Draft version 3.1

ATLAS NOTE

August 14, 2009

Prospects for the Top Pair Production Cross-section at $\sqrt{s}$ = 10 TeV in the Single Lepton Channel in ATLAS

ATL-PHYS-PUB-2009-087
14 August 2009

Abstract

The measurement of the top quark pair production cross-section at the LHC is considered for a center of mass collision energy of 10 TeV and an assumed 200 pb^{-1} of integrated luminosity. We present a detailed study of the prospects for measuring the cross-section with the ATLAS detector using final states consisting of a single lepton (e or μ) plus jets. We describe methods for obtaining the cross-section based on two simple but different event selections, neither of which require the ability to tag jets originating from a b-quark and one of which does not require the use of a missing energy cut. These event selections can be used to clearly demonstrate the existence of top quarks in the early data. A method to measure the dominant W+jets background from the data is developed and shown to significantly reduce the systematic uncertainties of the $t\bar{t}$ cross-section measurements. Assuming the Standard Model scenario, we expect a measurement of the cross-section with a relative uncertainty of less than 20% (excluding the luminosity uncertainty).

Contents

1 Introduction

The top quark, discovered at Fermilab in 1995 [1], completed the three generation structure of the Standard Model (SM) and opened up the new field of top quark physics.

After QCD jets, W and Z bosons, the production of top quarks is the dominant process in pp collisions at multi-TeV energies: in the first phase of LHC running, in which 10 fb^{-1} of data is expected, millions of $t\bar{t}$ pairs will be produced. Using these events, many questions, including questions about new physics can be answered. For a review of the potential of top quark physics at the LHC see [2].

The measurement of the $t\bar{t}$ cross-section in the single lepton channel implies the reconstruction of final states which include jets, leptons and missing transverse energy with the ATLAS detector. This will likely be the first measurement with these characteristics. Since many models of physics beyond the Standard Model predict events with similar signatures, this measurement is an essential stepping stone to new physics. Finally, $t\bar{t}$ events are a major background for many processes in the Standard Model (e.g. Higgs production) and beyond (e.g. supersymmetry).

Given its importance, several different techniques have been developed to measure the $t\bar{t}$ cross section in the single lepton channel. The studies reported here are focused on 200 pb^{-1} of ATLAS data, assuming 10 TeV pp collisions. Furthermore, these analyses have been developed *without* the use of b-tagging which might require more time to reach the required level of performance. We do, however, assume a functioning single lepton triggering system. The developed methods are sensitive to different sources of uncertainty, so that when the first data arrives one can choose the optimal method depending on how well the different sources of systematics are understood. The single muon and single electron channels are considered separately, giving independent determinations of the cross-section.

The first analysis, described in Section 4, begins with a simple but effective baseline event selection. Using these selected events, a simple reconstruction of the hadronically decaying top quark can be used to re-establish the top quark signal with the early data. Then, two different techniques are studied to extract the cross-section. The simplest is a Cut and Count method where one counts events according to an optimized selection criterion and simply estimates the background from Monte Carlo simulation. This method has the advantage that it does not rely on estimating correctly the shape of any distribution such as the top mass, but the disadvantage that it relies on a Monte Carlo simulation which will not be optimal at this stage. Related to this is the problem that the overall normalization of the dominant background, W+jets, suffers from a large theoretical uncertainty. This can be improved dramatically by measuring the W+jets normalization from data, which also significantly reduces the dependence of the final result on the jet energy scale (JES) uncertainty, Monte Carlo modelling uncertainties and the uncertainty of the luminosity of the LHC. The second technique takes the baseline selection and fits the reconstructed hadronic top mass peak, extracting the number of events. The main advantage of this Fit method is that it is less sensitive to the JES and the normalization of the backgrounds, which are subtracted in the fit procedure. The disadvantage is that it has a larger statistical uncertainty than the baseline counting method, since one is selecting events in the peak region only.

An alternative analysis (Variant analysis), presented in Section 6, is also considered. Unlike the baseline, the variant analysis uses tighter lepton requirements and *no* missing energy cut. This can be motivated by the fact that missing energy is a global observable of the ATLAS detector and might require substantial commissioning. The systematics of this analysis is compared to the baseline analysis and, in section 7, some general conclusions are drawn.

2 Top quark production and observables

2.1 Top quark production in 10 TeV collisions

Top quark pairs can be produced through both gluon-gluon and quark-antiquark scattering. The relative importance of these processes depends on the centre of mass energy and the particles involved in the collision. In pp collisions at the LHC $t\bar{t}$ production takes place at small momentum fractions of the proton $x \approx 10^{-2}$ where the gluon parton distribution function (PDF) is large; therefore the gluon scattering process dominates $t\bar{t}$ production (approx. 90% at 10–14 TeV). In the $p\bar{p}$ collisions at the Tevatron at 1.96 TeV, larger values of x are probed and valence antiquarks are available in the $\bar{p}$, so that $q\bar{q}$-scattering dominates (approx. 85%).

At the LHC the top pair production cross-section $\sigma_{t\bar{t}}$ is significantly larger than at the Tevatron, essentially because of the much larger parton luminosities, dominated by the gluon PDF. At 10 TeV $\sigma_{t\bar{t}}$ is around 50 times that of the Tevatron.

The cross-section for $t\bar{t}$ production at the LHC has been calculated both in fixed order calculations up to next-to-leading order (NLO) [3]; using resummation techniques at NLO up to next-to-leading logarithms (NLL) [4]; and at an approximate next-to-NLO (NNLO) with next-to-NLL (NNLL) resummation [5, 6]. The cross-section results for 10 TeV, assuming $m_t = 172.5$ GeV and using CTEQ6.6 and CTEQ6.5 PDFs [7], are 400 pb $\pm$ 11% at NLO and 400 pb $\pm$ 6% at NNLO, where the uncertainties come from varying the renormalization and factorization scales by factors of two and from uncertainties in the PDFs.

At 1.96 TeV, CDF and DØ each combined several analysis channels and obtained cross-sections of $7\ \text{pb}^{+9\%}_{-9\%}$ [8] and $8\ \text{pb}^{+12\%}_{-11\%}$ [9], respectively. All of these measurements are in good agreement with the predictions, and the experimental and theoretical uncertainties are of the same order.

2.2 Observables

In the Standard Model, the decay of top quarks takes place almost exclusively through the $t \to Wb$ decay mode. A W-boson decays in about 1/3 of the cases into a charged lepton and a neutrino. All three lepton flavors are produced at an approximately equal rate. In the remaining 2/3 of the cases, the W-boson decays into a quark of up-type and an anti-quark of down-type, e.g. $W^+ \to c\bar{s}$. Since the CKM matrix suppresses decays involving b-quarks as $|V_{cb}|^2 \simeq 1.7 \times 10^{-3}$, W-boson decay can be considered as a clean source of light quarks (u, d, s, c).

