ATLAS NOTE

ATLAS-CONF-2011-023

March 14, 2011

Top Quark Pair Production Cross-section Measurement in ATLAS in the Single Lepton+Jets Channel Without b-tagging

The ATLAS collaboration

Abstract

A measurement of the production cross-section for top quark pairs ($t\bar{t}$) in pp collisions at $\sqrt{s} = 7$ TeV is presented using data recorded with the ATLAS detector at the Large Hadron Collider. Events are selected in the lepton+jets topology by requiring a single lepton (electron e or muon μ), large missing transverse energy and at least three jets. No explicit identification of secondary vertices inside jets (b-tagging) is performed. In a data sample of 35 pb^{-1}, 2009 μ+jets and 1181 e+jets candidate events are observed. A multivariate method using three kinematic variables is employed to extract a cross-section measurement of

$$\sigma_{t\bar{t}} = 171 \pm 17(\text{stat.})^{+20}_{-17}(\text{syst.}) \pm 6(\text{lumi.})\ \text{pb}.$$

The measurement agrees with approximate NNLO perturbative QCD calculations. Cross-check measurements are performed with one-dimensional likelihood fits and "cut-and-count" methods which are found to be consistent with the main result.

1 Introduction

Top quark measurements are of central importance to the LHC physics programme. The production of top quark pairs in pp collisions at the LHC is the process which lies at the threshold between the Standard Model of particle physics and what may lie beyond it. Uncertainties on the theoretical predictions for the top quark pair production cross-section are now less than 10%, and comparisons with experimental measurements performed in different channels allow a precision test of the predictions of perturbative QCD. Top-quark pair production is an important background in many searches for physics beyond the Standard Model, and new physics may also give rise to additional $t\bar{t}$ production mechanisms or modifications of the top quark decay channels.

In the Standard Model (SM) the $t\bar{t}$ production cross-section in pp collisions is calculated to be $165\ ^{+11}_{-16}$ pb [1] at a centre of mass energy $\sqrt{s} = 7$ TeV assuming a top mass of 172.5 GeV. Top quarks are predicted to decay to a W boson and a b-quark ($t \to Wb$) nearly 100% of the time. Events with a $t\bar{t}$ pair can be classified as "single-lepton", "dilepton", or "all hadronic" according to the decays of the two W bosons: a pair of quarks ($W \to q_1\bar{q}_2$) or a lepton-neutrino pair ($W \to \ell\nu$). At the Tevatron the dominant production mechanism is $q\bar{q}$ annihilation, and the $t\bar{t}$ cross-sections at $\sqrt{s} = 1.8$ TeV and at $\sqrt{s} = 1.96$ TeV have been measured by CDF and DØ [4] in all channels. The production of $t\bar{t}$ at the LHC is dominated by gg fusion. Recently, ATLAS and CMS presented the first measurements of the $t\bar{t}$ cross-section at the LHC. CMS measured $\sigma_{t\bar{t}} = 194 \pm 72$ (stat.) ± 24 (syst.) ± 21 (lumi.) pb in the dilepton channel using 3.1 pb^{-1} of data [5] and ATLAS measured $\sigma_{t\bar{t}} = 145 \pm 31$ (stat.) ± 42 (syst.) ± 27 (lumi.) pb combining the single lepton and dilepton channels in 2.9 pb^{-1} of data [6].

This note describes measurements of the $t\bar{t}$ cross-section in the single lepton plus jets channel with 35 pb^{-1} of data collected by ATLAS in 2010 with high p_T electron and muon triggers. The results are based on analyses which do not require any b-tagged jets. This is a complementary approach to $t\bar{t}$ analyses that explicitly attempt to identify secondary vertices produced by b-jets. The analyses presented here are free of uncertainties related to b-tagging, such as the b-tagger calibration uncertainty and the heavy flavour content of the background, although they do suffer from a larger relative background contamination. The dominant backgrounds, namely W+jet and QCD multi-jet production, are normalized using data driven techniques since their overall rate is relatively difficult to predict theoretically. We take advantage of the increased data sample to extend the measurement techniques developed in [6] by employing likelihood fitting methods and extended "cut-and-count" methods to extract cross-section measurements. The method that yields the best expected uncertainty, and which is quoted as the main result, utilises a multivariate fit to three variables that discriminate $t\bar{t}$ events from the background. Complementary methods which use "cut-and-count" methods and simpler one-dimensional fits are also presented.

2 Detector, Data and Simulated Samples

The ATLAS detector [7] at the LHC covers nearly the entire solid angle[1] around the collision point. It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and an external muon spectrometer incorporating three large superconducting toroid magnet assemblies. Only data where all these subsystems are fully operational are used. Applying these requirements to all of the $\sqrt{s} = 7$ TeV pp collision data taken in stable beam conditions in 2010 results in a data sample of 35 pb^{-1}. This ATLAS luminosity uncertainty has recently improved from 11% to 3.4% [8], which results in a relative improvement of 20% in the $t\bar{t}$ cross-section uncertainty presented

[1] In the right-handed ATLAS coordinate system, the pseudorapidity η is defined as $\eta = -\ln[\tan(\theta/2)]$, where the polar angle θ is measured with respect to the LHC beamline. The azimuthal angle ϕ is measured with respect to the x-axis, which points towards the centre of the LHC ring. The y-axis points up. Transverse momentum and energy are defined as $p_{\mathrm{T}} = p\sin\theta$ and $E_{\mathrm{T}} = E\sin\theta$, respectively.

in this note.

Exactly the same Monte Carlo (MC) generator samples with their associated systematic uncertainties as developed for the previous $t\bar{t}$ cross-section measurement [6] are employed. However the detector simulation has been upgraded to reflect the improved knowledge of the detector material, alignment, geometry and calibration acquired since then. The presence of additional pp interactions superimposed on each collision event, i.e. "pile-up" interactions, is included in the MC simulation. A small pile-up uncertainty is considered to cover the remaining mismatch in the observed number of reconstructed primary vertices per event between data and MC.

3 Object and Event Selections

The reconstruction of $t\bar{t}$ events makes use of electrons, muons, jets and missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ which is a measure of the energy imbalance in the transverse plane. $E_{\mathrm{T}}^{\mathrm{miss}}$ is used as an indicator of undetected neutrinos. The same definition of these objects as in ref. [6] is used for the measurements presented here, except for the following refinements:

- To reduce further the fake electron background, tighter electron selections have been applied. In particular, high threshold hits in the Transition Radiation Tracker (TRT) are required to distinguish electrons from other particles.
- More stringent inner detector track quality cuts have been applied to the muon candidates.

The modelled acceptances and efficiencies of reconstructed particles are verified by comparing Monte Carlo and detector simulations with data in control regions. Lepton efficiencies are derived from data in the Z boson mass window and are validated by using them to estimate inclusive W and Z boson cross-sections. The acceptances for the jet multiplicity and $E_{\mathrm{T}}^{\mathrm{miss}}$ cuts are validated using a number of control regions surrounding the $t\bar{t}$ signal region in phase-space.

3.1 Systematic uncertainties for reconstructed objects

The uncertainties in the simulation modelling the lepton trigger, reconstruction and selection efficiencies are assessed using leptons from $Z \to ee$ and $Z \to \mu\mu$ events selected from the same data sample used for the $t\bar{t}$ analyses. Scale factors are then applied to Monte Carlo samples when calculating acceptances. The statistical and systematic uncertainties on these scale factors are included in the uncertainties on the acceptance values. The modelling of the lepton energy scale and resolution is studied using reconstructed Z boson mass distributions to adjust the simulation.

The jet energy scale (JES) and its uncertainty are derived by combining information from single particle response studies of collision and test-beam data, jet control samples and simulation [13]. The JES uncertainty varies in the range of approximately 4–8% as a function of jet p_T and η. Jet energy resolution and jet finding efficiency are measured in data and in simulation to extract systematic uncertainties. They have a minor impact on the cross-section measurement in comparison to the jet energy scale.

The lepton and jet energy scale and resolution uncertainties are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation when estimating their impact on the cross-section measurement.

3.2 Event selections

The single lepton $t\bar{t}$ final state is characterized by an isolated lepton with relatively high p_T, missing transverse energy arising from the neutrino from the leptonic W decay, two b quark jets and two light quark jets from the hadronic W decay. The selection of events for the single-lepton analysis consists of a

series of requirements on the reconstructed objects defined in Sec. 3, designed to select events with the above topology. For each lepton flavour, the following event selections are first applied:

- the appropriate single-electron or single-muon trigger has fired;
- the event contains one and only one reconstructed lepton (electron or muon) with $p_T > 20$ GeV, matching the corresponding high-level trigger object;
- in the muon channel, $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} + m_T(W) > 60$ GeV is required[2]. The cut on $E_{\mathrm{T}}^{\mathrm{miss}}$ rejects a significant fraction of the QCD multi-jet background. Further rejection can be achieved by applying a cut in the ($E_{\mathrm{T}}^{\mathrm{miss}}$, $m_T(W)$) plane; true $W \to \ell\nu$ decays with large $E_{\mathrm{T}}^{\mathrm{miss}}$ also have large $m_T(W)$, while mis-measured jets in QCD multi-jet events may result in large $E_{\mathrm{T}}^{\mathrm{miss}}$ but small $m_T(W)$. The requirement on the sum of $E_{\mathrm{T}}^{\mathrm{miss}}$ and $m_T(W)$ discriminates between these two cases;
- in the electron channel more stringent cuts on $E_{\mathrm{T}}^{\mathrm{miss}}$ and $m_T(W)$ are required because of the more important QCD multi-jet background, i.e. $E_{\mathrm{T}}^{\mathrm{miss}} > 35$ GeV and $m_T(W) > 25$ GeV;
- finally, the event is required to have ≥ 1 jet with $p_T > 25$ GeV and $|\eta| < 2.5$. The requirement on the p_T and the pseudorapidity of the jets is a compromise between the efficiency of the $t\bar{t}$ event selection, and the rejection of W+jets and QCD multi-jet background.

