Abstract

The production of a W or Z boson in conjunction with jets is an interesting process in its own right as well as a signal channel (and background) for many interesting standard model and beyond standard model physics signals. Final states with 2,3,4 or more jets accompanying a W/Z boson will be observable at the LHC and will serve as a crucial part of the ATLAS physics program. The variety of possible jet multiplicities allows for precision tests of jet reconstruction algorithms and techniques. In addition, the reconstruction of leptons and of missing transverse energy becomes more complex in the presence of a multi-jet final state. In this note, we will quantify the differences of lepton, missing transverse energy and jet reconstruction with respect to that observed in inclusive W and Z production.

The wide kinematic range for production of W/Z + jets allows serves as a testing ground for perturbative QCD predictions, both fixed order alone and in conjunction with parton shower Monte Carlos. As an example, the possible large rapidity separations between pairs of jets allows for tests for the applicability and importance of BFKL-type logarithms.
Contents

1 Introduction 3

2 Reference Cross Sections 3
   2.1 Corrections from Parton to Hadron Level 3

3 Monte Carlo Datasets 4

4 Definitions and Cuts 5
   4.1 Electrons 5
   4.2 Muons 5
   4.3 Missing Transverse Energy 5
   4.4 Jets 6

5 Trigger Paths 6

6 Tag and Probe Studies 6
   6.1 Global Results in Comparison with the Inclusive Pythia Generated Sample 7
   6.2 Differential Efficiencies 7
   6.3 Background Treatment 7

7 Measurement of Z + Jet Cross-Sections 7
   7.1 Introduction 7
   7.2 Z → e⁺e⁻ + Jets 8
      7.2.1 Signal and Background Distributions 8
      7.2.2 Background Estimation 9
      7.2.3 Unfolding of Detector Effects 9
      7.2.4 Background Subtraction 11
      7.2.5 Comparison to Theory 12
   7.3 Z → µ⁺µ⁻ + Jets 12
      7.3.1 Background Estimation 12
      7.3.2 Signal and Background Distributions 13
      7.3.3 Correction from Parton to Particle Level 14
      7.3.4 Comparison of Event Generators and MCFM at the Hadron Level 14
      7.3.5 Unfolding Detector Effects 16
      7.3.6 Systematic Errors 16
   7.4 Z → µ⁺µ⁻ + b Jets 17
      7.4.1 Event Selection 18

8 Measurement of W + Jet Cross-Sections 20
   8.1 W → eν + Jets 20
      8.1.1 Signal-Background Analysis 20
      8.1.2 Data-Driven Methods for the Background Extraction 22
      8.1.3 Jet Energy Scale Uncertainties 23
      8.1.4 PDF Reweighting Technique 24
      8.1.5 PDF Uncertainties 25
   8.2 W → µν+ Jets 25
      8.2.1 Signal-Background Analysis 25
1 Introduction

In this note, we will examine the triggering, reconstruction and analysis of events containing a W or Z boson plus jets in ATLAS. We will concentrate on channels with decays into electrons and muons, ignoring for the moment taus (except to account for backgrounds to the other decay channels). Much of the effort on the triggering and reconstruction of leptons and of missing transverse energy is in common with the inclusive $W/Z$ note [1]; thus, we will not reproduce all of the details from that note, but rather will comment on the impact of a multi-jet environment on these issues. Fully simulated signal and background event samples will be treated as pseudo-data. Comparisons of the reconstructed/corrected quantities, when possible, will be to truth-level hadron information (and in some cases to parton-level information with parton-to-hadron corrections applied). Our primary end-result will be hadron-level cross sections, similar to what we will have with the real ATLAS data. We will present expectations/yields/systematics scaled to 1 fb$^{-1}$, an integrated luminosity that should be accumulated within the first two years of running, but we will also comment on difficulties expected and analysis strategies to be adopted during the early running.

An important source of theoretical systematic uncertainties at the LHC is represented by the Parton Distribution Functions (PDFs). Thus, assessing the level of PDF uncertainties is crucial for studying both processes within and beyond the SM. Experimental systematic uncertainties are the main limitations to the possibility of improving our knowledge on PDFs. Among these, errors due to a miscalibration of the jet energy scale are expected to be the dominant systematics. A comparison of the impact of theoretical and experimental systematic uncertainties is therefore crucial. The impact of JES uncertainties of 1, 3 and 5% (10%) will be examined and compared to the level of PDF uncertainties.

2 Reference Cross Sections

Reference cross sections for the ATLAS CSC notes are collected in ATLAS note [2], but we briefly discuss here the cross sections relevant for this CSC note. NLO is the first order at which the $W/Z$ + jets cross sections have a realistic normalization (and realistic shapes for some kinematic distributions) [3]. The current state of the art for NLO calculations is for $W/Z$ plus 2 jets, although there is ongoing work for the calculation of 3 jet final states. Cross sections for $W/Z$ + 0, 1 and 2 (3) final states can be conveniently calculated at NLO (LO) using the MCFM [4] program, and it is from this program that we determine our reference cross sections. The MCFM cross sections were generated using the CTEQ6.1 PDFs and a renormalization/factorization scale of $M^2_Z + p_T^2$, with similar kinematic cuts on the leptons and jets as will be described in Section 4.

2.1 Corrections from Parton to Hadron Level

Cross section measurements in data and LO/NLO predictions are to be compared at the hadron level (particle level). Hence, the data have to be unfolded with respect to the detector response. For comparisons of data to NLO parton level predictions from MCFM, either the data needs to be corrected to the parton level or the theory corrected to the hadron level. We discuss the latter correction below for the specific case of $Z$ + jets (but which can also be applied without great error to the case of $W$ + jets). The MCFM predictions have to be corrected with respect to the non-perturbative effects of fragmentation and underlying event (UE). The fragmentation and underlying event corrections are extracted using Pythia [5].

1) For this note, we compare the MCFM predictions only for $Z$ + jets.
Monte Carlo samples by comparing the hadron level results, with the current ATLAS underlying event tune, to corresponding results in samples in which fragmentation and multiple-parton-interactions have been switched off. The corrections are determined by dividing the hadron-level distribution of the observables from standard Pythia by the respective distributions from Pythia with the non-perturbative effects switched off. To the extent to which the 2 partons that can comprise a jet in MCFM mimic the effects of the parton shower in Pythia, the corrections derived from the procedure above can be applied to the MCFM output [3]. We study the hadronization corrections for both cone (R=0.4 and 0.7; termed Cone04 and Cone07 hereafter) and $k_T$ jet algorithms (D=0.4 and 0.6; termed Kt04 and Kt06).

Figure 1 shows the correction from parton to hadron level from fragmentation alone (a) and from both fragmentation and underlying event (b) for Cone04 jets. The impact of fragmentation is to reduce the amount of energy in the jet cone. Thus, from fragmentation effects alone, jets at the hadron level tend to have lower $p_T$ than jets at the parton level. The impact of the underlying event is to add energy to the hadron level jet. In general, the underlying event tends to add more energy to the jet than lost by fragmentation, but the exact ratio depends on the radius of the jet. Whereas the fragmentation corrections for Cone07 jets are smaller than for Cone04 jets, the underlying event corrections are larger due to the larger cone size. Kt04 shows the lowest combined corrections since fragmentation and underlying event effects basically cancel out. The performance of Kt06 jets is comparable to the one of Cone04 jets. Except for Cone07 jets, the non-perturbative effects are negligible for jets with $p_T > 40$ GeV in the current Pythia tune.

![Figure 1: Ratio of Cone04 Jet $p_T$ distributions (a) between standard Pythia and Pythia without fragmentation and (b) between standard Pythia and Pythia without non-perturbative corrections.](image)

### 3 Monte Carlo Datasets

Most of the Monte Carlo data samples used in these studies were generated with Alpgen [6] interfaced with Herwig [7] using the leading order PDF set CTEQ6LL [8]. The cross sections were calculated using a renormalization/factorization scale of $M_Z^2 + p_T^2$, similar to that used for the MCFM predictions. The full datasets were obtained by merging samples of $W(Z) + 0$ up to 5 partons, weighted according to the expected cross-sections, with an MLM [6] matching cut at $p_T > 20$ GeV, and normalized to an integrated luminosity of 1 fb$^{-1}$. All datasets correspond to exclusive samples, with the exception of the highest multiplicity sample ($W(Z) + 5$ partons) which is inclusive. The samples correspond to the
decays, $W \rightarrow e\nu$, $W \rightarrow \mu\nu$, $Z \rightarrow ee$ and $Z \rightarrow \mu^+\mu^-$. A generator-level filter (VBF-loose) requires one truth Cone04 jet of $p_T > 20$ GeV and $|\eta| < 5.0$ and one (two for $Z$) electrons/muons with $p_T > 10$ GeV and $|\eta| < 2.7$ in the event. Pythia signal and background samples were also produced, using the standard ATLAS underlying event tune. Pythia $Z \rightarrow ee$ and $W \rightarrow e\nu$ events were generated with a generator-level filter requiring one truth electron with $p_T > 10$ GeV and $|\eta| < 2.7$. Pythia $Z \rightarrow \tau\tau$ events were generated with a filter requiring two electrons/muons with $p_T > 5$ GeV and $|\eta| < 2.8$; Pythia $W \rightarrow \tau\nu$ events were generated with a filter requiring electron/muon decays of the $\tau$, with the decay leptons having $p_T > 5$ GeV and $|\eta| < 2.8$. For $Z \rightarrow \tau\tau$ and $Z \rightarrow ee$, the dilepton mass is required to be larger than 60 GeV. A large QCD di-jet sample (3.3M events) was also produced, with a minimum hard-scattering transverse momentum of 17 GeV/c.