From an experimental point of view where one measures final states, one can characterise top pair events by the number of W-bosons that decay leptonically, 0, 1 or 2. In this paper the main channel of interest is that in which one W-boson decays leptonically and the other hadronically e.g. $t\bar{t} \to bW^+\bar{b}W^- \to bl^+\nu\bar{b}d\bar{u}$. This is a *semi-leptonic* $t\bar{t}$ decay. When the four quarks in a semi-leptonic $t\bar{t}$ event hadronise, the final state typically consists of four jets, a single charged lepton and missing energy from the neutrino. Initial and final state gluon radiation often increases the number of final state jets. Notice also that, if the three jets originating from one of the top quarks can be identified, the top quark can be reconstructed. These observations will provide the backbone of the analyses presented here.

The fully-leptonic channel, having two charged leptons, suffers from less background than the semi-leptonic channel. However, the presence of two neutrinos makes top reconstruction more difficult. The fully-hadronic channel suffers from a large QCD background and will be difficult to study in the early data.

3 Reconstruction of physics objects

In this section we define the basic reconstructed objects - electrons, jets, muons and missing transverse energy - used in the baseline analysis. For the variant analysis of Section 6, further quality cuts will be required. The definitions used are based upon the reconstruction algorithms standard and recommended in ATLAS [10].

Electrons

Electron candidates are reconstructed by the calorimeters and inner tracker of ATLAS, i.e. are electromagnetic energy deposits matched to tracks. The electron candidates which are used in the analyses here, and henceforth referred to as good electrons, are required to have pseudo-rapidity in the range $|\eta| \leq 2.47$ and $p_T > 20$ GeV. Furthermore, if a good electron candidate is found in the calorimeter crack region $1.37 < |\eta| < 1.52$, the candidate is discarded. Finally, good electrons are required to be isolated based on calorimeter energy: the additional transverse energy E_T in a cone with radius $\Delta R = 0.2$ around the electron axis is required to be less than 6 GeV.

The reconstruction efficiency expected for good electrons in $t\bar{t}$ events is shown in Fig. 1. The plots are produced by taking all the truth electrons coming from a W-boson, and counting how many are matched with ΔR = 0.05 to good reconstructed electrons (as described above): the efficiency is the ratio of the two numbers. The truth electrons were not required to pass the η or p_T cuts defined for good electrons. Hence, this plot also includes the acceptance.

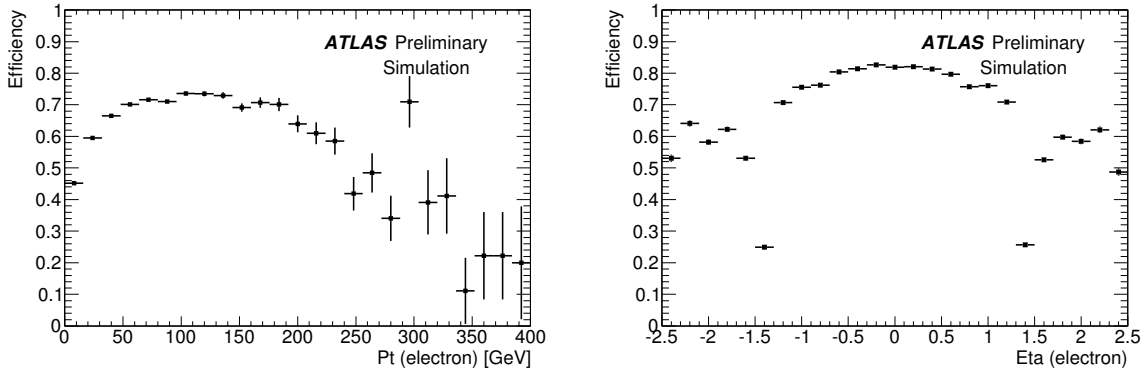

Figure 1: Left: Electron reconstruction efficiency versus true p_{T} and, right, versus η.

Jets

Jets are reconstructed with the standard ATLAS cone algorithm [10] in $\eta - \phi$ space, with a cone radius of 0.4, based on energy deposits in calorimeter towers of size $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$.

The jets used in the analyses in this paper are required to have $p_T > 20$ GeV, and the pseudorapidity restriction $|\eta|$ <2.5 . Finally, since many electron candidates are also reconstructed as jets, if a jet overlaps with a good electron within a cone of size $\Delta R < 0.2$, the jet is removed.

Muons

Muons are reconstructed by combining information from the muon spectrometer and inner detector, matching spectrometer hits to inner tracks. The muons used in the baseline analysis are required to have

a minimum transverse momentum, $p_T > 20$ GeV and to lie in the pseudo-rapidity range $|\eta| \leq 2.5$. They have to be isolated based on calorimeter energy: the additional transverse energy E_T in a cone with radius $\Delta R = 0.2$ around the muon is required to be less than 6 GeV. The expected reconstruction efficiency for muons from W-decays in $t\bar{t}$ events can be seen in Fig. 2.

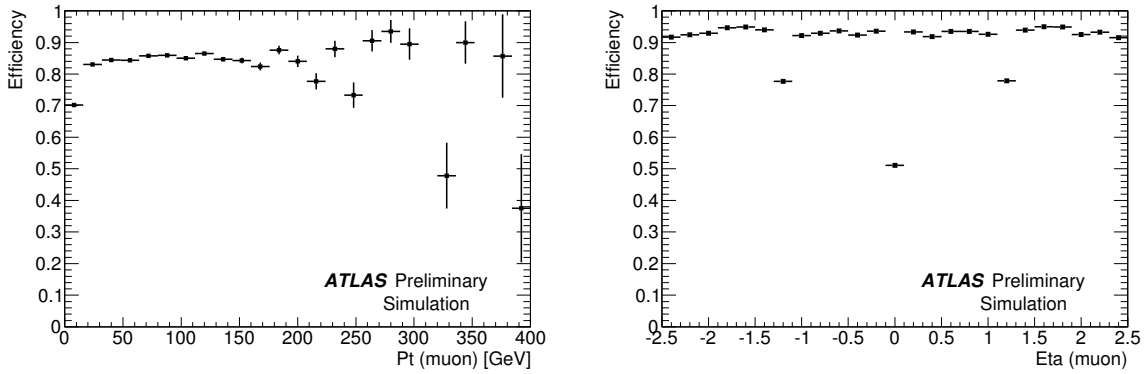

Figure 2: Left: Muon reconstruction efficiency versus true p_{T} and right, versus η.

Muons which are within a cone of radius $\Delta R < 0.3$ from a jet, are removed. This helps to remove those muons which arise from decays of hadrons inside jets, such as B-mesons originating from the b-quark in top decays.

Missing transverse energy

For the calculation of the missing transverse energy ($\not\!\!E_{\mathrm{T}}$) various contributions are taken into account – the cells in an identified electron or photon cluster, inside jets, or in topological clusters outside identified objects; muons and the cryostat correction [10].

4 Baseline analysis

The Baseline analysis presented in this section begins with a description of the event selection designed to select a $t\bar{t}$ event sample with high purity. We then describe the reconstruction of top quarks which undergo hadronic decays in these events. Based upon this selection and reconstruction, two methods are presented to determine the $t\bar{t}$ cross-section in the semi-leptonic decay mode. The first method is a simple Cut-based counting experiment where all backgrounds, excluding W+jets which is extracted from data, are estimated from Monte Carlo simulations. The second method is based upon fitting the reconstructed hadronic top mass peak.

