Data Driven Estimate of W+Jets Background in H->WW(*)->lvlv analysis using 4.7 fb^-1 of data collected with the ATLAS detector.

Aysha Abdel-Aziz

Supervisors: William Trischuk, Pierre Savard, Nicola Venturi

IPP SUMMER STUDENT REPORT

Introduction

The July 2012 discovery of a neutral boson compatible with the production and decay of a Standard Model Higgs boson at CERN [1] generated much excitement in the physics and broader community. Further studies of data being collected by the ATLAS experiment at the LHC are underway to investigate the properties of the newly discovered boson and to determine whether it is in fact a Standard Model Higgs boson whose discovery would confirm the Higgs Mechanism and complete the Standard Model (SM). One of the decay modes of the SM Higgs boson is to a W boson particle-antiparticle pair where one of the W bosons may be off-shell. The $H \rightarrow WW(*) \rightarrow l\nu l\nu \;\; (l = e, \mu)$ channel is one of the most sensitive channels for a SM Higgs mass of 126.0 GeV. One of the dominant backgrounds to this channel is the leptonic decay of W-bosons that are produced in association with jets. Unlike the other dominant backgrounds, the W+jets contribution is estimated by a data-driven method rather than by Monte Carlo (MC) simulations. This paper presents the nominal data-driven method and the results of modifying the method by using a Z+jets control region (CR) in place of a QCD multijet control region. The analysis was applied to 4.7 fb^{-1} of data collected with the ATLAS detector at $\sqrt{s}$ = 7 TeV. Monte Carlo simulations were used for ZZ, ZW, Wγ*, Z+(0-5) partons, and Z+$b\bar{b}$+(0-3) partons leptonic decay events.

Much gratitude is owed to my supervisors Nicola Venturi, Pierre Savard, and William Trischuk for their help and mentorship throughout the project. Additional thanks to Tatsuya Masubuchi and Keisuke Yoshihar from the University of Tokyo as well as to John Allison and Doug Schaefer from the University of Pennsylvania who

conduct W+jets background studies for the HS3G group and who met with me several times and provided me with essential information to ensure that this analysis employed a method consistent with that of one of the parallel H3G3 Z+jets fake factor analyses.

$H \rightarrow WW(*) \rightarrow l\nu l\nu$ Event Candidates

The signature of $H \rightarrow WW(*) \rightarrow l\nu l\nu \;\; (l = e, \mu)$ channel is a pair of isolated high P_T charged leptons. Events are required to have exactly two identified leptons of opposite electric charge, one of which must match a triggering object. The leading and sub-leading leptons must have a minimum P_T of 25 GeV and 15 GeV respectively. The cuts $\Delta\Phi_{ll} < 1.8$ and $M_{ll} < 50$ GeV ($M_{ll} < 80$ GeV for events with jet multiplicity greater than two) are applied to the dilepton pair to identify the event topology that arises from the Higgs boson's zero spin. Note that $\Delta\Phi_{ll}$ is the difference in the azimuthal angle of the leptons and M_{ll} is their invariant mass. Further cuts on M_{ll} and E^T_{miss}, the magnitude of the missing transverse momentum, are applied to suppress QCD multijet and Drell-Yan background. [2]

The leptons must satisfy nominal identification criteria to be included in the dilepton pair. Muon candidates have tracks that are reconstructed in the muon spectrometer (MS). A muon candidate is a *Staco* combined muon if tracks are independently reconstructed in the MS and in the inner detector (ID) and then successfully combined into a single track using the *Staco* reconstruction algorithms [3,4]. Identified muons are required to satisfy the ID hit requirements [3], to be *Staco* combined muons, and to be in the region $|\eta| < 2.4$. Electron candidates in the central region ($|\eta| < 2.47$) are reconstructed by matching fixed size *sliding window*

[5, 6] hit clusters in the electromagnetic (EM) calorimeter to tracks of charged particles in the inner detector. Central electrons are reconstructed by the standard electron algorithm designed for high P_T isolated electrons, by the soft electron algorithm designed for electrons in jets, or by both algorithms independently [7]. To distinguish isolated electrons from jets, electrons are categorized as loose, medium, or tight with increasing jet rejection efficiency based on a cumulative set of cuts on calorimeter, track, and track-cluster matching variables [5]. Identified electrons are required to be in the central region (excluding $1.37 < |\eta| < 1.52$), to be tight, and to be reconstructed as standard electrons.

All identified leptons must have a minimum P_T of 15 GeV and to satisfy isolation requirements to further reject low P_T leptons and leptons in jets. Track isolation requires the summed P_T of all charged particle tracks within $\Delta R = \sqrt{\Delta\varphi^2 + \Delta\eta^2} < 0.3$ of the lepton candidate be less than 13% (electrons) or 15% (muons) of the lepton P_T. Similarly, calorimeter isolation requires the summed E_T deposits in the calorimeter within $\Delta R < 0.3$ of the lepton be less than 14% of the lepton P_T; the summed E_T is corrected for pileup and for the lepton's energy losses. Finally, cuts on transverse impact parameter significance $\frac{d0}{\sigma\,(d0)} < 10$ (< 3 for muons) and on the longitudinal impact parameter $|z0| < 1$ mm with respect to the primary vertex suppress heavy flavour decays.

W + jets Background

The primary sources of background to the $H \rightarrow WW(*) \rightarrow l\nu l\nu$ channel are top ($t\bar{t}$ and Wt), WW, and W+jets events [8]. In the case of top and WW backgrounds, a pair of real high P_T isolated leptons with opposite electric charge can be produced from W-decays. In the case of the W + jets background, there is only one real isolated lepton from a W-decay but a misidentified object from a jet may "fake" the second lepton.

Prompt leptons are those that are produced at the primary event vertex. A "fake" lepton is used here to mean a lepton candidate that is a lepton but is non-prompt or a lepton candidate that is not in fact a lepton; a "real" lepton is used here to mean a lepton candidate that is a lepton and is both prompt and produced in isolation. The primary sources of electron candidates in ATLAS are hadrons followed by conversions, semileptonic b- and c-hadron decays, and then finally $W/Z/\gamma*$ leptonic decays [9]. Accordingly, fake electron sources—that is, sources of electron candidates which are not electrons or are non-prompt electrons—include misidentified charged hadrons, conversions, semileptonic heavy flavour decays, π^0 Dalitz decays, and muons. The three most dominant sources are b- and c-semileptonic decay, light flavour jets with a leading π^0 whose track overlaps with a charged particle, and photon conversions [10]. Muon candidates in ATLAS, and the majority of fake muons, come from in-flight kaons or pions and from semileptonic b- and c-hadron decays [11]. All of these sources provide lepton candidates at rates that decay with electron P_T and peak at the minimum lepton P_T (except for the $W/Z/\gamma*$ source which peaks near 40 GeV)[9, 11].

W + Jets Background Estimation

In the 2011 dataset, a data driven method for the W+jets background and Monte Carlo methods for the other backgrounds indicated W+jets constituted (16 ±

9) % of the total estimated background in the 0-jet bin for an expected Higgs mass (M_H) of 125 GeV [2]. The complications of jet-calorimeter interactions, jet shape, and lepton identification algorithms introduce large uncertainties in Monte Carlo simulations of the W + jets background such that the data driven estimate of the W + jets background is used [12]. This data-driven method is described here.

A lepton satisfying the nominal lepton identification criteria will hereafter be referred to as an identified (id-) lepton. A lepton satisfying a loosened set of identification criteria while failing the nominal set of identification criteria will be referred to as an anti-identified (anti-id) lepton. Tables 1 and 2 give detailed lists of these criteria. The observable W + jets control region is defined as the set of events with one identified lepton from a W-decay and one anti-identified lepton. The number of events in the observable W + jets control region is multiplied by the estimated ratio of id-leptons to anti-id leptons in the sample of jets from W+ jets events. The product is the number of W + jets events with two identified leptons—one from a real W decay and one from a jet. It is in this way that the contribution of W + jets events to the $H \rightarrow WW(*)$ background is estimated.

$$\boldsymbol{N_{1\ id(fromW)+1\ fake\ id(from\ jet)} = \frac{N_{id\ (QCD\ CR)}}{N_{anti-id\ (QCD\ CR)}} \cdot N_{1id(from\ W)+1\ anti-id(from\ jet)}}$$

The ratio of id-leptons to anti-id leptons in the W+jets jet sample is called the W+jets fake factor and is estimated by observing fake factors in another jet-rich control region (CR) that is well separated from the signal. In particular, the nominal fake factor is observed in the QCD multijet control region. The uncertainty on this multijet fake factor is the primary source of uncertainty on the W + jets background; the other source is the statistical error on the observable W+jets CR. In the 2011 dataset, the W+jets fake factor introduced 10% and 7% relative systematic uncertainties on total estimated background in the 0-jet and 1-jet bins respectively (M_H = 125 GeV) [2]. The dominant uncertainty on the multijet fake factor is systematic: primarily sample dependence uncertainty with smaller contributions from trigger bias, run dependence, and uncertainty in the cross section of electroweak processes that contaminate the multijet sample with real isolated leptons [2]. The sample dependence uncertainty is assigned as a result of the fake factor being estimated in a different data sample than the W + jets control region to which it is applied.

