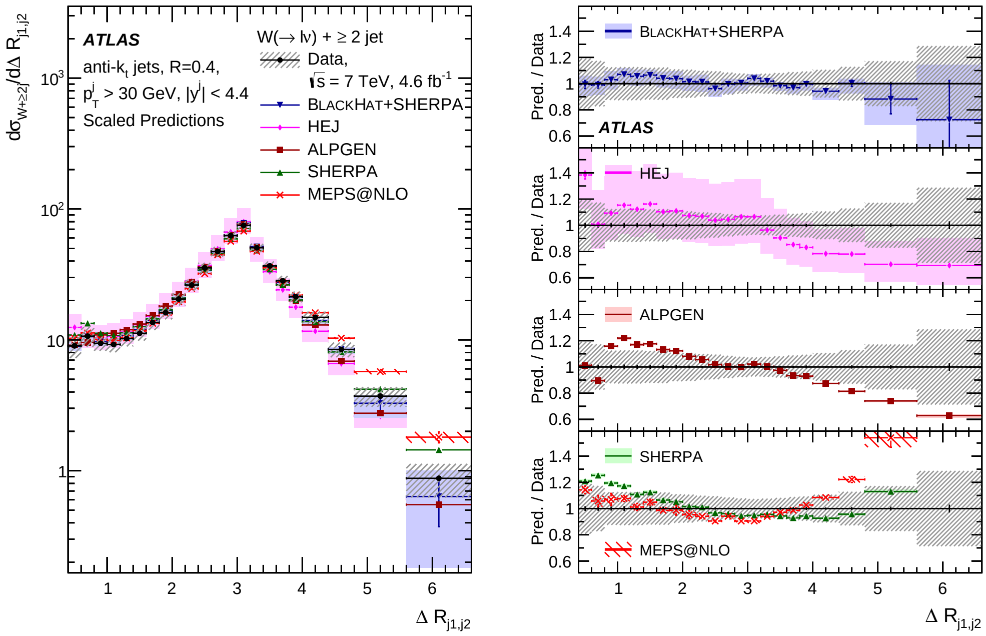


Figure 7.21: Cross section for the production of $W \to l\nu$ as a function of $\Delta R_{j1,j2}$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

Figure 7.22 shows the cross section as a function of the invariant mass of the two leading jets (m_{12}). SHERPA and MEPS@NLO struggle to predict the data at high invariant mass, both significantly overestimating the data in this region. ALPGEN underestimates the data at high invariant mass by one standard deviation and BLACKHAT+SHERPA shows a similar tendency although the prediction is only available up to 1 TeV. Finally, the HEJ prediction shows good agreement with the data across the range, again demonstrating the usefulness of resummation.

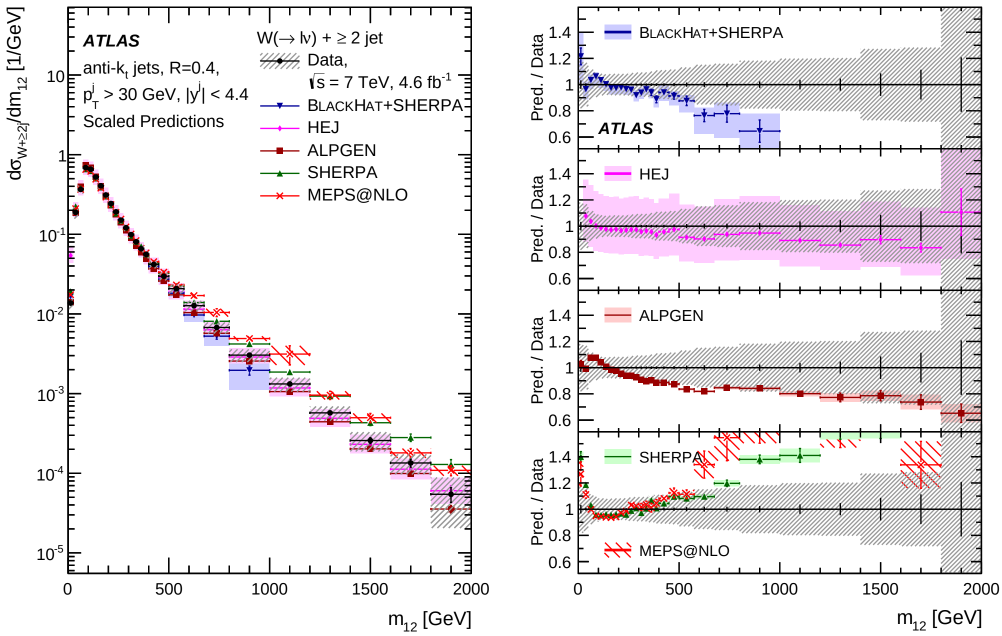
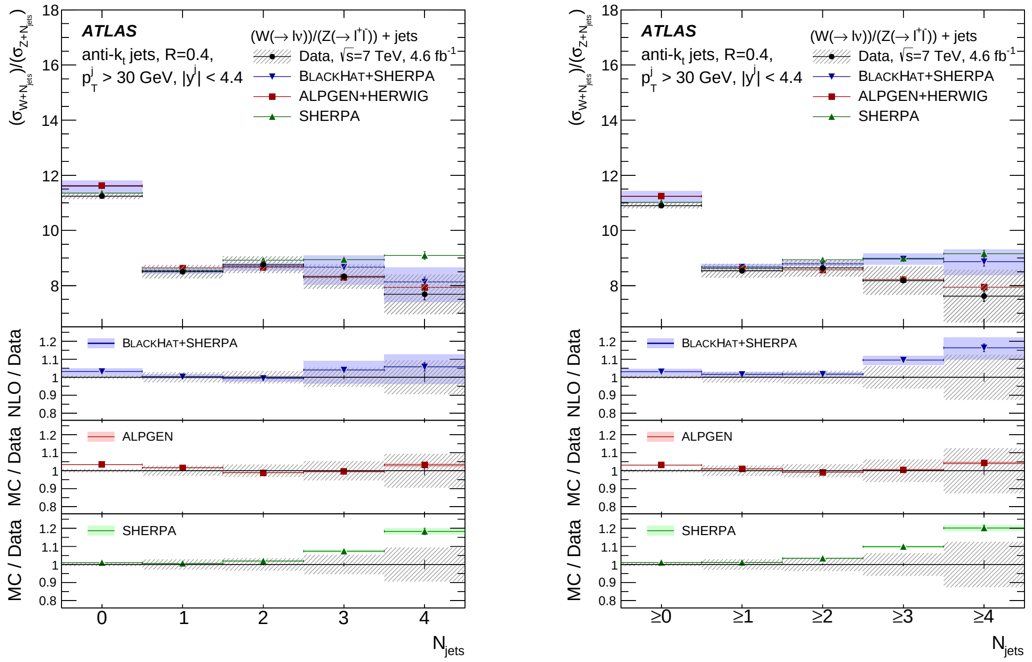


Figure 7.22: Cross section for the production of $W \to l\nu$ as a function of dijet invariant mass, m_{12}. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.



(a) (b)

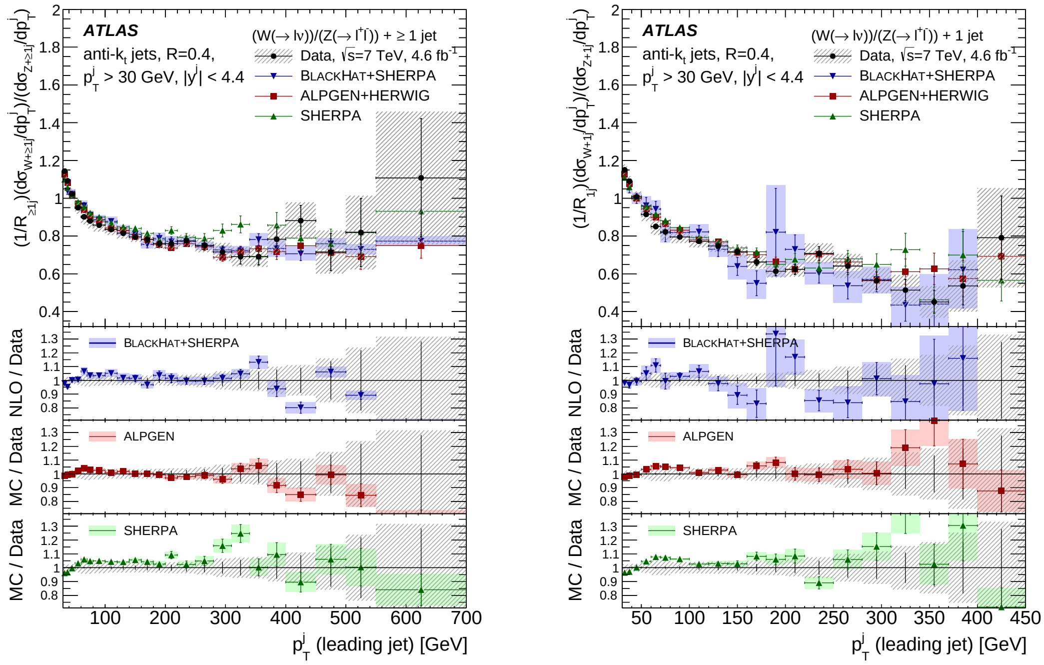
Figure 7.23: R_{jets} as a function of exclusive jet multiplicity (a) and inclusive jet multiplicity (b). Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty.

7.2 R_{jets}

7.2.1 Jet multiplicity

Figure 7.23 shows the ratio measurement as a function of exclusive and inclusive jet multiplicity. The theoretical predictions describe the data reasonably considering the experimental uncertainties. BLACKHAT+SHERPA has a tendency to overestimate the ratio at high jet multiplicity although only by one standard deviation. SHERPA also overestimates the ratio at high jet multiplicity by about 1.5 standard deviations.

In the following differential results the measurement is normalised to the ratio in the corresponding jet multiplicity bin of Figure 7.23 in order to decouple the normalisation differences of Figure 7.23 from the shape differences shown in the following plots.



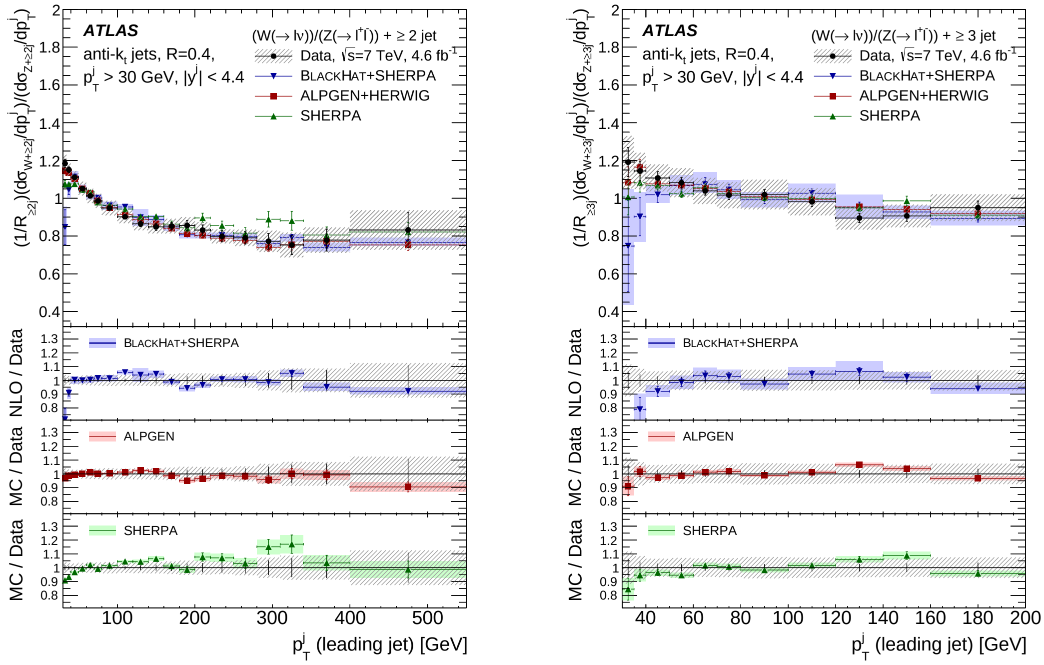
(a) (b)

Figure 7.24: R_{jets} as a function of leading jet p_{T} in events with $N_{\text{jet}} \geq 1$ (a) and leading jet p_{T} in events with $N_{\text{jet}} = 1$ (b). Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty.

7.2.2 Jet transverse momentum

Figure 7.24 shows the ratio measurement as a function of the leading jet p_{T} in events with at least one jet and exactly one jet. In the region below 200 GeV it is observed that R_{jets} falls as jet p_{T} increases due to the differences in the shapes of the W+jets and Z+jets distributions arising from W and Z mass and polarisation differences. These differences affect the parton radiation scale and the decay product kinematics respectively. Across the p_{T} range shown, BLACKHAT+SHERPA, ALPGEN and SHERPA all agree with the data although some small tensions are observed at low p_{T}. This demonstrates that the tensions shown in Figure 7.3 cancel out well when the ratio is formed.

Figure 7.25 shows the ratio measurement as a function of the leading jet p_{T} in events with at least two jets and at least three jets. At low p_{T} ($p_{\text{T}} < 50$ GeV)

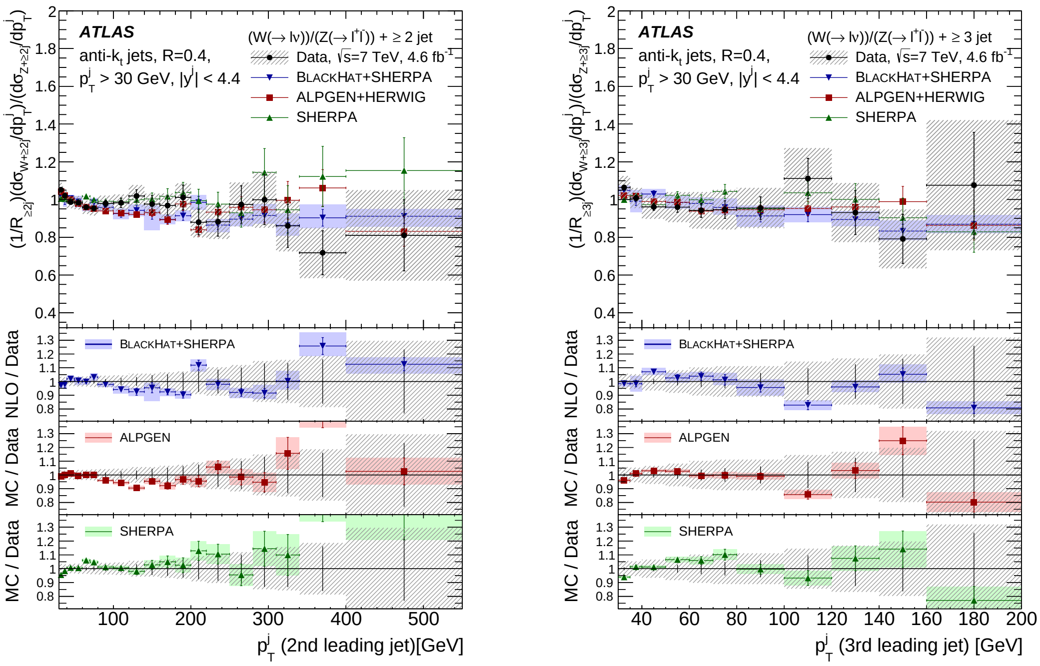


(a) (b)

Figure 7.25: R_{jets} as a function of leading jet p_{T} in events with $N_{\text{jet}} \geq 2$ (a) and leading jet p_{T} in events with $N_{\text{jet}} \geq 3$ (b). Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty.