4.1 Event selection

The identification of semi-leptonic $t\bar{t}$ events starts by requiring that either the single isolated electron or muon trigger, both with 15 GeV thresholds, have fired. The trigger efficiencies can be estimated either from Monte Carlo simulation or from data using clean samples of Z$\rightarrow ee$ or Z$\rightarrow \mu\mu$ events [10].

Next, we define a candidate $t\bar{t}$ event as containing:

- Exactly one lepton (electron or muon) with $p_{\mathrm{T}} > 20$ GeV.
- $\not\!\!E_{\mathrm{T}} > 20$ GeV.

- At least four jets with $p_{\mathrm{T}} > 20$ GeV.
- Of which at least three jets have $p_{\mathrm{T}} > 40$ GeV.

4.2 Reconstruction of $t\bar{t}$ events

In the candidate events, three of the (four or more) reconstructed jets are expected to have originated from a top quark which has decayed hadronically. In the absence of b-tagging information to identify the jets coming directly from the top quark decays, there is an additional ambiguity in selecting the correct three-jet combination from all of the reconstructed jets. We define the top quark candidate as the combination of three jets whose total p_T (i.e. the p_T of the sum of the three energy-momentum vectors) is the highest amongst all three jet combinations. This algorithm picks the correct combination about 35% of the time.

Additionally, one can exploit the fact that the three jets which originate from a hadronic top decay also contain two jets originating from a W-boson decay. There are three choices for this di-jet combination. Of these three di-jet combinations, if at least one of them has an invariant mass within 10 GeV of the value of the *reconstructed* W-boson mass M_W, we say that the event has passed the 'M_W-cut'. This algorithm is about 40% efficient. The reconstructed W-boson mass can be obtained by fitting the peak of the invariant mass distribution of these three di-jet combinations.

The distribution of the three-jet invariant mass with and without the M_W-cut is shown in Figure 3 with the major backgrounds included. Notice that the M_W-cut reduces the background significantly compared to the default selection and will ultimately lead to lower systematic uncertainties on the cross-section measurement.

4.3 Signal and background evaluation

The signal and major backgrounds have been estimated from Monte Carlo simulations which include a full simulation of the ATLAS detector. The dominant expected background is W+jets, but single top production, Z+jets and W$b\bar{b}$ (which was considered separately from W+jets) are also significant. Table 1 summarises the expected numbers of signal and background events for the electron and muon channel analyses. The first column of the two tables shows the event numbers obtained by applying the default selection, whilst the second column gives the corresponding numbers with the M_W-cut. All numbers are normalised to 200 pb^{-1} . Note that the '$t\bar{t}$' referred to in the table denotes all top pair events in which at least one top quark decays leptonically i.e. 'signal' includes fully-leptonic $t\bar{t}$ events as well as $t\bar{t}$ events which include W-bosons which decayed into τ's. From the numbers given one can conclude that the efficiency with which semi-electronic $t\bar{t}$ events pass the default (resp. default + M_W-cut) event selections is 22% (resp. 11%). Similarly, semi-muonic $t\bar{t}$ events have event selection efficiencies of 27% (resp. 14%).

The QCD production of jets is characterised by a cross-section many orders of magnitude larger than the $t\bar{t}$ signal and could therefore be a potentially important background. By requiring the presence of a high p_T isolated lepton and missing energy it should be possible to reduce its contribution to below that of W+jets, but since the cross-section enhancement relative to the signal is so large, there will be some QCD events with a fake lepton and/or poor missing energy reconstruction that pass the event selections as well. Though we expect it to be small compared to W+jets, the QCD contamination of the $t\bar{t}$ signal will ultimately be quantified from collision data and is not treated here. See [10] and Section 6 for further details.

Table 1: Number of events which pass the various event selection criteria for the $t\bar{t}$ signal and for the most relevant backgrounds, for the electron analysis (left) and muon analysis (right). The results are normalised to 200 pb^{-1} .

Numbers of Selected Events				
	Electron Analysis		**Muon Analysis**	
Sample	**default**	**+M_W-cut**	**default**	**+M_W-cut**
$t\bar{t}$	2600	1286	3144	1584
W+jets	1305	448	1766	628
single top	210	81	227	98
Z$\rightarrow ll$ +jets	148	43	144	49
hadronic $t\bar{t}$	16	10	11	5
W $b\bar{b}$	21	7	32	10
WW	11	6	14	7
WZ	3	1	5	2
ZZ	0.4	0.2	0.5	0.2
Signal	2600	1286	3144	1584
Background	1715	598	2199	799
S/B	1.5	2.1	1.4	2.0

4.4 Cut and Count method

The Cut and Count method is a very simple approach to determining the cross-section. The total number of events that pass the event selection are counted and all expected backgrounds are subtracted in order to get the yield of $t\bar{t}$ events in the sample. The total efficiency (ε) for signal events to pass the event selection criteria is estimated from Monte Carlo. From the observed events meeting the selection criteria (N_{obs}), the estimated background (N_{bkg}) and the integrated luminosity ($\mathscr{L}$), the cross-section is determined as:

$$\sigma = \frac{N_{\text{sig}}}{\mathscr{L} \times \varepsilon} = \frac{N_{\text{obs}} - N_{\text{bkg}}}{\mathscr{L} \times \varepsilon} \tag{1}$$

N_{bkg} can be estimated from a combination of Monte Carlo simulations and data-driven methods. ε includes the geometrical acceptance, the trigger efficiency and the event selection efficiency, and is slightly dependent on m_t. The advantage of using event counts is simplicity. A significant drawback is that some backgrounds can only be estimated from Monte Carlo and have large associated systematic uncertainties. To address this, we will describe in Section 5 a data driven estimate of the major background, W+jets.

4.5 The hadronic top mass fit method

This method extracts the $t\bar{t}$ cross-section from the baseline selection by fitting the hadronic top mass distribution, i.e. the highest p_T, three-jet invariant mass distribution after the M_W-cut. As for the Cut and Count method, the electron and muon channels are separated giving two independent determinations of the cross-section.

To extract the number of completely reconstructed $t\bar{t}$ events (after having applied the default + M_W-cut selection) a binned maximum likelihood fit is performed on the three-jet mass distribution with a Gaussian function to model the signal on top of the background, which is fitted with a 6th order Chebyshev polynomial, as shown in Figure 4. The integral of the fitted Gaussian is then used as an estimator of the number of correctly reconstructed $t\bar{t}$ events.