Events are then classified by the number of jets with $p_T > 25$ GeV and $|\eta| < 2.5$, being either 1, 2, 3 or at least 4. The number of events observed in data and predicted by simulation or by data-driven estimates (for QCD multi-jet as discussed in Sec. 4.1) are given in Table 1. The uncertainty on the number of expected events comes from the data-driven method, in the case of the QCD background, or from the theory predictions in the cases of the other processes. The number of observed and expected events are in good agreement for each jet bin and lepton flavor. The distribution of $m_T(W)$ in the 2-jet control region is shown in Fig. 1. A good agreement between data and predictions is observed in this region which is dominated by the W+jets background. The distribution of the reconstructed hadronic top quark mass, defined as the invariant mass of the three jets with the highest vector sum p_T [6], is shown in Fig. 2 for events with ≥4-jets. These events contain a significant fraction of $t\bar{t}$ events and again a good agreement between data and MC predictions is observed.

The estimated products of acceptance and branching fraction for $t\bar{t}$ events, measured from Monte Carlo samples, are 3.5% and 5.8% in the electron channel for events with exactly 3-jets and ≥4-jets, respectively, and 5.1% and 8.6% in the muon channel for events with exactly 3-jets and ≥4-jets, respectively.

4 QCD Data Driven Background Estimation

4.1 QCD background estimate in the μ+jets channel

In the μ+jets channel, the background to "real" (prompt) muons coming from non-prompt muons in QCD multi-jet events is predominantly due to heavy flavor jets containing hadrons decaying semileptonically. As all other processes in this channel ($t\bar{t}$, W+jets, Z+jets and single-top) feature a prompt muon from a W or Z boson decay, it is sufficient to estimate the number of events with a non-prompt muon to quantify the QCD multi-jet background.

The number of events in the sample with a non-prompt muon can be extracted from the data by considering the event count in the signal region with two sets of muon identification criteria. The "standard"

[2]Here $m_T(W)$ is the W-boson transverse mass, defined as $\sqrt{2p_T^{\ell}p_T^{\nu}(1-\cos(\phi^{\ell}-\phi^{\nu}))}$ where the measured missing E_{T} vector provides the neutrino information.

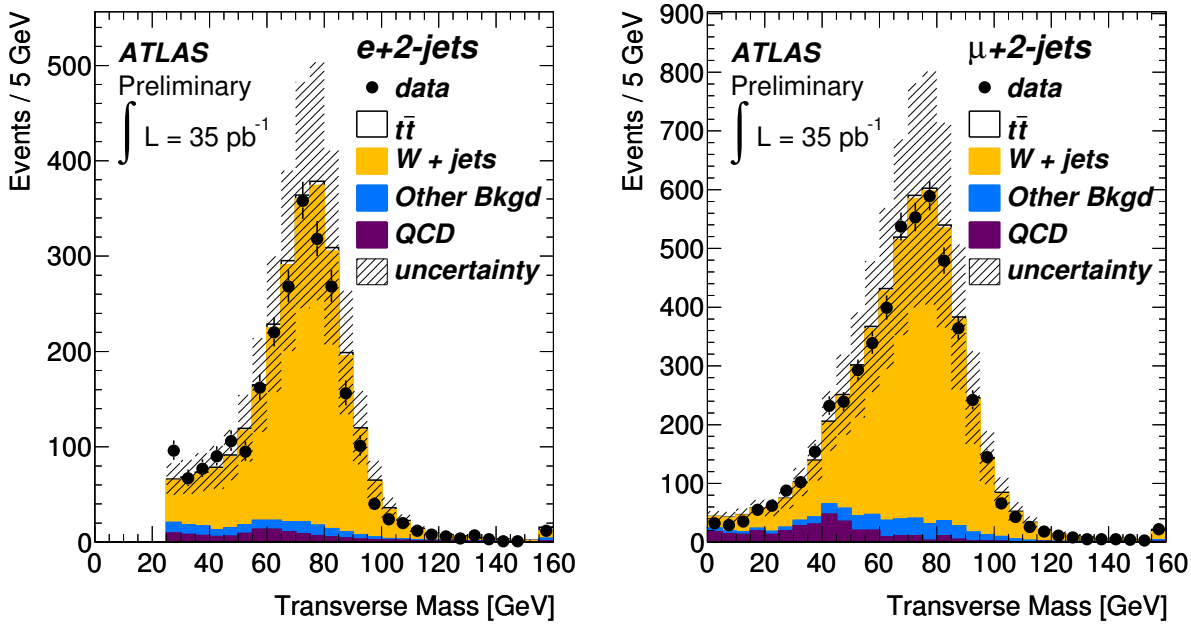
e+jets channel				
	1-jet	2-jet	**3-jet**	**≥4-jet**
$t\bar{t}$ (MC)	14 ± 3	61 ± 9	116 ± 13	194 ± 27
QCD (DD)	290 ± 140	123 ± 62	62 ± 31	22 ± 11
W+jets (MC)	9000 ± 1900	2300 ± 700	580 ± 250	180 ± 120
Z+jets (MC)	65 ± 14	62 ± 20	32 ± 14	18 ± 11
Single-top (MC)	36 ± 4	42 ± 5	22 ± 4	11 ± 3
Dibosons (MC)	35 ± 3	30 ± 2	9 ± 2	3 ± 1
Total background	9400 ± 1900	2600 ± 800	710 ± 250	240 ± 120
Total expected	9400 ± 1900	2700 ± 800	830 ± 250	430 ± 120
Observed	9481	2552	781	400

(a)

μ+jets channel				
	1-jet	2-jet	**3-jet**	**≥4-jet**
$t\bar{t}$ (MC)	19 ± 4	81 ± 12	161 ± 18	273 ± 38
QCD (DD)	520 ± 160	287 ± 86	121 ± 36	51 ± 15
W+jets (MC)	19000 ± 4000	4600 ± 1500	1100 ± 500	310 ± 200
Z+jets (MC)	770 ± 160	250 ± 80	69 ± 30	25 ± 16
Single-top (MC)	57 ± 7	64 ± 8	32 ± 6	15 ± 4
Dibosons (MC)	63 ± 5	55 ± 4	16 ± 3	4 ± 1
Total background	20000 ± 4000	5300 ± 1500	1300 ± 500	400 ± 200
Total expected	20000 ± 4000	5300 ± 1500	1500 ± 500	680 ± 200
Observed	20582	5228	1356	653

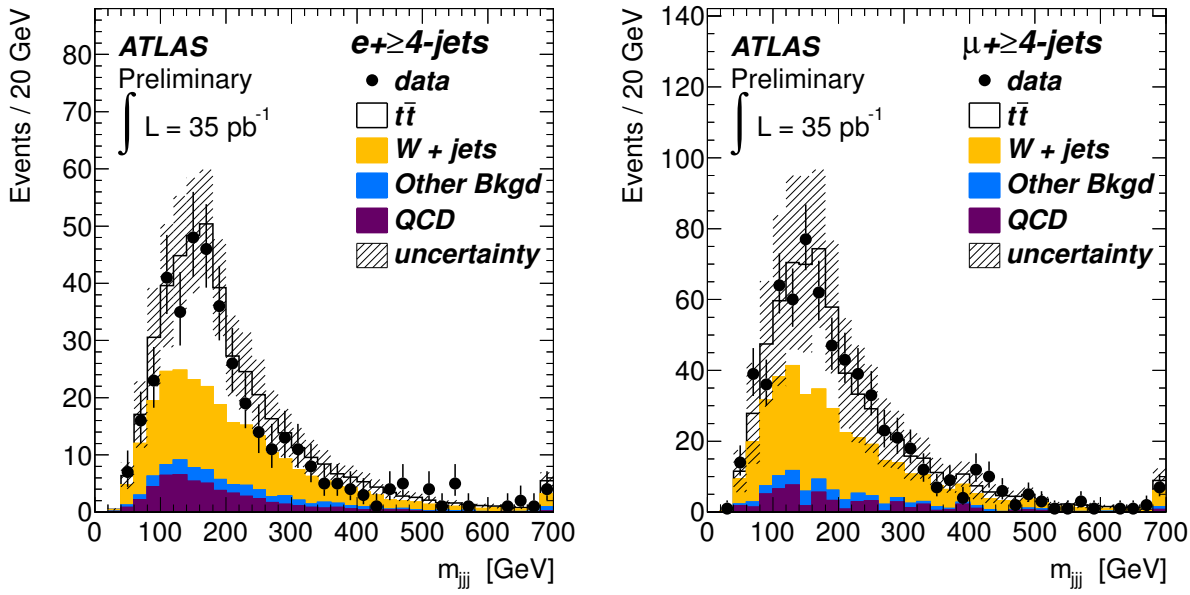
(b)

Table 1: Numbers of events with different jet multiplicities in the (a) electron and (b) muon channels. The observed number of events are shown, together with the Monte-Carlo simulation estimates for $t\bar{t}$, W+jets, Z+jets and single-top events, normalised to the data integrated luminosity of 35 pb^{-1}. The data-driven estimates (DD) for QCD multi-jet background (see Sec. 4.1) are also shown. The uncertainties on the data-driven background estimates include the statistical uncertainty and the (dominant) systematic uncertainties. The number of expected events and total background have been rounded to two significant digits.