<table>
<tr><td rowspan="4">Muon</td><td colspan="2">Common Id/Anti-Id Selection Cuts</td></tr>
<tr><td colspan="2">Staco Combined Muon P_T > 15 GeV $|z0| < 1mm$ $|\eta| < 2.4$ Inner Detector Hit Requirements</td></tr>
<tr><td>Identified</td><td>Anti-identified</td></tr>
<tr><td>P_TCone30Corr/P_T < 0.15
E_TCone30Corr/P_T < 0.14
$\frac{d0}{\sigma(d0)} < 3$</td><td>$E_T$Cone30Corr/$P_T$ < 0.3
Satisfies at least one of:
P_TCone30Corr/$P_T \geq 0.15$
E_TCone30Corr/$P_T \geq 0.14$
$\frac{d0}{\sigma(d0)} \geq 3$</td></tr>
</table>

Table 1: id- and anti-id muon criteria

	Common Id/Anti-Id Selection Cuts	
Electron	• Only standard or both standard and soft reconstruction • P_T > 15 GeV • $\|z0\| < 1mm$ • $\|\eta\|$ < 2.47 • 1.37 < $\|\eta\|$ < 1.52 excluded • P_TCone30Corr/P_T < 0.13	
	Identified	**Anti-identified**

	• E_TCone30Corr/P_T < 0.14
	• $\frac{d0}{\sigma\,(d0)} < 10$
	• Tight
		• E_TCone30Corr/P_T < 0.3
		• N_{Hits}(SCT + Pixel) ≥ 4
		• Satisfies at least one of:
		  ○ E_TCone30Corr/P_T ≥ 0.14
		  ○ $\frac{d0}{\sigma\,(d0)} \geq 10$
		  ○ Not Tight

Table 2: Id- and Anti-Id Electron Criteria

Z + Jets Fake Factor Method

In order to reduce the systematic error in the fake factor calculation, the fake factor can be observed in the Z + jets control region rather than in the QCD multijet control region. Due to the similar jet compositions in the dominant Z + jets and W + jets background processes, the sample dependence is expected to be reduced.

Specifically, since the Z+jets and W+jets processes have similar colour structure, the relative contributions to fake leptons from quark and gluon initiated jets is expected to be similar in the two samples. While the quark initiated and gluon initiated jets in the QCD multijet control region have comparable contributions to lepton fakes, the dominant fake lepton sources in the Z + jets and W+jets processes are quark initiated jets. The Z/W+jets samples also have larger fake contributions from photons and real leptons than the multijet sample. The similar Feynman diagrams for the dominant background Z + jets and W + jets production processes are displayed in Figure 1.

As the sample dependence is the largest contribution to the systematic error on the QCD multijet fake factor thus a Z+jets factor is an attractive alternative for estimating the W + jets background in the $H \rightarrow WW(*) \rightarrow l\nu l\nu$ channel. The limitation on using a Z+jets fake factor to reduce systematic error is an increase in the statistical error on the fake factor; the dominant error on Z+jets fake factors therefore will decrease with increasing data statistics whereas the dominant error on the QCD multijet fake factors, being systematic, has a non-zero lower bound that is inherent to the fake factor calculation method.

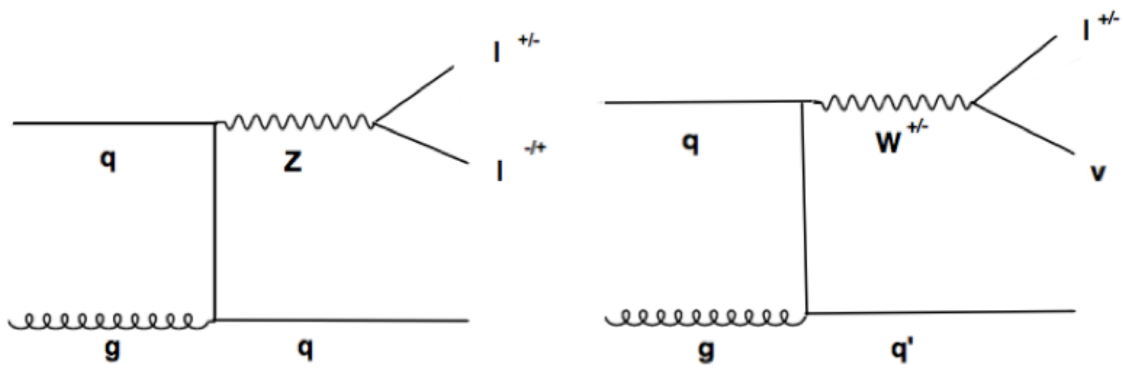

Figure 1: Feynman diagrams for production of Z (left) and W (right) bosons in association with a single quark initiated jet

Z + Jets Control Region Event Selection

The event selection and electroweak (EW) event veto used here were chosen to agree with those of one of the parallel Z+jets studies in the HSG3 group. The objectives of increasing Z+jets data statistics and removing all EW contamination of the Z+jets control region motivate these event selection and veto criteria.

To increase the selection efficiency of events with leptonic Z boson decays, all events tagged with a "medium" Z are included in the unfiltered Z+jets control region. The Z is "medium" in that the electron objects with which the Z is reconstructed need not satisfy all of the nominal id-electron criteria: the tight requirement for electrons is replaced with a medium requirement while the isolation cut is not applied at all. The muons used for Z reconstruction, however, are subject to the nominal id-muon criteria. An event is tagged if from the collection of such leptons a same-flavour and opposite-charge dilepton pair with invariant mass within 15 GeV of the Z mass is found. Furthermore, at least one of the leptons in the dilepton pair must have P_T greater than 25 GeV.

As in the multijet control region, electroweak events produce real leptons primarily from Z and W decays that can bias the fake factor. Accordingly, the Z + jets control region is filtered with event vetoes for ZZ→llll and ZW→lllν decays. Real isolated leptons from electroweak (EW) events contribute predominantly to the id-lepton categories. In the Z+jets control region the low statistics of the id-lepton samples result in fake factors that are extremely sensitive to even a few real leptons from EW events. Figure 2 shows that these few real leptons from EW events are in fact accepted into the unfiltered Z+jets control region.