4 Definitions and Cuts

We adopt as much as possible definitions and cuts in common with the other CSC notes, and in particular with the inclusive $W/Z$ note [1], with comparisons to alternate definitions/cuts where relevant. The reconstruction algorithms in this note are not necessarily optimal for the running conditions in the first two years, but provide a reasonable baseline. We expect further optimization as data becomes available.

4.1 Electrons

Both the $W$ and $Z$ analyses impose common electron selection criteria. It is required that the transverse momentum of the electron candidate, $p_{Te} > 25$ GeV, and that it lies in the range $|\eta| < 2.4$, excluding the barrel-to-endcap calorimeter crack region ($1.37 < |\eta| < 1.52$). The electrons are required to fulfill the medium electron-ID [9] (hereafter, referred to as medium IsEM), which consists of requirements on the calorimeter shower-shape and the matched track. In order to further suppress the contamination from fake electrons in the $W$ selection an isolation requirement within a cone $\Delta R = 0.2$ of the candidate is imposed, requiring the number of tracks with $p_T > 1$ GeV to be less than 4 and the scalar sum of the track $p_T$ to be less than 4 GeV. The $Z$ selection requires two electron candidates with an invariant mass of $81 < M_{ee} < 101$ GeV and $\Delta R > 0.2$. No calorimeter isolation cuts are applied for these analyses, due to simulation problems, but will be in the actual data analyses. There is an implicit isolation cut, however, present in the trigger [10].

4.2 Muons

As in the case of electrons, both the $W$ and $Z$ analyses impose common requirements on muon candidates. A muon candidate requires the combined reconstruction of an inner detector track and a track in the muon spectrometer [11]. The muon is required to have a transverse momentum of $p_{T\mu} > 20$ GeV. In the $Z$ analysis, the muon is required to have $p_{T\mu} > 15$ GeV and $0.1 < |\eta| < 1.2$ or $1.3 < |\eta| < 2.4$. Isolation is applied by requiring the energy deposition in the calorimeter to be the less than 15 GeV in a cone of $\Delta R = 0.2$ around the extrapolation of the muon track (hereafter referred to as cone0.2 isolation). The $Z$ analysis selects two muon candidates with an invariant mass of $81 < M_{\mu\mu} < 101$ GeV.

4.3 Missing Transverse Energy

In the $W$ analysis, a cut on the missing transverse transverse energy ($E_T^\ast$) is applied to suppress QCD backgrounds. The raw calorimeter-based $E_T^\ast$ is corrected for the presence of a muon or electron in the

2) General criteria corresponding to tight and loose electron ID are also available, with tight(loose) resulting in a lower (higher) efficiency, and a smaller (larger) background.
final state [12]. Studies were performed to establish the optimal value of the cut on $E_T$ that reduces QCD backgrounds to acceptable levels, while not introducing significant biases to the cross-section measurement. An optimal cut of $E_T > 25$ GeV was chosen.

4.4 Jets

Jet clustering algorithms can be divided into two main classes: cones and iterative recombination (e.g., the $k_T$ algorithm) [13]. Historically, cone algorithms have been used in hadron colliders, due to concerns about speed, especially at the trigger level, and of large systematic effects in busy multi-jet environments. Fast implementations of the $k_T$ algorithm [14], as well as detailed studies at the Tevatron performing precision measurements with the algorithm [15] argue for the use of the $k_T$ algorithm along with cone-based ones at the LHC.

For the analyses in this note, we use jets clustered with the standard ATLAS Seeded Cone algorithm with a radius of $R = 0.4$ (appropriate for the complex final states expected for the multi-jet environments explored in this note), built from calorimeter towers, and calibrated to the hadron level. In the $Z \rightarrow \mu^+ \mu^-$ channel we also examine the impact of using clusters as a jet constituent for the same algorithm. A detailed comparison with other jet algorithms, including the $k_T$ algorithm, is beyond the scope of this note, but can be found in Ref. [17]. The jet kinematic selection is common for the $Z$ and $W$ analyses. A cone radius of 0.4 is it is required that the jet transverse momentum be $p_{Tj} > 20$ GeV in the range $|\eta| < 3.0$ excluding the crack region ($1.37 < |\eta| < 1.52$). The lepton and jet candidates must be separated by $\Delta R_{lj} > 0.4$. The cross section measurement itself is prepared for jets with $p_{Tj} > 40$ GeV.

5 Trigger Paths

The trigger selection used here is the same as that used in the inclusive analyses [1]. In the electron channel, $W \rightarrow e\nu + \text{jet(s)}$ events are required to pass the isolated single-electron trigger (e25i); $Z \rightarrow ee + \text{jet(s)}$ events are required to pass the isolated di-electron trigger (2e15i) or the isolated single-electron trigger. In the muon channel, $W \rightarrow \mu\nu + \text{jet(s)}$ events are required to pass the isolated single-muon trigger; $Z \rightarrow \mu^+ \mu^- + \text{jet(s)}$ events are required to pass the isolated dimuon trigger.

The trigger efficiencies at the first, second and event filter levels were evaluated as a function of the jet multiplicity. The trigger efficiency was also studied as a function of the overall hadronic activity, the $P_T$ of the leading jet and the $Z,W$ transverse momentum. For this purpose, a MC truth and data-driven tag-and-probe method were implemented. Good agreement between the two methods was found.

The overall trigger efficiency, with respect to that for the off-line cuts, for the inclusive analysis has been compared to that obtained here. It was found that the trigger efficiency for the $Z,W+$jets analysis is, in general, $1.5 - 2\%$ lower than that of the inclusive sample.

6 Tag and Probe Studies

The ‘tag and probe’ method is used in $Z \rightarrow e^+e^-+$jet events to determine the single object e25i efficiency as well as the offline reconstruction and identification efficiencies from data. As described in detail in the tag and probe section of [18], a diagnostic sample of size $N_1$ is defined by requiring at least one electron candidate to satisfy tag requirements. A control sample of size $N_2$ is defined by additionally requiring that the second electron satisfies the probe requirements. $N_1$ and $N_2$ may be used to extract the single electron efficiency. Absolute reconstruction, identification and trigger efficiencies are calculated independently and may be multiplied together to give the overall efficiency for an electron. Results in this note are quoted at medium IsEM level and the cuts used are those taken as standard in this note. Results are quoted at $1 \text{fb}^{-1}$ of integrated luminosity.
6.1 Global Results in Comparison with the Inclusive Pythia Generated Sample

Comparisons have been made with an inclusive Pythia generated sample used in [1], which doesn’t have the generator level truth jet cuts that are present in this sample. The efficiencies measured from these events may be found in CSC notes [9] and [18]. The difference in global efficiencies between the two samples are of the order of a few percent or less.

6.2 Differential Efficiencies

The dependence of the efficiency on the electron parameters ($\eta$, $\phi$, $p_T$) in the Pythia sample is well documented in [9] and [18] and the dependence observed is similar in the Alpgen sample. The effects of jet parameters on efficiencies are shown in Figure 2. Some points noted:

- A lower efficiency is observed in regions of higher jet activity (higher jet multiplicity, hadronic activity and $p_T$ of the most energetic jet. These events are more likely to fail isolation cuts.

- The drop in efficiency at low hadronic $p_T$ is due to the lower electron $p_T$ in this region.

- There is a drop at higher $\Delta R$ (between the electron and jet) due to the space-point bug in the L2 tracking algorithm [18].