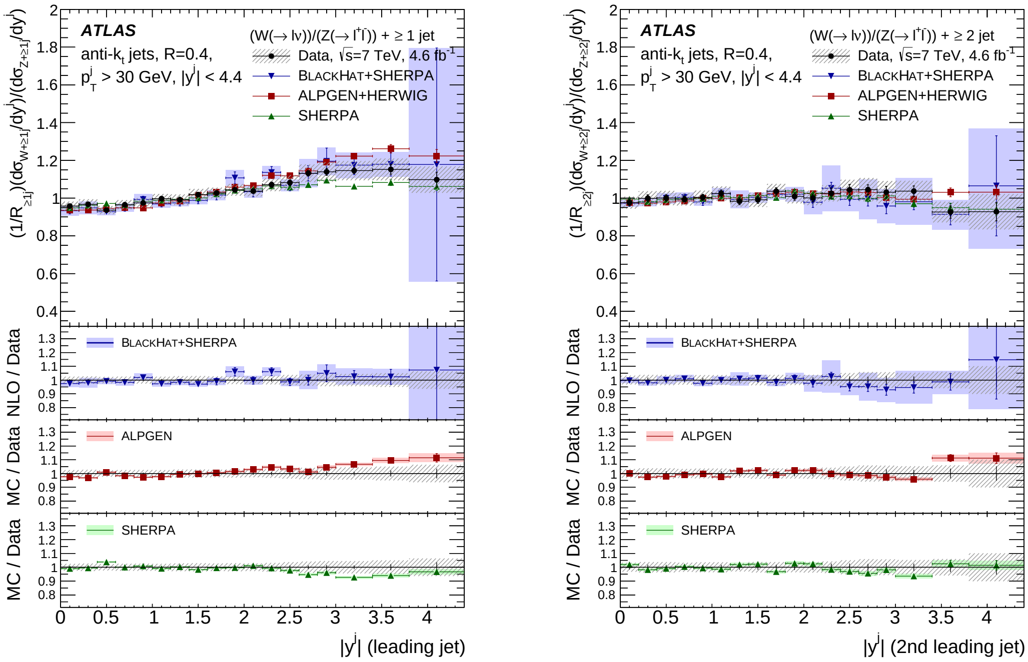
different behaviour of the theoretical predictions is observed with both BLACKHAT+SHERPA and SHERPA underestimating the data. This effect is most pronounced for BLACKHAT+SHERPA.

Figure 7.26 shows the second and third leading jet p_{T} in events with at least two and at least three jets respectively. All three predictions agree well with the measured ratio.



(a) (b)

Figure 7.26: R_{jets} as a function of second leading jet p_{T} in events with $N_{\text{jet}} \geq 2$ (a) and third leading jet p_{T} in events with $N_{\text{jet}} \geq 3$ (b). Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty.

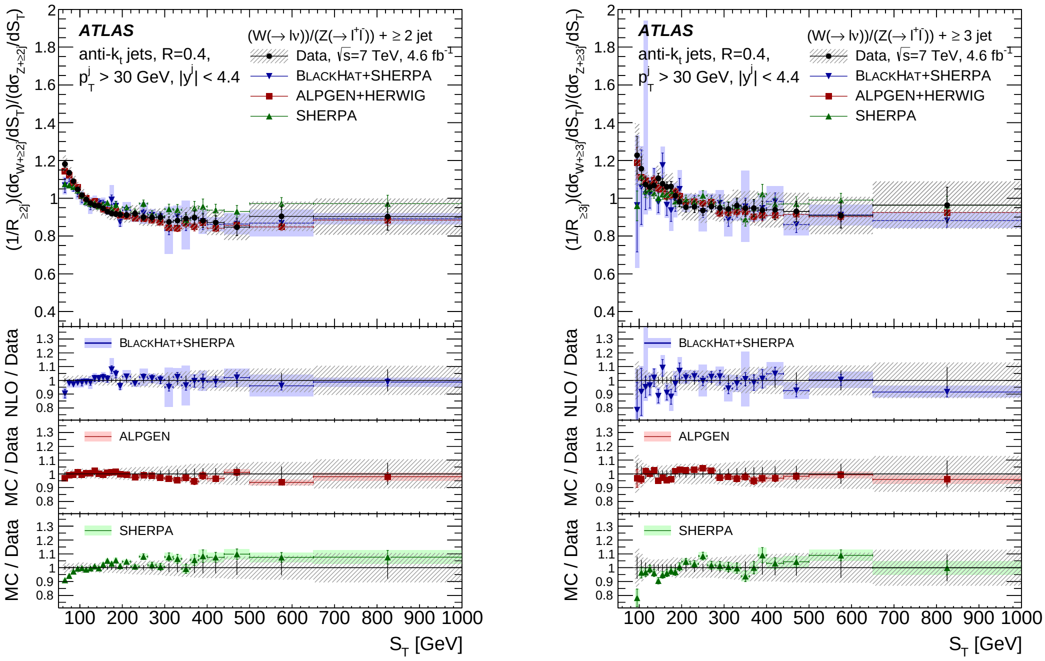


(a) (b)

Figure 7.27: R_{jets} as a function of leading jet rapidity in events with $N_{\text{jet}} \geq 1$ (a) and second leading jet rapidity in events with $N_{\text{jet}} \geq 2$ (b). Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty.

7.2.3 Jet rapidity

Figure 7.27 shows the ratio measurement as a function of the leading and second leading jet rapidities in events with at least one and at least two jets respectively. The predictions generally agree well with the measurement however there are significant differences in the trends exhibited by the predictions which may arise from the effects of different parton showers, the effects of which do not fully cancel in the ratio measurement. The ratio measurement as a function of the third leading jet rapidity in events with at least three jets is shown in Appendix B.



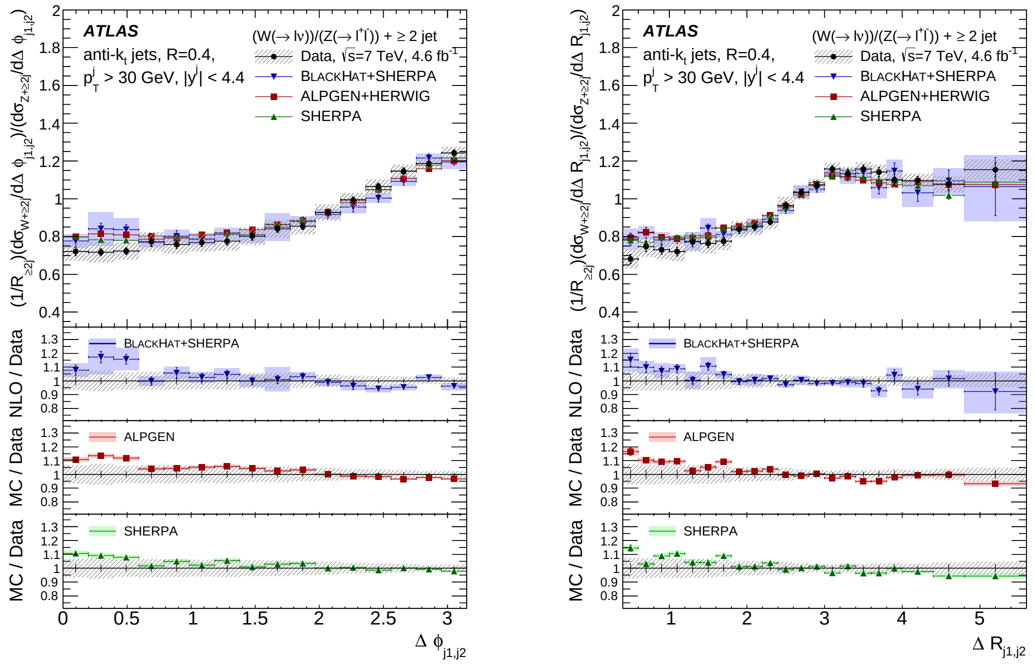
(a) (b)

Figure 7.28: R_{jets} as a function of S_{T} in events with $N_{\text{jet}} \geq 2$ (a) and S_{T} in events with $N_{\text{jet}} \geq 3$ (b). Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty.

7.2.4 Inclusive variables

Figure 7.28 shows the ratio measurement as a function of S_{T} (the scalar sum of jet p_{T} in the event) for events with at least two jets and at least three jets. All predictions agree well with the data up to high values of S_{T}, although tensions are observed at low S_{T}, particularly between the data and the prediction from SHERPA. In particular, the tension observed in Figure 7.15 (albeit for H_{T}) is removed in the ratio measurement, implying that the tensions affect W and Z events equally.

Further plots of the ratio measurement as a function of S_{T} in events with exactly two and exactly three jets are shown in Appendix B.



(a) (b)

Figure 7.29: R_{jets} as a function of $\Delta\phi_{j1,j2}$ (a) and $\Delta R_{j1,j2}$ (b). Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty.

7.2.5 Dijet variables

Figures 7.29 and 7.30 shows the ratio measurement as a function of the azimuthal angular separation ($\Delta\phi_{j1,j2}$), the total angular separation ($\Delta R_{j1,j2}$) and the invariant mass (m_{12}) of the two leading jets in events with at least two jets. Predictions agree well, although tensions at small angular separation and small invariant mass may hint at weak sensitivity to differences in the soft QCD radiation models employed in the various predictions. Again, it is observed that tensions in Figure 7.22 cancel in the ratio measurement.

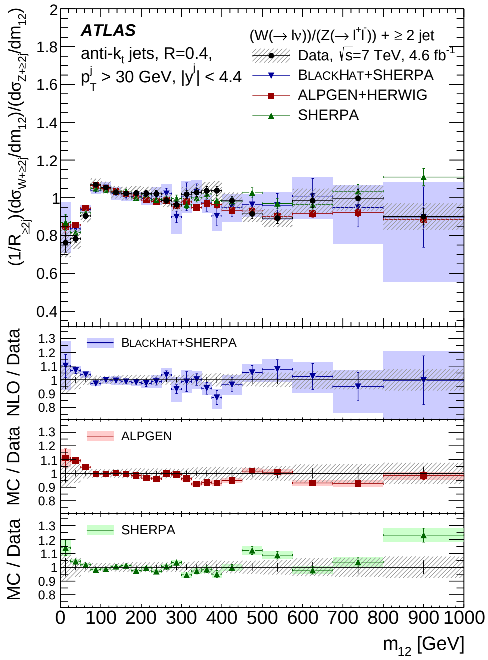


Figure 7.30: R_{jets} as a function of m_{12}. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty.

Chapter 8

Constraints on Parton Distribution Functions

As discussed in Section 1.2, an accurate knowledge of parton distribution functions (PDFs) is essential to predict the cross sections of processes at a hadron collider. Our knowledge of PDFs can be improved by inclusion of additional data which may add additional constraint to global fits. The measurements outlined in this thesis are good candidates for inclusion in such a fit and the possible effects and improvements are investigated in this Chapter.

In order to perform such studies, a reference PDF set must be used which, in this case, is chosen to be a HERAPDF set refitted with a slightly looser parameterisation than is used in the nominal fit. This reference PDF is discussed in Section 8.1. The inclusion of ATLAS $W+$ jets and R_{jets} measurements is described in Section 8.2 and the results are presented in Section 8.3.

8.1 Parton distribution functions from HERA

The starting point for the PDF studies presented here is a PDF fit based on the HERA I PDF set [38, 39]. This uses a parameterisation of the PDFs based on the gluon distribution, $xg(x)$, the valence quark distributions, $xu_v(x)$ and $xd_v(x)$, and the up-type and down-type anti-quark (sea) distributions, $x\bar{U}(x)$ and $x\bar{D}(x)$. In this case, $x\bar{U}(x) = x\bar{u}(x)$ and $x\bar{D}(x) = x\bar{d}(x) + x\bar{s}(x)$. The PDFs are then parameterised as [155]:

$$xg(x) = A_g x^{B_g}(1-x)^{C_g}, \tag{8.1}$$

$$xu_v(x) = A_{u_v} x^{B_{u_v}}(1-x)^{C_{u_v}}\left(1 + E_{u_v}x^2\right), \tag{8.2}$$

$$xd_v(x) = A_{d_v} x^{B_{d_v}}(1-x)^{C_{d_v}}, \tag{8.3}$$

$$x\bar{U}(x) = A_{\bar{U}} x^{B_{\bar{U}}}(1-x)^{C_{\bar{U}}}, \tag{8.4}$$

$$x\bar{D}(x) = A_{\bar{D}} x^{B_{\bar{D}}}(1-x)^{C_{\bar{D}}}. \tag{8.5}$$

The A parameters provide an overall normalisation, the B parameters define the behaviour at small x and the C parameters define the shape as $x \to 1$. The normalisation parameters, A_g, A_{u_v} and A_{d_v}, are constrained by the quark number sum rules,

$$\int_0^1 \mathrm{d}x u_v(x) = 2, \int_0^1 \mathrm{d}x d_v(x) = 1, \tag{8.6}$$

and the momentum sum rule,

$$\int_0^1 \mathrm{d}x x\left(\Sigma(x) + g(x)\right) = 1, \tag{8.7}$$

where $\Sigma(x)$ is the sum over all quark and anti-quark PDFs. In addition, $B_{\bar{U}}$ and $B_{\bar{D}}$ are set to be equal and the strange quark distribution is expressed as $x\bar{s}(x) = f_s x\bar{D}(x)$ with $f_s = 0.31$ as determined from neutrino-induced di-muon production [35, 156]. The constraint $A_{\bar{U}} = A_{\bar{D}}(1 - f_s)$ is applied. Finally, in the nominal parameterisation, $B_{u_v} = B_{d_v}$. This parameterisation defines the 10 parameter HERAPDF parameterisation.