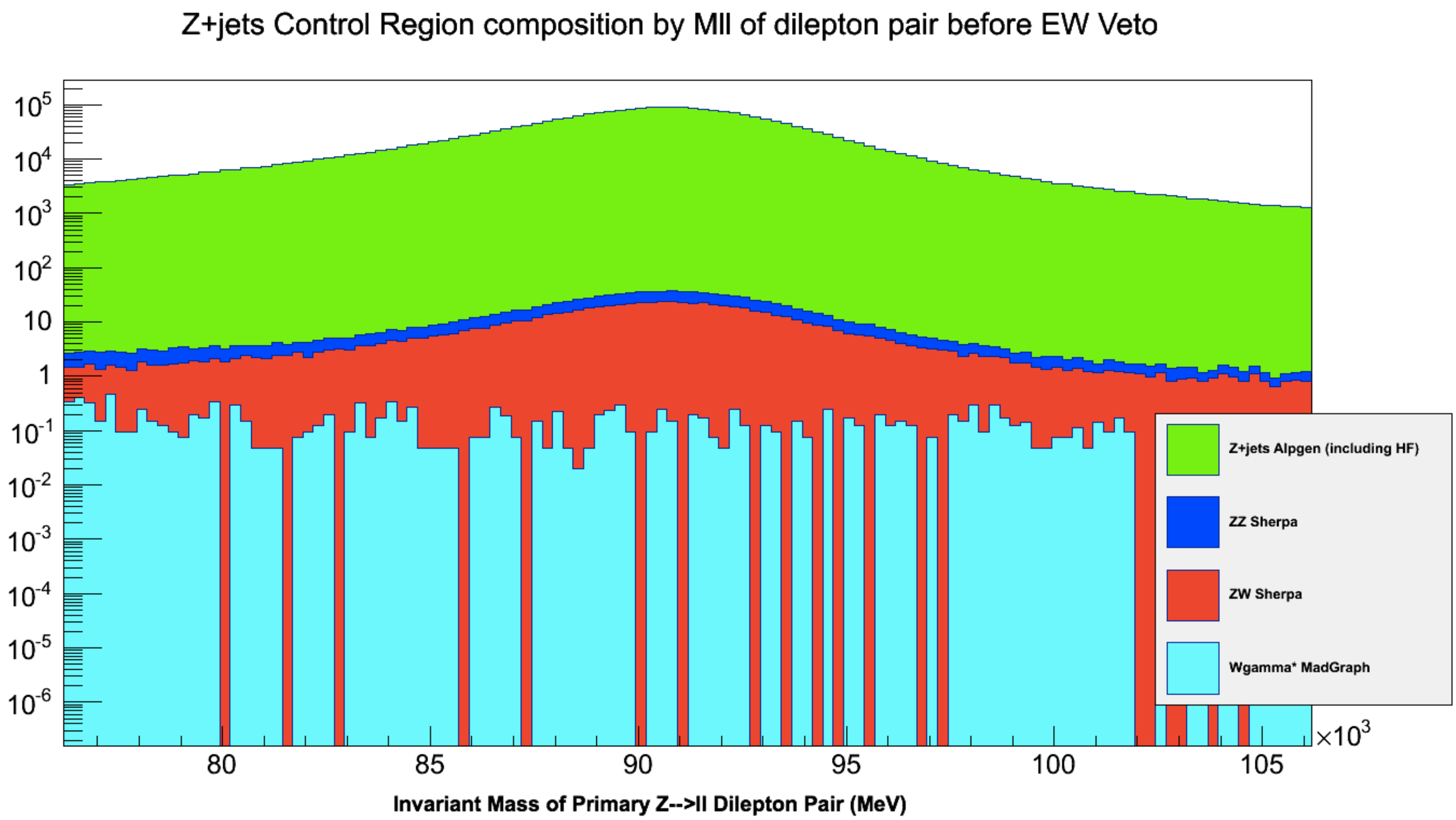

Figure 2: The composition of the Z+jets unfiltered control region using Z+jets, ZZ, ZW, and Wγ* MC samples

The EW event vetoes are therefore designed to apply even looser criteria to a second reconstructed Z or a reconstructed W than is applied to the primary reconstructed Z that tagged the event. The lepton objects with which the second Z or W are reconstructed have no selection criteria applied to them at all other than the removal of the leptons from the first reconstructed Z. In a single event there can be up to 30 such electron objects and up to 2 such muon objects. Any same-flavour opposite-charge dilepton pair with invariant mass within 15 GeV of the Z mass is used to reconstruct a "loose" Z and the event is tagged as a ZZ→llll event. Similarly, after the removal of the leptons from the primary Z→ll decay, a cut on $M_T^W = \sqrt{2|\boldsymbol{p_T^l}||\boldsymbol{p_T^{miss}}|(1-\cos\Delta\theta)}$ is applied to all remaining lepton objects; here, $\boldsymbol{p_T^l}$ is the transverse moment of the lepton, $\boldsymbol{p_T^{miss}}$is the missing transverse momentum, and $\Delta\theta$ is the difference in azimuthal angle between them. If any of the lepton candidate objects have $M_T^W > 30$ GeV the event is flagged as a ZW→lllν event and discarded.

Id- and Anti-id Lepton Selection

The Z+jets control region is composed of events with one and only one reconstructed Z boson and no reconstructed W bosons. As mentioned above, the fake factor is calculated by finding the ratio of fake id-leptons to fake anti-id leptons in the jet sample. The leptons from the Z decay are removed from the lepton containers before categorizing the remaining leptons into the id- and anti-id containers. Furthermore, hierarchal electron-muon overlap removal is applied such that an id-electron within a 0.1 solid angle of an id-muon is discarded while an anti-id electron with a 0.1 solid angle of either an id- or an anti-id muon is discarded.

The fake factors are calculated separately for electrons and muons as both the lepton identification criteria and the abundance of fake leptons in jets vary significantly with lepton flavour. The fake factor calculations are further separated into lepton P_T bins. This is because the fake factor is significantly dependent on P_T; such dependence is a result of the interplay between the fake factor's sensitivity to the source of fake leptons, the different P_T dependencies of dominant fake lepton sources, and the delicate relationship between the id- and anti-id selection criteria.

In Figures 3 and 4 are the P_T distributions of the fake lepton candidates and of the jets in the Z+jets control region for electrons and muons respectively. The variation amongst these distributions for jets, electrons, muons, id-leptons, and anti-id leptons indicates how sensitive fake factor calculations are to lepton flavour and P_T. Note that the minimum P_T for a jet is 25 GeV compared to 15 GeV for leptons.

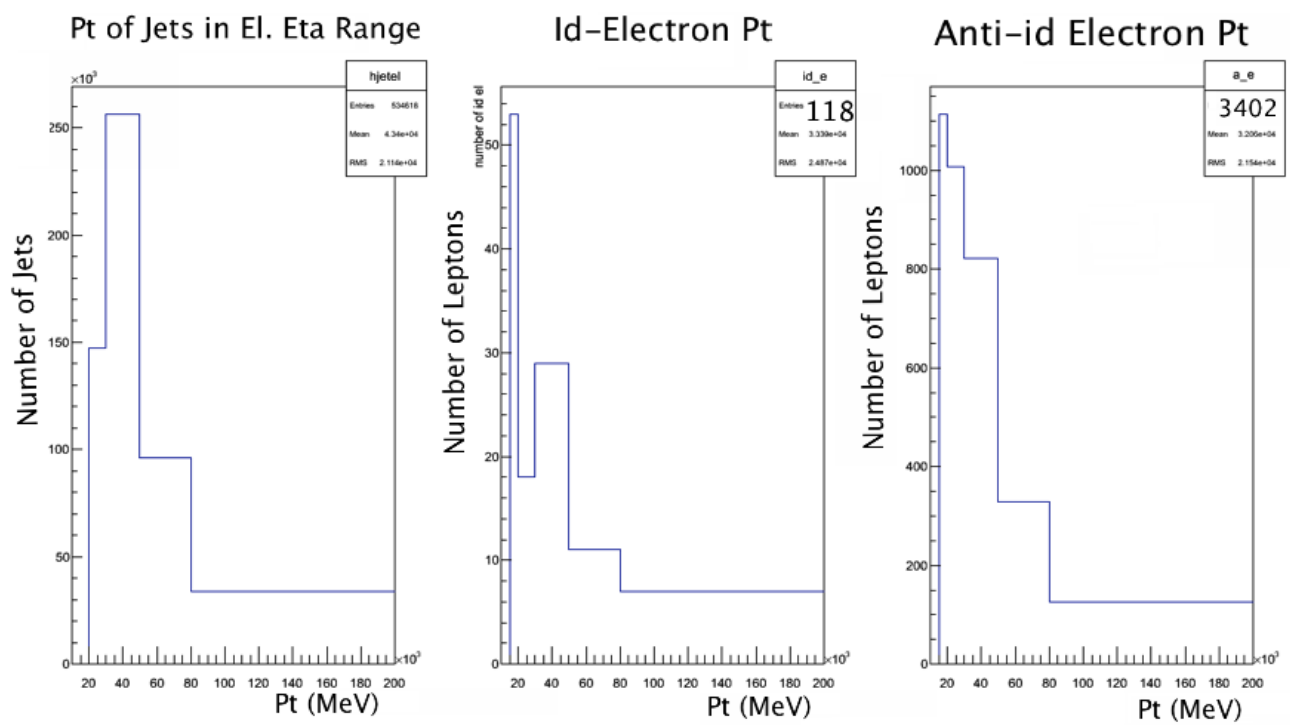

Figure 3: P_T distributions of id-electrons, anti-id electrons and jets in electron eta range. Standard fake factor P_T binning is applied.