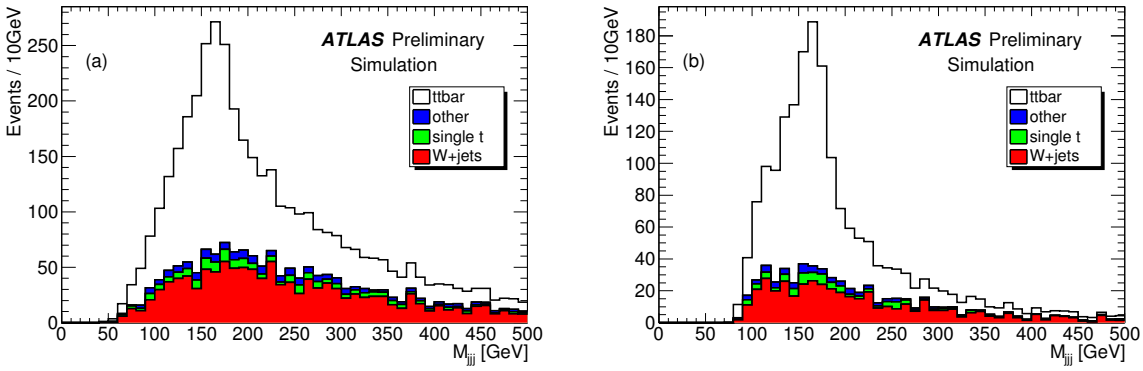

Figure 3: Left: Expected distribution of the three-jet invariant mass after the standard selection. Right: The same after the M_W-cut. Both plots are for the electron analysis, and the distributions are normalised to 200 pb^{-1} .

It has been verified that the background model correctly describes the combined $t\bar{t}$ combinatorial and background distribution in the signal region by comparing the fitted background to the non-resonant background as determined from information in the Monte Carlo generator. This check can also be made with data, on a background-enhanced sample, by selecting events which fail the M_W-cut.

The statistical uncertainty on the fit was evaluated by performing 10000 pseudo-experiments, using as input the hadronic top mass distribution from the fully simulated Monte Carlo events normalised to 200 pb^{-1} . The result of each pseudo-experiment was fitted and the σ of the Gaussian distribution of the results gives the statistical uncertainty on the fit.

For the muon channel, the average number of events in the peak, i.e. the correctly reconstructed semi-muonic $t\bar{t}$ events, is 396 with a statistical uncertainty of 59. This corresponds to an efficiency of $(3.42 \pm 0.52)\%$. For the electron analysis we observe 336 events in the peak with a statistical uncertainty of 48. This corresponds to an efficiency of $(2.90 \pm 0.41)\%$.

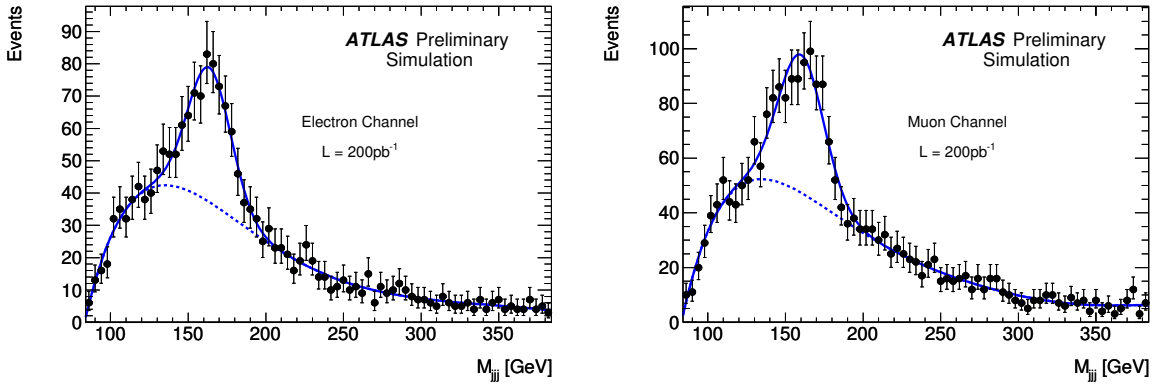

Figure 4: Left: The likelihood fit in the three-jet invariant mass for the electron channel. Right: The same for the muon channel. The statistical uncertainties correspond to an integrated luminosity of 200 pb^{-1} .

4.6 Systematic uncertainties in the baseline analyses

We now describe the evaluation of the main systematic uncertainties which affect both the Cut and Count and Fit methods for measuring the cross-section. The final results are collected in Table 2.

ISR/FSR modelling

To evaluate the cross-section uncertainty coming from the modeling of Inital/Final state radiation (ISR/FSR), we compared the results obtained with the $t\bar{t}$ ACERMC-generator [11] interfaced to PYTHIA [12] for showering and hadronisation. The Pythia parameters which control the ISR/FSR model were then varied in order to generate the maximum difference for the reconstructed hadronic top masses and two reference samples created. In the Cut and Count method, the relative differences between the numbers of selected events obtained with these samples, relative to the central sample, was taken as the uncertainty on the cross-section. One finds that this is one of the dominant sources of uncertainty for the fit method, which was redone with the two samples. Notice that the uncertainty for the fit method is one-sided (negative) because increasing the number of jets increases the combinatorial background, whilst decreasing the number of jets leads to fewer events in the Gaussian peak.

PDF uncertainties

Several different PDF sets were used to estimate the systematic effect of the uncertainties in the PDF parametrisations. In particular, we evaluated the spread of values obtained from both the MRST2006nnlo [13] and CTEQ6.6 [7] PDF sets. Since CTEQ6.6 showed a larger variation in the cross-section measurement, this is what is quoted in the final table.

Normalization of W+jets background

The uncertainty from the W+jets background normalization will be determined from Z+jets events using data driven methods, as will be described in Section 5. This justifies considering a 20% uncertainty on the W+jets background. Because it is determined by data, this component of the background does not enter in the JES and ISR/FSR systematic uncertainty estimates. For comparison, we also considered the effect of a 50% uncertainty in the W+jets cross-section.

JES uncertainty

The uncertainty with which the jet energy is measured affects the accuracy of the cross-section measurement. This jet energy scale (JES) uncertainty is evaluated in two cases. In the default case, the energy of reconstructed jets in the central region $|\eta|$ <3.2 are rescaled by $\pm$5% and $\pm$10% outside this region. In the second, pessimistic case, the JES uncertainty is assumed to be 10% for central jets and 20% otherwise. $\not\!\!E_{\mathrm{T}}$ is also rescaled accordingly. The cross-section σ is recalculated and leads to the following systematic uncertainty:

$$\Delta\sigma/\sigma = (S' + B' - B - S)/S \tag{2}$$

where S', B' are the numbers of signal ($t\bar{t}$) and background events after changing the JES, and B, S are the numbers for the default JES. The fit method is less sensitive to the uncertainty in the JES because, by fitting the signal and background, the effect of the JES is partially reabsorbed. As with the uncertainty with ISR/FSR modelling the effect is always one-sided (negative).

Monte Carlo generator uncertainties

Also considered was the uncertainty due to the use of different Monte Carlo generators. This was evaluated for the $t\bar{t}$ signal and the single top background by comparing the results obtained with the MC@NLO [14] and ACERMC [11] generators. The W+jets background is not considered in this context since it can be evaluated with a data driven technique.