(a) (b)

Figure 1: Transverse mass of the W boson in events with exactly two jets for (a) the electron channel, (b) the muon channel. The QCD multi-jet background is obtained from the data-driven methods described in the text whilst all the other backgrounds and $t\bar{t}$ are obtained from MC simulation. The uncertainty on the MC and data-driven predictions are shown.



(a) (b)

Figure 2: Distribution of the invariant mass m_{jjj} of the three jets with the highest vector sum p_T for the events passing the baseline event selection in (a) the electron channel and (b) the muon channel. The QCD multi-jet background is obtained from the data-driven methods described in the text whilst all the other backgrounds and $t\bar{t}$ are obtained from MC simulation. The uncertainty on the MC and data-driven predictions are shown.

and "loose" criteria comprise the standard muon definition with and without, respectively, the requirements on the isolation, i.e. that the corresponding calorimeter isolation energy in a cone of $\Delta R = 0.3$ be less than 4 GeV and that the scalar sum of track transverse momenta in a cone of $\Delta R = 0.3$ be less than 4 GeV after subtraction of the muon p_{T}.

The procedure followed at this point is the so-called "matrix method": the number of events selected by the loose and by the standard cuts, N^{loose} and N^{tight} respectively, can be expressed as linear combinations of the number of events with a prompt or a non-prompt muon:

$$\begin{aligned} N^{\mathrm{loose}} &= N^{\mathrm{loose}}_{\mathrm{prompt}} + N^{\mathrm{loose}}_{\mathrm{non-prompt}}, \\ N^{\mathrm{tight}} &= \epsilon_{\mathrm{prompt}} N^{\mathrm{loose}}_{\mathrm{prompt}} + \epsilon_{\mathrm{non-prompt}} N^{\mathrm{loose}}_{\mathrm{non-prompt}}, \end{aligned} \tag{1}$$

where $\epsilon_{\mathrm{prompt}}$ is the fraction of prompt muons in the loose selection that also pass the standard selection and $\epsilon_{\mathrm{non-prompt}}$ is the fraction of non-prompt muons in the loose selection that also pass the standard selection. If $\epsilon_{\mathrm{prompt}}$ and $\epsilon_{\mathrm{non-prompt}}$ are known, the number of events with non-prompt muons can be calculated from Equation 1 given a measured N^{loose} and N^{tight}. The relative efficiencies $\epsilon_{\mathrm{prompt}}$ and $\epsilon_{\mathrm{non-prompt}}$ are measured in data in control samples enriched in either prompt or non-prompt muons. The key issue in selecting these control regions is that they should be kinematically representative of the signal region so that the measured control-region efficiency can be applied in the signal region.

An inclusive $Z \rightarrow \mu^+\mu^-$ control sample was used to measure the prompt muon efficiency. The measurement of the non-prompt muon efficiency was performed in a control region with low transverse W mass and $E_{\mathrm{T}}^{\mathrm{miss}}$, i.e. $m_T(W) < 20$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} + m_T(W) < 60$ GeV. The prompt muon contamination in the control region by true W events has been subtracted using MC simulation. The matrix method is then applied to obtain the numbers shown in Table 1 as a function of jet multiplicity. This calculation has been cross-checked using a different control region to measure the non-prompt muon efficiency, i.e. $E_{\mathrm{T}}^{\mathrm{miss}} < 10$ GeV. The two methods agree within 30% or better, which is used as the systematic uncertainty on the QCD multi-jet estimate in the muon channel.

The shape of the QCD multi-jet background as a function of a given variable can be obtained by applying the matrix method in bins of the variable under consideration. A possible bias in this shape will affect the cross-section measurements based on likelihood fits that are presented in Secs. 5 and 6.1. The systematic uncertainty on the QCD shape in the muon channel is taken as the difference between the matrix method shape and the alternative electron shape model described in the next section[3].

4.2 QCD background estimate in the e+jets channel

In the e+jets channel, the non-prompt electron background consists of both real electrons and fake electrons where the latter includes both electrons from photon conversion and misidentified jets with high electromagnetic fractions. The relative magnitude of the real and fake lepton components is not well known, as it depends on the details of electron mis-reconstruction effects that are not perfectly modelled in the simulation. As the ratio also varies with the event kinematics, the method of Equation 1, which relies on a representative control region to measure the input values of $\epsilon_{\mathrm{non-prompt}}$, is not suitable for the electron channel.

A method based on a binned likelihood template fit of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is used for the background estimate. The data are fitted to a sum of four templates describing the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of the QCD multi-jet, $t\bar{t}$, W+jets and Z+jets components respectively. The fit is performed in the region with $E_{\mathrm{T}}^{\mathrm{miss}} < 35$ GeV which is complementary to the signal region. The QCD multi-jet template is extracted from the data as described in the next paragraph, while the templates for the other processes are taken from the simulation. The fraction of QCD multi-jet events in the signal region is then calculated by

[3]The QCD electron method has been applied to the muon channel and yields consistent QCD estimates. It is thus considered to constitute a realistic alternative QCD model.

extrapolating the expected fraction of events for each component to the signal region using the template shape. The output of the fit is the predicted fraction of QCD multi-jet events in the signal region, which is then multiplied by the observed event count.

The templates for the QCD multi-jet $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions are obtained in a data electron sample where all the standard event selections are applied, except that the electron candidate must fail one or several of the standard identification criteria. Several such "anti-electron" definitions have been explored and the one that provided the best predictions in data control regions is chosen. The choice corresponded to varying cuts on the limited energy leakage in the hadronic calorimeter. The method is then applied to data to obtain the numbers shown in Table 1. Several other "anti-electron" definitions were used which all agree within 50% or better. The differences were used to estimate the systematic uncertainty on the QCD multi-jet estimate in the electron channel.

The systematic uncertainty on the QCD shape in the electron channel is defined as the difference between the default "anti-electron" definition and the one that shows the largest shape difference, but still yields consistent QCD estimates.

5 Multivariate Cross-Section Measurement

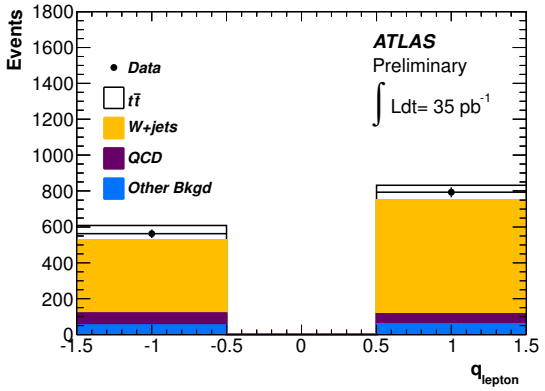
The $t\bar{t}$ production cross-section can be extracted by exploiting the different properties of $t\bar{t}$ events with respect to the dominant W+jets background. Three variables were selected which were as uncorrelated as possible and consistent with reduced statistical and jet energy scale uncertainties in the overall analysis. These variables are:

- the pseudorapidity of the lepton η_{lepton}, which exploits the fact that $t\bar{t}$ events produce more central leptons than W+jet events;
- the charge of the lepton q_{lepton}, which uses the fact that $t\bar{t}$ events produce charge-symmetric leptons while W+jet events produce an excess of positively charged leptons;
- the exponential of the aplanarity ($\exp(-8 \times \mathcal{A})$), where $\mathcal{A} = \frac{3}{2}\lambda_3$ and where λ_3 is the smallest eigenvalue of the normalized momentum tensor calculated using the selected jets and lepton in the event. This variable exploits the fact that $t\bar{t}$ events are more isotropic than W+jets.

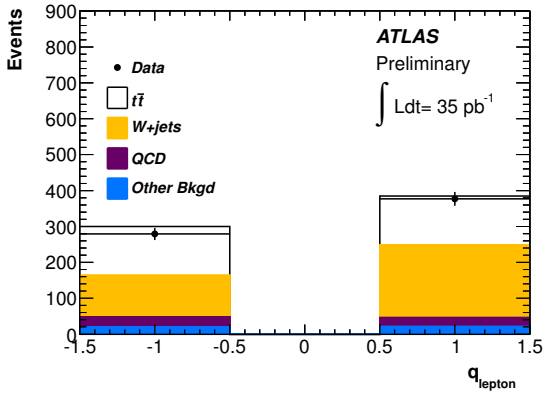
This measurement utilizes events from the μ+jets and e+jets channels, as defined in Sec. 3. In addition to using the usual $t\bar{t}$ signal region that requires four or more jets, events with exactly three jets are also used, but fitted separately, to increase the $t\bar{t}$ acceptance by approximately 37%. The number of observed and expected events for each process are given in Table 1. The distributions of these variables in data and simulated events are shown in Figs. 3 and 4 for the muon and electron channels, respectively. In these plots, the normalization of the simulated processes is taken from theoretical predictions (including for $t\bar{t}$) described in Sec. 2, except for QCD multi-jet that uses the normalization extracted in Sec. 4.1. Reasonable agreement has been found in the shape of these variables between data and MC using a Kolmogorov-Smirnov test. A probability of 8% is, however, observed in the worst case (muon η in the 3-jets sample).