Here, a slightly looser parameterisation is used. The gluon distribution is modified to be

$$xg(x) = A_g x^{B_g}(1-x)^{C_g} - A'_g x^{B'_g}(1-x)^{25}, \tag{8.8}$$

and the constraint $B_{u_v} = B_{d_v}$ is removed. This results in a 13 parameter fit. This parameterisation is used to derive a starting point for inclusion of the $W+\text{jets}$ and R_{jets} measurements using the combined HERA1 dataset as defined in Ref. [38]. This dataset comprises results on neutral and charged current scattering of e^-p and e^+p (a total of four datasets).

Fits are performed using the HeraFitter open source framework [38,39] interfaced to QCDNUM for DGLAP evolution [157] and Minuit for minimisation [158]. The data are fitted using a χ^2 minimisation with

$$\chi^2(m,b) = \sum_i \frac{\left[\mu_i - m_i\left(1-\sum_j \gamma^i_j b_j\right)\right]^2}{\delta^2_{i,\text{unc}} m_i^2 + \delta^2_{i,\text{stat}}\mu_i m_i} + \sum_j b_j^2, \tag{8.9}$$

where μ_i are the experimental data, m_i are the theory predictions, $\delta^2_{i,\text{unc}}$ and $\delta^2_{i,\text{stat}}$ are the relative uncorrelated and statistical uncertainties of the measurement, γ^i_j quantify the sensitivity of measurement i to the correlated systematic source j and b_j are nuisance parameters. The experimental uncertainties are propagated [45,159] to the final PDFs using a Hessian method which examines the shape of the χ^2 around the minimum [36].

The PDFs fitted using the parameterisation and prescription defined above form a starting PDF to which ATLAS measurements can be added to understand the possible additional constraints provided by the inclusion of the ATLAS measurements.

8.2 Inclusion of ATLAS measurements in PDFs

In this study, a subset of the $W+\text{jets}$ and R_{jets} measurements are considered for inclusion in a PDF fit. Leading jet p_{T} and rapidity results from both measurements have been investigated. The poor agreement between NLO QCD predictions and measurement for the $W+\text{jets}$ leading jet p_{T}, discussed in Section 7.1.2, motivates the immediate removal of $W+\text{jets}$ leading jet p_{T} from consideration. As such, these studies are performed including $W+\text{jets}$ leading jet rapidity and R_{jets} leading jet p_{T} and rapidity.

Fits are performed separately in the electron and muon phase spaces, defined in Table 6.2, as the physical origin of some of the correlations is lost in the combination procedure discussed in Section 6.4. The fits are carried out using fast NLO calculation, made possible by APPLGRID [160] using grids generated from MCFM [161,162]. The MCFM predictions are at parton level and so, the hadronisation, underlying event and QED FSR corrections outlined in Chapter 6 are applied.

The treatment of uncertainties as correlated or uncorrelated is summarised in Table 8.1 and is heavily influenced by the treatment of uncertainties in the combination method of Section 6.4. This table also defines the abbreviations assigned to each systematic source which will be used through the rest of this Chapter.

The results of the PDF fit to the ATLAS measurements are shown in Figure 8.1. Reasonable agreement is observed between the MCFM predictions and the measurements and only small variations are observed between the raw prediction and the prediction including the shifts, $\sum_j \gamma^i_j b_j$. Table 8.2 shows the χ^2 results with and without the ATLAS data whilst Figure 8.2 shows the shifts, b_j, which are returned by the χ^2 minimisation. The variation in these shifts is approximately consistent with what one would expect, as confirmed by a χ^2 test for consistency with $b_j = 0$ which returns $\chi^2/\text{ndf} = 1.05$.

Uncertainty	Correlated	Uncorrelated	Abbreviation
Electron	✓	-	ELEC
Muon	✓	-	MUON
JES (NP1)	✓	-	JESN1
JES (NP2)	✓	-	JESN2
JES (NP3)	✓	-	JESN3
JES (NP4)	✓	-	JESN4
JES (NP5)	✓	-	JESN5
JES (NP6)	✓	-	JESN6
JES (η-intercalibration)	✓	-	JESN7
JES (single-hadron response)	✓	-	JESN8
JES (non-closure)	✓	-	JESN9
JES μ dependence	✓	-	JESMU
JES N_{PV} dependence	✓	-	JESNPV
JES Close-by	✓	-	JESCB
JES Flavour Composition	✓	-	JESFC
JES Flavour Response	✓	-	JESFR
JER	✓	-	JER
MET	-	✓	MET
Multi-jet syst.	-	✓	QCD_NST
Multi-jet stat.	-	✓	QCD_STAT
$t\bar{t}$ syst.	✓	-	TOP
$t\bar{t}$ stat.	-	✓	TOP_STAT
Diboson	✓	-	DIB
$W \to \tau\nu$ and $Z \to \tau\tau$	✓	-	TAU
V+jets cross-talk	-	✓	VBKG
Single t	✓	-	STOP
Unfolding model	✓	-	UNFM
Unfolding stat.	-	✓	UNFS
Luminosity	✓	-	LUMI

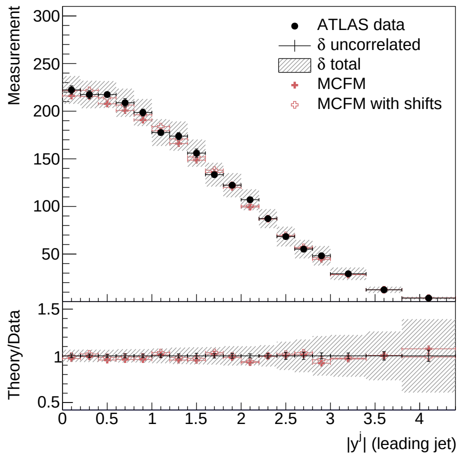
Table 8.1: Schematic of the treatment of systematics in the PDF fitting procedure. Each line represents a single uncertainty and defines whether the uncertainty is treated as correlated or uncorrelated in the PDF fit. Also shown is the abbreviation label assigned to each systematic.

Dataset	HERA only	HERA + ATLAS	HERA + ATLAS (JER unc.)
HERA I, e^-p NC cross section	104/145	105/145	105/145
HERA I, e^+p NC cross section	379/379	379/379	379/379
HERA I, e^-p CC cross section	21/34	21/34	21/34
HERA I, e^+p CC cross section	30/34	31/34	30/34
ATLAS, R_{jets} leading jet p_{T}, electron	-	23/23	16/23
ATLAS, R_{jets} leading jet p_{T}, muon	-	41/23	30/23
ATLAS, $W+$jets leading jet $\|y\|$, electron	-	21/18	5.4/18
ATLAS, $W+$jets leading jet $\|y\|$, muon	-	76/18	8.9/18
ATLAS, R_{jets} leading jet $\|y\|$, electron	-	27/18	17/18
ATLAS, R_{jets} leading jet $\|y\|$, muon	-	37/18	25/18
Correlated χ^2	40	66	53
Total χ^2/dof	573/579	825/697	689/697

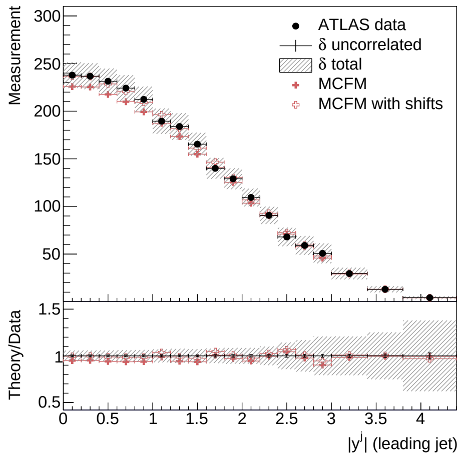
Table 8.2: The χ^2 values separated by dataset for fits to HERA I and ATLAS data. Results are shown fitting only HERA data, HERA and ATLAS data and HERA and ATLAS data but treating the jet energy resolution as uncorrelated rather than correlated.

There are only very small changes in the χ^2 values for HERA I data when including the ATLAS data, however the χ^2 values for ATLAS data are generally above one. Of particular note is the observation that the χ^2 values in the muon channel are systematically higher than the equivalent values in the electron channel. This phenomenon is under investigation and is, as yet, not understood.

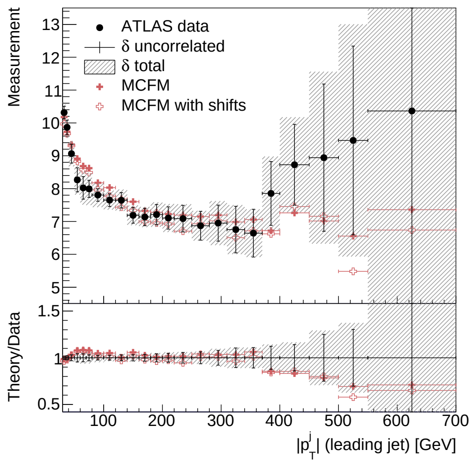
In addition to the results including and excluding ATLAS data, the fit has also been repeated treating the jet energy resolution (JER) as uncorrelated rather than correlated. The treatment as uncorrelated is less correct, however this study demonstrates that the χ^2 improves significantly when making this change. Similar observations have been reported when including inclusive jet [163], dijet [164] and trijet [165] measurements into a PDF fit. It must be noted, however, that the resulting parameters between the two different treatments of JER are entirely compatible within experimental uncertainties.



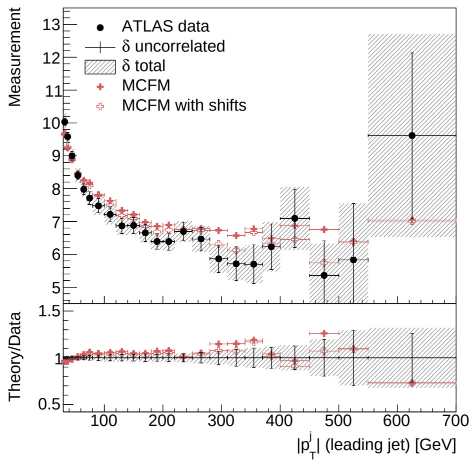
(a) $W+$jets leading jet rapidity, electron



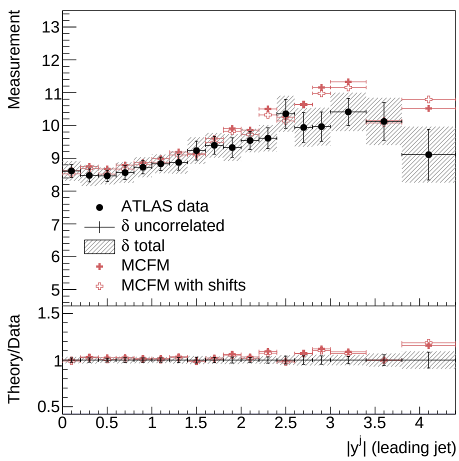
(b) $W+$jets leading jet rapidity, muon



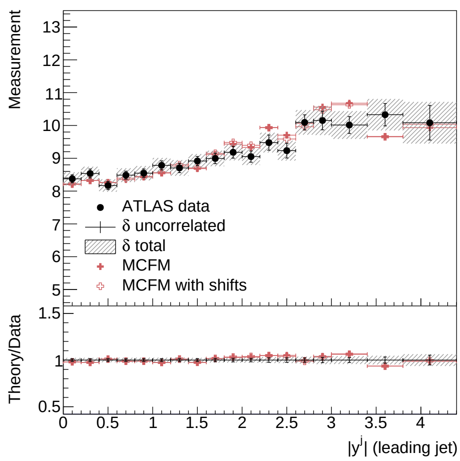
(c) R_{jets} leading jet p_{T}, electron



(d) R_{jets} leading jet p_{T}, muon



(e) R_{jets} leading jet rapidity, electron



(f) R_{jets} leading jet rapidity, muon

Figure 8.1: Comparison of theory predictions taken from MCFM and measured cross sections and ratios as a function of the leading jet p_{T} and rapidity. The theory is shown in its raw form and with the shifts, $\sum_j \gamma^i_j b_j$, of Equation 8.9 applied. Results are shown for electron and muon channels separately.

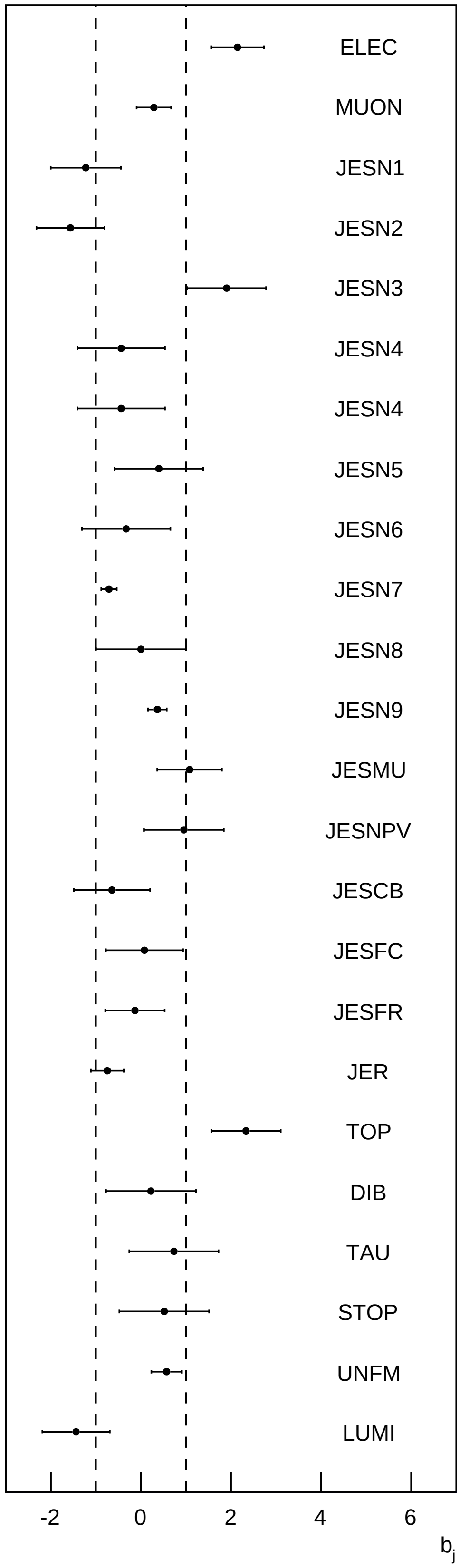


Figure 8.2: The shifts, b_j, returned by the χ^2 minimisation utilised in the HERA+ATLAS PDF fits discussed in the text. The shifts are found to be consistent with zero with $\chi^2/\mathrm{ndf} = 1.05$. The abbreviations used here for the systematic sources are listed in Table 8.1.