6.3 Background Treatment

The method of fake rates as explained in [18] was used to estimate the background magnitude. It has been observed that the signal-background ratio is lower in the Alpgen sample. The measured efficiencies after background subtraction (as opposed to earlier results which are from signal analysis only) are tabulated in Table 1. We see that the main discrepancy is seen at the reconstruction level, which is to be expected as this is where the background level is highest.

<table>
<thead>
<tr>
<th>Level</th>
<th>$\Delta$Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconstruction</td>
<td>1.00</td>
</tr>
<tr>
<td>LVL1</td>
<td>-0.74</td>
</tr>
<tr>
<td>LVL2</td>
<td>0.04</td>
</tr>
<tr>
<td>EF</td>
<td>0.02</td>
</tr>
<tr>
<td>Whole trigger</td>
<td>-0.63</td>
</tr>
</tbody>
</table>

7 Measurement of Z + Jet Cross-Sections

7.1 Introduction

This section discusses the measurement of the inclusive $Z(\rightarrow e^+e^-,\mu^+\mu^-) + \text{jets}$ cross sections as a function of the jet transverse momentum and jet multiplicity. The measurement is compared with LO and NLO perturbative QCD predictions. It will be used, in addition, to validate the event generators which are used to predict the W/Z+jets backgrounds for searches in ATLAS.

In order to prepare the measurement, we use fully-simulated signal and background Monte Carlo events as pseudo-data, with which we perform all steps of the analysis. We also compare the predictions
of fully simulated signal Monte Carlo sets from different event generators. The goal of the analysis simulation is to validate the lepton and jet reconstruction in high-multiplicity events, develop the necessary analysis techniques (unfolding, background extraction) and evaluate the statistic and systematic limitations in terms of probing the QCD predictions and discriminating between predictions of different event generators. Two separate studies were performed on the feasibility of the $Z \rightarrow e^+e^-$ and the $Z \rightarrow \mu^+\mu^-$ channels (Sections 7.2 and 7.3).

### 7.2 $Z \rightarrow e^+e^- + \text{Jets}$

#### 7.2.1 Signal and Background Distributions

The presence of additional jets in the event has an impact on the kinematics of both the leptons and jets. The distribution of electron $p_T$, jet $p_T$, $\Delta R$ between electrons and the minimum $\Delta R$ between each electron and the jets (for different jet multiplicities) were studied using fully-simulated $Z \rightarrow ee$ samples. As expected, the electrons are more boosted (larger $p_T$ and lower $\Delta R$ between electrons) in events with jets, and the distance between electrons and jets becomes smaller in high-multiplicity events. The average $p_T^{\text{jet}}$ increases with the number of jets.
7.2.2 Background Estimation

The most important backgrounds to the $Z \rightarrow ee + \text{jets}$ signal are processes with real electrons ($t\bar{t}$, $W \rightarrow ee$, $Z \rightarrow \tau\tau$) and QCD jet production. All backgrounds are estimated from Monte Carlo. $Z \rightarrow \tau\tau$ and $W \rightarrow ee$ are simulated with Pythia and $t\bar{t}$ is simulated with MCatNLO [19]. The QCD background is derived from a filtered Pythia dijet sample.

The combined distribution of the invariant mass for signal and background events is shown in Figure 3a-c for various jet multiplicities. Table 2 gives an overview of the accepted cross section expected from Monte Carlo from signal and backgrounds. With increasing jet multiplicity, $t\bar{t}$ replaces QCD as the dominant background source. Eventually the QCD background will be determined with data-driven methods. The simulation of the $t\bar{t}$ background also has to be validated separately with data. Figures 3d-f show the distribution of signal and backgrounds for the observables (jet multiplicity and $p_T$).

7.2.3 Unfolding of Detector Effects

The reconstructed data has to be unfolded from the detector level to the hadron level, correcting for efficiency, resolution and non-linearities in electron and jet reconstruction. In this note, the individual unfolding corrections are assumed to factorize in leading approximation, and the individual contributions are investigated separately. For the data analysis, an approach will be used consisting of unfolding the combined impact of jet reconstruction efficiency and jet resolution in a bin-by-bin unfolding procedure. The reconstructed data has to be unfolded from the detector level to the hadron level, correcting for shifts in the jet reconstruction efficiency. All corrections are derived for ten bins in $p_T$ due to the $p_T$ resolution, extracted from Alpgen $Z \rightarrow ee + \text{jets}$ Monte Carlo. The jet $p_T$ scale and resolution are determined using a matching window of $\Delta R(\text{truth} - \text{reco jet}) < 0.2$ and $0.5 < p_T(\text{reco})/p_T(\text{truth}) < 1.5$. The jet reconstruction efficiency is determined as the fraction of truth jets which are matched to reconstructed jets applying the same requirements. As expected, the largest bias is observed for low values of $p_T$. The impact of the resolution, shown in Figure 4b, is derived by comparing the $p_T$ distribution of truth jets before and after a gaussian smearing with the resolution as determined above. Due to the

<table>
<thead>
<tr>
<th>Process</th>
<th>$Z \rightarrow ee + \geq 1\text{jet}$</th>
<th>$Z \rightarrow ee + \geq 2\text{jets}$</th>
<th>$Z \rightarrow ee + \geq 3\text{jets}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$xsec$ [fb]</td>
<td>fraction [%]</td>
<td>$xsec$ [fb]</td>
</tr>
<tr>
<td>$Z \rightarrow ee$</td>
<td>23520±145</td>
<td>91.7±0.9</td>
<td>4894±45</td>
</tr>
<tr>
<td>QCD jets</td>
<td>1545±89</td>
<td>6.0±0.3</td>
<td>336±42</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>552±74</td>
<td>2.2±0.3</td>
<td>424±65</td>
</tr>
<tr>
<td>$Z \rightarrow \tau\tau$</td>
<td>3.2±1.2</td>
<td>0.01±0.005</td>
<td>(0.66±0.45)</td>
</tr>
<tr>
<td>$W \rightarrow ev$</td>
<td>(28±13)</td>
<td>(0.1±0.05)</td>
<td>(5.9±2.6)</td>
</tr>
</tbody>
</table>

Table 2: The accepted cross section [in fb] from signal and background in the $Z \rightarrow ee + \text{jets}$ analysis. The numbers in brackets are extrapolated from results obtained for a lower jet multiplicity.

The event weight is corrected for the electron reconstruction efficiency, derived as a function of generated $|\eta|$ in four bins of generated $p_T$. It is also corrected for the global efficiency for the electron trigger with respect to the offline selection, determined as: $\text{eff}_{\text{trig}} = (99.63 + -0.11)$%. The errors on deriving this efficiency, stemming from the limited Monte Carlo statistics, are taken into account as systematic errors for the unfolding procedure. The jet observables have to be corrected for shifts in the measured jet energy scale (mainly non-linearities at low $p_T$), the jet energy scale resolution and the jet reconstruction efficiency. All corrections are derived for ten bins in $p_T$ with a comparable number of events in each bin in order to avoid large statistical fluctuations.
Figure 3: The distribution of the di-electron mass for signal and background for (a) $Z \rightarrow ee+ \geq 1\text{jet}$, (b) $Z \rightarrow ee+ \geq 2\text{jets}$ and (c) $Z \rightarrow ee+ \geq 3\text{jets}$; the number of events with $\geq N\text{ jets}$ ($p_T > 40\text{ GeV}$) (d) and of the $p_T$ of the leading (e) and the next-to-leading (f) jet for $\int Ldt = 1\text{fb}^{-1}$. 
large statistical uncertainties, a global average correction of $98.4 \pm 0.5\%$ has been determined for jets with $p_T > 40$ GeV.

The reconstructed jet $p_T$ is corrected with the jet energy scale corrections, and the event is weighted for each required jet with the correction for efficiency-loss in reconstruction, and the overpopulation due to the jet $p_T$ resolution. The unfolding corrections are validated by comparing the distributions of truth jet variables with the corresponding ones for corrected reconstructed jets. Figure 5 compares the distribution of the $p_T$ of the leading and the next-to-leading jet, in different unfolding stages, with the $p_T$ distribution of truth jets. Within the statistical and systematic errors, the $p_T$ distributions of truth jets and corrected reconstructed jets are in agreement.

![Figure 4: Jet reconstruction efficiency (a), and the impact of resolution on the jet $p_T$ spectrum (b).](image)

![Figure 5: Comparison of the distribution of the $p_T$ of the leading jet (a) and the next-to-leading jet (b) in truth and corrected reconstructed MC](image)

### 7.2.4 Background Subtraction

The $Z \rightarrow \tau\tau, t\bar{t}$ and $W \rightarrow e\nu$ backgrounds are subtracted using the Monte Carlo estimates. This will also be done for the collision data. The QCD background is subtracted by weighting all events with a global factor, calculated as $1 - \text{QCD-fraction}$ (as in Table 2). In collision data, the factor can be derived by a combined fit for signal and background of the invariant mass peak (using sidebands). This method assumes a comparable distribution of the jets in QCD events with two electron fakes and in $Z \rightarrow ee$ events. Alternatively, the background sample can be derived by inverting selected electron-ID cuts. The
latter method is preferable but cannot be tested here for the lack of statistics of full simulated QCD MC events. For the final simulation of the measurement, signal and background samples are combined and unfolded to the hadron level as described above.