A likelihood discriminant is built from these input variables following the projective likelihood approach defined in the TMVA package [14]. The likelihood discriminant D_i for an event i is defined as the ratio of the signal ($\mathcal{L}_S(i)$) to the signal plus background ($\mathcal{L}_S(i) + \mathcal{L}_B(i)$) likelihoods:

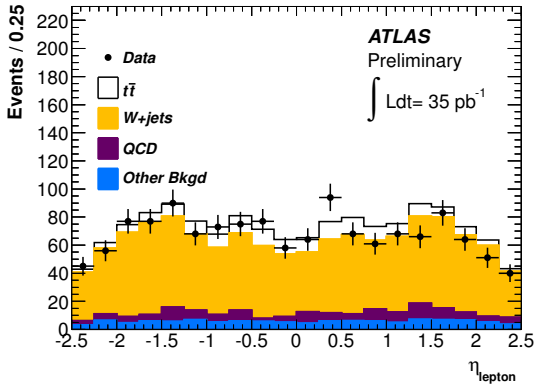
$$D_i = \frac{\mathcal{L}_S(i)}{\mathcal{L}_S(i) + \mathcal{L}_B(i)}. \tag{2}$$



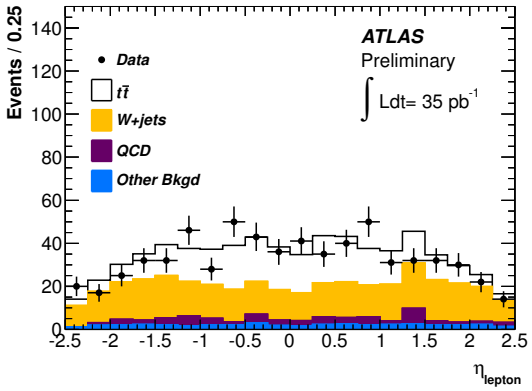
(a) 3-jets



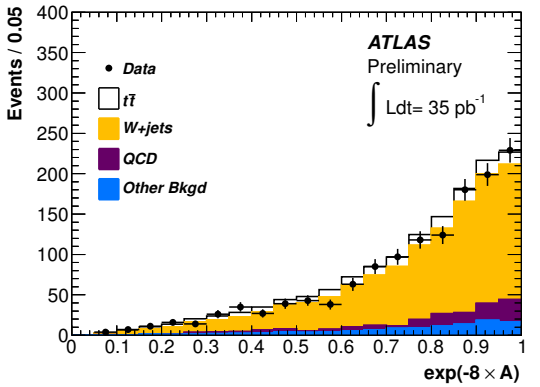
(b) ≥4-jets



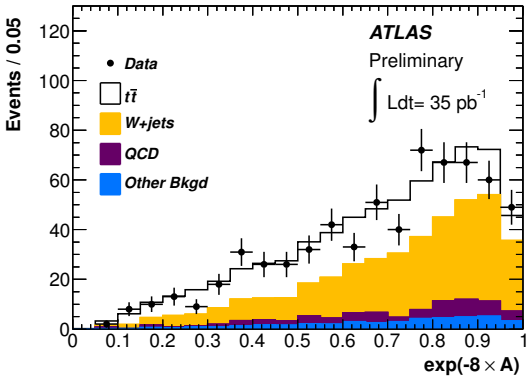
(c) 3-jets



(d) ≥4-jet

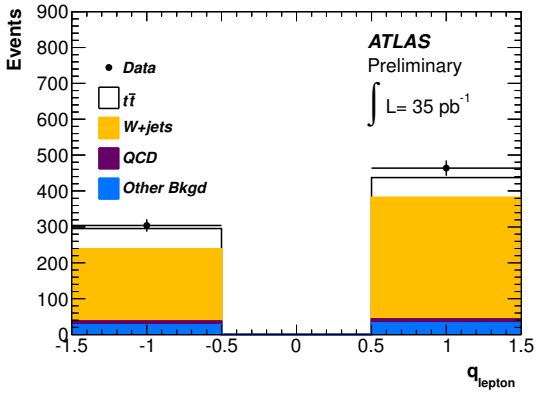


(e) 3-jets

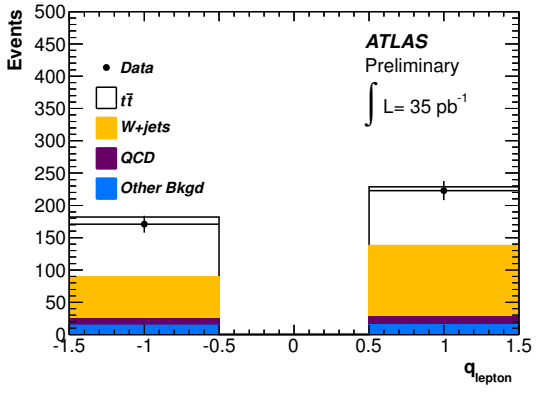


(f) ≥4-jet

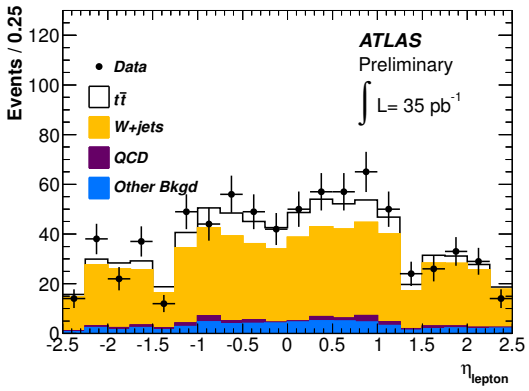
Figure 3: Input variables to the likelihood discriminant for the muon channel. The plots on the left and right are for events with exactly 3-jets and ≥4-jets, respectively. The variables are the lepton charge (top row), lepton η (middle row) and $\exp(-8 \times \mathcal{A})$, where $\mathcal{A}$ is the aplanarity of the jets and lepton (bottom row). The normalization of the simulated processes is taken from theoretical predictions (including for $t\bar{t}$) except for QCD multi-jet which uses the normalization extracted in Sec. 4.1.



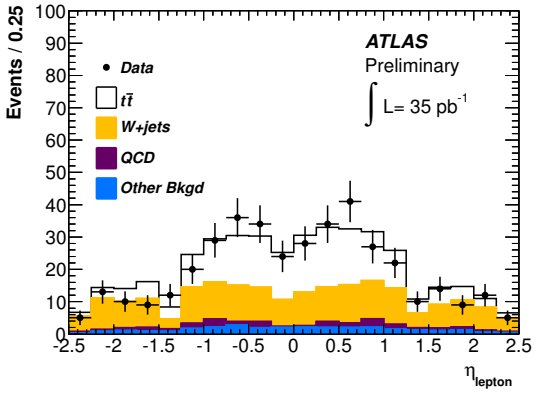
(a) 3-jets



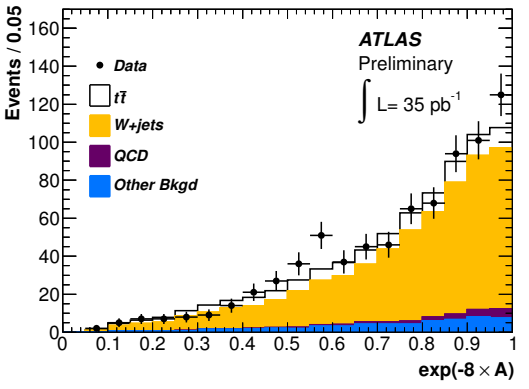
(b) ≥4-jets



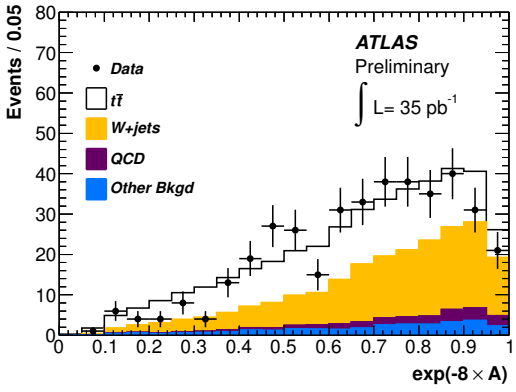
(c) 3-jets



(d) ≥4-jet



(e) 3-jets



(f) ≥4-jet

Figure 4: Input variables to the likelihood discriminant for the electron channel. The plots on the left and right are for events with exactly 3-jets and ≥4-jets, respectively. The variables are the lepton charge (top row), lepton η (middle row) and $\exp(-8 \times \mathcal{A})$, where $\mathcal{A}$ is the aplanarity of the jets and lepton (bottom row). The normalization of the simulated processes is taken from theoretical predictions (including for $t\bar{t}$) except for QCD multi-jet which uses the normalization extracted in Sec. 4.1.