The distributions for id-electrons and jets share a peak in the 30—50 GeV bin. This is consistent with the MC simulations demonstrating that hadrons are the primary source of fake electron candidates [9, 13]. Misidentified hadrons or other electron candidates in jets that have a large fraction of the total jet P_T are more likely to pass the isolation requirements. Indeed, studies of fake lepton candidates show a correlation between the ratio of electron candidate P_T to summed event E_T and the fraction of anti-id electrons promoted to id-electrons; the correlation is particularly strong in the region in which jets are most abundant ($P_T < 70$ GeV) [14]. Id-electrons in the very low P_T region ($P_T < 20$ GeV) may be isolated leptons with momentum perpendicular to the jet axis that are characteristic of semileptonic b-hadron decays; they may also be accidentally isolated electron candidates such as charged hadrons or non-prompt electrons from hadron decays in low pt jets or converted photons.

Conversely to id-electrons, anti-id electrons have loosened isolation requirements and thus contain more objects from jets that have electron-like properties but are not leading objects in jets. The jet peak in the 30-50 GeV P_T bin thus produces a peak shifted to lower P_T in the anti-id electron P_T distribution near the minimum P_T threshold of 15 GeV. Electron objects from 40 GeV P_T jets with P_T approaching 40 GeV are likely isolated and promoted to id-electrons while the abundance of 40 GeV P_T jets produces many non-isolated electron candidate objects with a fraction of the jet P_T that consequently fall into the lowest P_T bins. As well, the removal of the impact parameter significance requirement in conjunction with loosened isolation requirements allows for greater acceptance of electron candidates from heavy flavour jets such as semileptonic b- and c- hadron decays that produce electrons inside the jet cone.

Muon candidate objects are almost entirely non-prompt muons produced from semileptonic heavy flavour decays and in-flight decays of light flavour hadrons such as pions and kaons. Since id-muon candidates are non-prompt muons rather than misidentified leading hadrons they are much less likely to carry a large fraction of jet P_T than are electron candidates and thus usually only have high P_T if they originate from a much higher P_T jet. It is therefore not surprising id-muons do not share the same P_T distribution as jets.

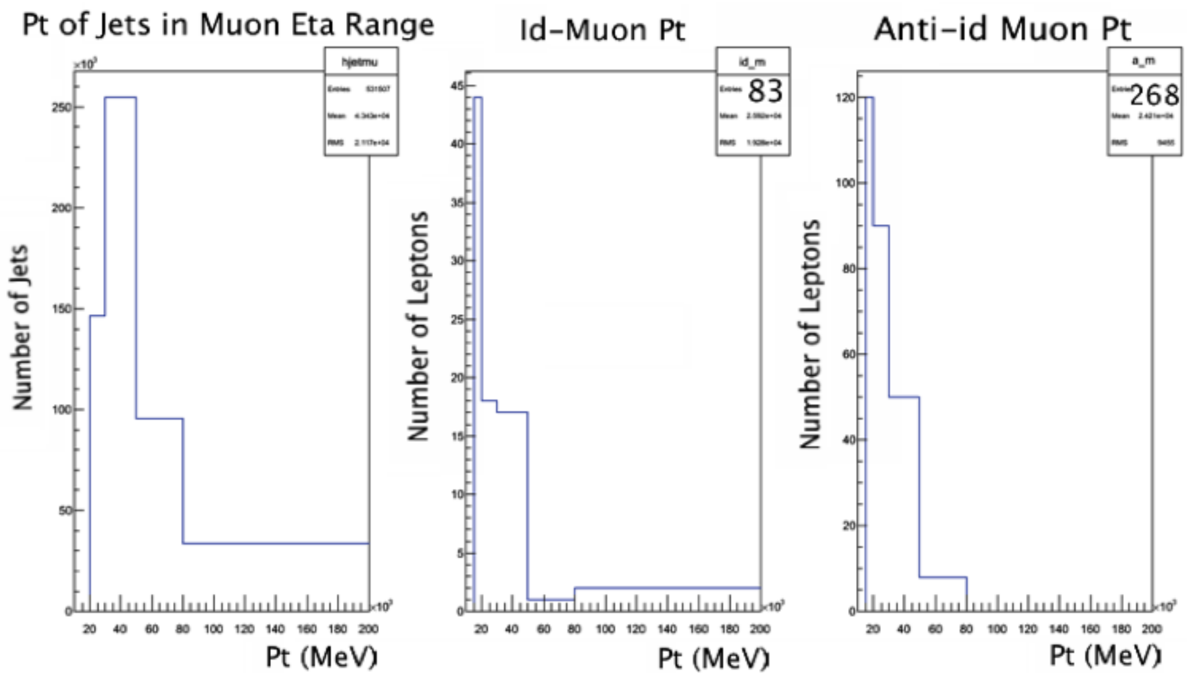

Figure 4: P_T distributions of id-muons, anti-id muons and jets in muon eta range. Standard fake factor P_T binning is applied.

Heavy flavour semileptonic decays are the dominant source of muon candidates above 15 GeV. Due to the small sample of fake muon candidates, the rare event of an isolated fake muon candidate is unlikely to occur in any P_T outside of the low P_T region where the muon candidates are most abundant [11]. In particular, the loosened isolation requirements for anti-id muons allow for non-prompt non-isolated muons from hadronic decays throughout the P_T<80 GeV range carrying a fraction of the jet P_T that extends to 180 GeV. The id-muon candidates lack the misidentified hadron contribution that is the primary source of isolated high P_T fake electrons; this is consistent with the insufficient statistics observed in the high P_T region for calculating muon fake factors. In fact, the major source of the already scarce id-muons—and the only source above 50 GeV—is the contribution of ZZ/ZW leptonic decay events that pass the EW veto.

Electroweak Background Subtraction

The EW event vetoes do not cut all of the electroweak background but rather only 75% of it at the expense of cutting 9% of the Z+jets events, as estimated using MC. The remaining real lepton contamination from EW events primarily contributes to the id-lepton samples, which biases the fake factor calculations. The low statistics in the id-lepton bins result in fake factors that are extremely sensitive to even contamination on the order of 10 events. To correct for this effect, the id- and anti-id lepton selection is applied to MC EW samples and subtracted from the id- and anti-id leptons selected in the data.

Figures 5 and 6 are plots of the P_T distributions of id- and anti-id leptons both before and after the EW background subtraction. Since leptonic Z and W decays produce real isolated leptons, the EW background subtraction has little effect on the anti-id lepton samples. The subtraction does however have a significant effect on the id-lepton samples. This is particularly true in the 20—50 GeV P_T range where the W/Z/γ* contribution to lepton candidates has a peak [9] and where the number of id-leptons are about halved by the subtraction. For muons the electroweak subtraction accounts for more than all of the high P_T (>50 GeV) id-muons. The high P_T bins for id-muons become negative after the electroweak subtraction indicating that no fake muons are produced above 50 GeV in this data sample and a fake factor cannot be calculated for muons in this P_T range.