7.2.5 Comparison to Theory

We consider the comparison of theory and measurement for theoretically well-defined quantities: the inclusive cross section for $Z \rightarrow \ell\ell+ \geq 1\text{jet}$, $Z \rightarrow \ell\ell+ \geq 2\text{jets}$ and $Z \rightarrow \ell\ell+ \geq 3\text{jets}$ and the differential cross sections with respect to the $p_T$ of the leading and the next-to-leading jets. The MCFM predictions are corrected for the residual energy loss due to non-perturbative effects for jets with $p_T > 40$ GeV, determined from the current Pythia tune of underlying event and fragmentation as $0.98 \pm 0.01$ for $Z \rightarrow ee+ \geq 1\text{jet}$, $0.98 \pm 0.02$ for $Z \rightarrow ee+ \geq 2\text{jets}$ and $0.95 \pm 0.06$ for $Z \rightarrow ee+ \geq 3\text{jets}$. PDF uncertainties on the MCFM predictions are calculated using the complete set of error PDFs. The inclusive Pythia and Alpgen $Z \rightarrow e^+e^-$ samples are normalized to the NLO inclusive (DrellYan) MCFM $Z \rightarrow e^+e^-$ cross section at the generator level.

Figure 6 shows the comparison of the distribution of the observables (jet multiplicities and the $p_T$ of the leading and next-to leading jets) at the hadron level for Alpgen and Pythia Monte Carlo with LO and NLO MCFM calculations. The error bars are calculated only from intrinsic MC quantities as the quadratic sum of statistical errors from the MC sample size and the systematic errors from the unfolding corrections derived for MC. The shape of the jet-$p_T$ distribution predicted by Alpgen agrees well with the shape predicted by MCFM. Due to the tuning of the leading soft radiation in the parton shower, Pythia predicts a larger inclusive cross section for $Z \rightarrow ee+ \geq 1\text{jet}$ but a clearly softer $p_T$ spectrum.

As a further step towards real Atlas data, the systematic errors are adjusted to the values expected from the collision data and propagated to the measured cross section. Figure 6 shows the relative systematic uncertainty on the cross section (normalized to 1) expected for different uncertainties on the jet energy scale for the production of a $Z$ with 1-4 jets ($p_T > 40$ GeV). Since the difference between the LO and NLO cross section predictions are on the order of 30%, with a 3% uncertainty on the jet energy scale this is not possible. The uncertainty on the jet energy resolution and its impact on the unfolding procedure is an additional source of systematic errors on the cross section measurement. They are investigated for jet resolutions up to two times the ones currently expected. An uncertainty of 50% on the jet resolution propagates to an error of 2-4% on the $Z \rightarrow ee+ \geq 1\text{jet}$ cross section measurement. A wrong assumption of the jet-$p_T$ distribution in calculating the unfolding corrections from the jet energy resolution can also lead to a systematic shift in the cross section measurement. A comparison between the unfolding corrections derived from Pythia and from Alpgen yields a systematic uncertainty of up to 1.5%.

7.3 $Z \rightarrow \mu^+\mu^- + \text{Jets}$

The impact of a high jet multiplicity environment on the lepton reconstruction efficiency has also investigated for the case of the muon channel. Muon reconstruction efficiencies for different isolation requirements have been studied.

7.3.1 Background Estimation

The important backgrounds for the $Z \rightarrow \mu^+\mu^- + \text{jets}$ analysis are processes with similar topologies ($t\bar{t}$, $W \rightarrow \mu\nu$, $Z \rightarrow \tau^+\tau^-$) with real muons, and QCD multi-jet production. For $t\bar{t}$, $W \rightarrow \mu\nu$, and $Z \rightarrow \tau^+\tau^-$, Pythia and Alpgen Monte Carlo samples are used. It can be assumed that the dominating QCD dijet contribution of highly energetic muons are from decays of $b\bar{b}$ mesons. Monte Carlo samples
Figure 6: The inclusive jet cross section as a function of jet multiplicity is shown in (a); the $p_T$ of the leading jet (b) and of the next-to-leading jet (c) are shown for the unfolded Pythia and Alpgen Monte Carlo predictions and for MCFM for an $\int L dt = 1$ fb$^{-1}$. In (d), the relative systematic uncertainties on the cross section (normalized to 1) are shown for different uncertainties on the jet energy scale.

generated with Pythia $b\bar{b}(\mu\mu)$ are thus used to increase the muon background statistics. Figure 7 shows the distribution of the invariant mass for signal and background, using Pythia (a) and Alpgen (b).

### 7.3.2 Signal and Background Distributions

Muons from $Z \rightarrow \mu\mu$ are highly energetic and isolated. Therefore, muons coming from background can be rejected by imposing isolation requirements. In this analysis, several different isolation cuts have been studied. Table 3 shows results for signal and background considering different isolation criteria, for Pythia and for Alpgen. As the size of the muon isolation cone is increased, the background is reduced, and in particular that from $b\bar{b}$. For our analysis, the cone0.20 isolation cut has been used (a muon isolation cone of R= 0.20 with $E_T < 15$ GeV inside the cone) is utilized; after the imposition of this isolation cut, the $b\bar{b}$ background is reduced, and no bias is observed for the muon reconstruction efficiency for high jet multiplicities. The invariant mass window considered is 81-101GeV in order to increase the signal/background ratio (in particular for the $t\bar{t}$ background). Table 4 shows the accepted cross section and the fraction of signal and background for different jet multiplicities using Pythia and Alpgen. As the jet multiplicity increases, the $t\bar{t}$ background increases. For $Z \rightarrow \mu\mu + \geq 3$ jets events, the $t\bar{t}$ background fraction increases to 12-20%.
Figure 7: Background estimation for Pythia (a) and Alpgen (b) requiring isolated muons (cone0.20, \(E_T\) isolation \(< 15\text{GeV}\)) and at least one jet \((E_T > 40\text{GeV}, \eta < 3.0)\).

Figure 8 shows jet multiplicities for the signal and backgrounds for Pythia (a) and Alpgen (b) Monte Carlo data samples. Figure 8 also shows jet \(p_T\) and \(\eta\) distributions for signal and backgrounds (Pythia (a),(b) and Alpgen (c),(d)).