Background cross-sections

The other backgrounds apart from W+jets and single top also have theoretical uncertainties on their normalisations e.g. due to the uncertainty in the value of α_s which leads to an 8% uncertainty in the background cross-sections. The fit method is less sensitive to the normalisation of the backgrounds, because they will be subtracted from the hadronic top mass plot.

Monte Carlo statistical uncertainty

The statistical uncertainty on the cross-section estimation due to limited Monte Carlo statistics is about 2% both in the electron and in the muon channel.

Fit systematic

The systematic uncertainty of the fit method is evaluated by changing the fit function for the background (e.g. by considering a polynomial of different order) and changing the binning of the histogram, in order to investigate the sensitivity of the method to the procedure employed for the fit.

Luminosity

An uncertainty of 20% on the integrated luminosity is assumed for the early data-taking period. Results are also shown for a 10% uncertainty. Since this has an impact on the background estimate, the contribution to the Cut and Count method is not simply proportional to the uncertainty. The W+jets background, being measured from the data, is not susceptible to this uncertainty.

Lepton ID and trigger efficiency

There is a small effect of about 2% [10] obtained by adding linearly the uncertainties in identifying the charged leptons (1%) and estimating the trigger efficiency (1%).

4.7 Cross-section evaluation with the baseline analysis

We have shown that it will be possible to observe a $t\bar{t}$ signal and determine its production cross-section with 200 pb^{-1} of data at 10 TeV. Assuming a functioning lepton trigger system, a 5% JES uncertainty and 20% uncertainty on the W+jets background, the Cut and Count method yields the following uncertainty on the $t\bar{t}$ cross-section using the baseline analysis plus the M_W-cut:

$$\texttt{ElectronCutandCount}\frac{\Delta\sigma}{\sigma} = (3(\texttt{stat})^{+14}_{-15}(\texttt{syst}) \pm 22(\texttt{lumi}))\% \quad (3)$$

$$\texttt{MuonCutandCount}\frac{\Delta\sigma}{\sigma} = (3(\texttt{stat})^{+12}_{-15}(\texttt{syst}) \pm 22(\texttt{lumi}))\% \quad (4)$$

where we added the systematic uncertainties from the table entries shown in bold in quadrature.

Table 2: Systematic uncertainties on the Cut and Count and Fit methods for the cross-section measurement, in percent, for electrons and muons. The entries in bold indicate the uncertainties which are finally quoted and constitute the total estimates. The final summed totals do not include the luminosity uncertainty which is quoted separately.

	Cut and Count method				**Fit method**	
Source	**e-analysis**		**μ-analysis**		**e-analysis**	**μ-analysis**
	default (%)	**+M_W-cut** (%)	**default** (%)	**+M_W-cut** (%)	**+M_W-cut** (%)	**+M_W-cut** (%)
Stat.	$\pm$ 2.5	$\pm$ 3.4	$\pm$2.3	$\pm$3.1	$\pm$ 14.1	$\pm$ 15.2
Lepton ID eff.	$\pm$1.0	$\pm$**1.0**	$\pm$1.0	$\pm$**1.0**	$\pm$ **1.0**	$\pm$ **1.0**
Lepton trig. eff.	$\pm$1.0	$\pm$**1.0**	$\pm$1.0	$\pm$**1.0**	$\pm$ **1.0**	$\pm$ **1.0**
50% W+jets	$\pm$25.1	$\pm$17.4	$\pm$28.1	$\pm$19.8	$\pm$ 3.3	$\pm$ 5.6
20% W+jets	$\pm$10.0	$\pm$**7.0**	$\pm$11.2	$\pm$**7.9**	$\pm$ **1.5**	$\pm$ **2.6**
JES (10%,-10%)	+24.8-23.4	+15.9-19.1	+20.5-22.3	+11.9-17.9	-14.4	-15.4
JES (5%,-5%)	+12.3-11.9	**+8.6-9.3**	+10.4-10.9	**+6.1-8.4**	**-3.7**	**-3.9**
PDFs	$\pm$1.6	$\pm$ **1.9**	$\pm$1.2	$\pm$ **1.4**	$\pm$ **1.9**	$\pm$ **1.4**
ISR/FSR	+9.1-9.1	**+7.6-8.2**	+8.2-8.2	**+5.2-8.3**	**-12.9**	**-12.9**
Signal MC	$\pm$3.3	$\pm$**4.4**	$\pm$0.3	$\pm$**2.8**	$\pm$ **4.5**	$\pm$ **1.4**
Back. Uncertainty	$\pm$0.6	$\pm$**0.4**	$\pm$0.5	$\pm$**0.4**	-	-
Fitting Model	-	-	-	-	$\pm$ **3.3**	$\pm$ **4.7**
10% Lumi.	$\pm$11.6	$\pm$11.2	$\pm$11.4	$\pm$11.1	$\pm$10	$\pm$10
20% Lumi.	$\pm$23.2	$\pm$**22.3**	$\pm$22.8	$\pm$**22.2**	$\pm$**20**	$\pm$ **20**
Tot. without Lumi.	+18.8-18.5	**+14.4-15.2**	+17.5-17.7	**+11.9-14.7**	**+6.4 -14.9**	**+6.0 - 14.8**

The Cut and Count method will thus allow a measurement of the $t\bar{t}$ cross-section with a less than 20% systematic uncertainty, excluding the luminosity uncertainty. The method is mainly sensitive to the JES, uncertainties in the modelling of ISR/FSR and the signal production process. With real data, these uncertainties are expected to be reduced significantly because the Monte Carlo can be tuned. Data can also be used to reduce correlations between different sources of uncertainty, which have not been considered here. Clearly, the analyses which include the M_W-cut yield a lower uncertainty, due to the higher S/B. The fact that the major background, W+jets, can be estimated from the data itself, is a major advantage to this analysis since its contribution is not susceptible to the uncertainties in the JES or luminosity.

For the Fit method, the relative uncertainties on the cross-section in the electron and muon channels are, respectively:

$$\texttt{ElectronFit}\frac{\Delta\sigma}{\sigma} = (14(\texttt{stat})^{+6}_{-15}(\texttt{syst}) \pm 20(\texttt{lumi}))\% \quad (5)$$

$$\texttt{MuonFit}\frac{\Delta\sigma}{\sigma} = (15(\texttt{stat})^{+6}_{-15}(\texttt{syst}) \pm 20(\texttt{lumi}))\% \quad (6)$$

which shows a similar systematic uncertainty in total, though with a larger statistical uncertainty when compared to the Cut and Count results.