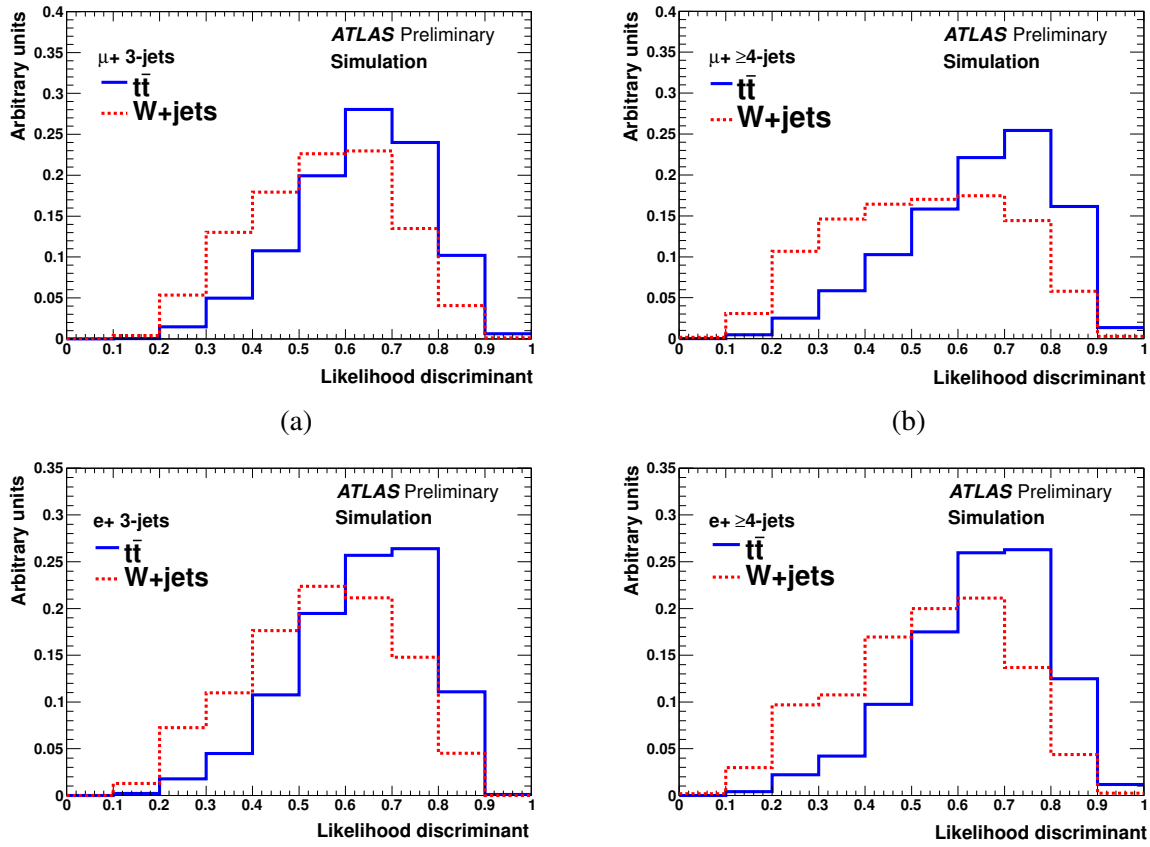


Figure 5: Likelihood discriminants for $t\bar{t}$ (solid blue) and W+jets (dashed red). Top row - muon channel: (a) events with 3-jets, (b) events with ≥4-jets. Bottom row - electron channel: (c) events with 3-jets, (d) events with ≥4-jets. The shapes have been normalized to the same area.

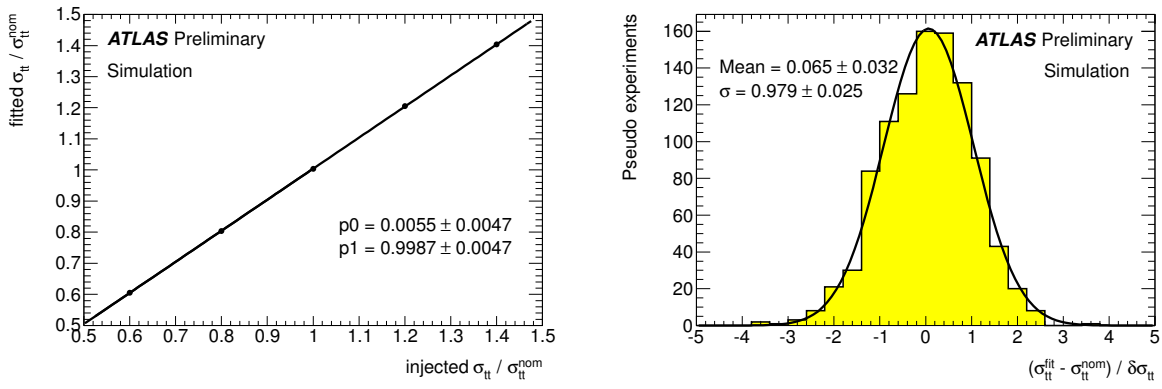
The individual likelihoods are products of the corresponding probability densities of the discriminating input variables x_k defined above, for example for the signal likelihood:

$$\mathcal{L}_S(i) = \prod_{k=1}^{3} p_{S,k}(x_k(i)). \tag{3}$$

This simple multivariate approach assumes that the variables x_k are uncorrelated. The variables used in the analysis have been found to have small correlation coefficients, up to 2% in the worst case. The resulting templates of the likelihood discriminants are shown in Fig. 5 for $t\bar{t}$ and W+jet events normalized to the same area.

A binned maximum likelihood fit is applied to the discriminant shapes described previously to extract the $t\bar{t}$ cross-section. Likelihood functions are defined for each of the four channels (e and μ, 3-jets and ≥4-jets) and are multiplied together in a combined fit to extract the total number of $t\bar{t}$ events. The QCD multi-jet and small backgrounds (single top, diboson. Z+jets) are fixed to their expected contributions in the fit. The $t\bar{t}$ cross-section is then extracted using the usual formula:

$$\sigma_{t\bar{t}} = \frac{N_{sig}}{\int \mathcal{L}dt \times \epsilon_{sig}}, \tag{4}$$



(a) (b)

Figure 6: (a) Ratio of the fitted to the nominal $t\bar{t}$ cross-section as a function of the input ratio, (b) Pull distribution of the likelihood fit at the nominal cross-section.

where N_{sig} is the number of $t\bar{t}$ events extracted from the fit, $\int \mathcal{L}dt$ is the integrated luminosity and ϵ_{sig} is the product of the signal acceptance, efficiency and branching ratio, which is calculated using the $t\bar{t}$ simulation augmented with data-to-simulation scale factors described in Sec. 3. The performance of the likelihood fit is estimated by performing pseudo-experiments. Signal and background events are drawn randomly from the discriminant templates and the same fit as used on the data is applied to this simulated data sample. This operation is repeated a thousand times. The pull distribution and test of the linearity of the fit as a function of the input $t\bar{t}$ cross-section demonstrate the validity of the method, as shown in Fig. 6. The expected statistical uncertainty of the likelihood fit is 9.7% with the current integrated luminosity.

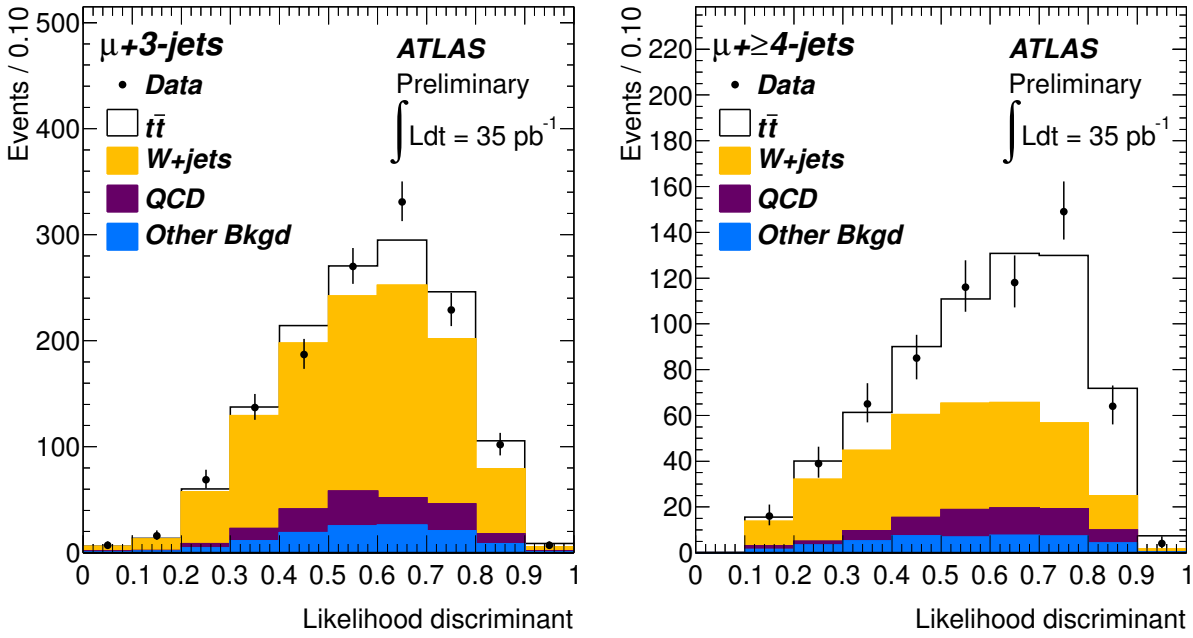
The systematic uncertainties of the measurement are extracted using the pseudo-experiments method described above. These are associated with the simulation, object definitions and the QCD multi-jet estimate as described in Secs. 2, 3 and 4.1, respectively. Some of the uncertainties, including the jet energy scale, will affect both the shape of the discriminant templates as well the $t\bar{t}$ acceptance, and thus both N_{sig} and ϵ_{sig} in Eq. 4. The correlation between these two effects is taken into account when calculating the systematic uncertainties on the $t\bar{t}$ cross-section. The list of all systematic uncertainties are provided in Table 2. They are added in quadrature to give a relative uncertainty on the cross-section of -10.2/+11.6%. The effects of the parton distribution function uncertainty on both the $t\bar{t}$ acceptance and the W charge asymmetry have been considered. The dominant systematic uncertainty is the jet energy scale uncertainty.