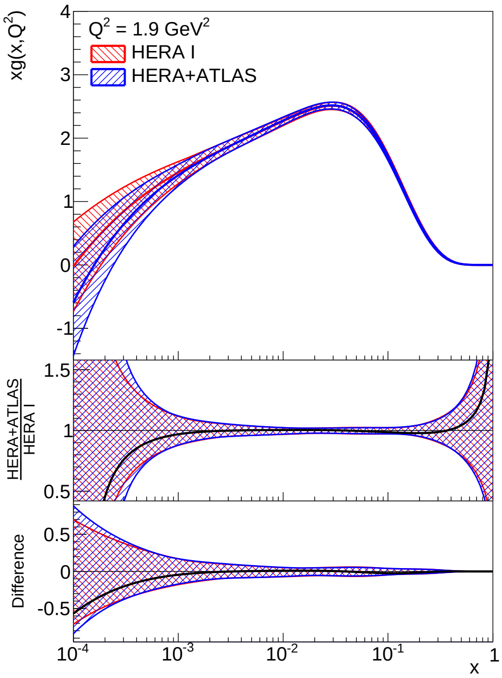


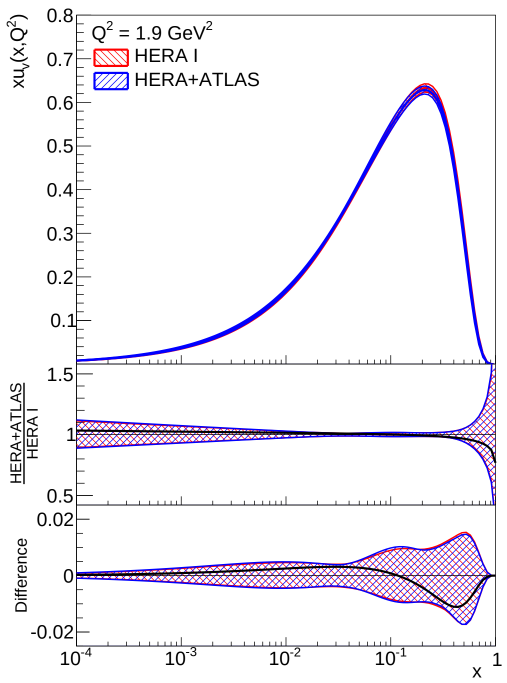
Figure 8.3: The gluon PDF as a function of X at the starting scale of $Q^2 = 1.9$ GeV^2. Results are shown fitting to HERA I data alone (red) and to HERA I and ATLAS $W +$ jets and R_{jets} data (blue) as discussed in the text. The uncertainty band includes the experimental uncertainties on the fitted parameterisation. Also shown is the ratio of the two fits with the relative uncertainties and the difference between the fits with the absolute uncertainties.

8.3 PDF constraints from ATLAS measurements

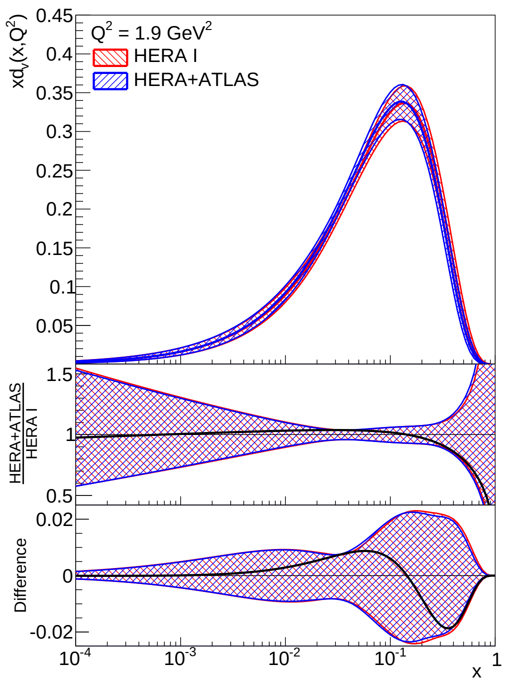
Following the procedures described above, the effects of the ATLAS measurements on PDF fits can be explored by comparing the results of the HERA I fit and the HERA+ATLAS fit. Figures 8.3, 8.4 and 8.5 show the gluon, valence quark and sea quark PDFs respectively, both before and after inclusion of the ATLAS data.

Some changes are observed in the gluon PDF (Figure 8.3) on the inclusion of the ATLAS data although all changes are well within the HERA I uncertainty band. At low x in particular, small tensions between the ATLAS and HERA I datasets serve to increase the uncertainty in the gluon PDF when ATLAS data are included.

Small changes are also observed in the valence quark distributions (Figure 8.4), however again these are easily covered by the HERA I uncertainty.



(a) $u_v(x)$

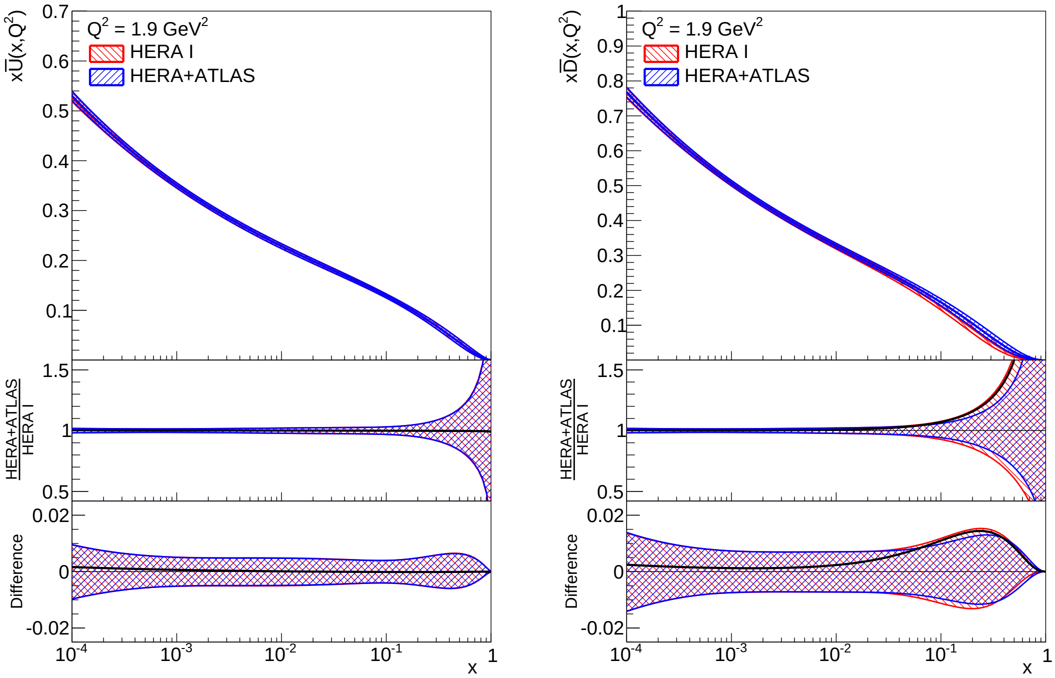


(b) $d_v(x)$

Figure 8.4: The u_v and d_v PDFs as a function of X at the starting scale of $Q^2 = 1.9$ GeV^2. Results are shown fitting to HERA I data alone (red) and to HERA I and ATLAS $W+$jets and R_{jets} data (blue) as discussed in the text. The uncertainty band includes the experimental uncertainties on the fitted parameterisation. Also shown is the ratio of the two fits with the relative uncertainties and the difference between the fits with the absolute uncertainties.

Finally, we look at the sea quark distributions (Figure 8.5). No difference is observed in the u-type sea distribution when including ATLAS data, however a reasonable difference is observed at high x ($x > 0.01$) in the d-type sea distribution. In this region, the inclusion of ATLAS data favours an increase in the $x\bar{D}(x)$ distribution to the level of about 1σ relative to the HERA I PDF. At the same time a small reduction is observed in the experimental uncertainty on the PDF.

These studies show that there is little additional constraint provided by ATLAS $W+$jets and R_{jets} measurements when included in the HERA I PDF fit. Some improvement is seen in $x\bar{D}(x)$ however this is the region that the HERA data does not constrain well. In regions where the HERA I data are constraining, no improvement is observed. The inclusion of these ATLAS measurements in full global fits is likely to provide even less improvement than has been demonstrated here as the global fits



(a) $\bar{U}(x)$ (b) $\bar{D}(x)$

Figure 8.5: The $\bar{U}$ and $\bar{D}$ PDFs as a function of X at the starting scale of $Q^2 = 1.9$ GeV^2. Results are shown fitting to HERA I data alone (red) and to HERA I and ATLAS $W +$ jets and R_{jets} data (blue) as discussed in the text. The uncertainty band includes the experimental uncertainties on the fitted parameterisation. Also shown is the ratio of the two fits with the relative uncertainties and the difference between the fits with the absolute uncertainties.

will already be constraining $x\bar{D}(x)$ better than HERA I data alone. Nevertheless, it is promising that it has been proven possible to include such measurements in a PDF fit without observing large tensions between the fits and the measurements. It should be noted that, although the χ^2 values shown here demonstrate reasonable agreement of QCD calculation with measurement, this is partially by construction given that poorly described distributions (such as leading jet p_{T} in $W +$ jets) have not been included in the fit.

Conclusions

In this thesis, two measurements of vector boson production in association with jets, using 4.6 fb^{-1} of data from the ATLAS detector at the LHC, have been described. The first, W+jets, constitutes a significant increase in the kinematic reach compared to earlier similar measurements, whilst a ratio measurement, R_{jets}, demonstrates significant cancellation of theoretical and experimental uncertainties and includes a wide array of differential distributions which have not been measured previously.

The W + jets measurement has been compared to many state-of-the-art perturbative QCD calculations. Multi-leg Monte Carlo calculations from ALPGEN and SHERPA demonstrate impressive agreement with the measurements up to high jet multiplicities of five or more jets, particularly in the case of ALPGEN. SHERPA is seen to struggle with the prediction of forward jet cross sections ($|y| > 3$) and overestimates inclusive variables such as H_{T} (the scalar sum of the p_{T} of objects in the event) at high values ($H_{\text{T}} > 500$ GeV). Both multi-leg predictions are seen to struggle to model dijet angular variables such as $\Delta\phi$, Δy and ΔR between the jets with the predictions showing very different shapes from both the data and each other. This is reflected in the dijet mass, m_{12}, which is modelled poorly, particularly by SHERPA.

Moving to next-to-leading order (NLO) predictions from BLACKHAT+SHERPA, it is observed that the calculation does a reasonable job modelling jet multiplicity however, a general trend is seen of underestimating the jet p_{T} at high p_{T}, particularly noticeable in the case of the leading jet p_{T}. Agreement improves with increasing jet multiplicity, however this is also where a significant increase in the measurement uncertainty is observed. BlackHat+Sherpa does a surprisingly good job of modelling the leading jet p_{T} in events with exactly one jet. BlackHat+Sherpa has a similar tendency to Sherpa to overestimate the cross section with jets in the forward region. Above a jet multiplicity of two, BlackHat+Sherpa predicts inclusive variables well but missing higher order terms result in an underestimation at high H_{T} for lower jet multiplicities. Finally, an improvement is observed for BlackHat+Sherpa over multi-leg predictions in the case of dijet angular variables however this improvement is not carried through to the dijet invariant mass which BlackHat+Sherpa underestimates at high values.

Approximate next-to-next-to-leading order (nNLO) predictions have also been investigated from BlackHat+Sherpa exclusive sum predictions and LoopSim. Particularly in the case of inclusive variables, such as H_{T}, these approximations are observed to go a long way to improving the agreement with data by the inclusion of otherwise missing higher order terms. This improvement is not reflected in the leading jet p_{T} where only minimal improvement is seen.

NLO Monte Carlo predictions are also shown from the Sherpa MEPS@NLO prescription. The difficulties of Sherpa to model jet rapidities and dijet angular variables are not improved by MEPS@NLO and in fact are exacerbated in some regions. In particular, the Δy and dijet invariant mass become significantly more discrepant from data when using the MEPS@NLO formalism. In contrast, the agreement in inclusive variables is a huge improvement over Sherpa with MEPS@NLO agreeing almost perfectly across the explored kinematic range up to 2 TeV.

Finally, the predictions from a resummation calculation (HEJ) are shown. Although jet p_{T} distributions are poorly predicted, jet rapidity and H_{T} distributions show fair agreement. The most striking result is in the dijet mass where HEJ is the only prediction explored which performs well up to the top of the kinematic range (up to 2 TeV). This high dijet mass region is exactly the region HEJ is designed for as its resummation procedure is valid in regions with large rapidity separations.

The ratio measurement, R_{jets}, is compared to a reduced set of these predictions with results shown for BlackHat+Sherpa, Alpgen and Sherpa. The measurement itself demonstrates impressive cancellation of uncertainties with the experimental uncertainty reduced by a factor of up to 10. Across the kinematic region explored, NLO and multi-leg predictions do a fair job, demonstrating that many of the mis-modelling issues identified in $W+\text{jets}$ are reduced in the formation of the ratio. Nevertheless, some small tensions have been observed at low jet p_{T} and high jet rapidity which may well be arising from differences in the parton showers of the various predictions. Particularly striking is the good agreement of the ratio in inclusive distributions, lending weight to the methods used in BSM searches which map $Z/\gamma^* + \text{jets}$ to $W + \text{jets}$ to provide data-driven estimates of the latter.

In this thesis, the agreement with various state-of-the-art predictions has been investigated. None of the predictions manage to predict the data well across the large range of distributions shown and it is clear that in many regions some predictions are more suitable than others whilst being less suitable elsewhere. It is intended that this measurement, the results of which are available on HepData[1], will be used by the theory community to further understand and develop the theoretical calculations (which will prove invaluable to extracting the most accurate and interesting information from the large datasets expected from the LHC) and provide a reference for experimentalists to better choose calculations suitable in their particular phase space.

Furthermore, a first study has also been performed investigating the possibility of including $W + \text{jets}$ and R_{jets} results in parton distribution function (PDF) fits.