5 Data driven estimate of W+jets background

The Monte Carlo predictions for the production of W+jet events have a large uncertainty. On the other hand, the W to Z ratio is predicted with a much smaller uncertainty [15, 16]. This ratio, as a function of jet multiplicity, is shown in Figure 5. Since the jet multiplicity distribution for Z events can be measured with data, this observation can be used to reduce the Monte Carlo uncertainty on the fraction of W+jets present in the selected sample of candidate top events. The idea is to extrapolate from a control region (CR) with zero or one jets into the top signal region (SR) with four or more jets and estimate the number of W+jets background events using the formula

$$(W^{\mathrm{SR}}/W^{\mathrm{CR}})_{\mathrm{data}} = (Z^{\mathrm{SR}}/Z^{\mathrm{CR}})_{\mathrm{data}} \cdot C_{\mathrm{MC}}, \qquad C_{\mathrm{MC}} = \frac{(W^{\mathrm{SR}}/W^{\mathrm{CR}})_{\mathrm{MC}}}{(Z^{\mathrm{SR}}/Z^{\mathrm{CR}})_{\mathrm{MC}}} \tag{7}$$

where W^{CR} and Z^{CR} represent the number of W and Z candidates reconstructed in the low jet multiplicity control region. Z^{SR} denotes the number of candidate Z events which pass the same selection criteria as those imposed in the default baseline analysis. This data driven technique is used for the estimation of the W$\rightarrow e\nu$ (W$\rightarrow \mu\nu$) background in the electron (muon) channel. The W$\rightarrow \tau\nu$ background will be estimated using the Monte Carlo to predict the ratio of $(\mathrm{W} \rightarrow \tau\nu)/(\mathrm{W} \rightarrow l\nu)$.

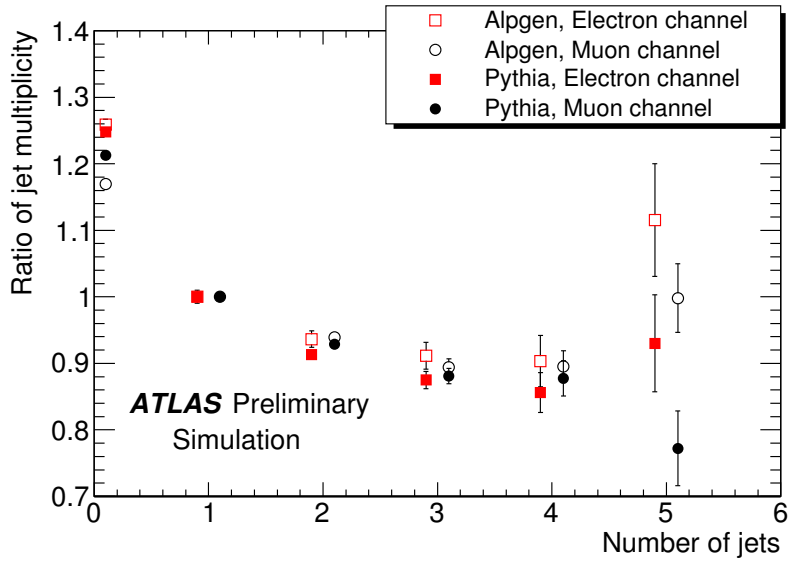

Figure 5: Ratio of reconstructed jet multiplicity for W$\rightarrow l\nu$+jets over Z$\rightarrow ll$+jets events. The ratio is taken after event selection cuts and normalizing the ratio to the 1 jet bin. Statistical errors are shown.

Z$\rightarrow ee$ (resp. Z$\rightarrow \mu\mu$) candidate events are selected (after the trigger) by requiring two electrons (resp. muons) of opposite charge, with an invariant mass between 80 GeV and 100 GeV. The signal-like Z sample is then selected by applying the default baseline analysis cuts on the jets, i.e. three jets with p_T above 40 GeV and a fourth with p_T greater than 20 GeV. The W$\rightarrow e\nu$ (resp. W$\rightarrow \mu\nu$) candidates are selected by requiring exactly one electron (resp. muon), zero muons (resp. electrons) and missing transverse energy greater than 20 GeV. The control region is defined to have exactly one jet with p_T greater than 20 GeV.

The statistics of the signal-like sample is much smaller (Table 3) than the low jet multiplicity control sample. The resulting statistical uncertainty represents a serious limitation on the precision of the final W+jets background estimation for low integrated luminosity. The table shows that the purity of selected Z samples is high. The background in this region will be estimated by fits to the invariant mass distribution or the Z peak sidebands.

In contrast, the contamination of the W control region from background processes is significant, the largest contribution being QCD dijet production. For the estimation of QCD background events in the electron channel a di-jet PYTHIA sample with a generator-level filter requiring two truth jets of $p_{\mathrm{T}} > 17$ GeV has been used. For the muon channel we assumed the main process for fake muons is heavy quark production. A PYTHIA $b\bar{b}$ sample was used to count the number of events surviving the cuts. Because of the huge cross-sections the QCD background will have to be estimated with data driven techniques. We assume that it will be possible to estimate it with an accuracy of 50%.

In Table 4 we show the various systematic uncertainties expected. We estimated the uncertainty of the Monte Carlo correction factor by comparing the ALPGEN [17] and PYTHIA predictions. These are 0.96 ± 0.03 for ALPGEN and 0.85 ± 0.04 for PYTHIA. By combining the difference between the two values and their uncertainties we quote an uncertainty on C_{MC} of 12.1%. The total uncertainty we estimate is 23.9% for W$\rightarrow e\nu$+jets and 19.6% for W$\rightarrow \mu\nu$+jets. Note that if we also apply the M_W-cut, the resulting drop in statistics increases the uncertainty by about 5%.

Table 3: Composition of the samples in the signal region (SR) and the control region (CR) for an integrated luminosity of 200 pb^{-1}.

Process	**Electron analysis**			**Muon analysis**		
	$Z \rightarrow ee$		$W \rightarrow e\nu$	$Z \rightarrow \mu\mu$		$W \rightarrow \mu\nu$
	CR (1 jet)	SR	CR (1 jet)	CR (1 jet)	SR	CR (1 jet)
Z(ee)	10210	82	1197	0.0	0.0	0
Z($\mu\mu$)	0.0	0.0	1.0	15750	150	8066
Z($\tau\tau$)	0.1	0.1	879	0.9	0.0	1130
W($e\nu$)	6.0	0.0	148700	0.0	0.0	0
W($\mu\nu$)	0.0	0.0	43	0.0	0.0	190300
W($\tau\nu$)	0.0	0.0	5570	0.0	0.0	6820
$t\bar{t}$	8.4	2.8	203	10.5	5.0	241
single top	2.9	0.0	272	2.3	0.0	308
Wbb	0.0	0.0	97	0.0	0.0	119
Diboson	24.9	0.5	427	40.0	1.0	557
QCD	110	0.4	42000	$\leq$50.0	$\leq$0.5	31000

Table 4: The expected relative uncertainties on the W+jets background estimation.