The results of the likelihood fit applied to the data are shown in Fig. 7, where the shape of the discriminant in the data is overlaid on the signal and background templates in the proportion returned by the fit. The combined fit yields $\sigma_{t\bar{t}} = 171 \pm 17$ pb (stat.). The results of the likelihood applied to the four individual channels are given in Table 3 and are found to be in good agreement with each other.

In conclusion a measurement of the $t\bar{t}$ cross-section has been performed using a simple multivariate fit analysis without b-tagging. The final result is

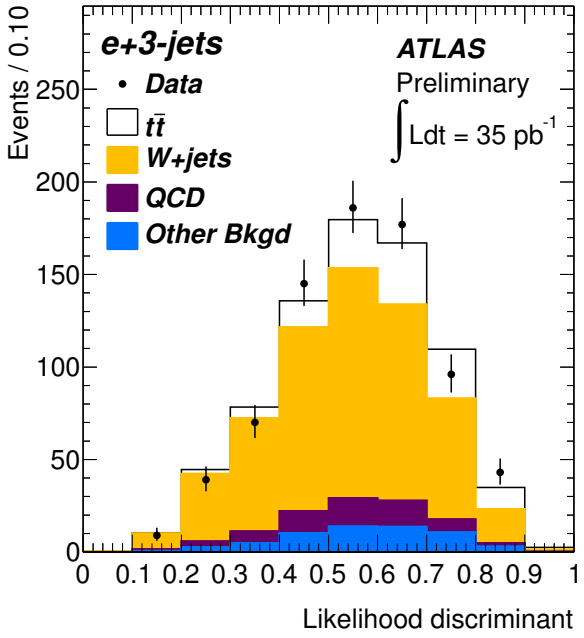
$$\sigma_{t\bar{t}} = 171 \pm 17(\text{stat.})^{+20}_{-17}(\text{syst.}) \pm 6(\text{lumi.})\ \text{pb}, \tag{5}$$

for a total relative uncertainty of -14.5/+15.5%. The measured cross-section is in good agreement with theoretical predictions.

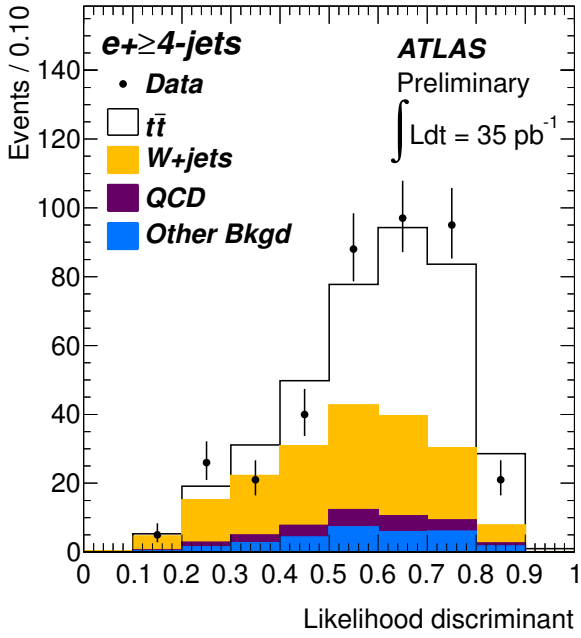


(a)

(b)



(c)



(d)

Figure 7: Multivariate discriminant for the data, as well as for the signal and background processes in the proportions returned by the fit for (a) μ+3-jets, (b) μ+≥4-jets, (c) e+3-jets, (d) e+≥4-jets.

Source	Relative cross-section uncertainty [%]
Object selection	
Lepton reconstruction, identification, trigger	-1.9 / +2.6
Jet energy scale and reconstruction	-6.1 / +5.7
Background rates and shape	
QCD normalisation	±3.9
QCD shape	±3.4
W+jets shape	±1.2
Other backgrounds normalisation	±0.5
Simulation	
Initial/final state radiation	-2.1 / +6.1
Parton distribution functions	-3.0 / +2.8
Parton shower and hadronisation	±3.3
Next-to-leading-order generator	±2.1
MC statistics	±1.8
Pile-up	±1.2
Total systematic uncertainty	-10.2 / +11.6

Table 2: Summary of individual systematic uncertainty contributions to the multivariate fit analysis.

Channel	$\sigma_{t\bar{t}}$ (pb)
e + 3 jets	225 ± 72
e + ≥4 jets	182 ± 29
μ + 3 jets	143 ± 67
μ + ≥4 jets	164 ± 24
Combined	171 ± 17

Table 3: Results of the likelihood fit to individual channels in the data (statistical uncertainty only).

6 Further Cross-Section Measurements

Several complementary methods (with slightly higher total uncertainties) have been used to measure the $t\bar{t}$ cross-section in the single lepton channel without b-tagging and are described in this section.

6.1 One-dimensional likelihood fits

Simpler fits using only one variable have also been used to extract the $t\bar{t}$ cross-section. These methods do not provide as precise a measurement as the one described in Sec. 5, but have the advantage of being simpler and sensitive to different systematics, and thus provide good cross-checks. Two variables are used for these cross-checks: $\max \Delta\eta$, the maximum pseudorapidity difference between the lepton and any of the three highest p_{T} jets, and the lepton pseudorapidity $|\eta_{\mathrm{lepton}}|$. Two different fitting methods are used to extract the top cross section from these two variables.

The $\max \Delta\eta$ uses a modified χ^2 function [15] to fit for 5 parameters; the number of $t\bar{t}$ events and the number of W+jet events are floated freely in the fit, the number of QCD events is fixed for all but the electron channel fit where it is constrained within its uncertainty (50%), the other sources: Z+jets, diboson and single-top are constrained to their predicted number of events within the uncertainty of the prediction. It is found that the small sources do not contribute significantly to the fit result, thus fixing them in the fit would give very similar results. Both 3-jets and ≥4-jets samples are used in the measurement.

The analysis methods for both cross-checks are very similar to those of the main analysis in terms

of systematic variations; the same sources of systematic uncertainties are considered and the resulting uncertainty on the cross section are obtained by computing the mean shift over 1000 pseudo-experiments.

In order to check the fitting methods, for both cross-check analyses the linearity and pull tests were performed. The fits were found to perform as expected with no significant biases and show that the errors obtained from the fit are in agreement with the expectations. Both analyses compared expected distributions to the observed ones in the non-signal regions as well as checking that the description of the data was reasonable for the variables used for the event selection.

Figure 8(top) shows the result of the combined fit to the max $\Delta\eta$ distribution for the simultaneous fit to the electron and muon channels. The final result is

$$\sigma_{t\bar{t}} = 168 \pm 21 \text{ (stat)} \pm 24 \text{ (sys)} \pm 6 \text{ (lumi) pb.}$$

The statistical uncertainty of the fit, 12.2%, is consistent with the expectation of 13.3%. The dominant systematics are the jet energy scale, the jet efficiency, and the modeling of the $t\bar{t}$ signal both in terms of the generator and the parton shower, and are all found to be 5-7%. The total uncertainty on the measurement is 19.1%.

The $|\eta_{\text{lepton}}|$ fit uses a binned maximum likelihood fit with 2 parameters: the $t\bar{t}$ cross section and the W+jets cross section. The signal region is defined to contain ≥4-jets. The Z+jets shape is included in the W+jets shape; the data driven QCD shape (see Sec. 4.1) is fixed in the fit; the shape and normalisation of the single-top and dibosons are extracted from the MC samples and theoretical calculations and are fixed in the fit. The fit to the $|\eta_{\text{lepton}}|$ distribution is performed simultaneously in the muon and electron channels. The results of the fit are shown in Fig. 8(bottom). The expected statistical uncertainty of the fit is 16.6% and the one obtained from the fit to the data is 12.3%. The final result for this method is a cross section of

$$\sigma_{t\bar{t}} = 204 \pm 25 \text{ (stat)} \pm 39 \text{ (sys)} \pm 7 \text{ (lumi) pb,}$$

giving a total uncertainty on the cross section of 23.0%. The dominant systematics are the jet energy scale, the initial and final state radiation uncertainty as well as the uncertainty on the QCD background shape.