[1] `http://hepdata.cedar.ac.uk/`.

Initial studies have shown a promising agreement between calculation and data, although this requires careful selection of the results to include. Comparing to the HERA I PDF, particular improvement of the constraints on the down-type sea have been shown however this improvement is expected to be decreased when including these data in a global PDF which already constrains this PDF using non-HERA data. To extract the most from these data, it is expected that further improvement in theoretical calculations will be required as the distributions most expected to effect PDFs, for example jet p_{T}, are currently too poorly modelled for inclusion in fits.

Extensions of such measurements are already possible (and underway) using the 20.3 fb^{-1} dataset taken with ATLAS in 2012 at $\sqrt{s} = 8$ TeV. Significant improvements in the $W +$ jets measurement are not expected over the phase space already explored as the measurement is already systematically limited. The larger dataset will however allow the measurement to be pushed up to higher p_{T} and therefore stretch theoretical predictions even further. By contrast, R_{jets} is a statistically limited measurement using the 2011 dataset. The uncertainty is statistically dominated above 250 GeV in leading jet p_{T} and so would benefit from a larger dataset.

In updated measurements it is also hoped that the separation of the $W +$ jets measurement into W^+ and W^- will be carried out and in addition the measurement of R_{jets}^{W}, the ratio of W^+ and W^- production in association with jets. Such a measurement is expected to give information on the u/d ratio in the proton PDF [166].

With the LHC about to start up again at $\sqrt{s} = 13$ TeV, early measurements of vector boson production in association with jets at detector level will form an important validation of the performance and simulation of the ATLAS detector and once a significant dataset has been collected a further extension of such measurements will be possible.

In particular, in Run I the region where expected electroweak (EW) corrections to predictions become large ($> 20\%$) is aligned with the region where data statistics began to run out and the size of EW corrections is, at most, the same as the measurement uncertainty [167, 168]. In Run II it is expected that such effects will become accessible and investigation of electroweak corrections in $V+$jets will be possible. This investigation is particularly interesting in $V+$jets as opposed to, for example inclusive Z, as in $V+$jets there is an interplay between NLO QCD and EW corrections. This interplay may also require the measurement of R_{jets} to fully disentangle QCD and EW corrections at high p_{T} as QCD corrections are expected to cancel more than EW corrections [166].

References

[1] ATLAS Collaboration. *Eur. Phys. J.*, **C74**, 12 (2014) 3168. http://arxiv.org/abs/1408.6510

[2] ATLAS Collaboration. *Eur. Phys. J.*, **C75**, 2 (2015) 82. http://arxiv.org/abs/1409.8639

[3] ATLAS Collaboration. Tech. Rep. ATLAS-CONF-2013-083, CERN, Geneva (2013). http://cds.cern.ch/record/1570994

[4] ATLAS Collaboration. *Phys. Lett.*, **B716** (2012) 1–29. http://arxiv.org/abs/1207.7214

[5] CMS Collaboration. *Phys. Lett.*, **B716**, 1 (2012) 30–61. http://www.sciencedirect.com/science/article/pii/S0370269312008581

[6] CDF Collaboration. *Phys. Rev.*, **D77** (2008) 011108. http://arxiv.org/abs/0711.4044

[7] D0 Collaboration. *Phys. Rev.*, **D88** (2013) 092001. http://arxiv.org/abs/1302.6508

[8] ATLAS Collaboration. *Phys. Rev.*, **D85** (2012) 092002. http://arxiv.org/abs/1201.1276

[9] CMS Collaboration. *Phys. Lett.*, **B741** (2015) 12–37. http://arxiv.org/abs/1406.7533

[10] E. Abouzaid & H. J. Frisch. *Phys. Rev.*, **D68** (2003) 033014. http://arxiv.org/abs/hep-ph/0303088

[11] ATLAS Collaboration. *Phys. Lett.*, **B708** (2012) 221–240. http://arxiv.org/abs/1108.4908

[12] CMS Collaboration. *JHEP*, **1201** (2012) 010. http://arxiv.org/abs/1110.3226

[13] M. E. Peskin & D. V. Schroeder. "An Introduction to Quantum Field Theory". Westview Press (1995).

[14] I. J. R. Aitchison & A. J. G. Hey. "Gauge Theories in Particle Physics". Taylor and Francis, third edition (2003).

[15] D. Griffiths. "Introduction to Elementary Particles". Wiley-VCH, second edition (2008).

[16] F. Halzen & A. D. Martin. "Quarks and Leptons". Wiley (1984).

[17] K. A. Olive et al. *Chin. Phys.*, **C38**, 9 (2014) 090001. http://iopscience.iop.org/1674-1137/38/9/090001/

[18] J. Goldstone, A. Salam & S. Weinberg. *Phys. Rev.*, **127** (1962) 965–970. http://journals.aps.org/pr/abstract/10.1103/PhysRev.127.965

[19] L. G. Sheldon. *Nuclear Physics*, **22**, 4 (1961) 579–588. http://www.sciencedirect.com/science/article/pii/0029558261904692

[20] A. Salam & J.C. Ward. *Phys. Lett.*, **13**, 2 (1964) 168–171. http://www.sciencedirect.com/science/article/pii/0031916364907115

[21] S. Weinberg. *Phys. Rev. Lett.*, **19** (1967) 1264–1266. http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.19.1264

[22] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes & R. P. Hudson. *Phys. Rev.*, **105** (1957) 1413–1415. http://journals.aps.org/pr/abstract/10.1103/PhysRev.105.1413

[23] F. Englert & R. Brout. *Phys. Rev. Lett.*, **13** (1964) 321–323. http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.13.321

[24] P.W. Higgs. *Phys. Lett.*, **12**, 2 (1964) 132–133. http://www.sciencedirect.com/science/article/pii/0031916364911369

[25] P. W. Higgs. *Phys. Rev. Lett.*, **13** (1964) 508–509. http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.13.508

[26] K. G. Wilson. *Phys. Rev.*, **D10** (1974) 2445–2459. http://journals.aps.org/prd/abstract/10.1103/PhysRevD.10.2445

[27] R. K. Ellis, W. J. Stirling & B. R. Webber. "QCD and Collider Physics". Cambridge University Press (2003).

[28] G. 't Hooft & M. Veltman. *Nuclear Physics*, **B44**, 1 (1972) 189–213. http://www.sciencedirect.com/science/article/pii/0550321372902799

[29] R. Devenish & A. Cooper-Sarkar. "Deep Inelastic Scattering". Oxford University Press (2004).

[30] D. J. Gross & F. Wilczek. *Phys. Rev.*, **D8** (1973) 3633–3652. `http://journals.aps.org/prd/abstract/10.1103/PhysRevD.8.3633`

[31] H. D. Politzer. *Phys. Rev. Lett.*, **30** (1973) 1346–1349. `http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.30.1346`

[32] G. Altarelli & G. Parisi. *Nucl. Phys.*, **B126** (1977) 298.

[33] Y. L. Dokshitzer. *Sov. Phys. JETP*, **46** (1977) 641–653. `http://inspirehep.net/record/126153`

[34] V. N. Gribov & L. N. Lipatov. *Sov. J. Nucl. Phys.*, **15** (1972) 438–450. `https://cds.cern.ch/record/427157?ln=en`

[35] A.D. Martin, W.J. Stirling, R.S. Thorne & G. Watt. *Eur. Phys. J.*, **C63** (2009) 189–285. `http://arxiv.org/abs/0901.0002`

[36] H.-L. Lai et al. *Phys. Rev.*, **D82** (2010) 074024. `http://arxiv.org/abs/1007.2241`

[37] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio et al. *Nucl. Phys.*, **B867** (2013) 244–289. `http://arxiv.org/abs/1207.1303`

[38] F. D. Aaron et al. *JHEP*, **1001** (2010) 109. `http://arxiv.org/abs/0911.0884`

[39] F. D. Aaron et al. *Eur. Phys. J.*, **C64** (2009) 561–587. `http://arxiv.org/abs/0904.3513`

[40] A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Hoche et al. *Phys. Rept.*, **504** (2011) 145–233. `http://arxiv.org/abs/1101.2599`

[41] J. M. Campbell, J. W. Huston & W. J. Stirling. *Rept. Prog. Phys.*, **70** (2007) 89. `http://arxiv.org/abs/hep-ph/0611148`

[42] M. L. Mangano et al. *JHEP*, **0307** (2003) 001. `http://arxiv.org/abs/hep-ph/0206293`

[43] G. Corcella et al. *JHEP*, **0101** (2001) 010. `http://arxiv.org/abs/hep-ph/0011363`

[44] J. Butterworth, Jeffrey R. Forshaw & M. Seymour. *Z. Phys. C*, **72** (1996) 637–646. `http://arxiv.org/abs/hep-ph/9601371`

[45] J. Pumplin et al. *JHEP*, **0207** (2002) 012. `http://arxiv.org/abs/hep-ph/0201195`

[46] ATLAS Collaboration. Tech. Rep. ATL-PHYS-PUB-2011-008, CERN, Geneva (2011). `http://cds.cern.ch/record/1345343`

[47] S. Hoeche, F. Krauss, N. Lavesson, L. Lonnblad, M. Mangano et al. *Proceedings of the HERA and the LHC Workshop, CERN/DESY 2004/2005* (2006). http://arxiv.org/abs/hep-ph/0602031

[48] T. Gleisberg et al. *JHEP*, **0902** (2009) 007. http://arxiv.org/abs/0811.4622

[49] S. Hoeche, F. Krauss, M. Schonherr & F. Siegert. *JHEP*, **1304** (2013) 027. http://arxiv.org/abs/1207.5030

[50] S. Catani, F. Krauss, R. Kuhn & B.R. Webber. *JHEP*, **0111** (2001) 063. http://arxiv.org/abs/hep-ph/0109231

[51] L. Lonnblad. *JHEP*, **0205** (2002) 046. http://arxiv.org/abs/hep-ph/0112284

[52] P. Golonka & Z. Was. *Eur. Phys. J.*, **C45** (2006) 97–107. http://arxiv.org/abs/hep-ph/0506026

[53] D. R. Yennie, S. Frautschi & H. Suura. *Annals Phys.*, **13** (1961) 379–452. http://www.sciencedirect.com/science/article/pii/0003491661901518

[54] Geant4 Collaboration, S. Agostinelli and others. *Nucl. Instrum. Meth. A*, **506** (2003) 250–303. http://www.sciencedirect.com/science/article/pii/S0168900203013688

[55] T. Sjöstrand, S. Mrenna & P. Skands. *JHEP*, **0605** (2006) 026. http://arxiv.org/abs/hep-ph/0603175

[56] C.F. Berger et al. *Phys. Rev.*, **D80** (2009) 074036. http://arxiv.org/abs/0907.1984

[57] C.F. Berger et al. *Phys. Rev. Lett.*, **106** (2011) 092001. http://arxiv.org/abs/1009.2338

[58] Z. Bern et al. *Phys. Rev.*, **D88** (2013) 014025. http://arxiv.org/abs/1304.1253

[59] J. Alcaraz Maestre et al. *Proceedings of the 2011 Les Houches workshop* (2012). http://arxiv.org/abs/1203.6803

[60] M. Rubin, G. P. Salam & S. Sapeta. *JHEP*, **1009** (2010) 084. http://arxiv.org/abs/1006.2144

[61] J. R. Andersen, T. Hapola & J. M. Smillie. *JHEP*, **1209** (2012) 047. http://arxiv.org/abs/1206.6763

[62] J. R. Andersen & J. M. Smillie. *JHEP*, **1001** (2010) 039. http://arxiv.org/abs/0908.2786

[63] L. Evans & P. Bryant. *JINST*, **3**, 08 (2008) S08001. http://iopscience.iop.org/1748-0221/3/08/S08001

[64] ATLAS Collaboration. *JINST*, **3** (2008) S08003. http://iopscience.iop.org/1748-0221/3/08/S08003

[65] CMS Collaboration. *JINST*, **3** (2008) S08004. http://iopscience.iop.org/1748-0221/3/08/S08004

[66] LHC*b* Collaboration. *JINST*, **3** (2008) S08005. http://iopscience.iop.org/1748-0221/3/08/S08005

[67] ALICE Collaboration. *JINST*, **3** (2008) S08002. http://iopscience.iop.org/1748-0221/3/08/S08002

[68] D. Hall & C. Hays. Ph.D. thesis, Oxford University, Oxford (2014). Presented 29 Aug 2014. http://cds.cern.ch/record/1950648

[69] ATLAS Collaboration. *Eur. Phys. J.*, **C73** (2013) 2518. http://arxiv.org/abs/1302.4393

[70] "ATLAS luminosity public results page". Accessed 19-02-2015. http://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults

[71] ATLAS Collaboration. Tech. Rep. ATLAS-CONF-2011-116, CERN, Geneva (2011). http://cds.cern.ch/record/1376384

[72] ATLAS Collaboration. *Eur. Phys. J.*, **C71** (2011) 1630. http://arxiv.org/abs/1101.2185

[73] ATLAS Collaboration. "ATLAS inner detector: Technical Design Report, 1". Technical Design Report ATLAS. CERN, Geneva (1997). http://cds.cern.ch/record/331063

[74] ATLAS Collaboration. "ATLAS inner detector: Technical Design Report, 2". Technical Design Report ATLAS. CERN, Geneva (1997). http://cds.cern.ch/record/331064

[75] ATLAS Collaboration. "ATLAS pixel detector: Technical Design Report". Technical Design Report ATLAS. CERN, Geneva (1998). http://cds.cern.ch/record/381263

[76] A. Abdesselam et al. *Nucl. Instrum. Meth.*, **A568**, 2 (2006) 642–671. http://www.sciencedirect.com/science/article/pii/S016890020601388X

[77] A. Abdesselam et al. *Nucl Instrum. Meth.*, **A575**, 3 (2007) 353–389. http://www.sciencedirect.com/science/article/pii/S0168900207003270

[78] ATLAS TRT Collaboration. *JINST*, **3**, 02 (2008) P02013. http://iopscience.iop.org/1748-0221/3/02/P02013

[79] B. Dolgoshein. *Nucl. Instrum. Meth.*, **A252** (1986) 137–144. http://www.sciencedirect.com/science/article/pii/0168900286911745

[80] "ATLAS liquid-argon calorimeter: Technical Design Report". Technical Design Report ATLAS. CERN, Geneva (1996). http://cds.cern.ch/record/331061

[81] "ATLAS tile calorimeter: Technical Design Report". Technical Design Report ATLAS. CERN, Geneva (1996). http://cds.cern.ch/record/331062