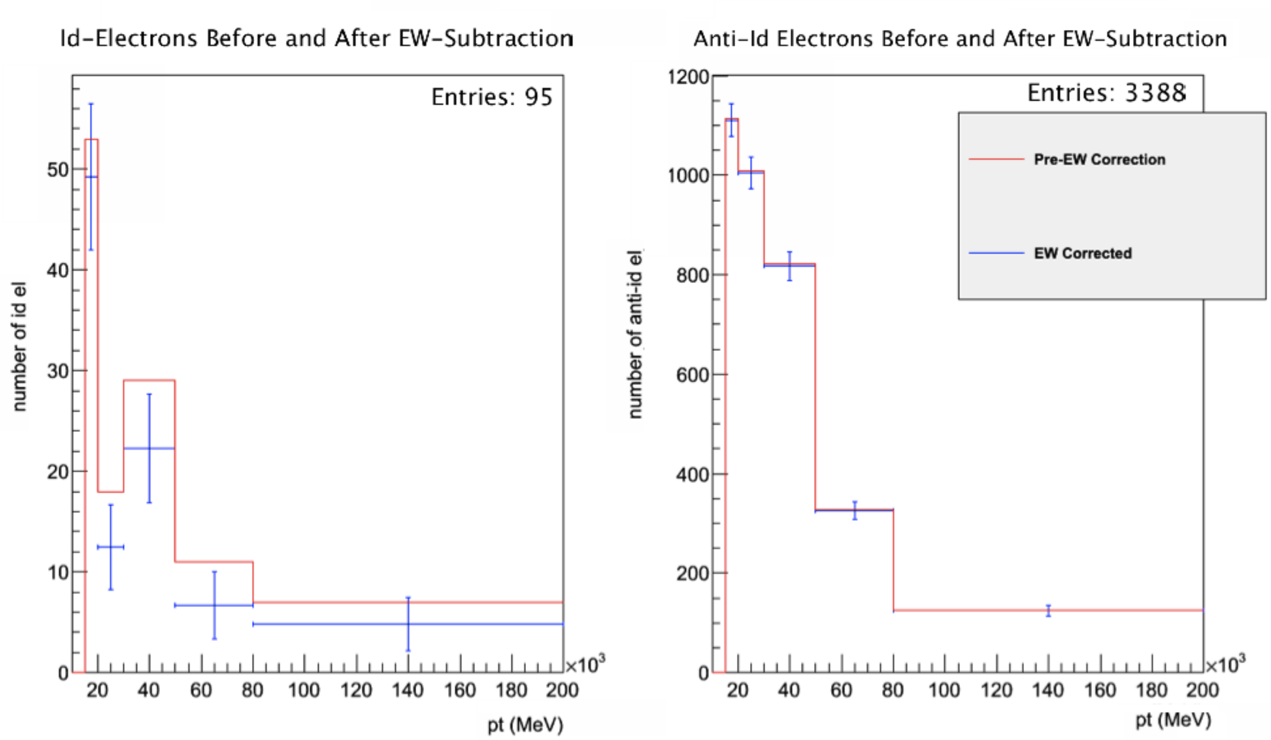

Figure 5: Id- (left) and Anti-id (right) Electrons before and after EW Subtraction

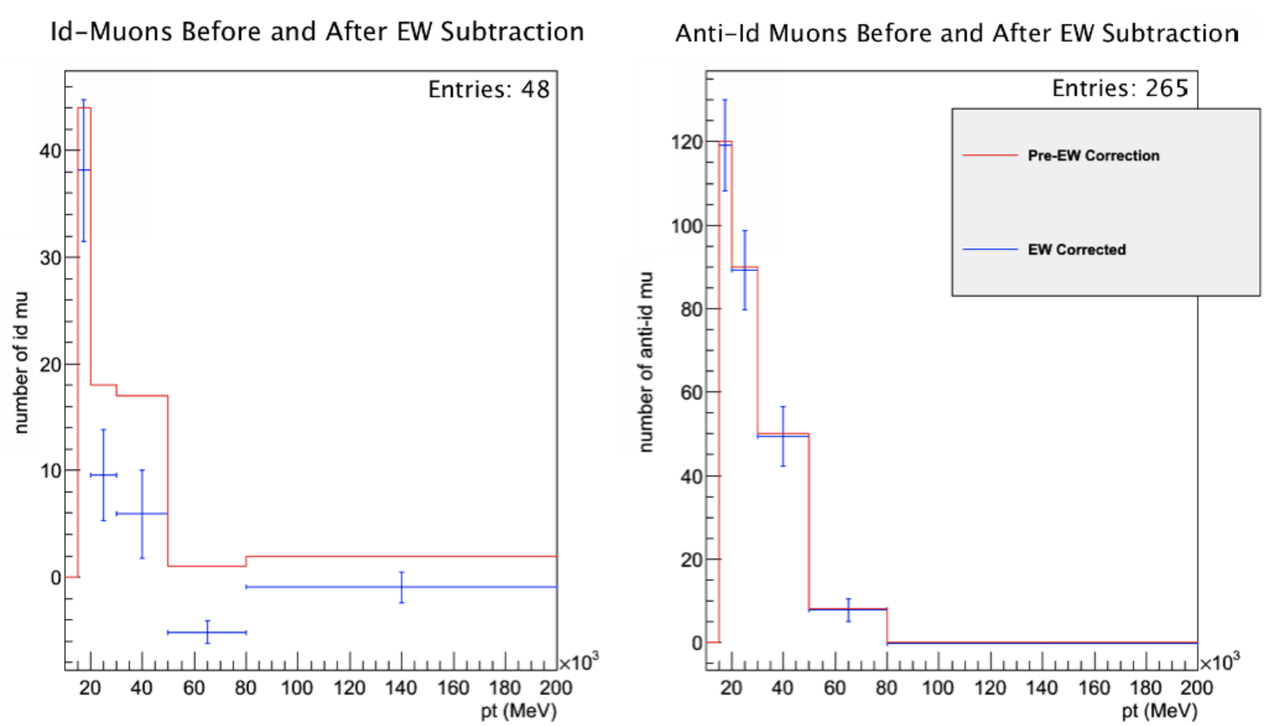

Figure 6: Id-Muons (left) and Anti-id Muons (right) before and after EW Subtraction

ATLAS WORK IN PROGRESS

Fake Factors

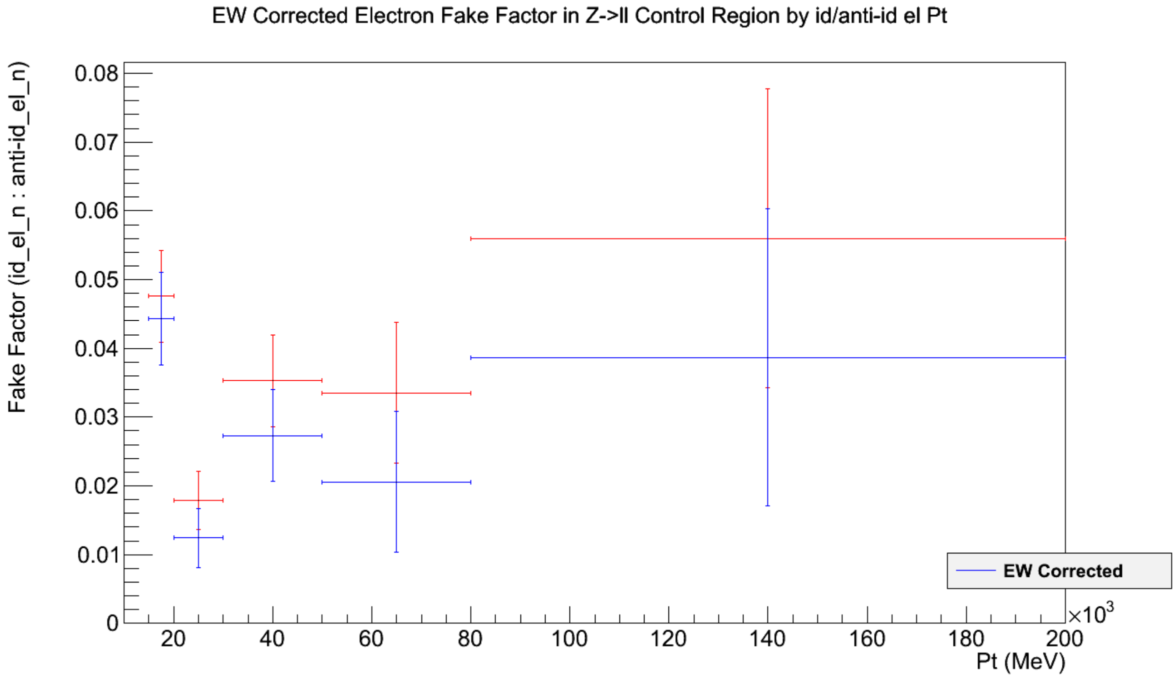

Figure 7: Electron Fake Factors before and after EW Subtraction (Statistical error only)

