Invariant mass distribution of jet pairs produced in association with one lepton and missing transverse energy at the ATLAS experiment

The ATLAS Collaboration

Abstract

This note presents a study of the invariant mass distribution of jet pairs produced in association with one identified lepton and missing transverse energy with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. Three recent ATLAS analyses have studied this final state in the electron and muon decay channels, using an integrated luminosity of about 33 pb$^{-1}$ collected in 2010. One is a measurement of W production in association with jets, and the others are searches for Higgs bosons and for leptoquark pair production. In the analysed data sample, no significant excess is observed in the dijet mass spectrum above an invariant mass of 120 GeV.
1 Introduction

The CDF Collaboration recently reported the results of a study of the invariant mass distribution of jet pairs produced in association with a W boson, with an excess in the 120–160 GeV mass range [1]. Three recent ATLAS analyses also examine event topologies with two jets produced in association with a high transverse momentum lepton (electron or muon) and large missing transverse energy. The measurement of the production cross section for W bosons in association with jets [2, 3] considers jet multiplicities up to five, and the subset with exactly two jets is used here to study the dijet invariant mass distribution. The search for Higgs boson production [4] using the decay $H \rightarrow WW \rightarrow \ell\nuqq$ looks for leptonic W decays in association with a jet pair. For the Higgs search, the dijet invariant mass, $m_{jj}$, is required to be consistent with the hadronic decay of a W boson, and the reconstructed invariant mass of the WW system gives an estimate of the mass of the Higgs boson candidate. Events where one or more of the jets are tagged as $b$-jets are excluded to reduce the contribution from $t\bar{t}$ decays. Final states with two jets, a charged lepton and $E_{T}^{miss}$ could also be a signature for pair-production of first or second generation leptoquarks [5]. This would not be expected to give a resonance in the dijet mass spectrum, but instead would be characterised by an excess of $\ell^{+}\ell^{-} jj$ or $\ell\nu jj$ events where the masses of the correct lepton-jet or neutrino-jet combinations correspond to the leptoquark mass.

All three ATLAS analyses examine the $m_{jj}$ distribution as one of the many checks that the data are well described by a combination of Monte Carlo simulation and data-driven background estimates. In view of the CDF results, these distributions are reproduced in this note to make them publicly available. Furthermore, the W+jets analysis has been repeated with selection criteria as close as possible to the CDF study and the results will be presented in the penultimate section (Section 7) below.

Inclusive $W$ production in association with jets (W+jets) is the dominant source of events with a lepton, missing transverse energy and two jets. Other contributions come from QCD multijet events with misidentified lepton candidates or leptons coming from heavy quark decays, from electroweak processes including $Z \rightarrow \ell\ell + \text{jets}$ where one lepton is not identified, $W \rightarrow \tau\nu + \text{jets}$, and $Z \rightarrow \tau\tau + \text{jets}$, from $t\bar{t}$ and single top decays with one real lepton, and from diboson ($WW$, $WZ$, $ZZ$) processes.

The CDF study pays particular attention to the $WW$ and $WZ$ diboson component of the dijet mass spectrum. Their first significant measurement of this small contribution was made with a sample of 3.9 fb$^{-1}$ [6]. With the present much smaller integrated luminosity, ATLAS cannot observe these diboson processes in the dijet mass spectrum. It is worth noting the relative scaling of the diboson and W+jets processes from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV to $pp$ collisions at $\sqrt{s} = 7$ TeV. The production cross section for $W + (n \geq 2) \text{jets}$ grows by approximately a factor of 20, from $\approx 22$ pb measured by CDF [7] to $\approx 440$ pb measured by ATLAS [2]$^1$, while the cross section for $WW/WZ \rightarrow \ell\nuqq$ production, calculated with MCFM [8], only grows by approximately a factor of 4, from 3.7 pb at the Tevatron to 15.3 pb at the LHC. However, without making model dependent assumptions on the possible production mechanism for the reported excess in the range of dijet masses from 120 to 160 GeV range, it is impossible to predict its importance relative to diboson production rate at the LHC.