**NOTE:** WE WILL REPLACE FIGURES 7 AND 8 BY SOMETHING SIMILAR TO FIGURE 3.

### Table 3: Accepted cross section from signal and background for \(Z \rightarrow \mu^+\mu^- + \text{jets}\).

<table>
<thead>
<tr>
<th>Process</th>
<th>no isolation xsec ([\text{fb}])</th>
<th>no isolation fraction [%]</th>
<th>cone0.20 isol. xsec ([\text{fb}])</th>
<th>cone0.20 isol. fraction [%]</th>
<th>cone0.40 isol. xsec ([\text{fb}])</th>
<th>cone0.40 isol. fraction [%]</th>
<th>cone0.60 isol. xsec ([\text{fb}])</th>
<th>cone0.60 isol. fraction [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>pythia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Z \rightarrow \mu^+\mu^-)</td>
<td>73663 \pm 788</td>
<td>90.6 \pm 0.9</td>
<td>72819 \pm 784</td>
<td>96.9 \pm 1.0</td>
<td>71182 \pm 775</td>
<td>98.7 \pm 1.1</td>
<td>67900 \pm 757</td>
<td>99.0 \pm 1.1</td>
</tr>
<tr>
<td>(W \rightarrow \mu\nu)</td>
<td>183 \pm 183</td>
<td>0.2 \pm 0.2</td>
<td>0 \pm 183</td>
<td>0.0 \pm 0.2</td>
<td>0.0 \pm 183</td>
<td>0.0 \pm 0.3</td>
<td>0.0 \pm 18.3</td>
<td>0.0 \pm 0.3</td>
</tr>
<tr>
<td>QCD((b\bar{b}))</td>
<td>5390 \pm 2201</td>
<td>6.6 \pm 2.7</td>
<td>1225 \pm 541</td>
<td>1.6 \pm 0.7</td>
<td>0.0 \pm 898</td>
<td>0.0 \pm 1.2</td>
<td>0.0 \pm 8980</td>
<td>0.0 \pm 1.3</td>
</tr>
<tr>
<td>(\tau^+\tau^-)</td>
<td>2091 \pm 149</td>
<td>2.6 \pm 0.2</td>
<td>1141 \pm 110</td>
<td>1.5 \pm 0.1</td>
<td>939 \pm 100</td>
<td>1.3 \pm 0.1</td>
<td>757 \pm 89</td>
<td>1.1 \pm 0.1</td>
</tr>
<tr>
<td>(Z \rightarrow \tau^+\tau^-)</td>
<td>0.0 \pm 0.0</td>
<td>0.0 \pm 0.0</td>
<td>0.0 \pm 0.0</td>
<td>0.0 \pm 0.0</td>
<td>0.0 \pm 0.0</td>
<td>0.0 \pm 0.0</td>
<td>0.2 \pm 0.0</td>
<td>0.0 \pm 0.0</td>
</tr>
<tr>
<td><strong>alpgen</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Z \rightarrow \mu^+\mu^-)</td>
<td>60117 \pm 69.2</td>
<td>88.9 \pm 1.0</td>
<td>59408 \pm 687</td>
<td>96.1 \pm 1.0</td>
<td>57702 \pm 647</td>
<td>98.4 \pm 1.1</td>
<td>54435 \pm 655</td>
<td>98.7 \pm 1.2</td>
</tr>
<tr>
<td>(W \rightarrow \mu\nu)</td>
<td>106 \pm 4.7</td>
<td>0.2 \pm 0.1</td>
<td>22 \pm 20</td>
<td>0.0 \pm 0.0</td>
<td>22 \pm 20</td>
<td>0.0 \pm 0.0</td>
<td>22 \pm 20</td>
<td>0.0 \pm 0.0</td>
</tr>
<tr>
<td>QCD((b\bar{b}))</td>
<td>5390 \pm 2201</td>
<td>6.6 \pm 2.7</td>
<td>1225 \pm 541</td>
<td>2.0 \pm 0.9</td>
<td>0.0 \pm 898</td>
<td>0.0 \pm 1.5</td>
<td>0.0 \pm 8980</td>
<td>0.0 \pm 1.6</td>
</tr>
<tr>
<td>(\tau^+\tau^-)</td>
<td>1961 \pm 177</td>
<td>2.9 \pm 0.3</td>
<td>1163 \pm 133</td>
<td>1.9 \pm 0.3</td>
<td>918 \pm 117</td>
<td>1.6 \pm 0.2</td>
<td>682 \pm 97</td>
<td>1.2 \pm 0.2</td>
</tr>
<tr>
<td>(Z \rightarrow \tau^+\tau^-)</td>
<td>12 \pm 11</td>
<td>0.0 \pm 0.0</td>
<td>5 \pm 3</td>
<td>0.0 \pm 0.0</td>
<td>5 \pm 3</td>
<td>0.0 \pm 0.0</td>
<td>2 \pm 2</td>
<td>0.0 \pm 0.0</td>
</tr>
</tbody>
</table>

#### 7.3.3 Correction from Parton to Particle Level

The MCFM predictions have been corrected with respect to non-perturbative effects (fragmentation and underlying event) using similar corrections as those determined for the electron case.

#### 7.3.4 Comparison of Event Generators and MCFM at the Hadron Level

Figure 9 (a) shows a comparison of the inclusive \(Z \rightarrow \mu^+\mu^- + \text{Njets}\), cross section (shown as the number of events expected for 1 fb\(^{-1}\)) as predicted by the event-generator Alpgen by Pythia, and by MCFM. Alpgen, as well as Pythia, predicts fewer events than MCFM in all jet multiplicity bins. Since it is
Figure 8: Distribution of signal and backgrounds jet multiplicities, for Pythia (a) and Alpgen (b), with no invariant mass applied. Jet $p_T$ and $\eta$ distributions for signal and backgrounds using Pythia (c),(d) and Alpgen (e),(f) data samples. $\eta$ distributions are for high-$E_T$ jets ($p_T > 40$GeV).
not the purpose of this analysis to determine the inclusive $Z \rightarrow \mu \mu$, cross section, the event generator output can be normalized globally to the inclusive $Z \rightarrow \mu \mu$ cross section observed in data. For now, the cross sections of all Alpgen $Z \rightarrow \mu \mu$ samples and of the Pythia $Z \rightarrow \mu \mu$ sample are normalized to the NLO inclusive cross sections. The derived normalization factors are: Alpgen, 1.328±0.056; Pythia, 1.049±0.056.

Figure 9 (b) shows the comparison of the inclusive $Z \rightarrow \mu \mu + \geq N$ jets cross sections after the normalization. Figure 9 also shows the distribution of the $p_T$ of the leading jet (c) and next-to-leading jet (d), predicted by Alpgen, by Pythia (at the truth level of fully-simulated samples, normalized to the inclusive cross section) and by MCFM, for an integrated luminosity of 1 fb$^{-1}$. Pythia predicts a higher jet multiplicity in the low-$p_T$ region ($p_T < 100$ GeV), while Alpgen predicts more high-$p_T$ jets.

### 7.3.5 Unfolding Detector Effects

The reconstructed data has been unfolded to the hadron level, correcting for efficiency losses, resolution and non-linearities in both muon and jet reconstruction, in a similar manner as for the electron channel. The unfolding procedure has been derived (and validated) with fully-simulated Monte Carlo. Figure 10 (a) shows the $p_T$ distribution of the unfolded reconstructed leading jet (Alpgen and Pythia Monte Carlo) compared to NLO and LO MCFM predictions at hadron level. For $p_T > 150$ GeV, the Pythia and Alpgen distributions can be clearly distinguished. Figure 10 (b) shows the unfolded $p_T$ distribution for the reconstructed next-to-leading jet (Alpgen and Pythia Monte Carlo) compared to NLO and LO MCFM predictions at hadron level. In this case, Pythia predicts a noticeably softer $p_T$ distribution than either Alpgen or MCFM.

### 7.3.6 Systematic Errors

Figure 11 shows the relative systematic error expected for the N jet total cross section for different uncertainties on the jet energy scale (JES) (1%, corresponding to the ultimate, but challenging Atlas goal, 3%, a level expected within the first two years and 10%, a possibility for the earliest running). If a 3% error on the JES is achieved, the result is a systematic error of up to 10% for events $\geq$4 jets. For the 10% scenario, the systematic errors are up to 40% for $\geq$ 4 jet and 25% for $\geq$ 3 jet events. If the JES uncertainty figure of 1% is achieved, the systematic errors decrease to less than 0.5%.
7.4 $Z \rightarrow \mu^+\mu^- + b$ Jets

The measurement of the $Z + b$ jet production cross section at the LHC will provide an important test of perturbative quantum chromodynamics (QCD). The cross section is sensitive to the $b$ quark density of the proton and its precise measurement will help in reducing the current uncertainty on the partonic heavy flavor content of the proton (PDF’s). Such uncertainties are currently affecting the potential for the discovery of a number of new physics channels at the LHC.

Table 5 shows the next-to-leading order cross-sections for $Z +$ jets at the Tevatron and LHC [20]. As we can see from this table, at the LHC, the process $gb \rightarrow Zb$ clearly dominates w.r.t. $q\bar{q} \rightarrow Zb$. The selection of $Z+b$ jet events should be easier at the LHC than at the Tevatron for two reasons: we will have a cross section for $Z+b$ jet a factor 50 larger than at the Tevatron, and the relative production of $Z+c$ jet will be less important at the LHC, thus reducing the probability of jet mis-tagging. With the present analysis, we provide a preliminary estimate of the selection efficiency for $Z \rightarrow \mu^+\mu^- + b$ jets events and for the main background channels. We also show the rejection power for $Z \rightarrow \mu^+\mu^- + c$ and light jets, as a function of the $b$-tagging efficiency.
Figure 10: The transverse momentum of the leading jet in unfolded Pythia and Alpgen signal Monte Carlo shown along with predictions from MCFM for $\int L dt = 1 fb^{-1}$.