	Electron analysis	**Muon analysis**
Statistical for 200 pb^{-1}	11.3%	8.3%
Purity of control samples	17.0%	12.7%
Monte Carlo correction factor	12.1%	12.1%
JES ($\pm$10%)	3.6%	2.3%
JES ($\pm$5%)	3.0%	0.7%
Lepton energy scale	0.4%	0.7%
total uncertainty	23.9%	19.6%

6 Variant analysis without a cut on Missing Transverse Energy

The analysis presented in this section does not cut on missing transverse energy, which might not be completely understood at the beginning of data taking. To compensate, lepton identification is tightened with respect to the baseline analysis, by requiring higher p_T leptons ($p_T^{el} > 40$ GeV and $p_T^{\mu} > 30$ GeV) and more central values of $|\eta|$ ($|\eta|^{el} < 1.37$ and $|\eta|^{\mu} < 1.5$). The HT2 variable (scalar sum of the transverse momenta of the lepton and the second, third and fourth jet) is required to exceed 160 GeV. HT2 is a useful discriminator between QCD dijet events and $t\bar{t}$ because, in QCD, the energy of the event tends to be dominated by the two leading jets, whereas in semileptonic $t\bar{t}$ events, the energy is shared more uniformly amongst the four leading jets. Exactly one tight electron (muon) is required though, in contrast to the baseline analysis, we do not require zero muons (electrons). Finally, in the muon channel, the fourth jet p_T is required to be greater than 30 GeV. Table 5 shows the expected number of events for a luminosity of 200 pb^{-1}. From the $t\bar{t}$ events selected in the electron (muon) channel, 86.0% (85.6%) are true semileptonic electron (muon) events. The rest are mainly dilepton and semileptonic tau events. Of the $t\bar{t}$ events in the electron (muon) channel selected in the baseline analysis, 52.2% (44.1%) are also selected in this analysis.

Table 5: Number of events which pass the different selection criteria for $t\bar{t}$ signal and the most relevant backgrounds, normalized to 200 pb^{-1}.

Electron Analysis		**Muon Analysis**	
Process	Events (200 pb^{-1})	Process	Events (200 pb^{-1})
$t\bar{t}$	1483	$t\bar{t}$	1611
W$\rightarrow e\nu$ +jets	597	W$\rightarrow \mu\nu$ +jets	585
W$\rightarrow \tau\nu$ +jets	36	W$\rightarrow \tau\nu$ +jets	26
W b$\bar{\text{b}}$	12	W b$\bar{\text{b}}$	12
Z$\rightarrow ll$ +jets	100	Z$\rightarrow ll$ +jets	102
hadronic $t\bar{t}$	1	hadronic $t\bar{t}$	2
single top	63	single top	61
WW, WZ, ZZ	9	WW, WZ, ZZ	6
Signal	1483	Signal	1611
non QCD background	817	non QCD background	794
QCD background	240	QCD background	14

Since there are not enough simulated QCD events to properly evaluate the QCD contribution to the total background, the selection has been split into hadronic requirements (jets and HT2) and leptonic requirements (trigger and lepton). The full selection efficiency is computed as the product of the hadronic and leptonic efficiencies. The method has been proven to be good within the statistical uncertainty.

Two different approaches have been studied to determine the cross-section of the $t\bar{t}$ process. The simplest one is to apply a cut-and-count method, as in the baseline analysis. From Table 5 the expected statistical uncertainty on the $t\bar{t}$ cross-section is 3.2 (3.0)% in the electron (muon) channel for an integrated luminosity of 200 pb^{-1}. The limited Monte Carlo statistics introduces an uncertainty on the cross-section estimation of about 2% in both the electron and muon channels.

The second method extracts the number of signal events by fitting reconstructed quantities to a combination of template distributions by means of a χ^2 fit. To evaluate the performance of this method a subsample of the Monte Carlo simulated signal and background events, mixed in the appropriate amounts,

has been used as pseudo-data. We show results for the top mass (combination of three jets which gives the highest transverse momentum) distribution, but the method could be extended to other distributions. For a given distribution, the observed data ($\mathscr{D}_{data}$) can be fitted to a weighted sum of $t\bar{t}$ ($\mathscr{D}_{t\bar{t}}$), W+jets plus QCD ($\mathscr{D}_{W,QCD}$), and other background ($\mathscr{D}_{other}$) distributions as $\mathscr{D}_{data} = A \times \mathscr{D}_{t\bar{t}} + B \times \mathscr{D}_{W,QCD} + C \times \mathscr{D}_{other}$. A χ^2 fit determines the values of the coefficients A, B and C. Both the $t\bar{t}$ signal and the third component which includes small background contributions like single top, Z+jets, diboson production, etc. are taken from Monte Carlo. For small integrated luminosities this component is kept fixed to its expected value in the fit. The mass distributions of the W+jets and QCD background for the events passing the $t\bar{t}$ selection are consistent with that from QCD events fulfilling the hadronic requirements only, as shown in Fig. 6 (left) for the electron channel. Hence the latter will account for the shape of these two sources of background. Fig. 6 (right) shows the result of the χ^2 fit to the pseudo data for the muon channel. Fig. 6 is made with the result of the template fit on a sample of pseudo-data with smaller integrated luminosity (135 pb^{-1}) and slightly different mix of signal and background samples with respect to the Cut and Count method, but is normalised to 200 pb^{-1} for comparison purposes. To estimate the expected statistical uncertainty of the method, the pseudo-data was divided into eight subsamples, of 17 pb^{-1} each, which were then fitted individually. The corresponding statistical uncertainty on the $t\bar{t}$ cross-section is 5.7% for the electron channel and 5.6% for the muon channel for an integrated luminosity of 200 pb^{-1}.

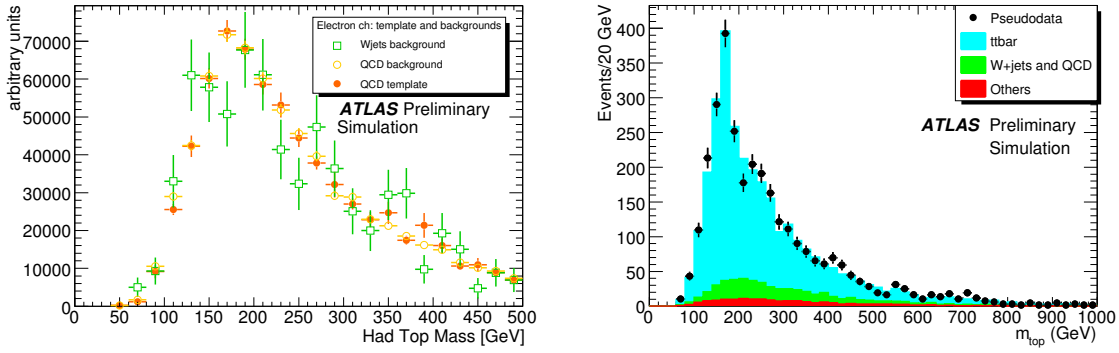

Figure 6: (Left) Reconstructed top invariant mass distribution for QCD events (passing the hadronic selection only, QCD template) superimposed to W+jets and QCD events (passing the full selection) for the electron channel. (Right) χ^2 fit to the top invariant mass distribution for the muon channel, normalised to 200 pb^{-1}.

Table 6 lists the main systematic uncertainties for the two cross-section extraction methods presented in this section. The jet energy scale is the dominant systematic for the χ^2 fit measurement, due to the displacement of the peak position of the signal. This uncertainty could be reduced by using several template options with different jet energy scales.