2 The ATLAS Detector

The ATLAS detector [9] consists of an inner tracking system (inner detector, or ID) surrounded by a thin superconducting solenoid providing a 2T magnetic field, electromagnetic and hadronic calorimeters and a muon spectrometer (MS) embedded in three large superconducting toroid magnets with eight coils each. The nominal $pp$ interaction point at the centre of the detector is defined as the origin of a right-handed coordinate system. The $x$-axis points to the centre of the LHC ring, with the $y$-axis pointing

$^1$These cross sections are measured in slightly different kinematic regions. Although both require jets with transverse momentum larger than 20 GeV, the jet pseudorapidity is limited to $|\eta| < 2.0$ for CDF and jet rapidity $|\eta| < 2.8$ for ATLAS.
upwards. The beam direction defines the z-axis. The azimuthal angle $\phi$ is measured around the beam axis and the polar angle $\theta$ is the angle from the z-axis. The pseudorapidity is defined as $\eta = - \ln \tan(\theta/2)$ and rapidity is defined as $y = 0.5 \times \ln[(E + p_T)/(E - p_T)]$.

3 Data and event selection

The event samples for the analyses considered here are selected by unprescaled lepton triggers, requiring either an electron with transverse energy, $E_T$, greater than 15 GeV, or a muon with transverse momentum, $p_T$, greater than 13 GeV. Only data taken during periods with stable proton-proton collisions, and where all relevant parts of the detector are operating normally are considered. The events are required to have a reconstructed primary vertex with at least three associated tracks.

The initial estimate of the corresponding integrated luminosity was 35 pb$^{-1}$, with an uncertainty of $\pm 11%$ [10], and this is used for the Higgs and leptoquark searches. The $W$+jets cross section measurement takes advantage of an improved luminosity analysis, which reduces the uncertainty to $\pm 3.4%$ and revises the estimated integrated luminosity to 33 pb$^{-1}$ [11] for the same sample.

Full details of the simulated event samples, including choices of parton density functions and of renormalisation and factorisation scales are given in Refs. [3–5]. In brief, inclusive W and Z production is simulated with PYTHIA 6.4.21 [12], while W or Z plus jets samples use ALPGEN 2.13 [13], SHERPA 1.1.3 [14] or MC@NLO 3.3.1 [15]. The shapes of the $W$+jets contributions to the $m_H$ distributions in this note are taken from ALPGEN. The MC@NLO generator is also used for $t\bar{t}$ and single top production, with additional $t\bar{t}$ samples from POWHEG-HVQ [16]. Diboson ($WW$, $WZ$, $ZZ$) production is simulated by HERWIG 6.510 [17]. PYTHIA 6.4.21 is used to study the background due to fake leptons or leptons from heavy flavour decay in dijet events, although this background is finally assessed directly from data. Samples were generated with minimum bias interactions overlaid on top of the hard-scattering event in order to account for the multiple $pp$ interactions in the same beam crossing (pile-up) experienced in the data. The detector simulation [18] is performed using GEANT4 [19].

Electron candidates are defined as an electromagnetic calorimeter cluster with an associated ID track. They are required to satisfy the ATLAS “medium” or “tight” identification quality requirements [2]. Their transverse energy must satisfy $E_T > 20$ GeV. Candidates are accepted with $|\eta| < 2.47$, excluding the crack region between barrel and endcap, $1.37 < |\eta| < 1.52$. The electron isolation is determined from the sum of calorimeter-cluster transverse energies, $E_T^{\text{cone}}$, in a cone of radius $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$ around the electron direction, not including energy from the electron itself.

Muon candidates are identified in both the ID and MS, and must satisfy $p_T > 20$ GeV and $|\eta| < 2.4$. They must also be consistent with originating at the primary vertex to reject cosmic ray background [20]. The muon isolation is defined either by $E_T^{\text{cone}}$, or the sum of track transverse momenta in a cone of radius $\Delta R$ around the muon, $\Sigma p_T^{\text{ID}}$. The exact lepton isolation criteria applied by each analysis are detailed below.

Jets are reconstructed from calorimeter clusters using the anti-$k_t$ algorithm [21] with radius parameter 0.4. To take into account the differences in calorimeter response to electrons and hadrons and to correct for experimental effects, a $p_T$ and $\eta$-dependent factor, derived from simulated events, was applied to each jet to provide an average energy scale correction [22]. Jets arising from detector noise or cosmic rays are rejected. Calibrated jets are required to have $p_T > 20$ GeV, and there are analysis-dependent requirements to ensure that the jets are well separated from lepton candidates. For the purposes of this note there must be at least two jets.

The missing transverse energy in the event, $E_T^{\text{miss}}$, is measured from all topological clusters in the calorimeter with $|\eta| < 4.5$, together with the muon transverse momentum for events in the muon channel. A local calibration is applied to each cluster [22]. The transverse mass of the lepton-neutrino system, $m_T$, is measured assuming that the neutrino transverse momentum is equal in magnitude and direction to the $E_T^{\text{miss}}$ vector. The exact requirements on $E_T^{\text{miss}}$ and $m_T$ are analysis dependent, but all require at least
$E_T^{\text{miss}} > 25$ GeV, and the $W$+jets and leptoquark selections require $m_T > 40$ GeV.