<table>
<thead>
<tr>
<th>Cross Section(pb)</th>
<th>Tevatron</th>
<th>LHC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process</td>
<td>ZQ inclusive</td>
<td></td>
</tr>
<tr>
<td>$g b \rightarrow Zb$</td>
<td>$13.4 \pm 0.9 \pm 0.8 \pm 0.8$</td>
<td>$1040^{+70+70+30}_{-60-100-50}$</td>
</tr>
<tr>
<td>$q \bar{q} \rightarrow Z\bar{b}b$</td>
<td>$6.83$</td>
<td>$49.2$</td>
</tr>
<tr>
<td>$g c \rightarrow Zc$</td>
<td>$20.3^{+1.8}<em>{-1.5} \pm 0.1^{+1.3}</em>{-1.2}$</td>
<td>$1390 \pm 100^{+60+40}_{-70-80}$</td>
</tr>
<tr>
<td>$q \bar{q} \rightarrow Zc^{\pm}c^{-}$</td>
<td>$13.8$</td>
<td>$89.7$</td>
</tr>
<tr>
<td>$g q \rightarrow Zq , q \bar{q} \rightarrow Zg$</td>
<td>$1010^{+44+9+7}_{-40-2-12}$</td>
<td>$15870^{+900+60+300}_{-600-300-500}$</td>
</tr>
</tbody>
</table>

Table 5: The next-to-leading-order inclusive cross sections (in pb) for $Z$ plus jets at the LHC ($\sqrt{s} = 14 TeV\, pp$) and at the Tevatron ($\sqrt{s} = 1.96 TeV\, p\bar{p}$). The calculations are for the case of a jet in a range $p_T > 15 GeV$ and $|\eta| < 2.5$ (LHC) or $|\eta| < 2.0$ (Tevatron).

### 7.4.1 Event Selection

We require two muons of opposite charge, with $p_T > 15$ GeV, $|\eta| < 2.4$ and $\Delta R > 0.4$ from any jet axis. The $Z$ selection is done by requiring $M_{\mu\mu}$ in the range $M_{Z} \pm 20$ GeV. We select Cone04 jets with $|\eta| < 2.5$ and require $N_{jets} \geq 1$. The analysis has been performed for two jet energy cuts: $E_{T}^{jet} > 20$ GeV and $E_{T}^{jet} > 40$ GeV. The b-tagging is performed cutting on the weight parameter, namely the combination of the secondary vertex (SV1) and impact parameter (IP3D) algorithms. In this analysis we cut at weight $\geq 5$.

Figure 12 (a) shows the transverse momentum spectrum for all the muons found in the various data samples, before any cut is applied, while Figure 12 (b) shows the invariant mass of the two muons, after all the muon identification cuts. The plots are normalized to the number of entries. Figure 13 (a) shows the b-tagging efficiency as a function of the weight parameter; cutting at weight $\geq 5$, as we do in this analysis, ensures 60% efficiency for b jets. Figure 13 (b) shows the rejection factor for c and light quarks as a function of the b-tagging efficiency; 60% efficiency gives a rejection factor of 7 for c jets and 110
Figure 11: Relative systematic errors expected on the measured cross section from different uncertainties on the Jet Energy Scale.

for light jets, in fairly good agreement with what is found for other physics channels.

Figure 12: Muon transverse momentum spectrum, before any cut (left). Invariant mass of the two muons, after all the muon cuts (right).

Figure 14 (a) shows the jet transverse energy spectrum for b, c and light jets in the $Z \rightarrow \mu \mu$ inclusive data sample, after all muon cuts, including the cut on $M_{\mu\mu}$, and for $|\eta^{jet}| < 2.5$ and $E_T^{jet} > 20$ GeV. Figure 14 (b) shows the same events of (a) when the cut on the b-tagging weight is also added. A summary of the efficiency for the various muons cuts, for all the samples is shown in Table 6.

The final efficiency after additional cuts on jets is shown in Table 7 for the Z inclusive, Z+b jets and $Z+c$ and light jets samples. Those efficiencies are calculated requiring the transverse energy of the jet, $E_T^{jet} > 20$ or $E_T^{jet} > 40$.

Given the final efficiencies quoted in this table, for an integrated luminosity of $1 fb^{-1}$ and a signal cross section of 41.25 pb, we should be able to select $5360 \pm 230$ events when cutting at $E_T^{jet} > 20$ GeV and $2890 \pm 170$ events when cutting at $E_T^{jet} > 40$ GeV. The contamination from $Z+c$ and light jets should be of the order of 30% while the background from $t\bar{t}$ and $W \rightarrow \mu \nu$ is negligible.
Figure 13: The b-tagging efficiency as a function of the weight parameter calculated for the sample Z+b jet (left). The rejection factor for c (black circles) and light (black stars) jets as a function of the b-tagging efficiency (right).

<table>
<thead>
<tr>
<th>Cuts</th>
<th>Z inclusive</th>
<th>Z+b jets</th>
<th>Z+c or light jets</th>
<th>ttbar</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>at least 1 ( \mu^+ ) and 1 ( \mu^- )</td>
<td>56.9 ± 0.3%</td>
<td>91 ± 2%</td>
<td>93 ± 3%</td>
<td>11.4 ± 0.1%</td>
<td>2.51 ± 0.09%</td>
</tr>
<tr>
<td>( p_T^{\mu} &gt; 15 ) GeV</td>
<td>32.6 ± 0.2%</td>
<td>51 ± 1%</td>
<td>53 ± 2%</td>
<td>1.24 ± 0.04%</td>
<td>0.09 ± 0.01%</td>
</tr>
<tr>
<td>(</td>
<td>\eta^{\mu}</td>
<td>&lt; 2.4 )</td>
<td>30.9 ± 0.2%</td>
<td>49 ± 1%</td>
<td>50 ± 2%</td>
</tr>
<tr>
<td>( \Delta R_{\mu-jet} &gt; 0.4 )</td>
<td>70 GeV &lt; ( M_Z ) &lt; 110 GeV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6: Efficiency after the cuts to identify the muons from Z decay

8 Measurement of W + Jet Cross-Sections

8.1 \( W \to e \nu \) + Jets

The following analysis focuses on \( W(\to e \nu) + \text{jets} \) events, especially in regards to an evaluation of the Jet Energy Scale (JES) and Parton Distribution Function (PDF) uncertainties. Observable distributions will be shown for signal and backgrounds at detector level, after a standard cut-based selection. A data driven method to remove backgrounds will also be discussed. The impact of the jet energy scale on jet multiplicity distributions will be examined. The PDF reweighting technique will then applied to predict PDF uncertainties on electron and jet distributions and the PDF uncertainties on the measurement of the cross section, as a function of the cumulative jet multiplicity, will be compared to the experimental systematic uncertainties.

8.1.1 Signal-Background Analysis

The number of signal events and the fraction of background events after the selection cuts of Section 4 have been applied are given in Table 8, using the datasets mentioned in Section 3 and assuming an integrated luminosity of 1 fb\(^{-1}\).

The table indicates that the largest contribution to the \( W \to e \nu \) background comes from QCD di-jet production. It has to be noted that this QCD background is, among all background processes, the least understood both theoretically and experimentally. A rather large uncertainty (factor ≈ of 2-3) on the level of this background is expected due to the large theoretical uncertainty. Furthermore, due to the
<table>
<thead>
<tr>
<th>Cuts</th>
<th>$Z$ inclusive</th>
<th>$Z+b$ jets</th>
<th>$Z+c$ or light jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of jets $&gt;0$</td>
<td>$18.9 \pm 0.1%$</td>
<td>$36 \pm 1%$</td>
<td>$25 \pm 1%$</td>
</tr>
<tr>
<td>$E_T^{jet} &gt; 20$ GeV &amp; $</td>
<td>\eta^{jet}</td>
<td>&lt; 2.5$</td>
<td>$10.26 \pm 0.08%$</td>
</tr>
<tr>
<td>b jet weight $&gt;5$</td>
<td>$0.24 \pm 0.01%$</td>
<td>$12.8 \pm 0.2%$</td>
<td>$0.2 \pm 0.1%$</td>
</tr>
<tr>
<td>$E_T^{jet} &gt; 40$ GeV &amp; $</td>
<td>\eta^{jet}</td>
<td>&lt; 2.5$</td>
<td>$3.94 \pm 0.05%$</td>
</tr>
<tr>
<td>b jet weight $&gt;5$</td>
<td>$0.072 \pm 0.005%$</td>
<td>$7.1 \pm 0.5%$</td>
<td>$0.2 \pm 0.1%$</td>
</tr>
</tbody>
</table>

Table 7: The final event reconstruction efficiency after all the cuts for the selection of $Z \rightarrow \mu^+\mu^- + b$ jets.

Figure 14: The jet transverse energy spectrum for $b$, $c$ and light jets in the $Z \rightarrow \mu\mu$ inclusive data sample, before (left) and after (right) the cut on the b-tagging weight.

large di-jet cross-section and to the high rejection power of the selection criteria, the available Monte Carlo statistics are not sufficient to evaluate the number of QCD background events by simply counting the events after all cuts to the Monte Carlo QCD background sample. To overcome this difficulty the QCD background contamination is obtained by using the product of the offline selection efficiency only, times the trigger selection efficiency. To properly take into account the correlation between these two quantities, the trigger selection efficiency is calculated as the ratio of the efficiency of the trigger and electron identification cuts and the electron identification efficiency.