The first 200 pb^{-1} of data collected will allow ATLAS to measure the $t\bar{t}$ production cross-section with a precision of:

Table 6: Systematic uncertainties in the analysis presented in section 6, in percent.

Source	Electron analysis		Muon analysis	
	Counting (%)	**Fit** (%)	**Counting** (%)	**Fit** (%)
Statistical	±3.2	±5.7	±3.0	±5.6
Lepton ID and trigger efficiency	±**3.0**	±**2.0**	±**3.0**	±**2.0**
50% more W+jets	±22	–	±21	–
20% more W+jets	±**9**	–	±**8**	–
100% more QCD	±**8**	–	±**1**	–
Shape background template	–	±**2.9**	–	±**2.3**
Muon fake rate (20%)	–	–	±**0.7**	–
Electron fake rate (50%)	±**9**	–	–	–
Sensitivity to fakes	–	±**3.5**	–	±**2.4**
Jet Energy Scale (10%)	+15-23	-21-14	+23 -24	-27-18
Jet Energy Scale (5%)	**+8-11**	**-11-4**	±**12**	**-12-7**
PDFs	±**1.8**		±**2.3**	
ISR/FSR	±**8**	±**7**	±**13**	±**5**
Signal MC	±**1.1**	±**6**	±**3.6**	±**7**
Background uncertainty	±**0.9**	–	±**0.8**	–
10% luminosity	±13	±10	±12	±10
20% luminosity	±**26**	±**20**	±**23**	±**20**
Total without lumi	**+19.2 - 20.6**	**+10.5 -15.2**	**+20.2 -20.2**	**+9.4 - 15.3**

$$\texttt{VariantAnalysis:ElectronCutandCount}\frac{\Delta\sigma}{\sigma} = (3(\texttt{stat})^{+19}_{-21}(\texttt{syst}) \pm 26(\texttt{lumi}))\% \quad (8)$$

$$\texttt{VariantAnalysis:MuonCutandcount}\frac{\Delta\sigma}{\sigma} = (3(\texttt{stat})^{+20}_{-20}(\texttt{syst}) \pm 23(\texttt{lumi}))\% \quad (9)$$

$$\texttt{VariantAnalysis:ElectronFit}\frac{\Delta\sigma}{\sigma} = (6(\texttt{stat})^{+11}_{-15}(\texttt{syst}) \pm 20(\texttt{lumi}))\% \quad (10)$$

$$\texttt{VariantAnalysis:MuonFit}\frac{\Delta\sigma}{\sigma} = (6(\texttt{stat})^{+9}_{-15}(\texttt{syst}) \pm 20(\texttt{lumi}))\% \quad (11)$$

where we added in quadrature the systematic uncertainties shown in bold in the table.

The Cut and Count method presented in this section provides a similar statistical precision as in the baseline analysis. The systematic uncertainty is of the same order for the muon channel while it is slightly degraded for the electron channel due to the QCD background. The statistical precision of the fitting method degrades approximately by a factor two compared to the cut and count method. On the other hand the sensitivity to the normalization of the main W+jet background and the presence of residual QCD background is reduced, as well as the impact of the luminosity uncertainty. The sensitivity to the Jet Energy Scale is reduced but it remains important. In the strategy of the fit method used with the baseline selection, there is no constraint coming from the reconstructed top mass, hence the sensitivity to the Jet Energy Scale is smaller. This comes at the cost of some statistical precision, since only the resonant part is used as signal, and a larger sensitivity to the shape of the background parametrization.

7 Conclusions

Several methods have been presented which show that it is possible to measure the $t\bar{t}$ cross-section in the single lepton plus jets channel with 200 pb^{-1} of data from 10 TeV pp collisions in ATLAS without the use of b-tagged jets. The different methods are sensitive to different sources of systematic uncertainties which will allow an optimal choice once the conditions of the ATLAS detector are understood during the early data taking period. Under the assumed detector conditions, these methods afford a systematic uncertainty of less than 20 % plus the luminosity uncertainty.

The Variant analysis further shows that, without using a missing energy cut, similar results to the baseline analysis are possible. A data-driven technique for the estimation of the W+jets background in the e and μ channels has also been presented. Using this method, it appears possible to control the W+jets background to within 20 %. This clearly demonstrates that systematic uncertainties can be substantially reduced with such data-driven background estimates. In principle, some of the sources of uncertainty, such as the modelling of ISR/FSR can be reduced further once collision data becomes available.

A Bibliography

References

[1] CDF Coll., Phys. Rev. D 50 (1994) 2966; CDF Coll., Phys. Rev. D 50 (1994) 2966; CDF Coll., Phys. Rev. D 50 (1994) 2966; CDF Coll., Phys. Rev. Lett. 74 (1995) 2626; CDF Coll. Phys. Rev. Lett. 73 (1994) 225;
D0 Coll., Phys. Rev. Lett. 74 (1995) 2632; Phys. Rev. Lett. 74 (1995) 2632.

[2] M. Beneke *et al.*, arXiv:hep-ph/0003033.

[3] P. Nason, S. Dawson and R. K. Ellis, Nucl. Phys. B 303, 607 (1988). W. Beenakker, H. Kuijf, W. L. van Neerven and J. Smith, Phys. Rev. D 40, 54 (1989).

[4] M. Cacciari, S. Frixione, M. M. Mangano, P. Nason and G. Ridolfi, JHEP 0809, 127 (2008).

[5] S. Moch and P. Uwer, Phys. Rev. D 78, 034003 (2008). S. Moch and P. Uwer, Nucl. Phys. Proc. Suppl. 183, 75 (2008).

[6] N. Kidonakis and R. Vogt, Phys. Rev. D 78, 074005 (2008).

[7] P. M. Nadolsky et al., Phys. Rev. D 78, 013004 (2008).

[8] CDF Collaboration, *Combination of CDF top pair production cross section measurements with 2.8 fb^{-1}*, CDF Note 9448, 2008.

[9] V. M. Abazov et al. (DØ Collaboration), arXiv:0903.5525, submitted to Phys. Rev. D.

[10] ATLAS Collaboration, arXiv:hep-ph/0901.0512v3, CERN-OPEN-2008-020, Geneva, 2008.

[11] B.P. Kersevan and E. Richter-Was, arXiv:hep-ph/0405247.

[12] T. Sjoestrand, et al. Computer Phys. Commun. 135 (2001) 238, arXiv:hep-ph/0010017.

[13] A. D. Martin it et al., Phys. Lett. B 652 (2007) 292.

[14] S. Frixione and B.R. Webber, JHEP 0206 (2002) 029, arXiv:hep-ph/0204244;
S. Frixione et al., JHEP 0308 (2003) 007, arXiv:hep-ph/0305252.

[15] F. A. Berends, W. T. Giele, H. Kuijf, R. Kleiss, W. J. Stirling, Physics Letters B 224, (1989) 237.

[16] E. Abouzaid, H. Frisch, Phys. Rev. D 68 (2003) 033014.

[17] M.L. Mangano et al., JHEP. 0307,(2003) 001.