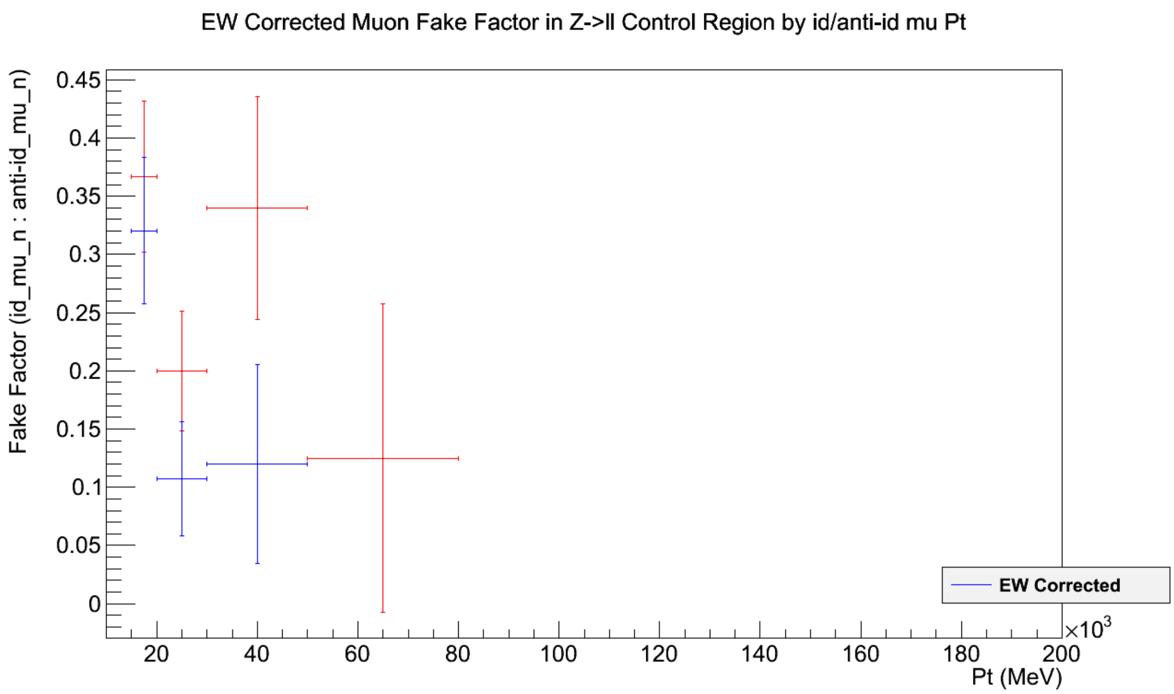

Figure 8: Muon Fake Factors before and after EW Subtraction (Statistical error only)

ATLAS WORK IN PROGRESS

Figures 7 and 8 display the fake factors before and after EW correction for with only statistical errors. Since EW subtraction primarily affects id-lepton samples, the fake factors are consistently reduced by EW corrections. As the number of fake leptons decays with P_T, the statistical errors grow with P_T and are as large as 56% for electrons. In the case of muons, the lack of statistics in the high P_T region is even more drastic. The EW background correction accounts for more than all of the id-muons above a P_T of 50 GeV while the number of anti-id muons vanishes above a P_T of 80 GeV. Evidently, this gives unphysical fake factors for high P_T muons. It is clear that a larger data sample is required to calculate the Z+jets muon fake factor with representative statistics in the high P_T region; however, the results that the id-muon count in the 50—80 GeV bin is not even non-negative within 5σ and that before EW-subtraction there are not even anti-id muons in the 80—200 GeV bin suggest the W+jets background may in fact contribute no high P_T fake muons.

Sources of Systematic Error

Two dominant sources of systematic uncertainty on the nominal QCD multijet fake factor are trigger bias and sample dependence. While the multijet control region uses a pre-scaled trigger, events in the Z+jets region are flagged using the same single lepton trigger that is used in the signal region [9]. Thus by using the Z+jets control region, trigger bias is eliminated and run dependence effects due to trigger type are reduced. Indeed, the fake factors were calculated for this analysis by run period and by period-dependent trigger settings to confirm the lack of potential run dependence in the Z+jets fake factors. The remaining sources of systematic uncertainty evaluated for Z+jets fake factors are EW cross sections and sample dependence.

Electroweak Background Cross Section

In the fake factor calculations for the $H \rightarrow WW(*) \rightarrow lvlv$ analysis the NNLO cross sections of EW background processes are varied conservatively by 20% and the resulting variation in fake factors are set as systematic uncertainties. The nominal fake factor is defined as that corrected using the nominal cross sections for EW background processes and is denoted below by FF_{σ}. The systematic uncertainty is then calculated as:

$$Systematic\ Uncertainty = \left| \frac{\mathrm{FF}(\sigma \pm 20\%) - \ \mathrm{FF}\sigma}{\mathrm{FF}\sigma} \right|$$

where the expression inside the absolute value bars is the *signed relative difference*. Figures 9 and 10 give the EW subtraction signed relative differences for electrons and muons respectively. For electrons the EW systematic uncertainty is less than 15% while for the muons it is significantly larger reaching nearly 40%. In both cases the uncertainty varies with the P_T bin and is larger in the P_T regions where the nominal EW subtraction accounts for a significant proportion of the id-leptons. For muons, the 80—200 GeV P_T bin is omitted since there are zero anti-id muons in this bin even before EW subtraction. Conversely, in the 50—80 GeV P_T bin there is a positive number of anti-id muons after nominal EW corrections and potential for a positive fake factor if the id-muons were not all accounted for by EW background. The id-muon count is negative in this bin even after reducing the EW cross section by 20%, confirming that there are no high P_T fake muons in this dataset and that a larger sample is necessary to understand the sources of leptons in this region.

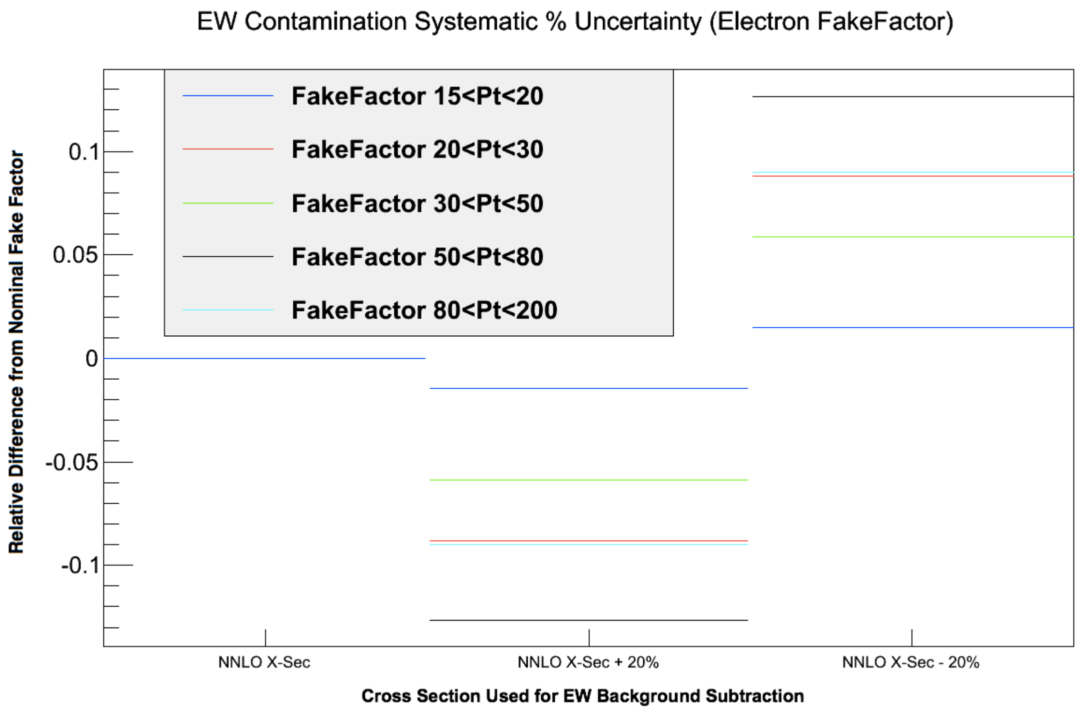

Figure 9: EW Contamination Signed Relative Difference (Electron)

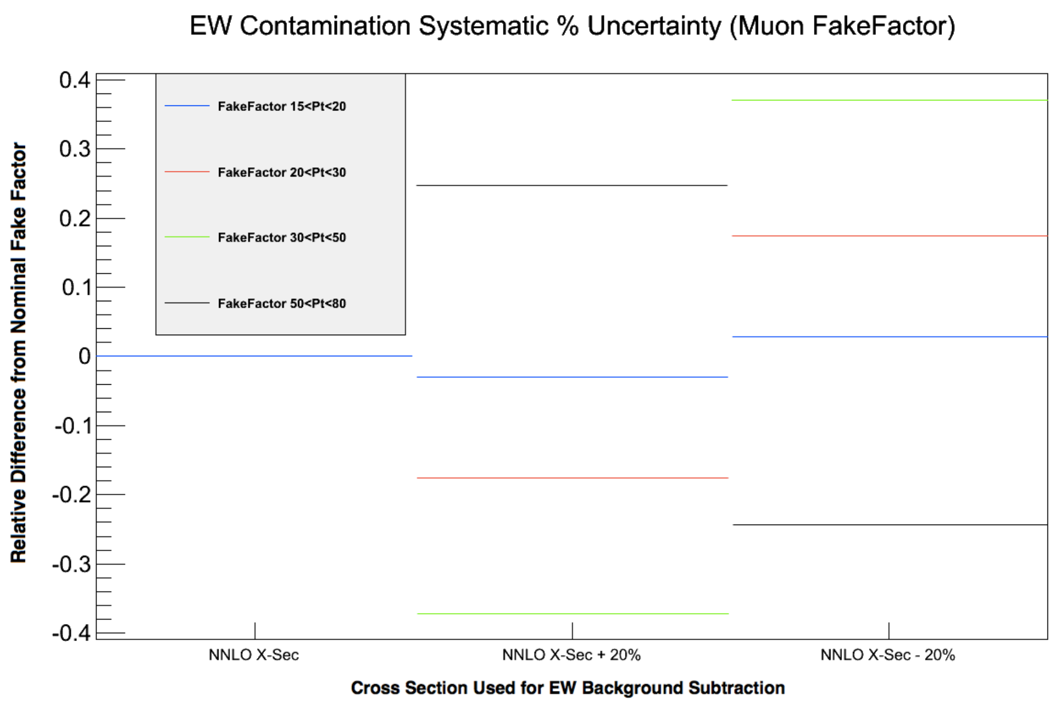

Figure 10: EW Contamination Signed Relative Difference (Muon)