[82] B. Aubert et al. *Nucl. Instrum. Meth.*, **A558**, 2 (2006) 388–418. http://www.sciencedirect.com/science/article/pii/S016890020502382X

[83] ATLAS Electromagnetic Liquid Argon Endcap Calorimeter Group. *Journal of Instrumentation*, **3**, 06 (2008) P06002. http://iopscience.iop.org/1748-0221/3/06/P06002

[84] M. L. Andrieux et al. *Nucl. Instrum. Meth.*, **479**, 23 (2002) 316–333. http://www.sciencedirect.com/science/article/pii/S0168900201009433

[85] M. J. Varanda et al. *IEEE Transactions on Nuclear Science*, **48**, 3 (2001) 367–371. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=940082

[86] J. Abdallah et al. Tech. Rep. ATL-TILECAL-PUB-2008-001, CERN, Geneva (2007). http://cds.cern.ch/record/1071921

[87] D. M. Gingrich et al. *Journal of Instrumentation*, **2**, 05 (2007) P05005. http://iopscience.iop.org/1748-0221/2/05/P05005

[88] A. Artamonov et al. *Journal of Instrumentation*, **3**, 02 (2008) P02010. http://iopscience.iop.org/1748-0221/3/02/P02010

[89] S. Palestini. *Nuclear Physics - Proceedings Supplements*, **B125**, 0 (2003) 337–345. http://www.sciencedirect.com/science/article/pii/S0920563203910139

[90] A. Aloisio et al. *Nucl. Instrum. Meth.*, **A535**, 12 (2004) 265–271. http://www.sciencedirect.com/science/article/pii/S0168900204016511

[91] ATLAS Collaboration. *Eur. Phys. J.*, **C72** (2012) 1849. http://arxiv.org/abs/1110.1530

[92] "ATLAS level-1 trigger: Technical Design Report". Technical Design Report ATLAS. CERN, Geneva (1998). http://cds.cern.ch/record/381429

[93] P. Jenni, M. Nessi, M. Nordberg & K. Smith. "ATLAS high-level trigger, data-acquisition and controls: Technical Design Report". Technical Design Report ATLAS. CERN, Geneva (2003). http://cds.cern.ch/record/616089

[94] ATLAS Collaboration. submitted to EPJC. http://arxiv.org/abs/1412.7086

[95] ATLAS Collaboration. *Eur. Phys. J.*, **C74**, 10 (2014) 3071. http://arxiv.org/abs/1407.5063

[96] T. Cornelissen, M. Elsing, I. Gavrilenko, W. Liebig, E. Moyse & A. Salzburger. *Journal of Physics: Conference Series*, **119**, 3 (2008) 032014. http://iopscience.iop.org/1742-6596/119/3/032014

[97] R. Fruhwirth. *Nucl. Instrum. Meth.*, **A262** (1987) 444–450. http://www.sciencedirect.com/science/article/pii/0168900287908874

[98] ATLAS Collaboration. Tech. Rep. ATLAS-CONF-2010-069, CERN, Geneva (2010). http://cds.cern.ch/record/1281344

[99] R. Fruhwirth, W. Waltenberger & P. Vanlaer. Tech. Rep. CMS-NOTE-2007-008, CERN, Geneva (2007). http://cds.cern.ch/record/1027031

[100] ATLAS Collaboration. Tech. Rep. ATLAS-CONF-2012-048, CERN, Geneva (2012). http://cds.cern.ch/record/1450089

[101] ATLAS Collaboration. *Eur. Phys. J.*, **C74**, 7 (2014) 2941. http://arxiv.org/abs/1404.2240

[102] ATLAS Collaboration. *Eur. Phys. J.*, **C72** (2012) 1909. http://arxiv.org/abs/1110.3174

[103] ATLAS Collaboration. Tech. Rep. ATL-PHYS-PUB-2011-006, CERN, Geneva (2011). http://cds.cern.ch/record/1345327

[104] M. Aharrouche et al. *Nucl. Instrum. Meth.*, **A568** (2006) 601–623. http://arxiv.org/abs/physics/0608012

[105] S. Laplace & J-B de Vivie. Tech. Rep. ATL-COM-PHYS-2012-467, CERN, Geneva (2012). http://cds.cern.ch/record/1444890

[106] ATLAS Collaboration. *Phys. Rev.*, **D85** (2012) 072004. http://arxiv.org/abs/1109.5141

[107] Tech. Rep. ATLAS-CONF-2012-099, CERN, Geneva (2012). http://cds.cern.ch/record/1462601/

[108] ATLAS Collaboration. *Eur. Phys. J.*, **C74**, 11 (2014) 3130. http://arxiv.org/abs/1407.3935

[109] ATLAS Collaboration. *Eur. Phys. J.*, **C74**, 9 (2014) 3034. http://arxiv.org/abs/1404.4562

[110] "ATLAS muon public results page". Accessed 09-01-2015. http://twiki.cern.ch/twiki/bin/view/AtlasPublic/MuonPerformancePublicPlotsOld

[111] G. P. Salam. *Eur. Phys. J.*, **C67** (2010) 637–686. http://arxiv.org/abs/0906.1833

[112] M. Cacciari, G. P. Salam & G. Soyez. *JHEP*, **0804** (2008) 063. http://arxiv.org/abs/0802.1189

[113] M. Cacciari & G. P. Salam. *Phys. Lett.*, **B641** (2006) 57. http://arxiv.org/abs/0802.1189

[114] ATLAS Collaboration. *Eur. Phys. J.*, **C73** (2013) 2304. http://arxiv.org/abs/1112.6426

[115] W. Lampl et al. Tech. Rep. ATLAS-LARG-PUB-2008-002, CERN, Geneva (2008). http://cds.cern.ch/record/1099735

[116] ATLAS Collaboration. *Eur.Phys.J.*, **C75**, 1 (2015) 17. http://arxiv.org/abs/1406.0076

[117] ATLAS Collaboration. *Eur. Phys. J.*, **C72** (2012) 1844. http://arxiv.org/abs/1108.5602

[118] ATLAS Collaboration. Tech. Rep. ATLAS-CONF-2012-123, CERN, Geneva (2012). http://cds.cern.ch/record/1473426/

[119] ATLAS Collaboration. Tech. Rep. ATLAS-CONF-2012-142, CERN, Geneva (2012). http://cds.cern.ch/record/1485531

[120] T Barillari et al. Tech. Rep. ATL-LARG-PUB-2009-001-2, CERN, Geneva (2008). http://cds.cern.ch/record/1112035

[121] B. P. Kersevan & E. Richter-Was. *Comput. Phys. Commun.*, **184** (2013) 919–985. http://arxiv.org/abs/hep-ph/0405247

[122] A. Sherstnev & R. Thorne. *Eur. Phys. J.*, **C55** (2008) 553–575. http://arxiv.org/abs/0711.2473

[123] K. Melnikov & F. Petriello. *Phys. Rev.*, **D74** (2006) 114017. http://arxiv.org/abs/hep-ph/0609070

[124] M. Cacciari et al. *Phys. Lett.*, **B710** (2012) 612–622. http://arxiv.org/abs/1111.5869

[125] P. Baernreuther, M. Czakon & A. Mitov. *Phys. Rev. Lett.*, **109** (2012) 132001. http://arxiv.org/abs/1204.5201

[126] M. Czakon & A. Mitov. *JHEP*, **1212** (2012) 054. http://arxiv.org/abs/1207.0236

[127] M. Czakon & A. Mitov. *JHEP*, **1301** (2013) 080. http://arxiv.org/abs/1210.6832

[128] M. Czakon, P. Fiedler & A. Mitov. *Phys. Rev. Lett.*, **110** (2013) 252004. http://arxiv.org/abs/1303.6254

[129] M. Czakon & A. Mitov. *Comput. Phys. Commun.*, **185** (2014) 2930. http://arxiv.org/abs/1112.5675

[130] N. Kidonakis. *Phys. Rev.*, **D83** (2011) 091503. http://arxiv.org/abs/1103.2792

[131] N. Kidonakis. *Phys. Rev.*, **D81** (2010) 054028. http://arxiv.org/abs/1001.5034

[132] N. Kidonakis. *Phys. Rev.*, **D82** (2010) 054018. http://arxiv.org/abs/1005.4451

[133] J. M. Campbell, R. K. Ellis & C. Williams. *JHEP*, **1107** (2011) 018. http://arxiv.org/abs/1105.0020

[134] S. Alioli, P. Nason, C. Oleari & E. Re. *JHEP*, **1006** (2010) 043. http://arxiv.org/abs/1002.2581

[135] ATLAS Collaboration. *JHEP*, **1307** (2013) 032. http://arxiv.org/abs/1304.7098

[136] ATLAS Collaboration. Tech. Rep. ATLAS-CONF-2011-102, CERN, Geneva (2011). http://cds.cern.ch/record/1369219

[137] ATLAS Collaboration. Tech. Rep. ATL-COM-PHYS-2013-590, CERN, Geneva (2013). http://cds.cern.ch/record/1545593

[138] ATLAS Collaboration. *Eur. Phys. J.*, **C72** (2012) 2211. http://arxiv.org/abs/1206.2135

[139] ATLAS Collaboration. *Phys. Lett.*, **B711** (2012) 244–263. http://arxiv.org/abs/1201.1889

[140] ATLAS Collaboration. Tech. Rep. ATLAS-CONF-2010-017, CERN, Geneva (2010). http://cds.cern.ch/record/1277648

[141] ATLAS Collaboration. *Eur. Phys. J.*, **C73**, 3 (2013) 2306. http://arxiv.org/abs/1210.6210

[142] G. D'Agostini. *Nucl. Instrum. Meth. A*, **362** (1995) 487–498. http://www.sciencedirect.com/science/article/pii/016890029500274X

[143] T. Adye. arXiv:1105.1160 (2011). http://arxiv.org/abs/1105.1160

[144] A. Glazov. *AIP Conf. Proc.*, **792** (2005) 237–240. http://scitation.aip.org/content/aip/proceeding/aipcp/10.1063/1.2122026

[145] H1 Collaboration. *Eur. Phys. J.*, **C63** (2009) 625–678. http://arxiv.org/abs/0904.0929

[146] P. Skands. *Phys. Rev.*, **D82** (2010) 074018. http://arxiv.org/abs/1005.3457

[147] ATLAS Collaboration. http://arxiv.org/abs/1412.2641

[148] M. Chiesa, G. Montagna, L. Barzè, M. Moretti, O. Nicrosini et al. *Phys. Rev. Lett.*, **111**, 12 (2013) 121801. http://arxiv.org/abs/1305.6837

[149] R. Boughezal, C. Focke & X. Liu. arXiv:1501.01059 (2015). http://arxiv.org/abs/1501.01059

[150] ATLAS Collaboration. *JHEP*, **1310** (2013) 130. http://arxiv.org/abs/1308.1841

[151] ATLAS Collaboration. *JHEP*, **1409** (2014) 015. http://arxiv.org/abs/1406.1122

[152] ATLAS Collaboration. submitted to EPJC. http://arxiv.org/abs/1503.08988

[153] ATLAS Collaboration. *JHEP*, **1504** (2015) 117. http://arxiv.org/abs/1501.04943

[154] ATLAS Collaboration. *JHEP*, **1404** (2014) 031. http://arxiv.org/abs/1401.7610

[155] S. Chekanov et al. *Eur. Phys. J.*, **C42** (2005) 1–16. http://arxiv.org/abs/hep-ph/0503274

[156] P. M. Nadolsky, H-L. Lai, Q-H. Cao, J. Huston, J. Pumplin et al. *Phys. Rev.*, **D78** (2008) 013004. http://arxiv.org/abs/0802.0007

[157] M. Botje. *Comput. Phys. Commun.*, **182** (2011) 490–532. http://arxiv.org/abs/1005.1481

[158] F. James & M. Roos. *Comput. Phys. Comm.*, **10**, 6 (1975) 343–367. http://www.sciencedirect.com/science/article/pii/0010465575900399

[159] J. Pumplin, D.R. Stump & W.K. Tung. *Phys. Rev.*, **D65** (2001) 014011. http://arxiv.org/abs/hep-ph/0008191

[160] T. Carli, D. Clements, A. Cooper-Sarkar, C. Gwenlan, G. P. Salam et al. *Eur. Phys. J.*, **C66** (2010) 503–524. http://arxiv.org/abs/0911.2985

[161] J. M. Campbell & R. K. Ellis. *Phys. Rev.*, **D65** (2002) 113007. http://arxiv.org/abs/hep-ph/0202176

[162] J. Campbell, R. Ellis & D. Rainwater. *Phys. Rev.*, **D68** (2003) 094021. http://arxiv.org/abs/hep-ph/0308195

[163] ATLAS Collaboration. Submitted to JHEP. http://arxiv.org/abs/1410.8857

[164] ATLAS Collaboration. *JHEP*, **1405** (2014) 059. http://arxiv.org/abs/1312.3524

[165] ATLAS Collaboration. Submitted to EPJC. http://arxiv.org/abs/1411.1855

[166] S. A. Malik & G. Watt. *JHEP*, **1402** (2014) 025. http://arxiv.org/abs/1304.2424

[167] A. Denner, S. Dittmaier, T. Kasprzik & A. Muck. *JHEP*, **1106** (2011) 069. http://arxiv.org/abs/1103.0914

[168] A. Denner, S. Dittmaier, T. Kasprzik & A. Muck. *JHEP*, **0908** (2009) 075. http://arxiv.org/abs/0906.1656