The various analyses differ in some details of the additional selection criteria, as described in the following sections. In comparing the data to Standard Model expectations, apart from the dominant $W$+jets process, other electroweak processes are normalised to NNLO calculations [23]. The cross section for $t\bar{t}$ production is taken from the NLO+NNLL prediction [24], and single top and diboson production are scaled to the NLO expectation [8]. Backgrounds from QCD multijet events are assessed by data-driven methods. For the purposes of examining the dijet invariant mass distribution in this note, the $W$+jets contribution is then normalised so that the sum of all sources matches the data in the region $m_{jj} < 120$ GeV. The resulting shift in normalisation with respect to the original ALPGEN prediction is at the few percent level only.

4 Dijet mass in the $W$ plus jets selection

Events are required to have exactly one isolated lepton, either a tight electron, or a muon, with $p_T > 20$ GeV. They are vetoed if there is a second lepton of the same flavour, including medium quality electrons. The electron isolation requirement is $E_{\text{cone}}^T < 4$ GeV (with the cone defined by $\Delta R = 0.2$), while the muon must satisfy $\Sigma p_T^\text{ID}/p_T^\mu < 0.1$ with a cone size $\Delta R = 0.2$.

Jets must have $p_T > 20$ GeV, $|y| < 2.8$ and must be separated from leptons by $\Delta R > 0.5$. To avoid spurious jets from pile-up events in the same bunch crossing, jets are removed if more than 25% of their associated tracks do not originate from the same primary vertex as the leptons. Only events with exactly two jets surviving this selection are used here.

The dijet invariant mass distribution for the electron and muon channels combined is shown in Figure 1. The expected shapes of the dijet invariant mass distributions for $W$+jets, $t\bar{t}$, single top, $Z$, and diboson decays are taken from the Monte Carlo simulation. The QCD multijet background includes
events with a semileptonic heavy flavour decay. In the electron channel, there is also a second component from events where a light flavour jet is misidentified as an electron. These backgrounds are estimated by Monte Carlo template fits to the $E_T^{\text{miss}}$ distribution for both channels. The normalisations of all other sources apart from the $W+$jets are calculated from the predicted cross sections, assuming an integrated luminosity of 33 pb$^{-1}$ [11]. The $W+$jets contribution is scaled so that the sum of all contributions matches the number of events in the data in the region $m_{jj} < 120$ GeV. The experimental uncertainties on the Monte Carlo expectation are dominated by the jet energy scale and resolution uncertainties, and pile-up removal. No significant excess in the 120 to 160 GeV region is seen.

The same data sample is also used to imitate the CDF analysis as closely as possible by applying additional cuts, as discussed in Section 7 and shown in Figure 2. In Figures 1 and 2, the systematic uncertainty band includes experimental uncertainties from the jet energy scale and resolution, lepton identification efficiency and energy or momentum resolution, modelling of pile-up and evaluation of the integrated luminosity. Theoretical uncertainties in the production cross sections, and shape differences from comparing ALPGEN [13] and SHERPA [14], have also been studied, and give much smaller variations than the experimental uncertainties.

5 Dijet mass in the Higgs analysis

The search for $H \rightarrow WW \rightarrow \ell\nu qq$ applies tighter kinematic cuts, requiring exactly one electron or muon with $p_T > 30$ GeV, and $E_T^{\text{miss}} > 30$ GeV, but makes no explicit additional cut on $m_T$.

The tight electron identification selection is applied, with $E_{T}^{\text{cone}} < 6$ GeV (for a cone size of $\Delta R = 0.3$). The muon isolation requirement is $\Sigma p_T^{\text{ID}} < 1.8$ GeV, with a cone size $\Delta R = 0.2$. Events are vetoed if there are additional leptons with $p_T > 20$ GeV, including medium quality electrons. Electrons within $\Delta R < 0.1$ of a muon candidate are ignored.

For the purpose of the study presented here, only events with exactly two jets with $p_T > 30$ GeV and $|\eta| < 2.8$ are considered. Candidate jets must also be separated from electrons by $\Delta R > 0.3$, and events with one or more $b$-tagged jets are rejected to reduce the background from $t\bar{t}$ [4].