ATLAS WORK IN PROGRESS

Sample Dependence

Although the jet composition in Z+jets and W+jets events are very similar in that they have similar proportions of quark and gluon initiated jets, there is still variability in the flavour compositions of quark initiated jets between the two samples. Furthermore, the dominant production mechanisms in events with higher jet multiplicity differ for Z+jets and W+jets events; for example gg→Z+$b\bar{b}/c\bar{c}$ only contributes at LO to the Z+jets sample.

In order to evaluate the systematic error associated with estimating the fake rate outside of the W+jets sample itself, the relative difference between fake factors calculated from Z+jets and W+jets MC samples is assigned as a systematic uncertainty on the data-driven fake factor. The systematic uncertainty is the magnitude of the sample dependence where sample dependence is calculated as:

$$Systematic\ Uncertainty = \left|\frac{FF^{MC}_{Zjets} - FF^{MC}_{Wjets}}{FF^{MC}_{Zjets}}\right|$$

where $FF^{MC}_{Z(W)jets}$ is the fake factor calculated using Z(W)+jets MC samples**.**

The sample dependences for electrons and for muons in figure 11 are preliminary results as the fake factor framework used for this analysis does not include a complete self-contained sample dependence calculation. The current framework calculates Z+jets fake factors from MC samples but is not yet equipped to calculate W+jets fake factors for comparison; thus the results shown here compare the Z+jets MC results to the W+jets MC results from the parallel Z+jets study (PS) on which this analysis was modeled. The data driven fake factors presented in this paper agree with the PS results within error and the central event and lepton selection methods were made to match those of the PS analysis. The separate ntuple production codes are nonetheless multifarious and have extensive analytical stages to which the central values of the fake factor are sensitive. Additionally, the W+jets MC results were calculated using truth information while the Z+jets MC fake factors results were obtained by applying the kinematic selection as is applied to the data. In both cases, the EW event veto was not applied and only Zee(mm)+(0-5) parton and Zee(mm)+bb+(0-3 parton) MC samples were considered. Lastly, the W+jets fake factors were calculated with the heavy flavour overlap removal between inclusive W+jets samples and W+bb(cc)+jets samples whereas the heavy flavour overlap removal is still in the process of being implemented into the Z+jets analysis. Therefore, there is some additional overlap in phase space for heavy flavour decays in the Z+jets MC fake factors used here.

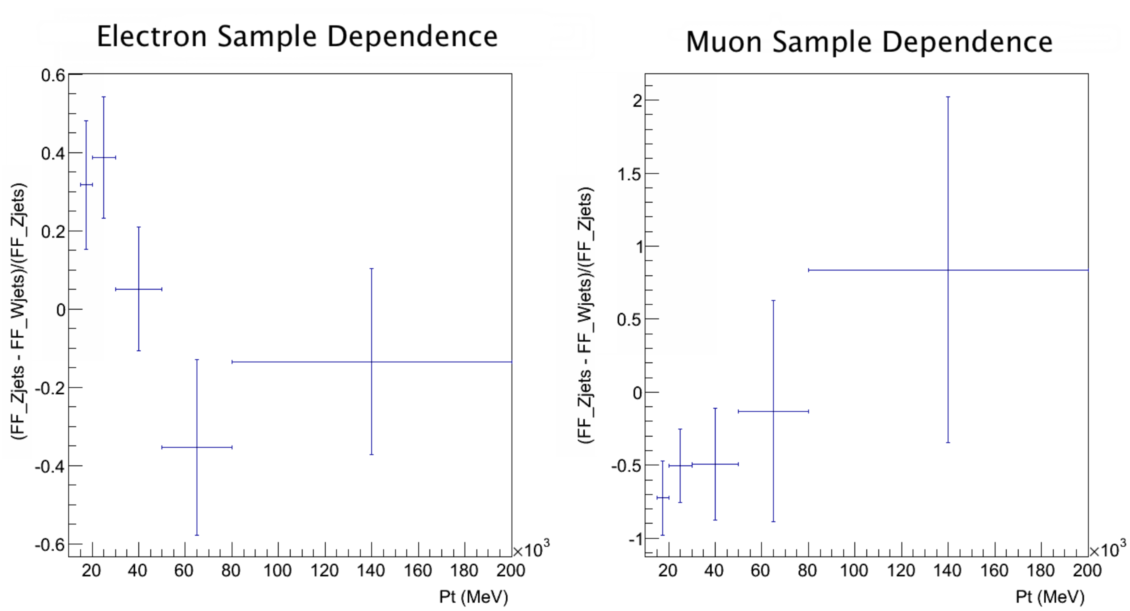

Figure 11: Preliminary Sample Dependence Results for Electrons (Left) and Muons (Right)

Final Fake Factors with statistical + systematic Error

Figures 12 and 13 present the EW-corrected fake factors with all statistical and systematic errors summed in quadrature. The statistical errors, drawn in blue, are the dominant uncertainty on Z+jets fake factors. The sample dependence is the second largest contribution to the error and in the low P_T region is comparable to the contribution of the statistical error. Conversely, the contribution from the systematic errors assigned for the EW subtraction, drawn in red, are typically an order of magnitude smaller than the statistical errors. The muon fake factors in the P_T > 50 GeV region are set to zero in place of unphysical fake factors with negative values of EW-corrected id- and anti-id muons.