In Figure 15, we see the $E_T$ spectrum for all jets in the events for the signal, $W \rightarrow e\nu$, and the backgrounds, after the full selection. Despite the poor Monte-Carlo statistics for QCD events (giving rise to the observed “spikiness??), the Figure clearly indicates the increasing importance of the QCD and $t\bar{t}$ backgrounds at higher jet $E_T$ with respect to the steeply falling signal distribution.

In Figure 16, we can see the expected number of events as a function of the jet multiplicity in any given event, for an integrated luminosity of 1 fb$^{-1}$. Figure 16 also shows the corresponding cumulative jet multiplicity. The term cumulative multiplicity refers to events with at least $n$ jets in the final state, so that $n = 0$ corresponds to the inclusive $W \rightarrow e\nu$ sample. The distributions show that the $t\bar{t}$ background dominates at large jet multiplicities, while QCD di-jet production is the largest background for small multiplicities.

As has been noted previously, however, the QCD background for this analysis has been estimated by using PYTHIA di-jet processes, which are known to underestimate the level of multi-jet production. The
Table 8: Number of signal events selected for an integrated luminosity of 1 fb$^{-1}$. For the background samples, the ratio $N_b/N_s$ in percent is indicated. The quoted uncertainties are due to the Monte-Carlo statistics.

<table>
<thead>
<tr>
<th>Selection</th>
<th>$W \rightarrow e\nu$ ($N$)</th>
<th>QCD ($f$ (%))</th>
<th>$W \rightarrow \tau\nu$ ($f$ (%))</th>
<th>$Z \rightarrow ee$ ($f$ (%))</th>
<th>$t\bar{t}$ ($f$ (%))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trigger/offline e-id + $E_T &gt; 25$ GeV</td>
<td>$(1.053 \pm 0.001) \cdot 10^6$</td>
<td>$10 \pm 3$</td>
<td>$8.0 \pm 0.1$</td>
<td>$1.57 \pm 0.05$</td>
<td>$6.25 \pm 0.04$</td>
</tr>
</tbody>
</table>

Figure 15: The plot in the left shows the W transverse mass distribution for signal and background normalized to 1 fb$^{-1}$: $W \rightarrow e\nu$ (white), QCD (yellow), $W \rightarrow \tau\nu$ (red), $Z \rightarrow ee$ (green), $t\bar{t}$ (blue). The plot on the right shows the jet $E_T$ distribution for signal and background normalized to 1 fb$^{-1}$: $W \rightarrow e\nu$ (white), QCD (yellow), $W \rightarrow \tau\nu$ (red), $Z \rightarrow ee$ (green), $t\bar{t}$ (blue). Histograms are cumulative.

The real amount of QCD background can only be reliably estimated from LHC data.

8.1.2 Data-Driven Methods for the Background Extraction

Whereas the $W \rightarrow \tau\nu$ and $Z \rightarrow ee$ backgrounds can be reliably estimated by Monte Carlo techniques, the $t\bar{t}$ (to some extent) and especially the QCD di-jet backgrounds can only be estimated by using LHC data. Thus, to minimize the uncertainty on the amount of background, an alternative, mostly data-driven, selection procedure is proposed. The $E_T$ cut is efficient in rejecting the backgrounds, but results in a loss of information of the QCD distribution shape. The amount and shape of the background under the signal area will be hard to estimate. Therefore, a data-driven approach to reject the background has been developed (see the inclusive $W/Z$ CSC note for a detailed description of this technique).

In this approach, $Z \rightarrow ee$ events are efficiently removed by cutting on the invariant mass of an $e^+e^-$ pair, as shown in Figure 16. The dominant QCD background can be parameterized and subsequently subtracted from the signal area by using a pure QCD di-jet sample. To accomplish this, photon candidates are selected which pass a photon trigger and the same calorimeter-based identification as electrons. Photon candidates are distinguished from electrons as no good quality track is found to be associated with the electro-magnetic cluster. Applying some additional requirements on track isolation, this sample is largely dominated by QCD jets faking a photon and can be parameterized as a function of the cell-based $E_T$. As the electron and photon samples are kinematically very similar, the $E_T$ spectrum for the QCD background has the same shape in both samples. The background parameterization coming from
Figure 16: Jet multiplicity (upper left) and cumulative jet multiplicity (upper right) distributions for signal and backgrounds. Transverse mass distributions before (lower left) and after (lower right) the rejection of the $Z \rightarrow ee$ background (in green). Plots are normalised to 1 fb$^{-1}$ integrated luminosity.

the photon sample is then normalized to the electron sample. The normalization factor is obtained using, as a side-band, the low $E_T$-region in the electron sample, where QCD is dominating. Only after having subtracted the QCD background, is a cut on $E_T$ applied to remove the side-band.

In order to obtain a measurement of the $W \rightarrow e\nu$ cross section as function of the jet multiplicity, the above QCD-subtraction method can be applied for each given jet multiplicity. However, the limited Monte Carlo statistics of QCD events prevents us from applying this method in this analysis, but the principle has been tested in the inclusive cross section analysis.

### 8.1.3 Jet Energy Scale Uncertainties

The dominant experimental systematic uncertainty for $W +$ jets production, as for the $Z +$ jets case, is given by the jet energy scale (JES) uncertainty. The effect of jet energy scale uncertainties on the cross section as a function of cumulative jet multiplicity is investigated in the following section. As before, a mis-calibration of $\pm 1\%$, $\pm 3\%$ and $\pm 5\%$ is assumed for the jet energy and its effect on the jet transverse momentum is calculated. Because of time constraints, the jet energy is miscalibrated after having run the jet reconstruction algorithm. The transverse momentum is then recalculated from the miscalibrated energy and the pseudorapidity. This simple method allows a quick estimate of the uncertainty on the cross section. It is sufficient for the scope of this study, where the aim is a comparison with the uncertainties due to PDFs. After the miscalibration, the event selection is applied and the jet multiplicity distribution is plotted for each value of the assumed miscalibration. From these distributions the error on the cross section is calculated.

Figure 17 shows the cross section as a function of the leading jet transverse momentum and the
cumulative jet multiplicity with the experimental uncertainty deriving from a ±1%, ±3% and ±5% energy miscalibration. Figure 17 indicates how the JES uncertainties increase significantly with the transverse momentum of the leading jet, rising above 30% for highly energetic jets.

The main effect on the cross section as a function of the jet multiplicity results from the selection cut of 20 GeV applied on the jet transverse momentum. If the energy is mis-calibrated by a positive correction, more jets pass the selection cut and therefore there will be more high jet multiplicity events and vice-versa when the energy is miscalibrated by a negative correction. The relative uncertainties on the measurement are shown in Figure 17. When the jet energy is miscalibrated by 1%, the experimental systematic uncertainties on the cross section remain within 5%, but they increase significantly with higher miscalibration, rising above 20%. Uncertainties show a tendency to increase with the multiplicity. The first bin with zero jets is biased by the event filter cut applied at generator level and is therefore not representative.

8.1.4 PDF Reweighting Technique

The effect of PDF uncertainties on physical distributions is investigated applying PDF re-weighting. This is a powerful technique used in Monte Carlo simulations to evaluate PDF uncertainties on a physical quantity in a fast way. In fact, applying an event weight to the hard process, predictions for one PDF set can be derived from a sample generated with a different PDF set. The event weight is a function of momentum fraction $x$ of the two incoming partons and the scale $Q^2$ of the interaction:

$$ EW = \frac{f_{pdf_1}(x_1, Q^2)}{f_{pdf_1}(x_1, Q^2)} \frac{f_{pdf_2}(x_2, Q^2)}{f_{pdf_1}(x_2, Q^2)} $$  (1)
The scale $Q^2$ is calculated from the mass ($M_W$) and transverse momentum ($p_{tW}$) of the $W$ boson, consistent with the choice used in the MCFM and ALPGEN cross sections used in this note:

$$Q^2 = M_W^2 + p_{tW}^2$$ (2)

The accuracy of PDF re-weighting on events generated with Herwig was investigated in previous studies and found to be better than 1%. A similar analysis was performed on Alpgen events to test the accuracy of reweighting from a LO to a NLO PDF set. Two samples of 500,000 $W + 3\text{ partons}$ events were generated with two different PDF sets (CTEQ6LL and MRST2001NLO). Events generated with CTEQ6LL were then reweighed to MRST2001NLO and compared to those generated directly with MRST2001NLO. It was shown that the PDF reweighting technique is accurate to be better than 1% in the measurable kinematic regions.