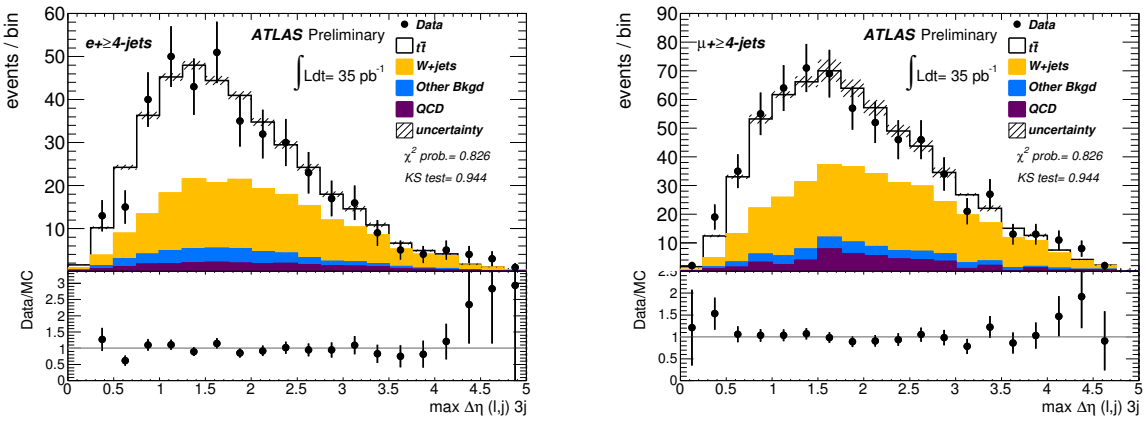
Both analyses also performed the fits separately in the electron and muon channels and found that the results were compatible with the combined results. Finally, both cross-check fit analyses find results compatible with the multivariate fit analysis presented in Sec.5.

6.2 Cross-section Measurement with Cut and Count Method

The most straightforward approach to measuring the $t\bar{t}$ cross-section is a simple "cut-and-count" procedure, in which one subtracts the number of expected background from the number of events observed in data and divides by the sample integrated luminosity and the total acceptance expected for the event selection cuts in $t\bar{t}$ events, obtained from MC simulation. The main challenge of this approach is to measure the dominant W+jets background without using shape information. Since the cross-section for this background is difficult to predict theoretically it is important to extract it from the data itself and two methods are described to accomplish this. Both rely on the fact that certain theoretically well understood cross-section ratios can be used to predict the W+jets contribution to the signal region.

The first approach relies on a theoretical prediction for the ratio of W and Z cross-sections. The W/Z ratio is better known than the inclusive W+jets rates, and it is approximately constant with jet multiplicity. The number of W events in the four-jet signal region can thus be estimated as

$$W^{\geq 4\text{jets}} = W^{1\text{jet}}_{\text{data}}(Z^{\geq 4\text{jets}}/Z^{1\text{jet}})_{\text{data}} \cdot C_{\text{MC}}, \qquad C_{\text{MC}} = \frac{(W^{\geq 4\text{jets}}/W^{1\text{jet}})_{\text{MC}}}{(Z^{\geq 4\text{jets}}/Z^{1\text{jet}})_{\text{MC}}} \tag{6}$$



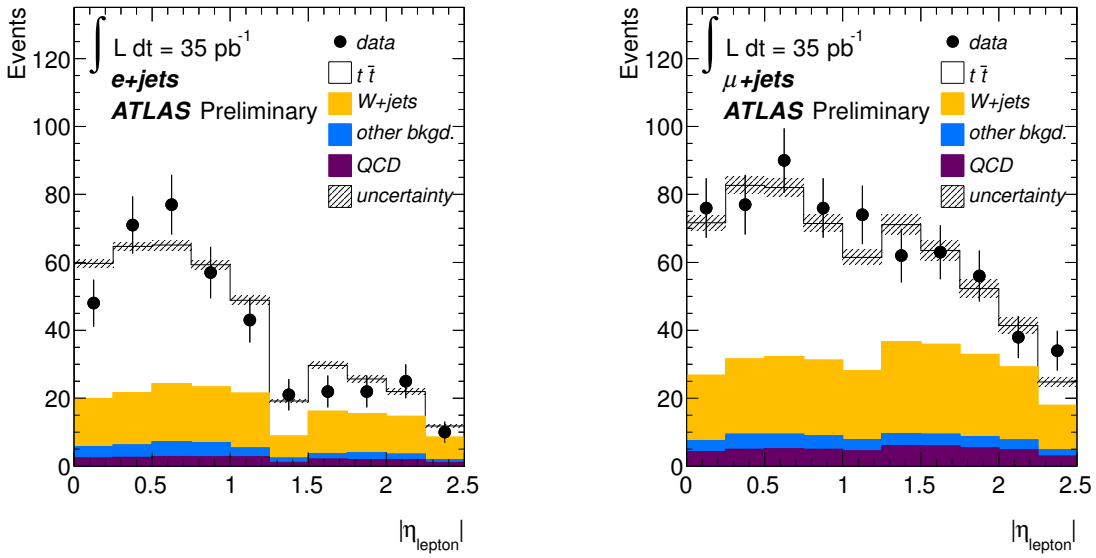


Figure 8: Top row: max $\Delta\eta$ distribution showing data compared to the sum of the signal and background predictions normalised to the output of the fit for (a) the electron and (b) the muon channels. Bottom row: Data and MC distributions of $|\eta_{\text{lepton}}|$ after the fit for (c) the electron and (d) the muon channels.

This method requires the selection of a W+1 jet control region that has the same selection as the signal region except that it requires only one jet with $p_T > 25$ GeV and $|\eta| < 2.5$. In this control region 9481 events are observed in the data in the electron channel and 20582 in the muon channel. The processes which are not W+jets can be subtracted by the amounts shown in Table 1 to yield a total of 8800 ± 200 W + 1-jet events in the electron channel and 18500 ± 300 in the muon channel. This is in good agreement with the W+jets MC prediction shown in Table 1.

A Z control sample is also required to apply Eq. 6. This sample selection uses the same object definitions as described in Sec. 3 but requires two electrons or muons with $p_T > 20$ GeV, with opposite charge and with an invariant mass between 80 GeV and 100 GeV. The W/Z ratio method requires the kinematical selection on the leptons from the Z to match those applied to the charged lepton and the neutrino in W and $t\bar{t}$ signal candidates. To match the muon selection, the sum of the transverse momentum of the negative lepton and the transverse mass of the two leptons is required to be larger than 60 GeV, in addition to the preselection cuts described above. In the electron channel, the negative lepton is required to have a p_T larger than 35 GeV and the transverse mass between the two leptons is required to be larger than 25 GeV. Finally, control region events are required to have at least one jet with $p_T > 25$ GeV and $|\eta| < 2.5$.

The QCD background estimate is the number of same sign leptons passing the selection. The other backgrounds were estimated from Monte Carlo simulation. The uncertainty on the $(Z + \geq 4\,\text{jet})\,/\,(Z + 1\,\text{jet})$ ratio is dominated by the statistical uncertainty of the numerator. Since this ratio is expected to be independent of the Z decay mode, the number of the $Z \to ee$ and $Z \to \mu\mu$ events is summed for each of the two selections to reduce the statistical uncertainty. The combined ratio is then measured to be

$$\frac{Z+ \geq 4j}{Z + 1j} = \frac{(39 \pm 6)}{(2530 \pm 50)} = ((1.52 \pm 0.25) \cdot 10^{-2} \tag{7}$$

for the muon channel selection and

$$\frac{Z+ \geq 4j}{Z + 1j} = \frac{(25 \pm 5)}{(1750 \pm 40)} = (1.42 \pm 0.29) \cdot 10^{-2} \tag{8}$$

for the electron channel selection.

The expected rate of W+jet events after the requirement of 4 jets has been evaluated using Eq. 6. The following two sources of systematic uncertainty have been considered in addition to the uncertainty on the measured $W \to l\nu$ and $Z \to ll$ rates discussed above. First, the uncertainty on the C_{MC} factor in Eq. 6, arising from the choice of matrix-element generator parameters and the parton distribution functions (PDF). The total theoretical uncertainty was found to be 12% in the electron channel and 9.4% in the muon channel[4]. The PDF uncertainty is estimated to be 3.2% in both channels. Second, the effects of detector reconstruction uncertainties have been evaluated; of these, the most important is that associated to the jet energy scale, which is about 3%.

In Table 4 the predicted number of $W \to l\nu$ events computed according to Eq. 6 is reported. The $W \to \tau\nu$ contribution, estimated from the data-driven $W \to l\nu$ rate and the Monte Carlo prediction of the ratio between $W \to \tau\nu$ and $W \to l\nu$ rates, is also reported in the table. The main sources of uncertainty are included.

The second approach for measuring the W+jets background is based on the fact that, while the top pair production is charge symmetric, W+jet production results in an excess of the positively charged leptons due to PDF effects. Since theoretically the charge asymmetry (which is determined by the W^+ and W^- cross-section ratios) in W+jets production is relatively well predicted, the total W+jet rate in a given sample can be estimated as

[4]They are different in the two channels due to the different event selections.

Channel	W/Z ratio **Electron**	W/Z ratio **Muon**	W^+/W^- **Electron**	W^+/W^- **Muon**
Estimated $W \to l\nu$	150	290	n.a.	n.a.
Estimated $W \to \tau\nu$	6	19	n.a.	n.a.
Statistical uncertainty	21%	17%	33%	27%
Purity of control samples	3%	2%		
Theoretical uncertainties	12%	9.4%	8.2%	7.0%
Jet energy scale	3%	3%	3.6%	3.6%
Total W+jets background	160 ± 40	310 ± 60	240 ± 80	380 ± 110

Table 4: Number of W background events estimated using the W/Z ratio method (2nd and 3rd columns) and the charge asymmetry method (last two columns)

$$N_{W^+} + N_{W^-} = \left(\frac{r_{MC}+1}{r_{MC}-1}\right)(D^+ - D^-), \tag{9}$$

where $D^+(D^-)$ are the total numbers of events in data with positively (negatively) charged lepton, and $r_{MC} \equiv \frac{\sigma(pp\to W^+)}{\sigma(pp\to W^-)}$ is evaluated for the signal region kinematic cuts from Monte Carlo simulation.