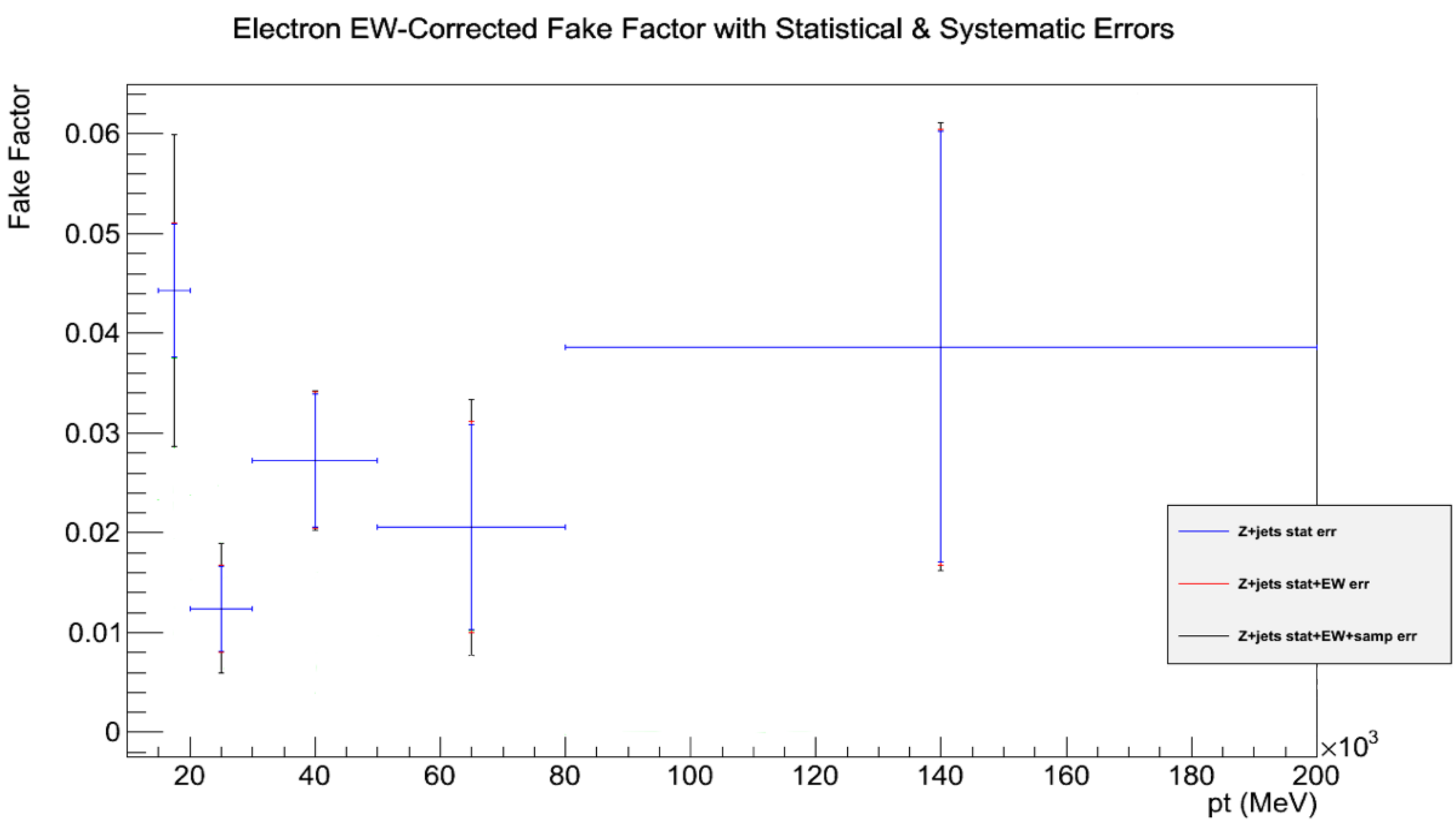

Figure 12: Electron Fake Factor with statistical (blue) and systematic uncertainties

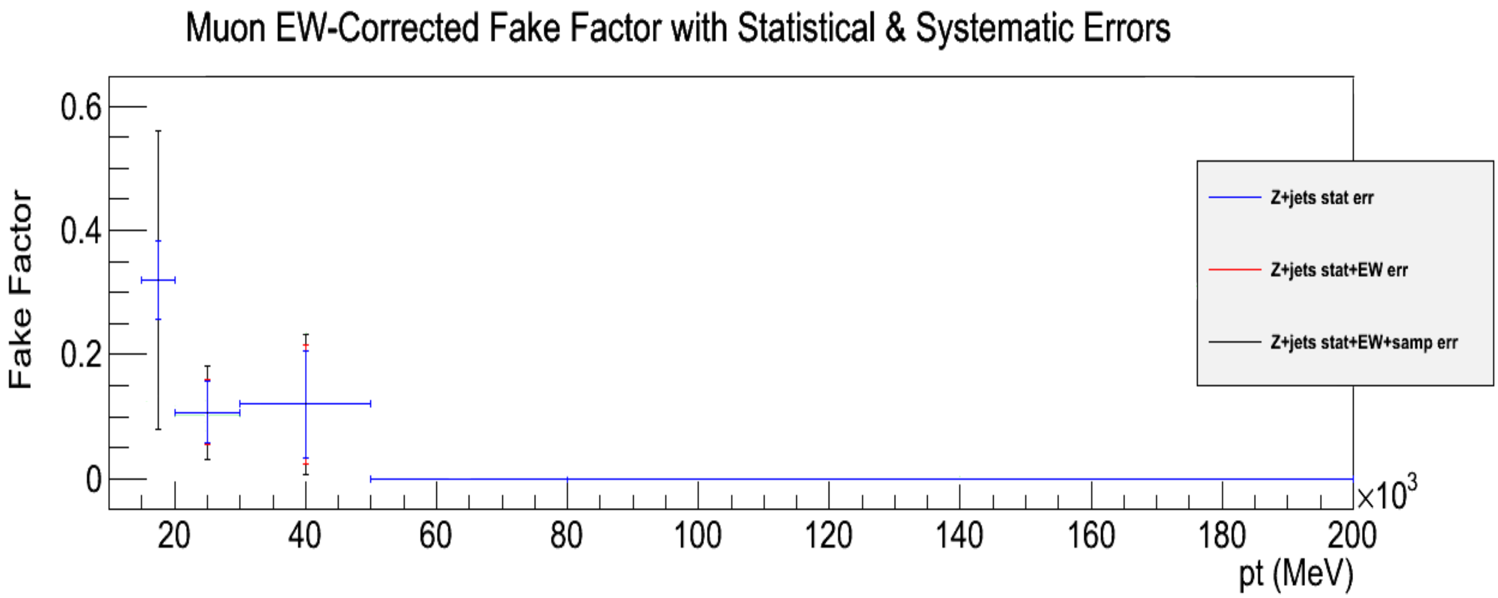

Figure 13: Muon Fake Factor with statistical (blue) and systematic uncertainties

ATLAS WORK IN PROGRESS

Conclusion and Next Steps

The motivation for studying Z+jets fake factors is the potential for lowering uncertainties on the fake factor and accordingly on the W+jets background estimate. The Z+jets fake factors obtained from the 2011 dataset are dominated by statistical errors; the systematic errors are below 41% for electrons and 75% for muons. The large statistical uncertainty limits the benefit of the Z+jets fake factor in current datasets as comparison with multijet fake factors show comparable relative uncertainties in most P_T bins; however, the bounds on the systematic errors indicate that in a dataset of about 14 fb^{-1} the Z+jets fake factors will have consistently lower relative uncertainties compared to the nominal fake factors. Higher statistics will also give a more representative estimate of the high P_T muon contribution and may confirm that the W+jets background contributes no high P_T fake id-muons. Finally, the development of the Z+jets framework to include W+jets fake factor calculation will complete the sample dependence calculation and allow for the optimization of the Z+jets fake factor anti-id lepton criteria that is already prepared in the Z+jets framework. In addition to the optimization of the anti-id lepton criteria, applying variations of this analysis to the data has indicated the optimization of the Z+jets control region event selection would be beneficial to both the reduction of statistical and systematic errors.

References

[1] G. Aad *et al.* [ATLAS Collaboration], "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," Phys. Lett. B **716**, 1 (2012) [hep-ex/1207.7214].

[2] G. Aad *et al.* [ATLAS Collaboration], "Search for the Standard Model Higgs boson in the H $\rightarrow$ WW(*) $\rightarrow$ l nu l ny decay mode with 4.7 fb^{-1} of ATLAS data at sqrt(s) = 7 TeV," Phys. Lett. B **716**, 62 (2012) [hep-ex/1206.0756].

[3] The ATLAS Collaboration, ATLAS-CONF-2011-063.

[4] The ATLAS Collaboration, ATLAS-CONF-2010-036.

[5] ATLAS Collaboration, "Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data," Eur. Phys. J. C **72**, 1909 (2012) [hep-ex/1110.3174].

[6] W. Lampl et al., Calorimeter Clustering Algorithms: Description and Performance, ATL-LARG-PUB-2008-002, CERN, Geneva, Apr 2008.

[7] The ATLAS Collaboration, Expected Performance of the ATLAS experiment: detector, trigger and physics, CERN-OPEN-2008-020, CERN, Geneva, 2009.

[8] G. Aad *et al.* [ATLAS Collaboration], "Limits on the production of the Standard Model Higgs Boson in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector," Eur. Phys. J. C **71**, 1728 (2011) [hep-ex/1106.2748].

[9] John Alison, *The road to discovery: TRT Alignment, Electron Identification, The Fake Factor Method, WW-Physics and the discovery of the Higgs*, University of Pennsylvania, Pennsylvania, 2012, http://higgs.hep.upenn.edu/~johnda/Papers/alison_thesis_fakes.pdf.

[10] G. Aad *et al*. [ATLAS Collaboration], "Search for same-sign top-quark production and fourth-generation down-type quarks in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector," JHEP **1204**, 069 (2012) [hep-ex/1202.5520].

[11] The ATLAS Collaboration, ATLAS-CONF-2010-075.

[12] The ATLAS Collaboration, ATLAS-CONF-2010-092.

[13] Scott Snyder [ATLAS Collaboration], "Electron and Photon Performance and Electron p_T Spectrum Measurement with ATLAS in pp Collisions at sqrt(s) = 7 TeV," PoS **ICHEP2010**, 024 (2010) [hep-ex/1012.0554].

[14] The ATLAS Collaboration, "Inclusive search for same-sign dilepton signatures in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector," JHEP **1110**, 107 (2011) [hep-ex/1108.0366].