Appendix A

Further W+jets Results Plots

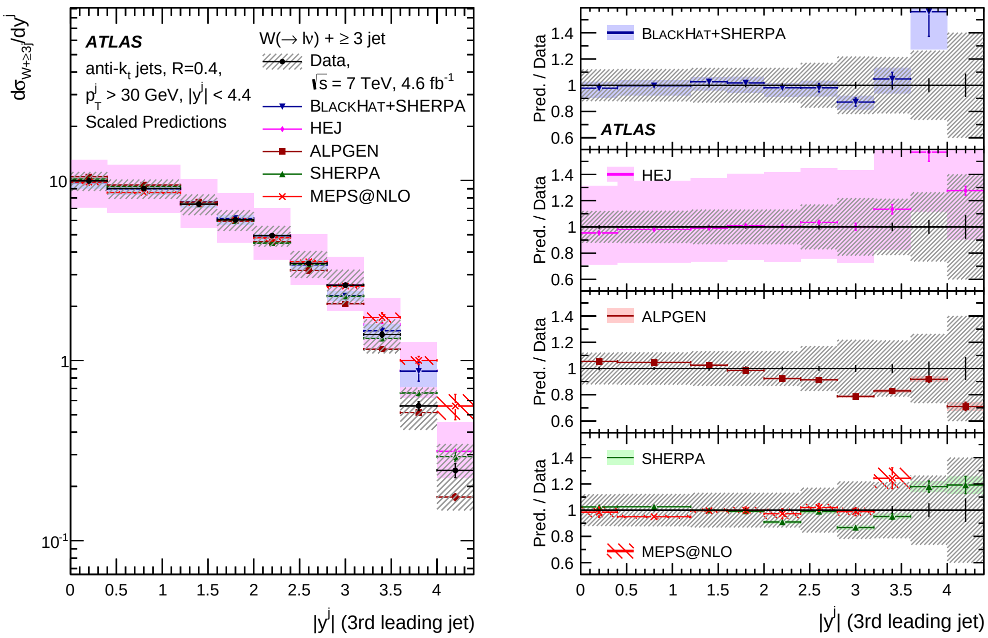


Figure A.1: Cross section for the production of $W \to l\nu$ as a function of third leading jet rapidity in events with $N_{\rm jet} \geq 3$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

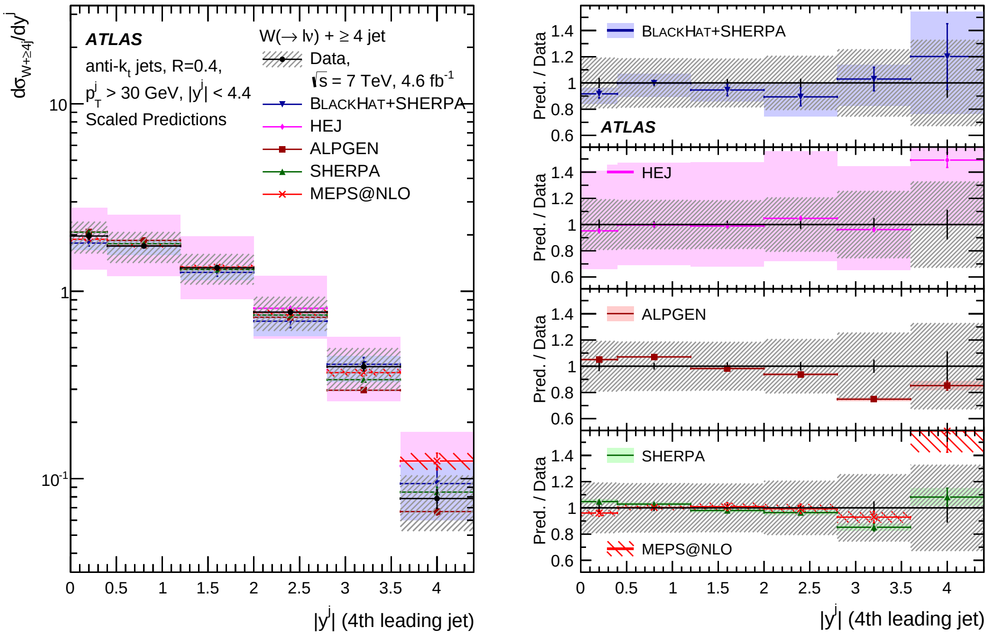


Figure A.2: Cross section for the production of $W \to l\nu$ as a function of fourth leading jet rapidity in events with $N_{\text{jet}} \geq 4$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

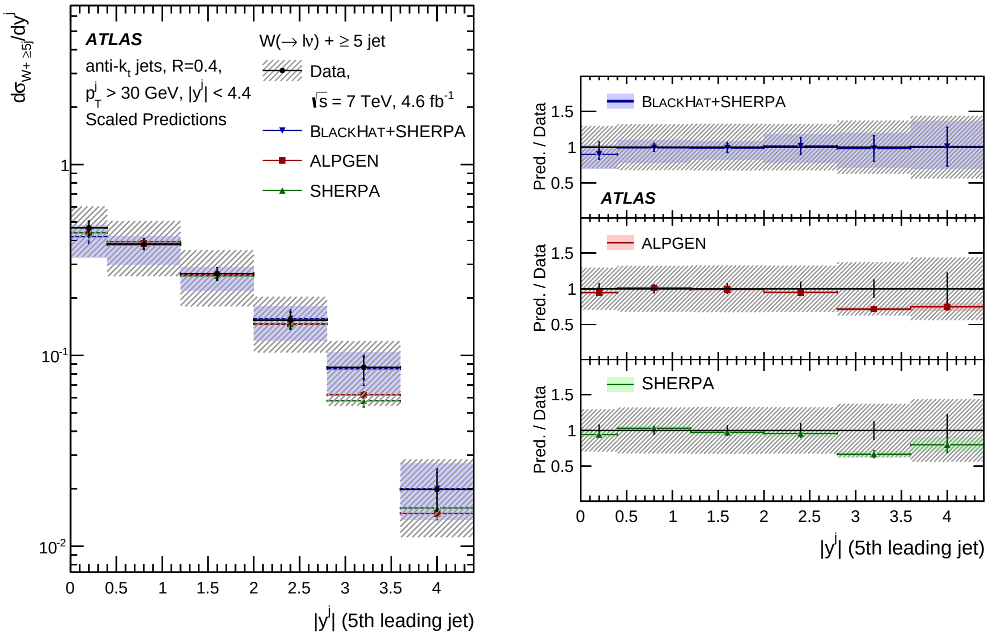


Figure A.3: Cross section for the production of $W \to l\nu$ as a function of fifth leading jet rapidity in events with $N_{\text{jet}} \geq 5$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

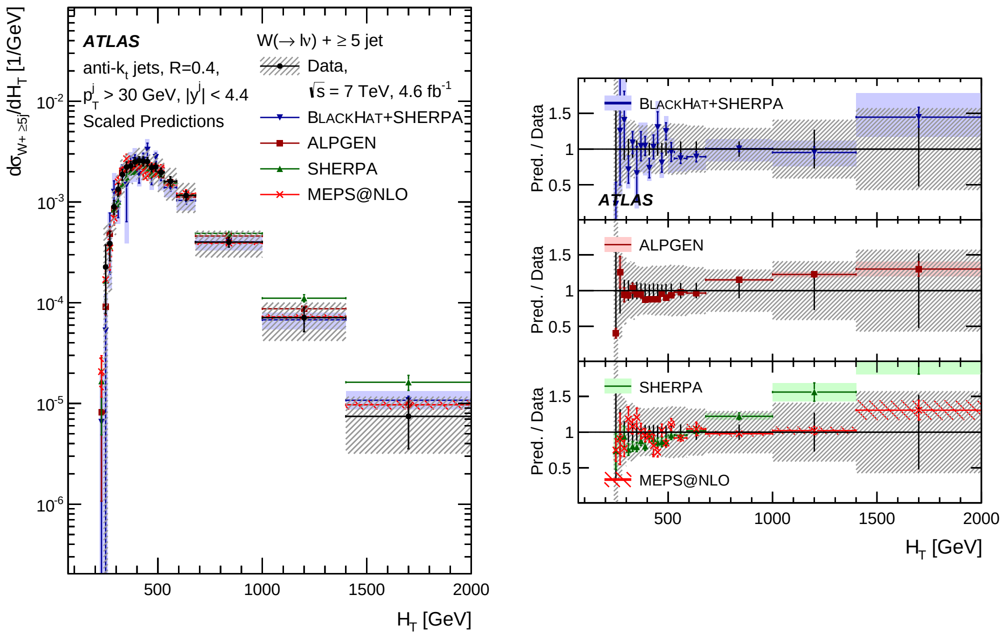


Figure A.4: Cross section for the production of $W \to l\nu$ as a function of H_T in events with $N_\mathrm{jet} \geq 5$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

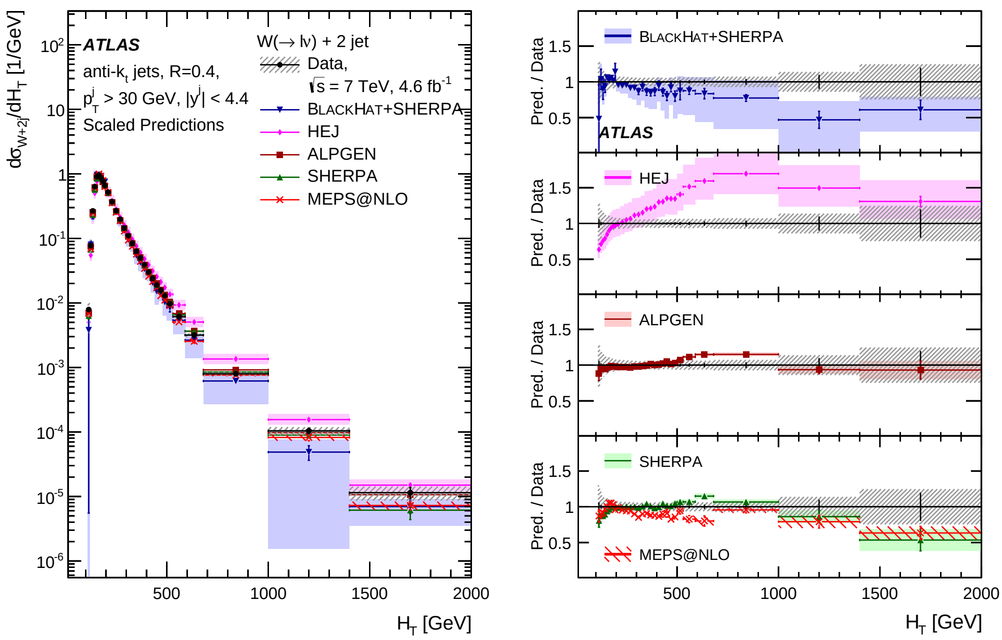


Figure A.5: Cross section for the production of $W \to l\nu$ as a function of H_T in events with $N_\mathrm{jet} = 2$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

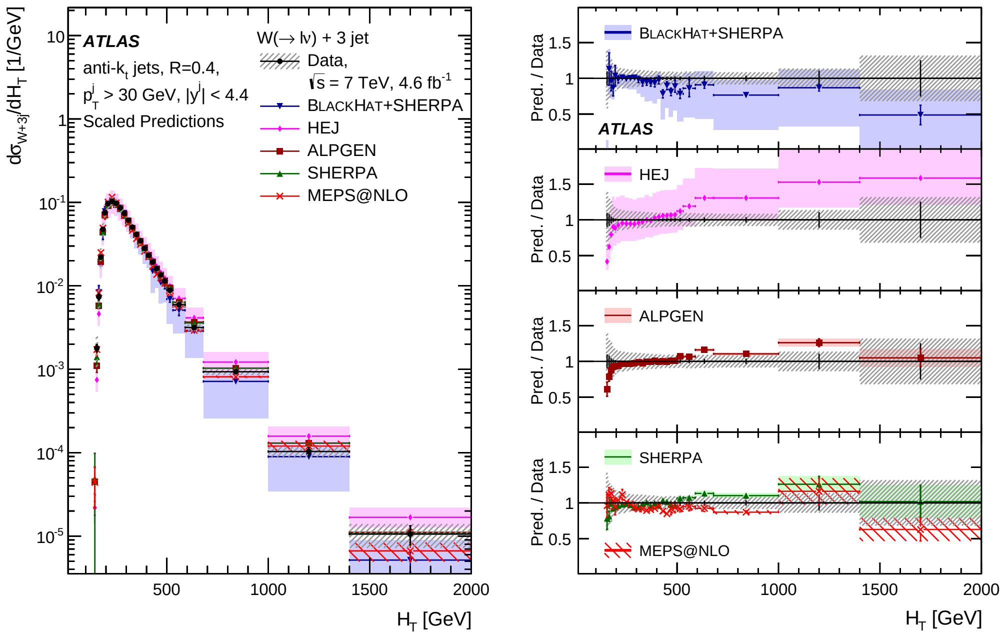


Figure A.6: Cross section for the production of $W \rightarrow l\nu$ as a function of H_{T} in events with $N_{\mathrm{jet}} = 3$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

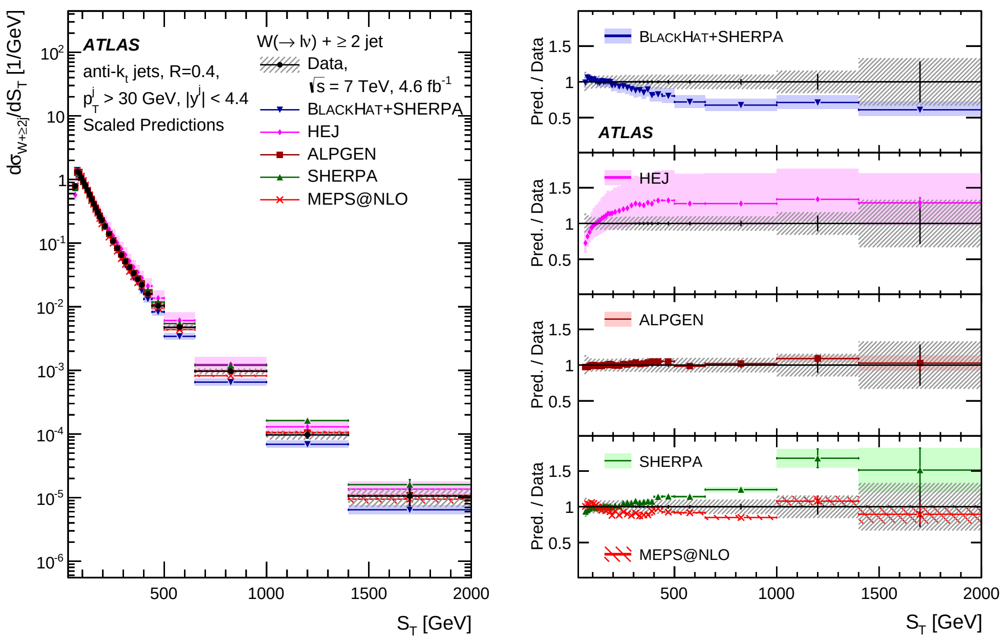


Figure A.7: Cross section for the production of $W \rightarrow l\nu$ as a function of S_{T} in events with $N_{\mathrm{jet}} \geq 2$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