The dijet mass spectrum for the electron and muon channels combined is shown in Figure 3. No Monte Carlo systematic uncertainty has been evaluated for this distribution, because in the Higgs search, the $WW$ invariant mass was the quantity of interest, and the background to the hypothetical Higgs signal was evaluated by fitting the distribution of that variable. The size and shape of the QCD multijet background is evaluated by a data-driven method. The normalisation of all other processes apart from $W+$jets is fixed to their expectation, assuming an integrated luminosity of 35 pb$^{-1}$ with an uncertainty of ±11% [10], and the normalisation of the $W+$jets Monte Carlo sample is scaled so that the sum of all contributions matches the number of events in the data in the region $m_{jj} < 120$ GeV. No significant excess is observed in the region of 120 to 160 GeV.

6 Dijet mass in the leptoquark search

The electron selection in the leptoquark analysis is slightly looser than in the other two analyses, requiring a medium quality electron, with a hit in the first layer of the pixel detector (the innermost layer of the inner detector) if an active module is traversed to reduce the contribution from converted photons. The electron isolation is defined relative to the electron $E_T$: $E_{T}^{\text{cone}}/E_T < 0.2$ (with a cone of $\Delta R = 0.2$). The muon isolation is also defined by comparing calorimeter activity to the muon $p_T$: $E_{T}^{\text{cone}}/p_T < 0.25$ (with a cone of $\Delta R = 0.2$). The analysis requires at least two jets, with $p_T > 20$ GeV and $|\eta| < 2.8$, which must be separated from leptons by $\Delta R > 0.5$. The two highest-$p_T$ jets are used to form the dijet invariant mass.
Figure 3: Dijet invariant mass distribution for the $H \rightarrow WW \rightarrow \ell\nu qq$ search. The errors are statistical only.

Figure 4: Dijet invariant mass distribution for the leptoquark selection. The errors on the data points are statistical, and the systematic uncertainties are described in the text.

The invariant mass distribution for the electron and muon channels combined is shown in Figure 4. The normalisation and shape of the QCD multijet background contribution is derived by data-driven methods. The expected shapes of the dijet invariant mass distributions of the other contributions are taken from the Monte Carlo simulation, and their normalisations are calculated from the predicted cross sections assuming an integrated luminosity of 35 pb$^{-1}$ with a $\pm 11\%$ uncertainty [10], except for the $W+$jets contribution, which is normalised to the data in the mass region below 120 GeV. The systematic uncertainty includes the luminosity uncertainty, and a $\pm 25\%$ uncertainty on the QCD multijet background. Additional uncertainties such as those arising from the jet energy scale or resolution, or lepton identification uncertainties are not included. No significant excess is visible in the 120 to 160 GeV region.

7 CDF style event selection

The CDF selection [1] requires that the lepton is isolated, and has $p_T > 20$ GeV. Events with an additional lepton are rejected. The event should have missing transverse energy $E_T^{miss} > 25$ GeV, transverse mass $m_T > 30$ GeV, and exactly two jets with $p_T > 30$ GeV. The transverse momentum of the dijet system is further required to satisfy $p_T > 40$ GeV. The jets must be separated by $|\Delta \eta| < 2.5$, and a cut on $|\Delta \phi| > 0.4$ between the most energetic jet and the direction of the missing transverse energy is applied to reduce QCD multijet background.

In order to imitate the CDF analysis, the ATLAS $W+$jets selection is modified with the additional cuts $p_T > 40$ GeV, an increased jet $p_T$ threshold of 30 GeV, and the angular cuts on $\Delta \eta$ and $\Delta \phi$ specified above. The cuts on $E_T^{miss}$ and $m_T$ are unchanged. The resulting dijet invariant mass distribution is shown in Figure 2. The normalisation of the $W+$jets contribution has been scaled so that the sum of all contributions matches the number of events in the data in the invariant mass region below 120 GeV, while the normalisation of the other MC processes is fixed. No significant excess is observed. Considering only
the statistical uncertainties in the data sample, the $\chi^2$/dof = 5.7/4 in the mass region between 120 and 160 GeV.

8 Summary

The invariant mass distribution of jet pairs produced in association with one identified lepton and missing transverse energy in three recent ATLAS analyses shows no significant disagreement between data and expectation. These analyses measure $W$+jets production, and search for Higgs boson or leptoquark pair production, using an integrated luminosity of about 33 pb$^{-1}$ collected in 2010. The $W$+jets selection criteria are also adjusted to match the CDF analysis as closely as possible, and no discrepancy between data and simulation is seen in this data sample in the mass region between 120 and 160 GeV.

References