8.1.5 PDF Uncertainties

The CTEQ group provides error sets for the Next to Leading Order PDF set CTEQ6M. Since the datasets used in this CSC note are generated with the PDF set CTEQ6LL, re-weighting is applied from CTEQ6LL to the central value of CTEQ6M and the 40 error sets corresponding to the upper and lower uncertainty on each of the 20 eigenvectors. The upper around the NLO central value are calculated using the asymmetric error formula recommended in [3].

Figure 18 shows the electron $p_T$ and $\eta$ distributions determined by the CTEQ6M central value with the upper and lower PDF uncertainties. The relative errors are shown in the same Figure. PDF uncertainties are of the order of 5% and exhibit a dependence on both $p_T$ and $\eta$. They increase at very low and very high transverse momentum, where statistical fluctuations also become important. Uncertainties are also larger in the central $\eta$ region, but remain well below 10%.

Similar results are obtained for jet variables. Figure 19 shows that the relative uncertainties on leading jet $p_T$ and jet $\eta$ are approximately 3-5% but can rise above 5% at very small and large $p_T$ and small $\eta$. An increase is also observed at the edges of the pseudorapidity spectra, outside the electron measurable $\eta$ region.

8.2 $W \rightarrow \mu\nu+$ Jets

This section describes the $W \rightarrow \mu\nu+$ jets analysis with a brief look at the relative contributions from signal and backgrounds. The $W \rightarrow \mu\nu+$ jets signal is simulated with the Alpgen generator followed by showering with the Jimmy program. Exclusive matching samples with up to 5 partons are combined in proportion to the relative cross sections to yield a combined signal sample. The calculated contributions are scaled to a 1 fb$^{-1}$ dataset. We focus on the following four background processes: di-jet events (the so-called QCD background), $t\bar{t}$, $W \rightarrow \tau\nu\tau$ and $Z \rightarrow \mu\mu$. The details of the signal and background datasets are described in Section 3.

Events consistent with $W \rightarrow \mu\nu+$ jet production are expected to pass the single muon 20 GeV trigger. Exactly one muon with $p_T > 20$ GeV and $|\eta| < 2.4$ must be isolated; specifically the energy in the $\Delta R = 0.2$ cone around the muon cannot exceed 15 GeV (cone0.20 isolation). True $W$ events should exhibit significant missing transverse energy from the neutrino; we require $E_T > 25$ GeV. Finally, we count jets with $E_T > 20$ GeV and $|\eta| < 5.0$, but we remove jets which are within $\Delta R = 0.4$ of the candidate muon.

8.2.1 Signal-Background Analysis

The number of signal and background events after the selection are presented in Table 9. Figure 20 shows both the expected total jet multiplicity and the jet $E_T$ distribution for selected signal and back-
Figure 18: PDF Uncertainties on the electron $P_t$ and $\eta$ at detector level, after a $P_t$ cut of 25 GeV. Upper left: Distribution before (blue) and after reweighting to the central CTEQ6M set (black) with upper (green) and lower (red) uncertainties. Upper Right: Relative uncertainties on the same distribution. Lower left: Upper (green) and lower (red) uncertainties on the $\eta$ spectra. Lower Right: Uncertainties expressed as percentage.

Table 9: Number of signal and background events selected for an integrated luminosity of 1 fb$^{-1}$. For the background samples, the ratio of the number of background over the number of signal events ($N_B/N_S$) is given in percent.

<table>
<thead>
<tr>
<th>$W \rightarrow \mu \nu$</th>
<th>QCD</th>
<th>$W \rightarrow \tau \nu$</th>
<th>$Z \rightarrow \mu \mu$</th>
<th>$t \bar{t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.490 $\cdot 10^6$</td>
<td>168.0</td>
<td>11.5</td>
<td>1.6</td>
<td>5.8</td>
</tr>
</tbody>
</table>

9 Conclusions

Final states containing $W/Z + \text{ jets}$ will serve as one of the Standard Model benchmarks for physics analyses at the LHC. These states form the signal channels (as well as the backgrounds) for known Standard Model processes such as $t \bar{t}$ production, as well as Beyond Standard Model signals such as supersymmetry. The events will be triggered by the presence of the energetic decay lepton(s) from the $W(Z)$. The presence of additional jets in the events will tend to boost the $W/Z$, and thus the decay leptons, to higher transverse momentum, leading to a larger acceptance. The presence of the jets will also tend to negatively affect the lepton analysis cuts, such as on isolation. The two effects roughly cancel out, resulting in total efficiencies in general a few percent lower than for the inclusive analyses.

$Z \rightarrow ee, \mu \mu + \text{ jets}$ events suffer from backgrounds involving real electrons and muons (from $t \bar{t}$,
Figure 19: PDF uncertainties on the leading jet transverse momentum and $\eta$ at detector level. Upper left: $P_t$ spectra with upper (green) and lower (red) uncertainties. Upper right: Relative uncertainties on jet $P_t$. Lower left: Upper (green) and lower (red) uncertainties on the $\eta$ spectra. Lower right: Uncertainties expressed as percentage.

$Z \rightarrow \tau\tau$ and $W \rightarrow e\nu(\mu\nu)$ and from backgrounds involving fake electrons and muons (QCD jet production). The former can be reliably estimated from current Monte Carlo predictions, although in situ verification with data will be needed, while the latter, dealing with rare fluctuations of large cross section processes can only be crudely estimated from current Monte Carlo simulation. An estimate for the QCD background has been made from the existing Monte Carlo event samples, and a scheme to determine the background from the real data has been outlined. The backgrounds increase with jet multiplicity, with $t\bar{t}$ being the largest background for high jet multiplicities, and with QCD jet production dominating the background for low jet multiplicities. The backgrounds for $W \rightarrow e\nu,\mu\nu + \text{jets}$ events result from similar sources as for the $Z$ case (with the $Z \rightarrow ee,\mu\mu$ background replacing the $W \rightarrow e\nu(\mu\nu)$ one). With the cuts used in the analyses presented in this note, the $W$ and $Z + \text{jets}$ signal is substantially larger than the sum of the backgrounds, while maintaining reasonable efficiency. Experience with actual data will help to improve this discrimination further.

In this note, we have considered cross section measurements for theoretically well-defined quantities. The cross sections for $W/Z + \text{jets}$ in ATLAS will be quoted at the hadron level, corrected for all detector measurement effects. An unfolding technique, from the detector to the hadron level, which can be used with actual data as well as with fully simulated Monte Carlo events, has been developed for this note. The corrections needed to between the parton and hadron level, necessary for comparisons to parton level predictions have also been determined. The two main non-perturbative effects, due to the underlying event and to the jet fragmentation, result in corrections in opposite directions that partially cancel, and in any case are very small for jets with transverse momenta larger than 40 GeV/c.

The dominant systematic error in the cross section measurements will be due to the uncertainty in the jet energy scale. We have considered the impact of jet energy scale uncertainties of 10%, as expected in the very early running, of 3-5%, as might be achieved after 2 years of running, and of 1%, an optimistic
Figure 20: Jet multiplicity (left) and jet $E_T$ (right) distributions for signal and background normalised to $1 \text{ fb}^{-1}$. The contribution from $W \rightarrow \mu \nu$ signal (blue) is shown with the contamination from jet events (green), $W \rightarrow \tau \nu$ (yellow), $Z \rightarrow \mu \mu$ (magenta) and $t\bar{t}$ (red).

Figure 21: $W$ transverse mass (left) and missing $E_T$ (right) distributions for signal and background normalised to $1 \text{ fb}^{-1}$. The contribution from $W \rightarrow \mu \nu$ signal (blue) is shown with the contamination from jet events (green), $W \rightarrow \tau \nu$ (yellow), $Z \rightarrow \mu \mu$ (magenta) and $t\bar{t}$ (red).

goal after the detector is well-understood. The resultant experimental uncertainties have been compared to one of the dominant theoretical uncertainties, that due to uncertainties in parton distribution functions. Within the timeframe considered in this note (the first two years of running), the JES errors will dominate over the PDF uncertainties for most of the cross sections to be measured.

References

[1] ATLAS Collaboration, W/Z inclusive cross section CSC note, this volume


[9] ATLAS Collaboration, Electron ID CSC note, this volume

[10] ATLAS Collaboration, L1 Calorimeter trigger performance CSC note, this volume


[12] ATLAS Collaboration, Missing transverse energy CSC note, this volume


[18] ATLAS Collaboration, Electron-Gamma trigger performance CSC note, this volume