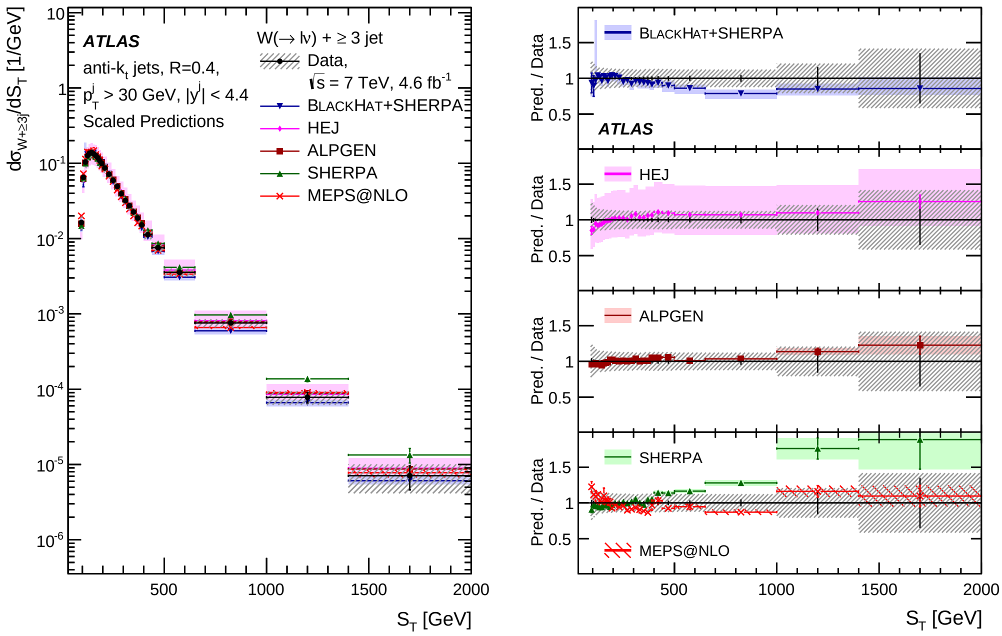


Figure A.8: Cross section for the production of $W \rightarrow l\nu$ as a function of S_{T} in events with $N_{\mathrm{jet}} \geq 3$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

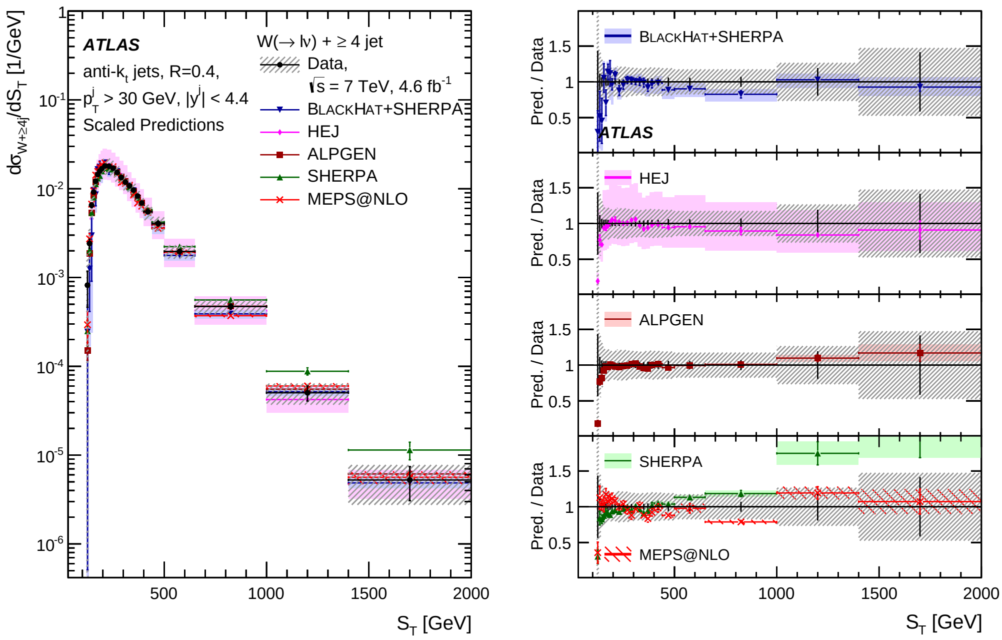


Figure A.9: Cross section for the production of $W \rightarrow l\nu$ as a function of S_{T} in events with $N_{\mathrm{jet}} \geq 4$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

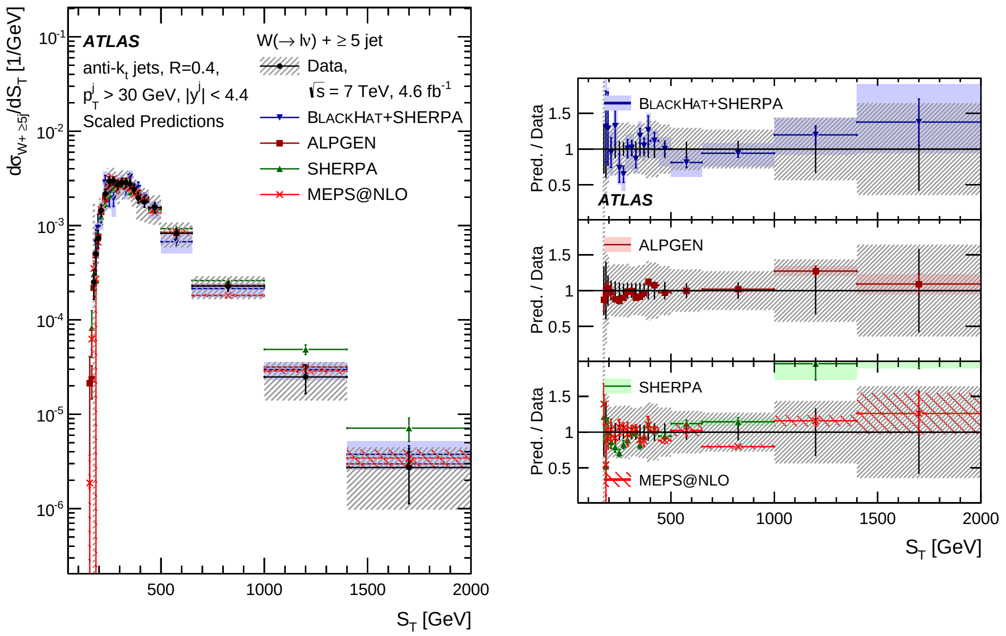


Figure A.10: Cross section for the production of $W \rightarrow l\nu$ as a function of S_{T} in events with $N_{\mathrm{jet}} \geq 5$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

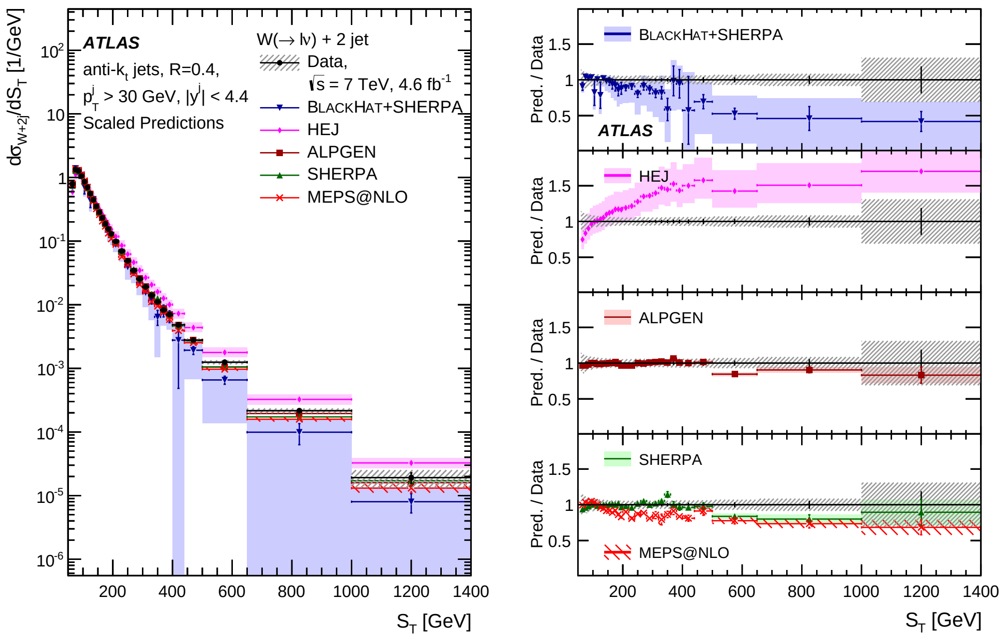


Figure A.11: Cross section for the production of $W \rightarrow l\nu$ as a function of S_{T} in events with $N_{\mathrm{jet}} = 2$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

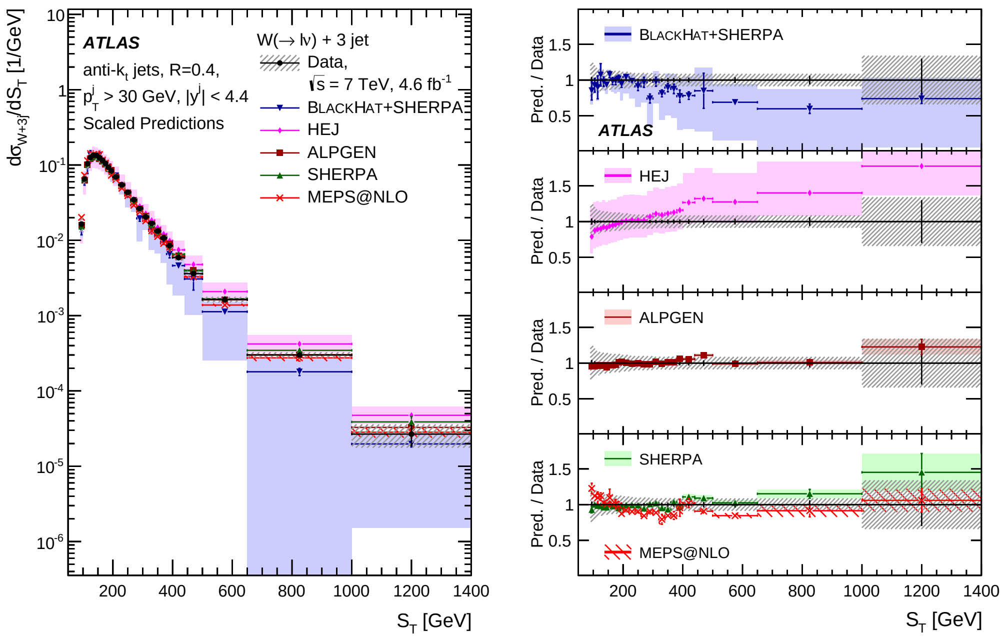


Figure A.12: Cross section for the production of $W \rightarrow l\nu$ as a function of S_{T} in events with $N_{\mathrm{jet}} = 3$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty. The left-hand plot shows the differential cross sections whilst the right-hand plot shows the ratio of the predictions to the measurement.

Appendix B

Further R_{jets} Results Plots

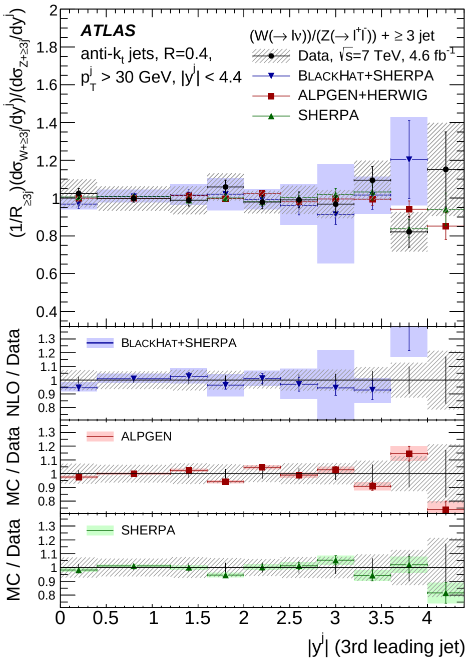
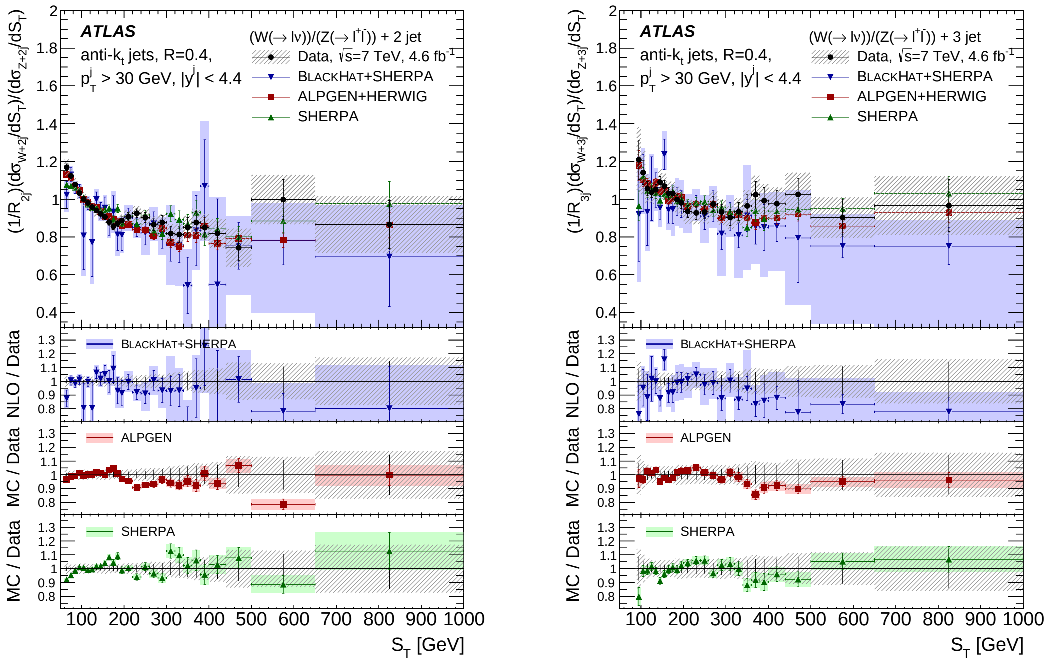


Figure B.1: R_{jets} as a function of third leading jet rapidity in events with $N_{\text{jet}} \geq 3$. Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty.



(a) (b)

Figure B.2: R_{jets} as a function of S_{T} in events with $N_{\text{jet}} = 2$ (a) and S_{T} in events with $N_{\text{jet}} = 3$ (b). Statistical uncertainties are shown by the vertical bars and the black hashed band shows the combined statistical and systematic uncertainty.