The formula is valid due to the fact that the processes $t\bar{t}$, QCD multi-jet and Z+jets are essentially charge symmetric, and so $N_{W^+} - N_{W^-} \approx D^+ - D^-$ to a very good approximation. Here $D^\pm$ denote the total number of events selected in the data with a positive or negative lepton and N_{W^+}, N_{W^-} are the numbers of W^+, W^- events in the signal region [5]. This technique has also been considered by CMS [12].

In the muon channel the 653 observed events in the signal region are composed of 374 anti-muons and 279 muons. The value of r_{MC} was calculated to be 1.67. The corresponding numbers in the electron channel are 230, 170 and 1.66. For the 35 pb^{-1} of data under consideration here the dominant source of uncertainty for this method is statistical. However, systematic uncertainties were also considered. The effects of Monte Carlo modelling on the value of r_{MC} were evaluated by comparing two different MC generators: Sherpa [10] and Alpgen [9]. PDF uncertainties were evaluated using the PDF4LHC prescription [11]. The impact of the uncertainty of the jet energy scale was also calculated. The uncertainty due to an incorrect lepton charge identification was found to be negligible. The results from the charge asymmetry method are given in Table 4 and are consistent with those from the W/Z ratio method.

Finally, note that the Berends scaling method used in [6] yields an estimated 180 ± 50 W+jet events in the electron channel and 320 ± 70 in the muon channel, in good agreement with the two methods presented here.

Having estimated the largest background to $t\bar{t}$, the $t\bar{t}$ cross-section can now be measured by subtracting all the other backgrounds similarly to the other cross-section measurements presented in this note. The W+jets estimate from the W/Z method is used since it yields the best uncertainty with the current dataset. The dominant source of systematic uncertainties on the cross-section measurement are due to uncertainties on the $t\bar{t}$ acceptance, i.e. the Jet Energy Scale/Reconstruction (13%), the modelling of Initial and Final State Radiation (8%), as well as the uncertainty on the total W+jets background from the data driven technique (21%). The final cross-section measurements obtained with the Cut-and-Count

[5] There is a small contribution to the asymmetry from the electroweak production of single top quarks, but this has been demonstrated to be negligible.

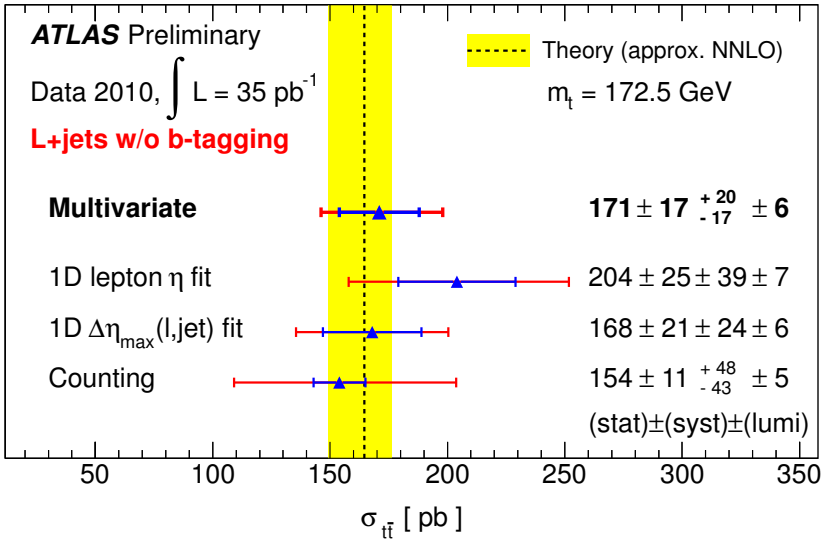


Figure 9: Summary of the $t\bar{t}$ cross-section measurements in the single lepton+jets channel without b-tagging using 35 pb^{-1} of data. The yellow band shows the approximate NNLO perturbative QCD prediction[1].

method are:

$$\sigma_{t\bar{t}}(e) = 159 \pm 17(\text{stat.})^{+50}_{-44}(\text{syst.}) \pm 5(\text{lumi.})\ \text{pb}, \quad (10)$$

$$\sigma_{t\bar{t}}(\mu) = 148 \pm 16(\text{stat.}) \pm 47(\text{syst.}) \pm 5(\text{lumi.})\ \text{pb}, \quad (11)$$

$$\sigma_{t\bar{t}}(\text{comb.}) = 154 \pm 11(\text{stat.})^{+48}_{-43}(\text{syst.}) \pm 5(\text{lumi.})\ \text{pb}. \quad (12)$$

7 Conclusions

Measurements of the $t\bar{t}$ production cross-section in the single-lepton channel without b-tagging using the ATLAS detector have been reported. The main analysis uses a multivariate fit of lepton charge, lepton rapidity and lepton-jets aplanarity to measure

$$\sigma_{t\bar{t}} = 171 \pm 17(\text{stat.})^{+20}_{-17}(\text{syst.}) \pm 6(\text{lumi.})\ \text{pb}.$$

Full $t\bar{t}$ cross-section measurements have been performed with simpler and complementary approaches to cross-check the main result. Two of them employ simpler one-dimensional fits of the lepton pseudorapidity and the maximum pseudorapidity difference between the lepton and one of the three highest p_{T} jets (max $\Delta\eta$), respectively. "Cut and count" measurements have also been performed where the W+jets background has been estimated using a data-driven approach.

The summary of the results presented in this note is presented in Fig. 9. The main result is in agreement with perturbative QCD calculations. The cross-check measurements are consistent with each other and with the main result.

References

[1] S. Moch and P. Uwer, Theoretical status and prospects for top-quark pair production at hadron colliders, Phys. Rev. D78 (2008) 034003;

U. Langenfeld, S. Moch, and P. Uwer, New results for $t\bar{t}$ production at hadron colliders, Proc. XVII Int. Workshop on Deep-Inelastic Scattering and Related Topics, dx.doi.org/10.3360/dis.2009.131, arXiv:0907.2527 [hep-ph].
Predictions in the paper are calculated with Hathor [2] with $m_{\text{top}} = 172.5$ GeV, CTEQ66 [3], where PDF and scale uncertainties are added linearly.

[2] M. Aliev *et al.*, HATHOR HAdronic Top and Heavy quarks crOss section calculatoR, arXiv:1007.1327 [hep-ph].

[3] J. Pumplin *et al.*, New generation of parton distributions with uncertainties from global QCD analysis, JHEP 07 (2002) 012.

[4] T. Affolder *et al.* (CDF Run I), Phys. Rev. D64 (2001) 032002, erratum-ibid. D67 (2003) 119901; CDF public note 10137 (2010) (Run II);
V. M. Abazov *et al.* (D0 Run I), Phys. Rev. D 67 (2003) 012004;
D0 note 6037-CONF (2010) (Run II).

[5] The CMS Collaboration, First Measurement of the Cross Section for Top-Quark Pair Production in Proton-Proton Collisions at sqrt(s)=7 TeV, Phys. Lett. B695 (2011).

[6] The ATLAS Collaboration, Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV, to be published by Eur. Phys J. C, arXiv:1012.1792 [hep-ex].

[7] The ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider, JINST 3 S08003 (2008).

[8] The ATLAS Collaboration, Updated Luminosity Determination in pp Collisions at $\sqrt(s) = 7$ TeV using the ATLAS Detector, ATLAS-CONF-2011-011, cdsweb.cern.ch/record/1334563.

[9] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, ALPGEN, a generator for hard multiparton processes in hadronic collisions, JHEP **0307** (2003) 001 [arXiv:hep-ph/0206293].

[10] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert and J. Winter, Event generation with SHERPA 1.1, JHEP **0902** (2009) 007 [arXiv:0811.4622 [hep-ph]].

[11] M. Botje, J. Butterworth, A. Cooper-Sarkar *et al.*, The PDF4LHC Working Group Interim Recommendations, [arXiv:1101.0538 [hep-ph]].

[12] CMS collaboration, Prospects for the first Measurement of the ttbar Cross Section in the Muon plus Jets Channel at sqrt(s)=10 TeV with the CMS Detector, CERN, CMS-PAS-TOP-09-003.

[13] The ATLAS Collaboration, Update on the jet energy scale systematic uncertainty for jets produced in proton-proton collisions at $\sqrt{s} = 7$ TeV measured with the ATLAS detector, ATLAS-CONF-2011-007.

[14] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, Toolkit for Multivariate Data Analysis with ROOT, arXiv:physics/0703039 (2007).

[15] T. Junk, Sensitivity, Exclusion and Discovery with Small Signals, Large Backgrounds, and Large Systematic Uncertainties, CDF/DOC/STATISTICS/PUBLIC/8128, `http://www-cdf.fnal.gov/~trj/mclimit/production/mclimit.html` (2007).